Environmental Risk Insurance under Dynamic Moral Hazard[*]

Bruno Biais[†]   Thomas Mariotti[‡]   Jean-Charles Rochet[§]   Stéphane Villeneuve[¶]

First Draft: February 2007
This Draft: April 2007

Abstract

A firm is subject to accident risk, which the manager can mitigate by exerting effort. An agency problem arises because effort is unobservable and the manager has limited liability. The occurrence of accidents is modelled as a Poisson process, whose intensity is controlled by the manager. We use martingale techniques to formulate the manager's incentive compatibility constraints. The optimal contract is characterized by a differential equation with delay. The manager receives cash transfers only if no accident occurs during a sufficiently long period of time, while the firm is downsized if accidents are too frequent. This can be implemented by cash reserves, along with insurance, financial, and compensation contracts. The insurance contract involves a deductible and a bonus-penalty system. The financial contract consists of bonds that pay constant coupons until the firm enters financial distress. Covenants requests that the firm be downsized when its liquidity ratio falls below a threshold. The manager's compensation policy promises incentive wages when the accumulated performance of the firm is strong enough. Our theoretical analysis also delivers new empirical implications about the dynamics of insurance premia and credit yield spreads.

[†]Université de Toulouse (GREMAQ/CNRS, IDEI, CRG) and CEPR.
[‡]Université de Toulouse (GREMAQ/CNRS, IDEI) and CEPR.
[§]Université de Toulouse (GREMAQ, IDEI) and CEPR.
[¶]Université de Toulouse (GREMAQ).

1. INTRODUCTION

On March 23, 2005, a series of explosions occurred at the BP Texas City refinery, resulting in 15 deaths, more than 170 injuries and significant economic losses.[1] The U.S. Chemical Safety and Hazard Investigation Board (CSB) as well as an independent panel, chaired by James A. Baker, investigated the case. Both concluded that the accident occurred because BP's management had not exerted sufficient risk prevention effort. In October 2006, CSB Chairman Carolyn W. Merritt stated, "BP's global management was aware of problems with maintenance, spending, and infrastructure well before March 2005. [...] Unsafe and antiquated equipment designs were left in place, and unacceptable deficiencies in preventative maintenance were tolerated."[2] In January 2007, the Baker Panel Report concluded that "BP executive and corporate refining management have not provided effective process safety leadership."[3] In addition to such dramatic events as the BP explosion, numerous smaller scale accidents occur regularly in industrial firms, generating serious damages.[4] Systematic analyses of these accidents concur with the above reported conclusions: Leadership and managerial failures are a major cause of risk.[5]

Under laisser-faire, firms will not fully internalize the externalities generated by such industrial risk. This will result in socially suboptimal risk prevention efforts. One way to stimulate risk prevention would be to make firms and managers bear the social cost of accidents. For example, they could have to compensate all the parties hurt in an accident and to clean up environmental damages. Yet, such Pigovian taxes are often impossible to enforce in practice. This is because the size of the total damages often exceeds the wealth of the managers and even the net worth of the firms, which are protected by limited liability and bankruptcy laws.[6] As a result of this, managers and firms have little incentives to lower the probability of accidents generating losses that greatly exceed the value of their assets.[7] Nevertheless, if risk prevention were observable, it would still be possible to provide managers with appropriate incentives to exert socially optimal risk prevention efforts. To a large extent, however, managerial risk prevention efforts are unobservable to external parties. This leads to a moral hazard problem.

Besides moral hazard, one important dimension of risk prevention in industrial projects is time. Indeed, severe accidents are rare and dramatic events. This contrasts with day-to-day firm operations and cash-flows. It is therefore natural to study moral hazard in risk prevention in a dynamic set-up, in which the timing of environmental risk differs from that of firm operations. To do so, we focus on the simplest possible model: operating cash-flows are constant per unit of time, while accidents follow a Poisson process whose intensity depends

---

[1]Other examples of large accidents include toxic gas leaks in Bhopal in 1989, oil spills such as Exxon Valdes in 1989, Erika in 2000 and Prestige in 2002, and the chemical explosion at AZF in Toulouse in 2001.

[2]"CSB Investigation of BP Texas City Refinery Disaster Continues as Organizational Issues are Probed," CSB News Release, October 30, 2006.

[3]"The Report of the BP U.S. Refineries Independent Safety Review Panel," January 16, 2007.

[4]Elliott, Wang, Lowe and Kleindorfer (2004) report that, out of 15,083 facilities storing hazardous material in the U.S. between 1994 and 1999, 4.4% had an accident causing worker or public injury.

[5]See for instance Leplat and Rasmussen (1984), Gordon, Flin, Mearns and Fleming (1996), or Hollnagel (2002).

[6]As Katzman (1988) reports, "In *Ohio v. Kovacs* (U.S.S.C. 83–1020), the U.S. Supreme Court unanimously ruled that an industrial polluter can escape an order to clean up a toxic waste site under the umbrella of federal bankruptcy."

[7]See Shavell (1984, 1986) for a discussion of how a party's inability of paying for the full magnitude of harm done dilutes its incentives to reduce risk.

on the level of risk prevention.

In this context, we study the optimal contract providing appropriate risk prevention incentives under dynamic moral hazard. The manager of the firm is risk-neutral and protected by limited liability. She can exert effort to reduce the instantaneous probability of accidents.[8] Effort is costly to the manager and unobservable to other parties. We characterize the set of implementable risk prevention policies. The optimal contract maximizes the expected benefit to society from an implementable risk prevention policy, subject to the constraint that the manager receives at least her reservation utility. The optimal contract relies on two instruments, non-negative transfers to the manager and irreversible downsizing of the firm. The former serve as a reward to motivate the manager, while the latter is used to punish her. We assume constant returns to scale, in that downsizing reduces by the same factor the size of operating profits, the social costs of accidents, and the manager's temptation to shirk. In the optimal contract, downsizing and payout decisions are functions of the entire past history of the accident process. However, we show that this complex history dependence can be summarized by two state variables: the size of the firm, reflecting past downsizing decisions, and the continuation utility of the manager, reflecting the promise of future transfers. The main features of the optimal contract are as follows.

C1. For a given firm size, incentive compatibility requires that the continuation utility of the manager be reduced by at least a certain amount following an accident. This punishment motivates the manager to exert effort. The greater the magnitude of the moral hazard problem, the greater the necessary punishment. The induced sensitivity of the manager's continuation utility to the random occurrence of accidents is socially costly because the value function is concave in this state variable. Therefore, it is optimal to set the reduction in the manager's continuation utility following an accident to the minimum level consistent with incentive compatibility.

C2. Irreversible downsizing is costly, since it reduces the scale of operation of a positive net present value project. Hence, the firm is downsized only when this cannot be avoided. This is the case when, following an accident, the continuation utility of the manager becomes so low that it cannot be further reduced without violating her limited liability constraint. In that situation, the threat that can be used to motivate the manager is limited. To cope with this limitation, it is necessary to lower the manager's temptation to shirk by reducing the scale of operation of the firm. Apart from such circumstances, and in particular when no accident occurs, the firm is never downsized.

C3. In addition to these threats, the promise of future compensation helps motivating the manager. If the initial utility of the manager is relatively low, there is a probation phase after the firm is set up during which the manager does not receive any compensation. This typically occurs if the frequency or the social cost of accidents are high. Then, if a sufficiently long period of time elapses with no accidents occurring, the manager starts receiving a constant wage per unit of time. But, as soon as an accident occurs, the contract reverts to a probation phase.

---

[8]Unlike in Shapiro and Stiglitz (1984) or Akerlof and Katz (1989), effort merely makes accidents less likely, but does not allow to avoid them altogether. As a result of this, accidents do occur on the equilibrium path, and it is no longer optimal to systematically fire the manager following an accident.

This abstract optimal mechanism can be implemented using instruments consistent with arrangements observed in practice. In the implementation, the firm is requested to contract with an insurance company.[9] The latter is liable for the accident cost and makes sure the manager has appropriate incentives to prevent risk. The implementation of the optimal contract is based on four main instruments.

I1. The firm holds cash reserves, that are held on its bank account and remunerated at the market rate. As in Biais, Mariotti, Plantin and Rochet (2007), cash reserves mirror the evolution of the continuation utility of the manager. In the implementation, the assets of the firm reflect its cash reserves and the size of its operations. The cash-flow statement of the firm is characterized, as a direct implication of the implementation of the optimal contract.

I2. Our implementation also involves contracting with an insurance company. The latter is liable for the cost of accidents, minus a deductible, which is paid by the firm out of its cash reserves. The payment of this deductible reflects the manager's incentive compatibility constraint. In any period, the firm pays an insurance premium, combining an actuarial component with an incentive component. The latter can be interpreted as a bonus-penalty score. If no accident occurs for a long period of time, the firm enjoys a high bonus, which reduces its insurance premium. By contrast, the firm pays a high premium when claims frequency is high. The incentive component of the insurance premium decreases with the cash reserves of the firm. Finally, the insurance contract involves a downsizing covenant, which stipulates that if the liquidity ratio of the firm falls below a certain threshold, downsizing must take place.

I3. The compensation of the manager also reflects the evolution of the firm's liquidity ratio. After a long period without accident, the firm holds large cash reserves and the liquidity ratio reaches a high water mark. At this point, the manager is compensated by cash transfers as long as no accident occurs. This compensation is designed so that the liquidity ratio of the firm stays constant. As soon as an accident occurs, the liquidity ratio of the firm drops down, as the deductible is paid out of its cash reserves, and one reverts to the regime without immediate managerial compensation.

I4. Finally, the implementation includes a financial component. The firm issues a bond, paying a coupon that is proportional to the size of the firm. Thus, bondholders are exposed to the risk of downsizing.

This implementation of the optimal contract rationalizes several regulatory and contractual features observed in practice, such as compulsory insurance, deductibles, and bonus-penalty systems for insurance premia. It also delivers new implications. There should be a decreasing relationship between a firm's liquidity and the insurance premium it pays. We obtain an analytic characterization of the dynamics of the bonus-penalty score of the firm and the insurance premium. Firms subject to greater moral hazard should have insurance premia that decrease more strongly when there are no accidents and increase more sharply following accidents. We also characterize the dynamics of the value of the optimal insurance policy

---

[9]Compulsory insurance is observed in practice. In the U.S., the Resource Conservation and Recovery Act and the Comprehensive Environmental Response, Compensation, and Liability Act request firms to insure against third-party damages, unless they are financially strong enough to bear liability risk (Katzman (1988)).

and relate it to the moral hazard parameters and the history of the firm. Finally, our model generates unpredictable credit risk for the bond issued by the firm, in line with empirical evidence, and allows us to derive an explicit formula for the link between the credit yield spread on the bond and the liquidity ratio of the firm.

Our model is related to previous analyses of dynamic moral hazard. Unlike Holmström and Milgrom (1987), Rogerson (1985), and Sannikov (2003), we consider a risk-neutral agent, but we assume that the agent has limited liability, as in DeMarzo and Fishman (2003), Clementi and Hopenhayn (2006), DeMarzo and Sannikov (2006), or Biais, Mariotti, Plantin and Rochet (2007). While these two last papers model operating cash-flows as a Brownian motion with drift, we suppose that the manager controls the intensity of a Poisson accident process. This leads us to extend the martingale methods of Sannikov (2003) to the case of an unpredictable process. This gives rise to substantial differences in the optimal contract. In Sannikov (2003), DeMarzo and Sannikov (2006), and Biais, Mariotti, Plantin and Rochet (2007), the optimal contract is characterized by an ordinary differential equation, reflecting that incentives are provided locally, through infinitesimal changes in the continuation value of the manager. By contrast, in our model, the optimal contract is characterized by a differential equation with delay, reflecting that incentives are provided non-locally, through jumps in the continuation value of the manager. Another difference is that the optimal transfer process in DeMarzo and Sannikov (2006) and Biais, Mariotti, Plantin and Rochet (2007) is singular, and characterized by a local time that reflects the diffusion process followed by the manager's continuation utility at a dividend boundary. By contrast, the optimal transfer process that emerges from our analysis is regular, as the manager receives a constant wage per unit of time when she is effectively compensated. Finally, unlike in Sannikov (2003), DeMarzo and Sannikov (2006), and Biais, Mariotti, Plantin and Rochet (2007), the size of the firm is a key variable in our optimal continuous-time contract. Furthermore, changes in firm size are discrete and unpredictable. This contrasts with Brownian motion models, in which no size adjustments are required, and liquidation is predictable.

Contemporaneous work by Myerson (2007) also analyzes dynamic moral hazard in a Poisson framework. While we focus on environmental risk, he considers a political economy model, in which a sovereign prince seeks to deter his governors from corruption and rebellion. The formal analysis is quite different in the two papers. Myerson (2007) considers the case where the principal and the agent have the same discount rates. As explained in Biais, Mariotti, Plantin and Rochet (2007), this case is not conducive to continuous-time analysis, as an optimal contract does not exist. To cope with this difficulty, Myerson (2007) imposes an exogenous upper bound on the continuation utility of the agent. In that constrained problem, existence is restored. By contrast, we do not impose such an exogenous bound, but we consider the case where the principal is more patient than the agent. While this makes the formal analysis more complex, this also restores existence of an unconstrained optimal contract. The boundedness of the continuation utility of the agent is an endogenous feature of this contract.[10]

_____

[10]Poisson processes have also proved useful in the theory of repeated games with imperfect monitoring. Abreu, Milgrom and Pearce (1991) use a Poisson signal structure to model the arrival of information in this class of games, and study the impact of varying the discount rate or the duration of a period on the set of equilibrium outcomes. More recently, Kalesnik (2005) offers a partial characterization of the set of equilibrium outcomes in a continuous-time model of repeated partnerships with Poisson signals, and Sannikov and Skrzypacz (2006) study a mixed model in which the monitoring process is a sum of a Brownian component and of a Poisson component. Our focus differs from these papers in that we consider a full commitment

Sannikov (2005) also uses a Poisson payoff structure. His setup allows for both adverse selection and ex-post moral hazard. An essential difference with our analysis lies in the way jumps affect output. In Sannikov's (2005) cash-flow diversion model, jumps correspond to positive innovations of the cash-flow process. By contrast, in our insurance model, jumps correspond to accidents and thus to negative shocks on total output. In particular, accidents are less likely to happen if the manager exerts the equilibrium level of effort.[11] This leads to qualitatively very different results. While downsizing is a key feature of our optimal contract, as it ensures that incentives can still be provided following a long sequence of accidents, it plays no role in Sannikov (2005). Liquidation in his model is still required to provide incentives, but it corresponds to a deterministic and predictable event: if a sufficiently long period of time elapses during which which the manager reports no cash-flow, the firm is liquidated. By contrast, downsizing in our model is unpredictable. The implementations of the optimal contract reflect these differences in the interpretation of jumps. While Sannikov (2005) focuses on the role of credit lines, we stress the role of insurance contracts to counter accident risk.

This paper is also related to the rich Law and Economics literature on accident law. Shavell (1986, 2000) points out that a firm unable to pay the full amount for which it is legally liable has too little incentives to exert preventive effort, and tends to engage in risky activities to a socially excessive extent. He further argues that the desirability of liability insurance depends on the ability of insurers to monitor the firm's prevention effort, and to link insurance premia to the observed level of care. The basic point is that if insurers cannot observe the firm's level of care, making full liability insurance mandatory results in no care at all being taken.[12] In our dynamic analysis, we study how different features of the insurance contract can be designed to mitigate the adverse effects of this lack of observability. In particular, insurance premia contain an incentive component and can be adjusted upward or downward according to the accident record. Moreover, the optimal insurance contract we derive ties the firm's allowed activity level to its accident record: following a series of accidents, the firm can be forbidden to engage at full scale in its risky activity. These instruments provide the manager of the firm with dynamic incentives to exert the appropriate risk prevention effort, although the latter is not observed by the insurance company.

The paper is organized as follows. In Section 2, we present the model. In Section 3, we characterize the optimal contract. In Section 4, we discuss the implementation of the optimal contract and we spell out the testable implications of our model. Section 5 concludes. All proofs are in the appendices.

2. The Model

Time is continuous, and indexed by $t \geq 0$. There are two agents, the manager of a firm and an insurance company. The insurance company is risk-neutral and has a discount rate $r > 0$. The manager is also risk neutral and has a discount rate $\rho > r$. She is thus more impatient than the insurance company. The manager has limited liability, that is, contracts cannot stipulate negative payments to her.

The manager runs an industrial project that can be continuously operated over an infinite

---

contracting environment, in which we explicitly characterize the optimal incentive compatible contract.

[11] Thus jumps in our model are "bad news" in the sense of Abreu, Milgrom and Pearce (1991).

[12] See Jost (1996) and Polborn (1998) for important extensions and qualifications of this argument.

5

horizon, but may be downsized or liquidated at any date. Downsizing is irreversible. When the firm is downsized, a fraction of its assets are liquidated, and the firm continues to operate at a reduced scale. For simplicity, we assume that the liquidation value of the firm's assets is 0. For each $t \geq 0$, denote by X_t the scale of the firm's operations at date t. We assume the project has positive net present value. Hence reducing the scale of the project or liquidating it outright is socially costly. However, as shown below, downsizing after bad performance can be useful as a threat to the manager. This is in line with DeMarzo and Fishman (2003), Clementi and Hopenhayn (2006), DeMarzo and Sannikov (2006), and Biais, Mariotti, Plantin and Rochet (2007). Without loss of generality, we normalize to 0 the set-up cost of the project, as well as the initial cash endowment of the manager.

Previous continuous-time analyzes of principal-agent interactions have typically modelled operating profits as diffusion processes (see for instance Holmström and Milgrom (1987), Sannikov (2003), DeMarzo and Sannikov (2006)). By contrast, we assume that instantaneous profits are deterministic. In addition, we assume constant returns to scale: given firm size X_t, operating profits at date t are equal to $X_t\mu$, for some $\mu > 0$. We hereafter refer to μ as size-adjusted operating profits.

