Dynamic Contracting:
Accidents Lead to Nonlinear Contracts

Agostino Capponi
Department of Industrial Engineering and Operations Research
Columbia University
New York NY 10027
USA
acappon1@jhu.edu

Christoph Frei
Department of Mathematical and Statistical Sciences
University of Alberta
Edmonton AB T6G 2G1
Canada
cfrei@ualberta.ca

This version: August 10, 2015

(SIAM Journal on Financial Mathematics 6 (2015), 959–983)

Abstract

We consider a dynamic multitask principal-agent model in which the agent allocates his resources on two tasks of different types: effort and accident prevention. We explicitly characterize the optimal contract as well as optimal effort and prevention actions applied by the agent. In contrast to the linear incentive scheme for effort, accident prevention leads to a log-linear punishment scheme if the agent is risk averse, becoming linear only if the agent is risk neutral. Both the sublinearity of the contract and the allocation of resources on the two tasks crucially depend on the risk aversion of the agent. Accident prevention ties up some of the agent's capacity and induces him to substitute resources away from effort to prevention.

1 Introduction

We extend the classical Holmström and Milgrom (1987) framework by introducing jumps in the output process, whose frequency can be controlled by the agent. This leads to a fundamentally different form for the optimal contract which becomes log-linear in the accident component, but remarkably the optimal contract can still be found explicitly despite its nonlinear form. The sublinearity of the contract as well as the optimal diversion of agent's resources from the continuous to the jump component of the output crucially depend on the risk aversion level of the agent. Our model is broadly applicable for analyzing contracting environments in which the output process can experience jumps of unknown size in addition to continuous innovations corresponding to the earnings process of any typical corporation. Hence, it allows analyzing the optimal incentive provisions of a principal to an agent who can execute two tasks, each having a *fundamentally different* impact on the output process. For example, jumps can be used to model accidents whose negative consequences usually have a longer-term impact on the output process. Indeed, this has become highly relevant in recent years given the growing interest in including sustainability performance as a part of executive compensation; see Cordeiro and Sarkis (2008). This stems from the fact that major environmental accidents are subject to severe moral hazard problems.[1]

Most of the literature on continuous-time contracting has focused on agents executing a single task. Starting with the pioneering paper by Holmström and Milgrom (1987), many extensions have been proposed; see for instance Schättler and Sung (1993), Schättler and Sung (1997), Hellwig and Schmidt (2002), Cvitanić et al. (2009), and Ju and Wan (2012). Capponi et al. (2013) allow for the agent to alter the fundamental value of the firm besides applying effort. Sannikov (2008) solves the principal-agent problem allowing for the agent to be paid continuously. He (2009) extends his framework to the case when the agent can control the size of the company.

Optimal contracting with multitasking has been initiated by Holmström and Milgrom (1991), who consider a single-period model and analyze how optimal incentives are related to measurability of the output. Follow-up works include Bond and Gomes (2009), who consider a static framework with a risk-neutral agent incurring linear effort and with no specific preference across tasks, and Manso (2011) who studies the tradeoff between two tasks interpreted as exploration and exploitation. A continuous-time multitasking model for the optimal design of mortgage backed securities is considered in Hartman-Glaser et al. (2012), where a mortgage underwriter must evaluate multiple defaultable loans, and may shirk in selecting them.

Our framework allows for jump risk to negatively affect revenue generation in a dynamic setting. We model operating profits using a continuous process, chosen to be a Brownian motion, and use a compound Poisson process to model the effect of accidents. This leads to a jump-diffusion driven outcome process. The agent can simultaneously apply two hidden actions, referred to as (1) *effort* to increase the instantaneous growth rate of the outcome process, and (2) *prevention* to decrease the frequency of accidents. Both actions are costly to the agent. Unlike the actions, the size of the accidents is publicly observable, verifiable, and hence contractible. Both principal and agent are risk averse with exponential utility.

[1] As reported by Palast (1994) in the Chicago Tribune, one of the main reasons for the Exxon Valdez oil spill disaster was that a sophisticated radar was turned off. The tanker's radar was left broken and disabled for more than a year before the disaster, and Exxon management knew it. It was, in their view, too expensive to fix it.

To the best of our knowledge, this paper is the first that explicitly characterizes the optimal contract in a dynamic framework with both effort and accident prevention. Crucially, the optimal contract is *sublinear* in the accident component. The principal charges a lower percentage penalty for big accidents than he does for small accidents. This is consistent with existing literature on managerial compensation, which has shown that payment schemes are often convex and in the form of call options, see for instance Jensen and Murphy (1990) and Murphy (1999). The typical explanation for the convexity is that this gives the manager more incentives to behave in a way that boosts the company's stock price. Our analysis highlights an additional reason: a small decline is penalized relatively (in percentage of the decline) more than a big decline. Indeed, the accident component in the contract serves the principal to give incentives to the agent for avoiding accidents. Because of the agent's risk aversion, already a small portion of large accidents gives the agent enough incentives to try to reduce their frequency while for small accidents a higher portion is required to incentivize him enough. Because the principal shifts a portion of accident costs to the agent, he has to compensate him by increasing his fixed-wage compensation. The latter exceeds the expected accident costs incurred by the agent throughout the life of the contract. The contract only becomes linear in the accident component if the agent is risk neutral, given that in this case the agent behaves in the same way as if he owned the company.

Our paper is related to Biais et al. (2010). Their focus is on analyzing the optimal incentives provided by an insurance company (principal) to a risk-neutral manager. The latter can exert costly effort to reduce the onset of accidents which can generate large losses. An important difference with their study is the crucial role played by task interaction in our framework in driving agent's incentives and determining *substitution effects*. When effort and prevention are substitutes and task interaction is strong, the principal finds it optimal to induce higher level of prevention and reduce effort incentives, regardless of whether the principal takes preventive measures to cover against part of the losses. All this reflects the fundamentally different impact of the two tasks on the final output. While applying little effort only results in a small instantaneous growth rate without affecting volatility risk of the output, applying little prevention generates high risk due to the increased accident costs. Pagès and Possamaï (2014) consider a contracting framework similar to Biais et al. (2010). They provide a comprehensive mathematical analysis of optimal contracting between competitive investors and an impatient bank monitoring a pool of long-term loans subject to default risk, i.e., the bank can reduce the default intensities through its monitoring activities.

Our study is also related to Sung (1997), who develops a continuous-time framework where the manager performs multiple tasks to reduce the expected number of accidents. Differently from ours, his model only allows for a finite number of categories to which accidents belong, and the cost of each category is assumed to be known. This leads to the optimal contract being linear in the account balance of accidents. Other studies have considered contractual frameworks where Poisson jumps affect the payoff process. DeMarzo et al. (2012) consider the case of unobservable productivity shocks, while Hoffmann and Pfeil (2010) allow for a persistent and publicly observable lucky shock to occur. A crucial difference with these studies, where jumps occur to the drift rate and are assumed to be positive and exogenously given, is that in our model jumps of random size negatively affect the total output process and not the drift rate. Moreover, the agent can control the accident frequency in our model.

We also make other contributions to the literature. We consider the corresponding risk-sharing model, where the principal has the entire bargaining power and decides both the contract and the actions applied by the agent. As opposed to the weak formulation arising in the moral hazard case, optimal contracting under risk sharing is usually modeled via a strong formulation of the principal-agent problem. Most of the previous studies on continuous-time contracting under a risk-sharing framework, see for instance Cadenillas et al. (2007) and references therein, only deal with continuous processes driven by Brownian motion. To the best of our knowledge, our study is the first to establish the strong formulation when the outcome process also consists of a discontinuous component, and achieves it via a time change of the Poisson process capturing the occurrence of accidents.

The rest of the paper is organized as follows. Section 2 sets up the principal-agent problem under moral hazard, which we solve in Section 3. Section 4 considers the risk-sharing formulation of our contracting problem. Section 5 illustrates effort substitution effects in the case of full accident exposure and partial preventive coverage in three contractual environments arising as specializations of our moral hazard framework. Section 6 concludes, and the Appendix contains all proofs.

2 Problem formulation

We define the dynamics of the output process under no effort and prevention action in Section 2.1. We model the hidden actions of the agent and the space of contracts considered in Section 2.2. We formulate the optimization problem of the agent in Section 2.3, and the corresponding problem for the principal in Section 2.4.

2.1 Output process dynamics

The outcome process consists of a continuous component modeling the outcome from daily activities, and of a jump component capturing the occurrence of accidents.

Continuous Component. Under no managerial effort, the outcome process has the dynamics

$$dx_t = c_t \, dt + \epsilon_t \, dB_t, \tag{2.1}$$

where B_t is a Brownian motion under the standard reference probability measure $\mathbb{P}$ defined on a measurable space $(\Omega, \mathcal{F})$. Here c_t and ϵ_t are, possibly time varying, coefficients modeling the exogenous drift rate and outcome volatility.[2]

[2] For the following computations to be valid, we assume that the function ϵ_t is positive and bounded away from zero and infinity, i.e., there exist constants $K \geq k > 0$ such that $K \geq \epsilon_t \geq k$. We also assume that c_t is integrable.

Jump Component. Accidents can occur and negatively affect the output process. Each accident is modeled as the jump of a compound Poisson process

$$J_t = \sum_{i=1}^{N_t} Y_i,$$

defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where (Y_i) is a sequence of bounded nonnegative i.i.d. random variables with cumulative distribution function F. The random variable Y_i quantifies the cost of the i-th accident. Under the reference probability measure $\mathbb{P}$, N_t is a Poisson process with some fixed intensity $\lambda_0 > 0$. Our results do not depend on the choice of λ_0 because in the optimization problems only the intensity chosen by the agent will be relevant.

For future purposes, we set $m = \int_0^\infty y\, dF(y) = \mathbb{E}[Y_i]$ as the average cost per accident and define the outcome process X_t as

$$X_t = x_t - J_t, \tag{2.2}$$

i.e., the cumulative output process inclusive of the impact of accidents occurred up to time t. We assume that B_t and J_t are independent and denote by $(\mathcal{F}_t)_{0\le t\le T}$ the augmented filtration generated by the processes, up to time $T > 0$.

2.2 Hidden actions and contract space

Through costly effort the agent can control the growth rate of the continuous component of the output. Moreover, by applying costly prevention, he can reduce the intensity at which accidents occur.

Effort. We model the effort applied by the agent, following a standard change of measure argument on the Brownian motion B_t. More precisely, an admissible effort action is a bounded $(\mathcal{F}_t)$-predictable process u_t, which induces a change of probability measure by

$$M_t^u = \exp\left(\int_0^t \frac{u_s}{\epsilon_s}\, dB_s - \frac{1}{2}\int_0^t \left(\frac{u_s}{\epsilon_s}\right)^2 ds\right) \quad \text{and} \quad \frac{d\mathbb{P}^u}{d\mathbb{P}} = M_T^u, \tag{2.3}$$

where M_t^u is an $(\mathcal{F}_t)$-adapted martingale. Then, by the Girsanov theorem, the process $B_t^u = B_t - \int_0^t \frac{u_s}{\epsilon_s}\, ds$ is a Brownian motion under the probability measure $\mathbb{P}^u$. Consequently, we can rewrite the dynamics (2.1) under $\mathbb{P}^u$ as

$$dx_t = (c_t + u_t)\, dt + \epsilon_t\, dB_t^u, \tag{2.4}$$

Above, notice that the agent chooses his effort u_t based on the output process x_t which is observable to the principal. However, although u_t is $(\mathcal{F}_t)$-adapted, the principal only observes x_t and does not know u_t, and hence he does not know the value of B^u, either. Therefore, this approach allows modeling the effort action u_t as a hidden action. From (2.4), we can see that the effect of the costly effort action of the agent is to increase the instantaneous growth rate of the output by u_t.

