Should Firms Be Allowed to Indemnify their Employees for Sanctions?

Wallace P. Mullin

*Department of Economics, George Washington University, Washington, DC 20052.*
*Email: wpmullin@gwu.edu.*

Christopher M. Snyder

*Department of Economics, Dartmouth College, Hanover, NH 03755.*
*Email: Christopher.M.Snyder@dartmouth.edu.*

June 18, 2008

**Abstract:** Policymakers have questioned whether firms should be allowed to indemnify their employees for personal sanctions for corporate crimes. This paper provides the first formal analysis of this form of indemnification. Indemnification should typically not be banned because of the relative benefits it offers to law-abiding firms: insuring employees against the risk of mistaken government prosecution, strengthening employees’ resolve to fight meritless suits. Only in limited circumstances do we find banning indemnification to be socially efficient, in particular providing prosecutors with the leverage to adjust the employee’s sanction in exchange for his cooperation against the firm.

**Keywords:** indemnification, corporate crime

**JEL Classification:** K22, D82, L20

**Acknowledgements:** The authors thank Nuno Garoupa, Daniel Klerman, Giacomo Ponzetto, Jennifer Reinganum, and seminar participants at Charles River Associates, Clemson University, Georgetown University, Harvard Law School, Hebrew University of Jerusalem, the Allied Social Sciences Association “Contracts and Organizations” session, the International Industrial Organization Conference (Chicago), M.I.T., Michigan State University, the N.B.E.R. Law and Economics Summer Institute, and the Securities and Exchange Commission for helpful comments.
1. Introduction

Sanctions may be levied on both a firm and its agents for violations of securities, antitrust, environmental, bribery, safety, and other laws. The incorporation laws of most U.S. states allow firms to reimburse agents’ legal costs and losses from settlements, judgments, and fines. Delaware law grants incorporating firms a broad ability to insure their agents, either through direct indemnification payments from the firm itself or through third-party Director and Officer (D&O) insurance (Easterbrook and Fischel 1991), even allowing corporations to include mandatory indemnification in their corporate charters or bylaws. According to a recent survey, 98 percent of U.S. firms with over 500 shareholders had D&O insurance (Tillinghast-Towers Perrin 2002).

While coverage under indemnification and D&O insurance is broad, there are exceptions. State laws forbid indemnification and D&O insurance coverage in the case of willful criminal misconduct (Harrington and Niehaus 1998). However, Stone (1980) argues that such de jure exclusions do not prevent de facto coverage for willful criminal misconduct. A number of federal crimes require only limited or no proof of intent or knowledge. State laws specify that conviction for such crimes “shall not, of itself, create a presumption that the person did not act in good faith” (Stone 1980, p. 49). Where to draw the line between indemnifiable and unindemnifiable actions is thus an important theoretical and public-policy question.

The debate over indemnification has been active. William Donaldson, then Chairman of the U.S. Securities and Exchange Commission (SEC), indicated: “I’m concerned about companies that, under permissive state laws, indemnify their officers and directors against disgorgement and penalties ordered by law enforcement agencies, including those brought by the Commission. In my mind, this just isn’t good public policy.”¹ This statement drew criticism from former SEC official Stanley Sporkin: “For the SEC to come out and say you can’t get insurance for these things, I think they are going pretty far.”²

In theory, a benevolent government authority that is a perfect law enforcer should divide sanctions between the firm and agent in just the right way to obtain optimal deterrence. Indemnification allows the firm to undo this balance by transforming employee sanctions into \textit{de facto} firm sanctions, perhaps impairing deterrence. This argues in favor of a wholesale ban on indemnification. In practice, however, indemnification is not only legal but ubiquitous, suggesting it provides some important social benefit.

This paper addresses the apparent contradiction between theory and practice as the first formal analysis of employee indemnification. The key is departing from the assumption that the government is perfect and benevolent. In the first of two related models, labeled the corporate-crime model, analyzed in Sections 2–5, we assume the authority is benevolent but is an imperfect law enforcer. The possibility of type-I enforcement errors—i.e., convicting law-abiding firms with some probability—provides a reason for even law-abiding firms to indemnify employees. This shifts the risk of sanctions from the high-cost bearer—the risk-averse agent—to the low-cost bearer—the risk-neutral principal.

It has been postulated (Stone 1980, Kraakman 1984, Privileggi, Marchese, and Cassone 2001) that banning indemnification magnifies the frictions in the principal-agent relationship, increasing the operating costs of a criminal firm. Our formal analysis confirms this postulate, but also demonstrates that the apparent policy implication, to ban indemnification, is erroneous. In our model, the government authority can always deter crime with a sufficiently high combination of fines on the firm and employee. The challenge is to deter crime at minimum social cost. We show that deterrence can typically be obtained at minimum social cost by sanctioning the firm alone. This maintains deterrence without exposing the agent to risk from sanctions or inducing the exit of productive, law-abiding firms.

Sanctioning the agent is valuable in limited circumstances. If deterrence is especially difficult, it may be optimal to hit the agent with a sanction large enough to bankrupt him. Although the \textit{de jure} sanctions cannot vary with actual guilt—imperfect enforcement prevents this—bankrupting the agent allows the \textit{de facto} agent sanction to vary with his wealth. The agent needs to be
paid a premium to induce him to commit a crime, and so the agent of the criminal firm ends up having more wealth to be seized than the agent of a law-abiding firm. Indemnification need not be explicitly banned for this strategy to work: the agent’s sanction can be set so high that the firm would not choose to indemnify the agent even if allowed by law.

Indeed, if sanctions are set appropriately, the government’s policy toward indemnification becomes moot. Either the agent should not be sanctioned at all, in which case there is nothing for the firm to indemnify, or the agent should be sanctioned so harshly that the firm chooses not to indemnify the agent even if it could. The government’s policy toward indemnification is not moot in an extension of the corporate-crime model in Section 5 in which the agent’s cooperation can help convict a criminal firm. The authority can offer to reduce the employee’s fine in return for his cooperation, an offer the firm can unravel by pledging to indemnify him fully.

In a second model, analyzed in Section 6, we assume that the action against the firm is a nuisance suit undertaken by a rent-seeking party (private or government) rather than a benevolent authority. In this nuisance-suit model, indemnification provides a strategic benefit to law-abiding firms. Absent indemnification, the firm’s director or officer would have an excessive incentive to settle the case to avoid personal liability. The party making the decision to settle is not a perfect agent of the firm because he bears 100 percent of his personal liability but only a share at best of the firm’s profits. This principal-agent friction is a chink in the firm’s armor which the plaintiff in a nuisance suit can exploit by aggressively targeting the agent. This vulnerability would encourage the entry of further baseless suits, potentially in unlimited supply. Such nuisance suits could lead to social-welfare losses from mounting court costs and exit of productive enterprises.

Section 7 provides a literature review, and Section 8 concludes with a discussion of the implications for recent controversies.

2. Corporate-Crime Model

The model has three players. Within the firm, there is a principal and an agent. The principal is the residual claimant of profit who designs the agent’s incentive scheme. The agent carries
out activities within the firm, including the possibility of committing a criminal act. We call
the principal simply the “firm” and the agent simply the “employee.” The third player is the
government, which sets and enforces sanctions against corporate crime.

The employee chooses action \( a \in \{0, 1\} \), an indicator for whether a crime is committed
\((a = 1)\) or not \((a = 0)\). Let \( c(a) \) be his disutility from working in the firm, with \( c(0) = 0 \) and
\( c(1) = C \). Thus \( C \) represents his cost of committing the crime, including any physical effort
required plus any psychic costs of violating a personal ethical code. Let \( r(a) \) be the firm’s gross
return, with \( r(0) = R \) and \( r(1) = R + X \). Thus \( R \) represents the firm’s baseline return and
\( X \) the extra return from the crime. Let \( h(a) \) be the external harm from the firm’s operations,
with \( h(0) = 0 \) and \( h(1) = H \). Thus \( H \) represents the external social harm generated by the
crime. Assume \( H > X - C \), implying that the first-best policy is to deter crime. Assume
\( C, R, X, H > 0 \).

The employee’s wage \( w(r) \) can be conditioned on the firm’s gross return. Since the firm’s
gross return \( r(a) \) is a deterministic function of the employee’s action, the wage can effectively
be conditioned on the criminal act. We will abuse notation slightly and write \( w(a) = w(r(a)) \).

The government makes type-I and type-II errors in enforcing corporate-crime laws, modeled
as follows. Let \( g(a) \) be the probability the government obtains a conviction. Thus \( g(0) \) is the
probability the government makes a type-I error, mistakenly convicting an innocent firm, and
\( 1 - g(1) \) is the probability the government makes a type-II error, failing to convict a criminal
firm. Assume the probability of conviction is higher if a crime is committed: \( g(1) > g(0) \).
Conviction rates are exogenously given.\(^3\) For conciseness, let \( g_0 = g(0) \) and \( g_1 = g(1) \).

Conditional on conviction, the government levies sanction \( s_f \geq 0 \) against the firm and \( s_e \geq 0 \)
against the employee. Let \( s = s_f + s_e \) be the total sanction. Sanctions are an endogenous choice
for the government. The employment contract may specify that the firm indemnifies the employee
for losses due to the sanction. Let \( s_i \in [0, s_e] \) be this indemnification payment, i.e., a payment

\(^3\)The structure of conviction probabilities and errors implicitly rules out the government’s using the wage scheme
to infer whether a crime was committed in equilibrium. This may be because the wage scheme is part of an
implicit contract unobservable to the government, or else because the contract and contracting environment are too
complicated for the government to make such inferences.
from the firm to the employee conditional on conviction.