While size-adjusted operating profits are constant, the firm is subject to accident risk. The occurrence of accidents is modelled as a point process $N = \{N_t\}_{t \geq 0}$, where for each $t \geq 0$, N_t is the number of accidents up to and including date t. When an accident occurs, it creates social and environmental costs. These costs are borne by society at large rather than by the manager of the firm. For example, an oil spill imposes huge damages on the environment and on the inhabitants of the affected region, but has no direct impact on the manager of the oil company. Since the manager has limited liability, she cannot be held responsible for these costs in excess of her current wealth. As for operating profits, we assume constant returns to scale: given firm size X_t, the social cost of an accident is $X_t C$, for some size-adjusted social cost $C > 0$. Overall, for each $t \geq 0$, the net output flow generated by the firm during the infinitesimal time interval $[t, t + dt)$ is equal to

$$X_t(\mu dt - C dN_t).$$

By exerting effort, the manager affects the probability with which accidents occur. The manager's risk prevention effort at date t is equal to the intensity of the process N at date t, Λ_t. A higher level of effort reduces the probability $\Lambda_t dt$ of an accident during the infinitesimal time interval $[t, t + dt)$. For simplicity, we consider only two levels of managerial effort, $\Lambda_t \in \{\lambda, \lambda + \Delta\lambda\}$, with $\lambda > 0$ and $\Delta\lambda > 0$. To model the cost of effort, we adopt the same convention as Holmström and Tirole (1997). If the manager exerts low effort at date t, that is $\Lambda_t = \lambda + \Delta\lambda$, she receives a private benefit $X_t B$, for some size-adjusted private benefit $B > 0$. By contrast, when the manager exerts high effort at date t, that is $\Lambda_t = \lambda$, she receives no private benefit. This formulation is equivalent to one in which the manager incurs a constant cost per unit of time and per unit of size of the firm when exerting effort, and no cost when shirking. We let $\mu > \lambda C$, so that the expected instantaneous net output flow is positive when the manager exerts high effort. Under these circumstances, operating the project is socially preferable to closing the firm.

Unlike profits and accidents, the manager's effort decisions are assumed to be unobservable to the insurance company. This leads to a moral hazard problem, whose key parameters are B and $\Delta\lambda$. The larger is B, the more attractive it is to shirk. The smaller is $\Delta\lambda$, the more difficult it is to detect shirking. We assume that $C > B/\Delta\lambda$, so that the private

benefits of shirking are lower than the social cost of increased accident risk. In the absence of moral hazard, this implies that it is socially optimal to always require high effort from the manager.[13] In Subsection 3.3, we derive the more restrictive conditions under which this maximal risk prevention policy remains optimal under moral hazard.

The firm is required to obtain insurance against the environmental risk, so that third parties are protected against the social costs of accidents. The insurance company has deep pockets and is liable for the social costs. It designs the compensation contract of the manager to give her incentives to adopt an appropriate risk prevention policy. Three contractual instruments can be used to cope with the moral hazard problem and provide incentives to the manager:

(i) First, the firm can be downsized. Denote by $X = \{X_t\}_{t \geq 0}$ the non-increasing and non-negative process describing the size of the firm. This process is bounded above by some maximal initial scale of operations $X_{0^-} > 0$.

(ii) Next, non-negative transfers can be made to the manager. Denote by $L = \{L_t\}_{t \geq 0}$ the non-decreasing and non-negative process describing the cumulative transfers to the manager.

(iii) Last, the firm can be liquidated. Denote by τ the random time at which liquidation occurs. We allow τ to be infinite and, without loss of generality, we assume that $\tau \leq \inf\{t \geq 0 \,|\, X_t = 0\}$.

The contract between the insurance company and the manager is designed and agreed upon at date 0, after which the firm operates and the contract is enforced. We assume the insurance company and the manager can fully commit to a long-term contract $\Gamma = (X, L, \tau)$. Thus we abstract from imperfect commitment problems and focus on a single source of market imperfection, namely moral hazard in risk prevention. The manager reacts to the contract Γ by choosing an effort process $\Lambda = \{\Lambda_t\}_{t \geq 0}$. At any date t prior to liquidation, the sequence of events in the infinitesimal time interval $[t, t + dt)$ can be heuristically described as follows:

1. The size of the firm X_t is determined.

2. The agent takes her effort decision Λ_t.

3. With probability $\Lambda_t dt$, an accident occurs, in which case $dN_t = 1$.

4. The agent receives a transfer dL_t.

5. The firm is either liquidated or continued.

According to this timing, the downsizing and effort decisions are taken before knowing the current realization of the accident process. This can be formalized by requiring X and Λ to be $\mathcal{F}^N$–predictable, where $\mathcal{F}^N = \{\mathcal{F}^N_t\}_{t \geq 0}$ is the filtration generated by N. By contrast, payout and liquidation decisions at any date are taken after observing whether or not an accident occurred at this date. Hence L and τ are $\mathcal{F}^N$–adapted.[14] As described in detail in

---

[13]This allocation can be implemented using a constant transfer l per unit of time to the manager, conditional on her exerting effort. When $C > B/\Delta\lambda$, one can choose l such that $l \geq B$ and $\mu - \lambda C - l \geq \mu - (\lambda + \Delta\lambda)C$, so that all parties find high effort desirable.

[14]See for instance Dellacherie and Meyer (1978, Chapter IV, Definitions 12 and 61) for precise definitions of these concepts.

the Appendix, an effort process Λ generates a unique probability distribution $\mathbb{P}^\Lambda$ over the paths of the accident process N. Denote by $\mathbb{E}^\Lambda$ the corresponding expectation operator.

Given a contract $\Gamma = (X, L, \tau)$ and an effort process Λ, the expected discounted utility of the manager is

$$\mathbb{E}^\Lambda\left[\int_0^\tau e^{-\rho t}(dL_t + 1_{\{\Lambda_t = \lambda + \Delta\lambda\}} X_t B\, dt)\right], \tag{1}$$

while the insurance company obtains an expected discounted profit

$$\mathbb{E}^\Lambda\left[\int_0^\tau e^{-rt}[X_t(\mu dt - C dN_t) - dL_t]\right]. \tag{2}$$

An effort process Λ is incentive compatible with respect to a contract Γ if it maximizes the manager's expected utility (1) given Γ. The problem of the insurance company is to find a contract Γ and an incentive compatible effort process Λ that maximize its expected discounted profit (2), subject to delivering to the manager a required expected discounted utility level. It is without loss of generality to focus on contracts Γ such that the present value of the payments to the manager is finite, that is:

$$\mathbb{E}^\Lambda\left[\int_0^\tau e^{-\rho t} dL_t\right] < \infty. \tag{3}$$

Indeed, by inspection of (2), if the present value of the payments to the manager were infinite, the fact that $\rho > r$ would imply infinitely negative expected discounted profits for the insurance company. The latter would be better off proposing no contract altogether.

Remark. In the presentation of the model, we found it convenient to directly introduce the insurance company as one of the parties, without waiting for the implementation to do so. This is without loss of generality as the insurance company is assumed to cover all the accident costs, and thus to represent the interests of society at large. Equivalently, one could reinterpret the model as one in which the manager of the firm, acting as an agent, contracts with the rest of society, acting as a principal.

3. The Optimal Contract

In this section, we first formulate the manager's incentive compatibility constraints. Next, we derive the optimal contract that induces the manager to always exert the high prevention effort. Last, we derive conditions under which it is indeed optimal to require this level of prevention from the manager.

3.1. Incentive Compatibility

In this subsection, we derive the incentive compatibility constraint of the manager, relying on martingale techniques similar to those introduced by Sannikov (2003). When deciding which effort decision to take at a date t, the agent considers how this decision will affect his continuation utility, defined as

$$W_t(\Gamma, \Lambda) = \mathbb{E}^\Lambda\left[\int_t^\tau e^{-\rho(s-t)}(dL_s + 1_{\{\Lambda_s = \lambda + \Delta\lambda\}} X_s B ds)\,|\,\mathcal{F}_t^N\right] 1_{\{t < \tau\}}. \tag{4}$$

Denote by $W(\Gamma, \Lambda) = \{W_t(\Gamma, \Lambda)\}_{t \geq 0}$ the manager's continuation utility process. Note that, by construction, $W(\Gamma, \Lambda)$ is $\mathcal{F}^N$–adapted. In particular, $W_t(\Gamma, \Lambda)$ reflects whether an accident

occurred or not at date t. To characterize how the manager's continuation utility evolves over time, it is useful to consider her lifetime expected utility, evaluated conditionally upon the information available at date t, that is:

$$U_t(\Gamma, \Lambda) = \mathbb{E}^{\Lambda}\left[\int_0^{\tau} e^{-\rho s}(dL_s + 1_{\{\Lambda_s = \lambda + \Delta\lambda\}} X_s B ds) \,|\, \mathcal{F}_t^N\right]$$

$$= \int_0^{t \wedge \tau} e^{-\rho s}(dL_s + 1_{\{\Lambda_s = \lambda + \Delta\lambda\}} X_s B ds) + e^{-\rho t} W_t(\Gamma, \Lambda).$$

(5)

Since $U_t(\Gamma, \Lambda)$ is an expectation conditional on $\mathcal{F}_t^N$, the process $U(\Gamma, \Lambda) = \{U_t(\Gamma, \Lambda)\}_{t \geq 0}$ is an $\mathcal{F}^N$–martingale under the probability measure $\mathbb{P}^{\Lambda}$. Its last element is $U_{\tau}(\Gamma, \Lambda)$, which is integrable by (3).

Relying on this martingale property, we now offer an alternative representation of $U(\Gamma, \Lambda)$. Consider the process $M^{\Lambda} = \{M_t^{\Lambda}\}_{t \geq 0}$ defined by

$$M_t^{\Lambda} = N_t - \int_0^t \Lambda_s \, ds$$

(6)

for all $t \geq 0$. Equation (6) is best understood when Λ is a constant process. In that case, M_t^{Λ} is simply the number of accidents up to and including date t, minus its expectation. More generally, a standard result from the theory of point processes implies that M^{Λ} is an $\mathcal{F}^N$–martingale under $\mathbb{P}^{\Lambda}$, see the Appendix. Changes in the effort process Λ induce changes in the distribution of accidents, which essentially amount to Girsanov transformations of the accident process N. The martingale representation theorem for point processes implies the following lemma.

Lemma 1. *The martingale $U(\Gamma, \Lambda)$ satisfies*

$$U_t(\Gamma, \Lambda) = U_0(\Gamma, \Lambda) - \int_0^{t \wedge \tau} e^{-\rho s} H_s(\Gamma, \Lambda) \, dM_s^{\Lambda}$$

(7)

for all $t \geq 0$, $\mathbb{P}^{\Lambda}$–almost surely, for some $\mathcal{F}^N$–predictable process $H(\Gamma, \Lambda) = \{H_t(\Gamma, \Lambda)\}_{t \geq 0}$.

Along with (6), (7) implies that the lifetime expected utility of the manager evolves in response to the jumps of the accident process N. At any date t, the change in $U_t(\Gamma, \Lambda)$ is equal to the product between a predictable function of the past, namely $e^{-\rho t} H_t(\Gamma, \Lambda)$, and a term $-dM_t^{\Lambda}$ reflecting the events occurring at date t. This term is equal to the difference between the instantaneous probability $\Lambda_t dt$ of an accident, and the instantaneous change dN_t in the total number of accidents, which is equal to 0 or 1. Thus $H_t(\Gamma, \Lambda)$ can be interpreted as the sensitivity of the manager's utility to the occurrence of accidents. Equations (5) and (7) imply that the continuation utility of the manager evolves as

$$dW_t(\Gamma, \Lambda) = [\rho W_t(\Gamma, \Lambda) - 1_{\{\Lambda_s = \lambda + \Delta\lambda\}} X_t B] dt + H_t(\Gamma, \Lambda)(\Lambda_t dt - dN_t) - dL_t$$

(8)

for all $t \in [0, \tau)$. Thus the higher is $H(\Gamma, \Lambda)$, the more sensitive to accidents is the continuation utility of the manager. Building on this analysis, and denoting $b = B/\Delta\lambda$, we obtain the following result.

Proposition 1. *A necessary and sufficient condition for the effort process Λ to be incentive compatible given the contract $\Gamma = (X, L, \tau)$ is that*

$$\Lambda_t = \lambda \ \text{ if and only if } \ H_t(\Gamma, \Lambda) \geq X_t b \tag{9}$$

for all $t \in [0, \tau)$, $\mathbb{P}^\Lambda$–almost surely.

It follows from (8) that, if an accident occurs at some date $t \in [0, \tau)$, the manager's continuation utility is instantaneously lowered by an amount $H_t(\Gamma, \Lambda)$.[15] Proposition 1 states that, to incite the manager to choose a high level of risk prevention, this reduction in the manager's continuation utility must be at least as large as $X_t b$. This is because $X_t b$ reflects the attractiveness of the private benefits obtained by the agent when shirking.

Our characterization of incentive compatibility in a model with jumps in the output process parallels that obtained in models where output is driven by a diffusion process. In such Brownian models, the continuation utility of the agent must display a minimal level of volatility in order to maintain incentive compatibility. However, there is no role for downsizing and the optimal contract only relies on liquidation and transfers (see for instance Sannikov (2003), DeMarzo and Sannikov (2006), or Biais, Mariotti, Plantin and Rochet (2007)). By contrast, a distinctive feature of our Poisson formulation of uncertainty is the role of firm size X in the provision of incentives, as can be seen from (9).

Now turn to the limited liability constraint. It is convenient to introduce the notation $W_{t^-}(\Gamma, \Lambda) = \lim_{s \uparrow t} W_s(\Gamma, \Lambda)$ to denote the left-hand limit of the process $W(\Gamma, \Lambda)$ at $t > 0$. While $W_t(\Gamma, \Lambda)$ is the continuation utility of the manager at date t after observing whether an accident occurred or not, $W_{t^-}(\Gamma, \Lambda)$ is the continuation utility of the manager evaluated before such knowledge is obtained.[16] Combining the fact that the continuation utility of the manager must remain positive according to the limited liability constraint, with the fact that it must be lowered by $H_t(\Gamma, \Lambda)$ after an accident according to (8), one must have

$$W_{t^-}(\Gamma, \Lambda) \geq H_t(\Gamma, \Lambda) \tag{10}$$

for all $t \in [0, \tau)$. Condition (10) states that the manager's continuation utility must always stay large enough to absorb a variation $-H_t(\Gamma, \Lambda)dN_t$ while remaining non-negative.

3.2. Derivation of the Optimal Contract under Maximal Risk Prevention

In this subsection, we characterize the optimal contract that induces maximal risk prevention, that is $\Lambda_t = \lambda$ for all $t \in [0, \tau)$. This contract can be described with the help of two state variables: the size of the firm, which results from past downsizing decisions, and the continuation utility of the manager, which reflects future transfer decisions. We first provide a heuristic derivation of the insurance company's value function. Next, we construct a contract that generates this value for the insurance company and incites the manager to always take the high effort decision. This delivers the desired optimal contract.

---

[15]In full generality, one should also allow for jumps in the transfer process. For incentive reasons, it is however never optimal to distribute transfers to the manager in case an accident occurs. Moreover, it will turn out that the optimal transfer process is absolutely continuous, so that transfers do not come in lump-sums. To ease the exposition, we therefore rule out jumps in the transfer process in the body of the paper. The possibility of allowing for such jumps is explicitly taken into account when we establish an upper bound for the insurance company's expected discounted profit, see the proof of Proposition 3.

[16]$W_{0^-}(\Gamma, \Lambda)$ is defined by (1). Note that while the process $W(\Gamma, \Lambda)$ is $\mathcal{F}^N$–adapted, the process $W_{\cdot^-}(\Gamma, \Lambda) = \{W_{t^-}(\Gamma, \Lambda)\}_{t \geq 0}$ is $\mathcal{F}^N$–predictable.

A Heuristic Derivation. In this heuristic derivation, we proceed in three steps.[17] First, we present the dynamics of the two state variables, that is, the size of the firm X_t, and the manager's continuation utility W_{t-}, both evaluated before the realization of uncertainty at date t. Next, we discuss the dynamics of the insurance company's value function $F(X_t, W_{t-})$. Last, we describe the main features of the resulting contract.

Consider first the manager's continuation utility. It follows from (8) that, under maximal risk prevention, W_{t-} evolves as

$$dW_{t-} = (\rho W_{t-} + \lambda H_t)dt - H_t dN_t - dL_t \tag{11}$$

at any date t prior to liquidation. Now consider the evolution of the size of the firm. Since the project has a positive net present value, it is suboptimal to downsize the firm, except after an accident, to improve the incentives of the manager. Correspondingly, a size adjustment should take place at date t only if an accident occurs at this date. That is:

$$dX_t = (X_{t+} - X_t)dN_t, \tag{12}$$

where $X_{t+} = \lim_{s \downarrow t} X_s \in [0, X_t]$ stands for the size of the firm just after the date t adjustment.

We now restate the constraints facing the insurance company. First, it must incite the manager to exert a high prevention effort at date t. By Proposition 1, this requires $H_t \geq X_t b$, or equivalently, letting $h_t = H_t/X_t$,

$$h_t \geq b. \tag{13}$$

Next, some downsizing may be necessary. To see why, consider the situation at the outset of date t, when the size of the firm is X_t and the continuation utility of the manager is W_t^-. If an accident occurs at date t, the manager's continuation utility must be reduced from W_t^- to $W_t = W_t^- - X_t h_t$. At this point, the question arises whether the occurrence of this accident implies that the firm should be downsized. Since a high prevention effort is still required from the manager, Proposition 1 implies that, if a new accident occurred, W_t would have to be reduced further by at least $X_{t+} b$. This would be consistent with limited liability only if $W_t^- - X_t h_t \geq X_{t+} b$, or equivalently, letting $w_t = W_{t-}/X_t$ and $x_t = X_{t+}/X_t$,

$$\frac{w_t - h_t}{b} \geq x_t. \tag{14}$$

Hence, downsizing is necessary, that is $x_t < 1$, only when the continuation utility of the manager is relatively low, so that $(w_t - h_t)/b < 1$. The last constraint facing the insurance company is that transfers to the manager at date t should be non-negative. Assuming that transfers are absolutely continuous with respect to time and that no transfer takes place after an accident, that is $dL_t = l_t 1_{\{dN_t=0\}}dt$, this amounts to

$$l_t \geq 0. \tag{15}$$

We are now in a position to characterize the evolution of the value function $F(X_t, W_{t-})$ of the insurance company. Since it discounts the future at rate r, the expected instantaneous change in its value function must be

$$rF(X_t, W_{t-})dt.$$

---

[17]For notational convenience, we drop the arguments Γ and Λ in what follows.

This must be equal to the sum of the expected instantaneous cash-flow it receives and of the expected change in its continuation value. The former is equal to the expected net cash-flow from the firm, minus the expected transfer to the manager,

$$[X_t(\mu - \lambda C) - l_t]dt + o(dt).$$

To compute the change in the insurance company's continuation value, we use the dynamics (11) of the manager's continuation utility along with the change of variable formula for processes of bounded variations, which is the counterpart of Itô's formula for these processes. This yields the following expected change in the insurance company's continuation value:

$$(\rho W_{t-} + \lambda X_t h_t - l_t)F_W(X_t, W_{t-})dt - \lambda dt[F(W_{t-}, X_t) - F(W_{t-} - X_t h_t, X_t x_t)] + o(dt).$$

The first term arises because of the drift of W_{t-}, while the second term reflects the possibility of jumps due to accidents. Putting these terms together, we obtain that the value function of the insurance company satisfies the Hamilton–Jacobi–Bellman equation

$$rF(X_t, W_{t-}) = (\mu - \lambda C)X_t + \max\left\{-l_t + (\rho W_{t-} + \lambda X_t h_t - l_t)F_W(X_t, W_{t-})\right.$$

$$\left. - \lambda[F(W_{t-}, X_t) - F(W_{t-} - X_t h_t, X_t x_t)]\right\}, \tag{16}$$

where the maximization in (16) is over the set of controls (h_t, x_t, l_t) that satisfy constraints (13) to (15).

