Efficient awards and standards of proof in judicial proceedings

Daniel L. Rubinfeld*

and

David E.M. Sappington**

We view the court system as an institution that enables defendants to signal their innocence or guilt, and we examine how the court can optimally minimize expected social losses from errors of type I and type II and from expenditures by defendants. Two of the policy instruments assumed available to the court are the standard of proof of one's innocence and the penalty imposed on a defendant who fails to meet the standard. Our analysis focuses on the effects that variations in the levels of these instruments have on the expenditures of defendants.

1. Introduction

Signalling models have gained popularity as a vehicle to a better understanding of a variety of economic institutions. (Cooper (1984) presents a survey.) The models demonstrate how institutions can serve, in part, to facilitate inferences about the unobservable qualities of individuals from observable outcomes. The purpose of this article is to suggest that the judicial process might profitably be viewed as an institution that enables defendants to signal their innocence or guilt to society. According to this view, trials (rather than settlements) can be socially valuable, despite the substantial litigation costs they entail. This effort is particularly important because the trial process has received relatively little attention in the growing law and economics literature. Instead, authors such as Shavell (1982a, 1982b) and Grossman and Katz (1983) have focused principally on the incentives of defendants to settle their case out of court or to plea bargain.1 The signalling approach described in this article provides some new conclusions about the design of an efficient trial system.

* University of California, Berkeley.
** Bell Communications Research.

We are grateful for helpful comments we received from Lucien Bebchuk, Ted Bergstrom, Joel Demski, Drew Fudenberg, Robert Inman, Richard McLean, Martin Perry, A. Mitchell Polinsky, Robert Pollak, Steven Shavell, Carl Simon, George Stigler, participants in workshops at Berkeley, Santa Cruz, Stanford, and the Federal Trade Commission, and, particularly, Al Klewski and anonymous referees. The authors would also like to thank the Exxon Foundation and the National Science Foundation for generous financial support (the latter through grants IST83-15690 and SES84-09858). This research began when both authors were at the University of Michigan. The views expressed in this article are not necessarily those of Bell Communications Research.

1 There are parallels between our framework and the conceptual framework adopted by other studies in the law and economics literature. Posner (1973) examines the behavior of parties involved in the litigation of accident cases. The court's objective in his analysis is to minimize the sum of costs associated with litigation and the failure
Our depiction of the trial process presumes that the trier of fact (a judge) faces a defendant whose guilt or innocence is known only to himself. After hearing the case, the judge assesses the probability of the individual's innocence. This assessment is an observable signal over which the defendant has some influence via his effort during the trial process. Greater effort (in, for example, preparing arguments and collecting evidence) increases the judge's estimate of the probability that the defendant is innocent. Defendants choose effort levels to minimize their expected losses, while taking into account both the monetary penalty they will incur if found guilty and the personal cost of effort.

The court's task is to select the optimal levels of the two policy instruments at its disposal: the standard of proof for innocence and the penalty imposed on a defendant who does not meet the standard. In setting its policy parameters the court's objective is to minimize the expected social costs from: (1) errors of type I (convicting an innocent party), (2) errors of type II (acquitting a guilty party), and (3) litigation effort by defendants. Issues of fairness, equity, and deterrence also bear consideration, but we abstract from these issues since our primary concern is with the signalling properties of the trial process. Our conclusions should therefore be interpreted as explaining how signalling considerations inform broader investigations of the judicial process, rather than as policy prescriptions in and of themselves.

In Section 2 we develop the formal model and carefully describe the signalling technology. Two benchmark problems are also presented, with which the court's actual problem is subsequently compared. After examining some extreme settings in Section 3, we turn to more realistic environments in Section 4. Our focus is on how defense expenditures respond to variations in the court's policy instruments. Concluding remarks appear in Section 5. Since our conclusions follow either from manipulation of first-order conditions or lengthy but straightforward calculations, most formal proofs are omitted. All proofs are available from the authors.

