

Should Firms be Allowed to Indemnify Their Employees for Sanctions?

Wallace P. Mullin*
George Washington University

Christopher M. Snyder**
Dartmouth College

Policymakers have questioned whether firms should be allowed to indemnify their employees for personal sanctions for corporate crimes. This article provides the first formal analysis of this form of indemnification. Targeting employees with unindemnifiable sanctions carries the social cost of exposing employees of law-abiding firms to the risk of mistaken government prosecution. Deterrence is typically achieved more efficiently by sanctioning the firm alone. We find the circumstances under which the government should additionally sanction employees to be quite limited and the circumstances under which the government should ban indemnification of these sanctions to be more limited still. One circumstance is when an unindemnifiable employee sanction provides prosecutors with leverage to adjust the employee's sanction in exchange for his cooperation against the firm. (*JEL* K22, D82, L20)

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 US 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

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

**Department of Economics, Dartmouth College. Email: chris.snyder@dartmouth.edu.

We 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, MIT, Michigan State University, University of British Columbia, University of Delaware, the Allied Social Sciences Association "Contracts and Organizations" session, the International Industrial Organization Conference (Chicago), the NBER Law and Economics Summer Institute, and the Securities and Exchange Commission for helpful comments. We are grateful to editors Tracy Lewis and Ian Ayres and to several anonymous referees for their suggestions, which improved the article considerably.

corporations to include mandatory indemnification in their corporate charters or bylaws. According to a recent survey, 98% of US firms with over 500 shareholders had D&O insurance (Tillinghast-Towers Perrin 2002).

Although 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: 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 US 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, reported in the *Chicago Sun Times* (June 17, 2003: 49): “For the SEC to come out and say you can’t get insurance for these things, I think they are going pretty far.”

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 *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 also ubiquitous, suggesting it provides some important social benefit.

This article 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 a perfect enforcer. The possibility of type-I enforcement errors—that is, 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 et al., 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

1. Chairman William H. Donaldson, “Remarks Before the New York Financial Writers Association,” June 5, 2003, www.sec.gov/news/speech/spch060503whd.htm.

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 *de jure* sanctions cannot vary with actual guilt—imperfect enforcement prevents this—bankrupting the agent allows the *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.

Section 2 sets up the model. We solve for equilibrium using backward induction, first determining the optimal employment contract taking government policy as given in Section 3 and then determining the optimal government policy (sanctions and policy toward indemnification) in Section 4. Section 5 extends the basic model to allow the employee to cooperate with the prosecution and provides an analysis of this extended model. Section 6 contains the literature review, and Section 7 concludes with a discussion of the implications for recent controversies and directions for future research.

2. 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.² 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, that is, a payment from the firm to the employee conditional on conviction.

The firm is risk neutral. The employee is risk averse. Let $u : \mathbf{R}^+ \rightarrow \mathbf{R}^+$ 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 ℓ_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 ℓ_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

2. 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.

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, and 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 the firm.

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

$$r(a) - w(a) - g(a)(s_f + s_i), \quad (1)$$

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

$$g(a)u(\max\{0, \ell_e + w(a) + s_i - s_e\}) + [1 - g(a)]u(\ell_e + w(a)) - c(a). \quad (2)$$

The employee adds the wage $w(a)$ to his existing wealth ℓ_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 ℓ_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, equation (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 both 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 e.g., 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 et al. (2001).

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-detering 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 ($s_e \leq \ell_e + w^*(0)$, where $w^*(0)$ is the equilibrium wage). 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 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.

Although the statement of Proposition 3 focuses on that part of the parameter space for which 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 is 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 nonempty 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, that is, 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 model in which the government also has a prosecutorial 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 $\alpha \in (0, 1)$ if the employee does not cooperate with the prosecutors and unity if he does.³ 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) αg_0 if no crime was committed, αg_1 if a crime was committed and the employee does not cooperate with prosecutors, and g_1 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 $s_c \in [0, s_e]$ be the amount of the sanction forgiven. If the firm fully indemnifies the employee, by setting $s_i = s_e - s_c$, the prosecutors’ strategy will not work since the employee will not care about reducing the sanction.⁴ 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. 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 nonempty set of parameters, in which the second-best sanction scheme requires a positive employee sanction.

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

4. In the basic model, we took s_i to be a constant without loss of generality. In the extended model of the present section, we will take s_i to be proportional to the employee’s realized liability, $s_e - s_c$. Full indemnification is equivalent to $s_i = s_e - s_c$.