Accident Prevention. We model the costly accident prevention action of the agent by letting him choose the intensity λ_t at which jumps occur. Such a process λ_t is admissible if it is $(\mathcal{F}_t)$-predictable and bounded away from zero and infinity. By reducing λ_t, the agent can decrease the frequency of accidents, but will face additional costs which we will later formalize. Similarly to (2.3), we model the choice of the intensity λ_t by a change of measure. We define a probability measure $\mathbb{P}^{u,\lambda}$ by

$$\frac{d\mathbb{P}^{u,\lambda}}{d\mathbb{P}} = \frac{d\mathbb{P}^{u}}{d\mathbb{P}} \exp\left(\int_0^T \log(\lambda_t/\lambda_0)\, dN_t - \int_0^T (\lambda_t - \lambda_0)\, dt \right)$$

so that under $\mathbb{P}^{u,\lambda}$, N_t is a Poisson process with instantaneous intensity λ_t and B_t is a drifted Brownian motion with instantaneous drift u_t.

A concrete application of our model, which outlines the importance of including the accident component when modeling manager's compensation, is sustainability. As reported by Novacovici (2013) in the Huffington Post, large companies are starting to link executive compensation to environmental sustainability. Intel started in 2008 to link 3% of all its employees' annual bonuses to environmental sustainability metrics and goals. Xcel Energy ties a third of its CEO's annual bonus to energy and greenhouse gas emission goals. Alcoa includes sustainability performance in its executive bonus plan; starting in 2010, it began linking 20% of the plan to non-financial metrics that include carbon dioxide reduction goals. Our contracting model captures these sustainability performance metrics, because incentives are provided to the agent not only to increase the daily operating profits through exertion of effort, but also to limit environmental damages through prevention. Another relevant application of our contracting framework is coinsurance. The principal (firm's owner) pays an insurer, or a guaranteed fund, a premium in exchange of a coinsurance contract offering coverage for a fraction of losses. For each realized loss, the firm uses its cash reserves to cover the loss fraction which is not covered by the insurer. Since part of the accident risk is borne by the firm, the principal still needs to provide incentives to the agent to apply prevention against accidents. The presence of coinsurance is also discussed in the context of optimal insurance under moral hazard by Winter (2000).

The contract space. At time 0, the principal (shareholders) offers a compensation C_τ to the agent which will be paid at a termination time τ optimally determined as a random variable by the principal. Moreover, the principal incurs termination costs which will depend on the considered contractual environment. The setting with a termination time determined by the principal allows us to consider different contractual possibilities detailed in Sections 5.2–5.4. Technically, we need some integrability conditions on C_τ. We will make them precise in Remark 3.2 in terms of a representation of C_τ.

2.3 The agent

The agent obtains utility from consuming C_τ. The payment C_τ is an $\mathcal{F}_\tau$-measurable random variable, to be interpreted as a payment in cash, the amount of which depends on random outcomes by time τ. Moreover, he incurs disutility for exerting costly effort and accident prevention up to time τ. We assume that this is of the form

$$E_\tau = \int_0^\tau e(u_t, \lambda_t)\, dt,$$

where $e : [0, \infty) \times (0, \infty) \to [0, \infty)$ is a twice continuously differentiable and strictly convex function with $\frac{\partial e}{\partial u} > 0$, $\frac{\partial e}{\partial \lambda} < 0$, $\lim_{u \nearrow \infty} \big(e(u, \lambda) - u\big) = \infty$ for all $\lambda > 0$, $\lim_{\lambda \nearrow \infty} \frac{\partial e}{\partial \lambda}(u, \lambda) = 0$ and $\lim_{\lambda \searrow 0} e(u, \lambda) = \infty$ for all $u \geq 0$.

Remark 2.1. **1)** *The cost function e captures the interaction between the two tasks of applying effort and reducing accident risk. Although our framework is general and can accommodate both complementary and substitute tasks, we will primarily focus on the case when they are substitutes. This is because our framework is designed to capture contractual scenarios related to, for example, sustainability performance. Here, executives are incentivized to dedicate resources toward set sustainability goals which may significantly reduce the level of effort devoted to the generation of operating profits. In other words, the marginal cost of the effort action is increasing in the amount of prevention applied by the agent, and vice versa. Since we measure the cost of accident prevention by $\frac{1}{\lambda}$ (the higher the prevention applied by the agent, the smaller the accident frequency λ is), the two tasks are substitutes if $\frac{\partial^2 e}{\partial u \partial \lambda} < 0$. A detailed analysis is reported in Section 5.*

2) *Sung (1997) restricts the intensity of accidents to be within a bounded interval. Although we do not impose any restriction and let the intensity vary on the whole interval $(0, \infty)$, our analysis carries through upon adjusting the incentive compatibility conditions to guarantee that the intensity stays within the desired interval. In a similar way, one could choose a reference level of accident prevention so that the agent doing nothing would correspond to the given reference level and any accident prevention would lead to an increase from the reference level.*

The objective of the agent is to maximize, over the choice of u and λ, the expected utility

$$V_1(C) := \sup_{u,\lambda} V_1(C, u, \lambda) := \sup_{u,\lambda} \mathbb{E}^{u,\lambda} \left[U_1(C_\tau - E_\tau) \right],$$

where $\mathbb{E}^{u,\lambda}$ is the unconditional expectation (we omit the initial information $\mathcal{F}_0$ from the expectation operator) corresponding to the effort and prevention strategy (u, λ) applied by the agent, and C_τ is the payment from the principal to the agent at time τ chosen by the principal. We choose

$$U_1(x) = \frac{1}{\gamma_1}(1 - \mathrm{e}^{-\gamma_1 x}), \tag{2.5}$$

i.e., the agent is risk averse with constant of absolute risk aversion $\gamma_1 > 0$. We say that a contract C is implementable if there exists u^f, λ^f such that

$$V_1(C) = V_1(C, u^f, \lambda^f).$$

2.4 The principal

Let R_0 be the reservation utility of the agent. The principal can terminate the contract with the agent prematurely at time $\tau \leq T$ and will then incur costs $d(\tau)$. It follows that the principal's optimization problem is to maximize his expected utility

$$V_2 := \sup_{C,\tau} \mathbb{E}^{u^*,\lambda^*}[U_2(X_\tau - C_\tau - d(\tau))] \tag{2.6}$$

over the class of implementable contracts C and exercise times τ subject to the following incentive compatibility constraint

$$V_1(C) = V_1(C, u^*, \lambda^*)$$

and individual rationality constraint

$$V_1(C) \geq R_0. \tag{2.7}$$

The principal is also risk averse with exponential utility function

$$U_2(x) = \frac{1}{\gamma_2}(1 - \mathrm{e}^{-\gamma_2 x}), \tag{2.8}$$

where $\gamma_2 > 0$ is the constant of absolute risk aversion of the principal.

Note that for fixed γ_i, choosing instead $-\mathrm{e}^{-\gamma_i x}$ as utility function leads to an equivalent optimization problem. However, our choice of U_i has the advantage that in the limit $\gamma_i \searrow 0$, which we will later study for the optimization problem, we obtain a linear utility $\lim_{\gamma_i \searrow 0} U_i(x) = x$.

3 Optimal contracting

We solve for the optimal contract in Section 3.1. We analyze the impact of vanishing risk aversion on the optimal contract in Section 3.2. All proofs are relegated to the Appendix.

3.1 Optimal contract and actions

We will derive the optimal contract in three steps. In the first step, stated in Lemma 3.1, we give the general form of the optimal contract. This step is based on the fact that the principal will not pay more than what the agent has as reservation utility R_0, which implies

$$R_0 = \mathbb{E}^{u,\lambda}[U_1(C_\tau - E_\tau)]$$

for the optimal contract and optimal actions. In the second step, Proposition 3.3 specifies which actions are incentive compatible. This uses the fact that for all possible actions (optimal or not), the agent's expected utility will be less than or equal to R_0. This allows us to prove Theorem 3.4, which completely characterizes the optimal contract. In this last step, we combine the previous results along with the consideration of the principal's optimization problem.

Lemma 3.1. *Given the agent's reservation utility R_0, the optimal contract is of the form*

$$C_\tau = \underbrace{\int_0^\tau \alpha_t \, dx_t}_{\text{continuous output component}} \underbrace{- \frac{1}{\gamma_1} \sum_{0<t\le\tau} \log(1 + \gamma_1 \beta_t \Delta J_t)}_{\text{accident penalty}} \tag{3.1}$$
$$\underbrace{- \frac{1}{\gamma_1} \log(1 - \gamma_1 R_0) + \int_0^\tau \left(\frac{\gamma_1}{2} \alpha_t^2 \epsilon_t^2 + \beta_t \lambda_t m + e(u_t, \lambda_t) - \alpha_t c_t - \alpha_t u_t \right) dt}_{\text{fixed wage}},$$

where α and β are predictable processes.

The above contract can be implemented using the following instruments:

- Cash amount given to the agent so as to meet his rationality constraint.
- Pay-for-performance compensation rewarding the agent for the daily profit generating activities.
- From the above contractual specification, we see that the principal charges the agent a portion of corporate liability. This may consist of disciplinary action in the form of fines, loss of personal assets or reputation damage, imposed by the company to its employees for negligent conduct causing, for example, environmental liability. However, the company compensates the agent for bearing accident risk. Using the inequality $\log(1 + x) \le x$, and setting $x = \gamma_1 \beta_t \Delta J_t$, we obtain

  $$\mathbb{E}^{u,\lambda} \left[\frac{1}{\gamma_1} \sum_{0<t\le\tau} \log(1 + \gamma_1 \beta_t \Delta J_t) \right] \le \mathbb{E}^{u,\lambda} \left[\int_0^\tau \beta_t \lambda_t m \, dt \right].$$

  Using (3.1), this indicates that the portion of the fixed wage compensating the agent for bearing accident risk exceeds the expected accident costs charged by the principal throughout the life of the contract.

Remark 3.2. *To derive the representation in Lemma 3.1, we did not need any specific integrability conditions on C_τ. This is because the agent's rationality constraint* (2.7) *automatically imposes an integrability condition on implementable contracts. Indeed, we have $-\infty < \mathbb{E}^{u,\lambda}[U_1(C_\tau - E_\tau)] < \infty$ for an admissible contract C_τ and the agent's actions for C_τ, where the former inequality follows from* (2.7) *while the latter inequality is due to the form* (2.5) *of U_1. To continue our calculations, we* *impose on the processes α_t and β_t in* (3.1) *that α_t is bounded and β_t is nonnegative and bounded. These conditions represent a restriction on the class of admissible contracts. For the following proofs, it would also be possible to choose the milder condition that $\int \alpha_t \, dB_t$ and $\int \beta_t(dJ_t - m\,dt)$ are BMO-martingales under $\mathbb{P}$. However, we choose the boundedness condition on α_t and β_t because this is in line with the boundedness conditions for u_t and λ_t so that allowed contracts induce the agent to perform allowed actions. This will be a consequence of Proposition 3.3 below.*

Lemma 3.1 also shows that the presence of accidents introduces a *nonlinearity* in the optimal contract. While the optimal compensation is still linear in the continuous component of the output with pay-for-performance sensitivity α_t, it is nonlinear in the accident component. From the term $\log(1 + \gamma_1 \beta_t \Delta J_t)$, we deduce that due to the concavity of the logarithmic function, the punishment of the principal is larger for small accidents relative to large accidents. As accidents become more costly, the additional punishment applied by the principal becomes smaller. This is in line with empirical evidence on managerial compensation using call options incentive schemes. If we think of the outcome process as the equity value of a firm (which holds approximately true for low levered firms), we obtain the sublinearity property from the convexity of the call option payoff. At a given level of stock price, a small decrease of the stock price translates to a larger relative (in percentage of the stock price decrease) change of the call option value compared to the one resulting from a large decrease of the stock price.