2. The basic model

In our simple view of the judicial process an individual, accused of a specific crime, appears before the court. The objective of the court is to minimize the sum of the expected social losses from type I errors, type II errors, and the costs incurred by defendants in the judicial process. Let $L_1$ and $L_2$ represent the social losses when type I errors and type II errors, respectively, are committed. Errors occur because the judge cannot observe directly whether a particular defendant is innocent or guilty. To capture the court's initial beliefs about a defendant's innocence or guilt, we let $p^0$ represent the fraction of defendants in the population who are guilty, and $p^r = 1 - p^0$ the corresponding fraction of innocent defendants.

Only the effort that defendants expend can lead the court to revise its initial beliefs about their innocence. We let $e^0$ represent the expenditure of effort by a guilty defendant and $e^i$ the corresponding expenditure by an innocent defendant. The probability that the judge assesses a defendant of type $i$ ($i = G$ for a guilty defendant and $i = I$ for an innocent defendant) to be innocent with probability $q$ when the defendant incurs effort $e^i$ in the course of his defense is denoted $f(q; e^i)$. Hence, $f(\cdot)$ is the probability density function to achieve efficient deterrence. Kleverick (1977) studies the problem of optimal jury size by using a social welfare function that includes the social costs of type I and type II errors and the administrative costs of the jury system. Joskow and Kleverick (1979) examine optimal judicial design with regard to predatory pricing, when the court's objective is to minimize the social costs associated with type I and type II errors and with the cost of implementing policies designed to inhibit predatory pricing. Our model also parallels that of Guasch and Weiss (1981), in which a firm structures employment test fees and wages to induce more able workers to signal their superior skills.

2 Two analyses that consider expenditures by two parties with conflicting interests (such as a defendant and prosecutor) are Braeutigam, Owen, and Panzar (1964) and Salant (1984). The former, however, focuses on the behavioral effects of alternative fee systems, not on the ability of the court to minimize type I and type II errors. The latter concentrates on the settlement-trial decision, not on effort at trial.
associated with an index, \( q_i \), of the likelihood of the defendant’s innocence, and we assume \( f(\cdot) \) is common knowledge. The probability that the defendant of type \( i \) who expends effort \( e_i \) will fail to convince the judge that he is innocent with a probability in excess of \( \bar{q} \) is 
\[
F_i(q, e_i) = \int_0^{\bar{q}} f(q, e_i) dq.
\]
The complement, \( 1 - F_i(q, e_i) \), is the associated probability that the defendant succeeds in convincing the judge of his innocence.

The basic goals of the adversarial process suggest that litigation effort by an innocent defendant generally should be more effective than an equal expenditure of effort by a guilty defendant. If this were not the case, litigation would serve no purpose, since it would not enable the court to distinguish more accurately the innocent from the guilty. Furthermore, since evidence to support one’s plea of innocence is more readily available and less costly to gather when one is innocent, the innocent will likely be able to support their plea more effectively than the guilty.

Although this assumption seems the most plausible, others are possible. For example, a small initial amount of effort by an innocent defendant could have very high productivity (as when a conclusive alibi is presented), but further expenditure might have very little effect on the judge compared with additional effort by a guilty defendant. Since this readily available evidence would likely be uncovered during the prosecutor’s pretrial inquiry, so that the judge would never hear the case, we shall not dwell on this alternative.

Formally, we adopt the following assumptions. Assumptions 1–3 hold for all nonnegative values of \( e_i \) and \( e_i \), and for all values of \( q \in (0, 1) \). Subscripts indicate partial derivatives.

**Assumption 1.** \( F_i(q, e_i) \leq 0, i = G, I \).

**Assumption 2.** \( F_i(q, e_i) \geq 0, i = G, I \).

**Assumption 3.** \( |F_i'(q, e)| \geq |F_i''(q, e_i)| \).

In words, increased effort increases at a decreasing rate—Assumptions 1 and 2—the probability that the judge’s final assessment will exceed any given level of \( q \). Ceteris paribus, this probability increases with effort more rapidly for the innocent defendant than for the guilty defendant—Assumption 3. We also maintain the following assumption, which states that with no expenditure of effort, the probability of achieving any prescribed standard of proof is the same for the innocent and guilty.