To get more insight into the structure of the solution, we impose further restrictions on the value function F, which will be checked to be without loss of generality in the verification theorem below. First, because of constant returns to scale, it is natural to require F to be homogenous of degree 1,

$$F(\xi, \omega) = \xi F\left(1, \frac{\omega}{\xi}\right) \equiv \xi f\left(\frac{\omega}{\xi}\right)$$

for all $(\xi, \omega) \in \mathbb{R}_{++} \times \mathbb{R}_+$. Intuitively, f maps the size-adjusted expected discounted utility of the manager into the size-adjusted expected discounted profit of the insurance company. Second, we require f to be globally concave, and linear over $[0, b]$,

$$f(w) = \frac{f(b)}{b}\,w$$

for all $w \in [0, b]$.

We can now derive several features of the optimal controls in (16). Optimizing with respect to $l_t \geq 0$ and using the homogeneity of F yields

$$f'(w_t) = F_W(X_t, W_{t-}) \geq -1, \tag{17}$$

with equality only if $l_t > 0$. Intuitively, the left-hand side of (17) is the decrease in the expected profit of the insurance company due to an increase in the manager's rent, while the right-hand side is the cost to the insurance company of an immediate transfer to the manager. It is optimal to delay transfers as long as they are more costly than rent promises, that is, as long as the inequality in (17) is strict. The concavity of f implies that this is the case when w_t is below a given threshold. The optimal contract thus satisfies the following property.

Property 1. *Transfers to the manager take place only if w_t is at or above a threshold w^m.*

Suppose that w_t is below the threshold w^m. Then, using the homogeneity of F, one can rewrite (16) as follows:

$$rf(w_t) = \mu - \lambda C + \max\left\{(\rho w_t + \lambda h_t)f'(w_t) - \lambda\left[f(w_t) - x_t f\left(\frac{w_t - h_t}{x_t}\right)\right]\right\}, \qquad (18)$$

where the maximization in (18) is over the set of controls (h_t, x_t) that satisfy (13) and (14). Since f is concave and vanishes at 0, the mapping $x_t \mapsto x_t f((w_t - h_t)/x_t)$ is non-decreasing. It is thus optimal to let x_t be as high as possible in (18), reflecting that downsizing is costly since the project is profitable. Using (14) along with the fact that $x_t \leq 1$ then leads to the second property of the optimal contract.

Property 2. *The optimal downsizing policy is given by*

$$x_t = \min\left\{\frac{w_t - h_t}{b}, 1\right\}. \qquad (19)$$

This property of the optimal contract reflects that downsizing is imposed only as the last resort, in order to maintain the consistency between the limited liability constraint and the incentive compatibility constraint. Using the linearity of f over $[0, b]$,[18] one can then substitute (19) into (18) to obtain

$$rf(w_t) = \mu - \lambda C + \max\left\{(\rho w_t + \lambda h_t)f'(w_t) - \lambda[f(w_t) - f(w_t - h_t)]\right\}. \qquad (20)$$

The concavity of f then implies that it is optimal to let h_t be as low as possible in (20), which according to (13) leads to the third property of the optimal contract.

Property 3. *The sensitivity of the manager's continuation utility to accidents is given by*

$$h_t = b. \qquad (21)$$

Because the expected discounted profit of the insurance company is a concave function of the manager's utility, it is optimal to reduce the manager's exposure to risk by letting h_t equal the minimal value b consistent with a high prevention effort at date t.

To summarize this heuristic derivation, our candidate for the insurance company's size-adjusted value function is the solution to

$$\begin{cases} f(w) = f(b)w/b & \text{if } w \in [0, b], \\\\ rf(w) = \mu - \lambda C + (\rho w + \lambda b)f'(w) - \lambda[f(w) - f(w - b)] & \text{if } w \in (b, w^m], \\\\ f(w) = w^m - w + f(w^m) & \text{if } w \in (w^m, \infty), \end{cases} \qquad (22)$$

for some transfer threshold w^m yet to be determined. The remainder of this subsection is organized as follows. We first show that (22) has a maximal solution in a space of suitably regular functions. Next, we argue that this maximal solution provides an upper bound for the insurance company's expected discounted profit when it incites the manager to maximal risk prevention and gives her at least her required expected discounted utility. Last, we show that this maximal solution can indeed be attained through an incentive compatible contract, so that it indeed coincides with the insurance company's optimal value function.

---

[18]Instead of assuming that f is linear over $[0, b]$, one could have first defined f over $[b, \infty)$ only with $f(b)/b \geq f'(b)$, and argue that when $x_t = (w_t - h_t)/b$, the term $x_t f((w_t - h_t)/x_t)$ in (18) becomes $f(b)(w_t - h_t)/b$, which we can then rewrite as $f(w_t - h_t)$ by conventionally letting f be linear over $[0, b]$.

The Maximal Solution. Three objects have to be jointly determined in problem (22). First, the slope $f(b)/b$ of the function f over the interval $[0, b)$. Second, the function f itself over $[b, w^m]$. Third, the threshold w^m above which the slope of f is equal to -1. To link these three objects, we impose that f be continuous over $\mathbb{R}_+$ and continuously differentiable over (b, ∞), which implies in particular that

$$f'(w^m) = -1$$

as long as $w^m > b$. Under this restriction, the choice of the slope of f over $[0, b)$ determines the value taken by the threshold w^m. Our goal is to show that there exists some choice of the slope of f over $[0, b)$ that makes f maximal among the solutions to (22) that are continuous over $\mathbb{R}_+$ and continuously differentiable over (b, ∞). To see this, fix some $\alpha \geq -1$ and consider the unique continuous solution ϕ_α to

$$\begin{cases} \phi_\alpha(w) = \alpha w & \text{if } w \in [0, b], \\ r\phi_\alpha(w) = \mu - \lambda C + (\rho w + \lambda b)\phi_\alpha'(w) - \lambda[\phi_\alpha(w) - \phi_\alpha(w - b)] & \text{if } w \in (b, \infty). \end{cases} \tag{23}$$

One then has the following result.

Proposition 2. *Whenever*

$$\mu - \lambda C \geq (\rho - r)b\left(2 + \frac{r}{\lambda}\right), \tag{24}$$

the following holds:

(i) $\phi_{\alpha_1} \geq \phi_{\alpha_2}$ *if and only if* $\alpha_1 \geq \alpha_2$.

(ii) *There exists a maximum value of* α, α^m, *such that* $\phi_\alpha'(w) = -1$ *has a solution.*

(iii) *The solution* w^m *to* $\phi_{\alpha_m}'(w) = -1$ *is unique, and strictly greater than* b.

(iv) *The function* ϕ_{α^m} *is concave over* $[0, w^m]$, *and strictly so over* $[b, w^m]$.

According to Lemma 2, three cases can occur. If $\alpha \in [-1, \alpha_m)$, the equation $\phi_\alpha'(w) = -1$ has at least one solution, but ϕ_α' is below -1 over some range. If $\alpha \in (\alpha_m, \infty)$, the equation $\phi_\alpha'(w) = -1$ has no solution, and ϕ_α' is always strictly above -1. Finally, if $\alpha = \alpha_m$, the equation $\phi_\alpha'(w) = -1$ has a unique solution w^m, and ϕ_{α^m}' is always greater than or equal to -1. One then define f as

$$f(w) = \begin{cases} \phi_{\alpha^m}(w) & \text{if } w \in [0, w^m], \\ w^m - w + f(w^m) & \text{if } w \in (w^m, \infty). \end{cases} \tag{25}$$

The function f defined by (25) is the maximal solution to (22) whose derivative at w^m is precisely equal to -1.

The assumption (24) is required for the value function associated with the optimal contract to be continuously differentiable over (b, ∞). We shall maintain this assumption in the remainder of the paper. Whenever (24) fails to hold, the optimal value function is piecewise linear. This somewhat degenerate case corresponds to the range of parameters studied in Proposition 5 of Biais, Mariotti, Plantin and Rochet (2004).

An Upper Bound for the Insurance Company's Profits. The second step of our argument consists in showing that the maximal solution f to (22) given by (25) provides an upper bound for the insurance company's expected profit from any incentive compatible contract that incites the manager to exert the high prevention effort in every contingency. Specifically, define $F(\xi, \omega) = \xi f(\omega/\xi)$ for all $(\xi, \omega) \in \mathbb{R}_{++} \times \mathbb{R}_{+}$ as in the heuristic derivation of the optimal contract. The following result holds.

Proposition 3. *For any contract $\Gamma = (X, L, \tau)$ that induces maximal risk prevention, that is, $\Lambda_t = \lambda$ for all $t \in [0, \tau)$, and delivers the manager an initial expected discounted utility W_{0-} given initial firm size X_0, one has*

$$F(X_0, W_{0-}) \geq \mathbb{E}^\Lambda \left[\int_0^\tau e^{-rt}[X_t(\mu dt - CdN_t) - dL_t] \right]. \tag{26}$$

That is, the insurance company's expected discounted profit at date 0 is at most $F(X_0, W_{0-})$.

In line with the heuristic derivation of Properties 1 to 3 of the optimal contract, the proof of this result relies in an essential way on the homogeneity of F and on the concavity of f. It should be noted that no restriction is made on contracts in Proposition 3, besides the fact that they induce the manager to always exert the high prevention effort. In particular, these contracts can exhibit arbitrarily complex history dependence, and can be contingent on other variables than the size of the firm and the continuation utility of the manager.

The Verification Theorem. We are now in a position to derive the optimal contract that induces maximal risk prevention. Along standard lines in optimal control theory, we provide a verification theorem. That is, we show that the upper bound for the insurance company's expected discounted profit derived in Proposition 3 can effectively be attained by an incentive compatible contract. One has the following result.

Proposition 4. *The optimal contract that induces maximal risk prevention, that is, $\Lambda_t = \lambda$ for all $t \in [0, \tau)$, and delivers the manager an initial expected discounted utility W_{0-} given initial firm size X_0, entails expected discounted profit $F(X_0, W_{0-})$ for the insurance company. The optimal contract involves two state variables, the size of the firm and the manager's continuation utility, which evolves as*

$$dW_{t-} = (\rho W_{t-} + X_t \lambda b)dt - X_t b dN_t - dL_t \tag{27}$$

for all $t \in [0, \tau)$, and $W_{t-} = 0$ for all $t > \tau$. For each $t \geq 0$, the manager's continuation utility after the realization of uncertainty at date t is $W_t = \lim_{s \downarrow t} W_{s-}$. The optimal contract $\Gamma = (X, L, \tau)$ can be described as follows:

(i) *The size of the firm is given by*

$$X_t = \sum_{n=0}^\infty \xi_n 1_{\{t \in (\tau_n, \tau_{n+1}]\}} \tag{28}$$

for all $t \in (0, \tau)$, where $\tau_0 = 0$, $\xi_0 = X_0$, and

$$\tau_{n+1} = \inf\{t > \tau_n \,|\, W_t < \xi_n b\}, \tag{29}$$

$$\xi_{n+1} = \frac{W_{\tau_{n+1}}}{b} \tag{30}$$

15

for all $n \geq 0$.

(ii) *Transfers are given by*

$$L_t = \max\{W_{0-} - X_0 w^m, 0\} + \int_0^t X_s(\rho w^m + \lambda b) 1_{\{W_s = X_s w^m\}} \, ds \tag{31}$$

for all $t \in [0, \tau)$.

(iii) *Liquidation occurs with probability zero on the equilibrium path,*

$$\tau = \inf\{t \geq 0 \,|\, W_t = 0\} = \infty, \tag{32}$$

$\mathbb{P}^\Lambda$*–almost surely.*

This result shows that the optimal contract that induces maximal risk prevention is only contingent on the size of the firm and on the continuation utility of the manager. The features of the optimal contract confirm the heuristic derivation of Properties 1 to 3. Let us examine each of these properties in turn, starting from the last one.

P3. According to (27), the sensitivity of the manager's continuation utility to accidents is equal to b in size-adjusted terms, as prescribed by Property 3.

P2. Consider next the evolution of the size of the firm, which is described in equations (28) to (30). Size adjustments take place at dates τ_1, τ_2, ..., and they successively lower the size of the firm by from ξ_0 to ξ_1, from ξ_1 to ξ_2, ... It follows from (27) to (29) that the firm is downsized at date τ_{n+1} if and only if an accident at this date lowers the manager's continuation utility by $X_{\tau_{n+1}} b$, and brings it at a level $W_{\tau_{n+1}}$ which lies itself below $X_{\tau_{n+1}} b$. Letting $w_t = W_{t-}/X_t$, and taking advantage of the fact that $W_{\tau_{n+1}} = X_{\tau_{n+1}}(w_{\tau_{n+1}} - b)$ and $X_{\tau_{n+1}} = \xi_n$, (30) then yields a downsizing factor

$$x_{\tau_{n+1}} = \frac{\xi_{n+1}}{\xi_n} = \frac{w_{\tau_{n+1}} - b}{b} < 1,$$

as prescribed by Property 2. By construction, if $W_{0-} > 0$, one has $w_t > b$ and thus $X_t > 0$ for all $t \geq 0$.

P1. Consider finally the transfer decisions, which are summarized by (31). For each $t > 0$, transfers take place at date t if and only if $W_t = X_t w^m$, and they are constructed in such a way that the manager's continuation utility stays constant at the level $X_t w^m$ until an accident occurs. Thus, in line with Property 1, transfers to the manager take place if only if her size-adjusted utility w_t before the realization of uncertainty at date t is at the threshold w^m, and no accident occurs at date t. By construction, $w_t \leq w^m$ for all $t \in (0, \tau)$. If $w_0 > w^m$, or equivalently $W_{0-} > X_0 w^m$, a lump-sum transfer $W_{0-} - X_0 w^m$ is immediately distributed to the manager, after which the above transfer policy is implemented.

It should be noted that liquidation plays virtually no role in the optimal incentive contract, as reflected by (32). Indeed, apart from exceptional circumstances, $w_t = W_{t-}/X_t$ always remains strictly greater than b. As a result of this, W_t, which is in the worst case equal to $W_{t-} - X_t b$

if an accident occurs at date t, always remains strictly positive.[19] This is in sharp contrast with the Brownian models studied by Sannikov (2003), DeMarzo and Sannikov (2006), and Biais, Mariotti, Plantin and Rochet (2007), in which the optimal contract relies crucially on liquidation and involves no downsizing. Admittedly, even in the context of our Poisson model, an alternative way to provide incentives to the manager in case of bad performance would be to allow for randomly liquidating the firm, as is customary in discrete-time models (see for instance DeMarzo and Fishman (2003), or Clementi and Hopenhayn (2006)). But in contrast with what happens in Brownian models, liquidation would then necessarily have to be both *stochastic* (as it would depend on the realization of a lottery at each potential liquidation date) and *unpredictable* (as it would take place only after an accident occurs). Modelling liquidation in this way would allow to achieve essentially the same outcome as under downsizing. This would however be both less tractable analytically, and less conducive to a realistic implementation of the optimal contract.

Initialization. Proposition 4 describes the optimal contract for a given initial firm size X_0 and a given initial promised utility W_{0-} for the manager. We now examine how these are determined at date 0. Consider for simplicity the case in which the insurance company is competitive. We then seek a pair (X_0, W_{0-}) that maximizes the utilitarian social welfare under the constraint that the insurance company breaks even on average. Letting $w_0 = W_{0-}/X_0$, the corresponding maximization problem is

$$\max X_0[f(w_0) + w_0], \tag{33}$$

under the break even and feasibility constraints:

$$X_0 f(w_0) \geq 0, \tag{34}$$

$$w_0 \geq 0, \tag{35}$$

$$X_{0-} \geq X_0, \tag{36}$$

where $X_{0-} > 0$ is the maximal initial scale of the project. Let η be the Lagrange multiplier of the break even constraint (34), and focus on the interesting case where $(1+\eta)f(w_0) + w_0 > 0$ at the optimum.[20] It immediately follows that it is optimal to start operating the project at maximum scale, $X_0 = X_{0-}$. This result hinges on the homogeneity of the insurance company's value function F. As shown in (33), this enables one to separate the determination of the firm's size from that of the manager's size-adjusted utility. Whenever f takes strictly positive values, it is optimal to start operating the project at full scale.

The initial size-adjusted rent of the manager is given by the first-order condition $f'(w_0) = -1/(1+\eta)$. Two cases may arise depending on whether the break even constraint is slack or binding at the optimum. If $f(w^m) \geq 0$, one has $\eta = 0$ and $w_0 = w^m$, and social welfare attains its unconstrained maximum in (33). If $f(w^m) < 0$, one has $\eta > 0$ and $w_0 < w^m$, and social welfare falls short of its unconstrained maximum in (33).

---

[19]Exceptions arise only with probability 0, for instance if $W_{0-} = X_0 b$ and an accident occurs at date 0, or if multiple accidents occurs at the same date.

[20]Otherwise the solution to problem (33) to (36) is $X_0 = W_{0-} = 0$ and the project is not operated.

The Long Run. Downsizing is a key component of the optimal contract. Indeed, it allows to punish the manager in case of poor cumulative performance, which is not achievable through transfers because the manager's utility is bounded below. Over time, the size of the firm decreases as downsizing activity takes place. The following result describes the asymptotic impact of downsizing.

Proposition 5. *In the long run, the size of the firm and the continuation utility of the manager tend to 0,*

$$\lim_{t\to\infty} X_t = \lim_{t\to\infty} W_{t^-} = 0, \tag{37}$$

$\mathbb{P}^\Lambda$–*almost surely.*

To maintain incentive compatibility, downsizing must take place in case of accident when the manager's size-adjusted utility is close to its lower bound b. Because the stochastic process describing the manager's size-adjusted utility is ergodic, with probability 1 this situation will prevail over an infinite collection of time intervals. As a result of this, the size of the firm and the continuation utility of the manager must eventually tend to 0.

The logic is different from that of Thomas and Worrall's (1990) classic immiseration result. In the principal-agent model they consider, the period utility function of the agent is concave and unbounded below. Consequently, providing incentives is cheaper, the lower the agent's promised utility. This reflects the fact that the cost of obtaining a given spread in the agent's continuation utility is lower. The principal thus has an incentive to let the agent's utility drift to $-\infty$, as this makes incentive compatibility cheaper to achieve on average. Instead, in our model, the cost of incentive compatibility is high when the manager's size-adjusted utility is close to its lower bound b. This is because limited liability then makes it more difficult to induce a large variability in the manager's continuation utility. Therefore, downsizing becomes necessary to induce the manager to exert effort.

