Optimal Contracting with Dynamic Multitasking

by

Wenjie Chen

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Mathematical Finance

Department of Mathematical and Statistical Sciences
University of Alberta

Abstract

We formulate a continuous-time principal-agent model in which the agent performs two tasks: accident prevention and effort exertion. The principal can design a contract for the agent consisting of three components: a lump-sum payment, penalties when accidents occur, and continuous payments depending on the daily production outcome. A patient principal induces the agent to do more prevention and less effort as time progresses so that the principal earns the benefit from extra accident reduction net extra lump-sum payment. The principal punishes a risk-averse agent on the same level regardless of the actual accident size. The principal gives incentives for more effort and less prevention if the agent is highly risk averse to sudden payment decreases because this allows the principal to avoid a massive lump-sum payment needed to compensate the agent. When a risk-neutral agent is protected by an absolute threshold on the penalty per accident, as a form of partially limited liability, he/she is punished more for small accidents than he/she is without the protection. For a risk-neutral agent, a suitably chosen threshold in percentage of the accident costs has the same effect on the optimal task levels as an absolute threshold. However, such a link does not exist when the agent is risk averse.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Dr. Christoph Frei, who has put considerable support and guidance in my master's coursework and research. His insightful teaching in the course Computational Finance inspired my passion for not only the knowledge itself, but also honed programming skills, which proved to be an asset in my research. I deeply appreciate his immense expertise, and also importantly, his great patience in assisting me in generating research ideas, polishing mathematical proofs and formulating a formal thesis. A special thanks also goes to him for his understanding and help in my application to the future PhD program.

Besides my supervisor, I wish to thank Dr. Felipe Aguerrevere, Dr. Philippe Marcoul and Dr. Alexander Melnikov for accepting to be members of my examination committee and for making helpful comments.

I would like to thank again Dr. Alexander Melnikov and Dr. Abel Cadenillas. They instilled their massive knowledge in the courses Stochastic Calculus and Mathematical Finance, which is the workhorse of my thesis.

A very special appreciation goes to Dr. Felipe Aguerrevere and Dr. David McLean. Their passionate teaching in the courses Advanced Asset Pricing and Advanced Corporate Finance inspired my deep interest in doctoral researches in finance while their down-to-earth advice helped me achieve the milestone of entering a prestigious program.

Financial support by the Natural Sciences and Engineering Research Council of Canada under grant RGPIN/402585-2011 is gratefully acknowledged.

Table of Contents

List of Figures

Chapter 1

Introduction

According to Holmström and Milgrom (1991), "multidimensional tasks are ubiquitous in the world of business". For example, a worker in an assembly line of hammers not only assembles the two parts of hammers — the head and the handle — but also maintains the machines that produce the two parts in good condition so that unexpected breakdowns occur rarely. Likewise, in addition to making operating, financing and investing decisions, a general manager supervises and trains all the employees so that they act according to the standards of regulatory compliance and do not release highly confidential business information. In these two examples, while assembling hammers and making daily business decisions increase the efficiency of the respective projects, machinery maintenance and employee training protect the projects from large losses incurred by accidents. Indeed, in the real world of business, multiple tasks are pervasive; and more importantly, those tasks can be of different nature.

In this thesis, we analyze the dynamics of the optimal incentive provision offered by the principal to the agent in a continuous-time payment framework.

Such a continuous-time framework allows us to distinguish between daily business fluctuations and the impact of accidents. Indeed, to characterize the existence of multiple tasks and their distinctive nature, the agent is assumed to perform two tasks, each having its unique impact on the output process of the project. The first task, called "effort exertion", affects the instantaneous growth rate of the output process; if the effort exertion of the agent is higher, the project grows on average at a higher rate. The second task, called "accident prevention", prevents accidents that damage the produced outputs; if the accident prevention is higher, accidents occur on average at a lower frequency.

The pioneering work of the continuous-time principal-agent problem is Holmström and Milgrom (1987), who analyze the optimal contract with the agent exerting effort and getting paid at the end of a finite time interval. Later, many extensions and variants have been studied. For example, Schättler and Sung (1993) as well as Schättler and Sung (1997) develop and analyze first-order approaches to principal-agent problems and relate them for discrete- and continuous-time models. Ou-Yang (2003) applies a continuous-time contracting formulation to study the relation between an individual investor and a professional portfolio manager. Ju and Wan (2012) consider the problem when the manager has constant relative risk aversion rather than constant absolute risk aversion (exponential utility) in Holmström and Milgrom (1987) while Cvitanić et al. (2009) study general utility functions. A good recent overview of these models with terminal payments can be found in the book by Cvitanić and Zhang (2013).

Sannikov (2008) introduces a new framework to analyze the optimal incentive provisions with intertemporal payments to the agent, rather than a terminal payment. Using the workhorse of (stochastic) differential equations and dynamic programming, the optimal contract is designed based on the agent's continuation value. The model and techniques of Sannikov (2008) have gained significant attention. For example, DeMarzo and Sannikov (2006) relate them to a firm's capital structure, and He (2009) studies them when the firm's size follows a geometric rather than arithmetic Brownian motion. A good overview is given in the survey by Sannikov (2013). While Sannikov (2008) covers many interesting aspects such as the agent's career path, he assumes that the agent performs only one task during the project. The first paper that studies optimal contracting with multiple tasks is Holmström and Milgrom (1991), who build a single-period model and analyze the conditions for a preferable fixed wage in the optimal contract.

This thesis is based on Capponi and Frei (2015), who, in a continuous-time model, first characterize explicitly the optimal contract signed by an agent who exerts effort and undertakes prevention. In their model, the risk-averse agent gets paid at the end of the contract period and the optimal contract is sublinear in the accident component. That is, the agent is charged for a lower percentage penalty for big accidents than for small accidents. The main difference of this thesis lies in the time of payment; the principal chooses to realize the lump-sum payment and the incentives at any time throughout the whole period, rather than just at the end. Besides giving the principal more room to design the optimal contract, this also leads to a totally different form of optimal punishment for accidents. Because the risk-averse agent does not know the size of an accident before it happens and the punishment for each accident happens separately in our model, the optimal penalty-for-accident is constant in accident size.

This thesis is related to Biais et al. (2010), who capture the optimal incentives provided by an insurance company — including downsizing, investment, and compensation policies in a dynamic principal-agent model. In their paper, the only task that the agent undertakes is the costly effort, which is a binary variable. The two states of the effort indicate whether the agent works hard or shirks. As opposed to their setting, the agent in our model performs two tasks, one affecting the efficiency of the project while the other being related to accident prevention. Moreover, rather than being binary variables, both effort exertion and accident prevention in this thesis take continuous values, allowing us more room to delineate their trends over time. Another major difference is that this thesis focuses more on how parameters such as task interaction and risk-aversion level affect the optimal incentive provisions and task levels while their paper discusses more the dynamics of the project size as a result of the optimal contract. An important feature of Biais et al. (2010) is limited liability, which means that the cumulative transfers to the agent are nonnegative and nondecreasing.

Other studies have also considered contractual frameworks with infrequent shocks. For example, DeMarzo et al. (2012) consider unobservable productivity shocks, whereas Hoffmann and Pfeil (2010) study the impact of publicly observable lucky shocks.

In this thesis, the building blocks of the project outcome are two parts: a drift-diffusion process capturing the continuous component of the project and a compound Poisson process representing the randomly occurring accidents. It is a compound Poisson process because the size of the accidents is not known in advance. The agent increases the drift of the continuous component with higher effort exertion and decreases the intensity of the Poisson process with higher accident prevention. Generally, the principal has the freedom to choose whatever contract he/she wishes to compensate the agent, be it linear or nonlinear. To gain more mathematical tractability, we first assume that the contract is linear in the accident component and then develop a more general form of contractual payments. To better fit the reality, later we also protect the agent by "partially limited liability" in that the agent is only responsible for up to a certain amount (parameterized by an absolute threshold) or percentage (parameterized by a percentage threshold) of accident costs. That means, no matter how destructive an accident is, the principal can only charge the agent for at most the specified amount or percentage, which are exogenously given and reflect the general regulation severity of the industry or company. It turns out that for a linear contract signed by a risk-neutral agent, the absolute and the percentage threshold at corresponding levels have the same impact on the optimal task levels. Unfortunately, there does not exist such an interesting relation between the two thresholds for a risk-averse agent.

The rest of this thesis is organized as follows. Chapter 2 sets up the principal-agent problem in detail. In Chapter 3, we find explicitly the optimal linear contract for a risk-neutral agent while Chapter 4 develops the situation with a general nonlinear contract signed by a risk-averse agent. We explore the effects of partially limited liability in Chapter 5 and Chapter 6 concludes. All proofs are presented in the Appendix.

Chapter 2

Problem Setting

A principal hires an agent to carry out a project. The contract payments can be made at any time between time 0 and the expiration date T. The outcome of the project does not only depend on the effort that the agent has put in the project, but may also suffer from accidents. For example, unexpected accidents in the assembly line will paralyze the production of a company and lead to repair costs, no matter how hard the employees work. In this kind of project, the agent is assumed to perform two tasks: increasing the efficiency of the project and doing accident prevention at a proper level.

The agent works for the principal for the contract payments while he/she feels pain when he/she works hard. Therefore, the contract payments are the revenues that the agent earns from the contract while the pain is the cost associated with effort exertion and accident prevention. Targeting at the maximal expected profits from the contract, the agent's goal is to find the optimal levels of effort exertion (to increase the efficiency properly) and accident prevention (to avoid accidents properly).

The principal makes contract payments to the agent and receives the production outcome in the form of revenues from the project. Therefore, the principal's profit is the production outcome minus the contractual payments to the agent. The principal observes only the outcome process and the accident sizes, but neither the agent's accident prevention nor the agent's effort exertion. Aiming at the maximal expected profit, the principal's goal is to design an optimal contract that specifies penalty-for-accident and reward-for-performance as incentives to induce the agent to execute proper levels of accident prevention and effort exertion.

Outcome Process. The outcome process consists of two components. One is a continuous process characterizing the daily production activities. The other is a pure-jump process modelling the occurrence of accidents.

Continuous Component. The dynamics of the continuous component is governed by

$$dx_t = u_t \, dt + \sigma_t \, dB_t^u,$$

where $u = (u_t)_{0 \le t \le T}$ characterizes the level of effort that the agent puts in the project, B_t^u is a Brownian motion under the probability measure $\mathbb{P}^u$ defined on a measurable space $(\Omega, \mathcal{F})$, and σ_t is an exogenously determined volatility of production, related to the economic environment of the industry.[1] The principal only observes x_t and does not know u_t, hence he/she does not know B_t^u, either. Mathematically, this is captured by modelling B_t^u as Brownian motion under a probability measure $\mathbb{P}^u$ depending on the chosen effort level u_t hidden to the principal.

[1] To ensure that the following computations are mathematically sound, we assume that the function σ_t is positive and bounded away from zero and infinity, i.e., there exist constants $K \ge k > 0$ such that $K \ge \sigma_t \ge k$.

Jump component. The damage of accidents reduces the outcome process. The cumulative sum of accident damage is represented by the compound Poisson process

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

defined on the same probability space, where (Y_i) is a sequence of bounded nonnegative i.i.d. random variables with average value $\mathbb{E}[Y_i] = m$. The random variable Y_i measures the size of accident i and the average size of accidents are assumed to be m. The counting process N_t jumps by one at time t if an accident occurs at that time t.

Hence, the outcome process is

$$X_t = x_t - J_t, \quad 0 \leq t \leq T,$$

i.e., the outcome process is the cumulative continuous component net the effect of accidents. We assume that B_t^u and J_t are independent, and the agent's information flow is given by the filtration $(\mathcal{F}_t)$ generated by the processes B_t^u and J_t.

Effort Exertion. The agent affects the continuous component by choosing the level of u_t, which is $(\mathcal{F}_t)$-adapted. Since the principal can only observe x_t at time t without knowing the value of B_t^u, he/she is unable to infer the precise value of u_t. Therefore, he/she can only design a contract based on the outcome process. By choosing a higher u_t, the agent increases the efficiency of the project at a cost of increased pain.

Accident Prevention. Additionally, the agent affects the outcome process by choosing the intensity λ_t of the process N_t at which accidents happen, where λ_t is predictable with respect to the filtration $(\mathcal{F}_t)$. The presence of hidden action lies in the fact that the agent, by choosing λ_t, can only affect the frequency with which accidents occur. Since more accidents can possibly occur with a lower λ_t, the principal cannot infer the precise value of λ_t by observing the number of accidents. However, by observing the number of accidents over time, the principal can make an estimation of the value of λ_t. By choosing a lower λ_t, the agent increases the prevention level $\frac{1}{\lambda_t}$ at a cost of increased pain.

Pain Function. The agent exerts effort and prevents accidents at a cost of pain $P(\lambda_t, u_t)$, where $P : (0, \infty) \times [0, \infty) \to [0, \infty)$ is a twice continuously differentiable function with basic properties $\frac{\partial P}{\partial \lambda} < 0$, $\frac{\partial P}{\partial u} > 0$ and convexity in λ and u as well as suitable properties for the limit behavior; see Appendix A.2 for details.

Chapter 3

A Risk-neutral Agent

3.1 Problem Formulation

The Agent. The agent maximizes his/her expected profits from the contract. Specifically, he/she enjoys the contract payments L_t as the revenues at costs of pain $P(\lambda_t, u_t)$ resulting from hard work. The agent is risk-neutral in the sense that he/she only cares about the expected value regardless of the potential risk. He/she cherishes time and the associated coefficient of time value is denoted by ρ. Therefore, the agent has the objective

$$\sup_{\lambda,u} \mathbb{E}\left[\int_0^T \mathrm{e}^{-\rho s}(dL_s - P(\lambda_s, u_s)\, ds) \right], \tag{3.1}$$

where $\mathrm{e}^{-\rho s}$ is the discount factor associated with the coefficient of time value ρ.

The Principal. The principal maximizes his/her expected profits from the project. Specifically, he/she enjoys the output $X_t = x_t - J_t$ of the project as the revenues while has to pay the agent the contract payment L_t as the costs. The principal is risk-neutral. He/she also cherishes time and the associated coefficient of time value is denoted by r. The parameter r may be different from ρ because the two players may have different views of time value. This difference plays a very important role in determining the time trends of the two actions and the corresponding incentives as well as the order of priority of the two actions.

Therefore, the principal has the objective

$$\sup_L \mathbb{E}\Big[\int_0^T \mathrm{e}^{-rs}(dx_s - dJ_s - dL_s)\Big], \tag{3.2}$$

where e^{-rs} is the discount factor associated with the coefficient of time value r. Moreover, the agent has some minimal requirements to enter the contract. Therefore, the principal needs to satisfy the least utility of the agent when the agent executes optimal levels of accident prevention and effort exertion. Hence, the principal has the constraint

$$\sup_{\lambda,u} \mathbb{E}\Big[\int_0^T \mathrm{e}^{-\rho s}(dL_s - P(\lambda_s, u_s)\,ds)\Big] \geq R_0, \tag{3.3}$$

where R_0 is the agent's reservation utility, measuring the least utility that the agent needs to enjoy in order to enter the contract.

Contract Space. The presence of hidden actions restricts the principal in the sense that he/she can design the contract only based on the observable performance of the agent, i.e., the realized continuous output x_t and the accidents J_t. Moreover, just before time t, the contract needs to specify the reward for dx_t, the infinitesimal increase of the continuous output as well as the penalty for $dJ_t = J_t - J_{t-}$, the possible accident that may occur at time t.

Therefore, we assume that the payment stream L_t is driven by

$$dL_t = \alpha_t \, dJ_t + \gamma_t \, dx_t + \beta_t \, dt + dA_t, \quad 0 \le t \le T.$$

In general, $\alpha = (\alpha_t)_{0 \le t \le T}$, $\beta = (\beta_t)_{0 \le t \le T}$, $\gamma = (\gamma)_{0 \le t \le T}$ could be stochastic processes, for the principal can choose whatever form of the contract he/she wants, based on the information available from x_t and J_t. Yet, in order to have a tractable problem, we analyze it with the assumption that α, β, γ are deterministic. Moreover, A_t is a deterministic pure-jump process that models the lump-sum payment as part of the contract payments to the agent. We impose that β_t has the same sign over the entire period and that A_t is monotonic. This excludes round-trip payments between principal and agent, which would lead to arbitrage opportunities as the following remark explains.

Remark: To have reasonable contracts, payments of the following form should be ruled out. Assume that $\rho > r$, and that the agent receives a payment of $\mathrm{e}^{-\rho T}$ at time 0 and loses 1 at time T. This means that the principal pays the agent $\mathrm{e}^{-\rho T}$ at time 0 and gets 1 from the agent at time T. Although this leads to a zero-sum result for the agent, there is a positive net present value for the principal: $1 * \mathrm{e}^{-rT} - \mathrm{e}^{-\rho T} = \mathrm{e}^{-rT} - \mathrm{e}^{-\rho T} > 0$, creating an arbitrage opportunity. In the case $\rho < r$, a similar arbitrage opportunity could be constructed. To avoid such arbitrage opportunities, we impose that β_t has the same for all t and that A_t is monotonic.