**Assumption 4.** \( F_i(q, 0) = F_i(q, 0) \forall q \in [0, 1] \).

We presume that there are two policy instruments available to the court: a standard of proof and a penalty for conviction. The standard of proof is a critical value, \( \bar{q} \), of the judge’s subjective probability. If the judge concludes that the defendant is innocent with probability \( q \leq \bar{q} \), the defendant is declared guilty and pays a penalty, \( D \). If the judge determines \( q \) to be greater than \( \bar{q} \), the defendant is declared innocent and incurs no penalty.

Defendants choose their effort levels to minimize expected losses: losses arise from litigation effort and, if convicted, from the penalty. Utility functions are separable and symmetric in effort and monetary penalties, as is indicated in the constraint in (1) below, which captures each defendant’s objective.

The court’s problem is the following:

\[
\min_{q, e_i} p^{iL} F_i(q, e_i) + p^{D} L^2 [1 - F_i(q, e_i)] + p^{D} q + p^{L} e_i
\]

(1)

3 There are other policy instruments that might be available to the court, including the fraction of total litigation expenses for which defendants deemed innocent or guilty are liable (Dewees, Prichard, and Trebilcock, 1980; Shavell, 1982a). The court might also determine rules for settling cases out of court as in Belchuk (1984), Png (1983), Salant (1984), and Shavell (1982a, 1982b).
subject to
\[ e^i \in \arg\min_{e \geq 0} \{ F^i(\bar{q}, e)[D + e] + [1 - F^i(\bar{q}, e)]e\}, \quad i = G, I \quad \text{and} \quad \bar{q} \in [0, 1]. \]

Before proceeding, we establish two benchmarks with which the solution to (1) can be compared. In the first benchmark situation the court is able to discern costlessly the defendant’s true “type” (i.e., his guilt or innocence). More precisely, in the first-best outcome: (i) each defendant is correctly classified as innocent or guilty, and (ii) no defendant exerts any effort on his defense.

A related benchmark is the second-best outcome, which is the solution to the court’s problem when the litigation effort of defendants (but not their types) is costlessly observed. Some features of the second-best outcome are summarized in the following observation.

Observation 1. If Assumptions 1–4 hold, then in the second-best outcome: (i) guilty defendants spend no money on their defense, and (ii) innocent defendants undertake expenditures until the marginal expected reduction in the loss from type I error is unity, i.e., \( L^1/F^i(\bar{q}, e^i) = 1. \)

3. Extreme judicial mechanisms

Before we characterize the general solution to (1) in detail in Section 4, we examine its properties for two extreme cases. In the first case the judicial process is very adept at distinguishing between guilty and innocent defendants. In the second case the judicial process adds no additional useful information concerning the guilt of the defendant.

To characterize the first extreme we introduce the following two assumptions.

Assumption 5. \( F^i(q, e^i) = -\infty \quad \forall q \in [0, 1] \quad \text{and} \quad \forall e^i \geq 0. \)

Assumption 6. \( F^0(q, e^0) = 0 \quad \forall q \in [0, 1] \quad \text{and} \quad \forall e^0 \geq 0. \)

If Assumptions 5 and 6 hold, the effort of the innocent defendant is infinitely productive at convincing the judge that the person before him is innocent, while that of the guilty defendant is completely unproductive.

Observation 2. If Assumptions 1–6 hold, then the court can achieve an outcome that approximates the first-best outcome arbitrarily closely by setting \( \bar{q} \) arbitrarily close to 1.

Diametrically opposite to this situation is the one described in Observation 3 in which litigation effort by innocent and guilty defendants provides no useful information to the court, because all defendants are equally productive at convincing the judge of their innocence (or alternatively, because the guilty are more productive than the innocent).