- (a) For the first set of parameters, the first best can be approached in the limit as $\varepsilon \rightarrow 0$ with a sanction scheme involving a small employee sanction $s_e = \varepsilon$ that is completely forgiven in exchange for cooperation ($s_c = \varepsilon$).
- (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 deadweight 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 article in which banning indemnification can be socially beneficial.

5.3 Other Forms of Cooperation

Thus far in Section 5, cooperation has been interpreted as providing prosecutors with additional evidence which increases the probability of conviction. The main result of the section will continue to hold if we broaden the interpretation of cooperation. Rather than the provision of new evidence, cooperation can be interpreted as a refusal to cover up existing evidence that would increase the probability of conviction or provide a better idea of the severity of the harm from the crime (and consequently the sanction if convicted). A related interpretation, adapting an idea from Arlen (1994), is that cooperation

relates to the maintenance of monitoring systems that increase the chance of alerting the government to the corporate crime.

If the firm is allowed to fully indemnify the employee, it can eliminate his incentive to cooperate under any of these new interpretations. As in Section 5.2, the government can encourage the employee's cooperation under these new interpretations by banning indemnification but then forgiving some of the employee's sanction if he is found to have cooperated. The only care that one needs to exercise when reinterpreting cooperation in these new ways is to understand the additional information burden that is placed on the government. Although prosecutors may automatically be able to verify that an employee has handed them new evidence in a case, they may have a harder time verifying that he has not covered up information or has maintained a monitoring system, especially in a setting in which type-I and type-II enforcement errors are being committed. Still, if prosecutors are able to verify these additional dimensions of employee cooperation with prosecution, this expands the circumstances under which banning indemnification becomes a useful prosecutorial tool.

6. 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 (2009) 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 our framework, 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 articles in the literature assume the crime benefits the firm. One exception is Privileggi et al. (2001). Our article 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 article.

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 articles 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 articles 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.⁵ In our work, the identity of the "cooperating" party is the opposite of theirs, the employee rather than the firm. More 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 articles.

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

5. Related issues arise in the literature on plea bargaining (e.g., Reinganum 1988; Kobayashi 1992) and on the use of leniency programs in the prosecution of cartels (e.g., Spagnolo 2000; Motta and Polo 2003; Aubert et al. 2006).

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.

7. Conclusion

This article has studied the private and social returns to indemnification when even law-abiding firms may face the risk of mistaken prosecution. The broad lesson to be drawn from the analysis is that authorities should be wary of sanctioning employees let alone banning their indemnification. Typically, firm sanctions deter crime more efficiently than unindemnifiable employee sanctions.

We uncovered only two circumstances under which the government should sanction the employee in addition to the firm. For a limited set of parameters, the government cannot deter crime using only a sanction against the firm without causing the shutdown of innocent firms (and consequent loss of these firms' socially valuable production). For a subset of these parameters, an employee sanction can help. The government can target the employee with a sanction high enough to bankrupt him. This sanction falls relatively harder on guilty than innocent firms because the employee must receive a higher wage to induce him to commit a crime, leaving him with more assets to be seized in bankruptcy. The government need not ban indemnification for this sanction scheme to work. The employee sanction can be set so high that the firm would prefer not to indemnify him even if allowed by law.

Indeed, the basic model analyzed in Sections 2–4 does not provide any rationale for the government to ban indemnification. To provide such a rationale, we extended the model in Section 5 to allow the probability of conviction to increase if the employee cooperates with prosecutors. In this extension, we found a limited set of parameters for which the socially optimal sanction scheme involved a small employee sanction which is forgiven if the employee cooperates with prosecutors. The firm would have an incentive to fully indemnify the employee to induce him not to cooperate, so indemnification must be banned for this sanction strategy to work.

For the broad set of remaining parameters, as proved in Proposition 2, targeting the employee and banning his indemnification is socially harmful. Indemnification reduces the deadweight loss from the mistaken conviction of innocent firms by allowing the risk-neutral firm an inexpensive way to insure the risk-averse employee against such errors. Our assumption that the agent is risk averse implicitly focused our analysis on indemnification of individuals (directors, officers, and other employees) within the firm. This is the leading case in policy debates concerning indemnification. Another interesting case, however, regards agency relationships a potentially criminal corporation may

have with other large, presumably risk-neutral, firms, for example, its input suppliers, investment bank, accounting firm, law firm, or insurance company, termed “gatekeepers” by Kraakman (1986). Since indemnification has little if any insurance value for such agents, the argument against banning their indemnification is weakened.