3.2 Incentive Compatibility

For payment stream L, intensity λ and effort level u, the continuation utility of the agent is given by

$$W_t(L,\lambda,u) = \mathbb{E}\bigg[\int_t^T \mathrm{e}^{-\rho(s-t)}(dL_s - P(\lambda_s,u_s)\,ds)\bigg|\mathcal{F}_t\bigg].$$

Hence the conditional expected utility U_t of the agent is

$$\begin{aligned} U_t(L,\lambda,u) &= \mathbb{E}\bigg[\int_0^T \mathrm{e}^{-\rho s}(dL_s - P(\lambda_s,u_s)\,ds)\bigg|\mathcal{F}_t\bigg] \\ &= \mathrm{e}^{-\rho t}W_t(L,\lambda,u) + \int_0^t \mathrm{e}^{-\rho s}(dL_s - P(\lambda_s,u_s)\,ds). \end{aligned}$$

Because U_t is a martingale, there exist predictable processes $H(L,\lambda,u)$ and $K(L,\lambda,u)$ such that

$$U_t(L,\lambda,u) = U_0(L,\lambda,u) - \int_0^t \mathrm{e}^{-\rho s}H_s(L,\lambda,u)\,dM_s^\lambda - \int_0^t \mathrm{e}^{-\rho s}K_s(L,\lambda,u)\,dB_s^u, \tag{3.4}$$

where $dM_s^\lambda = dJ_s - \lambda_s m\,ds$. Hence, the evolution of the continuation utility is given by

$$\begin{aligned} dW_t(L,\lambda,u) = {} & [\rho W_t(L,\lambda,u) + P(\lambda_t,u_t)]\,dt - dL_t \\ & + H_t(L,\lambda,u)[\lambda_t m\,dt - dJ_t] - K_t(L,\lambda,u)\,dB_t^u. \end{aligned} \tag{3.5}$$

Equation (3.5) states that net of the wage, the instantaneous expected relative change in the continuation utility of the agent is the discount factor ρ, provided that he/she performs little prevention and effort. Moreover, $H_t(L,\lambda,u)$ is the sensitivity of the loss of his/her continuation utility to the impact of unpredictable accidents.

Proposition 3.1: *The sufficient and necessary conditions for the intensity $\lambda = (\lambda_t)_{0\le t\le T}$ and effort $u = (u_t)_{0\le t\le T}$ to be incentive compatible are*

$$P_\lambda(\lambda_t, u_t) = -mH_t(L, \lambda_t, u_t) \text{ and } P_u(\lambda_t, u_t) = -K_t(L, \lambda_t, u_t). \tag{3.6}$$

Proposition 3.1 states that if the agent's goal is to maximizes his/her expected profits from the contract, he/she should work on such levels that the marginal pains from either of the actions is equal to the marginal gains in his/her utility from that action. The proof of Proposition 3.1 is given in the Appendix A.1.

3.3 Optimal Contracting

Recall from (3.4) that

$$U_t(L, \lambda, u) = U_0(L, \lambda, u) - \int_0^t \mathrm{e}^{-\rho s} H_s(L, \lambda, u)\, dM_s^\lambda - \int_0^t \mathrm{e}^{-\rho s} K_s(L, \lambda, u)\, dB_s^u.$$

On the other hand,

$$U_T(L, \lambda, u) = \int_0^T \mathrm{e}^{-\rho s}(dL_s - P(\lambda_s, u_s)\, ds).$$

Hence,

$$\begin{aligned}\int_0^T \mathrm{e}^{-\rho s}(dL_s - P(\lambda_s, u_s)\, ds) = & -\int_0^T \mathrm{e}^{-\rho s} H_s(L, \lambda, u)\, dM_s^\lambda + U_0(L, \lambda, u) \\ & - \int_0^t \mathrm{e}^{-\rho s} K_s(L, \lambda, u)\, dB_s^u.\end{aligned}$$

Moreover, assume that the principal has the correct information of the agent's reservation utility R_0, then he/she will always choose a contract such that (3.3) holds with equality, implying $\mathbb{E}[U_T(L,\lambda)] = R_0$. Therefore,

$$U_0(L,\lambda,u) - \mathbb{E}\left[\int_0^T \mathrm{e}^{-\rho s} H_s(L,\lambda,u)\, dM_s^\lambda + \int_0^t \mathrm{e}^{-\rho s} K_s(L,\lambda,u)\, dB_s^u\right] = R_0.$$

Since the processes $\int_0^t \mathrm{e}^{-\rho s} H_s(L,\lambda,u)\, dM_s^\lambda$ and $\int_0^t \mathrm{e}^{-\rho s} K_s(L,\lambda,u)\, dB_s^u$ are martingales under integrability conditions on $H_s(L,\lambda,u)$ and $K_s(L,\lambda,u)$, the above equation implies that $U_0(L,\lambda,u) = R_0$. Therefore,

$$\begin{aligned}\int_0^T \mathrm{e}^{-\rho s}(dL_s - P(\lambda_s,u_s)\, ds) = \; & -\int_0^T \mathrm{e}^{-\rho s} H_s(L,\lambda,u)\,(dJ_s - m\lambda_s\, ds) + R_0 \\ & -\int_0^T \mathrm{e}^{-\rho s} K_s(L,\lambda,u)\, dB_s^u.\end{aligned}$$

Plug $dL_s = \alpha_t\, dJ_t + \gamma_t\, dx_t + \beta_t\, dt + dA_t$ into the above equation. By the uniqueness of martingale representation, the optimal contract satisfies $\alpha_s = -H_s$ and $\beta_s = -K_s$. Therefore, we obtain the following result.

Theorem 3.1: *The coupled stochastic optimal contracting problem* (3.1)–(3.3) *is equivalent to the deterministic optimization problem*

$$\sup_{\alpha,\beta,\gamma,A} \int_0^T \mathrm{e}^{-rs}[(1-\gamma_s)u_s - (1+\alpha_s)\lambda_s m - \beta_s]\, ds - \int_0^T \mathrm{e}^{-rs} dA_s$$

subject to

$$\int_0^T \mathrm{e}^{-\rho s}[\alpha_s\lambda_s m + \beta_s + \gamma_s u_s - P(\lambda_s,u_s)]\, ds + \int_0^T \mathrm{e}^{-\rho s} dA_s = R_0,$$

$P_\lambda(\lambda_s,u_s) = \alpha_s m$ *and* $P_u(\lambda_s,u_s) = \gamma_s$.

The magnitude of the discount factors matters in solving the above problem. When the principal has a different view of time value than the agent, the lump-sum payment should be made such that the principal loses as least as he/she could while the agent gains the potentially most. Therefore, all deterministic payments should be made at either the beginning or the end of the contract depending on the relation between ρ and r. If $\rho > r$, the nonnegative lump-sum payment is made at time 0, the very beginning of the contract. If $\rho < r$, it is made at time T, the expiration date. In either case, β_s is chosen to be zero because intermediate payments are always suboptimal compared to the payments at the beginning or end.

We can bring Theorem 3.1 to the following form.

Proposition 3.2: *The optimal λ_s^* and u_s^* are the maximizers of*

$$\begin{aligned} G_s(\lambda_s, u_s) = \mathrm{e}^{-rs}[u_s - u_s P_u(\lambda_s, u_s) - m\lambda_s - \lambda_s P_\lambda(\lambda_s, u_s)] \\ + \mathrm{e}^{-\rho s}[\lambda_s P_\lambda(\lambda_s, u_s) + u_s P_u(\lambda_s, u_s) - P(\lambda_s, u_s)]. \end{aligned} \tag{3.7}$$

For $a = R_0 - \int_0^T \mathrm{e}^{-\rho t}[P_\lambda(\lambda_t^*, u_t^*)\lambda_t^* + P_u(\lambda_t^*, u_t^*)u_t^* - P(\lambda_t^*, u_t^*)]\, dt$, the optimal lump-sump payment is as follows:

- *If $\rho = r$, optimal β_s^* and A_s^* satisfy $A_T^* + \int_0^T \beta_s^*\, ds = a$, but they are not uniquely determined.*

- *If either $\rho > r$ and $a \geq 0$ or $\rho < r$ and $a < 0$, the optimal $A_s^* = a$ is constant, and the optimal β_s^* is identical to zero.*

- *If either $\rho < r$ and $a \geq 0$ or $\rho > r$ and $a < 0$, the optimal A_s^* is zero for all $s < T$ and $A_T^* = \mathrm{e}^{\rho T} a$, and the optimal β_s^* is identical to zero.*

Notice that the term a is the reservation utility net the amount that the agent earns from his/her two actions. We have $a \geq 0$ in the usual case, generating a lump-sum payment (serving as a "deposit") made at the beginning of the contract if the principal is more patient. However, in the case when $a < 0$, the agent is making on average more than what he/she is asking in terms of the reservation utility. Consequently, the principal will take away the lump-sum payment to benefit himself/herself. In this case, the lump-sum payment is retrieved at the expiration date T if the principal is more patient.

The incentives for the agent's proper actions are reflected in the terms α_t and γ_t, which punish the agent for more accidents while reward him/her for more continuous output. To create an optimal contract, the principal needs to figure out the desired levels of accident prevention $1/\lambda_t^*$ and effort exertion u_t^* for $0 < t < T$ so that he/she can design the optimal incentives α_t^* and γ_t^*. The optimal lump-sum payment is determined at last to adjust the agent's expected utility to his/her reservation utility R_0.

Proposition 3.3: *In the case $\rho > r$, an optimal contract exists if*

$$\mathop{lim\,sup}_{\frac{u_s}{\lambda_s}\to\infty} \frac{(\mathrm{e}^{-\rho s} - \mathrm{e}^{-rs})\lambda_s P_\lambda(\lambda_s, u_s) + \mathrm{e}^{-rs} u_s}{\mathrm{e}^{-\rho s} P(\lambda_s, u_s) - (\mathrm{e}^{-\rho s} - \mathrm{e}^{-rs}) u_s P_u(\lambda_s, u_s)} < 1. \tag{3.8}$$

The proof of Proposition 3.3 is given in the Appendix A.2.

Theorem 3.2: *The optimal actions $\lambda^* = (\lambda_t^*)_{0\leq t<T}$ and $u^* = (u_t^*)_{0\leq t<T}$ are the maximizers of the function* (3.7)*. The optimal penalty-for-accident sensitivity $\alpha^* = (\alpha_t^*)_{0\leq t<T}$ and reward-for-performance sensitivity $\gamma^* = (\gamma_t^*)_{0\leq t<T}$ are*

$$\alpha_t^* = \frac{1}{m} P_\lambda(\lambda_t^*, u_t^*) \text{ and } \gamma_t^* = P_u(\lambda_t^*, u_t^*).$$

3.4 Numerical Illustrations

To gain a more concrete capture of the effect of task interaction on the optimal contact, we provide here a numerical analysis of the model. We choose the same pain function

$$P(\lambda, u) = (k_1 u)^2 + (k_2/\lambda)^2 + 2\theta k_1 k_2 u/\lambda$$

with $\theta > 0$ as in Capponi and Frei (2015) to compare the results in settings with intermediate versus terminal payments. Here, k_1 and k_2 are the coefficients that convert the pain into monetary value while θ corresponds to the level of task interaction. Notice that a higher θ implies that an increase in accident prevention $1/\lambda$ will cause more pain than in the case of a smaller θ. This property indicates that the two actions are substitutes for each other; facing the different consequences of the actions, the agent has to put more resources in the more critical one in his/her decision.

Moreover, the difference between the nature of the two actions lies in the fact that effort exertion only increases the drift of the continuous output while accident prevention will avoid sudden loss of it. This difference in nature will be reflected when the principal has a different view of time value than the agent. If $\rho > r$, the principal is more patient than the agent. Hence future values for him/her will not be discounted a lot, indicating that the costs incurred by future accidents matter much to the principal. Therefore, he/she will design a contract that induces the agent to shifts more resources to accident prevention from effort exertion. If $\rho < r$, then the principal discounts the cost incurred by future accidents more than the agent. Therefore, he/she will design a contract that induces the agent to shift more resources to effort exertion from accident prevention, aiming to earn more from higher efficiency (higher drift u_s) from the continuous output component. In this sense, accident prevention is more critical than effort exertion when $\rho > r$ while effort exertion is the priority in the reverse case.

We choose $k_1 = k_2 = m = 1$ in the numerical illustrations. We consider the situation where the principal is more patient than the agent in four different scenarios: i) $r = 0.05$ and $\rho = 0.07$, ii) $r = 0.08$ and $\rho = 0.12$, iii) $r = 0.14$ and $\rho = 0.20$, iv) $r = 0.22$ and $\rho = 0.30$.

The Effects of Task Substitution. Figure 3.1 illustrates the relationships of the optimal effort exertion u_s^* and accident prevention $\frac{1}{\lambda^*}$ with task interaction θ. The optimal u_s^* decreases with θ until level 0 is reached while the optimal $\frac{1}{\lambda^*}$ has a U-shape before it levels off. The effects of higher θ are twofold. Firstly, higher task interaction implies overall higher pain for the agent. Therefore, the cumulative accident prevention and effort exertion are decreasing for fixed other factors. Secondly, since now prevention is more critical ($\rho > r$), higher θ makes the principal design a contract such that the agent shifts more resources to accident prevention from effort exertion. Either of these two effects decreases the optimal u_s^* until the level 0 is reached. The first effect dominates at lower θ, decreasing the optimal $\frac{1}{\lambda^*}$ while the second effect dominates at higher θ, increasing the optimal $\frac{1}{\lambda^*}$. The optimal $\frac{1}{\lambda^*}$ levels off since when $u^* = 0$, indicating that the agent pools all the resources in accident prevention. Moreover, this effect is stronger for higher difference between ρ and r.

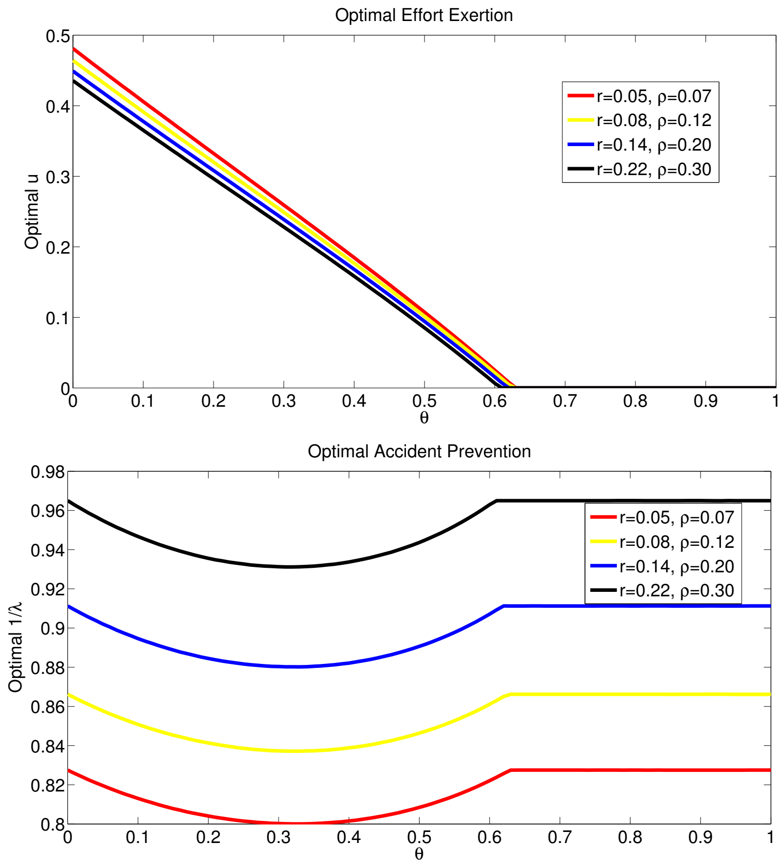


Figure 3.1: The Effects of Task Substitution

The Time Trend of Task Levels. Figure 3.2 illustrates the trend of optimal accident prevention $\frac{1}{\lambda^*}$ and effort exertion u_s^* over time. The pain from accident prevention not only affects the agent himself/herself, but also the principal in the following sense. The principal has to compensate the agent with a lump-sum payment for the pain incurred so that the agent still enjoys the reservation utility R_0. Therefore, the effects of overly-high pain from accident prevention on the principal are twofold: reduction in the average accident frequency (benefit for the principal) and increase in the lump-sum payment (cost for the principal). The average accident reduction is discounted by e^{-rs} while the lump-sum payment is discounted by $\mathrm{e}^{-\rho s}$. Since $\rho > r$, it is optimal for the principal to design a contract that induces the agent to increase the prevention level over time. Since the agent has limited resources and one task is substitute for the other, the increasing level of accident prevention will be accompanied by a decreasing level of effort exertion.

