ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

OPTIMAL DESIGN OF PRIVATE LITIGATION

Louis Kaplow

Discussion Paper No. 928

08/2017

Harvard Law School
Cambridge, MA 02138

Optimal Design of Private Litigation

Louis Kaplow*

Abstract

This article translates and extends Becker (1968) from public law enforcement to private litigation by examining optimal legal system design in a model with private suits, signals of case strength, court error, and two types of primary behavior: harmful acts that may be deterred and benign acts that may be chilled. The instruments examined are filing fees or subsidies that may be imposed on either party, damage awards and payments by unsuccessful plaintiffs (each of which may be decoupled), and the stringency of the evidence threshold (burden of proof). With no constraints, results arbitrarily close to the first best can be implemented. Prior analyses of optimal damage awards, decoupling, and fee shifting are shown to involve special cases. More important, previous results change qualitatively when implicit assumptions are relaxed. For example, introducing a filing fee can make it optimal to minimize what losing plaintiffs pay winning defendants and to reduce the evidence threshold as much as possible — even though the direct effect of these adjustments is to chill desirable behavior, a key feature absent in prior work.

JEL Classes: D82, H23, K13, K41, K42

Keywords: litigation, law enforcement, courts, fee shifting, decoupling, filing fees, burden of proof

---

*Harvard University and National Bureau of Economic Research. I am grateful to the editor, three referees, Jonathan Borowsky, Aaron Edlin, Steven Shavell, Holger Spamann, Kathryn Spier, and participants at Chicago, Harvard, NBER, and the Law and Economics Theory conference for comments; to Christopher Dunomes and Will Milliken for research assistance; and to the John M. Olin Center for Law, Economics, and Business at Harvard University for financial support. Disclaimer: I occasionally consult on antitrust cases, and my spouse is in the legal department of a financial services firm.

1. Introduction

Much law enforcement is accomplished using private suits, usually filed by victims, rather than through public enforcement. Inspired by Becker (1968), the problem of optimal system design in the latter context has received substantial attention. See the survey of Polinsky and Shavell (2007). By contrast, the majority of the literature on private litigation involves positive analysis of suit and settlement, with much less attention to optimal design. See Spier (2007).

The private enforcement context is qualitatively different and in some respects more challenging. In particular, plaintiffs' filing decisions are endogenous and, accordingly, are influenced by a variety of instruments that also affect actors' primary behavior. The social welfare problem is to maximize benefits from individuals' primary activity — which here includes both harmful and benign acts — net of harms caused and litigation costs.

The importance of exploring optimal system design in this setting is reinforced by the fundamental divergence identified by Shavell (1982), who shows how in even the most basic setting there is no tendency of plaintiffs' filing decisions to be socially optimal. In addition, we wish to understand the extent to which social welfare may be sacrificed by the fact that actual legal systems restrict the instrument set in ways that have received little attention in the literature and to determine how various limitations change the optimal use of those instruments that remain, which also has not been examined.

Section 2's model features three types of behavioral effects: the deterrence of harmful acts, the chilling of benign acts, and plaintiffs' inclination to sue. These may be influenced by several instruments: a fee (or subsidy) imposed on each of the plaintiff and the defendant at the time of suit, a damage award paid by losing defendants to successful plaintiffs (which may be decoupled, allowing the two amounts to differ), and a transfer paid by losing plaintiffs to successful defendants (which also may be decoupled). An extension adds an evidence threshold (akin to a burden of proof) that may be adjusted.

When individuals commit an act of either type, a signal of case strength is observed by a prospective plaintiff who then decides whether to sue. When there is a suit, each party pays a fee (which may be zero or negative) and incurs litigation costs. The tribunal then orders that transfers be paid, the amount and direction depending on whether liability is imposed. Individuals' decisions whether to commit their acts and plaintiffs' decisions whether to sue reflect private expected benefits and the expected costs of the legal system just described. The legal system's raison d'être is to deter harmful activity, but the planner also seeks to use unavoidably costly litigation as little as possible and, importantly, to enhance the system's diagnosticity, which promotes deterrence and reduces the chilling of benign conduct.

Before proceeding to the analysis of how this is accomplished, section 2 explores instrument redundancy. Two of the transfer instruments are redundant in a narrow accounting or mechanical sense, and an additional instrument is redundant in practice when others are sufficiently unconstrained. The relationships among these instruments help to illuminate prior literature wherein each paper allows only a couple particular instruments.

Section 3 demonstrates that if three instruments — a plaintiff filing fee, a standard (non-decoupled) damages award, and a (non-decoupled) transfer from losing plaintiffs to defendants — are unconstrained, we can to span many relevant choices of the three types of behavior (regarding harmful acts, benign acts, and suit). Moreover, it is feasible to implement a result arbitrarily close to the first best: first-best deterrence (individuals commit harmful acts if and only if their private benefit exceeds the external harm), no chilling of benign acts, and negligible litigation costs. The analysis is not straightforward because of the need to provide differential incentives to those who may commit harmful and benign acts, although in this unconstrained setting this is possible even though the evidence threshold is taken to be fixed.

In practice, instruments are typically restricted, by limits on feasibility (parties' payments cannot exceed their wealth) and by various legal institutions (perhaps for reasons outside this model). Section 4 analyzes a setting in which the available instruments are a filing fee that may be imposed on plaintiffs (only), a (non-decoupled, that is, ordinary) damages award that losing defendants pay to plaintiffs (which may be subject to a maximum), a (non-decoupled) transfer that losing plaintiffs pay to defendants, and the evidence threshold. The optimal sanction is maximal, the optimal transfer from losing plaintiffs to victorious defendants is minimal, and the optimal evidence threshold is as low as possible. The optimal filing fee — which trades off deterrence benefits and chilling as well as litigation costs — is characterized.

These results hold despite the presence of court error involving the imposition of sanctions on individuals who commit benign acts, which we wish to chill as little as possible. The core intuition behind these somewhat surprising conclusions is that plaintiffs' self-interested filing decisions themselves reflect the signal and hence sort cases by quality to that extent: plaintiffs file only those cases that are stronger than some critical level, which minimum can be thought of as akin to a de facto evidence threshold. By raising the filing fee — in conjunction with raising the sanction, lowering the transfer paid by losing plaintiffs, or reducing the de jure evidence threshold — one can, for example, maintain plaintiffs' filing decisions, holding fixed this de facto threshold. But there's more: each of the stated adjustments actually allows for an even higher filing fee and thus a more stringent filing threshold, which eliminates the weakest cases. It is thereby possible to simultaneously increase deterrence, hold chilling constant, and reduce filing rates and thus total litigation costs. Even the lower de jure evidence threshold enables a higher de facto threshold and thereby can better protect individuals who commit benign acts. A parallel intuition for these results is that each policy experiment makes the system harsher on defendants in inframarginal cases while making it more generous by removing marginal cases; even when the former bears relatively more heavily on benign acts, the latter relatively benefits benign acts to an even greater degree because it involves the weakest cases from among those previously filed.

The present analysis casts prior literature on the optimal design of private litigation in a different light.[1] The most relevant articles consider notably simpler environments along two different dimensions. First, no previous analyses of private litigation include the combination of ex ante signals of case strength, two types of court error, both harmful and benign acts, and legal system costs.[2] The omission of benign behavior that may appear to cause harm to prospective plaintiffs and hence be difficult for tribunals to distinguish from truly harmful acts is particularly important. For example, there may be suits for securities fraud when stock prices drop, such as after an IPO, and it may be difficult in a particular case to know whether there was really fraud or instead the drop was caused by other factors. The prospect of having to expend large sums to

[1]There is also some prior positive analysis of some of the elements at play here. See Polinsky and Shavell (1989).

[2]Kaplow (2011) includes this array but with public enforcement; also, fewer instruments are examined (some of which are moot when the government chooses enforcement effort by fiat). See also McAfee, Mialon, and Mialon (2008), who compare public and private enforcement of antitrust laws when private plaintiffs are defendants' competitors.

defend such suits (even if ultimately successful) and the possibility of erroneous liability may discourage IPOs, raise the cost of capital, and otherwise disrupt efficient corporate behavior. See Lin, Pukthuanthong, and Walker (2013). In antitrust, a competitor may sue a dominant firm for predatory pricing, which may be difficult to distinguish from efficient promotional pricing, moving down a learning curve, or a competitive response to new entry. Again, the prospect of litigation costs and mistaken assignments of liability can chill efficient behavior. Perhaps most familiar in this regard is medical malpractice, where there are concerns that the costs and errors of the legal system may reduce physician supply and induce defensive medicine. See Kessler, Sage, and Becker (2005) and Studdert et al. (2005).

Moreover, policies that might be optimal in models with only harmful acts, because they maximize deterrence, could be undesirable if one introduced benign acts that would be chilled. In the public enforcement setting, Kaplow's (2011) analysis of the optimal burden of proof shows that results are qualitatively quite different when the model allows for benign as well as harmful acts. Yet, in the context of private suits, Png's (1986) brief paper is the only investigation of two-way errors and benign as well as harmful acts. In his model, however, litigation is costless and all cases are filed, so few of the challenges investigated here are present.

A second limitation of prior work on private litigation is that models usually confine attention to (at most) two instruments: damages combined with one other. Polinsky and Che (1991), the seminal paper on decoupling, like subsequent work on the subject, examines a model with only harmful acts and thus one-way error, and it does not include filing fees or payments by losing plaintiffs. In a similarly limited framework, Kaplow (1993) addresses fee shifting. And when Polinsky and Rubinfeld (1996) introduce the possibility that losing plaintiffs can be made to pay defendants, they do not allow filing fees or decoupling and likewise do not model benign acts. Each setting is a special case in many respects. Moreover, as the foregoing explains, the optimal use of one instrument can change qualitatively when implicit restrictions on other instruments are relaxed — in some instances, by the mere introduction of a filing fee. Indeed, the sometimes surprising results in the present model are due to relaxing prior work's implicit assumptions, whose significance in generating contrary conclusions has not been appreciated.[3]

2. Model

A. Setup

There are two types of acts (indexed by a) that may be committed, a harmful one, H, and a benign one, B. The harmful type of act imposes an external social cost of h; the benign type of act involves no externality (although the analysis would be qualitatively the same if it causes a smaller negative externality or a positive externality). A mass of individuals normalized to 1 may commit the harmful type of act. Those who may commit the benign type of act have a mass

[3]This paper does not pursue an unrestricted mechanism design approach, under which the only limitations are derived from the information structure and underlying technology, because the purpose here is to explore the optimal design of private litigation of a form broadly employed and the subject of prior literature. The result in section 3 that it is possible to achieve an outcome arbitrarily close to the first best obviously would not differ under the pure mechanism design approach. In section 4, if instrument constraints are taken to depend, for example, on limits on parties' wealth, then any further limitations that make the outcome fall short of the first best might be mitigated if a broader class of instruments were allowed.

of γ. One interpretation is that γ indicates the relative mass of benign acts that may be undertaken in situations in which they might initially be confused with harmful acts (because other benign acts do not give rise to the possibility of suit and sanctions). Instead, one could imagine that the same individuals may commit both types of acts and that γ indicates the relative frequency of opportunities to commit the benign type of act.