Although the government was an imperfect enforcer in our model, it was still assumed to be benevolent. In future work, it would be useful to analyze an alternative model in which the legal action is brought by a nonbenevolent party, whether a government with objectives other than welfare maximization or a private party with the opportunity to bring a nuisance suit. The social benefit of indemnification would presumably be greater in this alternative model and the case against banning indemnification presumably stronger.

Another avenue for future work is to depart from the simplifying assumption that the principal is a unitary actor, breaking open the black box of the firm in various dimensions. In our model, the firm indemnifies the employee directly itself, whereas in practice indemnification is often provided by third-party D&O insurance. It would be useful to understand the equilibrium effects of moving from self to third-party indemnification. One possibility, following Holderness (1990), is that the third party functions as an additional monitor to ensure the indemnification payouts are only for acts taken in good faith. The black box of the firm can also be expanded by studying possible conflicts between shareholders and directors. Directors, who have more direct control over compensation and benefits packages, may protect themselves with stronger indemnification policies or more D&O insurance than shareholders would prefer. Similar agency conflicts could arise for employees that as officers in the firm can influence their own compensation packages (see Kuhnen and Zwiebel 2007). Accounting for these additional sources of principal-agent conflicts within the black box of the firm might change the calculus of optimal public policy toward indemnification.

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 ℓ_e is suppressed.

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

$$r(a) - w(a) - g(a)(s_f + s_i) \quad (\text{A1})$$

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

$$g(a) u(\max\{0, \ell_e + w(a) + s_i - s_e\}) + [1 - g(a)] u(\ell_e + w(a)) - c(a) \geq u(\ell_e). \quad (\text{A2})$$

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

$$s_e \geq \ell_e + w(a) + s_i, \quad (\text{A3})$$

$$[1 - g(a)] u(\ell_e + w(a)) - c(a) = u(\ell_e); \quad (\text{A4})$$

or both (A5) and (A6) hold:

$$s_e \leq \ell_e + w(a) + s_i, \quad (\text{A5})$$

$$g(a) u(\ell_e + w(a) + s_i - s_e) + [1 - g(a)] u(\ell_e + w(a)) - c(a) = u(\ell_e). \quad (\text{A6})$$

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:

$$\mathcal{L} = r(a) - w(a) - g(a)(s_f + s_i) + \lambda\{g(a) u(\ell_e + w(a) + s_i - s_e) + [1 - g(a)] u(\ell_e + w(a)) - c(a) - u(\ell_e)\}. \quad (\text{A7})$$

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

$$1/\lambda = g(a)u'(\ell_e + w(a) + s_i - s_e) + [1 - g(a)]u'(\ell_e + w(a)) \quad (\text{A8})$$

$$1/\lambda = u'(\ell_e + w(a) + s_i - s_e). \quad (\text{A9})$$

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

$$s_e > \frac{b(g(a), c(a)) - b(0, c(a))}{g(a)}. \quad (\text{A10})$$

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

$$u^{-1}(u(\ell_e) + c(a)) = u^{-1}\left(g(a)(0) + [1 - g(a)]\left[\frac{u(\ell_e) + c(a)}{1 - g(a)}\right]\right), \quad (\text{A11})$$

$$\leq g(a)u^{-1}(0) + [1 - g(a)]u^{-1}\left(\frac{u(\ell_e) + c(a)}{1 - g(a)}\right), \quad (\text{A12})$$

$$= [1 - g(a)]u^{-1}\left(\frac{u(\ell_e) + c(a)}{1 - g(a)}\right). \quad (\text{A13})$$

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 ,

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

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

$$g(a)u(\ell_e + y(a, s_e) - s_e) + [1 - g(a)]u(\ell_e + y(a, s_e)) - c(a) = u(\ell_e). \quad (\text{A15})$$

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,

$$R - y(0, s_e) + u(\ell_e) + g_0s_e. \quad (\text{A16})$$

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

$$s'_f \geq \frac{X - y(1, 0)}{g_1 - g_0}. \quad (\text{A17})$$

Without loss of generality, we will take the value of s'_f such that equation (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,

$$R + u(\ell_e). \quad (\text{A18})$$

Expected social surplus is higher under the new sanction scheme if and only if equation (A18) exceeds equation (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,

$$tu(x') + (1 - t)u(x'') < u(tx' + (1 - t)x'') \quad (\text{A19})$$

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

$$g_0u(\ell_e + g_0s_e - s_e) + (1 - g_0)u(\ell_e + g_0s_e) < u(\ell_e). \quad (\text{A20})$$