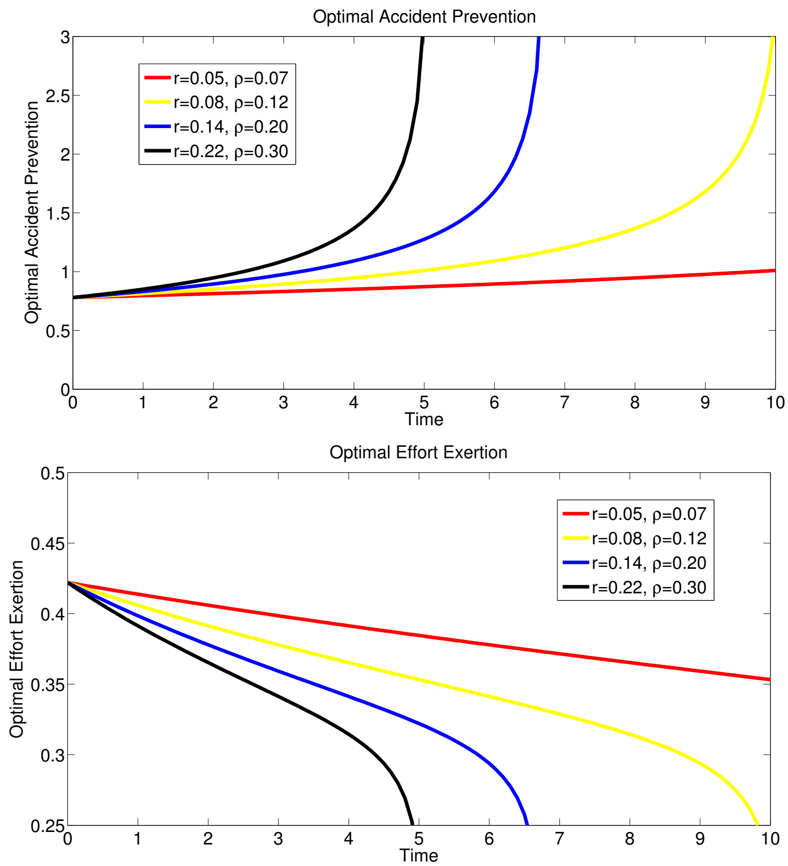


Figure 3.2: The Time Trend of Task Levels

The Optimal Contract. Figure 3.3 illustrates the incentives in the optimal contract: the penalty-for-accident sensitivity α^* and the reward-for-performance sensitivity γ^*. To induce the agent to increase the prevention level, the principal creates a contract such that the penalty-for-accident increases over time, implying that later accidents will be punished more severely. At time 0, the gap between the patience of the principal and agent has no effect. Since the agent is risk neutral, the contract actually behaves as if the principal sells the entire project ($\gamma_0^* = 1$) to the agent at the best price for an infinitesimal moment at time 0. After time 0, the reward-for-performance sensitivity is just a part of the contract payment. Therefore, it decreases below 1 for a reasonable time interval. Yet, since the agent shifts effort exertion to the more critical task of accident prevention as time progresses, the principal needs to give additional incentives in the form of increased reward-for-performance sensitivity so that the agent does not reduce too much effort exertion as substitution for a higher accident prevention.

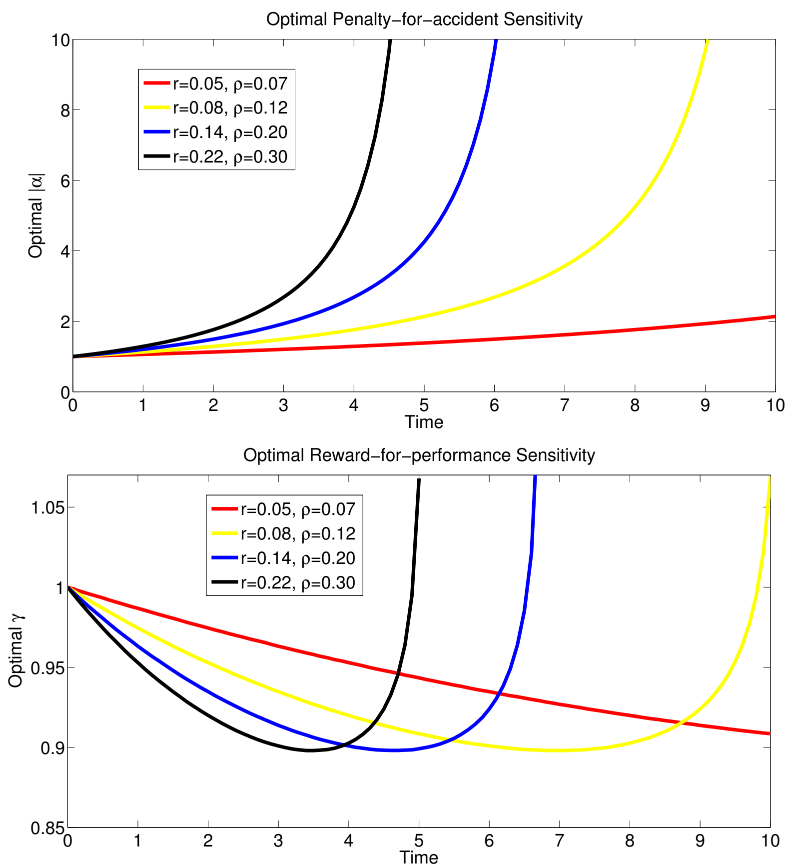


Figure 3.3: The Optimal Contract

Principal's Value Function. We define the principal's value function

$$V(t) = \sup_L \mathbb{E}\left[\int_t^T \mathrm{e}^{-r(s-t)}(dx_s - dJ_s - dL_s)\middle|\mathcal{F}_t\right], \quad 0 \le t \le T,$$

which shows the wealth path of the principal when the optimal contract is executed. When the principal designs the contract, he/she considers not only what the optimal contract is, but also how much he/she profits from the project when the agent performs the optimal actions.

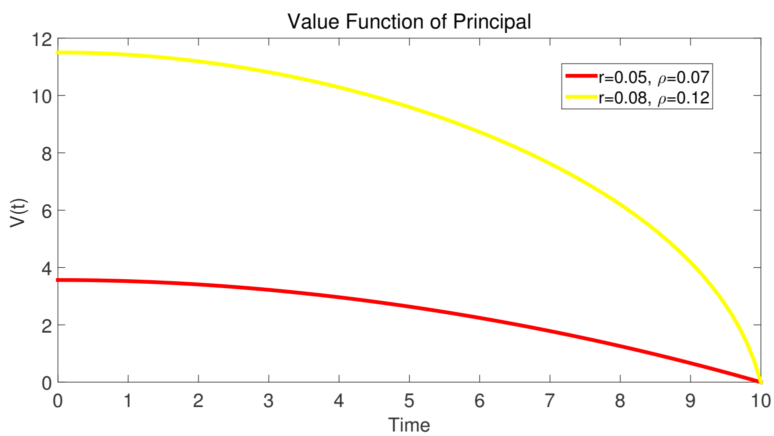


Figure 3.4: The Time Trend of Value Function of Principal

Figure 3.4 illustrates the trends of the principal's value function. Caring more about the value related to the optimal incentives, we choose the reservation utility R_0 such that the lump-sum payment is 0. Attention should be paid that we only depicts the trends of $\rho = 0.07, 0.12$ and $r = 0.05, 0.08$, where the optimal actions satisfy (3.8) for the entire period. If the principal acts such that the lump-sum payment is avoided, his/her value is always nonnegative, decreases over time and reaches 0 at the terminal time. The concavity of $V(t)$ is related to the discounting of future benefits: as t gets bigger (closer to T), the discounting of future benefits becomes less important and an increase in t leading to a shorter period for making future benefits results in a bigger decrease in $V(t)$. Mathematically, this can be seen as follows: Consider two pairs of time points (t_1, t_2) and $(\tilde{t}_1, \tilde{t}_2)$ with $t_1 < t_2$, $\tilde{t}_1 < \tilde{t}_2$, $t_2 < \tilde{t}_2$ and $t_2 - t_1 = \tilde{t}_2 - \tilde{t}_1$. To show concavity of $V(t)$, we will derive $V(t_1) - V(t_2) < V(\tilde{t}_1) - V(\tilde{t}_2)$. Since our contract optimization does not depend on the starting point, we have

$$V(t_1) = \sup_L \mathbb{E}\left[\int_{t_2}^{T+t_2-t_1} \mathrm{e}^{-r(s-t_2)}(dx_s - dJ_s - dL_s) \middle| \mathcal{F}_{t_2} \right].$$

Because the optimal actions and sensitivities do not depend on the time horizon, we obtain

$$V(t_1) - V(t_2) = \sup_L \mathbb{E}\left[\int_{T}^{T+t_2-t_1} \mathrm{e}^{-r(s-t_2)}(dx_s - dJ_s - dL_s) \middle| \mathcal{F}_{t_2} \right]$$

while

$$\begin{aligned} V(\tilde{t}_1) - V(\tilde{t}_2) &= \sup_L \mathbb{E}\left[\int_{T}^{T+\tilde{t}_2-\tilde{t}_1} \mathrm{e}^{-r(s-\tilde{t}_2)}(dx_s - dJ_s - dL_s) \middle| \mathcal{F}_{\tilde{t}_2} \right] \\ &= \sup_L \mathbb{E}\left[\int_{T}^{T+t_2-t_1} \mathrm{e}^{-r(s-\tilde{t}_2)}(dx_s - dJ_s - dL_s) \middle| \mathcal{F}_{\tilde{t}_2} \right]. \end{aligned}$$

Therefore, if the additional optimal benefits $(dx_s - dJ_s - dL_s)$ are positive in expectation, $t_2 < \tilde{t}_2$ leads to $V(t_1) - V(t_2) < V(\tilde{t}_1) - V(\tilde{t}_2)$, which implies the concavity of $V(t)$, as we observe it in Figure 3.4.

Chapter 4

A Risk-averse Agent

4.1 Problem Formulation

Unlike the risk-neutral agent, a risk-averse one maximizes his/her expected profit subject to bearable risk resulting from the accidents and the uncertainty of the continuous output component. This nature of risk-aversion affects the optimal contract; since the principal needs to compensate the agent by the lump-sum payment for the painful tasks, he/she will avoid the extra cost incurred by the relatively more painful one while induce the agent to perform the other. This strategy, which we call "induced task shift", will be illustrated in Section 4.4.

The principal has the freedom to design the contract in any form. Hence, a linear contract as given in Chapter 3 is not necessarily optimal in general. Accordingly, a nonlinear contract, regarded as a generalized form of that in Chapter 3, is provided and analyzed in this chapter. Perhaps surprisingly, the penalty-for-accident strategy in such a nonlinear contract is constant in accident size, implying that the principal punishes the agent on the same level regardless of the size of accidents.

The Contract Space. Since the principal has the discretion to create whatever form of the contract using information from the output process, the payment stream is not necessarily linear in accident costs and the continuous output component. Yet, the contract should still give proper incentives to induce desirable performance of the agent. Specifically, it should depend on the magnitude of accidents and the continuous output component. Therefore, the payment stream is assumed to be governed by

$$dL_t = g(t, dJ_t) + f(t)\,dx_t + \beta_t\,dt + dA_t, 0 \le t \le T$$

where g and f are deterministic functions. We impose that β_t has the same sign over the entire period and that A_t is monotonic for the same reason as in the remark on page 12.

The Agent. Now the agent is risk-averse so that he/she also cares about the risk associated with the profits. This nature of risk aversion is reflected in two concave functions $U_1 : (-\infty, 0] \to (-\infty, 0]$ and $U_2 : [0, \infty) \to [0, \infty)$ that measure how unhappy he/she is when facing the costs of accidents and the uncertainty of the continuous output component. Therefore, the agent solves the problem

$$\sup_{\lambda,u} \mathbb{E}\left[\int_0^T \mathrm{e}^{-\rho s}[U_1(g(s, dJ_s)) + U_2(f(s))\,dx_s + (\beta_s - P(\lambda_s, u_s))\,ds + dA_s]\right]. \quad (4.1)$$

Notice that the lump-sum payment A_t only serves as a “deposit” for the agent while the incentives $g(s, dJ_s)$ and $f(s)$ enable one to analyze the agent’s rational reaction to risk. In order to model this different nature of lump-sum payment and incentives, we only characterize the risk aversion by U_1 and U_2 on $g(s, dJ_s)$ and $f(s)$ while leave dA_s as is. Thus, one should interpret them as "relative utilities" as opposed to the utility from the lump-sum payment.

Remark: We separate the utility of the agent into U_1 and U_2 for two reasons. Firstly, we can derive mathematically such a separation under suitable assumptions because of the different nature of the underlying stochastic processes. Indeed, assume that the instantaneous change in the utility of the agent is dU_t at time t and that U_t is right-continuous with left limits. Then we can separate the jump and continuous components by writing $dU_t = \Delta U_t + dU_t^c$. Since the dynamics of the utility U_t is driven only by the contract payment L_t, the utility U_t jumps precisely when L_t jumps and U_t evolves continuously when L_t evolves continuously. Hence, ΔU_t should be a function of the jump-related compensation $g(s, dJ_s)$ while dU_t^c should depend on the continuous part $f(s)\, dx_s$. Therefore, we get $dU_t = U_1(g(t, dJ_t)) + U_2(f(t))\, dx_t$ for some functions U_1 and U_2. Economically, the separation results from the distinctive nature of the two tasks. While effort exertion increases the daily operation of the project (a continuous path), accident prevention avoids sudden huge losses (a jump process). Therefore, the utilities derived from the associated incentives are distinguishable for the agent. In conclusion, we formulate the agent's problem as (4.1).

The Principal. The principal still maximizes his/her expected profits subject to the constraint that he/she needs to keep the agent stay in the contract.

Therefore, he/she solves the problem

$$\sup_{L} \mathbb{E}\bigg[\int_0^T \mathrm{e}^{-rs}(dx_t - dJ_s - dL_s)\bigg] \tag{4.2}$$

subject to

$$\sup_{\lambda,u} \mathbb{E}\bigg[\int_0^T \mathrm{e}^{-\rho s}[U_1(g(s,dJ_s)) + U_2(f(s))\,dx_s + (\beta_s - P(\lambda_s,u_s))\,ds + dA_s]\bigg] \geq R_0. \tag{4.3}$$

4.2 Incentive Compatibility

For payment stream L, intensity λ and effort level u, we set

$$U_T = \int_0^T \mathrm{e}^{-\rho s}[U_1(g(s,dJ_s)) + U_2(f(s))\,dx_s + (\beta_s - P(\lambda_s,u_s))\,ds + dA_s].$$

Then for optimal λ^* and u^*, we have $\mathbb{E}[U_T(L,u^*,\lambda^*)] = R_0$ by (4.3) under the assumption that the principal has full information of the agent's reservation utility. Consequently,

$$\int_0^T \mathrm{e}^{-\rho s}(\mathbb{E}[U_1(g(s,dJ_s^*))] + \mathbb{E}[U_2(f(s))\,dx_s^*] + (\beta_s - \mathbb{E}P(\lambda_s^*,u_s^*))\,ds + dA_s) = R_0.$$

Notice that $\mathbb{E}[U_1(g(s,dJ_s^*))] = \mathbb{E}[\lambda_s^*]\,\mathbb{E}[U_1(g(s,Y_1))]\,ds$ and $\mathbb{E}[U_2(f(s))\,dx_s^*] = \mathbb{E}[U_2(f(s))(u_s^*\,ds + \sigma_s\;dB_s^*)] = U_2(f(s))\mathbb{E}[u_s^*]\,ds$. Therefore,

$$\begin{aligned}\int_0^T \mathrm{e}^{-\rho s}(\beta_s\,ds + dA_s) = R_0 &- \int_0^T \mathrm{e}^{-\rho s}\mathbb{E}[\lambda_s^*]\mathbb{E}[U_1(g(s,Y_1))]\,ds \\ &- \int_0^T \mathrm{e}^{-\rho s}U_2(f(s))\mathbb{E}[u_s^*]\,ds + \int_0^T \mathrm{e}^{-\rho s}\mathbb{E}[P(\lambda_s^*,u_s^*)]\,ds.\end{aligned}$$

Hence, for non-optimal λ and u,

$$
\begin{aligned}
U_T(L,\lambda,u) = &\int_0^T \mathrm{e}^{-\rho s}[U_1(g(s,dJ_s)) + U_2(f(s))\,dx_s + (\beta_s - P(\lambda_s,u_s))\,ds + dA_s] \\
&= R_0 + \int_0^T \mathrm{e}^{-\rho s}[U_1(g(s,dJ_s)) + U_2(f(s))\,dx_s - P(\lambda_s,u_s)\,ds] \\
&\quad - \int_0^T \mathrm{e}^{-\rho s}\mathbb{E}[\lambda_s^*]\mathbb{E}[U_1(g(s,Y_1))]\,ds - \int_0^T \mathrm{e}^{-\rho s}U_2(f(s))\mathbb{E}[u_s^*]\,ds \\
&\quad + \int_0^T \mathrm{e}^{-\rho s}\mathbb{E}[P(\lambda_s^*,u_s^*)]\,ds \\
&= R_0 + \int_0^T \mathrm{e}^{-\rho s}(U_1(g(s,dJ_s)) - \lambda_s\mathbb{E}[U_1(g(s,Y_1))]\,ds) \\
&\quad + \int_0^T \mathrm{e}^{-\rho s}(\lambda_s\mathbb{E}[U_1(g(s,Y_1))] - \mathbb{E}[\lambda_s^*]\mathbb{E}[U_1(g(s,Y_1))])\,ds \\
&\quad + \int_0^T \mathrm{e}^{-\rho s}(U_2(f(s))u_s - U_2(f(s))\mathbb{E}[u_s^*])\,ds \\
&\quad + \int_0^T \mathrm{e}^{-\rho s}(\mathbb{E}[P(\lambda_s^*,u_s^*)] - P(\lambda_s,u_s))\,ds \\
&\quad + \int_0^T \mathrm{e}^{-\rho s}U_2(f(s))\sigma_s\,dB_s^u.
\end{aligned}
$$

Therefore, the incentive compatibility conditions are given by the first-order conditions.

Proposition 4.1: *The sufficient and necessary conditions for the intensity $\lambda = (\lambda_t)_{0\le t\le T}$ and effort $u = (u_t)_{0\le t\le T}$ to be incentive compatible are*

$$P_\lambda(\lambda_t,u_t) = \mathbb{E}[U_1(g(t,Y_1))] \text{ and } P_u(\lambda_t,u_t) = U_2(f(t)).$$

The intuition for Proposition 4.1 is the same as that for Proposition 3.1.

Remark: The incentive compatibility conditions imply that the optimal $\lambda^* = (\lambda_s^*)_{0\le s\le T}$ and $u^* = (u_s^*)_{0\le s\le T}$ are deterministic.