The firm is risk neutral. The employee is risk averse. Let $u : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be the employee’s utility over wealth, with $u(0) = 0$, $u' > 0$, and $u'' < 0$. The cost of crime $c$ is additively separable from $u$ in the employee’s overall utility function.

To abstract away from firm judgment proofness, we assume the firm has an unlimited supply of liquifiable assets to pay its obligations. On the other hand, employee limited liability plays an integral role in one of our later results. To characterize employee limited liability, we assume the employee has a supply of liquifiable assets $\ell_e$ which in addition to his wage $w$ and indemnification payment $s_i$ can be used to pay the sanction $s_e$. The employee’s best option outside the firm provides no opportunity for crime, carries no risk of mistaken conviction, and pays a wage normalized to zero. If he takes this outside option, he consumes his liquifiable assets $\ell_e$, implying that his reservation utility is $u(\ell_e)$.

The timing is as follows. First the government sets the sanctions $s_f$ and $s_e$. These are observed by the firm. The firm then sets the employment contract $(w(0), w(1), s_i)$. The employee decides to accept the contract or pursue his outside option. Conditional on signing the contract, the employee then chooses whether or not to commit the crime. The state of the world determining whether the government convicts is realized, returns are realized, sanctions assessed, and wage and indemnification payments made.

3. Equilibrium Employment Contracts

We solve for the subgame-perfect equilibrium by backward induction. In the present section, we will take the government’s sanction scheme $(s_f, s_e)$ as given, we will solve for the employment contract $(w(0), w(1), s_i)$ maximizing the firm’s profit. Following Grossman and Hart (1983), we will separate the firm’s optimal-contracting problem into two steps. The first step is to solve for the optimal incentive-compatible and individually-rational contract implementing arbitrary employee action $a$. The second step is to compare the profits from the contract implementing no crime ($a = 0$) to that implementing crime ($a = 1$) and select the one yielding higher profit for
Consider the design of the optimal contract implementing action $a$, specifying three terms: an equilibrium wage $w(a)$, a wage $w(a')$ if the other action $a'$ is taken, and an indemnification payment $s_i$. The firm’s objective function is

$$r(a) - w(a) - g(a)(s_f + s_i),$$

(1)
equal to the firm’s gross return, less the wage payment, less the sanction and indemnification payments weighted by the probability the government convicts the firm. The employee’s expected surplus is

$$g(a) u(\max\{0, \ell_e + w(a) + s_i - s_e\}) + [1 - g(a)] u(\ell_e + w(a)) - c(a).$$

(2)
The employee adds the wage $w(a)$ to his existing wealth $\ell_e$ unless the government convicts. If the government convicts, the employee receives $w(a) + s_i$ from the firm. He puts this payment together with its other liquifiable assets $\ell_e$ and pays $s_e$ if it has sufficient funds. Otherwise it pays as much as it can and ends up with no wealth.

For the contract to be individually rational, the employee’s surplus in equation (2) must exceed his reservation utility $u(\ell_e)$. For the contract to be incentive compatible, (2) must exceed his surplus from choosing the “wrong” action $a'$. It turns out we can ignore the incentive-compatibility constraint. By reducing the wage for the “wrong” action to the lowest possible level, $w(a') = 0$, the firm can ensure the employee’s surplus from choosing $a'$ is no greater than $u(\ell_e) - c(a')$, which in turn is no greater than the employee’s reservation utility $u(\ell_e)$. Hence incentive compatibility is implied by individual rationality.

Since the firm is risk neutral and the employee risk averse, the optimal contract in many cases calls for the firm to insure the agent fully by indemnifying the full amount of the employee sanction: $s_i = s_e$. The exception arises when $s_e$ becomes large. Then rather than indemnifying this large loss, the firm finds it cheaper not to indemnify the employee at all, accepting the fact
that the sanction will bankrupt him, but taking advantage of his limited liability to cap his loss from the sanction, paying a higher wage to compensate for this loss. The next proposition fully characterizes the equilibrium contract.

**Proposition 1.** The profit-maximizing contract for the firm depends on the level of $s_e$. If $s_e$ is below a threshold, the firm fully indemnifies the employee ($s_i = s_e$). If $s_e$ is above this threshold, the firm offers no indemnification ($s_i = 0$), and $s_e$ bankrupts the employee. The firm optimally implements the action $a$ maximizing its objective function (1) given equilibrium wage and indemnification payments.

The proof of Proposition 1 in the appendix provides closed-form solutions for the threshold on $s_e$ determining whether the firm indemnifies the employee and for equilibrium wages.

Proposition 1 shows that the basic insights from the insurance literature apply to indemnification. The literature has shown (see for example Proposition 2 of Shavell (1986)) that a risk-averse agent with limited assets will purchase full insurance at actuarially fair rates if the insured loss is below a threshold and no insurance if the loss is above this threshold. The same principle applies to the self-insurance within the firm represented by indemnification.

### 4. Optimal Sanctions

Throughout the section we will distinguish between the first and second best. The first best is the outcome the government could achieve if it could directly set the employee contract and criminal action. The second best is the social-welfare-maximizing outcome subject to the constraints on the government assumed in the model: the government can set sanctions, not the employee contract or criminal action directly, and its enforcement ability is imperfect, with type-I and II errors. The second best will be the government sanction scheme observed in equilibrium.

#### 4.1. Alleged Benefit of Employee Sanctions

The literature suggests the following intuition for a possible benefit to the government of sanctioning the employee and banning his indemnification. Banning indemnification increases the friction
in the contracting process between the firm and employee, and this friction presumably harms the
criminal firm more because the probability of conviction—and the probability an indemnification
payment would have been made if it were allowed—is higher for the criminal firm. Intuition
along these lines was suggested by Stone (1980), Kraakman (1984), and Privileggi, Marchese,

Proposition 2 shows that this intuition is incorrect. While targeting the employee and banning
his indemnification increases the burden of a given sanction on a guilty firm, it increases the
burden on a law-abiding firm even more than would simply increasing sanctions against the
firm. As a result, targeting the employee and banning his indemnification is typically socially
inefficient.

**Proposition 2.** Consider any crime-deterring sanction scheme that (a) has a positive employee
sanction \((s_e > 0)\), (b) bans indemnification, \((s_i = 0)\) and (c) does not force the employee’s
limited-liability constraint to bind along the equilibrium path \((s_e \leq \ell_e + w^*(0))\). Social welfare
can be strictly increased by replacing this scheme with another that does not target the employee
with sanctions \((s_e = 0)\).

Proposition 2 follows from employee risk aversion. Banning indemnification prevents the
firm from insuring the employee against type-I enforcement error. Since the employee is risk
averse, such insurance would be socially valuable. Crime can be deterred more efficiently if
the firm alone were sanctioned because firm sanctions are just a transfer between the firm and
government involving no loss of social surplus.

Proposition 2 considers the case in which indemnification is banned. If indemnification is
allowed, employee sanctions are not necessarily inefficient. Given the firm has unlimited liability,
if the employee’s limited-liability constraint does not bind and indemnification were allowed,
employee sanctions would be equivalent to firm sanctions. Employee sanctions would be fully
passed through to the firm. Employee sanctions only become inefficient if indemnification is
banned as assumed in the conditions of Proposition 2.
4.2. True Benefit of Employee Sanctions

Proposition 2 leaves open a possible circumstance under which targeting the employee may be beneficial: if the employee sanction $s_e$ is so high that it forces his limited-liability constraint to bind.

An effective deterrence scheme should harm a guilty firm more than an innocent one. First, obviously, crime can only be deterred if the criminal firm’s surplus is reduced below that from innocent behavior. Second, conditional on deterring crime, the government prefers a scheme that harms law-abiding firms as little as possible. Unfortunately, harm to law-abiding firms cannot be avoided entirely because of type-I enforcement errors.

Proposition 3 shows that an employee sanction can be a useful deterrence tool if set so high that his limited-liability constraint binds. Such a high employee sanction will result in the seizure of all of the employee’s assets if there is a corporate-crime conviction. Since the employee of a criminal firm must be paid a higher wage to induce him to commit the crime, he has more assets to seize if a crime is committed than not, and so the employee sanction harms the firm relatively more if a crime is committed. Though the nominal employee sanction may be the same, the effective employee sanction is higher if a crime is committed than if not.

While the statement of Proposition 3 focuses on the possibility that employee sanctions are socially optimal, the proof is more comprehensive, fully characterizing the second-best sanction scheme for all parameters. A by-product of this full characterization are necessary and sufficient conditions for the case of interest in the statement of the proposition to arise. The interested reader is referred to the appendix for the proof and this expanded set of results.

**Proposition 3.** There exists a non-empty set of parameters for which the second-best sanction scheme requires a positive employee sanction, $s_e > 0$. For all these parameters, $s_e$ bankrupts the employee, and thus must be set sufficiently high in the second best; $s_e = \infty$ suffices.