The model also admits a third interpretation that involves a greater difference: one could allow individuals to choose one of the two acts or inaction, which reaches, for example, the case in which an act might be undertaken with a high level of care, a low level of care, or not at all. This variation would complicate the exposition but have only a modest effect on the qualitative results. Specifically, deterrence of harmful acts would induce some individuals to switch to benign acts rather than inaction, making deterrence more valuable, and chilling of benign acts would cause some individuals to switch to harmful acts rather than inaction, making chilling more detrimental. All of the below propositions would hold (but the terms in square brackets in the first line of the first-order condition in expression (4) would be modified accordingly).

An individual's benefit from committing an act is b, with density functions $f^a(b)$ (which are positive for positive values of b) and cumulative distribution functions $F^a(b)$, where, recall, $a = H, B$. Individuals know what type of act they are able to commit and its benefits to them, but the tribunal knows neither acts' types nor actors' benefits. Individuals commit their acts if and only if their private benefit b exceeds expected legal costs, as specified below. (All actors are taken to be risk neutral.[4])

After an act is committed, a prospective plaintiff observes the signal π, the probability of prevailing at trial. For each type of act, the densities and cumulative distribution functions for π are given by $g^a(\pi)$ and $G^a(\pi)$, respectively, which are assumed to satisfy the strict monotone likelihood ratio property: $\pi_1 > \pi_0$ implies that $g^H(\pi_1)/g^B(\pi_1) > g^H(\pi_0)/g^B(\pi_0)$, which is to say that higher values of π are relatively more likely to be generated by acts of type H than those of type B. (The extension in section 4.B models the tribunal's signals explicitly, allowing for an analysis of the evidence threshold as an additional instrument; the extended model collapses into that described here when that threshold is fixed.) The plaintiff may also be assumed to observe the type of act, although in the present model only knowledge of the signal π matters to the plaintiff's decision, which is now described.[5]

The prospective plaintiff (P) sues the prospective defendant (D) — usually referred to as the plaintiff and the defendant for simplicity — if and only if the expected gains from suit exceed the expected costs. The legal system operates as follows: First, when a suit is filed, each party pays the government a fee φ^j, where $j = P, D$; these fees may be negative (subsidies). (The superscript a, introduced above, refers to types of acts, H and B, and the superscript j to parties, P and D.) After a suit is filed, all cases proceed to trial and parties incur litigation costs, c^j.[6]

---

[4]If plaintiffs, defendants, or both were risk averse, then behavior would be influenced in familiar ways. Moreover, social welfare would differ because transfers would no longer be neutral. See, for example, Polinsky and Shavell (1979), and also note 10.

[5]This depiction may be viewed as a reduced form. For example, we may imagine that, before filing, the plaintiff observes an initial signal and the defendant's type, which jointly imply a further distribution of a subsequent signal that will emerge at trial. Here, the signal π is can be understood as a sufficient statistic for the ultimate probability of liability.

[6]The analysis throughout abstracts from settlement. With symmetric information and assuming that some costs are incurred before settlement (which is typical in reality), the qualitative results would be unchanged, as in Polinsky and Che (1991). As mentioned in the conclusion, one could allow the defendant to observe a different signal from that observed by the plaintiff, thereby introducing asymmetric information and making for more interesting explicit analysis

At trial, the tribunal assigns liability with probability π. If there is a finding of liability, D pays damages of δ^D and P receives damages of δ^P; that is, damages may be decoupled, although (for reasons given in section 2.C) most of the analysis will restrict attention to the case in which $\delta^D = \delta^P = \delta$. If there is a finding of no liability, P pays the transfer τ^P and D receives the transfer τ^D, although often it will be assumed that the transfers are not decoupled, with $\tau^D = \tau^P = \tau$. The values of the δ^j and τ^j are taken to be nonnegative unless otherwise indicated. Sometimes it will be supposed that δ or δ^D cannot exceed some maximum, δ^{Max}, as is common in much of the literature on the economics of public law enforcement.

It is useful to summarize much of the foregoing in a timeline:

1. The government sets all policy instruments: the parties' filing fees, φ^j; damages paid by losing defendants and received by victorious plaintiffs, δ^j ; and transfers paid by losing plaintiffs and received by victorious defendants, τ^j.
2. Defendants learn their type of act (H or B) and private benefit b.
3. Defendants decide whether to act.
4. For each act undertaken, the signal π is observed by a plaintiff.
5. Plaintiffs decide whether to sue.
6. If a suit is brought, the fees φ^j are paid to the government and the costs c^j are incurred by the parties.
7. With probability π, there is liability, in which event damages of δ^j are paid or received, as the case may be. With probability $1-\pi$, there is no liability, in which event transfers of τ^j are paid or received.

B. Behavior and Social Welfare

We can now characterize parties' actions. Plaintiffs sue if and only if the expected transfers they would thereby receive, net of the expected transfers they would have to pay (including their filing fee) and their costs of suit are positive, that is, if and only if: $\pi\delta^P - (1-\pi)\tau^P - \varphi^P - c^P > 0$. This may conveniently be rewritten in terms of a critical probability of liability:

$$(1)\ \pi > \frac{\tau^P + \phi^P + c^P}{\delta^P + \tau^P} \equiv \pi^*,$$

where π^* is the value of π that generates the probability indicated by the fraction on the right side of the inequality.[7] Note that π^* depends on the three transfers to or from plaintiffs and also on plaintiffs' litigation costs, whereas π^* does not depend on any of the defendants' instruments (or on defendants' litigation costs).

Turn now to primary behavior. Defendants will commit their harmful or benign acts, as the case may be, if and only if their private benefit from doing so exceeds the expected costs that arise in the event of suit, which are comprised of expected net transfers (those to be paid, minus

of settlement.

[7]The right side of the inequality can equal or exceed one for some parameter values, meaning literally that suit arises only when the probability of victory is in excess of one, which obviously holds for no value of π, meaning that no one sues. Also, without further restrictions on the transfers, the critical probability can be zero or negative, meaning that all plaintiffs sue. Attention is largely confined to cases in which $\pi^* \in (0, 1)$. See also note 13.

the value of any to be received) as well as legal costs. This condition is:

$$(2)\ b > \int_{\pi^*}^{1} \left[\phi^D + c^D + \pi\delta^D - (1-\pi)\tau^D\right] g^a(\pi)d\pi \equiv b^a,$$

where b^a denotes the private benefit of the prospective actor who is just indifferent about whether to commit an act of type a. Individuals who refrain from committing acts of type H will be described as *deterred*, and those who refrain from acts of type B will be described as *chilled*. Observe that actors' behavior is influenced by all of the plaintiffs' instruments (as well as plaintiffs' litigation costs) because π^* implicitly depends on all of these.

Social welfare, W, is taken to be the sum of individuals' benefits from committing both types of acts minus the harm caused by acts of type H and litigation costs incurred when either type of act is committed. Accordingly, social welfare is given by:

$$(3)\ W = \int_{b^H}^{\infty} \left[b - h - \left(1 - G^H(\pi^*)\right)(c^P + c^D)\right] f^H(b)db$$

$$+\gamma \int_{b^B}^{\infty} \left[b - \left(1 - G^B(\pi^*)\right)(c^P + c^D)\right] f^B(b)db.$$

The first term indicates all social benefits and costs with regard to individuals who commit harmful acts. These undeterred individuals, as indicated by expression (2), are those with benefits above the critical value, b^H. For each such act, there is the private benefit of the act, minus the harm it causes, and also minus the expected litigation costs associated with the act. To elaborate this latter expression, plaintiffs will sue when $\pi > \pi^*$ (expression 1), the probability of which equals $1 - G^H(\pi^*)$. And, when there is suit, both parties incur litigation costs. The second term regarding benign acts is analogous except that there is no direct harm associated with them. And this term is weighted by γ, the relative mass of benign acts. Finally, although none of our instruments appear directly in expression (3) for social welfare, all of them influence the b^a and the plaintiffs' instruments also determine π^*.

C. Instrument Redundancy

To avoid a priori restrictions and encompass prior work, the model here admits six transfer instruments: filing fees φ^j, damages (when liability is imposed) δ^j, and transfers (in the event of no liability) τ^j, $j = P, D$. This section explains that two of these are redundant in a mechanical, accounting sense and that another may sometimes be dispensed with as a practical matter. (The analysis here is largely independent of most of the model's structure – such as primary behavior, parties' signals, and court error – and hence is substantially more general.)

As a pure accounting matter, suppose, for example, that we increase δ^D by 1 and likewise reduce each of φ^D and τ^D by 1. A defendant pays 1 more (or, equivalently, receives 1 less) after trial, regardless of whether it loses or wins, and it pays 1 less (or, equivalently, receives 1 more) before trial, when the case is filed. Nothing really changes. This redundancy (and an analogous one for the plaintiff's transfers) reflects that we have two parties, P and D, and two outcomes, liability and no liability; hence, it is sufficient to specify four payoffs.[8] For example, we could take no liability as our baseline, define the two parties' payoffs in that situation via the two filing fees, and then further specify the difference in net outcomes when there is liability, which corresponds to decoupled damages payments. Furthermore, because this redundancy is mechanical, the use of a redundant instrument cannot relax real constraints. Thus, if δ^{Max} is the level of defendants' wealth, we cannot circumvent this limitation on δ^D by raise φ^D instead because the wealth-constrained defendant has that much less wealth available to pay δ^D.[9]

For convenience, let us pick four instruments in particular: the plaintiff's filing fee φ^P (the defendant's filing fee φ^D is taken to be fixed at 0), decoupled damages payments of δ^P and δ^D, and finally a non-decoupled transfer τ (which is to say that we are restricting attention to the case in which $\tau^D = \tau^P = \tau$). To implement the four payoffs, one can set τ to target D's baseline payoff, then φ^P to target P's baseline payoff, and finally δ^P and δ^D to generate each party's increment when there is liability.