3.3. Optimality of Maximal Risk Prevention

So far, we have focused on the optimal contract with maximal risk prevention. We now briefly investigate under which circumstances this high level of effort is indeed optimal. Note that the contract derived in Proposition 4 depends on B and $\Delta\lambda$ only through the ratio $b = B/\Delta\lambda$. Thus one has one degree of freedom in the parameters of the model, as one can move B and $\Delta\lambda$ while keeping b constant, leaving the optimal contract under maximal risk prevention unaffected. Intuition suggests that when $\Delta\lambda$ gets large, it is indeed optimal to prevent accidents as much as possible by requesting maximal risk prevention. To see why, observe first that if a contract were to call the manager to exert the low prevention effort at some date t, her continuation payoff at this date would no longer need to be affected by the occurrence of an accident. It would thus be optimal to let $H_t(\Gamma, \Lambda) = 0$ in (8). To determine whether requesting maximal risk prevention is optimal, we compare the value of the insurance company under high prevention effort to its counterpart under low prevention effort. The former is greater than the latter if

$$rf(w) \geq \mu - (\lambda + \Delta\lambda)C + (\rho w - B)f'_+(w). \tag{38}$$

The second term on the right-hand side of (38) is the expected social cost of accident under low effort. The third term reflects that the drift of $W(\Gamma, \Lambda)$ under low effort is equal to

18

$\rho W(\Gamma, \Lambda) - B$. Maximal risk prevention is optimal if (38) holds for all $w \geq b$. In Lemma C.7 in the Appendix, we show that

$$rf(w) \geq \mu - \lambda C + (\rho w + \lambda b)f'_+(w) - \lambda[f(w) - f(w - b)]$$

for all $w \geq b$. Hence a sufficient condition for (38) to hold is that

$$\Delta\lambda[C + bf'_+(w)] \geq \lambda[f(w) - f(w - b) - bf'_+(w)] \qquad (39)$$

for all $w \geq b$, where we have used the fact that $B = b\Delta\lambda$. The right-hand side of (39) is non-negative by concavity of f, and it is bounded as f is linear over (w^m, ∞). Consider next the left-hand side of (39). By assumption, maximal risk prevention is socially optimal in the first-best, so that $C > B/\Delta\lambda = b$. Since $f'_+ \geq -1$, this implies that the mapping $C + bf'_+$ is positive and bounded away from 0. Since f depends on B and $\Delta\lambda$ only through their ratio b, it follows that (39) is satisfied when $\Delta\lambda$ is high enough, while B is adjusted so as to keep b constant.[21] Thus, all other things being equal, it is optimal to request maximal risk prevention if the moral hazard parameters B and $\Delta\lambda$ are high enough.

4. Implementation

In this section, we show how realistic insurance and financial instruments can be used to implement the abstract optimal contract derived in Section 3. We show that these instruments are budget balanced, that is, that the net cash revenue generated by the firm is equal to the use of funds at each point in time. This implementation gives rise to the same production and distribution decisions as in the optimal contract, on and off the equilibrium path, which implies that it is incentive compatible. Finally, we derive several empirical implications of our analysis.

4.1. Insurance and Financial Contracts

There are three aspects to the relation between the insurance company and the manager. First, the insurance company is liable in case of damages. Next, incentives must be provided to the manager so that she always exerts the high prevention effort. Last, as the manager is more impatient than the insurance company, she would like borrow from it to finance consumption. While the first two features revolve around insurance issues, the third one is about finance. Correspondingly, the implementation we propose combines insurance and financial aspects.

Cash Reserves. A realistic feature of our implementation is that the firm must hold cash reserves. This parallels the corporate finance model of Biais, Mariotti, Plantin and Rochet (2007). These reserves are deposited on a bank account and earn interest at rate r. At any point in time, changes in this account's balance reflect the operating cash-flows of the firm, the transfers to the insurance company and to the manager, and the earned interest income. Cash reserves will thus be affected by the performance of the firm and the occurrence of accidents. In our implementation, the accumulated cash reserves held by the firm will be set equal to W_{t-} at the outset of any date t, and to W_t after the realization of uncertainty at

---

[21]Note that in the limit first-best case, the expected discounted profit of the insurance company is linear in the manager's expected discounted utility, with a slope equal to -1. Condition (39) then reduces to $C > b$.

date t.[22] It is convenient to interpret the ratio of cash reserves to the size of the firm as the liquidity ratio of the firm. The manager's compensation schedule as well as the downsizing policy of the firm are directly contingent on this measure of the firm's liquidity.

Insurance Contract. In line with clauses observed in practice, the insurance contract on which our implementation relies involves both a deductible and a bonus-penalty system. When an accident occurs at date t, the insurance company is liable for the entire size of the damage, $X_t C$, minus a deductible, paid by the firm out of its cash reserves. To provide appropriate incentives to the manager, the deductible is set equal to $X_t b$. In each period, the firm pays an insurance premium to the insurance company, which combines an actuarially fair component with an incentive component. Since accidents occur with an intensity λ under maximal risk prevention, the actuarially fair premium is given by

$$\lambda X_t (C - b) dt$$

during the infinitesimal time interval $[t, t + dt)$. During the same time interval, the incentive component of the insurance premium is given by

$$-(\rho - r) W_{t^-} dt.$$

This component works as a bonus-penalty system in that it adjusts the premium paid by the firm according to its claims frequency. As long as no accident occurs, W_{t^-} increases up to the threshold $X_t w^m$. This lowers the insurance premium, corresponding to a bonus. By contrast, when an accident occurs, W_{t^-} is lowered by $X_t b$. This raises the insurance premium, corresponding to a penalty.

Corporate Bond. To fund its initial cash reserves W_{0^-}, the firm issues a corporate bond at date 0, which is acquired by the insurance company. This bond first pays a constant coupon $X_0(\mu - \lambda b)$ per unit of time. If the firm subsequently incurs a large number of accidents, it must be downsized, which can be interpreted as a form of financial distress. When such events happen, the coupon is also downsized. Hence, in general, the coupon on the bond is given by

$$X_t(\mu - \lambda C) dt$$

during the infinitesimal time interval $[t, t + dt)$. Thus, while the size-adjusted coupon is constant and equal to $\mu - \lambda C$ per unit of time, the bond is exposed to the risk of downsizing. Thus credit risk arises endogenously in our model as a result of accidents and moral hazard. This risk is reflected in the bond price, as we will see below.

---

[22]This differs from Biais, Mariotti, Plantin and Rochet (2007), in which cash reserves are a multiple of the manager's continuation utility, reflecting the magnitude of the moral hazard problem. The key difference lies in the fact that they insist that financiers hold securities, defined as claims with limited liability, while we make the insurance company liable for the social costs generated by the firm's activity. The implementation in Biais, Mariotti, Plantin and Rochet (2007) can easily be transposed in our context. The firm would need to hold cash reserves $WC/b > W$ and would use these cash reserves to cover social costs, and to pay dividends on stocks and coupons on bonds. To ensure that coupons stay non-negative, one would need to have $\mu - \lambda C \geq w^m C/b$, which requires that C/b be close enough to 1. In this alternative implementation, the firm is liable for social costs on its cash reserves, which is why the latter have to be larger than those we need in our implementation. Thus there is no incompatibility between liability insurance and liability for harm in our model, as they only represent alternative ways to implement the optimal contract. Still, liability for harm requires higher cash reserves, and additional restrictions on the parameters of the model.

Managerial Compensation. If a sufficiently long period of time elapses without accidents occurring, the manager is compensated with cash transfers. The latter take place after the realization of uncertainty at date t if the liquidity ratio W_t/X_t of the firm is equal to the contractually specified threshold w^m. This requires in particular that no accident occurred in period t. Transfers to the manager are then drawn from the cash reserves of the firm so as to maintain these cash reserves constant. Whenever $W_t/X_t = w^m$, the transfers to the manager are thus given by

$$X_t(\rho w^m + \lambda b)dt$$

during the infinitesimal time interval $[t, t+dt)$. As long as no accident occurs, cash reserves then stay constant at the level $W_t = X_t w^m$. As soon as an accident occurs, the firm must pay the deductible $X_t b$, which reduces its cash reserves and its liquidity ratio. As a result of this, the firm reverts to the regime in which the manager receives no immediate cash compensation.

Downsizing Covenant. The bond and the insurance contract include a covenant. The latter stipulates that, if an accident at date t brings the liquidity ratio W_t/X_t of the firm below b, the firm is immediately downsized by a factor

$$x_t = \frac{W_t}{bX_t}.$$

This lowers the size of the firm to $X_{t+} = x_t X_t = W_t/b$. Downsizing does not alter the level W_t of the cash reserves, but it brings the liquidity ratio back to

$$\frac{W_t}{X_{t+}} = b.$$

Since the firm thereafter operates on a smaller scale, the size of the damage in case an other accident occurs is also reduced. Correspondingly, the deductible is lowered to

$$X_{t+}b = W_t.$$

The intuition is that, immediately after being downsized, the firm has just enough cash reserves to pay the deductible in case of an other accident, without violating the limited liability constraint.

4.2. Budget Balance

At any date t, the cash-flow statement of the firm is

Cash inflows	*Cash outflows*
Operating cash-flow $X_t\mu dt$	Coupon $X_t(\mu - \lambda C)dt$
Interest income $rW_{t^-}dt$	Insurance premium $\lambda X_t(C-b)dt - (\rho - r)W_{t^-}dt$
	Deductible $X_t b\, dN_t$
	Cash hoarding or wages

TABLE 1. Cash-flow statement of the firm.

On the left-hand side if Table 1 are the cash-flows generated by the firm, which consist of operating cash-flows and interest earned on cash reserves. The different uses of these cash-flows are displayed on the right-hand side of Table 1. While the coupon and the insurance

premium are continuously paid to the insurance company, the deductible is paid only in case an accident occurs. The last item can be interpreted as follows. (i) When $W_{t^-} < w^m$ and no accident occurs, it accounts for the cash hoarded by the firm and added to its cash reserves, $(\rho W_{t^-} + X_t \lambda b)dt$. (ii) When an accident occurs, it accounts for the change in cash reserves due to the payment of the deductible, $-X_t b$. (iii) When $W_{t^-} = X_t w^m$ and no accident occurs, it accounts for the wages paid to the manager, $X_t(\rho w^m + \lambda b)dt$.

Similarly, the cash-flow statement of the insurance company is

Cash inflows	Cash outflows
Coupon $X_t(\mu - \lambda C)dt$	Insurance liability $X_t(C - b)dN_t$
Insurance premium $\lambda X_t(C - b)dt - (\rho - r)W_{t^-}dt$	Profits

TABLE 2. Cash-flow statement of the insurance company.

In case an accident occurs, the insurance company must cover the social cost, net of the deductible. Its profits are therefore $-X_t(C - b)$. When no accident occurs, the profit of the insurance company amounts to $[X_t(\mu - \lambda b) - (\rho - r)W_{t^-}]dt$.

At date 0, the insurance company receives the bonds and commits to the insurance contract. It also transfers an initial amount of cash W_{0^-} to the firm. The latter hoards this as cash reserves. Throughout its history, the firm will use accumulated cash reserves to pay coupons, insurance premia, deductibles and transfers to the manager. Thus the present value of the insurance company's profits is[23]

$$
\mathbb{E}^\Lambda \left[\int_0^\infty e^{-rt} \{ [X_t(\mu - \lambda b) - (\rho - r)W_{t^-}]dt - X_t(C - b)dN_t \} \right]
$$

$$
= \mathbb{E}^\Lambda \left[\int_0^\infty e^{-rt}[X_t(\mu dt - CdN_t) - dL_t] \right] + \mathbb{E}^\Lambda \left[\int_0^\infty e^{-rt}[dL_t - (\rho - r)W_{t^-}dt] \right]
$$

$$
= F(X_0, W_{0^-}) - \mathbb{E}^\Lambda \left[\int_0^\infty e^{-rt}d(e^{-rt}W_{t^-}) \right]
$$

$$
= F(X_0, W_{0^-}) + W_{0^-},
$$

This identity states that the rent of the insurance company, $F(X_0, W_{0^-})$, is equal to the present value of its profits, minus the initial payment it makes to the firm.

The value of the bonds received by the insurance company exceeds the initial amount of cash it pays to the firm. The difference is equal to the sum of two terms. The first one is equal to the insurance company's initial rent. The second term reflects the payments that the insurance company receives and makes as a result of the insurance contract. By construction, the actuarially fair component of the insurance premium is on average equal to the net liabilities of the insurance company. By contrast, the incentive component of the insurance premium involves an expected discounted cost to the insurance company of

$$
\mathbb{E}^\Lambda \left[\int_0^\infty e^{-rt}(\rho - r)W_{t^-} dt \right].
$$

---

[23]The first equality follows from the fact that the process M^Λ is an $\mathcal{F}^N$–martingale under $\mathbb{P}^\Lambda$, the second from Proposition 4 along with the change of variable formula for processes of bounded variations and the dynamics of the manager's continuation utility, and the third from the fact that $\lim_{t\to\infty} e^{-rt}W_{t^-} = 0$.

Remark. In the implementation described above, the insurance company purchases the bonds issued by the firm. Alternatively, one could consider the case where the firm is dealing separately with the insurance company and a fringe of competitive, risk-neutral investors, with discount rate r. The firm would sell the bonds to these investors. The proceeds would be used to hoard cash reserves W_{0-} and pay a commitment fee to the insurance company. Note that the present value of the fees paid to the insurance company is lower than the present value of its liabilities. If the insurance company does not hold the bond, it thus makes negative expected discounted profits after date 0. The commitment fee initially paid by the firm to the insurance company is then a compensation for this cost.

4.3. Incentive Compatibility

We now verify that this implementation gives rise to the same decisions as in the optimal contract. First, the dynamics of the cash reserves and of the liquidity ratio resulting from the implementation exactly mirror those of the promised continuation utility and the size-adjusted utility of the manager in the optimal contract. Next, the downsizing covenant ensures that downsizing decisions are the same in the implementation and in the optimal contract. Thus, the real decisions arising in the implementation in response to the evolution of the liquidity ratio exactly parallel those requested in the optimal contract. Finally, the compensation package proposed in the implementation leads to the same transfers to the manager as in the optimal contract. As a result of this, the insurance and financial contracts we described are incentive compatible, and they implement the optimal allocation.

4.4. Empirical Implications

Moral Hazard, Deductible and Insurance Premia. The parameter b is large and moral hazard is severe when risk prevention involves very costly efforts which outsiders cannot observe. This is likely to be the case for technologically complex industrial processes involving hazardous substances, such as in the chemical or the nuclear industry. The parameter b is also likely to be large for projects involving a sequence of critical steps which, if not taken properly, can have dangerous consequences. Such situations typically arise in the energy sector. Our model implies that, in such cases, insurance contracts should have relatively large deductibles, and insurance premia should place relatively more weight on incentive considerations. In particular, they should increase sharply after accidents, but they also should decrease significantly after a long period without accidents. Overall, firms with greater moral hazard will thus have relatively more volatile insurance premia dynamics.

To go beyond these qualitative predictions, note that the different parameters in the model are related to several observable variables. First, observing the deductible allows one to estimate the moral hazard parameter b. Next, the intensity λ of the accident process can be estimated by observing the rate at which accidents occur. Finally, combining these estimations with the observation of the cash reserves' evolution, one can estimate the discount rate ρ. This offers a first opportunity to assess the fit of the model, by testing whether ρ is greater than the risk-free rate r. Furthermore, the evolution of the insurance premium during periods without accidents offers additional information about the key parameters of the model. To see this, recall that the actuarial component $p_t^a = \lambda X_t(C - b)$ of the insurance premium is constant as long as no downsizing takes place, and thus in particular as long as no accident occurs. Thus only the incentive component $p_t^i = -(\rho - r)W_{t-}$ of the insurance

premium varies during periods without accident, according to

$$dp_t^i = [\rho p_t^i - (\rho - r)X_t b]dt.$$

Combining this empirical restriction with those used to identify b, λ, and ρ, one obtains a set of overidentifying conditions. Hence one could use the generalized method of moments to estimate the key parameters and test the model.

Pricing Credit Risk. The implementation of the optimal contract involves an infinite maturity consol bond paying a stream of coupons $\{(\mu - \lambda C)X_t\}_{t \geq 0}$. At any date t, the market value of this bond depends on the current size X_t of the firm and on its current liquidity ratio w_t, which are the two state variables governing its future evolution. Because of risk-neutrality and homogeneity, the price of this bond at time t can be written as

$$\mathbb{E}_t\left[\int_t^\infty e^{-r(s-t)}(\mu - \lambda C)X_s \, ds\right] = X_t d(w_t)$$

where the function d solves:

$$\begin{cases} d(w) = d(b)w/b & \text{if } w \in [0, b], \\ rd(w) = \mu - \lambda C + (\rho w + \lambda b)d'(w) - \lambda[d(w) - d(w - b)] & \text{if } w \in (b, w^m), \\ rd(w^m) = \mu - \lambda C - \lambda[d(w^m) - d(w^m - b)]. \end{cases} \tag{40}$$

One can interpret $d(w)$ as the size-adjusted market value of the bond, given a liquidity ratio w. The last condition in (40) reflects that when the liquidity ratio reaches its maximum level w^m, it stays there until an accident occurs. Note that the function d satisfies the same functional equation as f, see (22), the difference between the two lying in the boundary condition at w^m, which is $f'(w^m) = -1$ for f, and $d'(w^m) = 0$ for d as can be seen from (40). Using this observation, it is straightforward to derive the following corollary along the lines of the proof of Proposition 2.

Corollary 1. *The function d is strictly increasing and strictly concave over $[0, w^m]$.*

Using the expression for $d(w^m)$ in (40), a direct implication of this result is that

$$d(w^m) < \frac{\mu - \lambda C}{r}. \tag{41}$$

The economic intuition underlying Corollary 1 is the following. If many accidents occur, the firm is downsized and repayments to bondholders are correspondingly scaled down. This is a form of partial default. Hence accident risk generates credit risk, which is reflected in the pricing of the bond. As the liquidity ratio of the firm increases, the risk of downscaling decreases and the bond price increases. Thus d is increasing. However, there is a limit to this process. Indeed, the liquidity ratio never exceeds w^m and this barrier is not absorbing. Like the risk of accident, the risk of a reduction in the liquidity ratio is thus never eliminated and the risk of an eventual downsizing always remains. Correspondingly, the maximal value of the debt is lower than its risk-free counterpart, as shown by (41). The concavity of d reflects that the bond price reacts less to changes in the liquidity ratio of the firm when its

24

accident record is low. By contrast, after a series of accidents, the liquidity ratio is low, and downsizing risk is high. As a result of this, the bond price reacts strongly to firm performance and ensuing changes in the liquidity ratio.