The high employee sanction in Proposition 3 is beneficial because it extracts more from the employee of a criminal than an innocent firm. A similar benefit can be obtained by conditioning the nominal fine on the employee’s income. Conard (1972) advocates such a scheme, in particular
advocating a cap on an employee’s liability equal to his after-tax net income from the firm in
the year of violation. In our model, if the fine were set equal to this cap, the fine would also
increase with the commission of a crime.

Employee sanctions generate a benefit in Proposition 3 whether or not indemnification is
banned by law. The employee sanction works by forcing the employee’s limited-liability constraint
to bind. If the employee’s limited-liability constraint binds anyway, the optimal scheme may as
well specify an unboundedly large employee sanction, in which case the firm would prefer not
to indemnify the employee whether or not indemnification is banned. As yet, there is no public-
policy rationale for banning indemnification. Such a rationale will be provided in Section 5.

5. Banning Indemnification as a Prosecutorial Tool

We extend the model to allow prosecutors to seek the cooperation of the employee in convicting
the firm. We show that for some parameters, the optimal scheme bans indemnification in order
to secure the employee’s cooperation with prosecutors, increasing the probability the firm is
convicted, and reducing the attractiveness of crime. This allows the government to deter crime
with lower fines. Lowering fines increases social welfare to the extent the fines can be reduced
below the bankruptcy threshold, i.e., the threshold above which a law-abiding firm is bankrupted
when the government commits a type-I enforcement error. Avoiding bankruptcy results in a
savings of social welfare amounting to the net value of the firm’s production.

5.1. Model Extension

Consider an extension of the corporate-crime model in which the government also has a pros-
cutorial function. The prosecutors can use the cooperation of the employee to increase the
probability the firm is convicted. We maintain the probabilities $g_0$ and $g_1$ but reinterpret them
as probabilities the government initiates an investigation of the crime rather than the probability
of conviction. Conditional on an investigation being initiated, the probability of conviction is
α ∈ (0, 1) if the employee does not cooperate with the prosecutors and unity if he does.⁴ One can interpret cooperation by the employee as revealing a piece of hard information proving the crime, a “smoking gun.” Consistent with this interpretation, the employee can only cooperate if a crime has actually been committed; if the government has committed a type-I enforcement error by investigating an innocent firm, it is impossible for the employee to increase the probability of conviction by cooperating since there is no “smoking gun” to reveal. Combining the probability of investigation with the probability of conviction conditional on investigation, the unconditional probability of conviction equals (in increasing order) αg₀ if no crime was committed, αg₁ if a crime was committed and the employee does not cooperate with prosecutors, and g₁ if a crime was committed and the employee cooperates with prosecutors.

5.2. Optimal Sanctions

Prosecutors induce the employee to cooperate by promising to forgive a portion of the sanction in return for cooperation. Let sᵅ ∈ [0, sₑ] be the amount of the sanction forgiven. If the firm fully indemnifies the employee, by setting sᵅ = sₑ − sᵅ, the prosecutors’ strategy will not work since the employee will not care about reducing the sanction.⁵ The government thus needs to ban full indemnification to induce the employee to cooperate. There are two ways for the government to do this. One is simply to set the employee sanction so high that the firm chooses not to indemnify the agent even if it were allowed to. Setting a high employee sanction may be inefficient if this increases the wages a law-abiding firm needs to pay so much that it shuts down in equilibrium. If the shutdown of law-abiding firms is a concern, it can be efficient for the government to prohibit indemnification directly.

The next proposition identifies different cases in which the second-best sanction scheme secures the employee’s cooperation with prosecutors by offering to forgive some of his sanction.

---

⁴It is sufficient to assume only that cooperation increases the probability of conviction; assuming it increases the probability from α < 1 to unity is a pedagogical simplification.

⁵In the basic model, we took sᵅ to be a constant without loss of generality. In the extended model of the present section, we will take sᵅ to be proportional to the employee’s realized liability, sₑ − sₑ. Full indemnification is equivalent to sᵅ = sₑ − sₑ.
The proof in the appendix provides closed-form solutions for second-best sanctions for all parameters. As part of the specification of second-best sanctions, the proof provides necessary and sufficient conditions for cases (a) and (b) in Proposition 4 to arise.

**Proposition 4.** Consider the extended model in which the employee can cooperate with prosecutors. There exist two different cases, each involving a non-empty set of parameters, in which the second-best sanction scheme requires a positive employee sanction.

(a) For the first set of parameters, the first best can be approached in the limit as $\epsilon \to 0$ with a sanction scheme involving a small employee sanction $s_e = \epsilon$ that is completely forgiven in exchange for cooperation ($s_c = \epsilon$).

(b) For a second set of parameters, social welfare in the second best is bounded away from the first best. The second best is obtained by sanctioning the employee with $s_e$ high enough to bankrupt him. If he cooperates, some of $s_e$ is forgiven, but the residual $s_e - s_c$ is bounded above 0.

In both case (a) and (b), second-best sanctions deter crime without shutting down the firm and induce the employee to cooperate with prosecution by offering to forgive some of the sanction.

In case (a), the government can increase the probability of conviction with virtually no dead-weight loss by levying a vanishingly small employee sanction which it forgives if the employee cooperates. Indemnification must be banned for the forgiveness strategy to work. Otherwise, because the employee sanction is so small, it would be virtually costless for the firm to indemnify the employee. If the employee is indemnified, he would not gain from cooperation.

The second best approaches but does not reach the first best in case (a). By assumption the employee cannot cooperate if a crime was not committed, since there is no “smoking gun” to offer. Thus, the employee of an innocent firm would face the full sanction. But the only gap in social welfare between first and second best is the risk borne by the employee because of the unindemnified sanction. The gap disappears as the sanction becomes vanishingly small.

The second-best sanction scheme in (b) bankrupts the employee with a large sanction. This large, unindemnified risk leads to a loss in social welfare that is bounded above 0. Out of equilibrium, if a crime is committed, just enough of the sanction is forgiven to induce the employee to cooperate, but a finite sanction is left unforgiven to enhance deterrence.
It is not necessary for the government to ban indemnification for the sanction scheme in case (b) to work. The employee sanction is sufficiently high that the firm would not choose to indemnify the employee fully even if it were allowed by law. Indemnification must be banned for the sanction scheme in (a) to work. Indeed, case (a) is the only case identified anywhere in the paper in which banning indemnification can be socially beneficial.

6. Nuisance Suits

We have found only limited circumstances under which the government strictly prefers to sanction the employee for corporate crime and even more limited circumstances under which prefers to ban indemnification. Viewed in this way, the results suggest that the government should adopt a lenient policy toward indemnification. However, one could view the results from a completely different perspective. While banning indemnification provides a strict benefit in only a few cases, in no cases does the ban result in a strict loss. Viewed in this way, the government may as well ban indemnification.

A review of the results so far confirms that indemnification can be banned in all cases in the second-best. In most cases, the employee is not sanctioned in the second best, and indemnification is irrelevant. In another set of cases, the employee is sanctioned so severely that the firm chooses not to indemnify him even if allowed. In these cases, there would be no loss in banning indemnification. In the remaining cases we found a strict benefit from banning indemnification, to secure the employee’s cooperation with prosecutors.

The theoretical result that the second best may as well ban indemnification presents us with something of a puzzle when compared against practical observation. Most corporations have policies of indemnifying their directors and officers and that, rather than banning indemnification, the government seems to permit or indeed encourage it. In this section we resolve the puzzle with an extension of the model in which the firm faces nuisance suits from private or government parties. Absent indemnification, the director or officer would have an excessive incentive to settle the case to avoid personal liability.
6.1. Model

Consider an extension of the basic model in which a plaintiff brings suit against the firm and its employee. (The plaintiff could be a rent seeking private party, or a non-benevolent government prosecutor.) Rather than government sanctions, reinterpret $s_f$ and $s_e$ as judgments awarded to the plaintiff from the firm and employee if the plaintiff wins the trial and $s = s_f + s_e$ as the total judgment. Let $t_f$ and $t_e$ be the trial costs borne by the firm and employee in addition to the judgment awards, assumed to be independent of the trial outcome, including costs from legal defense and reputation loss. Let $t = t_f + t_e$ be total such additional trial costs. The difference between $s$ and $t$ is that $s$ is a transfer between parties while $t$ is a deadweight loss to society. Let $m \in [0, 1]$ be the plaintiff’s private information about the merit of the case which is here indexed by the probability of prevailing at trial. Other parties (the firm and employee) only know the density $f(m)$ and distribution $F(m)$ of $m$. For simplicity assume $f(m) > 0$ for all $m \in [0, 1]$, i.e., all of the support of $m$ is all of $[0, 1]$. As before, the employee has concave utility over wealth $u$ and liquifiable assets $\ell_e$ in addition to any income from employment.\footnote{All the results in this section would go through if the employee is risk neutral, but we maintain the same utility specification from the basic model to maintain consistent notation.} To consolidate cases, assume $\ell_e$ is high enough that employee limited liability does not bind; it is sufficient to assume $s_e < \ell_e$ for him to be able to pay his sanction regardless of his employment earnings. The firm is risk neutral and has unlimited liability. The plaintiff is risk neutral.

We will take the employee to be a decision maker within the firm (possibly a director or officer). The employee has the authority to make a settlement offer to the plaintiff. We will assume that this decision has to be delegated to the employee and cannot be specified in a contract due to contractual incompleteness. The settlement offer involves a total payment $p$, the sum of a payment from the firm $p_f$ and a payment from the employee $p_e$. These payments are offered to the plaintiff in return for dropping the case.