4.3 Optimal Contracting

Now we solve for the optimal contract. Since $\mathbb{E}[g(s, dJ_s)] = \lambda_s \mathbb{E}[g(s, Y_1)]\, ds$, $\mathbb{E}[dB_s] = 0$, $\mathbb{E}[dx_s] = u_s\, ds$ and $\mathbb{E}[dJ_s] = m\lambda_s\, ds$, the principal's objective function (4.2) is reduced to

$$\sup_{g,f,\beta,A} \int_0^T \mathrm{e}^{-rs}\big(- \lambda_s(m + \mathbb{E}[g(s, Y_1)])\, ds + (u_s - f(s)u_s - \beta_s)\, ds\big) - \int_0^T \mathrm{e}^{-rs} dA_s.$$

Using $\mathbb{E}[U_2(f(s))(u_s\, ds + \sigma_s\, dB_s)] = u_s U_2(f(s))\, ds$ and $\mathbb{E}[U_1(g(s, dJ_s))] = \lambda_s \mathbb{E}[U_1(g(s, Y_1))]\, ds$, we can reduce the agent's reservation utility condition (4.3) to

$$\begin{aligned}&\sup_{\lambda,u} \int_0^T \mathrm{e}^{-\rho s}[\lambda_s \mathbb{E}[U_1(g(s, Y_1))]\, ds + U_2(f(s))u_s\, ds + (\beta_s - P(\lambda_s, u_s))\, ds] \\ &= R_0 - \int_0^T \mathrm{e}^{-\rho s} dA_s.\end{aligned}$$

Theorem 4.1: *The coupled stochastic optimal contracting problem is equivalent to the deterministic optimization problem*

$$\sup_{g,f,\beta,A} \int_0^T \mathrm{e}^{-rs}(-\lambda_s(m + \mathbb{E}[g(s, Y_1)])\, ds + (u_s - f(s)u_s - \beta_s)\, ds) - \int_0^T \mathrm{e}^{-rs} dA_s,$$

subject to

$$\begin{aligned}&\int_0^T \mathrm{e}^{-\rho s}[\lambda_s \mathbb{E}[U_1(g(s, Y_1))]\, ds + U_2(f(s))u_s\, ds + (\beta_s - P(\lambda_s, u_s))\, ds] \\ &= R_0 - \int_0^T \mathrm{e}^{-\rho s} dA_s, \\ &P_\lambda(\lambda_s, u_s) = \mathbb{E}[U_1(g(s, Y_1))] \text{ and } P_u(\lambda_s, u_s) = U_2(f(s)).\end{aligned}$$

Due to the role of the time values ρ and r, their values matter in the functional form of the optimal contract as well as in the time trends of optimal actions, the penalty-for-accident and reward-for-performance. The way to solve this problem is analogous to that in Chapter 3. Specifically, if $\rho > r$, the nonnegative lump-sum payment is made at time 0, the very beginning of the contract. If $\rho < r$, it is made at time T, the expiration date. In either case, intermediate payments are always suboptimal compared to the payments at the beginning or end, hence β_s is chosen to be zero.

Proposition 4.2: *The optimal λ_s^*, u_s^*, g^* and f^* are the maximizers of*

$$\begin{aligned} G_s(\lambda_s, u_s, g, f) = \mathrm{e}^{-rs}\big(- \lambda_s(m + \mathbb{E}[g(s, Y_1))] + (u_s - f(s)u_s)\big) \\ + \mathrm{e}^{-\rho s}\big(\lambda_s \mathbb{E}[U_1(g(s, Y_1))] + U_2(f(s))u_s - P(\lambda_s, u_s)\big) \end{aligned} \tag{4.4}$$

subject to

$$P_\lambda(\lambda_s, u_s) = \mathbb{E}[U_1(g(s, Y_1))] \text{ and } P_u(\lambda_s, u_s) = U_2(f(s)).$$

The optimal β_s^* and A_s^* are given as in Proposition 3.2.

Technically, this problem is much harder than the problem in Proposition 3.2, for here one needs to solve for g, a function of accident size Y_i in addition to λ and u while one only optimizes over λ and u in Proposition 3.2. However, with the help of Jensen's inequality, one can overcome this difficulty by choosing the function g in a smart way.

Theorem 4.2: *If U_1 is strictly concave, the optimal penalty-for-accident strategy g is constant in the size of accidents.*

Proof. Since U_1 is concave, $\mathbb{E}[U_1(g(s, Y_1))] \leq U_1(\mathbb{E}g(s, Y_1))$ by Jensen's inequality. Hence, $U_1^{-1}(P_\lambda(\lambda_s, u_s))$ gives a lower bound for $\mathbb{E}[g(s,Y_1)]$; i.e., $\mathbb{E}[g(s,Y_1)] \geq U_1^{-1}(P_\lambda(\lambda_s, u_s))$. Since $G_s(\lambda_s, u_s, g, f)$ is decreasing in $\mathbb{E}[g(s,Y_1)]$, $\mathbb{E}[g(s,Y_1)] = U_1^{-1}(P_\lambda(\lambda_s, u_s))$ in the optimum. Therefore, in the optimum, $U_1(\mathbb{E}g(s,Y_1)) = P_\lambda(\lambda_s, u_s) = \mathbb{E}[U_1(g(s,Y_1))]$. But U_1 is strictly concave, then the only class of functions g that admits equality has the property that g is constant in Y_1, the size of accidents. □

The fact that g is constant in the size of accidents has to do with the risk aversion of the agent and the independently determined punishment for each accident. Capponi and Frei (2015) concludes that in a terminal-payment contract signed by a risk-averse agent, big accidents are punished less than small accidents, a "sublinear contract" as they call it. In a continuous-time payment scenario, this effect is even stronger: as the agent does not know the size of each accident before it happens, there is no incentive for the principal to distinguish in the optimal penalty-for-accident strategy between big and small accidents. If the agent is punished more for big accidents, he/she will suffer much more, stimulating him/her to demand a larger lump-sum payment from the principal so as to reach his/her reservation utility. This chain reaction does not benefit the principal.

Proposition 4.3: *In the case $\rho > r$, an optimal contract exists if*

$$\mathop{lim\,sup}_{\frac{u_s}{\lambda_s}\to\infty} \frac{-\mathrm{e}^{-rs}\lambda_s U_1^{-1}(P_\lambda(\lambda_s,u_s)) + \mathrm{e}^{-rs}u_s + \mathrm{e}^{-\rho s}u_s P_u(\lambda_s,u_s)}{\mathrm{e}^{-rs}u_s U_2^{-1}(P_u(\lambda_s,u_s)) - \mathrm{e}^{-\rho s}\lambda_s P_\lambda(\lambda_s,u_s) + \mathrm{e}^{-\rho s}P(\lambda_s,u_s)} < 1, \tag{4.5}$$

and for all fixed $\bar{u}_s$,

$$\mathop{lim\,sup}_{\lambda_s\to\infty} -\lambda_s U_1^{-1}(P_\lambda(\lambda_s,\bar{u}_s))) < \infty. \tag{4.6}$$

The proof of Proposition 4.3 is given in the Appendix A.3.

Theorem 4.3: *The optimal actions* $\lambda^* = (\lambda_t^*)_{0\le t<T}$ *and* $u^* = (u_t^*)_{0\le t<T}$ *are the maximizers of the function* (4.4)*. Optimal penalty-for-accident strategy* g^* *and reward-for-performance strategy* f^* *are*

$$g^* = U_1^{-1}(P_\lambda(\lambda_s^*, u_s^*)) \text{ and } f^* = U_2^{-1}(P_u(\lambda_t^*, u_t^*)).$$

Remark: In the case where the agent is risk neutral, U_1 and U_2 are the identity mappings and then Theorem 4.1 corresponds to Theorem 3.1 with α_s defined as $\alpha_s = \mathbb{E}[g(s, Y_1)]/m$. In particular, this shows that for a risk-neutral agent, the optimal linear contract found in Chapter 3 is still an optimal contract in the bigger contract space of Chapter 4. However, it is not the unique optimal contract in this bigger contract space because there are also optimal contracts which are nonlinear in the accident component. Indeed, every contract which has the same expected punishment per accident and the same continuous component and lump-sum payment is also optimal.

4.4 Numerical Illustrations

We use the same pain function as in Section 3.4. The effects of task substitution, time-trends of task levels and the shape of the optimal contract and the intuition behind them are similar to the situation in Section 3.4.

More important in this section are the effects of the levels of relative risk-aversion, which will be analyzed in detail. Since accidents J_t dramatically decrease the outcome process while the uncertainty of the continuous output component only swings it within a controllable range governed by the exogenous volatility σ_t, the risks associated with them are of different nature. Accordingly, the agent may have two different levels of risk aversion towards the two different risks. The interesting point is how the optimal tasks shift in response to the shift of the level of relative risk aversion.

The different powers of the strictly concave utility functions U_1 and U_2 make it possible to compare the relative risk aversion. We first use a fixed value for the power of U_2 while change the power of U_1 within 4 different levels, and then fix the power of U_1 while change the power of U_2 within 4 different levels. Specifically, we first analyze the situation when $U_1(x) = -(-x)^p$, where $p = 1, 1.2, 1.4, 1.6$; $U_2(x) = x^q$ where $q = 0.8$. Hence, the level of relative risk aversion is changed; the higher p becomes, the more averse the agent is to the risk incurred by accidents. Then we focus on the case where $p = 1.1$ and $q = 1, 0.85, 0.7, 0.55$. In this case, the lower q becomes, the more averse the agent is to the risk incurred by the uncertainty of the continuous component.

Figures 4.1 and 4.2 show the effect of the relative risk aversion on task substitution in these two different scenarios, while Figure 4.3 illustrate the effects of relative risk aversion (with fixed q) on the time trends of task levels.

As in Section 3.4, we choose $k_1 = k_2 = m = 1$ in the situation where the principal is more patient than the agent with $\rho = 0.7$ and $r = 0.5$.

The Effect of Relative Risk-aversion Magnitude on Task Substitution with fixed q. Figure 4.1 illustrates the effect of relative risk-aversion magnitude on task substitution when q is fixed. The more relatively risk-averse the agent is to the accidents, the less accident prevention and more effort exertion he/she will perform. Although a severer penalty-for-accident induces the agent to do more prevention, it increases the lump-sum payment that the principal pays to reach the agent's reservation utility. A more relatively risk-averse agent will demand more lump-sum payment, a higher cost compared to the accident reduction that the principal gets as a benefit. Since the principal is risk-neutral, it is beneficial for him/her to bear more costs of accidents himself/herself, launching an "induced task shift" from accident prevention to effort exertion, even if it leads to less prevention by the agent.

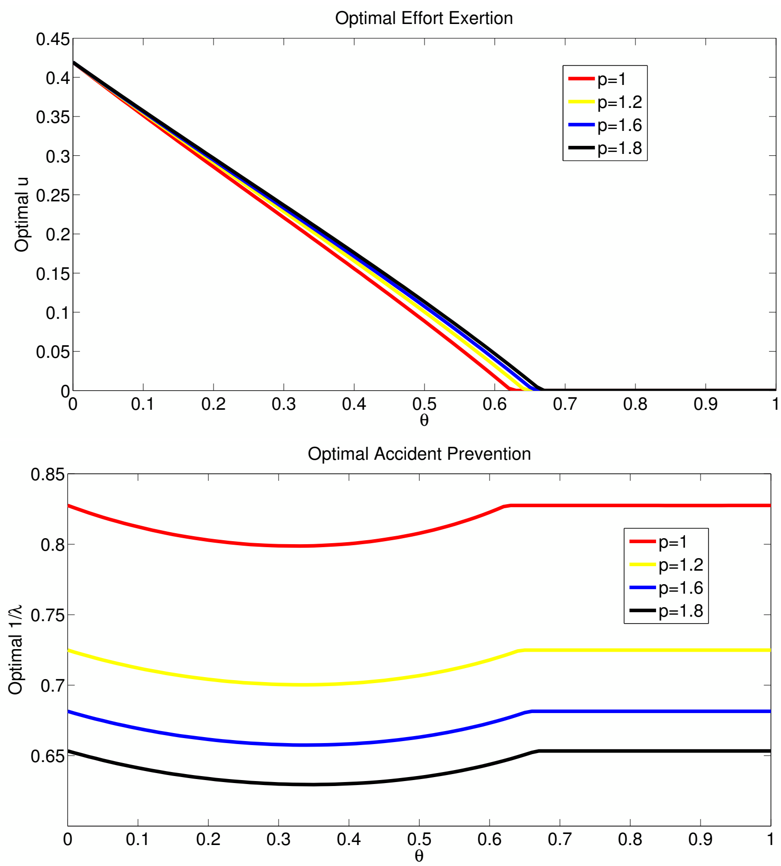


Figure 4.1: The Effect of Relative Risk-aversion Magnitude on Task Substitution with fixed q

The Effect of Relative Risk-aversion Magnitude on Task Substitution with fixed p. Figure 4.2 illustrates the effect of relative risk-aversion magnitude on task substitution when p is fixed. The more relatively risk-averse

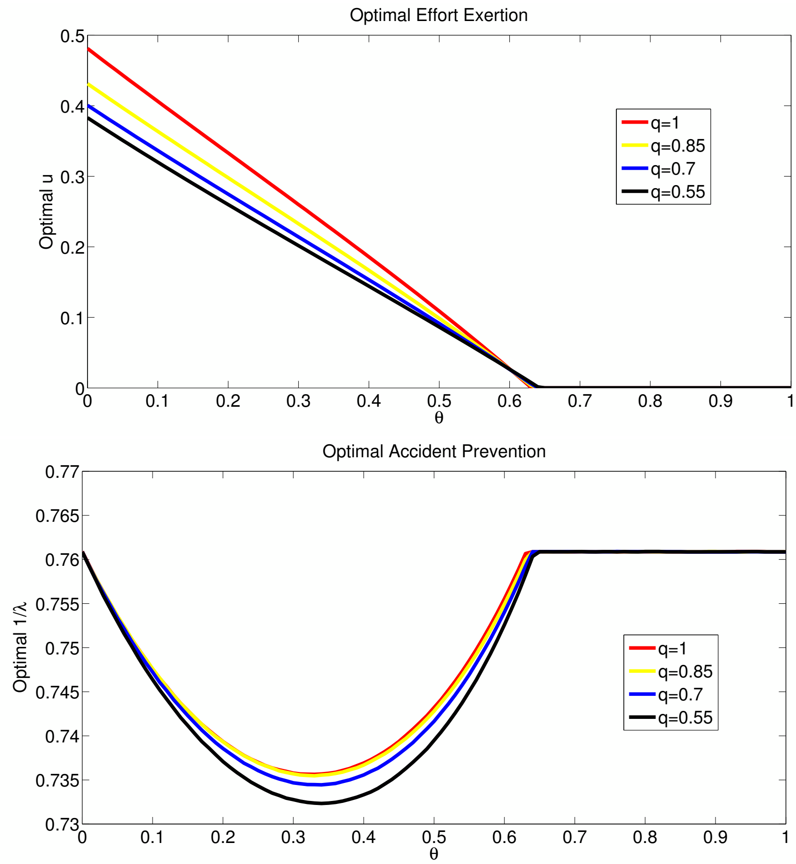


Figure 4.2: The Effect of Relative Risk-aversion Magnitude on Task Substitution with fixed p

the agent is to the uncertainty of the continuous output component, the less effort and — perhaps surprisingly — the less prevention he/she will perform. However, the differences in the prevention for varying q are small; for example, for $\theta = 0.3$, it varies somewhere between 0.732 and 0.736. For lower q, the agent reduces the effort as expected. However, a reduction in effort is negative for the principal. To counteract, the principal could increase (or not reduce in the required amount) the reward-for-performance compensation, but this may not be beneficial as it is costly. Another way is to slightly reduce the penalty-for-accident so that the agent undertakes slightly less prevention. This implies that the agent reduces the effort less because effort is slightly less costly due to the reduced prevention in the mixed term of the pain function.

To demonstrate the time-trend effects of the relative risk-aversion on task levels, we use the following graph with fixed q. Hence, one should interpret the relative risk-aversion in the following graph as the risk-aversion to the accidents against that to uncertainty of the continuous output component.

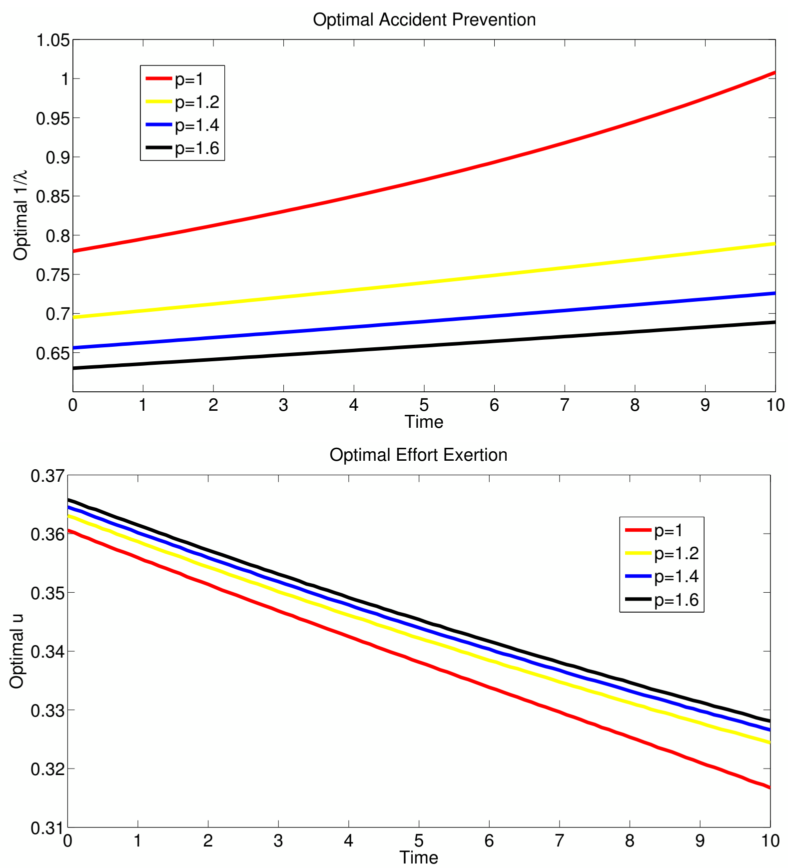


Figure 4.3: The Time Effect of Relative Risk-aversion Magnitude on Task Levels

The Time Effect of Relative Risk-aversion Magnitude on Task Levels. Figure 4.3 illustrates the effect of relative risk-aversion magnitude on the time trends of the optimal task levels. A more relatively risk-averse agent performs less accident prevention and more effort exertion over the whole contract period, due to the same reason as argued for Figure 4.1.

