Insurance: Mathematics and Economics 109 (2023) 69–93

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

# Optimal insurance contracts for a shot-noise Cox claim process and persistent insured's actions

Wenyue Liu, Abel Cadenillas *

*Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta T6G 2G1, Canada*

ARTICLE INFO

*Article history:*
Received May 2022
Received in revised form November 2022
Accepted 15 January 2023
Available online 20 January 2023

*JEL classification:*
C02
C61
C65
D86
G22
G52

*Keywords:*
Optimal insurance contract
Optimal risk sharing
Shot-noise Cox process
Persistent actions
Continuous-time stochastic control

ABSTRACT

We consider a continuous-time model in which an insurer proposes an insurance contract to a potential insured. Motivated by climate change and catastrophic events, we assume that the number of claims process is a shot-noise Cox process. The insurer selects the premium to be paid by the potential insured and the amount to be paid for each claim. In addition, the insurer can request some actions from the potential insured to reduce the number of claims. The insurer wants to maximize his expected total utility, while the potential insured signs the contract if his expected total utility for signing the contract is greater than or equal to his expected total utility when he does not sign the contract. We obtain an analytical solution for the optimal premium, the optimal amount to be paid for each claim, and the optimal actions of the insured. This leads to interesting managerial insights.



## 1. Introduction

We consider a finite-horizon, continuous-time model in which an insurer proposes an insurance contract to a potential insured.

It has been standard in the actuarial sciences literature to assume that the total claim amount process is a compound Poisson process with deterministic intensity, or equivalently that the number of claims process is a Poisson process with deterministic intensity. See, for example, Bühlmann (1970), Medhi (1982), Lindskog and McNeil (2003), and Moore and Young (2006). However, there are many important cases in which a Poisson process with deterministic intensity does not represent well the total number of claims. For instance, Beard et al. (1984) show that the standard Poisson process is not an appropriate model for the number of claims in catastrophe, fire, and some other types of insurance. Instead, Beard et al. (1984) suggest considering stochastic intensity.

The Cox process, also called doubly stochastic Poisson process, is a generalized Poisson process with stochastic intensity. We consider a Cox process where the intensity is a shot noise process. The shot noise process can be used to model the stochastic nature of catastrophic events. Due to climate change, natural disasters occur more frequently. The losses caused by catastrophes are usually enormous, so it is important to insure against losses caused by this type of events. Dassios and Jang (2003) explain that claims arising from catastrophic events depend on the intensity of natural disasters, and that one of the processes that can be used to measure the impact of catastrophic events is the shot noise process. Further, Dassios and Jang (2003) and Schmidt (2014) explain in detail the application of shot-noise Cox process in catastrophe insurance, although they do not study optimal insurance contracts. Following Dassios and Jang (2003) and Schmidt (2014), we adopt a shot-noise Cox process to count the number of claims. Besides catastrophe insurance, our model is also appropriate to other types of insurance. For example, Dassios et al. (2015) point out that the shot-noise Cox process models very well the number of traffic accidents if the rate of the event arrival is large.

* Corresponding author.
*E-mail addresses:* wenyue2@ualberta.ca (W. Liu), abel@ualberta.ca (A. Cadenillas).

https://doi.org/10.1016/j.insmatheco.2023.01.002
0167-6687/

We consider two cases: the insured does not intervene through his actions to reduce the number of claims, and the insured intervenes through his actions to reduce the number of claims. In the first case, we assume that the number of claims process is a shot-noise Cox process. In the second case, we assume that the number of claims process is a Cox process but the actions of the insured can affect the shot noise intensity. Equation (2) shows how the insured's actions $a = \{a_t; t \in [0, T]\}$ affect the intensity. The first case is the special case of the second case in which the actions of the insured are null. This is the first paper on optimal insurance contracts in which the number of claims is modeled by a Cox process with shot noise intensity.

We allow the actions of the insured to be persistent. That is, the actions of the insured at any point in time are effective until maturity. For instance, in flood insurance, the insurer may require the insured to bring the property up to some standards. See the national flood insurance program of the Federal Emergency Management Agency (2022). This action of the property owner will reduce the probability of having a loss caused by floods, and its protection against flood will last from the time of action. However, along with aging and wear, the flood-resistance equipment becomes less protective over time. Thus, we further assume that the action is discounted by time. We will discuss further details of persistent actions in Section 2. Hoffmann et al. (2021), Hopenhayn and Jarque (2010), Jarque (2010), and Mukoyama and Şahin (2005) have also considered persistent actions. We present a model in which persistent actions affect a Cox process.