The log-linearity property of the optimal contract may be explained in terms of the agent's risk aversion. While a small portion of the large accidents already gives the risk-averse agent enough incentives to try to reduce their frequency, a higher portion of the small accidents is required to give him enough incentives. Being the principal aware of that, he offers a *sublinear* punishment scheme to the agent.

Since it is the principal who chooses the contract, he determines α and β to maximize his expected utility. Using the agent's optimization problem, we first characterize the actions u and λ which are compatible with α and β in Proposition 3.3. As standard in the literature, actions are said to be incentive compatible with respect to some contract if they maximize the agent's total expected utility. In other words, incentive compatible actions are the responses of the agent to a contract offered by the principal.

Proposition 3.3. *Actions u and λ are incentive compatible with α and β if and only if*

$$\frac{\partial}{\partial u} e(u_t, \lambda_t) = \alpha_t, \qquad \frac{\partial}{\partial \lambda} e(u_t, \lambda_t) = -\beta_t m, \tag{3.2}$$

and this can be written as $u_t = i_t(\alpha_t, \beta_t)$ and $\lambda_t = j_t(\alpha_t, \beta_t)$ for suitably chosen functions i_t and j_t.

Combining the results given earlier, we can characterize the optimal contract in the main Theorem 3.4.

Theorem 3.4. *The optimal contract is given by*

$$C^\star_{\tau^\star} = -\frac{1}{\gamma_1}\log(1-\gamma_1 R_0) + \int_0^{\tau^\star} \alpha_t^\star \, dx_t - \frac{1}{\gamma_1}\sum_{0<t\le\tau^\star}\log(1+\gamma_1\beta_t^\star \Delta J_t)$$
$$+ \int_0^{\tau^\star}\Big(\frac{\gamma_1}{2}(\alpha_t^*)^2\epsilon_t^2 + \beta_t^\star\lambda_t^\star m + e(u_t^\star,\lambda_t^\star) - \alpha_t^\star c_t - \alpha_t^\star u_t^\star\Big)\,dt, \tag{3.3}$$

where $u_t^\star = i_t(\alpha_t^\star,\beta_t^\star)$, $\lambda_t^\star = j_t(\alpha_t^\star,\beta_t^\star)$, *and* $\alpha_t^\star$ *and* $\beta_t^\star$ *are the maximizers of*

$$p_t(\alpha_t,\beta_t) = -\,e\big(i_t(\alpha_t,\beta_t), j_t(\alpha_t,\beta_t)\big) - \frac{\gamma_1}{2}\alpha_t^2\epsilon_t^2 + c_t + i_t(\alpha_t,\beta_t) - \frac{\gamma_2}{2}\epsilon_t^2(\alpha_t-1)^2$$
$$- \left(\frac{1}{\gamma_2}\int_0^\infty (1+\gamma_1\beta_t y)^{-\gamma_2/\gamma_1} e^{\gamma_2 y}\, dF(y) - \frac{1}{\gamma_2} + \beta_t m\right) j_t(\alpha_t,\beta_t); \tag{3.4}$$

the optimal termination time $\tau^\star$ *is given by*

$$\tau^\star = \operatorname*{arg\,max}_{\tau\in[0,T]}\left(\int_0^\tau p_t(\alpha_t^\star,\beta_t^\star)\,dt - d(\tau)\right). \tag{3.5}$$

Recall that for a given contract (α,β), the principal's certainty equivalent CE is defined by

$$\mathbb{E}^{u=i(\alpha,\beta),\lambda=j(\alpha,\beta)}[U_2(X_\tau - C_\tau - d(\tau))] = U_2\,(\mathrm{CE})\,,$$

i.e., CE is the cash amount that will leave the principal indifferent in terms of expected utility between a guaranteed cash payment and what he receives after paying the agent and terminating the contract. The proof of Theorem 3.4 shows that

$$\mathrm{CE} = \int_0^\tau p_t(\alpha_t,\beta_t)\,dt - d(\tau) + \frac{1}{\gamma_1}\log(1-\gamma_1 R_0).$$

Because the contract is chosen so as to maximize the principal's certainty equivalent, from the above representation we can see that the principal will choose (α_t,β_t) at each time t to optimize the derivative p_t of the certainty equivalent. The latter can be decomposed as follows:

$$p_t(\alpha_t,\beta_t) = \underbrace{-\,e\big(i_t(\alpha_t,\beta_t), j_t(\alpha_t,\beta_t)\big)}_{\text{agent's costs}}\underbrace{-\frac{\gamma_1}{2}\alpha_t^2\epsilon_t^2 + c_t + i_t(\alpha_t,\beta_t) - \frac{\gamma_2}{2}\epsilon_t^2(\alpha_t-1)^2}_{\text{contribution from continuous output component}}$$
$$\underbrace{-\left(\frac{1}{\gamma_2}\int_0^\infty (1+\gamma_1\beta_t y)^{-\gamma_2/\gamma_1} e^{\gamma_2 y}\, dF(y) - \frac{1}{\gamma_2} + \beta_t m\right) j_t(\alpha_t,\beta_t)}_{\text{contribution from expected accident costs}}\,. \tag{3.6}$$

3.2 Vanishing risk aversion

We analyze how the form of the contract is affected by the risk aversion levels of principal and agent. We will see that the nonlinearity of the contract hinges on the agent's risk aversion, but not on that of the principal. The next corollary shows that when the agent is risk neutral, the principal finds it optimal to choose a punishment scheme which is *linear* in the accident component.

Corollary 3.5. *For $\gamma_1 \searrow 0$, the optimal contract converges in probability to*

$$C^{1;\star}_{\tau^\star} = R_0 + \int_0^{\tau^\star} e(u^\star_t, \lambda^\star_t)\, dt + \int_0^{\tau^\star} (dx_t - (c_t + u^\star_t)\, dt) - \int_0^{\tau^\star} (dJ_t - \lambda^\star_t m\, dt),$$

where $u^\star_t$ and $\lambda^\star_t$ are given by

$$\frac{\partial}{\partial u} e(u^\star_t, \lambda^\star_t) = 1, \qquad \frac{\partial}{\partial \lambda} e(u^\star_t, \lambda^\star_t) = -m. \tag{3.7}$$

It is best for the principal if the agent's incentive conditions are identical to the principal's optimality conditions, and hence it is optimal to charge the agent the entire randomness of the output process. Note, however, that the principal is still affected by the output process because due to the agent's risk neutrality, his expected compensation netted of action costs must still meet the agent's rationality constraint. Hence, the principal charges the agent only the randomness (martingale part) of the output process, but keeps the predictable part (drift and martingale compensator).

This behavior for vanishing agent's risk aversion is in line with the well-known fact that in a Holmström and Milgrom (1987) type framework, a risk-neutral agent behaves as if he owns the whole project. Indeed, if the agent is the sole owner of the project and is risk neutral, he will take the first-best actions given by (3.7) to maximize the firm value. With this choice of actions, the agent can pay the first-best firm value to the principal, which is the best value the principal can achieve. In contrast, if the agent is risk averse, he will ask the principal for a risk compensation which crucially depends on the size of the accidents. In this case, the principal will then offer a sublinear contract to minimize the risk compensation related to large accidents.

The next corollary shows that even if the principal were to become risk neutral, he would still offer a sublinear punishment scheme to the risk-averse agent.

Corollary 3.6. *For $\gamma_2 \searrow 0$, the optimal contract converges in probability to the same nonlinear contract* (3.3)*, but where $\alpha^\star_t$ and $\beta^\star_t$ are now the maximizers of*

$$\begin{aligned} p^2_t(\alpha_t, \beta_t) = &- e\big(i_t(\alpha_t, \beta_t), j_t(\alpha_t, \beta_t)\big) - \frac{\gamma_1}{2}\alpha_t^2 \epsilon_t^2 + c_t + i_t(\alpha_t, \beta_t) \\ &- \left(m + \beta_t m - \frac{1}{\gamma_1}\int_0^\infty \log(1 + \gamma_1 \beta_t y)\, dF(y)\right) j_t(\alpha_t, \beta_t). \end{aligned}$$

Moreover, we have $p^2_t(\alpha_t, \beta_t) \geq p_t(\alpha_t, \beta_t)$.

4 Risk sharing

This section studies the case of risk sharing, in which both principal and agent have the same information. This is also called first best and corresponds to the situation when the principal has the entire bargaining power so that he decides both the effort and prevention action of the agent. It can also be interpreted as the optimization of total welfare by a social planner. Before turning to the optimization problem, a technical remark is in order.

Remark 4.1. *There are two ways to construct a Brownian motion with drift. The first way is to start with a standard Brownian motion and then simply add a drift term to its dynamics. The second way is through Girsanov's theorem, i.e. by changing the probability measure so that under the new probability measure the Brownian motion has a drift term. The first method gives rise to the so-called strong formulation while the latter is related to the weak formulation of the principal-agent problem. The weak formulation corresponds to a model with moral hazard, also called second best, where the precise value of the instantaneous drift cannot be inferred from observing the output process. In the previous sections, we have used the weak formulation* (2.4) *because the drift was hidden to the principal. In this section, the action is observable and hence we choose the strong formulation for the dynamic contracting formulation.*

For the accident component, in the previous section we have used the intensity of the Poisson process as the control variable. This was also a weak formulation because a change in the intensity means considering the same process under a different probability measure. Like for the Brownian component, we want to have a strong formulation for the accident component in this section. We achieve this by performing a time change. Recall the standard result that if N_t is a Poisson process with intensity 1, then $\tilde{N}_t = N_{\Lambda(t)}$ is a Poisson process with instantaneous intensity $\lambda(t)$, where $\Lambda(t) = \int_0^t \lambda(s)\, ds$. Therefore, the prevention action in this section affects the accident component via a time change of the underlying Poisson process. The filtration, for which the actions are predictable, is here the augmented filtration generated by the Brownian motion and the time-changed compound Poisson process.