The timing of the game is as follows. First, that the firm offers a contract to the employee. Second, the plaintiff chooses whether to file suit. Next, the employee makes a settlement offer
on his and the firm’s behalf. Finally, if the settlement offer is not accepted, the case goes to trial and the plaintiff wins with probability $m$. Since the extension is more complicated than the basic model, involving the additional settlement stage, we will simplify the model in other dimensions. In particular, we will abstract from actions, criminal or otherwise, that the employee may have undertaken beforehand leading to the suit. Hence we will dispense with the notation for the social cost of crime $H$, firm benefit from crime $X$, and employee non-pecuniary cost of crime $C$. We will simply specify $R$ as the firm’s gross return outside of any payments to the plaintiff or employee.

In the basic model the employee had only two discrete options under his control, whether to engage in crime or not, and so fully general employment contracts could be captured with just two wage levels $w(0)$ and $w(1)$. Now, the employee has several continuous choices (settlement payments $p_f$ and $p_e$). To capture a rich set of contracts in a simple way, we will restrict attention to profit-sharing contracts, supposing that the employee receives a share $\beta \in [0, 1]$ of the firm’s profit. We will compare the equilibrium in which the firm indemnifies the employee’s total losses from trial, $s_e + t_e$, against that in which the firm does not indemnify any of these costs for the employee.

6.2. Settlement Offer

We solve for the subgame-perfect equilibrium using backward induction starting from the settlement offer. Clearly, given any total settlement payment $p$, the employee prefers the firm to pay it, so $p_f = p$ and $p_e = 0$. The employee will choose $p^*$ to maximize his expected utility from the settlement offer, trading off increases in the probability the offer is accepted against the increased payment conditional on acceptance. The employee’s evaluation of this tradeoff depends on whether he is indemnified. If so, he bears less direct liability if the offer is rejected. Indemnification thus leads to a reduction in $p^*$.

Proposition 5 uses monotone comparative statics techniques to show this result is general, holding for any density $f(m)$. The proof shows that indemnification strictly reduces $p^*$ unless $t$
is so high that the employee settles at all costs—i.e., chooses the corner solution $p^* = s$, which is sure to be accepted even by a plaintiff with the most meritorious case ($m = 1$).

We can also compare $p^*$ to the offer the firm would prefer if it did not have to delegate the decision to the employee. Since the firm is risk neutral, it values the insurance provided by settling less than the agent and thus would settle less often (i.e., offer a lower $p^*$). This result depends on settlement reducing risk compared to trial, which will be the case in equilibrium as long as the tails of $f(m)$ are not too fat, meaning that the density on cases in which the plaintiff is either very unlikely to win at trial ($m \approx 0$) or very likely to do so ($m \approx 1$) is low. A sufficient condition for the result is quasiconcavity of the objective function determining the firm’s settlement offer.

**Proposition 5.** Indemnification weakly reduces the employee’s equilibrium settlement offer $p^*$—strictly so if $t$ is not too high (sufficient condition in equation (A44)). If the firm were to make the settlement offer instead of delegating it to the employee, assuming the firm’s objective function is quasiconcave, it would offer a weakly lower $p^*$ than would even an indemnified employee—strictly so if $t$ is not too high (sufficient condition in equation (A47)).

Indemnification thus can make the employee a better bargainer in the settlement period for the firm and can increase firm profits. Banning indemnification eliminates socially valuable insurance. In addition, by lowering firm profit at the settlement stage, it could lead to the exit of the firm when $R$ is moderate and the firm is on the margin between operating or exiting. An offsetting social cost of indemnification is that there is less settlement, and therefore greater expenditure of trial costs. The employee and firm internalize their costs, but any trial costs for the plaintiff would be an externality that would lead to a distortion in the settlement decision. The next subsection shows this social loss is offset by the reduction in suits overall.

### 6.3. Endogenizing Plaintiff Behavior

We have so far taken the plaintiff’s decision to sue as exogenous. Suppose that the plaintiff has a fixed cost $k$ of bringing suit, jointly distributed with the merit of the suit $m$ according to some density function $f(m, k)$. Then making the employee a tougher bargainer by indemnifying him
could reduce the entry of suits, starting with those with low $m$ and high $k$. This would raise social welfare through fewer trials, lower expected trial expenditures, and higher firm profits. Higher profit for the firm can add to social welfare by increasing the chance that it undertakes socially valuable production.

Solving for equilibrium with endogenous plaintiff entry is quite complicated in general, since it leads to a game in which the plaintiff’s filing decision signals something about the merits of his case to the employee, who responds with a settlement offer. We can gain some intuition for the results and a glimpse of the complications involved by studying a simple numerical example below.

Another aspect of plaintiff behavior is the division of the total trial award $s$ into a component for the firm $s_f$ and for the employee $s_e$. While in practical settings, the division is not completely within the control of the plaintiff, he can affect the division through the damages that are sought when the suit is filed. Shifting as much of a fixed amount $s$ toward the employee has the benefit of making him a softer bargainer and increasing the generosity of his settlement offer. Employee limited liability will constrain how much of $s$ the plaintiff shifts to the employee and how much is directed at the firm, typically the party with “deeper pockets”. In the numerical example, we consider the extreme case in which the employee is not constrained by limited liability at all, so the plaintiff wants the employee to face the whole of $s$.

6.4. Numerical Example

Let $u(x) = \sqrt{x}$ be the employee’s utility function, $\ell_e = 10$ be his liquifiable assets, and $R = 100$ be the firm’s return. Suppose the plaintiff is successful in having the whole damage award $s = 10$ shifted against the employee and none against the firm, so $s_f = 0$ and $s_e = 10$. We assume the employee bears all the trial costs: $t_f = 0$ and $t_e = t = 5$. Assume the employee’s profit share, to take a round number, is $\beta = 1/2$.\footnote{We take $\beta$ as exogenous, set to address moral hazard considerations from outside the example. Within the model, the firm would optimally choose $\beta$ as close as possible to zero. Besides limiting the rent leaking out to the employee, reducing $\beta$ reduces the stakes involved in the trial “gamble”, and in the limit makes the indemnified employee bear all the costs.}
For the joint distribution of $m$ and $k$ for the plaintiff, there are three discrete realizations: with probability $1/2$, $m = 1$ and $k = 0$; with probability $1/4$, $m = 0$ and $k = 0$; and with probability $1/4$, $m = 0$ and $k = 4$. That is, there are two levels of merit for suits: certain losers and certain winners. All certain winners are costless to file. Half of the certain losers are costless to file and half involve a positive cost.

We solve for the perfect Bayesian equilibrium of this signaling game working backward from the settlement stage. With two types for plaintiff merit ($m = 0$ and $m = 1$), the optimal settlement offer will be one of two possibilities. The employee can offer $p = 0$ (here think of a very small amount, such as a penny), and have the offer accepted only by plaintiffs with $m = 0$, or $p = 10$, and have the offer accepted by all plaintiffs. Which strategy provides higher expected utility depends on the plaintiff’s filing behavior.

In equilibrium the plaintiff will file if it is costless to do so, regardless of merit. The only question is whether the plaintiff files if it is costly to do so. Suppose first that all plaintiffs file, even those with costly suits. It can be shown that the indemnified employee’s expected utility at the settlement stage is 7.42 if he offers $p = 10$ and 7.49 if he offers $p = 0$. Therefore, he prefers $p = 0$. The unindemnified employee’s analogous expected utilities are 7.42 and 7.40 respectively, so he prefers the opposite choice: $p = 10$. This calculation verifies the claim in Proposition 5 that indemnification makes the employee a tougher bargainer at settlement.

The calculations indicate a unique perfect Bayesian equilibrium when the employee is unindemnified: all plaintiffs file and the employee always settles with an offer of $p^* = 10$. The calculations also indicate that it cannot be an equilibrium for all plaintiffs to file when the employee is indemnified. Filing is dominated for the plaintiff with a meritless but costly-to-file case since the employee’s offer leaves him with no gross surplus to cover the cost of filing. The equilibrium must involve mixed strategies, with the employee randomizing between offering $p = 10$ and $p = 0$ at the settlement stage with probabilities $\gamma$ and $1 - \gamma$, and with the plaintiff randomizing between filing the costly suit and not with probabilities $\delta$ and $1 - \delta$. The equilib-
rium probabilities are calculated to make the opposing party indifferent between the two specified actions, leading to $\gamma^* = 0.400$ and $\delta^* = 0.034$.

Assessing how indemnification affects the equilibrium, it reduces the chance a suit is brought, from 100% to 76%. Indemnification also makes the employee a tougher bargainer, reducing the chance the employee offers $p = 10$ from 100% to 40%.

7. Literature Review

To our knowledge, ours is the first formal analysis of the question of whether indemnification should be banned, contributing to the literature studying the optimal division of corporate-crime sanctions between the principal and agent. See Mullin and Snyder (2008) for a review. Much of this literature (e.g., Newman and Wright 1990, Macey 1991, Arlen 1994, Chu and Qian 1995, Davis 1996, Arlen and Kraakman 1997, Shavell 1997, Arlen 1998, Garoupa 2000) analyzes the case in which the firm’s agent commits a corporate crime in his own, and against the firm’s, interest. In this setting, it is natural that the agent should be sanctioned in the socially-optimal legal regime; the interesting question is whether the principal should be as well. Drawing on the broader literature on vicarious liability (e.g., Sykes 1984, Shavell 1987), the articles show that sanctioning the firm increases deterrence if limits to the agent’s wealth prevent his paying sanctions sufficient to deter the crime; targeting the firm is particularly effective if it can monitor the agent’s actions better than can government authorities.