Furthermore, another of the plaintiff's instruments may be redundant in practice. Plaintiffs' instruments are relevant only through their effect on expression (1) for plaintiffs' filing decisions (they do not enter expression (2) except through their influence on π^*). Therefore, we can target a given level of π^* as long as there is a single unconstrained transfer instrument for plaintiffs Specifically, we can eliminate the decoupling of damages payments in the following manner: raise or lower δ^P, as the case may be, until it equals δ^D while, at the same time, raising or lowering φ^P (in the same direction) to the degree that keeps π^* in expression (1) constant. When φ^P may thus be adjusted, it is sufficient to analyze the three transfer instruments $\{\varphi^P, \delta, \tau\}$, where $\delta^D = \delta^P = \delta$ and $\tau^D = \tau^P = \tau$.[10]

The foregoing discussion casts new light on prior understandings of various instruments. Start with decoupling, which is to say allowing $\delta^D \neq \delta^P$ (studied, for example, in Polinsky and Che 1991), or $\tau^D \neq \tau^P$. Regarding the former, suppose that we start without decoupling, i.e., $\delta^D = \delta^P = \delta$, and then contemplate raising or lowering just δ^P. The earlier discussion in this section indicates that, as an accounting matter, this is equivalent to instead introducing (allowing the adjustment of) a filing fee, φ^P — and also adding and making a corresponding adjustment to τ^P. Moreover, the discussion just above shows how the latter is actually unnecessary since this way of eliminating decoupling only influences π^* (any effects on the b^a are only through this channel) — and, as long as we are free to adjust φ^P, any level of π^* can be induced. Understanding decoupling as tantamount to allowing filing fees (or subsidies) on plaintiffs helps to unify analysis and also to make intuitions more transparent.

The present analysis also illuminates fee shifting. Consider two-way fee shifting, such as under the English rule, and let us focus now on a basic setting without any decoupling. This rule is simply a special case of the present model in which δ — ordinarily in private damages actions

---

[8]Alternatively, for each outcome with regard to liability, we need one instrument to indicate the transfer between the parties and another instrument to indicate the transfer between the parties as a group and the government.

[9]In similar spirit, one could consider a restriction on the amount a plaintiff could potentially pay, which would place a maximum on φ^P and, as appropriate, on the sum of φ^P and τ^P. See note 10 and section 4.D.

[10]If there is a constraint on φ^P (say, a limit to plaintiffs' wealth), then permitting reductions in δ^P offers an alternative manner of discouraging suits. If plaintiffs are risk averse, the analysis of behavior is largely the same (a different adjustment of φ^P will be required when eliminating decoupling, but the same π^* can be maintained), but social welfare will differ due to risk-bearing costs: for example, raising δ^P and φ^P in a manner that holds π^* constant increases risk-bearing. Finally, note that this set of adjustments, while holding π^* constant, is favorable to plaintiffs with $\pi > \pi^*$ (they have an expected gain, at the expense of the government's budget).

being taken to equal the harm suffered, h — is augmented by c^P, and we also take τ (relative to the ordinary benchmark of 0) to equal c^D. Clearly, if δ and τ are not otherwise constrained to equal h and 0, respectively, there is no particular significance to fee shifting in the present sort of model. Only by chance would $\delta = h + c^P$ and $\tau = c^D$ be optimal. Moreover, if we started with $\delta = h$ and $\tau = 0$, only by chance would moving toward fee shifting raise rather than lower social welfare. In addition, fee shifting is often thought to be problematic for two reasons: a party's prospect of having its costs shifted to the other party reduces the marginal cost of expenditures, which makes litigation more expensive, and it can also be nontrivially costly in practice for the adjudicator to determine c^P and c^D.[11] See Katz (1987). Accordingly, if the main appeal of fee shifting is to improve plaintiffs' filing incentives or defendants' primary behavior, then it may make more sense to adjust directly the outcome-contingent transfers, δ and τ.[12]

Finally, observe that, contrary to what is assumed in most prior work on the economics of litigation, it is natural to entertain the use of nontrivial filing fees (or equivalents, like decoupling) when assessing of a broad range of policies. Consider any change that reduces the costs of adjudication: because this may make suit more attractive, total costs could rise. However, if one can impose (or raise) φ^P to keep filing behavior constant — substituting a transfer payment for a preexisting real resource cost — one might instead achieve an unambiguous improvement. As will now be seen, this degree of flexibility also has important implications for optimal system design in the present setting, which focuses on primary behavior as well as system costs.

3. Optimal System Design without Instrument Restrictions

The social welfare maximization problem — setting the instruments so as to maximize expression (3) — is concerned with three margins of behavior, two relating to defendants' primary activity (deterrence, b^H, and chilling, b^B), and one to plaintiffs' decisions whether to sue (π^*). We can view the planner's problem as choosing individuals' behaviors, subject to incentive constraints, in order to maximize the objective function.

In light of the discussion of instrument redundancy, here we will suppose that the planner uses the following three instruments: a filing fee paid (only) by the plaintiff, φ^P, non-decoupled damages paid by losing defendants to plaintiffs, δ, and a non-decoupled transfer paid by losing plaintiffs to victorious defendants, τ. Furthermore, the analysis will assume that there exists a further constraint that $\delta + \tau > 0$ — a constraint that is automatically satisfied in the ordinary situation in which damages paid by losing defendants are positive and transfers paid by losing plaintiffs are nonnegative. Further justifications for this restriction are noted in the margin.[13]

---

[11]Also, fee shifting raises stakes, which increases incentives to spend on litigation, but so would adjusting δ and τ in an analogous fashion. In addition, in some models fee shifting will influence the likelihood of settlement (notably, the implicit increase in stakes makes a given degree of asymmetric information more likely to disrupt settlement), which also affects total system costs. But adjusting δ and τ directly would, again, have the same effect.

[12]See also Kaplow (1993), where shifting victorious plaintiffs' costs has the further disadvantage of (relatively) favoring filing by plaintiffs with higher litigation costs.

[13]If $\delta + \tau < 0$, defendants would prefer to lose at trial, which might appear to make little sense with regard to deterrence. In addition to seeming bizarre, a regime, in which both litigants try to lose at trial, seems unsustainable. Note further that, if $\delta + \tau < 0$, the inequality in expression (1) for plaintiffs' filing decision would reverse and the negative denominator would then imply that no suits would be filed — unless the numerator were also negative, in which event the suits filed would be ones with low merit (including, outside the model, that countless cases with literally no merit, i.e.,

Because we have three targets, it is natural to consider whether these three instruments are sufficient to implement any outcome. The first result is that a broad although incomplete range of outcomes — values of $\{b^H, b^B, \pi^*\}$ — can be implemented, and the second, which follows immediately, is that the planner can implement a result arbitrarily close to the first best.

Proposition 1: If there are no extrinsic constraints on plaintiff filing fees, damage awards, and payments from losing plaintiffs to winning defendants — the instruments φ^P, δ, and τ — then:

- *a. It is feasible to implement any value of $\pi^* \in [0,1)$ (a critical filing threshold) and any nonnegative value of b^H (level of deterrence) and, for those values, any nonnegative value of b^B (degree of chilling) up to some upper limit.*
- *b. It is feasible to implement any value of $\pi^* \in [0,1)$ and any nonnegative value of b^B and, for those values, any nonnegative value of b^H down to some lower limit.*

The proof, which appears in the appendix, involves a direct construction: showing how various targeted values of $\{b^H, b^B, \pi^*\}$ can be implemented using the three instruments. First, for any given values of δ and τ, there exists a φ^P that can induce any desired plaintiff's filing threshold, π^*. Second, given any such π^*, we can find a combination of damages δ and transfers paid by losing plaintiffs τ that induces any desired b^H. In varying δ or τ to target b^H, φ^P is adjusted so as to keep π^* fixed. Third, for any $\{\pi^*, b^H\}$, we can vary the losing plaintiff's transfer τ — adjusting δ so as to keep b^H constant and, in light of the changes in both τ and δ, adjusting φ^P to keep π^* constant — to implement different levels of b^B. As Proposition 1 and its proof indicate, the only limit to complete spanning is that is that b^B cannot be too high relative to b^H. Specifically, we cannot induce $b^B > b^H$ because benign acts in expectation generate weaker evidence while winning defendants fare better than losing ones do.

The implication of Proposition 1 with regard to the maximization of social welfare is immediate. Notably, the only limits are on making b^B (chilling) high or b^H (deterrence) low, but these do not inhibit the minimization of chilling and the achievement of first-best deterrence.

Proposition 2: If there are no extrinsic constraints on plaintiff filing fees, damage awards, and payments from losing plaintiffs to winning defendants — the instruments φ^P, δ, and τ — it is feasible to implement a result arbitrarily close to the first best. Specifically, one can induce $b^H = h$ (first-best deterrence), $b^B = 0$ (first-best chilling, which is to say, none), and π^* arbitrarily high (implying negligible suits and hence negligible litigation costs).

The proof of Proposition 2 is straightforward, as is evident from the following sketch. First, Proposition 1 indicates that we can implement any b^H, so in particular we can implement $b^H = h$ — which is to say that individuals with acts of type *H* commit them if and only if $b > h$, the condition for their being socially desirable — by an appropriate choice of δ with $\tau = 0$ and some intermediate π^*. Second, for this instrument setting, $b^B > 0$, so we can now raise τ (and correspondingly raise δ so as to keep b^H fixed at h) until the point at which $b^B = 0$ — which is to say no chilling of acts of type *B*. Third, we can do all of this for any $\pi^* \in (0,1)$, so we can choose π^* to be as high as we like in that interval, driving litigation costs arbitrarily close to zero.

$\pi = 0$, would be filed).

4. Optimal System Design with Instrument Restrictions

Section 3 indicates that that the instrument set $\{\varphi^P, \delta, \tau\}$ is sufficient to span much of the space of the endogenous behavioral variables, π^*, b^H, and b^B, and to achieve a result arbitrarily close to the first best. Unfortunately, for institutional and practical reasons, there often are further restrictions on the instruments that can be employed. Much of the analysis here will continue to assume that it is possible to control π^* directly, in particular through choice of the plaintiff's filing fee.[14] Section A examines how each of the three instruments we have considered thus far — φ^P, δ, and τ — are optimally set. Section B extends the model to introduce an additional instrument, the evidence threshold, and to determine its optimal setting. Section C discusses some of the assumptions used in the derivations as well as the implications of the results for legal system design. Section D considers restrictions on the filing fee.