Chapter 5

Partially Limited Liability

We have successfully derived the optimal contracts and the optimal task levels in the scenarios where the agent is risk neutral and risk averse. However, one may notice that for most cases in the real world, the agent is responsible for only part of accident costs during a project. For instance, the general manager of a company may be fired because of his/her incompetence, but will never pay for all the losses of the company during his/her term of office.

We incorporate this feature into our model to fit the reality. Specifically, a threshold of penalty is exogenously given, reflecting the regulation severity of that industry (or company). Whenever an accident happens, the agent is fully punished if the cost for him/her is lower than the threshold while is partially punished up to the threshold if the cost is too high. To the extent that the agent is protected by the given "cap", we call it partially limited liability. We analyze its effect on the optimal contract and the time-trends of optimal tasks.

To better characterize the effects of the partially limited liability, we analyze the contract signed by a risk neutral agent in Section 5.1 and that signed by a risk averse agent in Section 5.2. In each scenario, the partially limited liability is further represented by an absolute threshold and a percentage threshold in terms of the occurring accident. After the analysis of the effects of the two thresholds, we present the link between them at the end of each scenario. It turns out that in a linear contract signed by a risk neutral agent, the optimal task levels are exactly the same as long as the two thresholds abide by a certain pattern. However, such rules do not exist in a more general contract.

5.1 A Risk-neutral Agent

5.1.1 Problem Formulation

We first derive the optimal contracts and task levels for the absolute threshold. At the end of this section, we compare the effects of an absolute threshold and a percentage threshold at corresponding levels on the optimal task levels. In the two scenarios, the optimal task levels are exactly the same under a linear contract if the agent is risk-neutral.

Contract Space. Based on the linear contract specified in Chapter 3, we add a threshold to protect the agent from being charged overly high costs of accidents. Since the uncertainty from the continuous output is controlled by the exogenous volatility σ_t and thus its instantaneous changes are often negligible compared to the uncertainty from accidents, we only cap the penalty-for-accident with the threshold and leave the continuous output as is. Thus we consider the contract of the form

$$dL_t = \max\{\alpha_t Y_{N_t}, k\}\, \Delta N_t + \gamma_t\, dx_t + \beta_t\, dt + dA_t, \quad 0 \le t \le T,$$

where the fixed $k \in (-\infty, 0)$ is the threshold and its absolute value specifies the highest level of punishment allowed. It is called the absolute threshold in that the agent is only responsible up to the amount of $|k|$ if the cost of an accident is too high. However, the agent will pay fully for small accidents. The parameters α_t, β_t, γ_t and A_t are deterministic as in Chapter 3. The agent and the principal solve the problems as specified by (3.1), (3.2) and (3.3).

5.1.2 Incentive Compatibility

The agent's utility at time T is

$$U_T = \int_0^T \mathrm{e}^{-\rho s}[\max\{\alpha_s Y_{N_s}, k\} \, \Delta N_s + \beta_s \, ds + dA_s + \gamma_s \, dx_s - P(\lambda_s, u_s) \, ds].$$

For optimal λ^* and u^*, we have $\mathbb{E}[U_T(L, u^*, \lambda^*)] = R_0$ by (3.3) under the assumption that the principal has full information of the agent's reservation utility. Hence,

$$\mathbb{E}\left[\int_0^T \mathrm{e}^{-\rho s}[\max\{\alpha_s Y_{N_s^*}, k\} \, \Delta N_s^* + \beta_s \, ds + dA_s + \gamma_s \, dx_s^* - P(\lambda_s^*, u_s^*) \, ds]\right] = R_0.$$

Notice that

$$\begin{aligned}\int_0^T \mathrm{e}^{-\rho s}(\beta_s \, ds + dA_s) = &- \int_0^T \mathrm{e}^{-\rho s}\mathbb{E}[\max\{\alpha_s Y_{N_s^*}, k\} \, \Delta N_s^*] + R_0 \\ &- \int_0^T \mathrm{e}^{-\rho s}\gamma_s u_s^* \, ds + \int_0^T \mathrm{e}^{-\rho s} P(\lambda_s^*, u_s^*) \, ds.\end{aligned}$$

Therefore, for the nonoptimal λ and u,

$$
\begin{aligned}
U_T(L,\lambda,u) &= \int_0^T \mathrm{e}^{-\rho s}[\max\{\alpha_s Y_{N_s}, k\}\,\Delta N_s + \beta_s\,ds + dA_s + \gamma_s\,dx_s \\
&\qquad - P(\lambda_s,u_s)\,ds] \\
&= \int_0^T \mathrm{e}^{-\rho s}[\max\{\alpha_s Y_{N_s}, k\}\,\Delta N_s + \gamma_s\,dx_s - P(\lambda_s,u_s)\,ds] \\
&\quad - \int_0^T \mathrm{e}^{-\rho s}\mathbb{E}[\max\{\alpha_s Y_{N_s^*}, k\}\,\Delta N_s^*] - \int_0^T \mathrm{e}^{-\rho s}\gamma_s u_s^*\,ds \\
&\quad + \int_0^T \mathrm{e}^{-\rho s} P(\lambda_s^*, u_s^*)\,ds + R_0 \\
&= \int_0^T \mathrm{e}^{-\rho s}\gamma_s(u_s - u_s^*)\,ds + \int_0^T \mathrm{e}^{-\rho s}(P(\lambda_s^*,u_s^*) - P(\lambda_s,u_s))\,ds \\
&\quad + \int_0^T \mathrm{e}^{-\rho s}(\max\{\alpha_s Y_{N_s}, k\}\,\Delta N_s - \mathbb{E}[\max\{\alpha_s Y_{N_s}, k\}\,\Delta N_s]) \\
&\quad + \int_0^T \mathrm{e}^{-\rho s}(\mathbb{E}[\max\{\alpha_s Y_{N_s}, k\}\,\Delta N_s] - \mathbb{E}[\max\{\alpha_s Y_{N_s^*}, k\}\,\Delta N_s^*]) \\
&\quad + \int_0^T \mathrm{e}^{-\rho s}\gamma_s\sigma_s dB_s^u + R_0.
\end{aligned} \tag{5.1}
$$

Lemma 5.1: *If $\alpha < 0$ and λ is deterministic, then*

$$
\mathbb{E}[\max\{\alpha_s Y_{N_s}, k\}\,\Delta N_s] = \left(k(1 - F(k/\alpha_s)) + \alpha_s \int_0^{k/\alpha_s} x\,dF(x)\right)\lambda_s\,ds.
$$

Proof. We compute

$$
\begin{aligned}
&\mathbb{E}[\max\{\alpha_s Y_{N_s}, k\}\,\Delta N_s] \\
&= \mathbb{E}[\max\{\alpha_s Y_{N_s}, k\}]\lambda_s\,ds \\
&= \mathbb{E}[\,\max\{\alpha_s Y_{N_s}, k\}\mathbb{1}_{\alpha_s Y_{N_s} < k} + \max\{\alpha_s Y_{N_s}, k\}\mathbb{1}_{\alpha_s Y_{N_s} \geq k}]\lambda_s\,ds \\
&= \mathbb{E}[k\mathbb{1}_{\alpha_s Y_{N_s} < k} + \alpha_s Y_{N_s}\mathbb{1}_{\alpha_s Y_{N_s} \geq k}]\lambda_s\,ds \\
&= \left(k(1 - F(k/\alpha_s)) + \alpha_s \int_0^{k/\alpha_s} x\,dF(x)\right)\lambda_s\,ds. \qquad \square
\end{aligned}
$$

Therefore, the incentive compatibility conditions are given by the first-order condition of (5.1).

Proposition 5.1: *The sufficient and necessary conditions for the intensity* $\lambda = (\lambda_t)_{0 \le t \le T}$ *and effort* $u = (u_t)_{0 \le t \le T}$ *to be incentive compatible are*

$$P_\lambda(\lambda_t, u_t) = k(1 - F(k/\alpha_s)) + \alpha_s \int_0^{k/\alpha_s} x \, dF(x) \ \textit{and} \ P_u(\lambda_t, u_t) = \gamma_t. \quad (5.2)$$

Remark: If $k = -\infty$, then $P_\lambda(\lambda_t, u_t) = \alpha_s m$, which corresponds to (3.6) in Chapter 3. When the threshold k is too small, it will not affect the magnitude of punishment. Therefore, the incentive compatibility conditions are exactly the same as in the case without the threshold.

5.1.3 Optimal Contracting

Plug the incentive compatibility conditions (5.2) into the principal's problem (3.2) and (3.3) and we have the following result.

Theorem 5.1: *The coupled stochastic optimal contracting problem* (3.1)–(3.3) *with partial limited liability is equivalent to the deterministic optimization problem*

$$\begin{aligned} \sup_{\alpha,\beta,\gamma,A} \int_0^T \mathrm{e}^{-rs} \Big[(1-\gamma_s) u_s - \Big(m + k(1 - F(k/\alpha_s)) + \alpha_s \int_0^{k/\alpha_s} x \, dF(x) \Big) \lambda_s \Big] \, ds \\ - \int_0^T \mathrm{e}^{-rs} (\beta_s \, ds + dA_s) \end{aligned}$$

subject to

$$\int_0^T \mathrm{e}^{-\rho s}\left[\left(k(1-F(k/\alpha_s))+\alpha_s\int_0^{k/\alpha_s} x\,dF(x)\right)+\beta_s+\gamma_s u_s - P(\lambda_s,u_s)\right]ds$$
$$= R_0 - \int_0^T \mathrm{e}^{-\rho s} dA_s,$$
$$P_\lambda(\lambda_s,u_s) = k(1-F(k/\alpha_s))+\alpha_s\int_0^{k/\alpha_s} x\,dF(x),$$
$$P_u(\lambda_s,u_s)=\gamma_s.$$

Note from the proof of Proposition 5.1 that $P_\lambda(\lambda_s,u_s)=\mathbb{E}[\max\{\alpha_s Y_{N_s},k\}]$. Moreover, $\mathbb{E}[\max\{\alpha_s Y_{N_s},k\}]\geq k$. As opposed to Chapter 3, $P_\lambda(\lambda_s,u_s)$ cannot take any nonpositive value, but must have $P_\lambda(\lambda_s,u_s)\geq k$, which gives the constraint of the following reduced principal's problem.

Proposition 5.2: *The optimal λ_s^* and u_s^* are the maximizers of*

$$\begin{aligned}G_s(\lambda_s,u_s) &= \mathrm{e}^{-rs}[u_s - u_s P_u(\lambda_s,u_s) - m\lambda_s - \lambda_s P_\lambda(\lambda_s,u_s)]\\ &\quad + \mathrm{e}^{-\rho s}[\lambda_s P_\lambda(\lambda_s,u_s)+u_s P_u(\lambda_s,u_s) - P(\lambda_s,u_s)], \qquad (5.3)\end{aligned}$$

subject to

$$P_\lambda(\lambda_s,u_s)\geq k.$$

For $a = R_0 - \int_0^T \mathrm{e}^{-\rho t}[P_\lambda(\lambda_t^*,u_t^*)\lambda_t^* + P_u(\lambda_t^*,u_t^*)u_t^* - P(\lambda_t^*,u_t^*)]\,dt$, the optimal lump-sump payment is as follows:

- *If $\rho=r$, optimal β_s^* and A_s^* satisfy $A_T^*+\int_0^T\beta_s^*\,ds=a$, but they are not uniquely determined.*
- *If either $\rho>r$ and $a\geq 0$ or $\rho<r$ and $a<0$, the optimal $A_s^*=a$ is constant, and the optimal β_s^* is identical to zero.*

- *If either $\rho < r$ and $a \geq 0$ or $\rho > r$ and $a < 0$, the optimal A_s^* is zero for all $s < T$ and $A_T^* = \mathrm{e}^{\rho T} a$, and the optimal β_s^* is identical to zero.*

After we solve for the optimal λ_s^* and u_s^* from Proposition 5.2, we plug them into (5.2) to get the optimal α_s^* and γ_s^*.

Theorem 5.2: *The optimal $\lambda^* = (\lambda_t^*)_{0 \leq t < T}$ and $u^* = (u_t^*)_{0 \leq t < T}$ are the maximizers of the function (5.2). If $P_\lambda(\lambda_t^*, u_t^*) > k$, the optimal α_t^* is given by*

$$P_\lambda(\lambda_t^*, u_t^*) = k(1 - F(k/\alpha_t^*)) + \alpha_t^* \int_0^{k/\alpha_t^*} x \, dF(x);$$

if $P_\lambda(\lambda_t^*, u_t^*) = k$, an optimal α_t^* is

$$\alpha_t^* = -\infty.$$

The optimal $\gamma_t^* = P_u(\lambda_t^*, u_t^*)$.

The optimal penalty-for-accident sensitivity is $\max\{\alpha_t^* Y_{N_t}, k\}$ and reward-for-performance sensitivity is $P_u(\lambda_t^*, u_t^*)$.

The case $\alpha^* = -\infty$ means that the agent is charged a constant penalty $|k|$ for every accident, even when the accident size is lower than $|k|$. The principal penalizes the agent excessively for small accidents because the agent can be punished only up to $|k|$. To give enough incentives for prevention, the principal punishes the agent independently of the accident size, taking into account that the agent's prevention reduces all possible accident sizes, since the sizes are independent of the accident occurrence.

While an agent protected by an absolute threshold is only responsible up to a fixed amount of accident cost, the agent with a percentage threshold only pays up to a fixed proportion of accident cost, regardless of accident size. The way in which the percentage protection influences the optimal task levels is presented in the following proposition, and it turns out that it corresponds well to the absolute protection.

Theorem 5.3: *If the agent is protected by the percentage threshold $c \in (-\infty, 0)$ in that the payment stream is governed by*

$$dL_t = \max\{\alpha_s, c\} Y_{N_s} \Delta N_s + \beta_s \, ds + dA_s + \gamma_s \, dx_s,$$

the optimal $\lambda^* = (\lambda_t^*)_{0 \le t < T}$ and $u^* = (u_t^*)_{0 \le t < T}$ are the same as in the case where the agent is protected by the absolute threshold $k = cm$.

The agent in such a "percentage-protection" contract is responsible for at most a proportion $|c|$ of the accident cost Y_{N_t} when the accident happens at any time t. Since the agent is risk neutral, he/she does not suffer more than the penalty incurred by the accident cost. Note that m is the average size of an accident. Therefore, the absolute protection $k = cm$ protects the agent just as well as the percentage protection c, leading to the same levels of optimal tasks. The proof of Proposition 5.3 is given in Appendix A.5.

5.1.4 Numerical Illustrations

To analyze the effects of the threshold k, we use the same pain function

$$P(\lambda, u) = (k_1 u)^2 + (k_2/\lambda)^2 + 2\theta k_1 k_2 u/\lambda$$

as in Chapters 3 and 4.

The value of k is critical in determining the shape of the optimal contract as well as the optimal task levels. Specifically, if $|k|$ is always higher than the amount (we call it "unregulated amount") by which the principal would charge the agent for an accident, then the threshold is not effective and the agent pays for the full amount of accident cost. However, if $|k|$ is lower than the unregulated amount sometime, then the agent will only pay the amount of $|k|$, a feature reflecting the partially limited liability. Therefore, we choose four different levels of the threshold $k = -200, -15, -10, -5$. The level $k = -200$ corresponds to the scenario when the threshold is so low that it is not effective while $k = -15, -10, -5$ illustrate the effects of narrowing the effective threshold.

The optimal α^* depends on the distribution of the accident size Y_i. Here we assume that Y_i is uniformly distributed on $[0, 2m]$. As in Sections 3.4 and 4.4, we choose $k_1 = k_2 = m = 1$ in the situation where the principal is more patient than the agent with $\rho = 0.07$ and $r = 0.05$.

The Time Trend of Task Levels. Figure 5.1 illustrates the trends of optimal accident prevention $\frac{1}{\lambda^*}$ and effort exertion u_s^* over time. When the threshold is low enough ($k = -200$), $\frac{1}{\lambda^*}$ and u_s^* behave as if there were no threshold. When the threshold is effective ($k = -15, -10, -5$), $\frac{1}{\lambda^*}$ increases until it reaches its highest level set by the threshold. Since the punishment is constant and equals $|k|$ from this point, the agent has no incentives to either increase or decrease the prevention level. The narrower the threshold is, the earlier the effect emerges. When $\frac{1}{\lambda^*}$ is constant, the agent has more resources for effort exertion compared to the case without the threshold. Therefore, although u_s^* still decreases, it does not decrease as fast as in the case without the threshold.