In the framework of our corporate-crime model in Section 2, the agent’s alleged conduct benefits the firm (at least in the absence of sanctions). This is the natural framework for studying our central issue—indemnification—because a firm would presumably not choose to indemnify its agent for crimes against itself. Our framework complements the existing corporate-crime literature because few other papers in the literature assume the crime benefits the firm. One exception is Privileggi, Marchese, and Cassone (2001). Our paper differs from theirs in many respects including that the level of the fine is exogenous in their model and can only be levied on one party or the other, so joint firm-employee liability is not allowed. Indemnification is also
exogenously ruled out in their model, whereas it is the focus of our paper.

Kornhauser (1982), Segerson and Tietenberg (1992), and Polinsky and Shavell (1993) consider the case of a corporate tort. The employee and perhaps the firm invest in care to prevent an accident. The authors find that the government authority should target the employee for sanctions when the government is better at monitoring care and/or when the government is better at levying sanctions because of its ultimate threat of imprisonment. Indemnification plays no role in these papers because the firm has no incentive to indemnify the agent in equilibrium. Indeed, Polinsky and Shavell (1993) demonstrate cases in which the firm prefers higher employee sanctions than the government. Our model of willful corporate crimes is quite different: indemnification reduces the cost of inducing the employee to commit the crime and hampers the prosecutor’s ability to reduce the employee’s sanction in return for his cooperation against the firm.

A number of the ideas formally developed here were first noted in law review articles by Stone (1980) and Kraakman (1984), including that type-I enforcement errors may provide a rationale for allowing employee indemnification and that forbidding indemnification can help secure the cooperation of employees in prosecuting the firm. These papers do not have models, however; our contribution is to provide a formal economic model and analysis. The formal analysis allows us to identify new reasons for targeting the employee. For example, we show it can be efficient to bankrupt the agent with a large sanction since the burden of this sanction falls more heavily on criminal than law-abiding firms. Some of the ideas in the law review articles do not withstand formal scrutiny. For example, we show that enhanced deterrence is not a reason to advocate a ban on indemnification because higher firm sanctions are a more efficient alternative.

Our result that forbidding indemnification helps secure the cooperation of the employee to increase the chances of successful prosecution of the firm is reminiscent of the work of Arlen (1994), Chu and Qian (1995), and Arlen and Kraakman (1997). They show that partially forgiving firm sanctions can increase the firm’s incentive to monitor the employee when such monitoring can increase the likelihood of uncovering criminal acts by employees. In both our work and theirs, the analysis is somewhat delicate because it is not obvious the “cooperating” party would
want to trade off a lower sanction for an increased chance of prosecution. In our work, the identity of the “cooperating” party is the opposite of theirs, the employee rather than the firm. The costs and benefits of securing cooperation are different in our model. Most importantly, our insight that forbidding indemnification plays a key role in allowing the government to trade reduced sanctions for cooperation did not appear in these previous papers.

The model of Section 6 contributes to the literature on the credibility of nuisance suits, including Bebchuk (1988), Katz (1990), and Schwartz and Wickelgren (2007). We study how delegating negotiations affects the settlement process and how the equilibrium is affected by changes in the agent’s compensation contract, including indemnification. Our paper is part of a broader literature on delegated bargaining including Jones (1989) and Fershtman, Judd, and Kalai (1991). Their applications are quite different from ours, and since their agents are not exposed to sanctions, the issue of indemnification does not arise.

We assume the firm has unlimited liability, thus abstracting from what Shavell (1986) and later authors term the problem of a “judgment proof” firm. We do this for two reasons. First, it is already well-understood from the literature on vicarious liability that if the liability of one party in a principal-agent setting is limited, it may only be possible to deter crime by also sanctioning the other party. Second, a model in which firm judgment-proofness provides the rationale for sanctioning the agent is not a fertile one for studying indemnification. Since sanctions bankrupt a judgment-proof firm, it would not have the funds to indemnify the agent whether or not allowed by law. So the government’s policy toward indemnification would be irrelevant.

8. Conclusion

This paper has studied the private and social returns to indemnification with regard to two significant business risks facing even law-abiding firms: the risk of mistaken prosecution (studied in the corporate-crime model of Sections 2–5) and the risk of nuisance suits (studied in nuisance-suit

8Related issues arise in the literature on plea bargaining (e.g., Reinganum 1988 and Kobayashi 1992) and on the use of leniency programs in the prosecution of cartels (e.g., Spagnolo 2000, Motta and Polo 2003, Aubert, Rey, and Kovacic 2006).
The broad lesson to draw from both models is that authorities should be wary of sanctioning employees let alone banning their indemnification. In the corporate-crime model, crime is typically deterred more efficiently with firm sanctions than with unindemnifiable agent sanctions. In the nuisance-suit model, indemnification provides a strategic benefit, “stiffening” the agent’s “spine” to fight against nuisance suits.

Delving more carefully into the results, we find benefits and costs of targeting the agent and banning his indemnification, depending on specific cases. The benefits of banning indemnification show up in the corporate-crime model. In that model, employee sanctions can be set in just the right way to remove any social loss—and in some cases provide a strict social benefit—from targeting the agent and banning his indemnification. To see this point, note that there are cases in which the efficient deterrence scheme involves such high sanctions that the employee is bankrupted. If the government is able to restrict employee sanctions just to these cases, then there is no loss in banning indemnification since the employee sanction is so high the firm chooses not to indemnify the employee even if the law allows. When the corporate-crime model is extended to allow prosecutors to bargain with the employee to secure his cooperation against the firm as in Section 5, in some cases there is a strict social benefit from banning indemnification. Therefore, the corporate-crime model can be used to justify a ban on indemnification. (It should be emphasized that the corporate-crime model justifies a ban on indemnification only if the government has the information needed to identify exactly which cases the agent should be targeted with large sanctions; in the absence of this information it may be more efficient to allow indemnification wholesale.)

The costs of banning indemnification show up in the nuisance-suit model since indemnification provides a strategic benefit to the firm, and this benefit is possibly of social value. Therefore, determining whether indemnification should be banned as a matter of policy requires weighing benefits versus costs, i.e., requires considering which is the larger social goal: securing cooperation with prosecutors, which is aided by banning indemnification, or limiting nuisance suits, which is aided by allowing indemnification. Under some conditions, a nuanced policy toward
indemnification could accomplish both goals: the government could allow indemnification for private suits but ban it for government actions against corporate crime. In particular, this nuanced policy would work if prosecutors are benevolent, prosecuting just the right cases for efficiency.

However, if prosecutors have objectives other than social-welfare maximization, then the analysis for nuisance suits could apply to government prosecution, and the nuanced policy which bans indemnification for government actions may no longer be optimal. A recent, controversial case in which prosecutorial power and its potential abuse were at issue is the U.S. Department of Justice’s prosecution of the Big-Six accounting firm KPMG for constructing allegedly illegal tax shelters. Following strategies laid out in the 2003 Thompson Memorandum, prosecutors threatened to indict KPMG unless it refrained from indemnifying the legal expenses of indicted employees. In July 2007, Judge Lewis Kaplan dismissed the indictments for 13 of 16 KPMG employees. The court found that “The government threatened to indict, and thus to destroy, the giant accounting firm, KPMG. It coerced KPMG to limit and then cut off its payment of legal fees of KPMG employees. KPMG avoided indictment by yielding to government pressure. Many of its personnel did not.”

The prosecutorial strategies laid out in the 2003 Thompson Memorandum have since been softened in subsequent guidelines issued by the Department of Justice, the 2007 McNulty Memorandum.

In future work we will adapt the model to consider the distinction between direct corporate indemnification and (third-party) D&O insurance. The present paper has assumed direct indemnification for simplicity. As Holderness (1990) notes, D&O insurance has the benefit of having another party, the insurance company, as a monitor to ensure the payouts are for acts taken in good faith. Our further work will also consider the costs of D&O insurance relative to indemnification.

---

9United States vs. Stein, S1 05 Crim. 0888 (LAK), South. District of New York, decided 17 July 2007, page 1.
Appendix

Proof of Proposition 1: Some new notation will help streamline the proofs. Let $b(g, c)$, a mnemonic for “bankruptcy wage,” be the wage that allows the employee to achieve his reservation utility $u(\ell_e)$ given that he exerts on-the-job effort $c$ and given that a government sanction sufficient to bankrupt him is levied with probability $g$. Formally, $b(g, c)$ implicitly solves

$$(1 - g) u(\ell_e + b) - c = u(\ell_e),$$

or, rearranging,

$$b(g, c) = u^{-1}\left(\frac{u(\ell_e) + c}{1 - g}\right) - \ell_e.$$

To keep the notation concise, the dependence of $b(g, c)$ on $\ell_e$ is suppressed.