In the risk-sharing problem, a social planner maximizes the joint welfare

$$\mathbb{E}^{\mathbb{P}}[U_1(C_\tau - E_\tau)] + \rho\, \mathbb{E}^{\mathbb{P}}\big[U_2\big(X_\tau - C_\tau - d(\tau)\big)\big], \tag{4.1}$$

where ρ is a given constant which represents the relative risk-sharing level. In this formulation, X is given by $X_t = x_t - J_{\Lambda(t)}$ where $\Lambda(t) = \int_0^t \lambda(s)\, ds$, $J_t = \sum_{i=1}^{N_t} Y_i$ is a compound Poisson process with intensity 1 under $\mathbb{P}$, and

$$dx_t = (c_t + u_t)\, dt + \epsilon_t\, dB_t$$

for a Brownian motion B_t under $\mathbb{P}$. Principal and agent share the contract C_τ. In addition, they have to find the optimal choice of effort u_t and prevention λ_t. Application of the first-order condition implies the classical *Borch rule* for risk sharing, which states that the ratio of marginal utilities is constant at the risk-sharing optimum, i.e.,

$$U_1'(C_\tau - E_\tau) = \rho\, U_2'\big(X_\tau - C_\tau - d(\tau)\big).$$

Using that both agent and principal have exponential utility functions given by (2.5) and (2.8), respectively, we find

$$C_\tau = \frac{\gamma_2}{\gamma_1+\gamma_2} X_\tau + \frac{\gamma_1}{\gamma_1+\gamma_2} E_\tau - \frac{1}{\gamma_1+\gamma_2} \log(\rho) - \frac{\gamma_2}{\gamma_1+\gamma_2} d(\tau). \tag{4.2}$$

Substituting this expression back into (4.1), the maximization problem becomes

$$-\rho^{\frac{\gamma_1}{\gamma_1+\gamma_2}}\Big(\frac{1}{\gamma_1}+\frac{1}{\gamma_2}\Big)\mathbb{E}^{\mathbb{P}}\Big[\exp\Big(-\frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2}\big(X_\tau - E_\tau - d(\tau)\big)\Big)\Big] + \frac{1}{\gamma_1} + \frac{\rho}{\gamma_2}.$$

We can write

$$\mathbb{E}^{\mathbb{P}}\Big[\exp\Big(-\frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2}\big(X_\tau - E_\tau - d(\tau)\big)\Big)\Big] = \mathbb{E}^{\mathbb{P}}\Big[M_\tau \mathrm{e}^{-\left(\int_0^\tau q_t(u_t,\lambda_t)\,dt - \frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2} d(\tau)\right)}\Big],$$

where we define

$$\begin{aligned} M_\tau = \exp\Big(&-\frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2}\int_0^\tau \epsilon_t\,dB_t - \frac{\gamma_1^2\gamma_2^2}{2(\gamma_1+\gamma_2)^2}\int_0^\tau \epsilon_t^2\,dt + \frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2} J_{\Lambda(\tau)} \\ &- \int_0^\tau \lambda_t\,dt \int_0^\infty \Big(\mathrm{e}^{\frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2}y} - 1\Big) dF(y)\Big), \\ q_t(u_t,\lambda_t) = {}& \frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2}\big(c_t + u_t - e(u_t,\lambda_t)\big) - \frac{\gamma_1^2\gamma_2^2}{2(\gamma_1+\gamma_2)^2}\epsilon_t^2 - \lambda_t \int_0^\infty \Big(\mathrm{e}^{\frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2}y} - 1\Big) dF(y). \end{aligned}$$

Note that q_t is a deterministic function and M is a martingale because ϵ_t, λ_t and Y_i are bounded. Consequently, we have

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}\Big[M_\tau \mathrm{e}^{-\left(\int_0^\tau q_t(u_t,\lambda_t)\,dt - \frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2} d(\tau)\right)}\Big] &\geq \mathbb{E}^{\mathbb{P}}\Big[M_\tau \mathrm{e}^{-\max_s\left(\int_0^s \max_{y,z} q_t(y,z)\,dt - \frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2} d(s)\right)}\Big] \\ &= \mathrm{e}^{-\max_s\left(\int_0^s \max_{y,z} q_t(y,z)\,dt - \frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2} d(s)\right)}\,\mathbb{E}^{\mathbb{P}}[M_\tau] \\ &= \mathrm{e}^{-\max_s\left(\int_0^s \max_{y,z} q_t(y,z)\,dt - \frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2} d(s)\right)} \end{aligned}$$

with equality when we choose

$$\tau^\star = \arg\max_\tau \Big(\int_0^\tau \max_{y,z} q_t(y,z)\,dt - \frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2} d(\tau)\Big)$$

and $u_t^\star$ and $\lambda_t^\star$ as the maximizers of $q_t(u_t,\lambda_t)$. Therefore, $u_t^\star$ and $\lambda_t^\star$ are given by

$$\frac{\partial}{\partial u} e(u_t^\star,\lambda_t^\star) = 1, \qquad \frac{\partial}{\partial \lambda} e(u_t^\star,\lambda_t^\star) = -\frac{\gamma_1+\gamma_2}{\gamma_1\gamma_2}\int_0^\infty \Big(\mathrm{e}^{\frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2}y} - 1\Big) dF(y).$$

This shows that the optimal actions are deterministic and constant in time. The optimal contract (4.2) is linear in both the accident part and the continuous component of the output process. The contract has pay-performance sensitivity $\frac{\gamma_2}{\gamma_1+\gamma_2}$, which is in agreement with standard results showing that the optimal risk-sharing fraction is determined by the relative size of the risk aversion parameters of principal and agent. In the limit case $\gamma_1 \searrow 0$ with fixed $\gamma_2 > 0$, we obtain that such sensitivity goes to one. This means that if the agent is indifferent toward accident and volatility risk, then the entire output process is transferred to him. As discussed in Section 3.2, the first-best actions of the agent would then satisfy (3.7). It can indeed be verified from the equation above that $\frac{\partial}{\partial \lambda} e(u_t^\star, \lambda_t^\star) = -m$ when $\gamma_1 \searrow 0$ with fixed $\gamma_2 > 0$, where we recall that m is the average cost per accident.

The higher the risk aversion level of the principal, and the the higher is the fraction of the output process that he wants to transfer to the agent to get rid off risk, a result consistent with the classical Holmström and Milgrom (1987) where accident risk is absent. If the principal becomes risk neutral, $\gamma_2 \searrow 0$ with fixed $\gamma_1 > 0$, the pay-performance sensitivity approaches zero as he wants to keep the entire risk himself. The optimal actions of the agent still satisfy (3.7).

5 Effort substitution effects

As optimal effort and prevention are defined via the maximization in (3.4), we provide a numerical study to assess the dependence of substitution effects on the variance of accident costs and levels of risk aversion. We choose an effort and prevention cost allocation function which makes the two tasks *substitutes*, given by

$$e(u, \lambda) = (\kappa_1 u)^2 + (\kappa_2/\lambda)^2 + 2\rho\kappa_1 u \kappa_2/\lambda \tag{5.1}$$

with the task interaction parameter $\rho > 0$ (see also Remark 2.1.1 for further discussion). Here, the quantity $1/\lambda$ captures the prevention action of the agent. Moreover, $\kappa_1, \kappa_2 > 0$ are constants converting effort and prevention into monetary levels. As in the seminal paper of Holmström and Milgrom (1987) dealing with single-task multiple agent problems, the cost function given in (5.1) is quadratic, but here in two arguments: effort and prevention actions. Clearly, higher ρ translates into higher marginal costs for both activities since

$$\frac{\partial e(u,\lambda)}{\partial u} = 2\kappa_1 \left(\kappa_1 u + \frac{\rho\kappa_2}{\lambda}\right) > 0 \quad \text{and} \quad \frac{\partial e(u,\lambda)}{\partial \lambda} = -\frac{2\kappa_2}{\lambda^2}\left(\frac{\kappa_2}{\lambda} + \rho\kappa_1 u\right) < 0.$$

Next, we demonstrate that the interaction parameter ρ plays a key role in deterministic substitution effects of agent's resources. We illustrate this under different contractual environments. We choose the distribution of the jump sizes Y_i so that $Y_i - \left(1 - \frac{1}{\varpi}\right)$ is exponential with parameter ϖ. This allows to control for the mean (identically equal to one), while varying the standard deviation $\frac{1}{\varpi}$ of accidents.

5.1 Case I: basic framework

The graphs in Figure 1 indicate that when ρ is increasing on a low level, the agent applies less prevention. However, if marginal costs become higher, the amount of prevention applied by the agent increases in ρ. This indicates that the agent is engaging in effort substitution, shifting resources away from effort to prevention, and reflects the fundamentally different nature of the two tasks. While applying little effort results in small instantaneous output growth, applying little prevention would result in accidents occurring at a large frequency, hence depressing significantly the outcome process. Having to choose where to devote his resources, the agent allocates them to the most critical task, especially if he is risk averse.

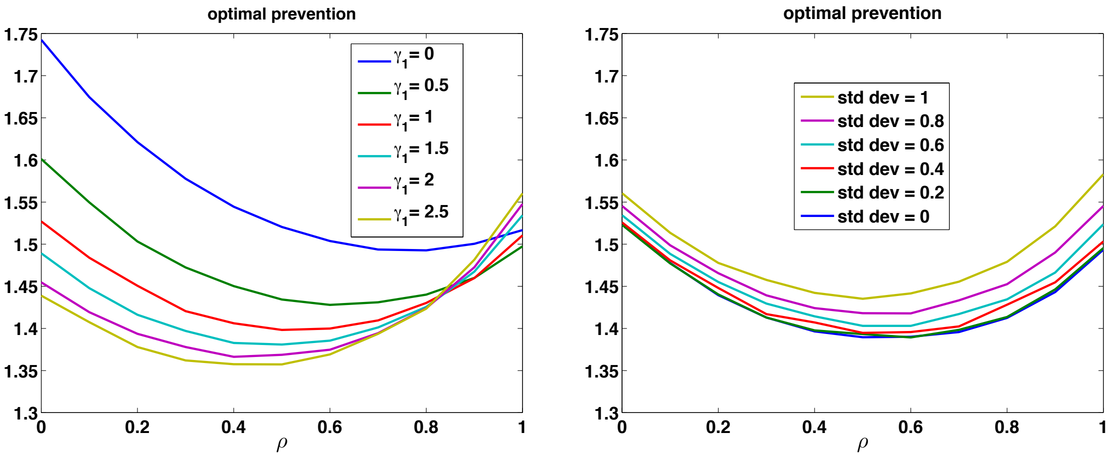

Figure 1: Optimal prevention action with respect to the strength of task interactions. We set $\epsilon = 1$, $\kappa_1 = 0.7$, $\kappa_2 = 0.3$, $\gamma_2 = 0.3$, and $\tau^\star = T = 1$. The left panel corresponds to accidents with standard deviation 0.5. In the right panel, we choose $\gamma_1 = 1$, and vary the standard deviation of accidents.

Like the sublinear form of the contract, these substitution effects depend crucially on the agent's risk aversion, as the left graph shows. The most risk-averse agent applies the highest prevention when the strength of task interaction is the largest. Moreover, when accident costs have higher variability, the risk averse agent applies more prevention given that the resulting negative effects are harder to predict.

5.2 Case II: accident coverage

We allow for the principal to take prevention measures which cover accident losses exceeding a threshold A. This means that the principal will be exposed to a maximum loss of A. Mathematically, the jump size Y_i modeling accident costs is replaced by $\tilde{Y}_i = \min(Y_i, A)$. The cost of these measures is detracted from the initial output process, and we fix $\tau = T$. In the absence of task interaction the incentive compatibility conditions are decoupled, hence the agent always applies the same level of effort regardless of the coverage threshold. However when tasks interact and full coverage is not guaranteed, Figure 2 shows that the agent progressively reduces the applied effort and increases the amount of accident prevention as the guaranteed coverage decreases.