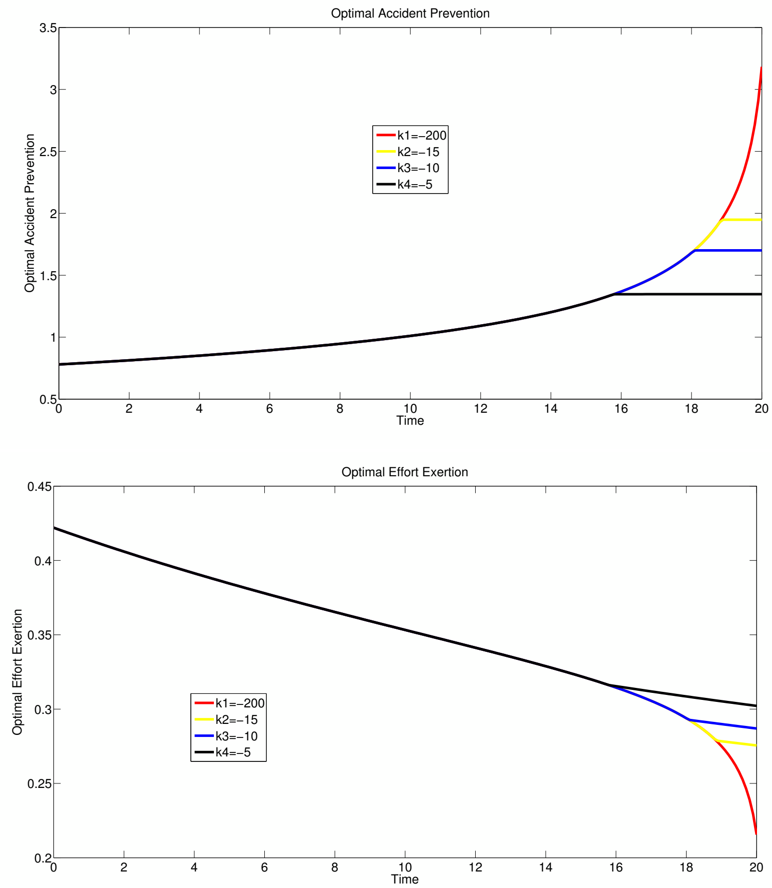


Figure 5.1: The Time Trend of Task Levels

The Optimal Contract. Figure 5.2 illustrates the incentives in the optimal contract: $|\alpha^*|$ and the reward-for-performance sensitivity γ^*. When the threshold is low enough ($k = -200$), $|\alpha^*|$ and γ^* behave as if there were no threshold. When the threshold is effective ($k = -15, -10, -5$), $|\alpha^*|$ surges dramatically, reflecting the fact that the principal charges the agent more for small accidents to gain a compensation (because he/she can only charge the agent for at most the amount of $|k|$) than in the case without threshold. Compared to the case without threshold, the principal does not have to worry about that the agent shifts too many resources to accident prevention (because prevention is constant and lower than the case without threshold), the reward-for-performance is reduced accordingly.

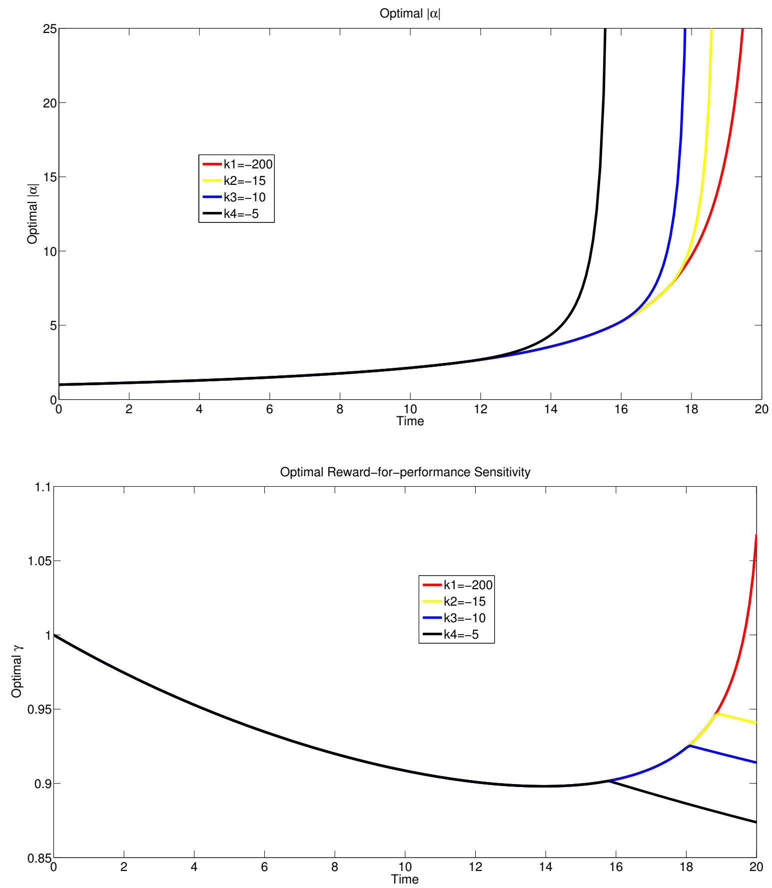


Figure 5.2: The Optimal Contract

5.2 A Risk-averse Agent

5.2.1 A Percentage Threshold

Contract space. Not restricted to the linear contract as in Section 5.1, the form of contract available to the agent is more general in the sense that the punishment policy is represented by a nonlinear function, since the principal has the freedom to choose whatever punishment policy as he/she prefers. This generalization allows us to analyze the penalty strategy specified by a broader class of functions as we do in Chapter 4. Moreover, as in Section 5.1, the agent is not fully responsible for the accident cost. Instead, he/she will only pay a percentage, reflecting the essence of partially limited liability. For the same reason as in Section 5.1, we only cap the penalty-for-accident strategy with the percentage threshold while leave the continuous output as is. Hence, the contract is of the form

$$dL_t = g(t, dJ_t) + f(t)\, dx_t + \beta_t\, dt + dA_t, \quad 0 \leq t \leq T,$$

where

$$g(., x) \geq cx. \tag{5.4}$$

The constant $c \in (-\infty, 0)$ is the percentage threshold and its absolute value specifies the highest proportion of accident cost for which the agent is responsible. It protects the agent in that he/she is only responsible up to the proportion $|c|$ if the cost of an accident is too high while the agent will pay fully for low accident cost. The function g and f, parameters β_t and A_t are deterministic as in Chapter 4. The agent and the principal solve the problems as specified by (4.1), (4.2) and (4.3) such that (5.4) holds.

Incentive Compatibility and Optimal Contracting

Not surprisingly, we arrive at the same incentive compatibility conditions, by applying the same method as in Chapter 4.

Proposition 5.3: *The sufficient and necessary conditions for the intensity $\lambda = (\lambda_t)_{0 \leq t \leq T}$ and effort $u = (u_t)_{0 \leq t \leq T}$ to be incentive compatible are*

$$P_\lambda(\lambda_t, u_t) = \mathbb{E}[U_1(g(t, Y_1))] \text{ and } P_u(\lambda_t, u_t) = U_2(f(t)).$$

We plug the conditions to the principal's problem and then have the following theorem.

Theorem 5.4: *The coupled stochastic optimal contracting problem specified by* (4.1), (4.2), (4.3) *and* (5.4) *is equivalent to the deterministic optimization problem*

$$\sup_{g,f,\beta,A} \int_0^T \mathrm{e}^{-rs}(-\lambda_s(m+\mathbb{E}[g(s,Y_1)])\,ds+(u_s-f(s)u_s-\beta_s)\,ds)-\int_0^T \mathrm{e}^{-rs}dA_s,$$

subject to

$$\begin{aligned}
&\int_0^T \mathrm{e}^{-\rho s}[\lambda_s\mathbb{E}[U_1(g(s,Y_1))]\,ds + U_2(f(s))u_s\,ds + (\beta_s - P(\lambda_s,u_s))\,ds] \\
&= R_0 - \int_0^T \mathrm{e}^{-\rho s}dA_s, \\
&P_\lambda(\lambda_s,u_s) = \mathbb{E}[U_1(g(s,Y_1))] \text{ and } P_u(\lambda_s,u_s) = U_2(f(s)) \\
&g(.,x) \geq cx.
\end{aligned}$$

The regular way to untangle this problem is to optimize over a class of functions g. Recall that we solved the similar problem in Chapter 4 by Jensen's inequality and reach that the optimal penalty-for-accident strategy g is constant in accident size. However, things are turned much more complicated here by the presence of the constraint $g(.,x) \geq cx$, because the constraint is not guaranteed to hold when g is constant in accident size. Now that it is hard (or perhaps impossible) to solve the problem directly, we have to resort to an alternative way to tackle the problem, namely, by parameterization. Before we present the details of the reasoning of parameterization, one may notice that Theorem 5.4 can be further reduced to the following proposition.

Proposition 5.4: *The optimal λ_s^*, u_s^*, g^* and f^* are the maximizers of*

$$G_s(\lambda_s, u_s, g, f) = \mathrm{e}^{-rs}\big(- \lambda_s(m + \mathbb{E}[g(s, Y_1)]) + u_s - U_2^{-1}(P_u(\lambda_s, u_s))u_s\big)$$
$$+ \mathrm{e}^{-\rho s}(\lambda_s \mathbb{E}[U_1(g(s, Y_1))] + P_u(\lambda_s, u_s)u_s - P(\lambda_s, u_s))$$

subject to

$$P_\lambda(\lambda_s, u_s) = \mathbb{E}[U_1(g(s, Y_1))]$$
$$g(., x) \geq cx.$$

The optimal β_s^* and A_s^* are given as in Proposition 3.2.

Notice that $G_s(\lambda_s, u_s, g, f)$ is decreasing in $\mathbb{E}[g(s, Y_1)]$. Hence, separating the other control variables λ_s, u_s and f, we consider the problem

$$\inf_g \mathbb{E}[g(s, Y_1)]$$

subject to

$$g(s, Y_1) \geq cY_1 \tag{5.5}$$

$$\mathbb{E}[g(s, Y_1)] \geq U_1^{-1}(P_\lambda(\lambda_s, u_s)). \tag{5.6}$$

If (5.5) were the only constraint, then we would choose $g^*(s, Y_1) = cY_1$ in optimum, indicating the optimal penalty-for-accident strategy is linear in accident size. Correspondingly, if (5.6) were the only constraint, we would choose $g^*(s, Y_1) = U_1^{-1}(P_\lambda(\lambda_s, u_s))$ in optimum, the very same situation as in Chapter 4 that indicates the optimal penalty-for-accident strategy is constant in accident size. Now that the constraints consist of both (5.5) and (5.6), an suitable approximation for the optimal $g^*(s, Y_1)$ is a combination of the properties of linear and constant functions. Hence, we restrict the class of functions g to

$$g(s, Y_1) = c_1(s) Y_1 \mathbb{1}_{Y_1 < c_2(s)} + c_1(s) c_2(s) \mathbb{1}_{Y_1 \geq c_2(s)}$$

such that the incentive compatibility conditions hold and

$$c_1(s) \geq c.$$

Remark: The parameterized $g(s, Y_1)$ has the properties: as $c_2(s) \to \infty$, $g(s, Y_1) \to c_1(s) Y_1$. As $c_2(s) \to 0$, $c_1(s) \to -\infty$ and $c_1(s) c_2(s) \to k$, $g(s, Y_1) \to k$ pointwise for all $Y_1 \neq 0$.

With the parameterized $g(s, Y_1)$, the following proposition reduces Proposition 5.4.

Proposition 5.5: *The optimal λ_s^*, u_s^*, $c_1(s)^*$ and $c_2(s)^*$ are the maximizers of*

$$\begin{aligned} G_s(\lambda, u, c_1, c_2) = \mathrm{e}^{-rs} \Big[& - \lambda_s \Big(m + c_1 \int_0^{c_2} x \, dF(x) + c_1 c_2 \int_{c_2}^{\infty} dF(x) \Big) \\ & + u_s - U_2^{-1}(P_u(\lambda_s, u_s)) u_s \Big] \\ & + \mathrm{e}^{-\rho s} (\lambda_s P_\lambda(\lambda_s, u_s) + u_s P_u(\lambda_s, u_s) - P(\lambda_s, u_s)) \end{aligned}$$

subject to

$$\begin{aligned} & P_\lambda(\lambda_s, u_s) = \mathbb{E}[U_1(g(s, Y_1))] \\ & c_1(s) \geq c \end{aligned}$$

Theorem 5.5: *The optimal penalty-for-accident strategy g^* and reward-for-performance strategy f^* are*

$$g(s, Y_1)^* = c_1^*(s) Y_1 \mathbb{1}_{Y_1 < c_2^*(s)} + c_1^*(s) c_2^*(s) \mathbb{1}_{Y_1 \geq c_2^*(s)} \text{ and } f^* = U_2^{-1}(P_u(\lambda_t^*, u_t^*)).$$

Numerical Illustration

To analyze the effects of the threshold k, we use the same pain function

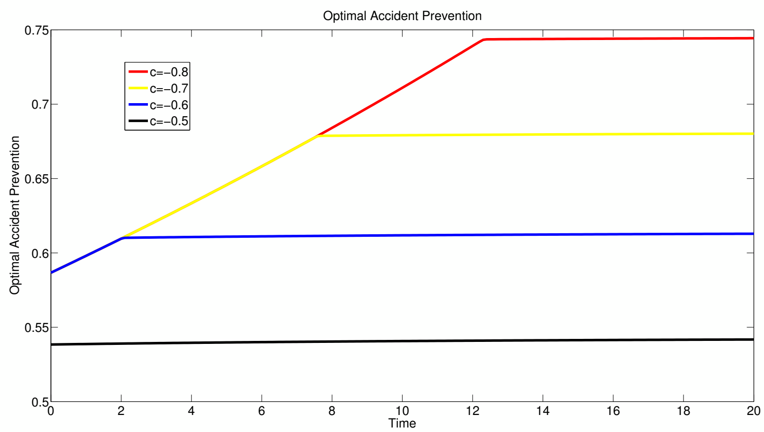
$$P(\lambda, u) = (k_1 u)^2 + (k_2/\lambda)^2 + 2\theta k_1 k_2 u/\lambda$$

as in Chapters 3 and 4.

The value of c is critical in determining the shape of the optimal contract as well as the optimal task levels. Specifically, if $|c|$ is so large that the amount $|cY_i|$ is always higher than the unregulated amount by which the agent is charged, then the threshold is not effective and the agent pays for the full amount of accident cost. However, if $|c|$ is small enough so that $|cY_i|$ is lower than the unregulated amount sometime, then the agent will only pay the amount of $|cY_i|$, a feature reflecting the partially limited liability. Therefore, we choose four different levels of the threshold $c = -0.8, -0.7, -0.6, -0.5$. The larger c goes, the strict the threshold is and thus the better the agent is protected.

Still we assume that Y_i is uniformly distributed on $[0, 2m]$. As in Sections 3.4 and 4.4, we choose $k_1 = k_2 = m = 1$, $p = 2$ and $q = 0.8$ in the situation where the principal is more patient than the agent with $\rho = 0.1$ and $r = 0.05$.

The Time Trend of Task Levels. Figure 5.3 illustrates the trends of optimal accident prevention $\frac{1}{\lambda^*}$ and effort exertion u_s^* over time. When the threshold is low ($c = -0.8$), $\frac{1}{\lambda^*}$ increases before it reaches its highest level set by the threshold. The narrower the threshold is, the earlier $\frac{1}{\lambda^*}$ reaches the highest level. Since the proportion of accident cost for which the agent is responsible is constant , he/she has no incentives to either increase or decrease the prevention level. Notice that $\frac{1}{\lambda^*}$ first increases mildly. Therefore, when it levels off, it does not have much impact on the trend of u^*. Therefore, u^* decreases smoothly throughout entire period as if it is not affected.



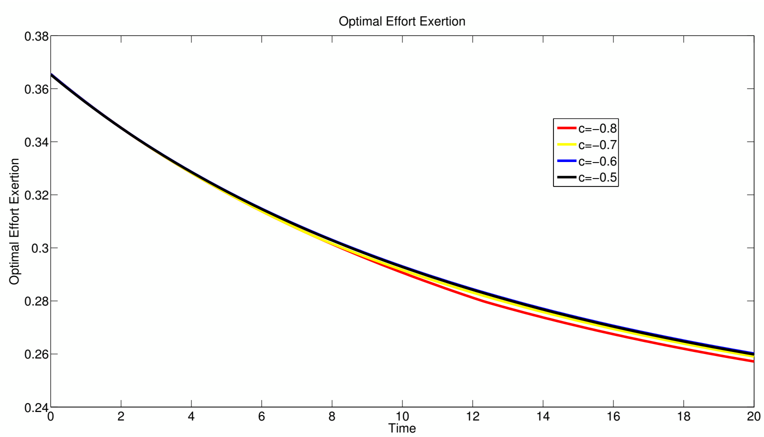


Figure 5.3: The Time Trend of Task Levels

The Optimal Contract. Figure 5.4 illustrates the optimal penalty-for-accident sensitivity c_1^* and the reward-for-performance sensitivity $f*$. It turns out that the original form of $g(s, Y_1) = c_1(s)Y_1 \mathbb{1}_{Y_1 < c_2(s)} + c_1(s)c_2(s)\mathbb{1}_{Y_1 \geq c_2(s)}$ is reduced to $g(s, Y_1) = c_1(s)Y_1$ in optimum. Hence c_1 represents the penalty-for-accident sensitivity. When the threshold is low ($c = -0.8$), c_1^* decreases before it reaches the threshold. The narrower the threshold is, the earlier c_1^* reaches the threshold. f^* decreases over the contract period as correspondence with u^*.