The profit-maximizing contract implementing action $a$ is the wage $w(a)$ and indemnification payment $s_i$ maximizing

$$r(a) - w(a) - g(a)(s_f + s_i)$$

subject to $w(a) \geq 0$, $s_i \geq 0$, and individual-rationality constraint

$$g(a) u(\max\{0, \ell_e + w(a) + s_i - s_e\}) + [1 - g(a)] u(\ell_e + w(a)) - c(a) \geq u(\ell_e).$$

It is obvious that (A2) binds. The non-differentiable max operator can be removed from (A2) by noting it is equivalent to the following set of constraints. Either both (A3) and (A4) hold:

$$s_e \geq \ell_e + w(a) + s_i$$

$$(1 - g(a)) u(\ell_e + w(a)) - c(a) = u(\ell_e);$$

or both (A5) and (A6) hold:

$$s_e \leq \ell_e + w(a) + s_i$$

$$g(a) u(\ell_e + w(a) + s_i - s_e) + [1 - g(a)] u(\ell_e + w(a)) - c(a) = u(\ell_e).$$

We will solve two separate constrained optimization problems for these two sets of constraints and compare the solutions.

To proceed, first consider the problem of maximizing (A1) subject to $w(a) \geq 0$, $s_i \geq 0$, (A3), and (A4). Put (A3) aside for now; we will return to this detail at the end of the proof. The solution obviously involves setting $s_i = 0$ since $s_i$ does not appear in (A4) and (A1) is decreasing in $s_i$. Solving (A4) yields equilibrium wage $w(a) = b(g(a), c(a))$.

Next, consider the problem of maximizing (A1) subject to $w(a) \geq 0$, $s_i \geq 0$, (A5), and (A6). Ignoring all constraints except (A6) yields the Lagrangian for equality-constrained optimization:

$$\mathcal{L} = r(a) - w(a) - g(a)(s_f + s_i) + \lambda\{g(a) u(\ell_e + w(a) + s_i - s_e)$$

$$[1 - g(a)] u(\ell_e + w(a)) - c(a) - u(\ell_e)\}$$

(A7)
The first-order conditions with respect to \( w(a) \) and \( s_i \) upon rearranging are

\[
\frac{1}{\lambda} = g(a) u'(\ell_e + w(a) + s_i - s_e) + [1 - g(a)] u'(\ell_e + w(a)) \tag{A8}
\]

\[
\frac{1}{\lambda} = u'(\ell_e + w(a) + s_i - s_e). \tag{A9}
\]

Setting the right-hand sides of (A8) and (A9) equal yields \( s_i = s_e \). Substituting \( s_i = s_e \) into (A6) yields the equilibrium wage \( w(a) = b(0, c(a)) \). It is easy to check that this solution satisfies the ignored constraints.

Next, we need to compare the two solutions. The firm will select the solution generating the highest value of the objective (A1) or equivalently the solution minimizing the total expected payment to the employee \( w(a) + g(a)s_i \). The first solution yields expected payment \( b(g(a), c(a)) \). The second yields expected payment \( b(0, c(a)) + g(a)s_e \). The first expected payment is lower, and thus the firm prefers the solution, if

\[
s_e > \frac{b(g(a), c(a)) - b(0, c(a))}{g(a)}. \tag{A10}
\]

A minor technical point remaining to be addressed is to verify that we were safe in ignoring constraint (A3) in the first maximization problem. We will do so by showing that (A10) implies (A3). As a preliminary step, note

\[
\begin{align*}
    u^{-1}(u(\ell_e) + c(a)) &= u^{-1} \left( g(a)(0) + [1 - g(a)] \left[ \frac{u(\ell_e) + c(a)}{1 - g(a)} \right] \right) \tag{A11} \\
    &\leq g(a) u^{-1}(0) + [1 - g(a)] u^{-1} \left( \frac{u(\ell_e) + c(a)}{1 - g(a)} \right) \tag{A12} \\
    &= [1 - g(a)] u^{-1} \left( \frac{u(\ell_e) + c(a)}{1 - g(a)} \right). \tag{A13}
\end{align*}
\]

Inequality (A12) follows from the concavity of \( u \), which implies the convexity of \( u^{-1} \). Equation (A13) follows from the assumption \( u(0) = 0 \), which implies \( u^{-1}(0) = 0 \). The right-hand side of (A10) is, after substituting the definition of \( b \),

\[
\frac{1}{g(a)} \left[ u^{-1} \left( \frac{u(\ell_e) + c(a)}{1 - g(a)} \right) - u^{-1}(u(\ell_e) + c(a)) \right] \geq u^{-1} \left( \frac{u(\ell_e) + c(a)}{1 - g(a)} \right), \tag{A14}
\]

where the inequality in (A14) follows from (A11) through (A13). But the right-hand side of (A14) equals the right-hand side of (A3) after substituting the solution \( s_i = 0 \) and \( w(a) = b(g(a), c(a)) \), and then substituting the definition of \( b \). Q.E.D.

**Proof of Proposition 2:** Suppose the government bans indemnification. Suppose further that it imposes sanction scheme \((s_f, s_e)\) with \( s_e > 0 \) such that, in the continuation equilibrium, (a) crime is deterred and (b) the employee’s limited-liability constraint does not bind. We will show that social surplus can be increased by moving to a new sanction scheme with no employee sanction.
Let \( y(a, s_e) \) denote the wage that is optimal for the firm to pay given it wants to implement action \( a \) and given the employee’s sanction is \( s_e \), which the firm is banned from indemnifying. As indicated in the proof of Proposition 1, the employee’s individual-rationality constraint binds at an optimum. Substituting the indemnification-ban condition \( s_i = 0 \) into equation (A2) and imposing the assumption that employee limited liability does not bind in equilibrium, we have that \( y(a, s_e) \) is the implicit solution to

\[
g(a)u(\ell_e + y(a, s_e) - s_e) + \left[1 - g(a)\right]u(\ell_e + y(a, s_e)) - c(a) = u(\ell_e). \tag{A15}
\]

Since \( c(0) = 0 \), equation (A15) implies \( y(0, 0) = 0 \).

Expected social surplus under the original sanction scheme \((s_f, s_e)\) equals expected firm profit \( R - y(0, s_e) - g_0s_f \) plus employee surplus \( u(\ell_e) \) (this simple expression follows because employee individual rationality is binding) plus expected government sanction revenue \( g_0(s_f + s_e) \), or, rearranging,

\[
R - y(0, s_e) + u(\ell_e) + g_0s_e. \tag{A16}
\]

For the new sanction scheme \((s'_f, s'_e)\) to deter crime, the firm’s profit from no crime \( R - y(0, 0) - g_0s'_f \) must weakly exceed its profit from crime \( R + X - y(1, 0) - g_1s'_f \), or, rearranging,

\[
s'_f \geq \frac{X - y(1, 0)}{g_1 - g_0}. \tag{A17}
\]

Without loss of generality, we will take the value of \( s'_f \) such that (A17) holds with equality.

Social surplus under the new scheme equals the sum of firm profit \( R - g_0s'_f \) plus employee surplus \( u(\ell_e) \) (again, employee individual rationality binds) plus expected government sanction revenue \( g_0s'_f \), or, rearranging,

\[
R + u(\ell_e). \tag{A18}
\]

Expected social surplus is higher under the new sanction scheme if and only if (A18) exceeds (A16) or equivalently if and only if \( g_0s_e < y(0, s_e) \).

We will show this last inequality follows from the concavity of \( u \). By definition of concavity,

\[
tu(x') + (1 - t)u(x'') < u(tx' + (1 - t)x'') \tag{A19}
\]

for \( t \in [0, 1] \). Substituting \( t = g_0 \), \( x' = \ell_e - (1 - g_0)s_e \), and \( x'' = \ell_e + g_0s_e \) into (A19) yields

\[
g_0u(\ell_e + g_0s_e - s_e) + (1 - g_0)u(\ell_e + g_0s_e) < u(\ell_e). \tag{A20}
\]

Substituting \( a = 0 \) into (A15) implies

\[
g_0u(\ell_e + y(0, s_e) - s_e) + (1 - g_0)u(\ell_e + y(0, s_e)) - c(a) = u(\ell_e). \tag{A21}
\]

Since \( u' > 0 \), (A20) and (A21) together imply \( g_0s_e < y(0, s_e) \) and thus that expected social welfare increases by moving to the sanction scheme with no employee sanction. \textit{Q.E.D.}
Proof of Proposition 3: In this proof we will solve for the second-best sanction scheme for arbitrary parameters. At the end we will show how the statement of Proposition 3 can be gleaned from these general results. The following expressions will help partition the characterization of the optimal scheme into subcases:

\[ b(0, C) + \left[ \frac{g_1 - g_0}{g_0} \right] R - X \]  
(A22)

\[ H - R - X + b(0, C) \]  
(A23)

\[ X - \left[ \frac{g_1 - g_0}{g_0} \right] R + \left( \frac{g_1}{g_0} \right) b(g_0, 0) - b(g_1, C) \]  
(A24)

\[ X - H + b(g_0, 0) - b(0, C). \]  
(A25)

By Proposition 2, without loss of generality, the socially optimal employee sanction can be taken to be either zero or so high it forces the employee’s limited-liability constraint to bind in equilibrium. In the latter event, without loss of generality, the socially optimal employee sanction can be taken to be \( s_e = \infty \). The proof proceeds by analyzing the \( s_e = 0 \) and \( s_e = \infty \) cases separately in two steps and then combining and extending the results in a final step.