Figure 3 shows the premiums at which it is beneficial for the principal to buy (area below the curve) or not to buy (area above the curve) accident coverage. On the

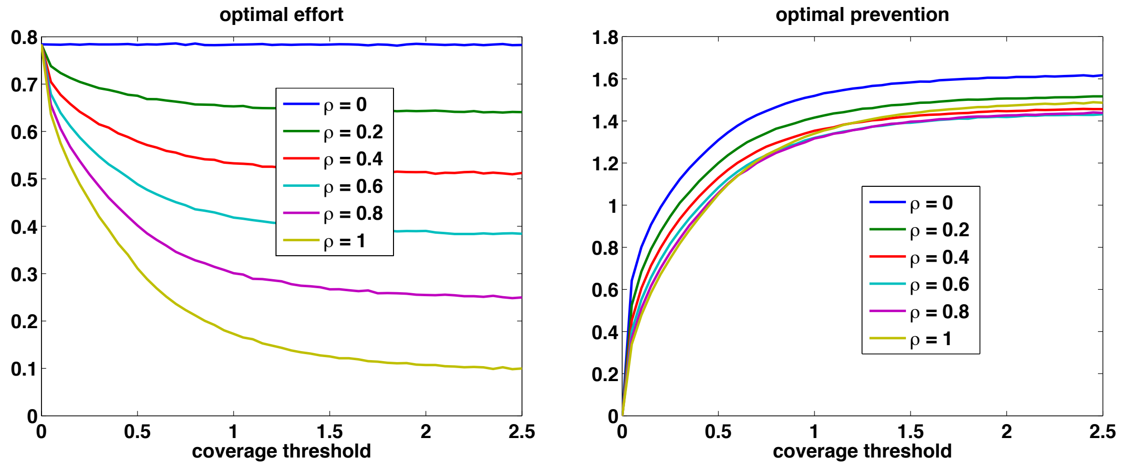

Figure 2: Optimal actions with respect to percentage coverage amount for different levels of ρ. The parameters are $\epsilon = 1$, $\varpi = 2$, $\gamma_1 = 0.4$, $\gamma_2 = 0.3$, $\kappa_1 = 0.7$, $\kappa_2 = 0.3$ and $\tau^\star = T = 1$.

curve, the principal is indifferent between having and not having accident coverage. The critical premium is defined so that the principal's certainty equivalent inclusive of the coverage at the critical premium equals the principal's certainty equivalent without the coverage. As ρ increases and the agent suffers higher costs from task interaction, the principal is willing to pay more to be insulated against accident losses.

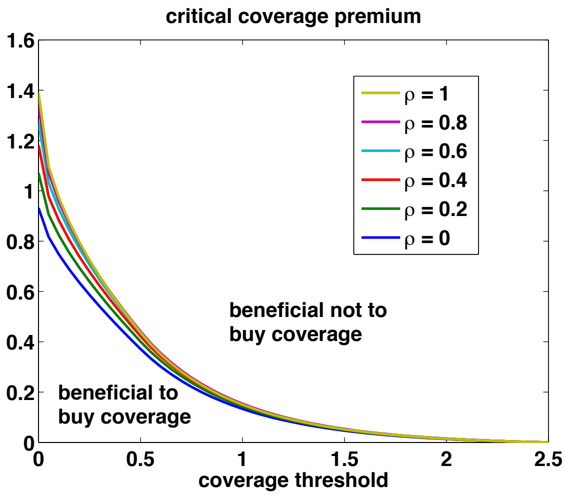

Figure 3: The critical coverage premium as a function of the coverage threshold, plotted for different levels of ρ. The parameters are the same as in Figure 2.

5.3 Case III: promoting the agent

We consider the situation in which the principal can train the agent at time ν upon incurring training costs $d(\nu)$. As in Sannikov (2008), we assume that training permanently increases the agent's productivity of effort from u_t to $\theta_1 u_t$. Additionally, we assume that it permanently increases the effectiveness of prevention, in that it reduces the frequency of occurring accidents from λ_t to λ_t/θ_2. Here, $\theta_1, \theta_2 \geq 1$ are constants. This is equivalent to changing the cost function to $\tilde{e}(u_t, \lambda_t) = e(u_t/\theta_1, \lambda_t\theta_2)$ after the promotion. For notational convenience, we set $d(T) = 0$ so that the principal does not train at all the agent when $\nu = T$. The principal's optimization problem is to maximize

$$\mathbb{E}^{u,\lambda}[U_2(X_T - C_T - d(\nu))]$$

over C_T and ν where

$$R_0 = \sup_{u,\,\lambda} \mathbb{E}^{u,\lambda}[U_1(C_T - E_{0,\nu,T})] \ \text{ with } E_{0,\nu,T} = \int_0^{\nu} e(u_t, \lambda_t)\,dt + \int_{\nu}^{T} \big(\tilde{e}(u_t, \lambda_t) + f(t)\big)\,dt.$$

Here, $f(t)$ models the opportunity costs of the agent at t deriving from his promotion. As his new skills become more valuable to other principals, the current principal would need to increase his incentives to keep him. This problem can be solved very similarly to Section 3.1, resulting in the optimal contract of the form (3.3) with τ^* replaced by T, and the function $e(u, \lambda)$ replaced by

$$e^*_t(u_t, \lambda_t) = \begin{cases} e(u_t, \lambda_t) & \text{for } t \leq \nu^\star \\ \tilde{e}(u_t, \lambda_t) + f(t) = e(u_t/\theta_1, \lambda_t\theta_2) + f(t) & \text{for } t > \nu^\star \end{cases};$$

the optimal promotion time ν^* is the maximizer of

$$\int_0^{\nu} p_t(\alpha^\star_t, \beta^\star_t)\,dt + \int_{\nu}^{T} \tilde{p}_t\big(\tilde{\alpha}^\star_t, \tilde{\beta}^\star_t\big)\,dt - d(\nu) - \int_{\nu}^{T} f(t)\,dt,$$

where $\tilde{p}_t$ is defined as p_t in (3.4) but with e replace by $\tilde{e}$. We next analyze how the opportunity costs of the agent affect the optimal promotion time for different levels of post-training efficiency and strength of task interaction. Clearly, an early promotion has higher opportunity costs than a later promotion, given that the agent would have more and better options in earlier stages of his career. This means that $f(t)$ should be decreasing in time.

As expected, the principal finds it optimal to promote the agent later if the interaction of the two tasks is stronger. Figure 4 shows that when task interaction is high, the promotion time is highly sensitive to the opportunity costs of the agent after promotion. A small increase in such costs may lead the principal to change his decision from immediately promoting the agent to never promoting him. By contrast, when task interaction is small the principal gradually postpones the promotion time if the opportunity costs increase. This happens because the efficiency gains after the promotion time outweigh the opportunity costs of the agent if he is promoted later.

5.4 Case IV: firm's liquidation

The principal can liquidate the firm and lay off the agent prematurely at some time $\tau \leq T$. The mathematical formulation is similar to the previous section with the

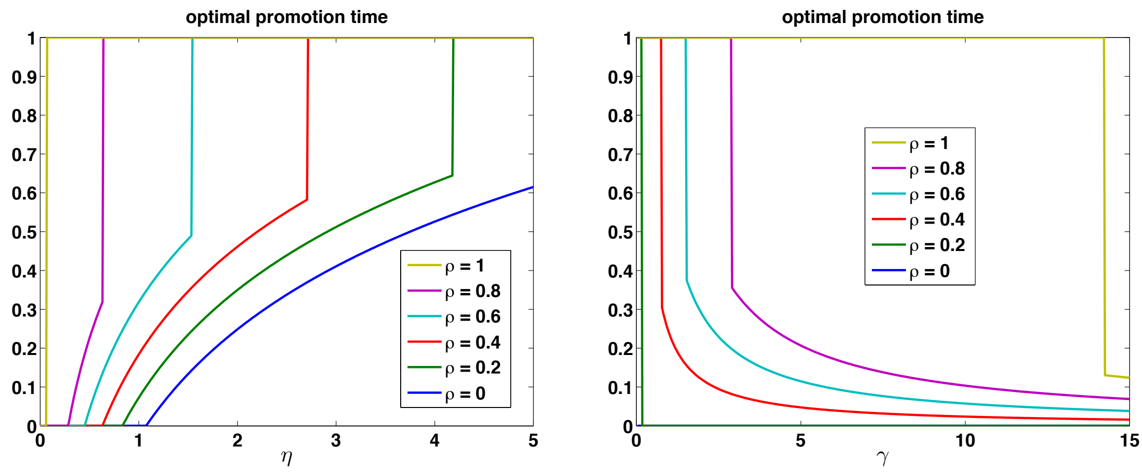

Figure 4: The optimal promotion time depending on the agent's opportunity cost and task interaction. We set $\epsilon = 1$, $\varpi = 2$, $\gamma_1 = 0.4$, $\gamma_2 = 0.3$, $\kappa_1 = 0.7$, $\kappa_2 = 0.3$, $T = 1$, $d(s) = 0.1$ for $s < T$, $\theta_1 = 1.2$ and $\theta_2 = 3.3$. We choose $f(t) = \eta e^{-\gamma t}$. In the left panel, we fix $\gamma = 2.5$, while we choose $\eta = 0.5$ in the right panel.

principal maximizing

$$\mathbb{E}^{u,\lambda}[U_2(X_\tau - C_\tau)]$$

over C_τ and τ where

$$R_0 = \sup_{u,\lambda} \mathbb{E}^{u,\lambda}[U_1(C_\tau - E_{0,\tau,T})] \quad \text{with } E_{0,\tau,T} = \int_0^\tau e(u_t, \lambda_t)\, dt + \int_\tau^T f(t)\, dt.$$

In particular, the termination costs are $d(\tau) = 0$. The term $\int_\tau^T f(t)\, dt$ reflects the benefit the agent has from a premature contract termination. Such benefits typically decrease with time as the opportunities for the agent would reduce with passage of time. The optimal contract is of a similar form as in (3.3) where $\tau^\star$ is now the maximizer of

$$\int_0^\tau p_t(\alpha_t^\star, \beta_t^\star)\, dt - \int_\tau^T f(t)\, dt.$$

As expected, the principal always liquidates earlier if task interaction is higher. Interestingly, a direct comparison of left and right panels of Figure 5 shows that when both task interaction and standard deviation of accidents are large, a small increase in accident uncertainty only slightly anticipates the liquidation time, but this action positively affects the principal's certainty equivalent gains. By contrast, when task interaction is small and standard deviation of accidents is large, the opposite happens: the principal anticipates the liquidation time, but his resulting certainty equivalent gains are negligible. This can be explained in terms of effort substitution. If task interaction is small, the agent applies high effort. As a consequence, when the principal liquidates prematurely he gains because accident losses are eliminated, but loses the short term profits resulting from the agent's effort action. Altogether, this leads to small net gains. Differently, when task interaction is high, the agent

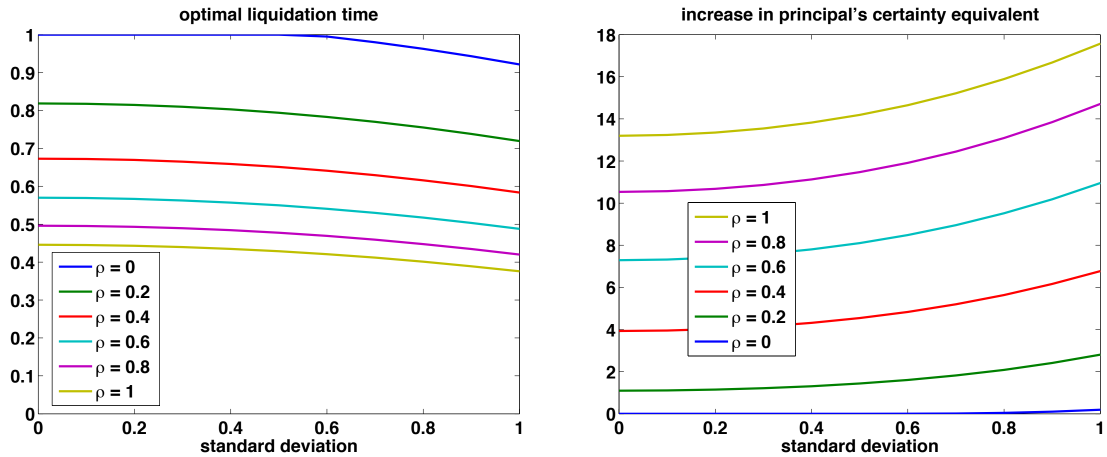

Figure 5: The optimal time to liquidate the firm and the corresponding certainty equivalent gains for the principal. We set $\epsilon = 1$, $\varpi = 2$, $\gamma_1 = 0.4$, $\gamma_2 = 0.3$, $\kappa_1 = 0.7$, $\kappa_2 = 0.3$ and $T = 1$. We choose $f(t) = \eta e^{-\gamma t}$ with $\gamma = 1$ and $\eta = 1.6$.

diverts most of his resources to preventing accidents and only applies little effort. Hence, when the principal liquidates earlier he gains from the absence of accidents, but does not lose much in terms of short-term profit generation.