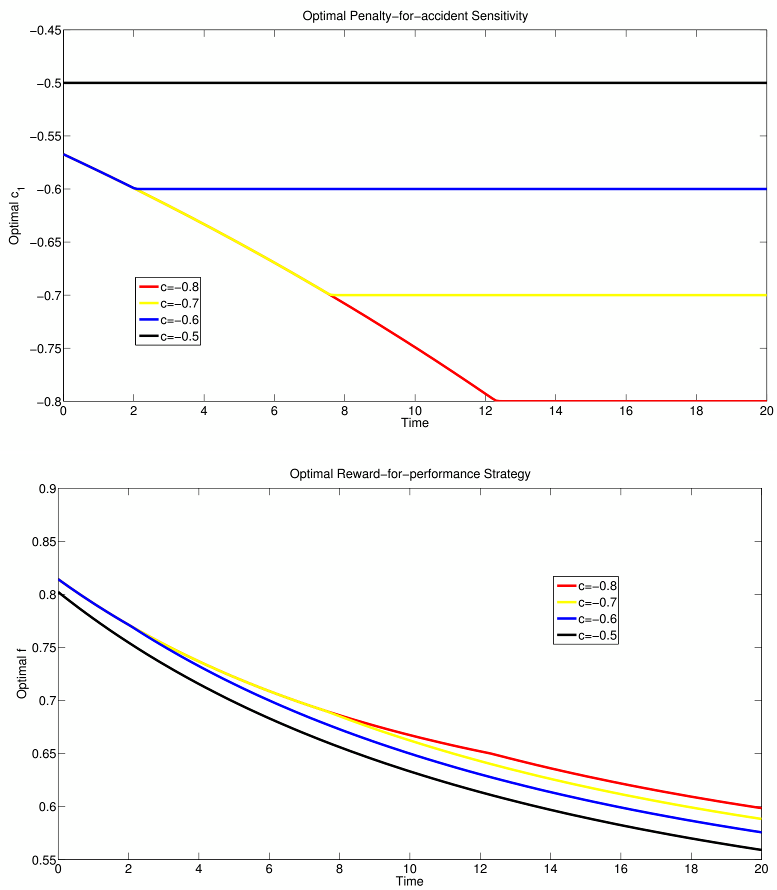


Figure 5.4: The Optimal Contract

5.2.2 An Absolute Threshold

Contract space. As in the Section 5.2.1, we generalize the linear penalty-for-accident strategy to a nonlinear one. However, as opposed to Section 5.2.1, the partially limited liability is reflected by an absolute threshold k, rather than a percentage threshold c. Therefore, the contract payment is of the form

$$dL_t = g(t, dJ_t) + f(t)\, dx_t + \beta_t\, dt + dA_t, \quad 0 \le t \le T,$$

where

$$g \ge k. \tag{5.7}$$

The constant $k \in (-\infty, 0)$ is the absolute threshold and its absolute value specifies the highest amount of accident cost for which the agent is responsible. It protects the agent in that he/she is only responsible up to the amount $|k|$ if the cost of an accident is too high while the agent will pay fully for low accident cost. The function g and f, parameters β_t and A_t are deterministic as in Chapter 4. The agent and the principal solve the problems as specified by (4.1), (4.2) and (4.3) such that (5.7) holds.

Incentive Compatibility and Optimal Contracting

We arrive at the same incentive compatibility conditions, by applying the same method as in Chapter 4.

Proposition 5.6: *The sufficient and necessary conditions for the intensity $\lambda = (\lambda_t)_{0\le t\le T}$ and effort $u = (u_t)_{0\le t\le T}$ to be incentive compatible are*

$$P_\lambda(\lambda_t, u_t) = \mathbb{E}[U_1(g(t, Y_1))] \text{ and } P_u(\lambda_t, u_t) = U_2(f(t)).$$

We plug the conditions to the principal's problem and then have the following theorem.

Theorem 5.6: *The coupled stochastic optimal contracting problem specified by (4.1), (4.2), (4.3) and (5.7) is equivalent to the deterministic optimization problem*

$$\sup_{g,f,\beta,A} \int_0^T \mathrm{e}^{-rs}(-\lambda_s(m+\mathbb{E}[g(s,Y_1)])\,ds+(u_s-f(s)u_s-\beta_s)\,ds)-\int_0^T \mathrm{e}^{-rs}dA_s,$$

subject to

$$\begin{aligned}&\int_0^T \mathrm{e}^{-\rho s}[\lambda_s\mathbb{E}[U_1(g(s,Y_1))]\,ds+U_2(f(s))u_s\,ds+(\beta_s-P(\lambda_s,u_s))\,ds]\\&=R_0-\int_0^T \mathrm{e}^{-\rho s}dA_s,\\&P_\lambda(\lambda_s,u_s)=\mathbb{E}[U_1(g(s,Y_1))] \text{ and } P_u(\lambda_s,u_s)=U_2(f(s))\\&g\geq k.\end{aligned}$$

Hence, the problem is further reduced by the following proposition.

Proposition 5.7: *The optimal λ_s^*, u_s^*, g^* and f^* are the maximizers of*

$$\begin{aligned}G_s(\lambda_s,u_s,g,f)=\mathrm{e}^{-rs}\big(&-\lambda_s(m+\mathbb{E}[g(s,Y_1))]+(u_s-f(s)u_s)\big)\\&+\mathrm{e}^{-\rho s}(\lambda_s\mathbb{E}[U_1(g(s,Y_1))]+U_2(f(s))u_s-P(\lambda_s,u_s))\end{aligned}$$

subject to

$$\begin{aligned}&P_\lambda(\lambda_s,u_s)=\mathbb{E}[U_1(g(s,Y_1))] \text{ and } P_u(\lambda_s,u_s)=U_2(f(s))\\&g\geq k.\end{aligned}$$

The optimal β_s^* and A_s^* are given as in Proposition 3.2.

Similarly to the remark on page 30, Proposition 5.7 corresponds to Proposition 5.2 when the agent is risk neutral.

Recall that we have applied Jensen's inequality in Chapter 4 to reach that $U_1^{-1}(P_\lambda(\lambda_s, u_s))$ gives the lower bound for $\mathbb{E}[g(s, Y_1))]$ and thus that the optimal g^* is constant in accident size. However, now the additional constraint $g(s, Y_1) \geq k$ complicates the problem, for it seems unclear whether $U_1^{-1}(P_\lambda(\lambda_s, u_s)) \geq k$ or the opposite. Nonetheless, we present the way of overcoming this difficulty in the following theorem.

Theorem 5.7: *The lower bound $U_1^{-1}(P_\lambda(\lambda_s, u_s))$ of $\mathbb{E}[g(s, Y_1))]$ is not smaller than the threshold k. If U_1 is strictly concave, the optimal penalty-for-accident strategy g^* is constant in accident size.*

Proof. Notice that the objective $G_s(\lambda_s, u_s, g, f)$ is decreasing in $\mathbb{E}[g(s, Y_1))]$. Suppose the threshold k is effective in that $k \geq U_1^{-1}(P_\lambda(\lambda_s, u_s))$, then we have $\mathbb{E}[g(s, Y_1))] = k$ in optimum, which corresponds to $g(s, Y_1) = k$. Hence, by the incentive compatibility $P_\lambda(\lambda_s, u_s) = \mathbb{E}[U_1(g(s, Y_1))]$, we have $P_\lambda(\lambda_s, u_s) = U_1(k)$, which implies $g(s, Y_1) = U_1^{-1}(P_\lambda(\lambda_s, u_s))$. Therefore, in the case $k \geq U_1^{-1}(P_\lambda(\lambda_s, u_s))$, we have $k = U_1^{-1}(P_\lambda(\lambda_s, u_s))$. Suppose the threshold is ineffective in that $k < U_1^{-1}(P_\lambda(\lambda_s, u_s))$, $g(s, Y_1) = U_1^{-1}(P_\lambda(\lambda_s, u_s))$ by Jensen's inequality. In summary, the lower bound $U_1^{-1}(P_\lambda(\lambda_s, u_s))$ of $\mathbb{E}[g(s, Y_1))]$ is not smaller than the threshold k and the optimal penalty-for-accident strategy $g^* = U_1^{-1}(P_\lambda(\lambda_s, u_s))$, which is constant in accident size. $\square$

Theorem 5.8: *Optimal penalty-for-accident strategy g^* and reward-for-performance strategy f^* are $g^* = U_1^{-1}(P_\lambda(\lambda_s^*, u_s^*))$ and $f^* = U_2^{-1}(P_u(\lambda_t^*, u_t^*))$.*

Comparison between the Percentage and Absolute Thresholds

As proved in Theorem 5.3, in a linear contract, the optimal task levels $\frac{1}{\lambda^*}$ and u^* are the same for a percentage threshold c and for an absolute one $k = mc$ if the agent is risk neutral. Thus one may expect that generally the two thresholds generate the same optimal task levels if we have the property

$$\mathbb{E}[U_1(cY_1)] = U_1(k). \tag{5.8}$$

It says that if the average disutility of the agent on the upper-bound proportion (the percentage threshold scenario) $|c|$ of accident cost is equal to the disutility of the agent on the upper-bound amount (the absolute threshold scenario) $|k|$ of accident cost, then the agent performs the same task levels.

However, although the property and its underlying intuition guarantee the result in Theorem 5.3, they cannot be applied to the case where the penalty-for-accident strategy is a nonlinear function. In fact, since the constraints (5.4) and (5.7) are inequalities and the optima may take place at interior points, condition (5.8) is far from enough. Therefore, we illustrate this disparity by the following example.

Our strategy is to first assume that $\mathbb{E}[U_1(cY_1)] = U_1(k)$ holds. Second, we take the optimal task levels for the percentage threshold where $c = -0.7$ and $c = -0.6$. Third, we plot the optimal task levels for the absolute threshold where k is such that $c = -0.7$ and $c = -0.6$. Last, we compare the corresponding optimal task levels.

To maintain the consistency, we use the same parameters as in Section 5.2.1. Specifically, we use our familiar pain function

$$P(\lambda, u) = (k_1 u)^2 + (k_2/\lambda)^2 + 2\theta k_1 k_2 u/\lambda$$

with $k_1 = k_2 = 1$ and $\theta = 0.1$. The accident size is assumed to be uniformly distributed over $[0, 2m]$ with $m = 1$. The risk averse agent is characterized by $p = 2$ and $q = 0.8$. The principal is more patient than the agent with $\rho = 0.1$ and $r = 0.05$.

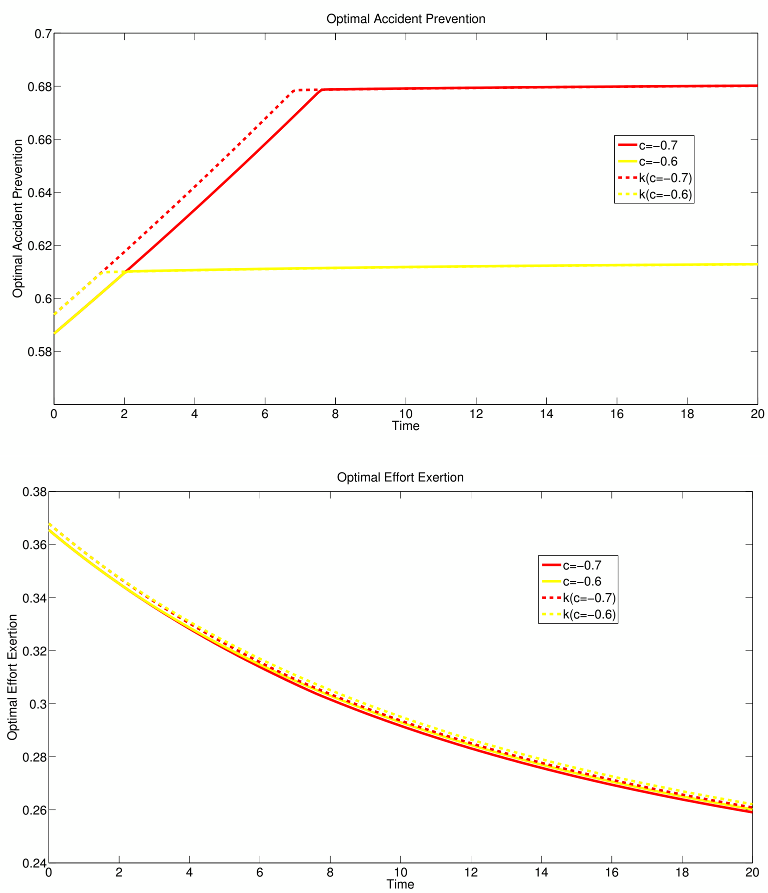


Figure 5.5: Comparison of Optimal Task Levels for Percentage and Absolute Thresholds

As shown in Figure 5.5, the optimal levels of accident prevention are different for the two kinds of thresholds that satisfy (5.8) before they reach their boundaries set by the thresholds. Similarly, the optimal levels of effort exertion are different throughout the entire time horizon.

Chapter 6

Conclusion

In a continuous-time payment framework, we have analyzed the optimal incentives in multitasking settings. The order of patience of the two players matters in the time trend of the optimal task levels and incentives. When the principal is more patient, he/she will induce the agent to do more prevention as time progresses, aiming to earn the benefit from extra accident reduction net the extra lump-sum payment. When the principal is more patient, the accident prevention has a U-shape as task interaction increases while effort exertion decreases. The principal punishes the risk-averse agent in the same magnitude no matter how big an accident is. The level of relative risk-aversion generates the "induced task shifts"; an agent who is highly risk-averse to sudden drops in income performs less prevention, resulting from the fact that the risk-neutral principal would rather bear the costs of accidents himself/herself than compensate the agent with a much higher lump-sum payment.

The partially limited liability protects the agent in that he/she pays only up to a exogenously given amount or percentage of accident costs. In a linear contract, since the risk-neutral agent only pays up to the absolute threshold, the principal charges him/her more for small accidents to gain a compensation. In such a contract, if the absolute and percentage threshold are chosen correspondingly, then they have the same effects on the optimal task levels, as opposed to the case where the contract is nonlinear and the agent is risk averse.

Appendix A

Proofs

A.1 Proof of Proposition 3.1

Suppose before time t, the agent applies any non-optimal actions $\tilde{\lambda}_s$ and $\tilde{u}_s$, $s \le t$, and applies the optimal actions $\hat{\lambda}_s$ and $\hat{u}_s$, $s > t$ after time t. Hence, the agent's continuation value at time t is

$$U_t(\tilde{\lambda}, \tilde{u}; \hat{\lambda}, \hat{u}) = \int_0^t \mathrm{e}^{-\rho s}[dL_s - P(\tilde{\lambda}_s, \tilde{u}_s)\, ds] + \mathrm{e}^{-\rho t} W_t(L, \hat{\lambda}_t, \hat{u}_t).$$

The corresponding evolution is

$$\begin{aligned} dU_t(\tilde{\lambda}, \tilde{u}; \hat{\lambda}, \hat{u}) &= \mathrm{e}^{-\rho t}[dL_t - P(\tilde{\lambda}_t, \tilde{u}_t)\ dt] + d(\mathrm{e}^{-\rho t} W_t(L, \hat{\lambda}_t, \hat{u}_t)) \\ &= \mathrm{e}^{-\rho t}[dL_t - P(\tilde{\lambda}_t, \tilde{\lambda}_t)\, dt] - \mathrm{e}^{-\rho t}[dL_t - P(\hat{\lambda}_t, \hat{u}_t)\, dt] \\ &\quad - \mathrm{e}^{-\rho t} H_t(L, \hat{\lambda}_t, \hat{u}_t)\, d\hat{M}_t - \mathrm{e}^{-\rho t} K_t(L, \hat{\lambda}_t, \hat{u}_t)\, d\hat{B}_t \\ &= \mathrm{e}^{-\rho t}[P(\hat{\lambda}_t, \hat{u}_t) - P(\tilde{\lambda}_t, \tilde{u}_t)]\, dt - \mathrm{e}^{-\rho t} H_t(L, \hat{\lambda}_t, \hat{u}_t)\, d\hat{M}_t \\ &\quad - \mathrm{e}^{-\rho t} K_t(L, \hat{\lambda}_t, \hat{u}_t)\, d\hat{B}_t. \end{aligned}$$

Then by changing the measure, the martingales are changed to $d\hat{M}_t = d\tilde{M}_t - m(\hat{\lambda}_t - \tilde{\lambda}_t)\,dt$ and $d\hat{B}_t = d\tilde{B}_t - (\hat{u}_t - \tilde{u}_t)\,dt$. Hence,

$$\begin{aligned} dU_t(\tilde{\lambda}, \tilde{u}; \hat{\lambda}, \hat{u}) = \mathrm{e}^{-\rho t}\{[&P(\hat{\lambda}_t, \hat{u}_t) + H_t(L, \hat{\lambda}_t, \hat{u}_t)m\hat{\lambda}_t + K_t(L, \hat{\lambda}_t, \hat{u}_t)\hat{u}_t \\ &- P(\tilde{\lambda}_t, \tilde{u}_t) - H_t(L, \hat{\lambda}_t, \hat{u}_t)m\tilde{\lambda}_t - K_t(L, \hat{\lambda}_t, \hat{u}_t)\tilde{u}_t]\,dt \\ &- H_t(L, \hat{\lambda}_t, \hat{u}_t)\,d\tilde{M}_t - K_t(L, \hat{\lambda}_t, \hat{u}_t)\,d\tilde{B}_t\}. \end{aligned}$$

Since $\hat{\lambda}_t$ and $\hat{u}_t$ are the optimal choices, $\tilde{\lambda}_t = \hat{\lambda}_t$ and $\tilde{u}_t = \hat{u}_t$ maximize $U_t(\tilde{\lambda}, \tilde{u}; \hat{\lambda}, \hat{u})$. Then for any action $\tilde{\lambda}_t$, $\tilde{u}_t$,

$$\begin{aligned} &P(\hat{\lambda}_t, \hat{u}_t) + H_t(L, \hat{\lambda}_t, \hat{\lambda}_t)m\hat{\lambda}_t + K_t(L, \hat{\lambda}_t, \hat{\lambda}_t)\hat{u}_t \\ &\leq P(\tilde{\lambda}_t, \tilde{u}_t) + H_t(L, \hat{\lambda}_t, \hat{u}_t)m\tilde{\lambda}_t + K_t(L, \hat{\lambda}_t, \hat{u}_t)\tilde{u}_t. \end{aligned}$$

In general, the drift term is nonpositive, and is zero when the optimal actions are attained. The first-order conditions yield the incentive compatibility conditions (3.6). These conditions are also sufficient because the maximizers are unique by the assumptions on the pain function. □

A.2 Properties of the Pain Function and Proof of Proposition 3.3

We restrict the pain function to the class of power functions. The pain function $P(\lambda, u)$ displays the following limit behavior, resulting from either mathematical properties or economic intuition.