Step 1: Compute the socially optimal sanction scheme constraining \( s_e = 0 \). We first compute the firm’s maximum profit as a function of \( s_f \) for each action \( a \) the firm can induce. Suppose the firm decides to induce \( a = 0 \). Then applying Proposition 1 with \( a = 0 \) and \( s_e = 0 \), we have that the firm optimally pays the employee wage \( b(0, 0) = 0 \). Consequently, the firm’s maximum profit is

\[ R - g_0 s_f. \]  
(A26)

Suppose the firm decides to induce \( a = 1 \). Then applying Proposition 1 with \( a = 1 \) and \( s_e = 0 \), we have that the firm optimally pays the employee wage \( b(0, C) \). Consequently, the firm’s maximum profit is

\[ R + X - b(0, C) - g_1 s_f. \]  
(A27)

Because firm sanctions are frictionless transfers, social welfare is independent of \( s_f \) except to the extent that \( s_f \) affects action \( a \). Because (A26) and (A27) are linear in \( s_f \), the socially optimal firm sanction can be taken, without loss of generality, to be one corner, \( s_f = 0 \), the other corner, \( s_f = \infty \), or the value at which the law-abiding firm is just indifferent between shutting down and not, \( s_f = R/g_0 \). If (A22) is positive, substituting \( s_f = R/g_0 \) implies (A26) is both nonnegative and greater than (A27). Thus, the sanction scheme \((s_f, s_e) = (R/g_0, 0)\) deters crime with no deadweight loss.

On the other hand, if (A22) is negative, deterring crime involves a deadweight loss. Among schemes with \( s_e = 0 \), the two possibilities are that crime is not deterred or that crime is deterred by shutting down the firm. (Among schemes with \( s_e > 0 \), it is possible that crime is deterred without shutting down the firm, but there is still a deadweight loss because the employee will bear some risk. This last possibility is discussed in step 3.) If (A23) is positive, the social surplus from setting \( s_f = \infty \) and thereby shutting down the firm, \( u(l_e) \), exceeds that from setting \( s_f = 0 \)
and thereby allowing crime, \( R + X - H - b(0, C') + u(\ell_e) \).

**Step 2:** Compute the socially optimal sanction scheme constraining \( s_e = \infty \). Throughout this step, maintain the assumption that (A22) is negative. In step 1, we already found the first best scheme if (A22) is positive. As in step 1, we begin by computing the firm’s maximum profits as functions of \( s_f \) for each action \( a \) it can induce. Suppose the firm decides to induce \( a = 0 \). Then applying Proposition 1 with \( a = 0 \) and \( s_e = \infty \), we have that the firm optimally pays the employee wage \( b(g_0, 0) \). Consequently, maximum firm profit is

\[
R - b(g_0, 0) - g_0 s_f. \tag{A28}
\]

Suppose the firm decides to induce \( a = 1 \). Then applying Proposition 1 with \( a = 1 \) and \( s_e = \infty \), we have that the firm optimally pays the employee wage \( b(g_1, C) \). Consequently, maximum firm profit is

\[
R + X - b(g_1, C) - g_1 s_f. \tag{A29}
\]

By similar logic to that in step 1, the socially optimal firm sanction can be taken, without loss of generality, to be one corner, \( s_f = 0 \), the other corner, \( s_f = \infty \), or the value at which the law-abiding firm is just indifferent between shutting down and not:

\[
s_f = \frac{R - b(g_0, 0)}{g_0} \tag{A30}
\]

If (A24) is positive, even if \( s_f \) is set to the value in (A30), (A29) exceeds (A28), implying that the value of \( s_f \) in (A30) cannot deter crime; hence, the optimal firm sanction is either \( s_f = \infty \) (if, as shown in step 1, (A23) is positive) or \( s_f = 0 \) (if, as shown in step 1, (A23) is negative).

If (A24) is negative, setting \( s_f \) to the value in (A30) is sufficient to deter crime without shutting the firm down; hence, \( s_f = \infty \) is suboptimal.

**Step 3: Comparison and extension.** The preceding steps provide a complete characterization of the socially-optimal sanction scheme in all but one subcase. If (A22) and (A24) are negative, we showed \( s_f = s_e = \infty \) is suboptimal. Thus, the optimal scheme must either involve \( s_f = s_e = 0 \) and allow crime or involve \( s_f \) equal to the value in (A30) and \( s_e = \infty \) and deter crime as efficiently as possible. (It is impossible to deter crime with \( s_e = 0 \) because this would lead to the first best, contradicting the previous finding that the first best cannot be obtained if (A22) is negative.) If (A25) is positive, the scheme that allows crime generates higher social welfare and if (A25) is negative, the reverse is true.

Synthesizing the analysis, we have the following four cases. To make the statements of the conditions more elegant, ignore knife-edge cases in which conditions (A22) through (A25) exactly equal zero.

- If (A22) is positive, then the government can obtain the first best, deterring corporate crime with no deadweight loss, with a sanction scheme that does not target the employee. In particular, the scheme \( s_f = R/g_0 \) and \( s_e = 0 \) suffices.
- If (A22) is negative and (A23) and (A24) are positive, then the socially optimal sanction
scheme deters crime by shutting down the firm. In particular, the scheme $s_f = \infty$ and $s_e = 0$ suffices.

- If (A22) and (A23) are negative and (A25) is positive, then the socially optimal sanction scheme does not deter crime. In particular, the scheme $s_f = s_e = 0$ suffices.

- For the remaining cases in which expressions (A22) through (A25) are nonzero, the socially optimal sanction scheme, which deters crime and does not lead to the shutdown of the firm, must involve a positive employee sanction. In particular, the scheme in which $s_f$ equals (A30) and $s_e = \infty$ suffices.

This completes the full characterization of the optimal sanction scheme and continuation equilibrium for all parameters. The last bullet point provides necessary and sufficient conditions for case of interest in the statement of the proposition, viz., the case in which the optimal sanction scheme requires $s_e > 0$. Q.E.D.

**Proof of Proposition 4:** Consider the extended model in which the employee can increase the probability of conviction by cooperating with prosecutors. In addition to (A22) and (A23), the following expressions will help partition subcases in our characterization of the socially optimal sanction scheme:

$$b(0, C) + \left[ \frac{g_1 - \alpha g_0}{\alpha g_0} \right] R - X \quad \text{(A31)}$$

$$X - \left[ \frac{g_1 - \alpha g_0}{\alpha g_0} \right] R + \frac{g_1}{\alpha g_0} b(\alpha g_0, 0) - b(\alpha g_1, C) \quad \text{(A32)}$$

$$X - H + b(\alpha g_1, C) - b(0, C). \quad \text{(A33)}$$

Following the calculations in the proof of Proposition 3, it can be verified that the first best can be obtained if (A22) is positive, using the same scheme as in the proof of Proposition 3. If (A22) is negative and (A31) is positive, the first best can be approached arbitrarily closely using the scheme in case (b) of Proposition 4. Arguments analogous to those in the proof of Proposition 2 can be used to show that if the firm operates and crime is deterred, an interior value of $s_e$ is suboptimal. Thus, if the firm operates and crime is deterred, the sanction scheme must either set an arbitrarily small value of $s_e$ or value large enough to force the employee’s limited-liability constraint to bind.

Therefore, if (A22) and (A31) are negative, we are left with three strategies for the optimal sanction scheme. The scheme either shuts the firm down, allows crime, or deters crime without shutting the firm down by forcing the employee’s limited-liability constraint to bind. We computed the maximum social surplus from the first two strategies in the proof of Proposition 3. It remains to compute the social surplus from the optimal scheme using the third strategy.

We first compute the firm’s maximum profit from the third strategy if no crime is induced. In this situation, the probability of conviction is $\alpha g_0$. The firm optimally pays a wage forcing employee individual rationality to bind. Given $s_e$ bankrupts the employee by assumption, this
wage is $b(\alpha g_0, 0)$. Thus the firm’s profit is

$$R - b(\alpha g_0, 0) - \alpha g_0 s_f.$$  \hspace{1cm} (A34)

Next, we compute the firm’s maximum profit if a crime is induced. The firm’s profit depends on whether or not the employee cooperates with prosecutors. The socially optimal sanction scheme obviously induces cooperation. We need to see what this implies for the employee’s compensation and the structure of the optimal sanction scheme. If the employee cooperates, he earns

$$g_1 u(\ell_e - s_e + s_e + w^*) + (1 - g_1)u(\ell_e + w^*) - C,$$  \hspace{1cm} (A35)

where $w^*$ is the equilibrium wage for which we will shortly solve. If he does not cooperate, he earns

$$\alpha g_1 u(0) + (1 - \alpha g_1)u(\ell_e + w^*) - C$$  \hspace{1cm} (A36)

since $s_e$ forces the employee’s limited-liability constraint to bind, so that all the employee’s assets are seized if there is a conviction. The optimal value of $s_e$ forces (A35) to equal (A36), implying

$$s_e = s_e - w^* - \ell_e + u^{-1}((1 - \alpha)u(\ell_e + w^*)),$$

in turn implying the employee’s surplus from cooperating is

$$(1 - \alpha g_1)u(\ell_e + w^*) - C$$  \hspace{1cm} (A37)

since $u(0) = 0$. The employee cooperates in equilibrium, so (A37) characterizes the employee’s equilibrium surplus. The firm optimally pays $w^*$ forcing the employee individual rationality to bind or, equivalently, forcing (A37) to equal $u(\ell_e)$, implying $w^* = b(\alpha g_1, C)$. Firm profit if it induces crime is thus

$$R + X - b(\alpha g_1, C) - g_1 s_f.$$  \hspace{1cm} (A38)

It is feasible for the sanction scheme considered above to deter crime if (A38) is negative when evaluated at the highest value of $s_f$ for which (A34) is nonnegative. Expression (A34) equals zero for the value of $s_f$ in part (b) of the statement of the proposition. Substituting into (A38) and rearranging, (A38) is negative if (A32) is positive. If it is feasible for this scheme to deter crime, this scheme generates higher social surplus than shutting the firm down. Straightforward calculations show that this scheme generates higher social surplus than allowing crime if (A33) is negative.