A. Optimal Damages, Transfers, and Filing Fees

The propositions advanced here are that extreme values of the damages and plaintiff transfer instruments are optimal; specifically, δ should be as high as possible, and τ as low as possible. These results are demonstrated in an analogous manner (with proofs in the appendix), using the following approach. Supposing that the instrument under investigation is not at its maximum or minimum feasible value, as the case may be, the instrument is adjusted slightly in the indicated direction while simultaneously adjusting the filing threshold π^* (by the requisite adjustment of the filing fee φ^P) so as to keep b^B constant. That is, the overall expected burden on benign acts will be held constant, so that chilling is unchanged.[15] It is shown that this policy experiment involves a higher filing threshold and thus, on that account, lower litigation costs. Finally, it is demonstrated that the experiment also necessarily increases deterrence, which is to say, raises b^H. Under the assumption that will be conditionally maintained here that, at the optimum, greater deterrence, ceteris paribus, would raise social welfare, this increase in deterrence, along with the litigation cost savings, implies that social welfare rises. An increase in social welfare can also be demonstrated in a similar fashion under the alternative assumption that, at the optimum, greater chilling is welfare reducing. Attention is confined to optima in which some suits are filed (otherwise marginal adjustments to the instruments have no effects) and where defendants prefer that the marginal suit not be filed. These assumptions (which cover all cases of interest, but, as will be explained, not all possible cases) are revisited in section 4.C.

---

[14]In practice, negative values of φ^P may be problematic due to frivolous suits or collusion. A partial remedy may be to pay any filing subsidy only at the end of trial. Whether an unconstrained optimum will involve a filing subsidy depends, of course, on all the parameters of the model, as analyzed below. Note that the results in Propositions 3–5 (and also Proposition 2 in section 3) involve raising φ^P to values that may be extremely high, in which case the present concern would be moot; one might instead face a plaintiff wealth constraint, discussed in note 10 and in section D, below.

[15]The reader may wonder why a seemingly easier-to-analyze experiment, which holds π^* fixed, was not undertaken. That experiment holds the direct contribution to litigation costs constant and also zeros out one of the channels by which changing the other instruments influences both deterrence and chilling (because, from expressions A9, A14, and A15 in the appendix, one of the two channels by which instruments otherwise influence the commission of acts in this setting is through their effect on π^*). However, this value of π^* — which is lower than that implied by the experiment undertaken in the text — implies that chilling rises, and one cannot derive unambiguous results regarding social welfare even if it can be demonstrated that deterrence increases to a greater extent.

Proposition 3: If the optimum involves some suits ($\pi^* < 1$)*, defendants prefer that the marginal suit not be filed, and deterrence is welfare increasing (or chilling is welfare reducing) at the margin, then the optimal level of damages, δ, cannot be interior. Specifically, if it were, it would be possible to raise δ and raise φ^P such that welfare increases, so that the optimal δ equals δ^{Max}, if such a maximum exists.*

The optimality of raising damages (akin to the sanction in cases with public enforcement) to its maximum feasible level is reminiscent of the result based on Becker (1968) that one can raise the sanction and reduce enforcement effort so as to maintain deterrence and save enforcement costs. The present model differs not only in having private enforcement (making enforcement effort endogenous) but also in a number of other respects: most importantly, the introduction of two types of error (specifically, including the erroneous imposition of sanctions) and, relatedly, of benign activity that may be chilled as a consequence. One might have been concerned that raising δ would worsen chilling, which it does. Nevertheless, the same result is obtained, although the just-mentioned extensions make the argument much more involved. The core idea is that, even ignoring litigation cost savings, it is better to employ higher damages while employing a larger filing fee that reduces suits because the latter eliminates the weakest cases from the system (which have the worst tradeoff between deterrence and chilling) whereas the former raises defendants' payments in the remaining, stronger cases (which have a better tradeoff between deterrence and chilling).

We next turn to the transfers paid by losing plaintiffs, τ, which has received very little attention, and none in this sort of setting. Here, we will assume that $\tau \geq 0$, which is to say that losing plaintiffs may have to pay winning defendants but are never paid by them.[16]

Proposition 4: If the optimum involves some suits ($\pi^* < 1$)*, defendants prefer that the marginal suit not be filed, and deterrence is welfare increasing (or chilling is welfare reducing) at the margin, then the optimal payment from losing plaintiffs to winning defendants, τ, cannot be interior. Specifically, if it were, it would be possible to reduce τ and raise φ^P such that welfare increases, so that the optimal τ equals 0.*

Proposition 4's indication that it is optimal to reduce the amount that losing plaintiffs pay defendants as much as possible may seem surprising. The present model, unlike most prior work (including what little allows for transfers when defendants win), introduces benign conduct that tends to be chilled as a side-effect of the attempt to regulate harmful acts. This would seem to make attractive a requirement that losing plaintiffs pay winning defendants, for two reasons: such payments occur more often in weak cases, which are more likely to involve benign acts, and the prospect of having to make such payments discourages the filing of weak cases. How, then, can the contrary conclusion be explained?

The core intuition actually is the same as that favoring maximal damages, δ. Plaintiffs' filing decisions, out of their own self-interest, result in their selection of the strongest cases. As will be elaborated when discussing the optimal evidence threshold (Proposition 5, in section 4.B), the filing cutoff π^* serves as something akin to a de facto evidence threshold. Any sort of

[16]The proof of Proposition 4 implies that, had we not restricted τ to be nonnegative, τ should be as negative as possible: specifically, it should approach $-\delta$ (recall the constraint in section 3 that $\delta + \tau > 0$).

transfer, including τ, that rewards victorious defendants occurs only in suits with high values of π, specifically, ones in excess of π^*. Furthermore, the experiment used to prove Proposition 4 — which parallels that used in the proof of Proposition 3 — reduces τ in a manner that holds b^B constant, which implies an increase in π^* (which is implemented by raising φ^P).

The consequences for benign acts are as follows: First, although defendants receive less when they win, this occurs only in filed cases, which are the stronger ones, those for which $\pi > \pi^*$. Second, defendants gain by having fewer suits filed. Moreover, this reduction in filings, caused by the increase in π^* due to the higher φ^P, involves elimination of the weakest cases from among those previously filed. Taken together, therefore, the policy experiment is relatively favorable to defendants who committed benign acts rather than harmful ones. The intuition that a lower τ has the opposite relative effect is countered by the second component of the experiment, which operates not merely in the reverse direction but to an even greater degree. Defendants' gains from the experiment arise in cases ($\pi = \pi^*$) that are weaker than all of those ($\pi > \pi^*$) in which they suffer a detriment, producing relatively greater gains for those who commit benign acts rather than harmful ones.

Finally, consider the optimal filing fee, φ^P. To derive the first-order condition that characterizes an optimum involving suits, we can differentiate expression (3) for social welfare with respect to φ^P, noting that the b^a are determined by expression (2) and π^* by expression (1).

$$(4)\quad -\frac{db^H}{d\phi^P} f^H(b^H)\left[h+\kappa^H - b^H\right] = -\gamma \frac{db^B}{d\phi^P} f^B(b^B)\left[b^B - \kappa^B\right]$$
$$+\frac{d\pi^*}{d\phi^P}\left[\left(1-F^H(b^H)\right)g^H(\pi^*) + \gamma\left(1-F^B(b^B)\right)g^B(\pi^*)\right](c^P + c^D),$$

where

$$(5)\ \kappa^a = \left(1-G^a(\pi^*)\right)(c^P + c^D).$$

A marginally higher filing fee will generate a deterrence cost, the left side of expression (4), and two types of benefits shown on the right side: reduced chilling and a cost savings from fewer cases being filed. For the deterrence cost term, $db^H/d\varphi^P$ is negative (a higher filing fee reduces the marginal type willing to commit the harmful act), and the factors before the square brackets indicate how many more individuals commit the harmful act. Each marginal harmful act committed imposes the harm h and gives rise to an expected litigation cost, indicated by κ^H, as determined by expression (5); the benefit of the act, b^H, is an offset. The first term on the right side of (4) is the reduction in the chilling cost, which has a similar interpretation (except that there is no external harm). The second term (the second row of (4)) is the savings in litigation costs from the marginal cases that are no longer filed.

Observe from expression (4) that the deterrence cost from additional harmful acts can be negative (if the benefit of the marginal harmful act exceeds the sum of the harm and the associated expected litigation costs). Similarly, the chilling benefit can be negative (if the benefit of the marginal benign act is exceeded by the associated expected litigation costs). These observations give insight into some of the conditions in Propositions 3–5, which are elaborated in section 4.C, below.

B. Extension: Optimal Evidence Threshold

This section analyzes the evidence threshold (related to the notion of a burden of proof).[17] The core model introduced in section 2 abstracted from this policy instrument, implicitly holding it constant in stating that plaintiffs' signals translated directly to probabilities of liability. The model may be extended to make evidence, and thus the operation of the evidence threshold, explicit in the following manner. (It will be apparent that this extended model collapses into the original model, with representations regarding π constituting a reduced form, when the evidence threshold is held constant.)

When an act is committed, assume now that the plaintiff observes the signal x. For each type of act, the densities and cumulative distribution functions for x are given by $z^a(x)$ and $Z^a(x)$, respectively, which are assumed to satisfy the strict monotone likelihood ratio property: $x_1 > x_0$ implies that $z^H(x_1)/z^B(x_1) > z^H(x_0)/z^B(x_0)$. At trial, the tribunal assigns liability if and only if the evidence it observes, χ, exceeds the evidence threshold, θ. The tribunal's signal is related to the plaintiff's signal (which the tribunal does not observe) as follows: $\chi = x + \varepsilon$, where ε varies according to the density function $z(\varepsilon)$ on the support $(-\infty, \infty)$.

In order to derive the analogue to expression (1) for the plaintiff's filing decision, it is useful to introduce the notation $\pi(x|\theta)$, which refers to the probability that a tribunal will assign liability when the initial signal is x, given an evidence threshold θ. Note that the liability rule $\chi > \theta$ can be expressed as $x + \varepsilon > \theta$, which is equivalent to $\varepsilon > \theta - x$. Hence,

$$(6)\ \ \pi(x|\theta) = \int_{\theta - x}^{\infty} z(\varepsilon)d\varepsilon.$$

As one would expect, for a given evidence threshold θ, $\pi(x|\theta)$ is increasing in x (the lower limit of integration falls). And, for any x, $\pi(x|\theta)$ is decreasing in θ. Expression (1), the condition for when plaintiffs sue, can now be rewritten as:

$$(7)\ \ \pi(x|\theta) > \frac{\tau^P + \phi^P + c^P}{\delta^P + \tau^P} \equiv \pi(x^*|\theta),$$

where $x^*|\theta$ is the critical value of x that generates the probability indicated by the fraction on the right side of the inequality when the evidence threshold is θ. Because, as remarked, $\pi(x|\theta)$ is increasing in x, this means that plaintiffs sue if and only if $x > x^*|\theta$. Note that the value of x^* now depends (implicitly) on four instruments — the three payments to or from plaintiffs (as before) and also the evidence threshold. The following result is proved in the appendix:

Proposition 5: If the optimum involves some suits ($\pi^* < 1$)*, defendants prefer that the marginal suit not be filed, and deterrence is welfare increasing (or chilling is welfare reducing) at the margin, then, for any finite evidence threshold, θ, it is possible to reduce θ in a manner that raises welfare.*[18]

[17]The burden (or standard) of proof, as it is understood in the legal sphere, ordinarily refers in essence to a specified Bayesian posterior probability rather than to a cutoff on the strength of the evidence (a likelihood ratio), i.e., an evidence threshold. These concepts turn out to differ in important, not widely appreciated ways. See Kaplow (2011).