6 Conclusions

We have developed a continuous-time contracting framework to analyze optimal incentives provision in multitasking settings. The agent can maximize short-term output growth through effort, and perform accident prevention. The accident prevention task breaks the linearity of the contract, and induces the principal to offer a sublinear punishment scheme to the risk-averse agent. Such a punishment scheme only becomes linear if the agent is risk neutral, in which case he behaves in the same way as if he owned the company. The influence of the agent's risk aversion is also reflected in his resource allocation decisions, where the critical impact of accidents on the outcome process induces him to divert resource away from effort to prevention. This substitution effect is significant regardless of whether accident prevention measures are taken by the principal.

A Proofs

Proof of Lemma 3.1. Using that the agent's reservation utility is R_0, we have $R_0 = \mathbb{E}^{u,\lambda}[U_1(C_\tau - E_\tau)]$ or, equivalently,

$$1 - \gamma_1 R_0 = \mathbb{E}^{u,\lambda}[\exp(-\gamma_1 C_\tau + \gamma_1 E_\tau)] \tag{A.1}$$

at the optimum. Indeed, it holds that $1 - \gamma_1 R_0 \geq \mathbb{E}^{u,\lambda}[\exp(-\gamma_1 C_\tau + \gamma_1 E_\tau)]$ by the individual rationality constraint (2.7), and if we assume to the contrary of (A.1) that

$1-\gamma_1 R_0 > \mathbb{E}^{u,\lambda}[\exp(-\gamma_1 C_\tau + \gamma_1 E_\tau)]$, the principal could replace C_τ by $\tilde{C}_\tau = C_\tau - c$ with

$$c = \frac{1}{\gamma_1} \log \frac{1-\gamma_1 R_0}{\mathbb{E}^{u,\lambda}[\exp(-\gamma_1 C_\tau + \gamma_1 E_\tau)]}.$$

Because $c > 0$, we have $\tilde{C}_\tau < C_\tau$, and hence the principal gets a strictly greater expected utility from using $\tilde{C}_\tau$ instead of C_τ while $\tilde{C}_\tau$ satisfies the individual rationality constraint (2.7) with equality by the choice of c. Hence, C_τ cannot be an optimal contract, which in turn proves the equality in (A.1). This implies that we can write

$$\exp(-\gamma_1 C_\tau + \gamma_1 E_\tau) = (1-\gamma_1 R_0) M_\tau \tag{A.2}$$

at time τ, where M is a strictly positive martingale under the measure $\mathbb{P}^{u,\lambda}$ with $M_0 = 1$. Indeed, we can define M by $M_t = \mathbb{E}^{u,\lambda}[\exp(-\gamma_1 C_\tau + \gamma_1 E_\tau)|\mathcal{F}_t]/(1-\gamma_1 R_0)$, using that $\exp(-\gamma_1 C_\tau + \gamma_1 E_\tau)$ is $\mathbb{P}^{u,\lambda}$-integrable due to (A.1). By the martingale representation theorem (see Theorem 13.19 of He et al. (1992)), there exist predictable processes $\tilde{\alpha}$ and $\tilde{\beta}$ such that

$$\frac{dM_t}{M_{t-}} = \tilde{\alpha}_t \, dB_t^u + \tilde{\beta}_t \, (dJ_t - \lambda_t m \, dt), \tag{A.3}$$

which is equivalent to

$$M_\tau = \exp\left(\int_0^\tau \tilde{\alpha}_t \, dB_t^u + \sum_{0<t\leq\tau} \log\left(1 + \tilde{\beta}_t \Delta J_t\right) - \int_0^\tau \left(\frac{1}{2}\tilde{\alpha}_t^2 + \tilde{\beta}_t \lambda_t m\right) dt \right)$$

by the formula for the stochastic exponential; see Theorem II.37 of Protter (2005). Combining this with (A.2) yields

$$\begin{aligned} C_\tau = &-\frac{1}{\gamma_1}\log(1-\gamma_1 R_0) - \frac{1}{\gamma_1}\int_0^\tau \tilde{\alpha}_t \, dB_t^u - \frac{1}{\gamma_1}\sum_{0<t\leq\tau} \log\left(1 + \tilde{\beta}_t \Delta J_t\right) \\ &+ \int_0^\tau \left(\frac{1}{2\gamma_1}\tilde{\alpha}_t^2 + \frac{1}{\gamma_1}\tilde{\beta}_t \lambda_t m + e(u_t, \lambda_t) \right) dt, \end{aligned}$$

from which we deduce (3.1) by defining $\alpha = -\frac{\tilde{\alpha}_t}{\gamma_1 \epsilon_t}$, $\beta = \frac{\tilde{\beta}_t}{\gamma_1}$ and using the dynamics (2.4) of x_t. ■

Proof of Proposition 3.3. For the optimal contract (3.1) and any not necessarily optimal actions $\tilde{u}$ and $\tilde{\lambda}$, we have $R_0 \geq \mathbb{E}^{\tilde{u},\tilde{\lambda}}\big[U_1\big(C_\tau - \int_0^\tau e\big(\tilde{u}_t, \tilde{\lambda}_t\big)\, dt\big)\big]$ or, equivalently,

$$1 - \gamma_1 R_0 \leq \mathbb{E}^{\tilde{u},\tilde{\lambda}}\left[\exp\left(- \gamma_1 C_\tau + \gamma_1 \int_0^\tau e\big(\tilde{u}_t, \tilde{\lambda}_t\big)\, dt\right)\right]$$

with equality for $\tilde{u} = u$ and $\tilde{\lambda} = \lambda$. We set

$$\begin{aligned}
Y_s &= \frac{1}{1-\gamma_1 R_0} \exp\left(- \gamma_1 C_s + \gamma_1 \int_0^s e\big(\tilde{u}_t, \tilde{\lambda}_t\big)\, dt\right) \\
&= \exp\Bigg(- \int_0^s \gamma_1 \alpha_t \epsilon_t \, dB_t^u + \sum_{0<t\leq s} \log(1 + \gamma_1 \beta_t \Delta J_t) - \int_0^s \Big(\frac{\gamma_1}{2}\alpha_t^2 \epsilon_t^2 + \gamma_1 \beta_t \lambda_t m\Big)\, dt \\
&\qquad\qquad + \gamma_1 \int_0^s \big(e\big(\tilde{u}_t, \tilde{\lambda}_t\big) - e(u_t, \lambda_t)\big)\, dt\Bigg).
\end{aligned}$$

This satisfies $Y_0 = 1$ and

$$\begin{aligned}
\frac{dY_t}{Y_{t-}} = &- \gamma_1 \alpha_t \epsilon_t \, dB_t^{\tilde{u}} + \gamma_1 \beta_t \left(dJ_t - \tilde{\lambda}_t m\, dt\right) \\
&+ \gamma_1 \big(\alpha_t (u_t - \tilde{u}_t) + \beta_t m \big(\tilde{\lambda}_t - \lambda_t\big) + e\big(\tilde{u}_t, \tilde{\lambda}_t\big) - e(u_t, \lambda_t)\big)\, dt.
\end{aligned}$$

If the drift is zero, Y is a martingale thanks to the boundedness assumptions on α_t and β_t; see Remark 3.2. For the optimal choices $\tilde{u} = u$ and $\tilde{\lambda} = \lambda$, the drift indeed vanishes so that $Y = M$ is a martingale, with M defined in (A.3). Because this is the optimal choice, the drift needs to be nonnegative in general. Hence, $\tilde{u}_t = u_t$ and $\tilde{\lambda}_t = \lambda_t$ minimize

$$\alpha_t (u_t - \tilde{u}_t) + \beta_t m \big(\tilde{\lambda}_t - \lambda_t\big) + e\big(\tilde{u}_t, \tilde{\lambda}_t\big) - e(u_t, \lambda_t),$$

which is equivalent to (3.2) as first-order condition, using the assumptions on e. ■

Proof of Theorem 3.4. From Lemma 3.1 and Proposition 3.3, the contract is of the form

$$\begin{aligned}
C_\tau = &- \frac{1}{\gamma_1} \log(1 - \gamma_1 R_0) + \int_0^\tau \alpha_t \, dx_t - \frac{1}{\gamma_1} \sum_{0<t\leq\tau} \log(1 + \gamma_1 \beta_t \Delta J_t) \\
&+ \int_0^\tau \Big(\frac{\gamma_1}{2}\alpha_t^2 \epsilon_t^2 + \beta_t \lambda_t m + e(u_t, \lambda_t) - \alpha_t c_t - \alpha_t u_t\Big)\, dt,
\end{aligned} \tag{A.4}$$

with $u_t = i_t(\alpha_t, \beta_t)$ and $\lambda_t = j_t(\alpha_t, \beta_t)$. Using this expression for C_τ in (2.6) and