These statements about the value of the bond can be translated in terms of yield spreads. The yield $y(w_t)$ on the consol bond is defined by

$$d(w_t) = \int_t^\infty e^{-y(w_t)(s-t)}(\mu - \lambda C)\, ds.$$

Hence $y(w_t)$ is the ratio of the size-adjusted coupon to the size-adjusted bond price:

$$y(w_t) = \frac{\mu - \lambda C}{d(w_t)}.$$

It follows from Corollary 1 that the credit risk yield spread $y(w_t) - r$ is a decreasing and convex function of the liquidity ratio w_t of the firm. From (41), this spread remains strictly positive, even when the liquidity ratio of the firm reaches its maximal value w^m:

$$y(w^m) = \frac{\mu - \lambda C}{d(w^m)} > r.$$

Instead of considering a consol bond paying a perpetual stream of appropriately downscaled coupons, one can analyze the value of the corresponding stream of zero-coupon bonds. To do this, consider the zero-coupon bond paying one dollar at date T for each unit of operation of the firm at that date. Because of risk-neutrality and homogeneity, the price of this bond at time $t < T$ can be written as

$$\mathbb{E}_t\left[e^{-r(T-t)}X_T\right] = X_t d_T(w_t),$$

while the corresponding yield $y_T(w_t)$ is defined by

$$d_T(w_t) = e^{-y_T(w_t)(T-t)}.$$

It follows from these two expressions that the credit yield spread $y_T(w_t) - r$ on the zero-coupon bond is strictly positive, reflecting that downsizing takes place with positive probability between dates t and T. Similarly to the consol bond, this spread decreases with the firm's liquidity ratio w_t, but remains strictly positive even when the firm's liquidity ratio reaches its maximal level w^m.

Now consider what happens when the maturity $T - t$ of the zero-coupon bond tends to 0. Two cases then arises. If $w_t > 2b$, then even if an accident occurs the liquidity ratio remains above the downsizing boundary b. As a result of this, there is no credit risk and the credit yield spread is 0 at 0 maturity. By contrast, if $b < w_t \leq 2b$ then, with probability λdt an accident occurs in the infinitesimal time interval $[t, t + dt)$ and the firm is downsized by a factor $w_t/b - 1$. A first-order approximation of the bond price and yield formulas implies

$$[1 - y_{t+dt}(w_t)dt] = (1 - rdt)\left[1 - \lambda dt + \lambda dt\left(\frac{w_t}{b} - 1\right)\right] + o(dt).$$

Hence, when $b < w_t \leq 2b$ the credit yield spread for 0 maturity bonds is

$$y_t(w_t) - r = \lambda\left(2 - \frac{w_t}{b}\right).$$

25

This spread is maximum in the neighborhood of the downsizing boundary b, at which it is equal to the accident intensity λ.[24]

To summarize, our model generates several empirical implications for credit risk in a context where contracts, and in particular the reliance on bonds and their covenants are endogenous. First, bond prices increase and credit risk yield spreads decrease with the liquidity ratio of the firm. Next, credit yield spreads on consol bonds or on positive maturity zero-coupon bonds remain strictly positive even when the firm is very liquid. Last, credit yield spreads on zero-coupon bonds also remain strictly positive when the maturity of the bond goes to 0 if the liquidity ratio of the firm is below a threshold. Above that threshold, the spread goes to 0 with the maturity of the bond.

5. Conclusion

We consider a dynamic setting where managers must exert costly unobservable effort to reduce industrial accident risk. We study how to cope with the moral hazard problem arising in this environment. The occurrence of accidents is modelled as a Poisson process. We use martingale techniques to characterize the optimal contract. We show that it can be implemented with realistic contractual instruments: compulsory insurance with a deductible and a bonus-penalty system, managerial compensation after good performance, risky bonds, and corporate downsizing when accidents are too frequent.

Our model generates several implications for endogenous optimal insurance and financial contracts. Firms with greater moral hazard problems have greater deductible and more volatile insurance premia. Credit risk decreases with the liquidity ratio of the firm, but credit default spreads remain positive even when the firm is very liquid. When the liquidity ratio is below a given threshold, zero coupon credit spreads remain bounded away from zero as maturities go to zero. Because it is explicitly dynamic, our model also provides an appropriate framework for empirical analyses of the joint time series of accidents, insurance premia and corporate bond prices.

Our analysis also delivers several policy implications. The insurance company and the firm should be liable for damages, and should not be allowed to escape liability by passing it to small uninsured subcontractors. Managerial compensation should clearly be negatively linked to accidents. Corporate downsizing should occur if the frequency of accidents is too high. Deviations from such guidelines would lead to socially irresponsible corporate behavior.

---

[24]As shown by Zhou (2001), positive credit yield spreads at 0 maturity are a natural outcome of a model with Poisson risk. However, while jumps in the firm's value are simply postulated in Zhou (2001), they are an endogenous feature of the optimal contract in our setup. A related difference is that we endogenize the financial distress threshold. Duffie and Lando (2001) provide a structural model of credit risk that generates strictly positive credit yield spreads at 0 maturity. As in Duffie and Lando (2001), default occurs in our framework at some intensity. While this arises in their model because of incomplete accounting information, this results in our model from imperfect effort observation and Poisson risk.

In this Appendix, we provide a precise description of the stochastic environment. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space over which is defined a Poisson process $N = \{N_t\}_{t \geq 0}$ of intensity λ. We denote by $\mathcal{F}^N = \{\mathcal{F}_t^N\}_{t \geq 0}$ the filtration generated by N, suitably augmented by the $\mathbb{P}$–null sets. This filtration satisfies the usual conditions (Dellacherie and Meyer (1978, Chapter IV, Definition 48)). The process $M = \{M_t\}_{t \geq 0}$ defined by

$$M_t = N_t - \lambda t$$

for all $t \geq 0$, is an $\mathcal{F}^N$–martingale under $\mathbb{P}$. As in the text, let $\Lambda = \{\Lambda_t\}_{t \geq 0}$ be an $\mathcal{F}^N$–predictable process with values in $\{\lambda, \lambda + \Delta\lambda\}$, and denote by $Z^\Lambda = \{Z_t^\Lambda\}_{t \geq 0}$ the unique solution to the stochastic differential equation

$$dZ_t^\Lambda = Z_{t-}^\Lambda \left(\frac{\Lambda_t}{\lambda} - 1 \right) dM_t$$

for all $t \geq 0$, where $Z_{0-}^\Lambda = 0$. By the exponential formula for Stieltjes–Lebesgue calculus (Brémaud (1981, Appendix A4, Theorem T4)), one has

$$Z_t^\Lambda = \left(\prod_{n=1}^\infty \frac{\Lambda_{T_n}}{\lambda} 1_{\{T_n \leq t\}} \right) \exp \left(\int_0^t (\lambda - \Lambda_s) \, ds \right)$$

for all $t \geq 0$, where $(T_n)_{n=1}^\infty$ is the sequence of dates at which the process N jumps. From Brémaud (1981, Chapter VI, Theorem T2), Z^Λ is a non-negative $\mathcal{F}^N$–local martingale under $\mathbb{P}$. Moreover $\mathbb{E}[Z_t^\Lambda] = 1$ for all $t \geq 0$. A standard extension argument implies that there exists a unique probability measure $\mathbb{P}^\Lambda$ over $(\Omega, \mathcal{F})$ that is defined by the Radon-Nikodym derivatives

$$\frac{d\mathbb{P}^\Lambda}{d\mathbb{P}}_{|\mathcal{F}_t^N} = Z_t^\Lambda$$

for all $t \geq 0$. It then follows from Brémaud (1981, Chapter VI, Theorem T3) that the process M^Λ defined by (6) is an $\mathcal{F}^N$–martingale under $\mathbb{P}^\Lambda$.

APPENDIX B: THE INCENTIVE COMPATIBILITY CONSTRAINT

Proof of Lemma 1. The process $U(\Gamma, \Lambda)$ defined by (5) is a non-negative $\mathcal{F}^N$–martingale under $\mathbb{P}^\Lambda$, of the form $U_t(\Gamma, \Lambda) = \mathbb{E}^\Lambda[U_\tau(\Gamma, \Lambda) \,|\, \mathcal{F}_t^N]$ for all $t \geq 0$, where $U_\tau(\Gamma, \Lambda)$ is integrable by (3). Since the filtration $\mathcal{F}^N$ satisfies the usual conditions, $U(\Gamma, \Lambda)$ admits a right-continuous modification with left-hand limits (Dellacherie and Meyer (1982, Chapter VI, Theorem 4)). The predictable representation (7) then follows directly from Brémaud (1981, Chapter III, Theorems T9 and T17). ∎

Proof of Proposition 1. Following Sannikov (2003, Lemma 2), consider the manager's lifetime expected utility, evaluated conditionally upon the information available at some date t, when she acts according to $\Lambda' = \{\Lambda_t'\}_{t \geq 0}$ until date t and then reverts to $\Lambda = \{\Lambda_t\}_{t \geq 0}$:

$$U_t' = \int_0^{t \wedge \tau} e^{-\rho s} (dL_s + 1_{\{\Lambda_s' = \lambda + \Delta\lambda\}} X_s B ds) + e^{-\rho t} W_t(\Gamma, \Lambda). \tag{B.1}$$

First, we show that if $U' = \{U_t'\}_{t \geq 0}$ is an $\mathcal{F}^N$–submartingale under $\mathbb{P}^{\Lambda'}$ that is not a martingale, then Λ is suboptimal for the manager. Indeed, in that case there exists some $t \geq 0$ such that

$$U_{0-}(\Gamma, \Lambda) = U_{0-}' < \mathbb{E}^{\Lambda'}[U_t'],$$

so that by (B.1) the manager is strictly better off acting according to Λ' until date t and then reverting to Λ. The claim follows.

Next, we show that if U' is a $\mathcal{F}^N$–supermartingale under $\mathbb{P}^{\Lambda'}$, then Λ is at least as good as Λ' for the manager. From (5) and (B.1) along with the fact that $U(\Gamma, \Lambda')$ is a $\mathbb{P}^{\Lambda'}$–martingale, one has

$$\mathbb{E}^{\Lambda'}[U'_t] = U_{0-}(\Gamma, \Lambda') + \mathbb{E}^{\Lambda'}\left[1_{\{t<\tau\}} \int_t^\tau e^{-\rho s}(1_{\{\Lambda_s = \lambda + \Delta\lambda\}} - 1_{\{\Lambda'_s = \lambda + \Delta\lambda\}})X_s B\, ds\right],$$

so that the mapping $t \mapsto \mathbb{E}^{\Lambda'}[U'_t]$ is right-continuous. Since the filtration $\mathcal{F}^N$ satisfies the usual conditions, U' admits a right-continuous modification with left-hand limits (Dellacherie and Meyer (1982, Chapter VI, Theorem 4)). Moreover, since U' is non-negative, it admits 0 as a last element. Hence, by the optional sampling theorem (Dellacherie and Meyer (1982, Chapter VI, Theorem 10)),

$$U'_{0-} \geq \mathbb{E}^{\Lambda'}[U'_\tau] = U_{0-}(\Gamma, \Lambda'),$$

from which the claim follows as $U'_{0-} = U_{0-}(\Gamma, \Lambda)$ by (B.1).

Now, for each $t \geq 0$,

$$U'_t = U_t(\Gamma, \Lambda) + \int_0^{t \wedge \tau} e^{-\rho s}(1_{\{\Lambda'_s = \lambda + \Delta\lambda\}} - 1_{\{\Lambda_s = \lambda + \Delta\lambda\}})X_s B\, ds$$

$$= U_0(\Gamma, \Lambda) - \int_0^{t \wedge \tau} e^{-\rho s} H_s(\Gamma, \Lambda)\, dM_s^\Lambda + \int_0^{t \wedge \tau} e^{-\rho s}(1_{\{\Lambda'_s = \lambda + \Delta\lambda\}} - 1_{\{\Lambda_s = \lambda + \Delta\lambda\}})X_s B\, ds$$

$$= U_0(\Gamma, \Lambda) - \int_0^{t \wedge \tau} e^{-\rho s} H_s(\Gamma, \Lambda)\, dM_s^{\Lambda'} - \int_0^{t \wedge \tau} e^{-\rho s} H_s(\Gamma, \Lambda)(\Lambda'_s - \Lambda_s)\, ds$$

$$+ \int_0^{t \wedge \tau} e^{-\rho s}(1_{\{\Lambda'_s = \lambda + \Delta\lambda\}} - 1_{\{\Lambda_s = \lambda + \Delta\lambda\}})X_s B\, ds$$

$$= U_0(\Gamma, \Lambda) - \int_0^{t \wedge \tau} e^{-\rho s} H_s(\Gamma, \Lambda)\, dM_s^{\Lambda'}$$

$$+ \int_0^{t \wedge \tau} e^{-\rho s} \Delta\lambda(1_{\{\Lambda'_s = \lambda + \Delta\lambda\}} - 1_{\{\Lambda_s = \lambda + \Delta\lambda\}})[X_s b - H_s(\Gamma, \Lambda)]\, ds,$$

where the first equality follows from (5) and (B.1), the second from (7), the third from (6), and the fourth from a straightforward computation. Since $H(\Gamma, \Lambda)$ is $\mathcal{F}^N$–predictable and $M^{\Lambda'}$ is an $\mathcal{F}^N$–martingale under $\mathbb{P}^{\Lambda'}$, the drift of U' has the same sign as

$$(1_{\{\Lambda'_t = \lambda + \Delta\lambda\}} - 1_{\{\Lambda_t = \lambda + \Delta\lambda\}})[X_t b - H_t(\Gamma, \Lambda)]$$

for all $t \in [0, \tau)$. If (9) holds for the effort process Λ, then this drift remains non-positive for all $t \in [0, \tau)$ and all choices of $\Lambda'_t \in \{\lambda, \lambda + \Delta\lambda\}$. This implies that for any effort process Λ', U' is an $\mathcal{F}^N$–supermartingale under $\mathbb{P}^{\Lambda'}$, and thus that Λ is at least as good as Λ' for the manager. If (9) does not hold for the effort process Λ, then choose Λ' such that for all $t \in [0, \tau)$, $\Lambda'_t = \lambda$ if and only if $H_t(\Gamma, \Lambda) \geq X_t b$. Then the drift of U' is everywhere non-negative and strictly positive on a set of $\mathbb{P}^{\Lambda'}$–positive measure. As a result of this, U' is an $\mathcal{F}^N$–submartingale under $\mathbb{P}^{\Lambda'}$ that is not a martingale, and thus Λ is suboptimal for the manager. This concludes the proof. ∎

APPENDIX C: THE VALUE FUNCTION

Proof of Proposition 2. In this Appendix, we shall work with the size-adjusted value social value function, v, instead of the size-adjusted value of the insurance company, f. These two functions are

related by $v(w) = f(w) + w$ for all $w \geq 0$, so that the system (22) can be rewritten in terms of v as

$$
\begin{cases}
v(w) = v(b)w/b & \text{if } w \in [0, b], \\
rv(w) = \mu - \lambda C - (\rho - r)w + v'(w)(\rho w + \lambda b) - \lambda[v(w) - v(w - b)] & \text{if } w \in (b, w^m], \\
v(w) = v(w^m) & \text{if } w \in (w^m, \infty).
\end{cases}
\tag{C.1}
$$

We assume throughout that $\mu - \lambda C \geq (\rho - r)b$, and that $\mu - \lambda C \geq (\rho - r)b(2 + r/\lambda)$ in Lemmas C.3 to C.7 below. Let $\mathcal{U}$ be the space of continuous functions over $\mathbb{R}_+$ that are continuously differentiable over (b, ∞), and consider the following linear first-order differential operator with delay:

$$
\mathcal{L}u(w) = (\rho w + \lambda b)u'(w) - \lambda[u(w) - u(w - b)],
\tag{C.2}
$$

for all $u \in \mathcal{U}$ and $w > b$. Define two auxiliary functions u_1 and u_2 in $\mathcal{U}$ that solve

$$
\begin{cases}
u_1(w) = 0 & \text{if } w \in [0, b], \\
ru_1(w) = \mu - \lambda C - (\rho - r)w + \mathcal{L}u_1(w) & \text{if } w \in (b, \infty),
\end{cases}
\tag{C.3}
$$

and

$$
\begin{cases}
u_2(w) = w & \text{if } w \in [0, b], \\
ru_2(w) = \mathcal{L}u_2(w) & \text{if } w \in (b, \infty).
\end{cases}
\tag{C.4}
$$

Given their initial conditions over the interval $[0, b]$, the solutions to (C.3) and (C.4) are recursively constructed over the intervals $(b, 2b]$, $(2b, 3b]$, and so on. Repeated applications of the Cauchy-Lipschitz theorem ensure that (C.3) and (C.4) both have a unique continuous solution. Neither u_1 and u_2 are differentiable at b. Indeed, using the operator $\mathcal{L}$ and the definitions of u_1 and u_2, it is straightforward to verify that

$$
u'_{1+}(b) = \frac{(\rho - r)b - \mu + \lambda C}{(\rho + \lambda)b} < 0 = u'_{1-}(b),
\tag{C.5}
$$

$$
u'_{2+}(b) = \frac{r + \lambda}{\rho + \lambda} < 1 = u'_{2-}(b).
\tag{C.6}
$$

However, the continuity of u_1 and u_2 ensures that they are both continuously differentiable over (b, ∞). This implies in turn that they are twice continuously differentiable over $(b, \infty) \setminus \{2b\}$, thrice continuously differentiable over $(b, \infty) \setminus \{2b, 3b\}$, and so on. One has the following results.

Lemma C.1. $\liminf_{w \to \infty} u'_1(w) \geq 1$.