[18]This result is like those in Propositions 3 and 4 in that, in each case, it is optimal to set the instrument at an extreme level. In the former Propositions, it was supposed that there was a finite boundary, whereas here there is not, which explains the different phrasing of the claim in Proposition 5.

Proposition 5 presents what may seem to be another surprising result. In a setting such as the present one, when there are benign acts, so chilling is a concern, a higher evidence threshold may seem appealing on the two aforementioned grounds: it is relatively generous with regard to weaker cases that are before the tribunal, which is favorable to benign acts, and it discourages the filing of weaker cases. But the apparent implication does not follow for the same reasons as with Proposition 4 on the optimal transfer from losing plaintiffs to winning defendants..

Specifically, our present instrument, θ, can be thought of as the legal, or de jure, evidence threshold. As mentioned, π^* (here, $\pi(x^*|\theta)$, or simply x^*) is akin to a practical, or de facto, evidence threshold. Plaintiffs only file stronger cases, those for which $\pi > \pi^*$. And if one wishes to increase this de facto evidence threshold, one can do so directly, by raising φ^P. Adjusting the de jure threshold, θ, is unnecessary if one wishes merely to adjust the de facto threshold, π^*.

But there is more: Our experiment, when reducing the de jure evidence threshold, θ, also adjusts φ^P so as to keep b^B constant. That adjustment, in turn, directly influences the de facto threshold, π^*. And here's the kicker: This second adjustment moves the de facto evidence threshold in the opposite direction from that of our manipulation of θ in the sense that it requires a stronger evidence signal before a plaintiff will be willing to sue, and this opposite force is of greater magnitude. The latter, as the proof demonstrates, is true because the change in filings needs to wash out, with regard to actors' expected costs of committing acts of type *B*, both the entire change in the filing threshold caused directly by the change in θ — which alone would produce an exact offset — and also the change in expected transfers in inframarginal cases (those filed regardless) — which requires a further opposing movement in the filing threshold.

To summarize, the net effect of our experiment on the de facto evidence threshold is opposite in direction to our movement of the de jure evidence threshold, θ. In light of this fact, we can see why the results are reversed from those we might have anticipated. The practical lesson is that, in the present model, if one wants to focus the legal system's burden most heavily on acts that generate strong signals — which are relatively more often generated by harmful acts than by benign ones — the best way to do so is to raise the de facto evidence threshold, not the de jure evidence threshold, θ. Moreover, once one does this, the latter is actually counterproductive: when the system, through informed plaintiffs' filing decisions,[19] is already confronting a smaller but stronger set of cases, it is attractive to require defendants to pay damages more often (and receive transfers from plaintiffs less often). And this is accomplished by reducing, not increasing, the de jure evidence threshold, θ.

C. Discussion

Initially, let us revisit the three conditions in Propositions 3–5. The first is that the optimum involves some suits. Clearly, if it does not, marginal adjustments to δ, τ, and θ will be inconsequential. Moreover, it is possible for an optimum to involve no suits. Specifically, suppose that *h* is very small whereas the c^j are large. In addition, assume that γ is large and the distributions of individuals' benefits from acts are such that few have low benefits. Then the

[19]This point is complementary to the argument that private enforcement has an advantage over public enforcement when potential plaintiffs information about defendants' acts is superior to the governments'. See Shavell (1993).

deterrence gain from some enforcement would be very small, whereas aggregate litigation costs would be relatively high (because so few acts are deterred or chilled). When litigation — or, more generally, enforcement — is costly and deterrence benefits are small, the optimum may involve no action. Accordingly, this article, like the pertinent prior literature, addresses optimal legal system design when some enforcement is desirable.

Second, it is assumed that defendants prefer that the marginal suit not be filed.[20] It might be supposed that such an assumption was either unnecessary or that it would in any event be satisfied at an optimum. Regarding the latter, recall that Propositions 3–5 derive conditions on individual instruments holding the others fixed at any given levels, including possibly nonoptimal ones (which may be of interest if some are institutionally constrained). Moreover, perhaps surprisingly, it turns out that there are not straightforward sufficient conditions that rule out the possibility that defendants benefit from the marginal suit. This is so despite the surprising nature of this situation: for defendants to prefer the marginal suit requires that t be high relative to other instruments and parameters (notably δ and c^D) and that plaintiffs are nevertheless willing to sue, which further requires a negative filing fee φ^P (a subsidy, perhaps a large one) to cover the net expected transfer to defendants in the marginal case as well as plaintiffs' own litigation costs. In such a scenario, it would seem appealing to raise the filing fee, which (see expression 4) would advantageously reduce litigation costs through the reduction in suits and, under these assumptions, also increase deterrence. However, chilling costs would rise as well, so without assuming that chilling is desirable at the margin, one cannot directly rule out the optimality of this scenario. Proposition 4's demonstration that the optimal τ equals 0 might be seen as another avenue to rule out this case, but the proof of that proposition itself makes use of this assumption. In any event, such a configuration would probably be unsustainable as a practical matter because it would encourage collusive suits since the parties collectively gain at the expense of the government. (See note 14.)

Third, the propositions assume (for the first branch) that deterrence is welfare increasing at the margin. For this to be so, the left side of expression (4) tells us that the sum of the harm avoided from deterring a harmful act and the expected litigation costs associated with an act exceeds the private benefit of the marginal harmful act. In simple models of law enforcement, deterrence is always beneficial at the margin at an optimum because, if it were detrimental or even neutral, reducing enforcement effort would be desirable or neutral with regard to behavior and also save enforcement resources.

In richer models such as this one, however, less enforcement (say, induced by increasing φ^P) has an additional behavioral effect, a reduction in the chilling of benign acts. For this reason, one might have supposed that it must likewise be true that deterrence is necessarily welfare increasing at the margin. (In addition, in all of our experiments, as explained at the conclusion of the proof of Proposition 3, we could instead have held deterrence, that is to say, b^H, constant, and the same sort of analysis shows that chilling, b^B, would fall, which on this account would be beneficial.) However, incremental chilling is not necessarily detrimental at an optimum. From the first term on the right side of expression (4), the social cost of chilling the marginal act is the forgone benefit minus the expected litigation cost it would have generated. Consider, then, a set of parameters and a setting of the instruments such that the chilling effect is very small, so that

[20]Note that this condition pertains to the marginal suit, which is the weakest case among those filed, not to all suits or to the average suit.

the forgone benefit b^B from the marginal chilled act is likewise small. Moreover, assume further that π^* is low and the distribution function $G^B(\pi)$ is such that benign acts often generate signals (values of π) that are above π^*. Finally, suppose that litigation costs, the c^j, are high. In that event, chilling the marginal benign act could be beneficial: little private benefit is lost, but the reduction in expected litigation costs is significant.

Taken together, we cannot a priori rule out the case in which, at an optimum, deterrence is welfare reducing at the margin and chilling is welfare increasing.[21] Note, however, that there is an important sense in which such an optimum is odd. Specifically, it would raise social welfare to reverse the outcome in some cases, as follows: for a small mass in which the evidence (see the extension in section 4.B) is extremely strong, find no liability rather than liability, and for a small mass in which the evidence is extremely weak, find liability rather than no liability.[22] This seemingly perverse adjustment to the decision rule would slightly lessen the expected burden on harmful acts (which is desirable since deterrence is imagined to be welfare reducing at the margin) and slightly increase the expected burden on benign acts (which is desirable since chilling is welfare increasing at the margin).

Having remarked on the propositions' three conditions, let us step back briefly to consider the implications of these results when they do hold, which the foregoing suggests may include most cases of practical interest. All of the propositions in this section involve extreme prescriptions, and the latter two seem particularly jarring. In this regard, however, we should recall the original implication from Becker (1968) that, in a very simple setting with public enforcement, the optimal sanction is extreme, even when the level of harm is quite low. The reaction of scholars in the field was to labor long and hard to identify factors outside the basic model — such as risk aversion, marginal deterrence, unobservable heterogeneity in ability to pay, and perverse litigant incentives — that might generate less extreme implications. See the survey by Polinsky and Shavell (2007). Much was learned from these efforts. In particular, making models more realistic in various ways brought additional, important dimensions of behavior and possible instruments to the fore. Moreover, the extensions often had additional implications, sometimes for the setting of other instruments.

It should be remarked that, relative to the handful of prior models involving optimal system design with private litigation, the results derived here were generated not by adding restrictions but by expanding the scope in a number of ways. First, benign acts were introduced, the chilling of which is an important concern when attempting to design an optimal legal system. Second and related, errors of both types were included (which little prior literature considers, particularly in connection with benign acts); relatedly, in the extension used to analyze the setting of the evidence threshold, the probability of a plaintiff victory is modeled explicitly in terms of signals (evidence), allowing for differences in what the parties and the tribunal observe. Third, the present analysis allows more instruments than does prior work.

Perhaps surprisingly, broadening the model actually contributed to the extreme results derived here. The presence of benign acts and error might, as discussed previously, have led us to suspect that rewarding victorious defendants and raising the evidence threshold would be attractive, but this is not so. In addition, introducing additional instruments — merely allowing a

---

[21]Even then, the results in the propositions would hold if the direct savings in litigation costs (from raising π^*) were sufficiently large so as to outweigh welfare reductions from changes in deterrence or chilling.