Substituting $a = 0$ into (A15) implies

$$g_0u(\ell_e + y(0, s_e) - s_e) + (1 - g_0)u(\ell_e + y(0, s_e)) - c(a) = u(\ell_e). \quad (\text{A21})$$

Since $u' > 0$, (A20) and (A21) together imply $g_0 s_e < y(0, s_e)$ and thus that expected social welfare increases by moving to the sanction scheme with no employee sanction. 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:

$$b(0, C) + \left(\frac{g_1 - g_0}{g_0} \right) R - X, \quad (\text{A22})$$

$$H - R - X + b(0, C), \quad (\text{A23})$$

$$X - \left(\frac{g_1 - g_0}{g_0} \right) R + \left(\frac{g_1}{g_0} \right) b(g_0, 0) - b(g_1, C), \quad (\text{A24})$$

$$X - H + b(g_0, 0) - b(0, C). \quad (\text{A25})$$

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

$$R - g_0 s_f. \quad (\text{A26})$$

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

$$R + X - b(0, C) - g_1 s_f. \quad (\text{A27})$$

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, $u(\ell_e)$, 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

$$R - b(g_0, 0) - g_0 s_f. \quad (\text{A28})$$

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

$$R + X - b(g_1, C) - g_1 s_f. \quad (\text{A29})$$

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:

$$s_f = \frac{R - b(g_0, 0)}{g_0}. \quad (\text{A30})$$

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 αg_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

$$R - b(\alpha g_0, 0) - \alpha g_0 s_f. \quad (\text{A34})$$

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

$$g_1 u(\ell_e - s_e + s_c + w^*) + (1 - g_1) u(\ell_e + w^*) - C, \quad (\text{A35})$$

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

$$\alpha g_1 u(0) + (1 - \alpha g_1) u(\ell_e + w^*) - C, \quad (\text{A36})$$

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_c forces (A35) to equal (A36), implying

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

in turn implying the employee's surplus from cooperating is

$$(1 - \alpha g_1) u(\ell_e + w^*) - C, \quad (\text{A37})$$

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

$$R + X - b(\alpha g_1, C) - g_1 s_f. \quad (\text{A38})$$

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 non-negative. 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, that is, $s_e = \infty$, 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 does not seek employee cooperation with prosecutors. In particular, the scheme $s_f = R/g_0$ and $s_e = s_c = 0$ suffices.
- If (A22) is negative and (A31) is positive, a socially optimal sanction scheme does not exist because of an “open set” problem, but the first best can be approached arbitrarily closely in the limit as $\varepsilon \rightarrow 0$ with sanctions $s_f = (R - \varepsilon)/\alpha g_0$ and $s_c = s_e = \varepsilon$.
- If (A22) and (A31) are negative and (A23) and (A32) are positive, then the socially optimal sanction scheme deters crime by shutting down the firm. In particular, the scheme $s_f = \infty$ and $s_e = s_c = 0$ suffices.
- If (A22), (A23), and (A31) are negative and (A33) is positive, then the socially optimal sanction scheme does not deter crime. In particular, the scheme $s_f = s_e = s_c = 0$ suffices.
- For the remaining cases in which (A22), (A23), (A31), (A32), and (A33) are nonzero, the following sanction scheme obtains the second best:

$$s_f = \frac{1}{\alpha g_0} \left[R - u^{-1} \left(\frac{u(\ell_e)}{1 - \alpha g_0} \right) + \ell_e \right], \quad (\text{A39})$$

$$s_c = M - u^{-1} \left(\frac{u(\ell_e) + C}{1 - \alpha g_1} \right) + u^{-1} \left(\frac{(1 - \alpha)[u(\ell_e) + C]}{1 - \alpha g_1} \right), \quad (\text{A40})$$

and $s_e = M$, where M is a sufficiently large number (e.g., $M = R + X$ suffices). Q.E.D.