(2.8), we need to minimize

$$\mathbb{E}^{u,\lambda}\Big[\exp\Big(-\gamma_2 X_\tau - \frac{\gamma_2}{\gamma_1}\log(1-\gamma_1 R_0) + \gamma_2\int_0^\tau \alpha_t\,dx_t - \frac{\gamma_2}{\gamma_1}\sum_{0<t\le\tau}\log(1+\gamma_1\beta_t\Delta J_t)$$
$$+\gamma_2 d(\tau) + \gamma_2\int_0^\tau\Big(\frac{\gamma_1}{2}\alpha_t^2\epsilon_t^2 + \beta_t\lambda_t m + e(u_t,\lambda_t) - \alpha_t c_t - \alpha_t u_t\Big)\,dt\Big)\Big]$$

over α, β and τ with $u_t = i_t(\alpha_t,\beta_t)$ and $\lambda_t = j_t(\alpha_t,\beta_t)$. Recalling (2.2) and (2.4) for the dynamics of X, the optimization problem becomes

$$\min_{\alpha,\beta,\tau}\ \mathbb{E}^{u,\lambda}\Big[M_\tau \mathrm{e}^{-\gamma_2\left(\int_0^\tau p_t(\alpha_t,\beta_t)\,dt - d(\tau)\right)}\Big],$$

where we use the abbreviations

$$\begin{aligned}
M_t &= (1-\gamma_1 R_0)^{-\gamma_2/\gamma_1}\exp\Big(-\gamma_2\int_0^t \epsilon_s(1-\alpha_s)\,dB_s^u - \frac{\gamma_2^2}{2}\int_0^t \epsilon_s^2(1-\alpha_s)^2\,ds\Big)\\
&\quad\times\prod_{0<s\le t}(1+\gamma_1\beta_s\Delta J_s)^{-\gamma_2/\gamma_1}\mathrm{e}^{\gamma_2\Delta J_s}\\
&\quad\times\ \exp\Big(-\int_0^t\Big(\int_0^\infty(1+\gamma_1\beta_s y)^{-\gamma_2/\gamma_1}\mathrm{e}^{\gamma_2 y}\,dF(y) - 1\Big)\lambda_s\,ds\Big),\\
p_t(\alpha_t,\beta_t) &= -\,e\big(i_t(\alpha_t,\beta_t), j_t(\alpha_t,\beta_t)\big) - \frac{\gamma_1}{2}\alpha_t^2\epsilon_t^2 + c_t + i_t(\alpha_t,\beta_t) - \frac{\gamma_2}{2}\epsilon_t^2(\alpha_t-1)^2\\
&\quad-\Big(\frac{1}{\gamma_2}\int_0^\infty(1+\gamma_1\beta_t y)^{-\gamma_2/\gamma_1}\mathrm{e}^{\gamma_2 y}\,dF(y) - \frac{1}{\gamma_2} + \beta_t m\Big) j_t(\alpha_t,\beta_t).
\end{aligned}$$

Using that p_t is a deterministic function and M is a martingale because α and β are bounded for an admissible contract, we have

$$\begin{aligned}
\mathbb{E}^{u,\lambda}\Big[M_\tau\mathrm{e}^{-\gamma_2\left(\int_0^\tau p_t(\alpha_t,\beta_t)\,dt - d(\tau)\right)}\Big] &\ge \mathbb{E}^{u,\lambda}\Big[M_\tau\mathrm{e}^{-\gamma_2\max_s\left(\int_0^s \max_{y,z} p_t(y,z)\,dt - d(s)\right)}\Big]\\
&= \mathrm{e}^{-\gamma_2\max_s\left(\int_0^s \max_{y,z} p_t(y,z)\,dt - d(s)\right)}\ \mathbb{E}^{u,\lambda}[M_\tau]\\
&= \mathrm{e}^{-\gamma_2\max_s\left(\int_0^s \max_{y,z} p_t(y,z)\,dt - d(s)\right)}(1-\gamma_1 R_0)^{-\gamma_2/\gamma_1}
\end{aligned}$$

with equality when choosing $\alpha_t^\star$ and $\beta_t^\star$ to be the maximizers of $p_t(\alpha_t,\beta_t)$ and $\tau=\tau^\star$ as in (3.5). This shows that $\alpha^\star$, $\beta^\star$ and $\tau^\star$ are the maximizers of

$$\begin{aligned}
\mathbb{E}^{u,\lambda}[U_2(X_\tau - C_\tau - d(\tau))] &= \mathbb{E}^{u,\lambda}\Big[-\frac{1}{\gamma_2}\Big(M_\tau\mathrm{e}^{\int_0^\tau p_t(\alpha_t,\beta_t)\,dt - \gamma_2 d(\tau)} - 1\Big)\Big]\\
&= -\frac{1}{\gamma_2}\Big((1-\gamma_1R_0)^{-\gamma_2/\gamma_1}\mathrm{e}^{-\gamma_2\left(\int_0^\tau p_t(\alpha_t,\beta_t)\,dt - d(\tau)\right)} - 1\Big)\\
&= U_2\Big(\int_0^\tau p_t(\alpha_t,\beta_t)\,dt - d(\tau) + \frac{1}{\gamma_1}\log(1-\gamma_1R_0)\Big),
\end{aligned}$$

and the optimal contract is given by (A.4) with $\alpha = \alpha^\star$, $\beta = \beta^\star$ and $\tau = \tau^\star$. To check that this contract and the corresponding actions are indeed admissible, we first fix some arbitrary $\tilde{u} \geq 0$ and $\tilde{\lambda} > 0$ with corresponding $\tilde{\alpha} = \frac{\partial}{\partial u} e\big(\tilde{u}, \tilde{\lambda}\big)$ and $\tilde{\beta} = -\frac{1}{m}\frac{\partial}{\partial \lambda} e\big(\tilde{u}, \tilde{\lambda}\big)$. Because ϵ_t is a bounded function, $p_t\big(\tilde{\alpha}, \tilde{\beta}\big) - c_t$ is also a bounded function so that there exists K such that $p_t\big(\tilde{\alpha}, \tilde{\beta}\big) - c_t \geq K$ for all t. Using that

$$p_t(\alpha_t, \beta_t) - c_t \leq -\, e(u_t, \lambda_t) + u_t - \bigg(\frac{1}{\gamma_2}\int_0^\infty \Big(1 - \frac{\gamma_1 y}{m}\frac{\partial}{\partial \lambda} e(u_t, \lambda_t)\Big)^{-\gamma_2/\gamma_1} \mathrm{e}^{\gamma_2 y}\, dF(y) \\ - \frac{1}{\gamma_2} - \frac{\partial}{\partial \lambda} e(u_t, \lambda_t)\bigg)\lambda_t,$$

we check that the right-hand side (and thus also $p_t(\alpha_t, \beta_t) - c_t$) diverges to $-\infty$ if $u_t \nearrow \infty$ or $\lambda_t \nearrow \infty$ or $\lambda_t \searrow 0$. In turn, this yields the existence of constants $k_1, k_2, k_3 > 0$ such that $p_t(\alpha_t, \beta_t) - c_t < K$ if $u_t > k_1$ or $\lambda_t < k_2$ or $\lambda_t > k_3$. Therefore, the optimal u_t^* and λ_t^* are valued in bounded sets $[0, k_1]$ and $[k_2, k_3]$, respectively, for all $t \in [0, T]$, and hence also the optimal α_t^* and β_t^*, which are related to u_t^* and λ_t^* by (3.2), are bounded. This shows that the optimal contract and actions are admissible. ■

Proof of Corollary 3.5. We first note that $i_t(\alpha_t, \beta_t)$ and $j_t(\alpha_t, \beta_t)$ do not depend on γ_1; see (3.2). Therefore, to consider $p_t(\alpha_t, \beta_t)$ as $\gamma_1 \searrow 0$, it is enough to deduce from

$$\lim_{\gamma_1 \searrow 0} (1 + \gamma_1 \beta_t y)^{-\gamma_2/\gamma_1} = \lim_{z \to \infty} \Big(1 + \frac{\gamma_2 \beta_t y}{z}\Big)^{-z} = \mathrm{e}^{-\gamma_2 \beta_t y},$$

that

$$\lim_{\gamma_1 \searrow 0} p_t(\alpha_t, \beta_t) = -\, e\big(i_t(\alpha_t, \beta_t), j_t(\alpha_t, \beta_t)\big) + c_t + i_t(\alpha_t, \beta_t) \\ - \frac{\gamma_2}{2}\epsilon_t^2(\alpha_t - 1)^2 - \bigg(\frac{1}{\gamma_2}\int_0^\infty \mathrm{e}^{\gamma_2 y(1-\beta_t)}\, dF(y) - \frac{1}{\gamma_2} + \beta_t m\bigg) j_t(\alpha_t, \beta_t)$$

and denote $p_t^1(\alpha_t, \beta_t) = \lim_{\gamma_1 \searrow 0} p_t(\alpha_t, \beta_t)$. This convergence is uniform in (α_t, β_t) for bounded values. We note that in (3.3), we have

$$\lim_{\gamma_1 \searrow 0} \Big(-\frac{1}{\gamma_1}\log(1 - \gamma_1 R_0)\Big) = R_0,$$
$$\lim_{\gamma_1 \searrow 0} \bigg(\frac{1}{\gamma_1}\sum_{0<t\leq\tau^\star} \log(1 + \gamma_1 \beta_t^\star \Delta J_t)\bigg) = \sum_{0<t\leq\tau^\star} \beta_t^\star \Delta J_t = \int_0^{\tau^\star} \beta_t^\star\, dJ_t \quad \text{a.s.}$$

Therefore, using that pointwise convergence of an integrand gives convergence in probability of the corresponding stochastic integral, we deduce that the optimal contract converges in probability to

$$C_{\tau^\star}^{1;\star} = R_0 + \int_0^{\tau^\star} e(u_t^\star, \lambda_t^\star)\, dt + \int_0^{\tau^\star} \alpha_t^\star(dx_t - (c_t + u_t^\star)\, dt) - \int_0^{\tau^\star} \beta_t^\star(dJ_t - \lambda_t^\star m\, dt),$$

where $u_t^\star = i_t(\alpha_t^\star, \beta_t^\star)$, $\lambda_t^\star = j_t(\alpha_t^\star, \beta_t^\star)$, and $\alpha_t^\star$ and $\beta_t^\star$ are the maximizers of

$$\begin{aligned} p_t^1(\alpha_t, \beta_t) = &- e\big(i_t(\alpha_t, \beta_t), j_t(\alpha_t, \beta_t)\big) + c_t + i_t(\alpha_t, \beta_t) - \frac{\gamma_2}{2}\epsilon_t^2(\alpha_t - 1)^2 \\ &- \left(\frac{1}{\gamma_2}\int_0^\infty \mathrm{e}^{\gamma_2 y(1-\beta_t)}\, dF(y) - \frac{1}{\gamma_2} + \beta_t m\right) j_t(\alpha_t, \beta_t). \end{aligned}$$

Using the incentive compatibility condition (3.2), we can write $p_t^1(\alpha_t^\star, \beta_t^\star)$ as a function of $u_t^\star = i_t(\alpha_t^\star, \beta_t^\star)$, $\lambda_t^\star = j_t(\alpha_t^\star, \beta_t^\star)$ in the form

$$\begin{aligned} \tilde{p}_t^1(u_t^\star, \lambda_t^\star) = &- e(u_t^\star, \lambda_t^\star) + c_t + u_t^\star - \frac{\gamma_2}{2}\epsilon_t^2\left(\frac{\partial}{\partial u}e(u_t^\star, \lambda_t^\star) - 1\right)^2 \\ &- \left(\frac{1}{\gamma_2}\int_0^\infty \mathrm{e}^{\gamma_2 y\left(1+\frac{1}{m}\frac{\partial}{\partial\lambda}e(u_t^\star,\lambda_t^\star)\right)}\, dF(y) - \frac{1}{\gamma_2} - \frac{\partial}{\partial\lambda}e(u_t^\star, \lambda_t^\star)\right)\lambda_t^\star. \end{aligned}$$