Proof. We first show that $\liminf_{w \to \infty} u'_1(w) \neq -\infty$. Otherwise, there exists an increasing divergent sequence $(w_n)_{n \geq 0}$ in $(2b, \infty)$ such that $\lim_{n \to \infty} u'_1(w_n) = -\infty$ and $w_n = \arg\min_{w \in [0, w_n]}\{u'_{1+}(w)\}$. For each $n \geq 0$, one can find some $\tilde{w}_n \in (w_n - b, w_n)$ such that

$$
(\rho w_n + \lambda b)u'_1(w_n) = \lambda[u_1(w_n) - u_1(w_n - b)] + ru_1(w_n) + (\rho - r)w_n - \mu + \lambda C
$$

$$
= \lambda b u'_1(\tilde{w}_n) + ru_1(w_n) + (\rho - r)w_n - \mu + \lambda C,
$$

where the first equality follows from (C.2) and (C.3) and the second from the mean value theorem. This may be conveniently rewritten as

$$
u'_1(\tilde{w}_n) = \frac{w_n}{\lambda b}\left[\rho u'_1(w_n) - \frac{r}{w_n}u_1(w_n)\right] + \frac{\mu - \lambda C}{\lambda b} + u'_1(w_n) - \frac{(\rho - r)w_n}{\lambda b}.
\tag{C.7}
$$

Since $u_1(0) = 0$, one has $u_1(w_n) \geq w_n u_1'(w_n)$ by construction of the sequence $(w_n)_{n \geq 0}$. Moreover, $u_1'(w_n) < 0$ for n large enough. It then follows from (C.7) that for any such n,

$$u_1'(\tilde{w}_n) \leq \frac{(\rho - r)w_n u_1'(w_n)}{\lambda b} + \frac{\mu - \lambda C}{\lambda b}.$$

Therefore, since $u_1'(w_n) < 0$,

$$\frac{u_1'(\tilde{w}_n)}{u_1'(w_n)} \geq \frac{(\rho - r)w_n}{\lambda b} + \frac{\mu - \lambda C}{\lambda b u_1'(w_n)},$$

so that the ratio $u_1'(\tilde{w}_n)/u_1'(w_n)$ goes to ∞ as n goes to ∞. Using again the fact that $u_1'(w_n) < 0$ for n large enough, one obtains that eventually $u_1'(\tilde{w}_n) < u_1'(w_n)$, which, since $\tilde{w}_n < w_n$, contradicts the fact that $w_n = \arg\min_{w \in [0, w_n]}\{u_{1+}'(w)\}$. Thus $\liminf_{w \to \infty} u_1'(w)$ is a finite number, that we denote l. We now show that necessarily $l \geq 1$. Consider an increasing divergent sequence $(w_n)_{n \geq 0}$ in $(2b, \infty)$ such that $\lim_{n \to \infty} u_1'(w_n) = l$. There exists a constant C_1 such that $u_1(w_n) \geq l w_n + C_1$ for all $n \geq 0$. Constructing $\tilde{w}_n \in (w_n - b, w_n)$ as in (C.7), it follows that

$$\rho u_1'(w_n) + \lambda b \frac{u_1'(w_n)}{w_n} \geq \lambda b \frac{u_1'(\tilde{w}_n)}{w_n} + rl + \frac{rC_1 - \mu + \lambda C}{w_n} + \rho - r$$

for all $n \geq 0$. Letting n go to ∞, one obtains that

$$(\rho - r)(l - 1) \geq \lambda b \limsup_{n \to \infty} \frac{u_1'(\tilde{w}_n)}{w_n}.$$

If $l < 1$, this implies that $\limsup_{n \to \infty} u_1'(\tilde{w}_n) = -\infty$, which in turn contradicts the fact that $\liminf_{w \to \infty} u_1'(w) = l$ is a finite number. Hence $l \geq 1$, and the result follows. □

Lemma C.2. $u_2'(w) > 0$ for all $w \in (b, \infty)$.

Proof. One has $u_{2+}'(b) = (r + \lambda)/(\rho + \lambda) > 0$. Now suppose that u_2' vanishes over (b, ∞) and let $\tilde{w} > b$ be the first point at which it does so. Then, using (C.2) and (C.4), one obtains that

$$-\lambda[u_2(\tilde{w}) - u_2(\tilde{w} - b)] - ru_2(\tilde{w}) = 0,$$

which is impossible as u_2 is strictly increasing and strictly positive over $(0, \tilde{w}]$. Hence the result. □

Consider now the ratio $-u_{1+}'(w)/u_{2+}'(w)$, which is a continuous function of w over $[b, \infty)$. This quantity is strictly positive at $w = b$, and ultimately becomes strictly negative as w gets large enough by Lemmas C.1 and C.2. Thus $-u_{1+}'/u_{2+}'$ has a maximum over $[b, \infty)$. We denote by w^m the smallest point at which this maximum is reached over $[b, \infty)$. The function u defined by

$$u(w) = u_1(w) - \frac{u_{1+}'(w^m)}{u_{2+}'(w^m)} u_2(w) \tag{C.8}$$

for all $w \in \mathbb{R}_+$ then satisfies $u_+'(w^m) = 0$, and is non-decreasing as $u_+' \geq 0$ over $\mathbb{R}_+$. By (C.3), (C.4) and (C.8), one has $u(b)/b = -u_{1+}'(w^m)/u_{2+}'(w^m)$, and u satisfies

$$\begin{cases} u(w) = u(b)w/b & \text{if } w \in [0, b], \\ ru(w) = \mu - \lambda C - (\rho - r)w + \mathcal{L}u(w) & \text{if } w \in (b, \infty). \end{cases} \tag{C.9}$$

Note that since u_2 is strictly positive over $\mathbb{R}_{++}$, u can alternatively be characterized as the highest function of the form $u_1 + \beta u_2$ whose right derivative vanishes over $[b, \infty)$. Such functions form an increasing family ordered by their slope β at 0, and they satisfy the analogue of (C.9) with β instead

of $u(b)/b$. This immediately implies the statement in Proposition 2(i). For $\beta > u(b)/b$, the right derivative of $u_1 + \beta u_2$ is always strictly positive, while it takes strictly negative values for $\beta < u(b)/b$. Letting $\alpha^m = u(b)/b - 1$ and $\phi_{\alpha^m}(w) = u(w) - w$ for all $w \geq 0$, one obtains that $\phi'_{\alpha^m+}(w^m) = -1$, and that the equation $\phi'_\alpha(w) = -1$ admits no solution for any $\alpha > \alpha^m$. To derive Proposition 2(ii), we need only to check that ϕ_{α^m} is actually differentiable at w^m, which results from the following lemma.

Lemma C.3. *If $\mu - \lambda C > (\rho - r)b(2 + r/\lambda)$, then $u''_+(b) < 0$ and $w^m > b$.*

Proof. We first prove that

$$\frac{u(b)}{b} \geq \frac{\mu - \lambda C - (\rho - r)b}{(r + \lambda)b}. \tag{C.10}$$

Indeed, substituting the explicit values (C.5) and (C.6) for $u'_{1+}(b)$ and $u'_{2+}(b)$ in the expression for $u'_+(b)$ yields

$$u'_+(b) = u'_{1+}(b) + \frac{u(b)}{b} u'_{2+}(b) = \frac{(\rho - r)b - \mu + \lambda C + (r + \lambda)u(b)}{(\rho + \lambda)b}, \tag{C.11}$$

and (C.10) follows from the fact that $u'_+ \geq 0$ over $\mathbb{R}_+$ and thus in particular $u'_+(b) \geq 0$. We next prove that $u''_+(b) < 0$. Differentiating (C.9) to the right of any $w \geq b$ leads to

$$(\rho w + \lambda b)u''_+(w) = \lambda[u'_+(w) - u'_+(w - b)] - (\rho - r)[u'_+(w) - 1].$$

Applying this formula at b and using (C.10) and (C.11), one then obtains that

$$(\rho + \lambda)bu''_+(b) = \lambda\left[u'_+(b) - \frac{u(b)}{b}\right] - (\rho - r)[u'_+(b) - 1]$$

$$= \frac{(\lambda - \rho + r)[(\rho - r)b - \mu + \lambda C] + (r - \rho)(r + 2\lambda)u(b)}{(\rho + \lambda)b} + \rho - r$$

$$\leq \frac{\lambda[(\rho - r)b - \mu + \lambda C]}{(r + \lambda)b} + \rho - r,$$

which is strictly negative under the assumption of the lemma. Hence the claim. We finally prove that $w^m > b$. A sufficient condition is that the right derivative of $-u'_{1+}/u'_{2+}$ at b be strictly positive. Differentiating (C.3) and (C.4) to the right of b leads to

$$(\rho + \lambda)bu''_{1+}(b) = (\lambda - \rho + r)u'_{1+}(b) + \rho - r, \tag{C.12}$$

$$(\rho + \lambda)bu''_{2+}(b) = (\lambda - \rho + r)u'_{2+}(b) - \lambda. \tag{C.13}$$

Substituting the explicit values (C.5) and (C.6) for $u'_{1+}(b)$ and $u'_{2+}(b)$ in (C.12) and (C.13) yields

$$-u''_{1+}(b)u'_{2+}(b) + u''_{2+}(b)u'_{1+}(b) = -\frac{(\rho - r)u'_{2+}(b) + \lambda u'_{1+}(b)}{(\rho + \lambda)b}$$

$$= \frac{\lambda}{b^2(\rho + \lambda)^2}\left[\mu - \lambda C - (\rho - r)b\left(2 + \frac{r}{\lambda}\right)\right],$$

which is strictly positive under the assumption of the lemma. This implies the result. $\square$

Under the assumption of Lemma C.3, u is differentiable at w^m and w^m is the smallest point at which u' vanishes, and at which ϕ_{α^m} equals -1. We now show that u is concave over $[0, w^m]$, and strictly so over $[b, w^m]$. Differentiating (C.9) to the right of $2b$ and using the inequalities (C.5)

and (C.6), one can verify that $u''_+(2b) > u''_-(2b)$. Since u is twice continuously differentiable over $(b, \infty) \setminus \{2b\}$, u''_+ is upper semicontinuous over $[b, \infty)$, and hence the set $\{w \geq b \mid u''_+(w) \geq 0\}$ is closed. Denote by w^c its smallest element. Since u is non-decreasing and $u'(w^m) = 0$, one must have $u''_+(w^m) \geq 0$, and thus $w^m \geq w^c$. By Lemma C.3, $w^c > b$ and $u''_+ < 0$ over (b, w^c), so that u is strictly concave over $[b, w^c]$. Moreover, along with the inequalities (C.5) and (C.6), the representation (C.8) implies that $u'_+(b) < u'_-(b)$. Thus, as u is linear over $[0, b]$, it is globally concave over $[0, w^c]$. In order to derive similar properties of u on the interval $[0, w^m]$, we now prove that w^c actually coincides with w^m. One first has the following result.

Lemma C.4. If $\mu - \lambda C > (\rho - r)b(2 + r/\lambda)$, then $w^c \geq 2b$.

Proof. Suppose by way of contradiction that $w^c \in (b, 2b)$. Since u is twice continuously differentiable over $(b, 2b)$, $u''(w^c) = 0$ and $u'' < 0$ over (b, w^c). We consider three cases in turn.

Case 1. Suppose first that $\lambda \leq \rho - r$. Then, since $w^c - b < b$ and $u''(w^c) = 0$, differentiating (C.9) at w^c yields

$$\lambda \left[u'(w^c) - \frac{u(b)}{b} \right] - (\rho - r)[u'(w^c) - 1] = 0,$$

which implies that $\lambda u(b)/b - \rho + r = (\lambda - \rho + r)u'(w^c) \leq 0$. By (C.10), it follows that

$$\frac{\lambda[\mu - \lambda C - (\rho - r)b]}{b(r + \lambda)} \leq \rho - r,$$

or equivalently $\mu - \lambda C \leq (\rho - r)b(2 + r/\lambda)$, which contradicts the assumption of the lemma.

Case 2. Suppose next that $\lambda \geq 2\rho - r$. Differentiating (C.9) twice over $(b, 2b)$ yields

$$(\rho w + \lambda b)u'''(w) = \lambda[u''(w) - u''(w - b)] - (2\rho - r)u''(w) = (\lambda - 2\rho + r)u''(w)$$

for all $w \in (b, 2b)$, where the second inequality follows from the fact that u is linear over $(0, b)$. Since $\lambda \geq 2\rho - r$ and $u'' < 0$ over (b, w^c), one has $u''' \leq 0$ over this interval. This implies that $u''(w^c) \leq u''_+(b)$, which is impossible since $u''(w^c) = 0$ by construction and $u''_+(b) < 0$ by Lemma C.3.

Case 3. Suppose finally that $\rho - r < \lambda < 2\rho - r$. Differentiating (C.9) twice as in Case 2 and using the fact that $\lambda - 2\rho + r < 0$ shows that u''' and u'' have opposite signs over $(b, 2b)$. It follows that $u''' > 0$ and hence $u'' > u''_+(b)$ over (b, w^c). Using again the fact that $\lambda - 2\rho + r < 0$, one obtains that

$$u'''(w) = \frac{(\lambda - 2\rho + r)u''(w)}{\rho w + \lambda b} < \frac{(\lambda - 2\rho + r)u''_+(b)}{\rho w + \lambda b}$$

for all $w \in (b, w^c)$. One then has

$$u''(w^c) = u''_+(b) + \int_b^{w^c} \frac{(\lambda - 2\rho + r)u''(w)}{\rho w + \lambda b} \, dw < \left[1 + \int_b^{w^c} \frac{\lambda - 2\rho + r}{\rho w + \lambda b} \, dw \right] u''_+(b) \equiv C_2 u''_+(b).$$

Since $u''(w^c) = 0$ and $u''_+(b) < 0$, one obtains a contradiction if $C_2 > 0$. Note that

$$\int_b^{w^c} \frac{1}{\rho w + \lambda b} \, dw < \int_b^{2b} \frac{1}{\rho w + \lambda b} \, dw < \frac{1}{\rho + \lambda}.$$

Since $\rho - r < \lambda < 2\rho - r$, this implies that

$$C_2 > 1 + \frac{\lambda - 2\rho + r}{\rho + \lambda} > 0,$$

32

and the result follows. $\qquad\square$

The following lemma then implies that u is concave over $[0, w^m]$, and strictly so over $[b, w^m]$.

Lemma C.5. *If $\mu - \lambda C > (\rho - r)b(2 + r/\lambda)$, then $w^c = w^m$.*

Proof. We first show that $u'' > 0$ in an interval $(w^c, w^c + \varepsilon)$ for some $\varepsilon > 0$. Whenever $w^c = 2b$ and $u''_+(2b) > 0$, this is immediate. Otherwise $u''_+(w^c) = 0$. Differentiating (C.9) twice to the right of w^c then yields

$$(\rho w^c + \lambda b)u'''_+(w^c) = \lambda[u''_+(w^c) - u''_+(w^c - b)] - (2\rho - r)u''_+(w^c) = -\lambda u''_+(w^c - b) > 0$$

where the strict inequality follows from the fact that $w^c - b \in [b, w^c)$ by Lemma C.4, and that $u''_+ < 0$ over $[b, w^c)$. Since $u''_+(w^c) = 0$ and $u'''_+(w^c) > 0$, $u'' > 0$ in an interval $(w^c, w^c + \varepsilon)$ for some $\varepsilon > 0$, as claimed. Suppose by way of contradiction that $w^c \neq w^m$, so that actually $w^m > w^c$. Then $u'(w^c) > 0$ as w^m is the smallest point at which u' vanishes. Since $u'(w^m) = u'(w^c) + \int_{w^v}^{w^m} u''(w)\,dw$, this implies that u'' cannot be non-negative over the whole interval (w^c, w^m). Let $\bar{w}^c = \inf\{w > w^c \,|\, u''(w) < 0\} \in (w^c, w^m)$. One has $u'' \geq 0$ over $(w^c, \bar{w}^c)$ and $u''(\bar{w}^c) = 0$ since $\bar{w}^c > w^c \geq 2b$ by Lemma C.4 and u is twice continuously differentiable over $(2b, \infty)$. We now show that $\bar{w}^c \geq w^c + b$. Note that one must have $u'''_+(\bar{w}^c) \leq 0$, since otherwise u'' would be strictly positive in an interval $(\bar{w}^c, \bar{w}^c + \eta)$ for some $\eta > 0$. Differentiating (C.9) twice to the right of $\bar{w}^c$ then yields

$$0 \geq (\rho \bar{w}^c + \lambda b)u'''_+(\bar{w}^c) = \lambda[u''(\bar{w}^c) - u''_+(\bar{w}^c - b)] - (2\rho - r)u''(\bar{w}^c) = -\lambda u''_+(\bar{w}^c - b),$$

and thus $u''_+(\bar{w}^c - b) \geq 0$. Now, $u''_+ < 0$ over (b, w^c). Since $\bar{w}^c > 2b$ and thus $\bar{w}^c - b > b$, it follows that $\bar{w}^c - b \geq w^c$, which implies the claim. Since $u'' \geq 0$ over $(w^c, \bar{w}^c)$, u is convex over $[\bar{w}^c - b, \bar{w}^c]$. Then, since

$$0 = (\rho \bar{w}^c + \lambda b)u''(\bar{w}^c) = \lambda[u'(\bar{w}^c) - u'(\bar{w}^c - b)] - (\rho - r)[u'(\bar{w}^c) - 1]$$

by differentiation of (C.9) at $\bar{w}^c$, one obtains that $u'(\bar{w}^c) \geq 1$. One then has

$$\rho \bar{w}^c + \lambda b u'(\bar{w}^c) \leq (\rho \bar{w}^c + \lambda b)u'(\bar{w}^c)$$

$$= \lambda[u(\bar{w}^c) - u(\bar{w}^c - b)] + ru(\bar{w}^c) + (\rho - r)\bar{w}^c - \mu + \lambda C \qquad \text{(C.14)}$$

$$\leq \lambda b u'(\bar{w}^c) + ru(\bar{w}^c) + (\rho - r)\bar{w}^c - \mu + \lambda C,$$

where the first inequality follows from $u'(\bar{w}^c) \geq 1$, the second from (C.9) and the third from the convexity of u over $[\bar{w}^c - b, \bar{w}^c]$. As a result of (C.14), $u(\bar{w}^c) \geq \bar{w}^c + (\mu - \lambda C)/r$. Since $w^m > \bar{w}^c$ and u is non-decreasing, one must have $u(\bar{w}^m) > (\mu - \lambda C)/r$. However, writing (C.9) at w^m yields

$$0 = (\rho w^m + \lambda b)u'(\bar{w}^m) = \lambda[u(w^m) - u(w^m - b)] + ru(w^m) + (\rho - r)w^m - \mu + \lambda C,$$

which, since u is non-decreasing, implies that $u(\bar{w}^m) < (\mu - \lambda C)/r$. This contradiction establishes that $w^c = w^m$, as claimed. $\qquad\square$

Finally, similar arguments can be used to show that u' vanishes only at w^m, so that u is strictly increasing over $\mathbb{R}_+$.