For fixed $\bar{\lambda}_t$ and $\bar{u}_t$,

1. $\lim_{\lambda_t \to 0} P(\lambda_t, \bar{u}_t) = \infty$,
2. $\lim_{u_t \to \infty} P(\bar{\lambda}_t, u_t) = \infty$,
3. $\lim_{\lambda_t \to \infty} P(\lambda_t, \bar{u}_t) = P(\bar{u}_t)$,
4. $\lim_{u_t \to 0} P(\bar{\lambda}_t, u_t) = P(\bar{\lambda}_t)$,
5. $\lim_{\lambda_t \to 0} P_\lambda(\lambda_t, \bar{u}_t) = -\infty$,
6. $\lim_{u_t \to \infty} P_u(\bar{\lambda}_t, u_t) = \infty$,
7. $\lim_{\lambda_t \to \infty} P_\lambda(\lambda_t, \bar{u}_t) = 0$,
8. $\lim_{u_t \to 0} P_u(\bar{\lambda}_t, u_t) = P_u(\bar{\lambda}_t)$,
9. $\lim_{\lambda_t \to 0} P_u(\lambda_t, \bar{u}_t) = \infty$,
10. $\lim_{u_t \to \infty} P_\lambda(\bar{\lambda}_t, u_t) = -\infty$,
11. $\lim_{\lambda_t \to \infty} P_u(\lambda_t, \bar{u}_t) = P_u(\bar{u}_t)$,
12. $\lim_{u_t \to 0} P_\lambda(\bar{\lambda}_t, u_t) = P_\lambda(\bar{\lambda}_t)$.

Properties 5–12 reflect the intuition of increasing marginal pain. Properties 13–20 below are derived from the properties of power functions, Properties 1–12 and basic properties specified in the introduction of pain function.

13. $\lim_{\lambda_t \to 0} \lambda_t P_\lambda(\lambda_t, \bar{u}_t) = -\infty$, 14. $\lim_{u_t \to \infty} u_t P_u(\bar{\lambda}_t, u_t) = \infty$

respectively by 5, 6 and basic properties.

15. $\lim_{\lambda_t \to \infty} \lambda_t P_\lambda(\lambda_t, \bar{u}_t) = 0$, 16. $\lim_{u_t \to 0} u_t P_u(\bar{\lambda}_t, u_t) = 0$

respectively by 7, 8 and basic properties.

17. $\lim_{\lambda_t \to 0} \bar{u}_t P_u(\lambda_t, \bar{u}_t) = \infty$, 18. $\lim_{u_t \to \infty} \bar{\lambda}_t P_\lambda(\bar{\lambda}_t, u_t) = -\infty$

respectively by 9 and 10.

19. $\lim_{\lambda_t \to \infty} \bar{u}_t P_u(\lambda_t, \bar{u}_t) = \bar{u}_t P_u(\bar{u}_t)$, 20. $\lim_{u_t \to 0} \bar{\lambda}_t P_\lambda(\bar{\lambda}_t, u_t) = \bar{\lambda} t P_\lambda(\bar{\lambda}_t)$

respectively by 11 and 12.

Proof of Proposition 3.3. For a given finite $\bar{u}_s$, as λ_s tends to ∞, $\lim_{\lambda_s \to \infty} G_s < \infty$ is automatically satisfied because of 3, 15 and 19.

For a given finite $\bar{\lambda}_s$, as u_s tends to 0, $\lim_{u_s \to 0} G_s < \infty$ is automatically satisfied because of 4, 16 and 20.

For a given finite $\bar{u}_s$, as λ_s tends to 0, $\lim_{\lambda_s \to 0} G_s < \infty$ is satisfied only if

$$\limsup_{\lambda_s \to 0} \frac{(\mathrm{e}^{-\rho s} - \mathrm{e}^{-rs})\lambda_s P_\lambda(\lambda_s, \bar{u}_s) + \mathrm{e}^{-rs}\bar{u}_s}{\mathrm{e}^{-\rho s} P(\lambda_s, \bar{u}_s) - (\mathrm{e}^{-\rho s} - \mathrm{e}^{-rs})\bar{u}_s P_u(\lambda_s, \bar{u}_s)} < 1 \tag{A.1}$$

by 1, 13 and 17.

For a given finite $\bar{\lambda}_s$, as u_s tends to ∞, $\lim_{u_s \to \infty} G_s < \infty$ is satisfied only if

$$\limsup_{u_s \to \infty} \frac{(\mathrm{e}^{-\rho s} - \mathrm{e}^{-rs})\bar{\lambda}_s P_\lambda(\bar{\lambda}_s, u_s) + \mathrm{e}^{-rs} u_s}{\mathrm{e}^{-\rho s} P(\bar{\lambda}_s, u_s) - (\mathrm{e}^{-\rho s} - \mathrm{e}^{-rs}) u_s P_u(\bar{\lambda}_s, u_s)} < 1 \tag{A.2}$$

by 2, 14 and 18.

As λ_s tends to 0 and u_s tends to ∞, $\lim_{\frac{u_s}{\lambda_s} \to \infty} u_s P_u(\lambda_s, u_s) = \infty$ by 9 and 14, $\lim_{\frac{u_s}{\lambda_s} \to \infty} \lambda_s P_\lambda(\lambda_s, u_s) = -\infty$ by 10 and 13, and $\lim_{\frac{u_s}{\lambda_s} \to \infty} P(\lambda_s, u_s) = \infty$ by 1 and 2. Then $\lim_{\frac{u_s}{\lambda_s} \to \infty} G_s < \infty$ is satisfied only if

$$\limsup_{\frac{u_s}{\lambda_s} \to \infty} \frac{(\mathrm{e}^{-\rho s} - \mathrm{e}^{-rs})\lambda_s P_\lambda(\lambda_s, u_s) + \mathrm{e}^{-rs} u_s}{\mathrm{e}^{-\rho s} P(\lambda_s, u_s) - (\mathrm{e}^{-\rho s} - \mathrm{e}^{-rs}) u_s P_u(\lambda_s, u_s)} < 1.$$

Since (3.8) implies (A.1) and (A.2), then the optimal contract exists if (3.8) is satisfied. □

A.3 Proof of Proposition 4.3

We need to solve

$$\sup_{\lambda_s, u_s} G_s(\lambda_s, u_s),$$

where

$$G_s(\lambda_s, u_s) = \mathrm{e}^{-rs}[-\lambda_s(m + U_1^{-1}(P_\lambda(\lambda_s, u_s))) + u_s(1 - U_2^{-1}(P_u(\lambda_s, u_s)))]$$
$$+ \mathrm{e}^{-\rho s}[\lambda_s P_\lambda(\lambda_s, u_s) + u_s P_u(\lambda_s, u_s) - P(\lambda_s, u_s)].$$

Using the properties of the pain function, the limit behavior of G is analyzed below.

Given the properties of 3, 7, 11, 15 and 19, as λ_s tends to ∞ with fixed $\bar{u}_s$, $\lim_{\lambda_s \to \infty} G_s(\lambda_s, u_s) < \infty$ is satisfied if $\limsup_{\lambda_s \to \infty} -\lambda_s(m + U_1^{-1}(P_\lambda(\lambda_s, \bar{u_s}))) < \infty$, which is implied by $\limsup_{\lambda_s \to \infty} -\lambda_s U_1^{-1}(P_\lambda(\lambda_s, \bar{u_s}))) < \infty$.

Given the properties 4, 8, 12, 16 and 20, as u_s tends to 0 with fixed $\bar{\lambda}_s$, $\lim_{u_s \to 0} G_s(\lambda_s, u_s) < \infty$ is automatically satisfied.

Given the properties 1, 5, 9, 13 and 17, as λ_s tends to 0 with fixed $\bar{u}_s$, $\lim_{\lambda_s \to 0} G_s(\lambda_s, u_s) < \infty$ is satisfied if

$$\limsup_{\lambda_s \to 0} \frac{-\mathrm{e}^{-rs}\lambda_s U_1^{-1}(P_\lambda(\lambda_s, \bar{u}_s)) + \mathrm{e}^{-rs}\bar{u}_s + \mathrm{e}^{-\rho s}\bar{u}_s P_u(\lambda_s, \bar{u}_s)}{\mathrm{e}^{-rs}\bar{u}_s U_2^{-1}(P_u(\lambda_s, \bar{u}_s)) - \mathrm{e}^{-\rho s}\lambda_s P_\lambda(\lambda_s, \bar{u}_s) + \mathrm{e}^{-\rho s}P(\lambda_s, \bar{u}_s)} < 1. \quad \text{(A.3)}$$

Given the properties 2, 6, 10, 14 and 18, as u_s tends to ∞ with fixed $\bar{\lambda}_s$, $\lim_{u_s \to \infty} G_s(\lambda_s, u_s) < \infty$ is satisfied if

$$\limsup_{u_s \to \infty} \frac{-\mathrm{e}^{-rs}\bar{\lambda}_s U_1^{-1}(P_\lambda(\bar{\lambda}_s, u_s)) + \mathrm{e}^{-rs}u_s + \mathrm{e}^{-\rho s}u_s P_u(\bar{\lambda}_s, u_s)}{\mathrm{e}^{-rs}u_s U_2^{-1}(P_u(\bar{\lambda}_s, u_s)) - \mathrm{e}^{-\rho s}\bar{\lambda}_s P_\lambda(\bar{\lambda}_s, u_s) + \mathrm{e}^{-\rho s}P(\bar{\lambda}_s, u_s)} < 1. \quad \text{(A.4)}$$

Hence, combining (A.3) and (A.4), we find the following sufficient conditions for the existence of a maximum of $G_s(\lambda_s, u_s)$

$$\limsup_{\frac{u_s}{\lambda_s}\to\infty} \frac{-\mathrm{e}^{-rs}\lambda_s U_1^{-1}(P_\lambda(\lambda_s,u_s)) + \mathrm{e}^{-rs}u_s + \mathrm{e}^{-\rho s}u_s P_u(\lambda_s,u_s)}{\mathrm{e}^{-rs}u_s U_2^{-1}(P_u(\lambda_s,u_s)) - \mathrm{e}^{-\rho s}\lambda_s P_\lambda(\lambda_s,u_s) + \mathrm{e}^{-\rho s}P(\lambda_s,u_s)} < 1,$$

and

$$\limsup_{\lambda_s\to\infty} -\lambda_s U_1^{-1}(P_\lambda(\lambda_s,\bar{u}_s))) < \infty. \qquad \square$$

A.4 Proof of Theorem 5.2

We first calculate

$$\begin{aligned}
P_\lambda(\lambda_s,u_s) &= \mathbb{E}[\max\{\alpha_s Y_{N_s}, k\}] \\
&= \mathbb{E}\Big[\max\{\alpha_s Y_{N_s}, k\}\mathbb{1}_{\alpha_s Y_{N_s}\le k} + \max\{\alpha_s Y_{N_s}, k\}\mathbb{1}_{\alpha_s Y_{N_s}> k}\Big] \\
&= \mathbb{E}\Big[k\mathbb{1}_{\alpha_s Y_{N_s}\le k} + \alpha_s Y_{N_s}\mathbb{1}_{\alpha_s Y_{N_s}> k}\Big] \\
&= k\mathbb{P}(\alpha_s Y_{N_s} \le k) + k\frac{\alpha_s}{k}\mathbb{E}\Big[Y_{N_s}\mathbb{1}_{\alpha_s Y_{N_s}> k}\Big] \\
&= k\mathbb{P}\Big(Y_{N_s} \ge \frac{k}{\alpha_s}\Big) + k\frac{\alpha_s}{k}\mathbb{E}\Big[Y_{N_s}\mathbb{1}_{Y_{N_s}<\frac{k}{\alpha_s}}\Big].
\end{aligned}$$

Now suppose $P_\lambda(\lambda_s,u_s) = k$, then

$$\mathbb{P}\Big(Y_{N_s} \ge \frac{k}{\alpha_s}\Big) + \frac{\alpha_s}{k}\mathbb{E}\Big[Y_{N_s}\mathbb{1}_{Y_{N_s}<\frac{k}{\alpha_s}}\Big] = 1. \tag{A.5}$$

To show that this implies $\alpha_s = -\infty$, we distinguish two cases.

Suppose first $\mathbb{P}\Big(Y_{N_s} < \frac{k}{\alpha_s}\Big) > 0$, then

$$\frac{\alpha_s}{k}\mathbb{E}\Big[Y_{N_s}\mathbb{1}_{Y_{N_s}<\frac{k}{\alpha_s}}\Big] < \frac{\alpha_s}{k}\mathbb{E}\Big[\frac{k}{\alpha_s}\mathbb{1}_{Y_{N_s}<\frac{k}{\alpha_s}}\Big] = \mathbb{P}(Y_{N_s} < \frac{k}{\alpha_s})$$

for finite α_s. Note that $\mathbb{P}\Big(Y_{N_s} \geq \frac{k}{\alpha_s}\Big) + \mathbb{P}\Big(Y_{N_s < \frac{k}{\alpha_s}}\Big) = 1$ by definition. Hence, for (A.5) to hold, the only value that α_s can take is $-\infty$.

Suppose $\mathbb{P}\Big(Y_{N_s} < \frac{k}{\alpha_s}\Big) = 0$, then $Y_{N_s} \geq \frac{k}{\alpha_s}$ a.s.. Hence $\max\{\alpha_s Y_{N_s}, k\} = k$ a.s.. We can equally well choose $\alpha_s = -\infty$ because $\max\{\alpha_s Y_{N_s}, k\} = k$ a.s. still holds and only the value of $\max\{\alpha_s Y_{N_s}, k\}$ and not α_s itself is relevant.

In summary, an optimal α_s is $\alpha_s = -\infty$ if $P_\lambda(\lambda_s, u_s) = k$. □

A.5 Proof of Theorem 5.3

Now we consider the following form of payment stream

$$dL_s = \max\{\alpha_s, c\} Y_{N_s}\, \Delta N_s + \beta_s\, ds + dA_s + \gamma_s\, dx_s,$$

where $\Delta N_s = N_s - N_{s-}$ and $c \in (-\infty, 0)$ is fixed. The agent is protected in that he/she only pays at most for a proportion c of the accident costs.

The agent's problem is

$$\sup_{\lambda, u} \mathbb{E}\bigg[\int_0^T \mathrm{e}^{-\rho s}(dL_s - P(\lambda_s, u_s)\, ds)\bigg].$$

The principal's problem is

$$\sup_{L} \mathbb{E}\bigg[\int_0^T \mathrm{e}^{-rs}(dx_t - dJ_s - dL_s)\bigg]$$

under the constraint that

$$\sup_{\lambda, u} \mathbb{E}\bigg[\int_0^T \mathrm{e}^{-\rho s}(dL_s - P(\lambda_s, u_s)\, ds)\bigg] \geq R_0.$$

Applying the method in Section 5.1.2, we have the incentive compatibility conditions

$$P_\lambda(\lambda_s, u_s) = \max\{\alpha_s, c\}m, \quad P_u(\lambda_s, u_s) = \gamma_s.$$

Hence, the reduced principal's problem is

$$\sup_{\lambda,u} G_s(\lambda_s, u_s)$$

subject to

$$P_\lambda(\lambda_s, u_s) \geq cm$$

where

$$G_s(\lambda_s, u_s) = \mathrm{e}^{-rs}[u_s - u_s P_u(\lambda_s, u_s) - \lambda_s m - \lambda_s P_\lambda(\lambda_s, u_s)]$$
$$+ \mathrm{e}^{-\rho s}[\lambda_s P_\lambda(\lambda_s, u_s) + u_s P_u(\lambda_s, u_s) - P(\lambda_s, u_s)].$$

Set $k = cm$, and we have the result of Theorem 5.3. $\square$

Bibliography

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

A. Capponi and C. Frei: Dynamic Contracting: Accidents Lead to Nonlinear Contracts. Preprint (2015) available at `www.math.ualberta.ca/~cfrei`

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.

J. Cvitanić and J. Zhang: *Contract Theory in Continuous-Time Models.* Springer Finance, Springer-Verlag Berlin, 2013.

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

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

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.

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

H. Ou-Yang: Optimal Contracts in a Continuous-Time Delegated Portfolio Management Problem. *The Review of Financial Studies* **16**, 173–208, 2003.

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

Y. Sannikov: Dynamic Security Design and Corporate Financing. In: G. M. Constantinides, M. Harris and R. Stulz (eds.), *Handbook of the Economics of Finance*, Elsevier, Vol. **2**, 71–122, 2013.

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.