The insurer selects the premium to be paid by the potential insured, the amount to be paid for each claim, and also requests some actions from the potential insured. The potential insured has a cost associated with his actions. Section 3 presents details on the utility and cost functions of the insurer and the potential insured. The insurer wants to maximize his expected total utility, while the potential insured signs the contract if his expected total utility for signing the contract is greater than or equal to his expected total utility when he does not sign the contract. Thus, the problem studied in our paper is different from other papers (such as Zou and Cadenillas (2014), and Zou and Cadenillas (2017)) in which an insurer has already designed an insurance contract (which might not be the optimal insurance contract) and decides its optimal liability. To the best of our knowledge, we obtain, for the first time in the literature, an analytical solution for the optimal premium, the optimal amount to be paid for each claim, and the optimal actions of the insured when the number of claims process is a Cox process. The analytical solution leads to interesting managerial insights. For instance, we show that the optimal expected action decreases over time. Furthermore, the insured will perform less expected action over time to reach the reservation utility when he does not enter the insurance market. Jarque (2010) presents the same trend of the optimal action only through a numerical example while we prove it with an analytical solution in a general setting. Our result challenges the assumption of Mukoyama and Şahin (2005) that the principal prefers the agent to insert the highest action all the time. The decreasing trend of the optimal actions results from action persistence, where the earlier action reduces the loss further because it is effective for a relatively long period. We also present an example.

Section 2 presents the total claim amount model and Section 3 presents the problem that we study in this paper. The solution is presented in Section 4. Section 5 discusses the reservation utility. An example is presented in Section 6. The conclusions are presented in Section 7. The proofs of the theorems, propositions, corollaries, and lemmas are presented in Appendix A.

## 2. The total claim amount process

We consider a finite time horizon $[0, T]$. There are two possibilities: the insured does not affect the risky external environment and the insured affects the risky external environment.

If the insured does not affect the risky external environment, then the total claim amount process $S = \{S(t); t \in [0, T]\}$ is given by

$$S(t) = \sum_{i=1}^{N(t)} L_i = L_1 + L_2 + \cdots + L_{N(t)},$$

where $N(t)$ is the number of claims up to time $t \in [0, T]$ and $\{L_1, L_2, \cdots, L_{N(t)}\}$ are the amounts claimed until time $t$. We make the following assumptions.

a) The random variables $\{L_1, L_2, L_3, \cdots\}$ are independent and identically distributed. Furthermore, their range is $R_L$ and $\inf R_L > 0$.
b) The sequence of random variables $\{L_1, L_2, L_3, \cdots\}$ are independent of the stochastic process $N = \{N(t); t \in [0, T]\}$.
c) The stochastic process $N = \{N(t); t \in [0, T]\}$ is a shot-noise Cox process with stochastic intensity rate $I = \{I(t); t \in [0, T]\}$ given by

$$I(t) = \theta \sum_{i=0}^{M(t)} Y_i e^{\delta(\tau_i - t)} = \theta \sum_{i=0}^{M(t)} Y_i e^{-\delta(t-\tau_i)}. \tag{1}$$

In the above equation, $\theta$ represents the risk level of the insured, $M(t)$ counts the number of risky events exposed to the insured from time 0 to time $t$, $Y_i$ is the jump size caused by the $i$-th random risky event, $\tau_i$ is the time when the $i$-th risky event occurs, and $\delta$ is the rate of decay. The effect of a risk event happening at time $\tau$ lasts in the time period $[\tau, T]$ but is discounted by $\delta$ at time $t \in [\tau, T]$. We make the following assumptions about the stochastic process $I$:

c1) $\theta$ is a positive constant.
c2) $M = \{M(t); t \in [0, T]\}$ is a Poisson process with a deterministic intensity process $\rho(t) \geq 0$, $t \in [0, T]$. If the frequency of exposures is high, then $\rho(\cdot)$ is large.
c3) $\{Y_i\}_{i=1,2,3,\dots}$ is a sequence of i.i.d. random variables and independent of $M$. We suppose they are the images of a random variable $Y$ that is positive and finite almost surely. $Y_0 > 0$ is a constant known at time 0. We denote $\mu = E[Y]$.
c4) $\{\tau_i\}_{i=1,2,3,\dots}$ is a sequence of non-decreasing stopping times. In the above equation, $\tau_0 = 0$, and for every $i \in \{1, 2, \cdots, M(t)\}$: $\tau_i \leq t$.
c5) $\delta$ is a positive constant.