Lemma C.6. *If $\mu - \lambda C > (\rho - r)b(2 + r/\lambda)$, then $u' > 0$ over $(b, \infty) \setminus \{w^m\}$.*

Proof. Since $w^c = w^m$, it follows as in the proof of Lemma C.5 that $u'' > 0$ in an interval $(w^m, w^m + \varepsilon)$ for some $\varepsilon > 0$. Since $u'(w^m) = 0$, one must have $u' > 0$ over $(w^m, w^m + \varepsilon)$. Suppose now that $u'(w) = 0$ for some $w \geq w^m + \varepsilon$, and let $\bar{w}^m = \inf\{w \geq w^m + \varepsilon \,|\, u'(w) = 0\}$. Because u is continuously differentiable over (b, ∞), $u'(\bar{w}^m) = 0 = u'(w^m)$. Since $u'' > 0$ over $(w^m, w^m + \varepsilon) \subset (w^m, \bar{w}^m)$, this implies that u'' cannot be non-negative over the whole interval $(w^m, \bar{w}^m)$. Let $\bar{\bar{w}}^c = \inf\{w > w^m \,|\, u''(w) < 0\} \in (w^m, \bar{w}^m)$. One has $u'' \geq 0$ over $(w^m, \bar{\bar{w}}^c)$ and $u''(\bar{\bar{w}}^c) = 0$ since $\bar{\bar{w}}^c > w^m = w^c \geq 2b$ by Lemmas C.4 and C.5. Proceeding as for $\bar{w}^c$ in the proof of Lemma C.5, one can show that $\bar{\bar{w}}^c \geq w^m + b$, so that u is convex over $[\bar{\bar{w}}^c - b, \bar{\bar{w}}^c]$, and that $u'(\bar{\bar{w}}^c) \geq 1$. One can then deduce similarly that $u(\bar{\bar{w}}^c) \geq \bar{\bar{w}}^c + (\mu - \lambda C)/r$, which yields a contradiction as $u(\bar{w}^m) \geq u(\bar{\bar{w}}^c)$ must be strictly smaller than $(\mu - \lambda C)/r$, just as $u(w^m)$. Hence the result. $\qquad\square$

The statements in Proposition 2(iii) and (iv) then follow from the fact that $\phi_{\alpha^m}(w) = u(w) - w$ for all $w \geq 0$. This completes the proof of Proposition 2. $\qquad\blacksquare$

The value function v that results from the optimal contract can then be defined as

$$v(w) = \min\{u(w), u(w^m)\} \qquad (\text{C.15})$$

for all $w \geq 0$. It is linear over $[0, b]$, globally concave and non-decreasing. It is strictly increasing over $[0, w^m]$, flat above w^m, and strictly concave over $[b, w^m]$. The corresponding value function f for the insurance company, defined by $f(w) = v(w) - w$ for all $w \geq 0$ or equivalently by (25), is linear over $[0, b]$ and globally concave. It has a slope -1 above w^m, and is strictly concave over $[b, w^m]$. The next lemma is crucial in establishing the verification theorem. Note that $f' = f'_+$ over (b, ∞).

Lemma C.7. *If $\mu - \lambda C > (\rho - r)b(2 + r/\lambda)$, then*

$$\mathcal{L}v(w) - rv(w) \leq (\rho - r)w - \mu + \lambda C$$

for all $w \in (b, \infty)$. As a result of this,

$$(\rho w + \lambda b)f'_+(w) - \lambda[f(w) - f(w - b)] - rf(w) \leq -\mu + \lambda C$$

for all $w \in [b, \infty)$, with equality if $w \in [b, w^m]$.

Proof. For $w \in [b, w^m]$, the result is a consequence of (C.9) and (C.15), the case $w = b$ following by continuity. For any $w > w^m$, one has

$$\mathcal{L}v(w) - rv(w) - (\rho - r)w + \mu - \lambda C = -\lambda[v(w^m) - v(w - b)] - rv(w^m) - (\rho - r)w + \mu - \lambda C,$$

$$= \lambda[v(w - b) - v(w^m - b)] - (\rho - r)(w - w^m)$$

$$\leq [\lambda v'_+(w^m - b) - \rho + r](w - w^m),$$

where the first equality follows from the fact that v is flat above w^m, the second from substituting $\mathcal{L}v(w^m) - rv(w^m) = (\rho - r)w^m - \mu + \lambda C$ into the second expression, and the inequality from the concavity of v. By construction, $v'_+(w^m - b) = u'_+(w^m - b)$ so that we need only to prove that

$$\lambda u'_+(w^m - b) - \rho + r \leq 0.$$

Differentiating (C.9) twice to the right of w^m and taking advantage from $u'(w^m) = 0$ yields

$$\lambda u'_+(w^m - b) - \rho + r = -(\rho w^m + \lambda b)u''_+(w^m),$$

which is non-positive as $u''_+(w^m) \geq 0$. Hence the result. $\qquad\blacksquare$

34

Proof of Proposition 3. Fix an arbitrary contract $\Gamma = (X, L, \tau)$ that induces maximal risk prevention, $\Lambda_t = \lambda$ for all $t \in [0, \tau)$, and delivers the manager an expected discounted utility W_{0-} given initial firm size X_0. For simplicity, let us drop the mention of the contract Γ and of the effort process Λ in the remainder of the proof. The manager's continuation utility follows a process W whose dynamics is described by (11). In line with our assumption that X is $\mathcal{F}^N$–predictable while W is $\mathcal{F}^N$–adapted, one can assume without loss of generality that X has left-continuous paths, while W has right-continuous paths. Now, observe that, by construction, the function f is continuously differentiable over (b, ∞), so that the function F is continuously differentiable over $\{(\xi, \omega) \in \mathbb{R}_{++} \times \mathbb{R}_+ \mid \omega/\xi > b\}$. Since f is continuous at b and $f'_+(b)$ is finite, one can continuously extend the derivative of F to the set $\{(\xi, \omega) \in \mathbb{R}_{++} \times \mathbb{R}_+ \mid \omega/\xi = b\}$. As limited liability and incentive compatibility imply that $W_{t-}/X_t \geq b$ for all $t \in [0, \tau)$, applying the change of variable formula for multidimensional processes of bounded variations (Dellacherie and Meyer (1982, Chapter VI, Section 92)) yields

$$e^{-rT} F(X_{T+}, W_T) = F(X_0, W_{0-}) + \int_0^T e^{-rt}[(\rho W_{t-} + \lambda H_t) F_W(X_t, W_{t-}) - rF(X_t, W_{t-})] \, dt$$

$$+ \int_0^T e^{-rt} F_X(X_t, W_{t-}) \, dX_t^c - \int_0^T e^{-rt} F_W(X_t, W_{t-}) \, dL_t^c \quad \text{(D.1)}$$

$$+ \sum_{t \in [0,T]} e^{-rt}[F(X_{t+}, W_t) - F(X_t, W_{t-})]$$

for all $T \in [0, \tau)$, where X^c and L^c stand for the pure continuous parts of X and L. For each $t \in [0, T]$, one has the following decomposition of the jump in $F(X_t, W_{t-})$ at date t:

$$F(X_{t+}, W_t) - F(X_t, W_{t-}) = F(X_{t+}, W_t) - F(X_t, W_t)$$

$$+ F(X_t, W_{t-} - H_t \Delta N_t - \Delta L_t) - F(X_t, W_{t-})$$

$$= F(X_{t+}, W_t) - F(X_t, W_t) \quad \text{(D.2)}$$

$$+ F(X_t, W_{t-} - H_t \Delta N_t - \Delta L_t) - F(X_t, W_{t-} - H_t \Delta N_t)$$

$$+ F(X_t, W_{t-} - H_t \Delta N_t) - F(X_t, W_{t-}).$$

To derive (D.2), we have used the fact that $W_t = W_{t-} - H_t \Delta N_t - \Delta L_t$, where $\Delta N_t = N_t - N_{t-}$ and $\Delta L_t = L_t - L_{t-}$ for all $t \in [0, T]$, with $N_{0-} = L_{0-} = 0$ by convention. Now fix $T \in [0, \tau)$ and, as in Appendix A, let $M_t = N_t - \lambda t$ for all $t \geq 0$. Using (D.2) and

$$\sum_{t \in [0,T]} e^{-rt}[F(X_t, W_{t-} - H_t \Delta N_t) - F(X_t, W_{t-})] = \int_0^T e^{-rt}[F(X_t, W_{t-} - H_t) - F(X_t, W_{t-})] \, dN_t,$$

one can then rewrite (D.1) as

$$e^{-rT} F(X_{T+}, W_T) = F(X_0, W_{0-}) + \int_0^T e^{-rt}[F(X_t, W_{t-} - H_t) - F(X_t, W_{t-})] \, dM_t + A_1 + A_2 + A_3, \quad \text{(D.3)}$$

where A_1 is a standard integral with respect to time,

$$A_1 = \int_0^T e^{-rt}\{(\rho W_{t-} + \lambda H_t)F_W(X_t, W_{t-}) - \lambda[F(X_t, W_{t-}) - F(X_t, W_{t-} - H_t)] - rF(X_t, W_{t-})\}\,dt,$$

A_2 accounts for changes in the size of firm,

$$A_2 = \int_0^T e^{-rt}F_X(X_t, W_{t-})\,dX_t^c + \sum_{t\in[0,T]} e^{-rt}[F(X_{t+}, W_t) - F(X_t, W_t)],$$

and A_3 accounts for changes in cumulative transfers,

$$A_3 = -\int_0^T e^{-rt}F_W(X_t, W_{t-})\,dL_t^c + \sum_{t\in[0,T]} e^{-rt}[F(X_t, W_{t-} - H_t\Delta N_t - \Delta L_t) - F(X_t, W_{t-} - H_t\Delta N_t)].$$

We treat each of these terms in turn.

Consider first A_1. For each $t \in [0, T]$, let $w_t = W_{t-}/X_t$ and $h_t = H_t/X_t$. The homogeneity of F implies that $F_W(X_t, W_{t-}) = f'_+(w_t)$ for all $t \in [0, T]$. Thus

$$\begin{aligned}
A_1 &= \int_0^T e^{-rt}X_t\{(\rho w_t + \lambda h_t)f'_+(w_t) - \lambda[f(w_t) - f(w_t - h_t)] - rf(w_t)\}\,dt \\
&\leq \int_0^T e^{-rt}X_t\{(\rho w_t + \lambda b)f'_+(w_t) - \lambda[f(w_t) - f(w_t - b)] - rf(w_t)\}\,dt \qquad \text{(D.4)} \\
&\leq \int_0^t e^{-rt}X_t(-\mu + \lambda C)\,dt
\end{aligned}$$

where the first and second inequalities respectively follow from the concavity of f and from Lemma C.7, along with the fact that $w_t \geq h_t \geq b$ for all $t \in [0, T]$ by limited liability and incentive compatibility.

Consider next A_2. The homogeneity of F implies that $F_X(X_t, W_{t-}) = f(w_t) - w_t f'(w_t)$ for all $t \in [0, T]$. One can then rewrite A_2 as

$$A_2 = \int_0^T e^{-rt}[f(w_t) - w_t f'(w_t)]\,dX_t^c + \sum_{t\in[0,T]} e^{-rt}W_t\left[\frac{X_{t+}}{W_t}f\left(\frac{W_t}{X_{t+}}\right) - \frac{X_t}{W_t}f\left(\frac{W_t}{X_t}\right)\right] \leq 0, \qquad \text{(D.5)}$$

where the inequality can be justified as follows. Since f is concave and vanishes at 0, $f(w) - wf'(w) \geq 0$ for all $w \geq 0$. Because $X^c = \{X_t^c\}_{t\geq 0}$ is a non-increasing process, this implies that the first term on the right-hand side of (D.5) is non-positive. The aforementioned properties of f also imply that $f(w)/w$ is a non-increasing function of w. Since $W_t/X_{t+} \geq W_t/X_t$ for all $t \in [0, T]$, this implies that the second term on the right-hand side of (D.5) is non-positive. As a result of this, $A_2 \leq 0$.

Consider finally A_3. The homogeneity of F and the concavity of f imply that, for each $t \in [0, T]$,

$$\begin{aligned}
&F(X_t, W_{t-} - H_t\Delta N_t - \Delta L_t) - F(X_t, W_{t-} - H_t\Delta N_t) \\
&= X_t\left[f\left(\frac{W_{t-} - H_t\Delta N_t - \Delta L_t}{X_t}\right) - f\left(\frac{W_{t-} - H_t\Delta N_t}{X_t}\right)\right] \\
&= -f'\left(\frac{W_{t-} - H_t\Delta N_t}{X_t}\right)\Delta L_t \\
&\leq \Delta L_t,
\end{aligned}$$

where the last inequality follows from $f' \geq -1$. Using the fact that $-F_W(X_t, W_{t-}) = -f'(w_t) \leq 1$ for all $t \in [0, T]$, along with the definition of A_3, one therefore obtains that

$$A_3 \leq \int_0^T e^{-rt} \, dL_t^c + \sum_{t \in [0,T]} e^{-rt} \Delta L_t = \int_0^T e^{-rt} \, dL_t. \tag{D.6}$$

Using (D.3) along with the upper bounds (D.4), (D.5) and (D.6) for A_1, A_2 and A_3, it follows that

$$F(X_0, W_{0-}) \geq e^{-rT} F(X_{T+}, W_T) + \int_0^T e^{-rt} [X_t(\mu - \lambda C) dt - dL_t]$$

$$+ \int_0^T e^{-rt} [F(X_t, W_{t-} - H_t) - F(X_t, W_{t-})] \, dM_t$$

$$\tag{D.7}$$

$$= e^{-rT} F(X_{T+}, W_T) + \int_0^T e^{-rt} [X_t(\mu dt - C dN_t) - dL_t]$$

$$+ \int_0^T e^{-rt} [F(X_t, W_{t-}) - F(X_t, W_{t-} - H_t) + X_t C] \, dM_t$$

Using the fact that $M = \{M_t\}_{t \geq 0}$ is an $\mathcal{F}^N$–martingale under maximal risk prevention, and that the process defined by $t \mapsto e^{-rt}[F(X_t, W_{t-}) - F(X_t, W_{t-} - H_t) + X_t C]$ is $\mathcal{F}^N$–predictable, taking expectations in (D.7) yields

$$F(X_0, W_{0-}) \geq \mathbb{E}\left[e^{-rT \wedge \tau} F(X_{T \wedge \tau+}, W_{T \wedge \tau}) + \int_0^{T \wedge \tau} e^{-rt} [X_t(\mu dt - C dN_t) - dL_t] \right]$$

$$= \mathbb{E}\left[\int_0^\tau e^{-rt} [X_t(\mu dt - C dN_t) - dL_t] \right]$$

$$- \mathbb{E}\left[1_{\{T < \tau\}} \left\{ \int_T^\tau e^{-rt} [X_t(\mu dt - C dN_t) - dL_t] - e^{-rT} F(X_{T+}, W_T) \right\} \right]$$

$$= \mathbb{E}\left[\int_0^\tau e^{-rt} [X_t(\mu dt - C dN_t) - dL_t] \right] \tag{D.8}$$

$$- e^{-rT} \mathbb{E}\left[1_{\{T < \tau\}} \left\{ \mathbb{E}\left[\int_T^\tau e^{-r(t-T)} [X_t(\mu dt - C dN_t) - dL_t] \,|\, \mathcal{F}_T^N \right] - F(X_{T+}, W_T) \right\} \right]$$

$$\geq \mathbb{E}\left[\int_0^\tau e^{-rt} [X_t(\mu dt - C dN_t) - dL_t] \right]$$

$$- e^{-rT} \mathbb{E}\left[1_{\{T < \tau\}} \left[\frac{X_T(\mu - \lambda C)}{r} - W_T - F(X_{T+}, W_T) \right] \right]$$

for all $T \in \mathbb{R}_+$, where the first equality follows from the fact that $W_\tau = 0$ by (4), and the second from the fact that $\rho > r$ along with the definition (4) of W_T and the monotonicity of X. Now, observe that for each $T \geq 0$, $F(X_{T+}, W_T) + W_T = X_{T+} v(W_T / X_{T+})$, which is non-negative and bounded above by $X_{0-} v(w^m)$. Taking limits as T goes to ∞ in (D.8) then implies (26). $\blacksquare$

Appendix E: The Verification Theorem

Proof of Proposition 4. For simplicity, let us drop the mention of the effort process Λ in the remainder of the proof. It follows from (28) to (30) that $W_{t-} \geq X_t b$ for all $t \geq 0$. Actually, using (27), it is

37

easy to verify that the strict inequality $W_{t-} > X_t b$ holds for all $t \geq 0$. This implies that the processes $W = \{W_t\}_{t\geq0}$ and $X = \{X_t\}_{t\geq0}$ defined by (27) and (28) remain strictly positive, so one can set $\tau = \infty$ as in (32). Fix some $T > 0$. Proceeding as for equality (D.3), one obtains that

$$e^{-rT}F(X_{T+}, W_T) = F(X_0, W_{0-}) + \int_0^T e^{-rt}[F(X_t, W_{t-}-X_t b) - F(X_t, W_{t-})]\, dM_t + A_1 + A_2 + A_3, \quad \text{(E.1)}$$

where $M_t = N_t - \lambda t$ for all $t \geq 0$ and A_1, A_2, and A_3 are defined as in the proof of Proposition 3, with $H_t = X_t b$ for all $t \geq 0$. We treat each of these terms in turn.

Consider first A_1. For each $t \in [0, T]$, let $w_t = W_{t-}/X_t$, which lies in $[b, w^m]$ by construction. The homogeneity of F implies that $F_W(X_t, W_{t-}) = f'_+(w_t)$ for all $t \in [0, T]$. One can then rewrite A_1 as

$$
\begin{aligned}
A_1 &= \int_0^{\tau_1} e^{-rt} X_t \{(\rho w_t + \lambda b) f'_+(w_t) - \lambda[f(w_t) - f(w_t - b)] - r f(w_t)\}\, dt \\[4pt]
&\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \text{(E.2)} \\[4pt]
&= \int_0^T e^{-rt} X_t (-\mu + \lambda C)\, dt,
\end{aligned}
$$

where the second equality follows from Lemma C.7 and from the fact that $w_t \in [b, w^m]$ for all $t \in [0, T]$.

Consider next A_2. Let $\nu = \sup\{n \geq 1 \,|\, \tau_n \leq T\}$. Since the process X is purely discontinuous,

$$
\begin{aligned}
A_2 &= \sum_{t \in [0,T]} e^{-rt}[F(X_{t+}, W_t) - F(X_t, W_t)] \\[4pt]
&= \sum_{n=1}^{\nu} e^{-r\tau_n}[F(\xi_n, W_{\tau_n}) - F(\xi_{n-1}, W_{\tau_n})] \\[4pt]
&= \sum_{n=1}^{\nu} e^{-r\tau_n}\left[\xi_n f(b) - \xi_{n-1} f\left(\frac{W_{\tau_n}}{\xi_{n-1}}\right)\right] \quad\quad\quad\quad\quad\quad\quad \text{(E.3)} \\[4pt]
&= \sum_{n=1}^{\nu} e^{-r\tau_n} f(b)\left[\xi_n - \frac{W_{\tau_n}}{b}\right] \\[4pt]
&= 0,
\end{aligned}
$$

where the second equality follows from (28) to (30), the third from (30) and the homogeneity of F, the fourth from the fact that $W_{\tau_n}/\xi_{n-1} < b$ and that f is linear over $[0, b]$, and the fifth from (30).

Consider finally A_3. Since the process $L = \{L_t\}_{t\geq0}$ is continuous except perhaps at date 0,

$$
\begin{aligned}
A_3 &= -\int_0^T e^{-rt} F_W(X_t, W_{t-})\, dL_t^c + F(X_0, W_{0-} - L_0) - F(X_0, W_{0-}) \\[4pt]
&= -\int_0^T e^{-rt} f'(w_t) X_t (\rho w^m + \lambda b) 1_{\{W_t = X_t w^m\}}\, dt + \max\{W_{0-} - X_0 w^m, 0\} \quad \text{(E.4)} \\[4pt]
&= \int_0^T e^{-rt}\, dL_t,
\end{aligned}
$$

where the second equality follows from (31) and the homogeneity of F, and the third from (31) along with the fact that $W_t = X_t w^m$ implies $w_t = w^m$ and thus $f'(w_t) = -1$.