[22]And, if necessary, adjust φ^P so as to hold the filing threshold constant.

filing fee — was central in demonstrating these unexpected outcomes. Therefore, we can now see that some prior analyses that reach seemingly more moderate and intuitively appealing conclusions are fragile, the results being artifacts of restrictive assumptions that were implicit in the setup and have not previously been discussed in the literature. Therefore, although the present findings should be regarded as exploratory rather than prescriptive, they are perhaps more informative than are prescriptions in prior work that are not robust to the introduction of realistic features such as the possibility of error involving innocent behavior and the availability of additional, simple instruments.[23]

D. Additional Instrument Restrictions: When Suits Cannot Be Regulated Directly

Propositions 3–5 and the first-order condition (4) tell us how to set the instruments δ, τ, θ, and φ^P when we are able to regulate suits (the filing threshold) directly and the only other restrictions we face are a possible maximum on δ or minimum on τ. Note that these results hold even if only a single instrument (one of δ, τ, or θ) can be adjusted along with φ^P, with institutional constraints tying down the levels of other instruments. That is, none of the results on these four instruments assumed that any of the others were set optimally.

Importantly, section 2.C explained how suits (the filing threshold, π^*) can alternatively be regulated — without any feedback effects on the b^a — through φ^P, the decoupling of δ (the ability to set δ^P independently of δ^D), or the decoupling of τ (the ability to set τ^P independently of τ^D). This flexibility matters because, just as we may have a constraint like $\delta \leq \delta^{Max}$, due perhaps to limits on defendants' wealth, so there might be an upper limit on φ^P (and, if relevant, on $\varphi^P + \tau^P$) due to limits on plaintiffs' wealth. Moreover, many actual legal systems do not impose significant filing fees. However, even with φ^P fixed, say, at 0, we can instead induce as high a π^* as we wish by reducing δ^P below δ^D. Furthermore, plaintiffs' wealth poses no constraint, for setting $\delta^P = 0$ is sufficient when, say, $\tau = 0$, to guarantee that no suits are filed. Hence, any limitation on the results in Propositions 3–5 are due to constraints outside the model and, in particular, other than limitations on plaintiffs' wealth.

For these reasons, consider briefly more restricted systems in which it is not possible to regulate plaintiffs' filing decisions directly. Specifically, assume that φ^P is fixed — say, at 0 — and that there is no decoupling of either sort. Furthermore, continue to assume that $\varphi^D = 0$. In such a regime, results like those in Propositions 3–5 do not arise. Sparing the reader extensive derivations, a number of observations can be offered. To begin, raising (or lowering) each of our remaining instruments, δ, τ, and θ, one at a time, presents the welfare tradeoffs like those depicted in the first-order condition (4) for the optimal φ^P: Raising δ hurts both types of

[23] It is also useful to compare this section's results to Proposition 2, where it was shown how, without restrictions on the magnitudes of the instruments, it is feasible to implement a result arbitrarily close to the first best. Regarding Proposition 3, that it is optimal to raise δ ever higher (until limited by any maximum, δ^{Max}), the construction showing how to implement Proposition 2's result was much the same. The connection with Proposition 4, that it is optimal to reduce τ, is less apparent. The supposition for this proposition is that deterrence is welfare increasing or chilling is welfare reducing at the margin, but neither is true with respect to Proposition 2, for, if there are no restrictions on instruments, both deterrence and chilling are at first-best levels ($b^H = h$ and $b^B = 0$). Additionally, the proof of Proposition 4 entails that reducing τ enables a reduction in litigation costs, but when there is no constraint on δ, we can already reduce them arbitrarily close to zero. Proposition 5 indicates that, with an interior solution, it is always optimal to reduce the evidence threshold θ (on the same grounds that it is optimal to reduce τ). For the reasons just stated, this too is unnecessary in demonstrating Proposition 2.

defendants and helps plaintiffs, so it increases deterrence and chilling as well as the level of lawsuits. That is, b^H and b^B rise, and π^* falls. The latter is always welfare reducing, ceteris paribus. As discussed just above, enhanced deterrence tends to be good and additional chilling bad, although either or both could be the opposite. The only thing that we can say with confidence regarding the optimum in the present setting is that it cannot be true that both deterrence and chilling are undesirable: If they were, then reducing δ would unambiguously increase social welfare. The effects of raising τ and the evidence threshold θ are, of course, the opposite in all respects from those of raising δ, so analogous reasoning applies

We can also ask whether further characterizations might be possible with the sorts of policy experiments we have examined previously wherein two instruments are simultaneously adjusted in particular, coordinated ways. For example, we might increase δ while simultaneously increasing τ by just the amount that keeps chilling, b^B, fixed. When this is done, it is possible to show that π^* rises. So far, chilling is unchanged and litigation costs are saved. However, the sign of the effect on deterrence, the level of b^H, is ambiguous, and no simple parameter restrictions remove the ambiguity. For the other two pairings of instruments — raising δ and θ in a manner that holds b^B constant, and raising τ while reducing θ in a manner that holds b^B constant — it is not possible to sign the effect on either π^* or b^H. In summary, even when we consider a broader array of policy experiments, no additional, sharp characterizations can be offered.

5. Conclusion

In a model with endogenous filing decisions, court error, and two types of behavior (harmful and benign), this article analyzes the optimal levels of filing fees, damages, and payments by losing plaintiffs — each of which may differ for each party — as well as the stringency of the evidence threshold. If there are no constraints, three instruments suffice to span a broad range of possible outcomes for the three endogenous behaviors: the commission of harmful and benign acts, and filing decisions. In particular, one can achieve an outcome arbitrarily close to the first best. Additional analysis examines instrument redundancy and the manner in which the present model incorporates prior work on decoupling and fee shifting as special cases.

The remainder of the article explores the optimal choice of particular instruments when there are restrictions of various sorts. As long as there is some free instrument that can regulate suits directly, such as a filing fee charged to plaintiffs, each other instrument considered is optimally set at an extreme level: the damage award as high as possible, the transfer paid by losing plaintiffs as low as possible, and the evidence threshold as low as possible. The core explanation for the latter two (more surprising) results is that informed plaintiffs' self-interest in making filing decisions already entails selection of the strongest cases. As a consequence, a filing fee can target a cutoff for suit that serves as something akin to a de facto evidence threshold, and it is demonstrated that greater generosity toward plaintiffs in filed cases, which are all stronger than the marginal case (just at the filing threshold) — even through reducing the de jure evidence threshold at trial — enables a tougher de facto threshold up front, the combination of which is relatively favorable to benign acts versus harmful ones. These results, however, may change when additional external constraints are imposed on the availability of instruments, which seems quite common in actual legal systems although for reasons that often are not apparent.

Many of the present conclusions differ qualitatively from those in the small body of prior literature on optimal legal system design with private enforcement, with the divergences partly attributable to unappreciated implicit instrument restrictions in previous models. It is also notable that the allowing for two types of court error that influence two types of behavior — including, importantly, the possible chilling of benign conduct — fundamentally changes much of the analysis but need not generate the sorts of results that one might have anticipated, such as the optimality of making losing plaintiffs pay or utilizing tougher evidence thresholds.

Although the present framework is more general and encompassing than those in prior work, the results are best viewed as outlining a partial and preliminary conceptual understanding of optimal legal system design with private enforcement. First, a number of familiar complications considered in the literature on public enforcement, such as risk aversion, socially costly sanctions, and marginal deterrence would influence the conclusions, perhaps in familiar ways. See notes 4 and 10, the survey by Polinsky and Shavell (2007), and Kaplow (2011).

Second, although the analysis here illustrates the importance of explicitly modeling signals of case strength, more can be examined regarding what is observed by plaintiffs and by defendants; how that may differ from what is observed by the tribunal; and how these signals are generated by different types of behavior.[24] Of particular interest would be variations where parties' initial signals differed from each other's, which would introduce asymmetric information that would allow for more interesting analysis of settlement and pre-trial discovery.[25]

Third, the signal observed by the tribunal could be made a function not only of parties' initial signals but also of their endogenously chosen expenditures on developing and presenting evidence.[26] In addition to introducing a source of asymmetric information that may influence settlement, this endogenous process will be influenced by how various instruments are set and in turn will influence decisions on liability. As in Katz (1987), higher damages or higher transfers paid by losing plaintiffs to victorious defendants would raise the stakes at trial, inducing greater expenditures and thus higher costs per case, which in itself would reduce social welfare. On the other hand, if greater expenditures enhance accuracy, as in Kaplow and Shavell (1994), the net effect on social welfare could be positive because the increased diagnosticity of liability findings makes it possible to achieve a given level of deterrence with fewer suits and also less chilling of benign behavior.

[24]One avenue would be to explore (in the extension appearing in section 4.B) the value of accuracy of the tribunal's noisy observation of the evidence signal, a question examined by Kaplow and Shavell (1994) in a simpler setting. Preliminary analysis has not yielded sharp results. Another variation of that extension would be to model the tribunal as observing a sharper signal and the parties observing a noisier signal (perhaps the process of trial yields greater clarity, although the opposite assumption employed in section 4.B, that the tribunal never learns as much as the parties observe directly, may be more realistic in many settings). A conjecture is that some results will differ because the tribunals' final decisions are more diagnostic than plaintiffs' filing decisions. Also, plaintiffs who observe a noisier signal would need to form beliefs about the underlying distribution of what the tribunal will subsequently observe, and rational beliefs will reflect Bayesian inferences about underlying primary behavior, which is endogenous.

[25]On settlement with symmetric information in the present model, see note 6.

[26]Another sort of extension would introduce additional dimensions of heterogeneity, such as in parties' litigation costs (which makes selection at the time of filing contingent on particular plaintiffs' costs in addition to their signals of case strength, as in Kaplow 1993) or in types of underlying behavior wherein the tribunal (or perhaps plaintiffs as well) cannot observe pertinent differences.

References

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76: 169-217.

Kaplow, Louis. 1993. "Shifting Plaintiffs' Fees Versus Increasing Damage Awards," *Rand Journal of Economics* 24: 625–30.

__________. 2011. "On the Optimal Burden of Proof," *Journal of Political Economy* 119: 1102–40.

__________ and Steven Shavell. 1994. "Accuracy in the Determination of Liability." *Journal of Law and Economics* 37: 1–15.

Katz, Avery. 1987. "Measuring the Demand for Litigation: Is the English Rule Really Cheaper?" *Journal of Law, Economics, and Organization* 3: 143–76.

Kessler, Daniel P., William M. Sage, and David J. Becker. 2005. "Impact of Malpractice Reforms on the Supply of Physician Services." *Journal of the American Medical Association* 293: 2618–25.

Lin, Hui Ling, Kuntara Pukthuanthong, and Thomas John Walker. 2013. "An International Look at the Lawsuit Avoidance Hypothesis of IPO Underpricing." *Journal of Corporate Finance* 19: 56–77.