Applications of Cox processes with shot noise intensity to insurance can be found in Albrecher and Asmussen (2006), Macci and Torrisi (2011), Schmidt (2014), and Zhu (2013). The number of claims from catastrophic events depends on the stochastic intensity of natural disasters. The above intensity process $I$ measures the frequency of external risky events (by $M$), their magnitude (by $Y_i$), and their time (by $\tau_i$) to determine the effect of catastrophic events. As time passes, the magnitude decreases (by $\delta$). We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ together with a filtration $\mathbb{F} := \{\mathcal{F}_t, t \in [0, T]\}$ that is the $\mathbb{P}$-augmentation of the natural filtration

$$\sigma\left(N(s), M(s), s \in [0,t]; L_i, i \in \{0,1,\cdots,N(t)\}; Y_j, \tau_j, j \in \{0,1,\cdots,M(t)\}\right).$$

If the insured affects the risky external environment, the total claim amount process $S = \{S(t); t \in [0,T]\}$ is given by

$$S(t) = \sum_{i=1}^{N^a(t)} L_i = L_1 + L_2 + \cdots + L_{N^a(t)},$$

where the number of claims process $N^a = \{N^a(t); t \in [0,T]\}$ is a Cox process with stochastic intensity rate $\lambda = \{\lambda(t); t \in [0,T]\}$ given by

$$\lambda(t) := \theta\left(\sum_{i=0}^{M(t)} Y_i e^{\delta(\tau_i - t)}\right)\left(1 - e^{-t}(A_0 + \int_0^t a_s r_s e^s ds)\right). \tag{2}$$

Here, the process $a = \{a_t; t \in [0,T]\}$ represents the actions to reduce the magnitude of external risk events and $A_0$ is a constant that represents the measures to reduce the magnitude of risk events taken before the contract is implemented. We assume that $a$ is adapted to the filtration $\mathbb{F}$. We also assume that $0 \le a_t \le K$ for $t \in [0,T]$ and $A_0 \in [0,K]$, where $K \in [0,1]$ is a constant that represents the proportion of the intensity that can be cleared through actions. The remaining $1-K$ proportion of the intensity is not avoidable through actions. $r_s$ is the effectiveness of action $a_s$. The process $r = \{r_t; t \in [0,T]\}$ is called the productivity of action in the principal-agent problem (Williams (2009)). Demarzo and Learning (2017) and Cvitanić and Zhang (2013) also introduce the coefficient $r_s$ to adjust for the action $a_s$. For example, the precaution against flood is more effective in the rainy season than in the dry season. Correspondingly, in flood insurance, $r_s$ is generally larger in rainy seasons. We assume that $r_s \in [0,1]$ for every $s \in [0,T]$. If the action and the effective rate take their highest values $K$ and 1 respectively at every time in $[0,T]$, then (2) becomes $\theta\left(\sum_{i=0}^{M(t)} Y_i e^{\delta(\tau_i - t)}\right)(1-K)$. Under the conditions that $0 \le A_0 \le 1$, $0 \le a_s \le 1$, and $0 \le r_s \le 1$ for $s \in [0,T]$, we have that $\lambda(t)$ is nonnegative for $t \in [0,T]$. In other words, the intensity of the random variable $N^a(t)$ is nonnegative for $t \in [0,T]$. In the special case where $A_0 = 0$ and $a_s = 0$ for every $s \in [0,T]$, we have for every $t \in [0,T]$: $I(t) = \lambda(t)$. Hence, the case in which the insured affects the external risk environment is more general than the case in which the insured does not affect the external risk environment.