The first-order conditions yield

$$\begin{aligned} \frac{\partial}{\partial u}\tilde{p}_t^1(u_t^\star, \lambda_t^\star) = &- \frac{\partial}{\partial u}e(u_t^\star, \lambda_t^\star) + 1 - \gamma_2\epsilon_t^2\left(\frac{\partial}{\partial u}e(u_t^\star, \lambda_t^\star) - 1\right)\frac{\partial^2}{\partial u^2}e(u_t^\star, \lambda_t^\star) \\ &- \left(\int_0^\infty y\mathrm{e}^{\gamma_2 y\left(1+\frac{1}{m}\frac{\partial}{\partial\lambda}e(u_t^\star,\lambda_t^\star)\right)}\frac{1}{m}\frac{\partial^2}{\partial\lambda\partial u}e(u_t^\star, \lambda_t^\star)\, dF(y) - \frac{\partial^2}{\partial\lambda\partial u}e(u_t^\star, \lambda_t^\star)\right)\lambda_t^\star \\ = &- \left(1 + \gamma_2\epsilon_t^2\frac{\partial^2}{\partial u^2}e(u_t^\star, \lambda_t^\star)\right)\left(\frac{\partial}{\partial u}e(u_t^\star, \lambda_t^\star) - 1\right) \\ &- \lambda_t^*\frac{\partial^2}{\partial\lambda\partial u}e(u_t^\star, \lambda_t^\star)\left(\int_0^\infty y\mathrm{e}^{\gamma_2 y\left(1+\frac{1}{m}\frac{\partial}{\partial\lambda}e(u_t^\star,\lambda_t^\star)\right)}\frac{1}{m}\, dF(y) - 1\right) \\ = &\ 0, \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial\lambda}\tilde{p}_t^1(u_t^\star, \lambda_t^\star) = &- \frac{\partial}{\partial\lambda}e(u_t^\star, \lambda_t^\star) - \gamma_2\epsilon_t^2\left(\frac{\partial}{\partial u}e(u_t^\star, \lambda_t^\star) - 1\right)\frac{\partial^2}{\partial\lambda\partial u}e(u_t^\star, \lambda_t^\star) \\ &- \left(\int_0^\infty y\mathrm{e}^{\gamma_2 y\left(1+\frac{1}{m}\frac{\partial}{\partial\lambda}e(u_t^\star,\lambda_t^\star)\right)}\frac{1}{m}\frac{\partial^2}{\partial\lambda^2}e(u_t^\star, \lambda_t^\star)\, dF(y) - \frac{\partial^2}{\partial\lambda^2}e(u_t^\star, \lambda_t^\star)\right)\lambda_t^\star \\ &- \left(\frac{1}{\gamma_2}\int_0^\infty \mathrm{e}^{\gamma_2 y\left(1+\frac{1}{m}\frac{\partial}{\partial\lambda}e(u_t^\star,\lambda_t^\star)\right)}\, dF(y) - \frac{1}{\gamma_2} - \frac{\partial}{\partial\lambda}e(u_t^\star, \lambda_t^\star)\right) \\ = &-\gamma_2\epsilon_t^2\left(\frac{\partial}{\partial u}e(u_t^\star, \lambda_t^\star) - 1\right)\frac{\partial^2}{\partial\lambda\partial u}e(u_t^\star, \lambda_t^\star) \\ &- \int_0^\infty \left(\mathrm{e}^{\gamma_2 y\left(1+\frac{1}{m}\frac{\partial}{\partial\lambda}e(u_t^\star,\lambda_t^\star)\right)} - 1\right)\left(\frac{1}{\gamma_2} + \frac{x}{m}\frac{\partial^2}{\partial\lambda^2}e(u_t^\star, \lambda_t^\star)\right) dF(y) \\ = &\ 0. \end{aligned}$$

These conditions are satisfied for $\frac{\partial}{\partial u}e(u_t^\star, \lambda_t^\star) = 1$ and $\frac{\partial}{\partial\lambda}e(u_t^\star, \lambda_t^\star) = -m$, which follows from the incentive compatibility condition (3.2) by choosing $\alpha_t^\star = 1$ and $\beta_t^\star = 1$. ∎

Proof of Corollary 3.6. The optimal contract in Theorem 3.4 depends on γ_2 only through $p_t(\alpha_t, \beta_t)$. Using the uniform convergence for bounded values in

$$\begin{aligned}\lim_{\gamma_2 \searrow 0} \frac{1}{\gamma_2}\big((1+\gamma_1\beta_t y)^{-\gamma_2/\gamma_1}\mathrm{e}^{\gamma_2 y} - 1\big) &= \lim_{\gamma_2 \searrow 0} \frac{1}{\gamma_2}\big(\big((1+\gamma_1\beta_t y)^{-1/\gamma_1}\mathrm{e}^{y}\big)^{\gamma_2} - 1\big)\\ &= \log \frac{\mathrm{e}^y}{(1+\gamma_1\beta_t y)^{1/\gamma_1}} \end{aligned} \tag{A.5}$$

and

$$\begin{aligned}\int_0^\infty \log \frac{\mathrm{e}^y}{(1+\gamma_1\beta_t y)^{1/\gamma_1}}\, dF(y) &= \int_0^\infty \Big(y - \frac{1}{\gamma_1}\log(1+\gamma_1\beta_t y)\Big) dF(y)\\ &= m - \frac{1}{\gamma_1}\int_0^\infty \log(1+\gamma_1\beta_t y)\, dF(y),\end{aligned}$$

we deduce the uniform convergence for bounded values of $p_t(\alpha_t, \beta_t)$ to $p_t^2(\alpha_t, \beta_t)$ as $\gamma_2 \searrow 0$.

To prove $p_t^2(\alpha_t, \beta_t) \geq p_t(\alpha_t, \beta_t)$, it is enough to show

$$\frac{1}{\gamma_2}\big((1+\gamma_1\beta_t y)^{-\gamma_2/\gamma_1}\mathrm{e}^{\gamma_2 y} - 1\big) \geq \log \frac{\mathrm{e}^y}{(1+\gamma_1\beta_t y)^{1/\gamma_1}} \quad \text{for all } \gamma_2 > 0.$$

Because of (A.5), we can establish this if we show that the function $x(y^{1/x} - 1)$ is decreasing in $x > 0$ for every fixed $y > 0$. We have

$$\frac{d}{dx} x(y^{1/x} - 1) = y^{1/x} - 1 + x\log(y) y^{1/x}\frac{-1}{x^2} = y^{1/x}\big(1 - y^{-1/x} + \log(y^{-1/x})\big) \leq 0,$$

where we applied $1 - z + \log(z) \leq 0$ for all $z > 0$, which follows from $\mathrm{e}^z \geq z\mathrm{e}$. ∎

Acknowledgment

We would like to thank two anonymous referees for valuable comments helping us improve the paper. We would also like to thank Thomas Weber and Tian Ni for helpful suggestions. Christoph Frei gratefully acknowledges financial support by the Natural Sciences and Engineering Council of Canada under grant RGPIN/402585-2011.

References

B. Biais, T. Mariotti, J.-C. Rochet and S. Villeneuve: Large Risks, Limited Liability, and Dynamic Moral Hazard. *Econometrica* **78**, 73–118, 2010.

P. Bond and A. Gomes: Multitask Principal-Agent Problems: Optimal Contracts, Fragility, and Effort Misallocations. *Journal of Economic Theory* **144**, 175–211, 2009.

A. Cadenillas, J. Cvitanić and F. Zapatero: Optimal Risk-Sharing with Effort and Project Choice. *Journal of Economic Theory* **133**, 403–440, 2007.

A. Capponi, J. Cvitanić and T. Yolcu: A Variational Approach to Contracting under Imperfect Observations. *SIAM Journal on Financial Mathematics* **3**, 605–638, 2012.

J. Cordeiro and J. Sarkis: Does Explicit Contracting Effectively Link CEO Compensation to Environmental Performance? *Business Strategy and the Environment* **17**, 304–317, 2008.

J. Cvitanić, X. Wan and J. Zhang: Optimal Compensation with Hidden Action and Lump-Sum Payment in a Continuous-Time Model. *Applied Mathematics and Optimization* **59**, 99–146, 2009.

P.M. DeMarzo, M. Fishman, Z. He and N. Wang: Dynamic Agency and the q Theory of Investment. *Journal of Finance* **67**, 2295–2340, 2012.

P. Goldsmith and R. Basak: Incentive Contracts and Environmental Performance Indicators. *Environmental and Resource Economics* **20**, 259–279, 2001.

B. Hartman-Glaser, T. Piskorski and A. Tchistyi: Optimal Securitization with Moral Hazard. *Journal of Financial Economics* **104**, 186–202, 2012.

S. He, J. Wang and J. Yan: *Semimartingale Theory and Stochastic Calculus.* Science Press, Beijing, 1992.

Z. He: Optimal Executive Compensation when Firm Size Follows Geometric Brownian Motion. *Review of Financial Studies* **22**, 859–892, 2009.

M. Hellwig and K.M. Schmidt: Discrete-Time Approximations of Holmström-Milgrom Brownian-Motion Model of Intertemporal Incentive Provision. *Econometrica* **70**, 2225–2264, 2002.

F. Hoffmann and S. Pfeil: Reward for Luck in a Dynamic Agency Model. *The Review of Financial Studies* **23**, 3329–3345, 2010.

B. Holmström and P. Milgrom: Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica* **55**, 303–328, 1987.

B. Holmström and P. Milgrom: Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, and Organization* **7**, 24–52, 1991.

M. Jensen and K.J. Murphy: Performance Pay and Top-Management Incentives. *Journal of Political Economy* **98**, 225–264, 1990.

N. Ju and X. Wan: Optimal Compensation and Pay-Performance Sensitivity in a Continuous-Time Principal-Agent Model. *Management Science* **58**, 641–657, 2012.

G. Manso: Motivating Innovation. *Journal of Finance* **66**, 1823–1860, 2011.

K.J. Murphy: Executive Compensation. *Handbook of Labor Economics* **3**, 2485–2563, North-Holland, 1999.

A. Novacovici: Linking Sustainability Performance to Compensation: A Must for Success. *Huffington Post*, published October 8, 2013, available at `http://www.huffingtonpost.com/anca-novacovici/post_5817_b_4060647.html`

H. Pagès and D. Possamaï: A Mathematical Treatment of Bank Monitoring Incentives. *Finance and Stochastics* **18**, 39–73, 2014.

G. Palast: Broken Promises and the Exxon Valdez. *Chicago Tribune*, published September 21, 1994, available at `http://articles.chicagotribune.com/1994-09-21/news/9409210193_1_exxon-valdez-alyeska-exxon-and-other-oil`

P. Protter: *Stochastic Integration and Differential Equations.* Second Edition, Volume 21 of *Stochastic Modelling and Applied Probability.* Springer, New York, 2005.

Y. Sannikov: A Continuous-Time Version of the Principal-Agent Problem. *The Review of Economic Studies* **75**, 957–984, 2008.

H. Schättler and J. Sung: The First-Order Approach to Continuous-Time Principal-Agent Problem with Exponential Utility. *Journal of Economic Theory* **61**, 331–371, 1993.

H. Schättler and J. Sung: On Optimal Sharing Rules in Discrete- and Continuous-Time Principal-Agent Problems with Exponential Utility. *Journal of Economic Dynamics and Control* **21**, 551–574, 1997.

J. Sung: Corporate Insurance and Managerial Incentives. *Journal of Economic Theory* **74,** 297–332, 1997.

R. Winter: Optimal Insurance under Moral Hazard. In: G. Dionne (ed.), *Handbook of Insurance*, Springer, 155–183, 2000.