Observation 3. Under Assumption 4 if Assumptions 1 and 2 hold as strict inequalities but Assumption 3 holds as an equality, then at the solution to (1) \( \bar{q} = 1 \) if \( p^0L^2 > p^1L^1 \), and \( \bar{q} = 0 \) otherwise.

Under the conditions of Observation 3, the effort of defendants does not improve the court’s ability to distinguish the innocent from the guilty. Consequently, there should be

---

4 In the second-best outcome we allow the court to link the penalty for conviction to the defendant’s expenditure of effort. The standard of guilt could also be so linked, but we ignore this possibility in our definition of the second-best outcome to facilitate subsequent comparisons.

5 Suppose that Assumptions 1, 2, and 4 hold, but that Assumption 3 is replaced by Assumption 3a: \( F^i(q, e) < [F^i(q, e)] \) for all \( q \in (0, 1) \) and \( e > 0 \), so that a guilty party is more productive at convincing the judge of his innocence than is an innocent party. Then, in the second-best outcome neither type of defendant will undertake any defense effort.
no trials, and all defendants should be classified as guilty or innocent, depending upon whether the unconditional expected loss from a type II or a type I error is greater.\textsuperscript{6}

Observation 3 suggests that when signalling considerations dictate the court's policy, extreme standards of proof will occur when it is difficult to distinguish between innocent and guilty parties. This conclusion is also consistent with our view that the adversarial system can be socially valuable when trials better enable the innocent to signal their innocence.

4. More general properties of the optimal court system

In this section we discuss the properties of the optimal court system when innocent defendants are better able to convince the judge of their innocence than are guilty defendants (Assumption 3 holds as a strict inequality), but the advantage is not so extreme that the judge can discern perfectly a defendant's identity (Assumption 1 holds as a strict inequality for \( t = G \), and \( -\infty < F_G^t(q, e^I) < 0 \forall q \in [0, 1], \forall e \geq 0 \)). We focus on interior solutions to (1), in which both innocent and guilty defendants exert a strictly positive amount of effort in their defense.\textsuperscript{7}

Observation 4. At the (interior) solution to (1), \( e^G > e^G \); the innocent devote greater effort to their defense than do the guilty.

Observation 4 follows from the fact that the marginal expected return to effort is greater for innocent than for guilty defendants, \textit{ceteris paribus}. Because the innocent exert more effort on their defense and are more productive, the court can use \( q \) as an informative signal about the defendant's guilt or innocence. Because the court cannot observe the defendant's effort, however, the court has less control than it would like over how informative a signal \( q \) is.

Observation 5. At the (interior) solution to (1), guilty defendants exert more effort on their defense and innocent defendants less than in the second-best outcome.

Observation 5, which is proved in the Appendix, indicates how the court's inability to observe \( e \) affects its calculus. In the second-best outcome the court links the penalty for conviction to the defendant's level of effort. An appropriate choice of penalties ensures that the guilty defendant spends nothing, and the innocent defendant's level of effort equates the social marginal cost of additional effort with the marginal expected reduction in the social loss from a type I error. (Recall Observation 1.) In the court's actual problem, as expressed in (1), the effort expended by defendants cannot be observed. As a result, the court cannot achieve the second-best outcome with \( D \) and \( \bar{q} \) as its only policy tools. By setting \( D > 0 \) and \( 0 < \bar{q} < 1 \), the court induces the level of effort put forth by the innocent defendant to rise toward its second-best level. But the expenditures of the guilty defendant also rise, and excessively rapidly from the social perspective. Thus, as a “compromise,” \( D \) is established at a level below that which induces the “correct” (second-best) level of effort from the innocent defendant. This result, which we state as Observation 6, is also demonstrated in the Appendix.

Observation 6. At an interior solution to (1), \( D < L^I \); the optimal penalty for conviction is strictly less than the social loss incurred when a type I error is committed.