References

- Arlen, Jennifer. 1994. "The Potentially Perverse Effects of Corporate Criminal Liability," 23 *Journal of Legal Studies* 833–67.
- . 1998. "Corporate Crime and its Control," in P. Newman, ed., *The New Palgrave Dictionary of Economics and the Law*. New York: Macmillan.
- Arlen, Jennifer, and Reiner Kraakman. 1997. "Controlling Corporate Misconduct: An Analysis of Corporate Liability Regimes," 72 *New York University Law Review* 687–779.
- Aubert, Cécile, Patrick Rey, and William Kovacic. 2006. "The Impact of Leniency and Whistleblowing Programs on Cartels," 24 *International Journal of Industrial Organization* 1241–66.
- Chu, C. Y., and Yingyi Qian. 1995. "Vicarious Liability under a Negligence Rule," 15 *International Review of Law and Economics* 305–22.
- Conard, Alfred F. 1972. "A Behavioral Analysis of Directors' Liability for Negligence," 21 *Duke Law Journal* 895–919.
- Davis, Michael L. 1996. "The Impact of Rules Allocating Legal Responsibilities Between Principals and Agents," 17 *Managerial and Decision Economics* 413–20.
- Easterbrook, Frank H., and Daniel R. Fischel. 1991. *The Economic Structure of Corporate Law*. Cambridge, MA: Harvard University Press.
- Garoupa, Nuno. 2000. "Corporate Criminal Law and Organization Incentives: A Managerial Perspective," 21 *Managerial and Decision Economics* 243–52.
- Grossman, Sanford, and Oliver Hart. 1983. "An Analysis of the Principal-Agent Problem," 51 *Econometrica* 7–45.
- Harrington, Scott E., and Gregory R. Niehaus. 1998. *Risk Management and Insurance*. New York: McGraw Hill.
- Holderness, Cliff. 1990. "Liability Insurers as Corporate Monitors," 10 *International Review of Law and Economics* 115–29.
- Kornhauser, Lewis. 1982. "An Economic Analysis of the Choice Between Enterprise and Personal Liability for Accidents," 70 *California Law Review* 1345–92.
- Kobayashi, Bruce H. 1992. "Deterrence with Multiple Defendants: An Explanation for 'Unfair' Plea Bargains," 23 *Rand Journal of Economics* 507–17.
- Kraakman, Reinier. 1984. "Corporate Liability Strategies and the Costs of Legal Controls," 93 *Yale Law Journal* 857–98.
- . 1986. "Gatekeepers: The Anatomy of a Third-Party Enforcement Strategy," 2 *Journal of Law, Economics, & Organization* 53–104.
- Kuhnen, Camelia M., and Jeffrey Zwiebel. 2007. "Executive Pay, Hidden Compensation and Managerial Entrenchment." Working Paper, Northwestern University Kellogg School of Management.
- Macey, Jonathan R. 1991. "Agency Theory and the Criminal Liability of Organizations," 71 *Boston University Law Review* 315–40.
- Motta, Massimo, and Michele Polo. 2003. "Leniency Programs and Cartel Prosecution," 21 *International Journal of Industrial Organization* 347–79.
- Mullin, Wallace, and Christopher Snyder. 2009. "Corporate Crime," in N. Garoupa, ed., *Encyclopedia of Law and Economics, Volume 11: Criminal Law and Economics*. Cheltenham, UK: Edward Elgar.
- Newman, Harry A., and David W. Wright. 1990. "Strict Liability in a Principal-Agent Model," 10 *International Review of Law and Economics* 219–31.
- Polinsky, A. Mitchell, and Steven Shavell. 1993. "Should Employees Be Subject to Fines and Imprisonment Given the Existence of Corporate Liability?" 13 *International Review of Law and Economics* 239–57.
- Privileggi, Fabio, Carla Marchese, and Alberto Cassone. 2001. "Agent's Liability Versus Principal's Liability When Attitudes Toward Risk Differ," 21 *International Review of Law and Economics* 181–95.
- Reinganum, Jennifer. 1988. "Plea Bargaining and Prosecutorial Discretion," 78 *American Economic Review* 713–28.

- Segerson, Kathleen, and Tom Tietenberg. 1992. "The Structure of Penalties in Environmental Enforcement: An Economic Analysis," 23 *Journal of Environmental Economics and Management* 179–200.
- Shavell, Steven. 1986. "The Judgment Proof Problem," 6 *International Review of Law and Economics* 45–58.
- . 1987. *Economic Analysis of Accident Law*. Cambridge, MA: Harvard University Press.
- . 1997. "The Optimal Level of Corporate Liability Given the Limited Ability of Corporations to Punish Their Employees," 17 *International Review of Law and Economics* 203–13.
- Spagnolo, Giancarlo. 2000. "Optimal Leniency Programs." Working Paper 42.00, Fondazione Eni Enrico Matteing.
- Stone, Christopher D. 1980. "The Place of Enterprise Liability in the Control of Corporate Conduct," 90 *Yale Law Journal* 1–77.
- Sykes, Alan O. 1984. "The Economics of Vicarious Liability," 93 *Yale Law Journal* 1231–80.
- Tillinghast-Towers Perrin. 2002. *2002 Directors and Officers Liability Survey*. Chicago: Tillinghast-Towers Perrin.