McAfee, R. Preston, Hugo M. Mialon, and Sue H. Mialon. 2008. "Private v. Public Antitrust Enforcement: A Strategic Analysis." *Journal of Public Economics* 92: 1863–1875.

Png, Ivan P.L. 1986. "Optimal Subsidies and Damages in the Presence of Judicial Error," *International Review of Law and Economics* 6: 101–05.

Polinsky, A. Mitchell, and Yeon-Koo Che. 1991. " Decoupling Liability: Optimal Incentives for Care and Litigation," *Rand Journal of Economics* 22: 562–70.

__________ and Daniel L. Rubinfeld. 1996. "Optimal Awards and Penalties When the Probability of Prevailing Varies Among Plaintiffs," *Rand Journal of Economics* 27: 269–80.

__________ and Steven Shavell. 1979. "The Optimal Tradeoff Between the Probability and Magnitude of Fines," *American Economic Review* 69: 880–91.

__________ and __________. 1989. "Legal Error, Litigation, and the Incentive to Obey the Law," *Journal of Law, Economics, and Organization* 3: 99–108.

__________ and __________. 2007. "The Theory of Public Enforcement of Law," in A. Mitchell Polinsky and Steven Shavell, *Handbook of Law and Economics*, vol. 1: 403–54, Amsterdam: North-Holland.

Shavell, Steven. 1982. "The Social versus the Private Incentive to Bring Suit in a Costly Legal

System," *Journal of Legal Studies* 11: 333–39.

__________. 1993. "The Optimal Structure of Law Enforcement," *Journal of Law and Economics* 36: 255–87.

Spier, Kathryn E. 2007. "Litigation," in A. Mitchell Polinsky and Steven Shavell, *Handbook of Law and Economics*, vol. 1: 259–342, Amsterdam: North-Holland.

Studdert, David M., et al. 2005. "Defensive Medicine Among High-Risk Specialist Physicians in a Volatile Malpractice Environment." *Journal of the American Medical Association* 293: 2609–17.

Appendix

Proof of Proposition 1: First , taking as given any initial values of δ and τ, we can choose φ^P to induce any given $\pi^* \in [0,1)$. For any admissible δ + τ, we can induce $\pi^* = 0$ by choosing $\varphi^P = -(\tau + c^P)$, and we can induce $\pi^* = 1$ by choosing $\varphi^P = \delta - c^P$. Because π* rises continuously with φ^P, we can span the interval [0,1] and thus induce any π*. The remainder of the analysis, which holds π* constant when we are changing δ or τ, entails adjusting φ^P the requisite amount. (Because φ^P only influences the b^a through its effect on π*, there are no feedbacks on the b^a that need to be taken into account.)

Second, taking π* as given in the manner just described, we can implement any chosen level of b^H. Using our current instrument set, we can use expression (2) to state

$$(A1)\ b^a = \int_{\pi^*}^{1} \left[c^D + \pi\delta - (1-\pi)\tau\right] g^a(\pi)d\pi.$$

Starting from a given φ^P, δ, and τ, expression (A1) determines a value of b^H. Raising δ increases b^H (at a constant rate). Accordingly, we can implement any b^H above its initial value. To achieve lower values of b^H, we can correspondingly reduce δ. However, in this instance we have our constraint that δ + τ > 0, so it might not be possible to drive b^H down to 0. If we cannot, we can instead fix δ and increase τ: this will decrease b^H at a constant rate, allowing us to span all values down to zero.

Third is implementation of a specified b^B. Start with a given π* and b^H, supported by some given φ^P, δ, and τ, which, from expression (A1), implies some value of b^B. Consider a policy experiment, parameterized by α, that manipulates b^B by adjusting τ. We will need, however, simultaneously to adjust δ so as to keep b^H fixed because changes in τ also influence b^H. Specifically, letting τ = α and recognizing that δ is now a function of α, we can derive how this function behaves as follows:

$$(A2)\ \frac{db^H}{d\alpha} = \int_{\pi^*}^{1} \left[\pi \frac{d\delta}{d\alpha} - (1-\pi)\right] g^H(\pi)d\pi = 0.$$

This implies:

$$(A3)\ \frac{d\delta}{d\alpha} = \frac{\int_{\pi^*}^{1} (1-\pi) g^H(\pi)d\pi}{\int_{\pi^*}^{1} \pi g^H(\pi)d\pi} > 0.$$

Note further that the value of this derivative is constant, given that we are holding π* constant.

Like b^H, b^B also rises in δ and falls in τ. Accordingly, our experiment — which involves, let us say, increasing τ, which implies from expression (A3) that δ also increases — will have conflicting effects on b^B. We will now demonstrate that increasing τ in this experiment causes b^B to fall because increasing τ is relatively favorable to type *B* (benign) acts whereas increasing δ is relatively unfavorable to type *H* (harmful) acts. First, observe that

$$(A4)\ \frac{d\delta}{d\alpha}\int_{\pi^*}^{1}\pi\frac{g^H(\pi)}{1-G^H(\pi^*)}d\pi > \frac{d\delta}{d\alpha}\int_{\pi^*}^{1}\pi\frac{g^B(\pi)}{1-G^B(\pi^*)}d\pi.$$

The leading derivative on each side, as indicated by expression (A3), is positive. Regarding the integrands, note that the choice of denominators implies that we are integrating over the same total mass. Moreover, because the strict monotone likelihood ratio property implies first-order stochastic dominance (a sufficient condition for this proof), the normalized density on the left side, $g^H(x)/(1-G^H(\pi^*))$, is an upward shift relative to that on the right side, $g^B(\pi)/(1-G^B(\pi^*))$. Finally, π is positive and increasing over the range of integration. Therefore, the inequality holds. The same logic implies that

$$(A5)\ \int_{\pi^*}^{1}(\pi-1)\frac{g^H(\pi)}{1-G^H(\pi^*)}d\pi > \int_{\pi^*}^{1}(\pi-1)\frac{g^B(\pi)}{1-G^B(\pi^*)}d\pi.$$

Here, the term $\pi - 1$ is negative, but it is also increasing in π. Because the normalized density on the left is an upward shift, it places more weight on the less negative values, so we again have the same inequality. Next, let us sum the two terms on the left sides of expressions (A4) and (A5) and compare that total to the sum of the two terms on the right sides:

$$(A6)\ \frac{d\delta}{d\alpha}\int_{\pi^*}^{1}\pi\frac{g^H(\pi)}{1-G^H(\pi^*)}d\pi + \int_{\pi^*}^{1}(\pi-1)\frac{g^H(\pi)}{1-G^H(\pi^*)}d\pi$$

$$> \frac{d\delta}{d\alpha}\int_{\pi^*}^{1}\pi\frac{g^B(\pi)}{1-G^B(\pi^*)}d\pi + \int_{\pi^*}^{1}(\pi-1)\frac{g^B(\pi)}{1-G^B(\pi^*)}d\pi.$$

Comparison of the left side of inequality (A6) with expression (A2) indicates that the former simply equals $(db^H/d\alpha)/(1-G^H(\pi^*))$. Likewise, the right side of inequality (A6) equals $(db^B/d\alpha)/(1-G^B(\pi^*))$. Therefore, we can write:

$$(A7)\ \frac{1-G^B(\pi^*)}{1-G^H(\pi^*)}\frac{db^H}{d\alpha} > \frac{db^B}{d\alpha}.$$

Because our experiment adjusts δ as we increase τ so as to hold b^H constant, the left side of expression (A7) equals 0. Hence, the right side is negative, which is to say that b^B falls as τ rises (and δ and φ^P are adjusted as specified). Furthermore, from inspection of the right side of expression (A6), it is apparent that the magnitude of this derivative is constant (π^* is held constant; neither δ nor τ appears anywhere; and, from expression (A3), $d\delta/d\alpha$ is constant). Therefore, as asserted above, we can raise τ to achieve any level of b^B from its initial level to as low as we would like. We can also lower τ (and δ, etc.), to achieve higher levels of b^B, but in this instance, the constraint that $\delta + \tau > 0$ will at some point bind.[27] This establishes Proposition 1.a.

[27] To offer a concrete illustration of these points, suppose that we have $\pi^* = 0$, that initially $\tau = 0$, and we take a positive value of δ that is arbitrarily close to 0. This implies that $b^H \approx c^D(1-G^H(0))$. Likewise, we have an initial value of $b^B \approx c^D(1-G^B(0))$, which, note, is lower than b^H. How much can we raise b^B, keeping in mind that we must keep both π^* and b^H constant? To raise b^B, we need to lower τ, and as per our experiment, δ as well. But we are bound by the constraint that $\delta + \tau > 0$, which from our initial conditions is almost binding already. Hence, this initial value of b^B is the approximate upper bound.

Finally, note that we could have undertaken our spanning construction in a different order, as follows: First, as before, we could peg π^*. Second, given this π^*, we could (setting b^H to the side), implement any level of b^B we wished. Third, given π^*, b^B, and an initial level of b^H, we can raise δ, adjusting τ so as to keep b^B fixed. This will allow us to achieve as high a level of b^H as we like. However, if we wish to implement lower levels of b^H, then we need to reduce δ and τ, and at some point our constraint will bind.[28] This establishes Proposition 1.b.■

Proof of Proposition 3: Suppose we are at an optimum in which $\delta < \delta^{Max}$ (and the other conditions hold). Raise δ while adjusting φ^P so as to keep b^B fixed. Parameterize this policy experiment by α, set $\delta = \alpha$, and also take φ^P to be a function of α. The effect of this experiment on social welfare (3) is given by:

$$(A8)\ \frac{dW}{d\alpha} = \frac{db^H}{d\alpha}\left(\left[h + \left(1 - G^H(\pi^*)\right)(c^P + c^D) - b^H\right] f^H(b^H)\right)$$
$$+ \int_{b^H}^{\infty} g^H(\pi^*)\frac{d\pi^*}{d\alpha}(c^P + c^D) f^H(b)db$$
$$-\gamma \frac{db^B}{d\alpha}\left(\left[b^B - \left(1 - G^B(\pi^*)\right)(c^P + c^D)\right] f^B(b^B)\right)$$
$$+\gamma \int_{b^B}^{\infty} g^B(\pi^*)\frac{d\pi^*}{d\alpha}(c^P + c^D) f^B(b)db.$$

The first term on the right side shows the effect on welfare due to the induced change in deterrence, the third term is the effect on welfare due to increased chilling, and the second and fourth terms indicate the change in the cost of suits due to any change in π^*.