\textsuperscript{6} It follows immediately from Observation 3 and footnote 5 that when Assumption 3 is replaced by Assumption 3c, the second-best outcome is a feasible solution to (1).

\textsuperscript{7} At interior solutions Assumption 2 holds as a strict inequality, \( F_G^t(q, 0) = -\infty \) and \( F_G^t(q, \infty) = 0 \) for \( q \in (0, 1) \), and the conditions cited in Observation 3 do not hold. We also assume that the interior solutions to (1), as written, are identical to the interior solutions to (1) where the defendants' first-order condition replaces the constraint in (1) (Roberson, 1985).
The court’s choice of \( \bar{q} \) also diverges from its second-best level when the constraint in (1) is in effect, but the direction of this divergence cannot be determined in general. The ambiguity arises because there are two opposing effects, either of which can dominate. First, because the guilty exert excessive effort relative to the second-best outcome, a higher critical standard of proof can reduce the expected loss from type II errors. Second, a lower value of \( \bar{q} \) can compensate for the fact that the innocent exert insufficient effort. Which effect dominates depends on the court’s beliefs and also on the marginal productivity of effort by innocent and guilty defendants.

The factors that affect the optimal deviations from the second-best outcome are best identified by a comparative-static analysis. Because of the generality of the model, however, we cannot unconditionally sign the relevant comparative-static derivatives. We can only explain the factors that affect the signs of these derivatives.

The ambiguity of general comparative-static results in the model is illustrated by considering an increase in \( L \). Such an increase will cause the court to adjust its policy parameters to reduce the expected incidence of type I error. There are a variety of ways in which this can be accomplished. For example, the court might increase the penalty for conviction, which would induce the innocent to spend more on their defense and thus reduce the probability that they would be convicted. It is possible, however, that an increase in \( D \) would cause the guilty to increase their litigation effort more than the innocent and thereby diminish the value of \( q \) as an informative signal of the defendant’s type.8

Thus, the court may prefer to lower both \( D \) and \( \bar{q} \) to reduce the incidence of type I error. A lower \( \bar{q} \) will reduce the probability of conviction and therefore of type I error, ceteris paribus. When a reduction in \( \bar{q} \) serves to increase the expenditure of effort by defendants, such a reduction may be optimally coupled with a decrease in \( D \). For analogous reasons, however, the court may find it optimal to increase \( \bar{q} \) when \( L \) rises if a decrease in \( \bar{q} \) would cause an excessive reduction in effort by the innocent relative to the guilty.

Hence, comparative-static derivatives cannot be signed in general. Many of the ambiguities are eliminated, though, if the court is limited to a single policy instrument. The case in which \( D \) is fixed is relevant because penalties are commonly tied directly to the nature of the defendant’s crime (e.g., mandatory minimum sentences for armed felonies). Yet in such cases, the court may have latitude in applying the standard of proof. The case in which \( \bar{q} \) is fixed is also relevant, as the court will often adhere strictly to a standard of proof such as “beyond a reasonable doubt,” yet exercise considerable discretion when penalties are imposed. Observations 7 and 8 address these two cases in turn.

Observation 7. If \( \bar{q} \) is fixed and beyond the control of the court, then the optimal penalty at an interior solution to (1) increases (i) as \( L \) or \( p^f \) increases and (ii) as \( L^2 \) or \( p^g \) decreases.

Observation 8. If \( D \) is fixed and beyond the control of the court, then at the interior solution to (1): (i) if \( F^L_{\bar{q}}(\cdot) > 0 \), the optimal level of \( \bar{q} \) decreases as \( L \) increases; and (ii) if \( F^L_{\bar{q}}(\cdot) < 0 \), the optimal level of \( \bar{q} \) increases as \( L \) increases.