Taking expectations in (E.1) and taking advantage of (E.2), (E.3) and (E.4), one obtains that

$$F(X_0, W_{0-}) = \mathbb{E}\left[e^{-rT} F(X_{T+}, W_T) + \int_0^T e^{-rt}[X_t(\mu dt - C dN_t) - dL_t]\right], \quad \text{(E.5)}$$

where we used the fact that $M = \{M_t\}_{t \geq 0}$ as defined in Appendix D is an $\mathcal{F}^N$–martingale under maximal risk prevention, and that the process defined by $t \mapsto e^{-rt}[F(X_t, W_{t-}) - F(X_t, W_{t-} - X_t b)]$ is $\mathcal{F}^N$–predictable. By construction, $F(X_{T+}, W_T)$ is uniformly bounded in T as $X_{T+} \in [0, X_{0-}]$ and $W_T \in [0, X_0 w^m]$ for all $T > 0$. Letting T go to ∞ in (E.5) and using the fact that $\tau = \infty$ yields

$$F(X_0, W_{0-}) = \mathbb{E}\left[\int_0^\tau e^{-rt}[X_t(\mu dt - C dN_t) - dL_t] \right]. \tag{E.6}$$

By Proposition 3, the insurance company's expected discounted profit from any contract that induces maximal risk prevention is at most $F(X_0, W_{0-})$. Thus (E.6) implies that the contract characterized in (27) to (32) is optimal in this class of contracts. ∎

APPENDIX F: THE LONG RUN

Proof of Proposition 5. For each $t \geq 0$, let $w_t = W_{t-}/X_t$ be the manager's size-adjusted utility at date t. It follows from (27) to (31) that the process $\{w_t\}_{t \geq 0}$ evolves according to

$$dw_t = (\rho w_t + \lambda b)dt - \min\{w_t - b, b\}dN_t - dl_t \tag{F.1}$$

for all $t \geq 0$, where $\{l_t\}_{t \geq 0}$ defined by

$$l_t = \max\{w_0 - w^m, 0\} + \int_0^t (\rho w^m + \lambda b) 1_{\{w_{s+} = w^m\}} \, ds$$

for all $t \geq 0$ is the size-adjusted transfer process. By construction, $\{w_t\}_{t \geq 0}$ is a Markov process that satisfies the Feller property. Let $P : \mathbb{R}_+ \times [b, w^m] \times \mathcal{B}([b, w^m]) \to [0, 1]$ be its transition function. To each time duration t, initial position w and Borel subset A of $[b, w^m]$, it associates the probability $P(t, w, A)$ of transiting from w to a position in A following an interval of time t.

It is straightforward to check from (F.1) that the minimum amount of time that it takes $\{w_t\}_{t \geq 0}$ to transit from b to w^m is

$$\underline{t} = \frac{1}{\rho} \ln\left(\frac{\rho w^m/b + \lambda}{\rho + \lambda} \right). \tag{F.2}$$

Moreover, $P(\underline{t}, w, \{w^m\}) \geq e^{-\lambda \underline{t}}$ for all $w \in [b, w^m]$. Hence the $\underline{t}$–transition function $P(\underline{t}, \cdot, \cdot)$ satisfies Condition M in Stokey and Lucas (1989, Chapter 11, Section 4). Specifically, for each $A \in \mathcal{B}([b, w^m])$ the following holds. Either $w^m \in A$ and $P(\underline{t}, w, A) \geq e^{-\lambda \underline{t}}$ for all $w \in [b, w^m]$, or $w^m \notin A$ and $P(\underline{t}, w, [b, w^m] \setminus A) \geq e^{-\lambda \underline{t}}$ for all $w \in [b, w^m]$. Let $T_{\underline{t}}^* : \Delta([b, w^m]) \to \Delta([b, w^m])$ be the adjoint operator associated with $P(\underline{t}, \cdot, \cdot)$ on the set of Borel probability measures on $[b, w^m]$, defined by

$$(T_{\underline{t}}^* \pi)(A) = \int P(\underline{t}, w, A) \, \pi(dw)$$

for all $(\pi, A) \in \Delta([b, w^m]) \times \mathcal{B}([b, w^m])$. Condition M as stated above implies that $T_{\underline{t}}^*$ is a contraction of modulus $1 - e^{-\lambda \underline{t}}$ on the space $\Delta([b, w^m])$ endowed with the total variation norm (Stokey and Lucas (1989, Lemma 11.11)). Because this is a complete metric space, it follows that $T_{\underline{t}}^*$ has a unique invariant measure $\pi_{\underline{t}}^*$, that is, a unique fixed point $\pi_{\underline{t}}^* = T_{\underline{t}}^* \pi_{\underline{t}}^*$. Using the fact that Condition M is stronger than Doeblin's condition (Condition D in Stokey and Lucas (1989, Chapter 11, Section 4)), one can deduce from the uniqueness of the invariant measure $\pi_{\underline{t}}^*$ that there exists a unique ergodic set, and that for each $\pi_0 \in \Delta([b, w^m])$, the following holds:

$$\lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^K T_{\underline{t}}^{*k} \pi_0 = \pi_{\underline{t}}^* \tag{F.3}$$

39

in the total variation norm (Stokey and Lucas (1989, Theorem 11.9)).

Consider now the behavior of the process $\{w_t\}_{t\geq 0}$ evaluated at dates $k\underline{t}$, $k \in \mathbb{N}$. The induced discrete-time process $\{w_{k\underline{t}}\}_{k=0}^{\infty}$ is Markov, with a transition function $P(\underline{t}, \cdot, \cdot)$ that satisfies the Feller property. Since (F.3) holds in the total variation norm, it is a fortiori true in the topology of weak convergence. The strong law of large numbers for Markov processes (Stokey and Lucas (1989, Theorem 14.7)) then implies that, for any continuous function $g : [b, w^m] \to \mathbb{R}$,

$$\lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} g(w_{k\underline{t}}) = \int g(w)\, \pi_{\underline{t}}^*(dw) \tag{F.4}$$

$\mathbb{P}$–almost surely. It follows that for any $\varepsilon > 0$, $w_{k\underline{t}} \in [b, b+\varepsilon)$ infinitely often, $\mathbb{P}$–almost surely, that is $\mathbb{P}[\limsup\{w_{k\underline{t}} \in [b, b+\varepsilon)\}] = 1$. To see why, note that $\{w_t\}_{t\geq 0}$ is such that, for each $k \in \mathbb{N}$ and $w \in (b, w^m]$, there is a strictly positive probability that $w_{(k+1)\underline{t}} < w$ given that $w_{k\underline{t}} \geq w$. Since $\pi_{\underline{t}}^*$ is invariant under $T_{\underline{t}}^*$, this implies that the lower bound of the support of $\pi_{\underline{t}}^*$ is b, and thus that $\pi_{\underline{t}}^*([b, b+\varepsilon)) > 0$ for all $\varepsilon > 0$. Fix $\varepsilon > 0$ and suppose that g in (F.4) is strictly positive on $[b, b+\varepsilon)$ and equal to 0 elsewhere. Then $\int g(w)\, \pi_{\underline{t}}^*(dw) > 0$. If in some state there exists some $k_0 \in \mathbb{N}$ such that $w_{k\underline{t}} \notin [b, b+\varepsilon)$ for all $k \in \mathbb{N}$ such that $k \geq k_0$, the limit on the left-hand side of (F.4) is 0. Since $\int g(w)\, \pi_{\underline{t}}^*(dw) > 0$, this can only happen on a set of states of measure 0, and the claim follows.

For the remainder of the proof, fix $\varepsilon \in (0, b/2)$, so that $b + 2\varepsilon < w^m$ as $w^m \geq 2b$ by Lemmas C.4 and C.5, and let $\inf\{\emptyset\} = \infty$ and $\infty - \infty = 0$. Define an increasing random sequence in $\mathbb{N} \cup \{\infty\}$ inductively by $K_0 = \inf\{\kappa \in \mathbb{N} \,|\, w_{\kappa\underline{t}} \in [b, b+\varepsilon)\}$ and $K_{k+1} = \inf\{\kappa \in \mathbb{N} \,|\, \kappa > K_k \text{ and } w_{\kappa\underline{t}} \in [b, b+\varepsilon)\}$ for all $k \in \mathbb{N}$. Note that $\{K_k \leq K\} \in \mathcal{F}_{K\underline{t}}^N \subset \mathcal{F}$, so that $\{K_k < \infty\} = \bigcup_{K=0}^{\infty}\{K_k \leq K\} \in \mathcal{F}$ for all $(k, K) \in \mathbb{N}^2$. In particular, $\bigcap_{k=0}^{\infty}\{K_k < \infty\} = \limsup\{w_{k\underline{t}} \in [b, b+\varepsilon)\}$. Using the fact that $b + 2\varepsilon < w^m$ and proceeding as for (F.2), it is straightforward to check that the minimum amount of time that it takes $\{w_t\}_{t\geq 0}$ to transit from $b + \varepsilon$ to $b + 2\varepsilon$ is

$$\underline{t}' = \frac{1}{\rho} \ln\left(\frac{\rho(1 + 2\varepsilon/b) + \lambda}{\rho(1 + \varepsilon/b) + \lambda}\right) < \underline{t}. \tag{F.5}$$

As a result of this, conditional on the event $\bigcap_{k=0}^{\infty}\{K_k < \infty\}$, for each $k \in \mathbb{N}$ it takes $\{w_t\}_{t\geq 0}$ strictly more than $\underline{t}'$ to exit the interval $[b, b+2\varepsilon]$ starting from date $K_k\underline{t}$. For each $k \in \mathbb{N}$, define $\tau_k = \inf\{t > K_k\underline{t} \,|\, \Delta N_t = 1\} - K_k\underline{t}$. It follows from the properties of the Poisson process that, conditional on $\bigcap_{k=0}^{\infty}\{K_k < \infty\}$, $\{\tau_k\}_{k\in\mathbb{N}}$ is a sequence of independent and exponentially distributed random variables with parameter λ. In particular, the events $\{\tau_k \leq \underline{t}'\}$ have probability $1 - e^{-\lambda\underline{t}'}$ conditional on $\bigcap_{k=0}^{\infty}\{K_k < \infty\}$, so that $\sum_{k=0}^{\infty}\mathbb{P}[\tau_k \leq \underline{t}' \,|\, \bigcap_{k=0}^{\infty}\{K_k < \infty\}] = \infty$, and they are independent conditional on $\bigcap_{k=0}^{\infty}\{K_k < \infty\}$. Therefore $\mathbb{P}[\limsup\{\tau_k \leq \underline{t}'\} \,|\, \bigcap_{k=0}^{\infty}\{K_k < \infty\}] = 1$ by the Borel–Cantelli lemma. Since the event $\bigcap_{k=0}^{\infty}\{K_k < \infty\}$ itself has $\mathbb{P}$–probability 1, it follows that $\mathbb{P}[\bigcap_{k=0}^{\infty}\{K_k < \infty\} \cap \limsup\{\tau_k \leq \underline{t}'\}] = 1$. Now fix some state in $\bigcap_{k=0}^{\infty}\{K_k < \infty\} \cap \limsup\{\tau_k \leq \underline{t}'\}$. In this state, the process $\{w_t\}_{t\geq 0}$ visits $[b, b+\varepsilon)$ infinitely often at dates $\{K_k\underline{t}\}_{k\in\mathbb{N}}$, and the process N jumps infinitely often at dates $K_{k_i}\underline{t} + \tau_{k_i}$ in the time intervals $[K_{k_i}\underline{t}, K_{k_i}\underline{t} + \underline{t}']$, where $\{k_i\}_{i\in\mathbb{N}}$ is some strictly increasing sequence in $\mathbb{N}$. Since $w_{K_{k_i}\underline{t}+\tau_{k_i}} \in [b, b+2\varepsilon)$ for all $i \in \mathbb{N}$, and since $\varepsilon < b/2$, each jump of N at date $K_{k_i}\underline{t} + \tau_{k_i}$ induces downsizing, with a downsizing factor given by $x_{K_{k_i}\underline{t}+\tau_{k_i}} = (w_{K_{k_i}\underline{t}+\tau_{k_i}} - b)/b < 2\varepsilon/b$. Thus, in the state under consideration, one has, for each $t \geq 0$,

$$X_t \leq \prod_{\{i\in\mathbb{N} | K_{k_i}\underline{t}+\tau_{k_i}\leq t\}} x_{K_{k_i}\underline{t}+\tau_{k_i}} \leq \left(\frac{2\varepsilon}{b}\right)^{\text{Card}\{i\in\mathbb{N}|K_{k_i}\underline{t}+\tau_{k_i}\leq t\}},$$

which goes to 0 as t goes to ∞ since $2\varepsilon/b < 1$ and $\lim_{t\to\infty}\text{Card}\{i \in \mathbb{N} \,|\, K_{k_i}\underline{t} + \tau_{k_i} \leq t\} = \infty$ by construction. Because $W_{t-} = X_t w_t$ is positive and bounded above by $X_t w^m$, it also goes to 0 as t goes to ∞. Hence the result since $\mathbb{P}[\bigcap_{k=0}^{\infty}\{K_k < \infty\} \cap \limsup\{\tau_k \leq \underline{t}'\}] = 1$. ∎

<div align="center">REFERENCES</div>

[1] Abreu, D., P. Milgrom, and D. Pearce (1991): "Information and Timing in Repeated Partnerships," *Econometrica*, 59, 1713–1733.

[2] Akerlof, G.A., and L.F. Katz (1989): "Workers' Trust Funds and the Logic of Wage Profiles," *Quarterly Journal of Economics*, 104, 525–536.

[3] Biais, B., T. Mariotti, G. Plantin, and J.-C. Rochet (2004): "Dynamic Security Design," CEPR Discussion Paper No. 4753.

[4] Biais, B., T. Mariotti, G. Plantin, and J.-C. Rochet (2007): "Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications," *Review of Economic Studies*, 74, 345–390.

[5] Brémaud, P. (1981): *Point Processes and Queues: Martingale Dynamics*, Berlin, Heidelberg, New York, Springer-Verlag.

[6] Clementi, G.L., and H. Hopenhayn (2006): "A Theory of Financing Constraints and Firm Dynamics," *Quarterly Journal of Economics*, 121, 229–265.

[7] Dellacherie, C., and P.-A. Meyer (1978): *Probabilities and Potential*, Volume A, Amsterdam, North-Holland.

[8] Dellacherie, C., and P.-A. Meyer (1982): *Probabilities and Potential*, Volume B, Amsterdam, North-Holland.

[9] DeMarzo, P.M., and M.J. Fishman (2003): "Optimal Long-Term Financial Contracting with Privately Observed Cash Flows," Mimeo, Stanford University.

[10] DeMarzo, P.M., and Y. Sannikov (2006): "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model," *Journal of Finance*, 61, 2681–2724.

[11] Duffie, D., and D. Lando (2001): "Term Structures of Credit Spreads with Incomplete Accounting Information," *Econometrica*, 69, 633–664.

[12] Elliott, M.R., Y. Wang, R.A. Lowe, and P.R. Kleindorfer (2004): "Environmental Justice: Frequency and Severity of US Chemical Industry Accidents and the Socioeconomic Status of Surrounding Communities," *Journal of Epidemiology and Community Health*, 58, 24–30.

[13] Gordon, R.P.E., R.H. Flin, K. Means, and M.T. Fleming (1996): "Assessing the Human Factors Causes of Acccidents in the Offshore Oil Industry," in M. Ognedal, R.J. Edwardes and J.P. Visser (Eds.), *Proceedings of the 3rd International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production*, pp. 635–644, Richardson, Society of Petroleum Engineers.

[14] Hollnagel, E. (2002): "Understanding Accidents–From Root Causes to Performance Variability," in J.J. Persensky, B.P. Hallbert and H.S. Blackman (Eds.), *New Century, New Trends*: *Proceedings of the 7th IEEE Conference on Human Factors and Power Plants*, pp. 1–6, New York, Intitute of Electrical and Electronics Engineers.

[15] Holmström, B., and P. Milgrom (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55, 303–328.

[16] Holmström, B., and J. Tirole (1997): "Financial Intermediation, Loanable Funds, and the Real Sector," *Quarterly Journal of Economics*, 112, 663–691.

[17] Jost, P.-J. (1996): "Limited Liability and the Requirement to Purchase Insurance," *International Review of Law and Economics*, 16, 259–276.

[18] Kalesnik, V. (2005): "Continuous Time Partnerships with Discrete Events," Mimeo, University of California, Los Angeles.

[19] Katzman, M.T. (1988): "Pollution Liability Insurance and Catastrophic Environmental Risk," *The Journal of Risk and Insurance*, 55, 75–100.

[20] Leplat, J., and J. Rasmussen (1984): "Analysis of Human Errors in Industrial Incidents and Accidents for Improvements of Work Safety," *Accident Analysis and Prevention*, 16, 77–88.

[21] Myerson, R.B. (2007): "Leadership, Trust, and Power," Mimeo, University of Chicago.

[22] Polborn, M.K. (1998): "Mandatory Insurance and the Judgment-Proof Problem," *International Review of Law and Economics*, 18, 141–146.

[23] Sannikov, Y. (2003): "A Continuous-Time Version of the Principal-Agent Problem," Mimeo, Stanford University.

[24] Sannikov, Y. (2005): "Agency problems, Screening and Increasing Credit Lines," Mimeo, University of California, Berkeley.

[25] Sannikov, Y., and A. Skrzypacz (2006): "The Role of Information in Repeated Games with Frequent Actions," Mimeo, University of California, Berkeley.

[26] Shapiro, C., and J.E. Stiglitz (1984): "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74, 433–444.

[27] Shavell, S. (1984): "Liability for Harm versus Regulation of Safety," *Journal of Legal Studies*, 13, 357–374.

[28] Shavell, S. (1986): "The Judgment Proof Problem," *International Review of Law and Economics*, 6, 45–58.

[29] Shavell, S. (2000): "On the Social Function and the Regulation of Liability Insurance," *Geneva Papers on Risk and Insurance*, 25, 166–179.

[30] Stokey, N.L., and R.E. Lucas, Jr., with E.C. Prescott (1989): *Recursive Methods in Economic Dynamics*, Cambridge, Harvard University Press.

[31] Thomas, J. and T. Worrall (1990): "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem," *Journal of Economic Theory*, 51, 367–390.

[32] Zhou, C. (2001): "The Term Structure of Credit Spreads with Jump Risks," *Journal of Banking and Finance*, 25, 2015–2040.