Next, we can inquire into how this experiment influences behavior, specifically, the b^a. By construction, it does not affect chilling. By examining the chilling effect and setting it equal to zero, we will be able to sign both $d\pi^*/d\alpha$ and $db^H/d\alpha$. Effects on the b^a are determined by differentiating expression (A1):

$$(A9)\ \frac{db^a}{d\alpha} = -\frac{d\pi^*}{d\alpha}\left[c^D + \pi^*\delta - (1-\pi^*)\tau\right] g^a(\pi^*) + \int_{\pi^*}^{1} \pi g^a(\pi)d\pi.$$

The first term indicates that, for a given increase in π^* due to the experiment, $d\pi^*/d\alpha$, the b^a will fall, which is to say that behavior will be encouraged. The second term reflects that, for the postulated increase in δ, all suits that are filed in any event now have a higher sanction being applied, which sanction is imposed with probability π.

If our experiment had simply increased δ and left φ^P unchanged, we know that the b^a, and, in particular, b^B, would rise on both accounts: from expression (1), a higher δ, ceteris paribus, reduces π^*, making the first term in expression (A9) positive, and the second term is positive. Therefore, if our experiment is to hold b^B constant, it must raise π^*. That is, φ^P must be elevated by enough not only to offset the reduction in π^* directly caused by the increase in δ

[28] Consider a simple variation of the example in the preceding footnote.

(which would make the first term in (A9) zero, leaving the positive second term), but also to actually raise π^* by just enough to generate a negative first term with a magnitude equal to that of the (positive) second term. Because $d\pi^*/d\alpha > 0$, the experiment results in positive values for the second and fourth terms in expression (A8) for the change in social welfare.

We can now show that $db^H/d\alpha > 0$. The core idea is that the influence of raising δ due to its direct effect on expected sanctions will be relatively greater for harmful acts than for benign acts; hence, deterrence rises relative to chilling; and, since chilling is unchanged, deterrence therefore rises absolutely. The subtlety concerns the term "relatively," which can be made explicit through a particular normalization, specifically, with respect to the values of the densities, $g^a(\pi)$, when each are evaluated at π^*. To begin, consider the validity of the following inequality:

$$(A10)\quad \int_{\pi^*}^{1} \left[\frac{g^H(\pi) \big/ g^B(\pi)}{g^H(\pi^*) \big/ g^B(\pi^*)} \right] \pi g^B(\pi)d\pi > \int_{\pi^*}^{1} \pi g^B(\pi)d\pi.$$

The term in brackets on the left side is the likelihood ratio evaluated at π divided by the likelihood ratio evaluated at π^*. For $\pi > \pi^*$, the strict monotone likelihood ratio property indicates that this term exceeds 1. Furthermore, π and $g^B(\pi)$ are both positive. Hence, the integrand on the left side exceeds that on the right side for all $\pi > \pi^*$, which establishes the inequality. Next, rearrange expression (A10), multiplying both sides by $1/g^B(\pi^*)$ and cancelling the $g^B(\pi)$'s on the left side to yield:

$$(A11)\quad \int_{\pi^*}^{1} \pi \left[\frac{g^H(\pi)}{g^H(\pi^*)} \right] d\pi > \int_{\pi^*}^{1} \pi \left[\frac{g^B(\pi)}{g^B(\pi^*)} \right] d\pi.$$

Now, add a common term to each side:

$$(A12)\quad -\frac{d\pi^*}{d\alpha}\left[c^D + \pi^*\delta - (1-\pi^*)\tau\right] + \int_{\pi^*}^{1} \pi \left[\frac{g^H(\pi)}{g^H(\pi^*)} \right] d\pi >$$
$$-\frac{d\pi^*}{d\alpha}\left[c^D + \pi^*\delta - (1-\pi^*)\tau\right] + \int_{\pi^*}^{1} \pi \left[\frac{g^B(\pi)}{g^B(\pi^*)} \right] d\pi.$$

Finally, multiply both sides by $g^H(\pi^*)$ and, on the right side, factor out $1/g^B(\pi^*)$, to yield:

$$(A13)\quad -\frac{d\pi^*}{d\alpha}\left[c^D + \pi^*\delta - (1-\pi^*)\tau\right] g^H(\pi^*) + \int_{\pi^*}^{1} \pi g^H(\pi)d\pi >$$
$$\frac{g^H(\pi^*)}{g^B(\pi^*)}\left(-\frac{d\pi^*}{d\alpha}\left[c^D + \pi^*\delta - (1-\pi^*)\tau\right] g^B(\pi^*) + \int_{\pi^*}^{1} \pi g^B(\pi)d\pi \right).$$

Expression (A13) establishes our result, as follows: From expression (A9), we can see that the left side is $db^H/d\alpha$ and the term in large parentheses on the right side is $db^B/d\alpha$. Moreover, our experiment adjusts φ^P as we change α so that $db^B/d\alpha = 0$. Therefore, $db^H/d\alpha > 0$.

To summarize, the experiment of raising δ while adjusting φ^P so as to keep b^B constant raises b^H (deterrence) along with reducing the level of suits (because holding b^B fixed when δ is increased requires increasing φ^P sufficiently to increase π^*). The argument is complete if deterrence is welfare increasing at the margin. If it is not, then Proposition 3's alternative condition holds: chilling is welfare reducing at the margin. In that case, simply alter the experiment such that, instead of having $db^B/d\alpha = 0$, we adjust φ^P such that $db^H/d\alpha = 0$. Then everything follows until the last sequence of analysis, which now implies that $db^B/d\alpha < 0$, which completes the argument for this alternative premise.■

<u>Proof of Proposition 4</u>: The argument closely follows that given for Proposition 3, so the presentation will be abbreviated. Suppose that $\tau > 0$. Our analogous experiment is to lower τ while adjusting φ^P so as to keep b^B fixed. However, to ease the exposition, let us instead analyze an experiment that increases our variable of interest, τ, while adjusting φ^P so as to keep b^B fixed. We will show that this experiment unambiguously reduces social welfare, implying that the reverse experiment, in which we indeed lower τ, raises welfare.

The effect of this experiment on social welfare (3) is given by an expression identical to expression (A8), but in our parameterization using α we now set $\tau = \alpha$. Accordingly, we can proceed directly to analyze how the experiment influences behavior. Differentiating expression (A1), we have:

$$(A14)\ \frac{db^a}{d\alpha} = -\frac{d\pi^*}{d\alpha}\left[c^D + \pi^*\delta - (1-\pi^*)\tau\right]g^a(\pi^*) - \int_{\pi^*}^{1}(1-\pi)g^a(\pi)d\pi.$$

Comparing expression (A14) to expression (A9), note that the second term is subtracted rather than added because raising τ rewards actors whereas raising δ punishes them. In addition, the direct effect of raising τ on π^* is to increase it (whereas when raising δ, the direct effect was to decrease π^*). Therefore, raising τ by itself reduces b^B on account of both terms in expression (A14). Accordingly, to hold b^B constant we need more suits: a lower π^*, which is implemented by reducing φ^P. As a consequence, this experiment raises litigation costs.

To determine the effect on b^H, we proceed as before. Because the analysis is virtually identical, it is largely omitted. The key difference concerns the analogue to inequality (A10). The corresponding inequality indeed holds. (It suffices to observe that, just as π being positive was used to establish the previous inequality, the fact that $1-\pi$ is positive can be used to establish the analogous inequality here.) However, here we wish to be characterizing the second term in expression (A14), which is subtracted (whereas the corresponding term in expression (A9) was added). Hence, to implement the preceding series of manipulations (that ultimately must mimic, on each side of the inequality, what is now expression (A14)), we need first to multiply both sides by negative one. This reverses the inequality. Therefore, when the corresponding analysis is complete, it shows that the experiment of increasing τ while holding b^B constant causes deterrence (b^H) to fall.

Taken together, the experiment that increases τ raises litigation costs and reduces deterrence, both of which are welfare decreasing. Hence, the reverse experiment, involving a reduction in τ, raises social welfare.[29]■

<u>Proof of Proposition 5</u>: The argument follows very closely that given for Proposition 4. Our

[29]For this proposition and the next, the proofs omit the final step in Proposition 3's proof for the alternative condition under which chilling is welfare reducing at the margin.

experiment is to lower θ while adjusting φ^P so as to keep b^B fixed. Again, it will ease the exposition to consider an experiment that instead increases θ while adjusting φ^P so as to keep b^B fixed. We will show that this experiment reduces social welfare, implying that the reverse experiment, in which we lower θ, raises welfare. The effect of this experiment on social welfare (3) is given by an expression analogous to expression (A8); in our parameterization using α, we now set θ = α. Accordingly, we can proceed directly to analyze how the experiment influences behavior.

Differentiating a close analogue to expression (A1), in which the variable of integration is now x rather than π, keeping in mind that π is now a function of x as well as, implicitly, a function of θ, we have:

$$(A15)\frac{db^a}{d\alpha} = -\frac{d(x^*|\theta)}{d\alpha}\left[c^D + \pi(x^*|\theta)\delta - \left(1-\pi(x^*|\theta)\right)\tau\right]z^a(x^*|\theta)$$
$$-\int_{x^*|\theta}^{\infty} z(\theta - x)(\delta+\tau)z^a(x)dx.$$

Expression (A15) is quite similar to expression (A14). A few comments are in order. Regarding the first term, we again have a situation in which the direct effect of increasing our variable of interest, here θ, is to increase $x^*|\theta$ (as noted, an analogue to π^* in the base model). Raising the hurdle for victory at trial makes suit less attractive. Regarding the second term, the $\delta + \tau$ in the integrand reflects that the stakes in a case — the amount parties' prospects change depending on whether liability is found — are the sum of the two transfers. From expression (6), we can see that raising θ changes the probability of liability $\pi(x|\theta)$ by $-z(\theta - x)$, which indicates that the probability falls. This explains why the second term is also negative. As a result, the direct effect of raising θ is to reduce b^B on account of both terms in expression (A15). Therefore, if we wish to hold b^B constant, we need more suits: a lower $x^*|\theta$, which is implemented by reducing φ^P. Therefore, this experiment raises litigation costs.

To determine the effect on b^H, we perform the same sort of manipulations as before. The pertinent integrand, that in the second term of expression (A15), is again positive. As with the proof of Proposition 4 (expression A14) but unlike in the case of Proposition 3 (expression A9), this term is subtracted, so a slight modification of our previous analysis indicates that the experiment of increasing θ while holding b^B constant causes deterrence (b^H) to fall.

Taken together, the experiment that increases θ raises litigation costs and reduces deterrence, both of which are welfare decreasing. Hence, the reverse experiment, involving a reduction in θ, raises social welfare.■