Observation 7 implies that when the court wishes to reduce the incidence of type I error, the innocent will be induced to increase their litigation effort to reduce the probability that they will be found guilty according to the fixed standard, \( \bar{q} \). Although the guilty will also increase their effort as \( D \) increases, the social cost of the resulting increase in the incidence of type II error is outweighed by the social gain resulting from fewer type I errors, owing to the lower marginal signalling costs of the innocent. To illustrate Observation 8 we can show that when \( F^L_{\bar{q}}(\cdot) > 0 \) at the solution to (1), so that the marginal productivity of the innocent

---

8 Of course, if \( L \) increased with \( D \), there would be an additional reason for the court to lower \( D \) as \( L \) increased.
defendants' effort increases as $\bar{q}$ declines, a decrease in $\bar{q}$ induces the innocent to increase their defense efforts. The increase reduces the incidence of type I errors as $L^1$ increases.

5. Concluding remarks

Our analysis has focused on the optimal use of two policy instruments available to the court: the standard of proof and the penalty imposed on those defendants who fail to meet this standard. We have shown that the socially efficient magnitudes of these instruments are very sensitive to the exact relationship between the litigation efforts of defendants and the judge’s ultimate assessment of a defendant’s guilt (i.e., the signalling technology).

Our analysis has taken the signalling technology to be exogenous. More generally, however, the court might be able to undertake expenditures to reduce the degree of randomness in the judicial system. Such expenditures can be a valuable policy instrument if they induce patterns of spending by innocent and guilty defendants that cannot be replicated by setting $D$ and $\bar{q}$ alone. One possibility is that an increase in court expenditure will increase the ratio of expenditures by innocent defendants to expenditures by guilty defendants. In this case the optimal level of court expenditure will balance the court’s increased ability to distinguish innocent from guilty against its increased direct costs. Another possibility is that a decrease in court expenditure, resulting in more randomness, will increase spending by innocent defendants relative to that by guilty defendants, at least over some range of expenditures. In this case increased randomness can result in both lower direct social costs and a judicial system that is better able to discriminate between innocent and guilty defendants.

In closing we emphasize that our concern has been entirely with signalling considerations relating to the trial process. Some of the findings need to be reexamined when other objectives of the judicial system are taken into account. For example, although a reduction in the penalty imposed on those judged guilty can increase the value of $\bar{q}$ as an informative signal of the defendant’s type, such a reduction in the penalty may encourage more crimes to be committed. Hence, when deterrence is included as a goal, it may no longer be optimal to reduce the penalty. Similarly, a reduction in the penalty may not be optimal when the alleged crime is serious and retribution is a concern. Nevertheless, our findings point out some factors that have largely been overlooked but should inform a general, normative analysis of the court system.

Appendix

The proof of Observations 5 and 6 follows.

Proof of Observations 5 and 6. At an interior solution to (1), $e^i > 0$. Hence, the first statement in Observation 5 follows from Observation 1.

Among the necessary conditions for an interior solution to (1) are the following:

$$
\lambda^i F^i(q, e^i) + \lambda^j F^j(q, e^j) = 0, \tag{A1}
$$

$$
p^i L^i F^i(q, e^i) - p^j - \lambda^i DF^i(q, e^i) = 0, \tag{A2}
$$

$$
p^i L^i F^i(q, e^i) - p^i - \lambda^i DF^i(q, e^j) = 0, \tag{A3}
$$

$$
DF^i(q, e^i) = -1 DF^j(q, e^j). \tag{A4}
$$

Here $N, i = G, I$, are the Lagrange multipliers associated with the constraint in (1).

From (A4) and (A1) $\lambda^i$ and $\lambda^j$ are of opposite sign but have the same absolute value. From (A2) and Assumption 2 $\lambda^G < 0$. Hence, $\lambda^I > 0$. Therefore, from (A3), $L^i F^i(q, e^i) > 1$, so that from Observation 1 and Assumption 2 the second statement in Observation 5 follows.

Finally, note from (A3) and (A4) that $F^G(q) = p^i L^i - D^i \lambda^i D^i$. Hence, by Assumption 2 $D < L^1$, which proves Observation 6.
References