Our usual specification of a high employee sanction, i.e., $s_e = \infty$, will not work here since forgiving a finite amount from an infinite sanction leaves an infinite sanction. Instead, we will set $s_e$ to a finite number $M$ that is large enough that the firm chooses not to indemnify the employee even if allowed by law. For example, it suffices to set $M = R + X$.

The proof is concluded by providing a synthesis of the preceding results into a complete characterization of the socially optimal sanction scheme. There are five exhaustive cases.

- If (A22) is positive, then the government can obtain the first best, deterring corporate crime with no deadweight loss, with a sanction scheme that does not target the employee and
settling

To understand (A41), consider each term in turn. If \( m \) accepted with probability \( p/s \) higher than \( m \) values of \( \beta \), the plaintiff wins, and the employee obtains utility \( \tilde{u}(m) \). For these values the plaintiff rejects the offer and goes to trial. With probability \( m \) values of \( \beta \), the firm and employee win, and his utility is \( \tilde{u}(m) \). Integrating over \( m \), the firm indemnifies the employee’s sanction \( s_e \), but since this lowers the firm’s net return he indirectly bears \( \beta \) of it. With probability \( 1 - m \), the firm and employee win, and his utility is \( \tilde{u}(m) \). Integrating over \( m \) for which the plaintiff rejects the settlement offer gives the second term in (A41).

The derivative of (A41) with respect to \( p \) is

\[
\frac{f(p/s)}{s} \left[ \tilde{u}(R - p) - \left( \frac{p}{s} \right) \tilde{u}(R - s - t) - \left( 1 - \frac{p}{s} \right) \tilde{u}(R - t) \right] - F(p/s)\tilde{u}'(R - p).
\]  

Proof of Proposition 5: As a preliminary step, we derive a sufficient condition for \( p_i^* \), the indemnified employee’s equilibrium settlement offer, to be an interior solution, i.e., \( p_i^* \in (0, s) \). Letting \( \tilde{u}(x) = u(\beta(R - x) + \ell_e) \) for conciseness, his objective function can be written

\[
F(p/s)\tilde{u}(R - p) + \int_{p/s}^{1} [m\tilde{u}(R - s - t) + (1 - m)\tilde{u}(R - t)] f(m) dm.
\]  

To understand (A41), consider each term in turn. If \( m \leq p/s \), the plaintiff earns more from settling \( p \) than from a trial in expectation \( (ms) \) so accepts the offer. The settlement offer is thus accepted with probability \( F(p/s) \), in which case the employee earns his share \( \beta \) of the firm’s net return \( R - p \), providing him with utility \( \tilde{u}(R - p) \). The second term of (A41) reflects values of \( m \) higher than \( p/s \). For these values the plaintiff rejects the offer and goes to trial. With probability \( m \), the plaintiff wins, and the employee obtains utility \( \tilde{u}(R - s - t) \). Note that the firm indemnifies the employee’s sanction \( s_e \), but since this lowers the firm’s net return he indirectly bears \( \beta \) of it. With probability \( 1 - m \), the firm and employee win, and his utility is \( \tilde{u}(R - t) \). Integrating over values of \( m \) for which the plaintiff rejects the settlement offer gives the second term in (A41).

The derivative of (A41) with respect to \( p \) is

\[
\frac{f(p/s)}{s} \left[ \tilde{u}(R - p) - \left( \frac{p}{s} \right) \tilde{u}(R - s - t) - \left( 1 - \frac{p}{s} \right) \tilde{u}(R - t) \right] - F(p/s)\tilde{u}'(R - p).
\]  

31
At $p = 0$, (A42) is positive: $f(0)[\tilde{u}(R) - \tilde{u}(R - t)]/s > 0$. Since 0 is in the support of $m$, $f(0) > 0$. At $p = s$, (A42) equals

$$\frac{f(1)}{s} [\tilde{u}(R - s) - \tilde{u}(R - s - t)] - \tilde{u}'(R - s).$$  \hspace{1cm} (A43)

Since 1 is in the support of $m$, $f(1) > 0$. Rearranging (A43), we see the expression is negative for

$$t < R - s - u^{-1}\left(\tilde{u}(R - s) - \frac{s}{f(1)}\tilde{u}'(R - s)\right).$$  \hspace{1cm} (A44)

Condition (A44) is sufficient for $p_f^* \in (0, s)$, since the indemnified employee’s objective function is increasing at the left endpoint and decreasing at the right.

As a second preliminary step, we derive a sufficient condition for $p_f^*$, the firm’s settlement offer in the counterfactual case in which it does not delegate settlement to the employee, to be an interior solution, i.e., $p_f^* \in (0, s)$. In the absence of the employee, the firm’s objective function for its settlement offer would be

$$F(p/s)(R - p) + \int_{p/s}^{1} [m(R - sf - tf) + (1 - m)(R - t_f)]f(m)dm.$$  \hspace{1cm} (A45)

The derivative is

$$\frac{t_s f(p/s) - F(p/s)}{s}.$$  \hspace{1cm} (A46)

At $p = 0$, (A46) is positive: $tf(0)/s > 0$. At $p = s$, (A46) equals $tf(1)/s - 1$, which is negative if

$$t < \frac{s}{f(1)}.$$  \hspace{1cm} (A47)

Condition (A47) is sufficient for $p_f^* \in (0, s)$, since the firm’s objective function is increasing at the left endpoint and decreasing at the right.

With these preliminaries done, we next turn to showing that indemnification lowers the employee’s choice of $p^*$. His objective function when he is indemnified, (A41), can be nested with that when he is not as follows:

$$F(p/s)\tilde{u}(R - p) + \int_{p/s}^{1} [m\tilde{u}(R - s - \theta t) + (1 - m)\tilde{u}(R - \theta t)]f(m)dm.$$  \hspace{1cm} (A48)

where $\theta = 1$ under indemnification and $\theta = 1/\beta$ under no indemnification. The cross partial derivative of (A48) with respect to $p$ and $\theta$ is

$$\left(\frac{t_s}{s} f(p/s) \left[\left(\frac{P}{s}\right) \tilde{u}'(R - s - \theta t) + \left(1 - \frac{P}{s}\right) \tilde{u}'(R - \theta t)\right].$$  \hspace{1cm} (A49)

Since (A49) is positive, Milgrom and Shannon’s (1994) Monotone Selection Theorem implies that $p^*$ is weakly higher when $\theta = 1/\beta$ than when $\theta = 1$. If (A44) holds, implying, as we have seen, that $p^*$ is an interior solution when $\theta = 1$, by Edlin and Shannon’s (1998) Strict
Monotonicity Theorem 1, \( p^* \) is strictly higher for \( \theta = 1/\beta \) than for \( \theta = 1 \). Matching up the values of \( \theta \) with the relevant cases, we have that \( p^* \) is weakly higher when the agent is not indemnified than when he is, and strictly so if (A44) holds.

We next show that \( p^*_i > p^*_f \). Since the firm’s objective function (A45) is assumed to be quasiconcave, the first derivative is non-negative for \( p \leq p^*_f \); i.e.,

\[
\frac{t}{s} f(p/s) - F(p/s) \geq 0 \quad \text{for } p \leq p^*_f. \quad (A50)
\]

Turning to the employee’s first-order condition, we have that (A42) exceeds

\[
\frac{f(p/s)}{s} \left[ \tilde{u}(R - p) - \left( \frac{p}{s} \right) \tilde{u}(R - s - t) - \left( 1 - \frac{p}{s} \right) \tilde{u}(R - t) - t\tilde{u}'(R - p) \right] > \frac{f(p/s)}{s} \left[ \tilde{u}(R - p) - \tilde{u}(R - p - t) - t\tilde{u}'(R - p) \right]. \quad (A51)
\]

Condition (A51) follows from substituting (A50) into (A42) and rearranging. Condition (A52) follows from the concavity of \( \tilde{u} \). The concavity of \( \tilde{u} \) also implies \( t\tilde{u}'(R - p) < \tilde{u}(R - p) - \tilde{u}(R - p - t) \). Hence (A52) is positive, implying (A42) is positive for all \( p \leq p^*_f \). Therefore, \( p^*_i \) cannot be in \([0, p^*_f]\) because (A42) is 0 at \( p^*_f \). Hence \( p^*_i > p^*_f \).

The proof concludes with a remark about conditions for the quasiconcavity of the firm’s objective function (A45). It can be shown that (A45) is quasiconcave if \( m \) follows a beta \((r, s)\) distribution for \( r = s \geq 1 \), which nests the uniform \([0, 1]\) case. For sufficiently leptokurtic beta \((r, s)\) distributions, for example taking \( r = s \ll 1 \), it can be shown that (A45) is no longer quasiconcave. Q.E.D.
References