Therefore, we assume that the insured can affect the external risk environment. In other words, we assume that the number of claims is represented by the stochastic process $N^a = \{N^a(t); t \in [0,T]\}$, which is a Cox process with the stochastic intensity rate $\lambda = \{\lambda(t); t \in [0,T]\}$ defined in (2).

We can understand the actions $a = \{a_t; t \in [0,T]\}$ in the intensity process from the following four aspects. First, the more actions inserted, the smaller the intensity is. Second, $a_s$ has an effect on $\lambda(t)$ for every $t \in [s,T]$. Thus, an earlier action can play a role for a long time while a late action plays a role only for a short time. Particularly, $a_T$ is effective for almost zero duration. Third, the ratio between the weights of $a_{s'}$ and $a_s$ in (2) is $e^{(s'-s)}$ if $0 < s', s \le t$. If it is closer to time $t$ when an action is implemented, the action is more effective at time $t$. Fourth, the action $a_s$ is made at time $s$. As time passes by, the contribution of $a_s$ shrinks by $e^{s-t}$ at time $t \in (s,T]$.

In the case of flood insurance, the insured is a property owner and the risk event is a flood. We denote by $Y_i$ the magnitude of the $i$-th flood. The risk events can affect the frequency of claims, so we represent them in the intensity rate process $\lambda = \{\lambda(t); t \in [0,T]\}$. The effect of each risk event lasts for some time, but it is discounted (by $\delta$) as time passes. For instance, the destructive power of a flood lasts from the time of flood rising to the time of cleaning up. However, the effect of the flood is weaker as time goes by. The process $a$ represents actions, like using flood-resistance materials, that the property owner is required to take to reduce the frequency of claims.

## 3. The insurance problem

We assume symmetric information, in the sense that all the information is transparent and accessible to both insurer and insured. The information structure is denoted by $\mathbb{F} = \{\mathcal{F}_t, t \in [0,T]\}$ and the model is constructed on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Following the principal-agent literature that considers a representative principal and a representative agent (see, for instance, Section 4.1 of Bolton and Dewatripont (2005), Cadenillas et al. (2007) and Continuous (2008)), we consider a representative insurer and a representative insured.

The insurer selects the premium rate and the compensation. During the contract period, the client will pay the premium continuously. The company commits to compensate the insured immediately after he faces a loss. The compensation can cover partially or completely the loss. The insurer observes all the information, in particular, the insured's actions. The insurer requires the amount of action in the contract, and that must be followed by the insured. That is consistent with many papers on optimal contract theory. Under the full information case, Cvitanić and Zhang (2013) points out that the principal offers the contract and dictates the agent's actions. In the full information section, Williams (2015) also said the principal decides the actions. The first-best models in Chapter 4 of Bolton and Dewatripont (2005) expressed the same ideas. In practice, to reduce losses, the insurance company may write down provisions that require the insured to take designated actions in catastrophe and other insurance contracts. For example, the catastrophe insurance policy may require the insured to do necessary maintenance on the property. Otherwise, the insurer is entitled to deny compensation for the loss directly or indirectly caused by the lack of maintenance. See, for instance, Flex Insurance Company (2022). Thus, we suppose the actions are taken to maximize the insurer's utility in this paper. On the other hand, the insured will sign the contract if his participation constraint is satisfied. We denote by

$$(a, q, P) = \{(a_t, q_t, P_i); t \in [0,T] \text{ and } i = 1, 2, \cdots\}$$

the contract offered by the insurer. After signing the contract, the insured pays continuously the premium rate $q_t$ and takes action $a_t$ at time $t$. When the i-th loss happens, the insurer compensates the insured with the amount $P_i$. We do not assume that $P_i$ is equal to $L_i$.

We assume that the insured and the insurer have Von Neumann-Morgenstern utility functions $U_1 : \mathbb{R} \mapsto \mathbb{R}$ and $U_2 : \mathbb{R} \mapsto \mathbb{R}$, respectively. These utility functions are strictly increasing, concave, and twice differentiable with the following properties:

$$
\begin{aligned}
&U_1(0), U_2(0) \le 0, \\
&U_1'(-\infty) = \lim_{x\to-\infty} U_1'(x) = +\infty, \;\; U_1'(+\infty) = \lim_{x\to+\infty} U_1'(x) = 0, \\
&U_2'(-\infty) = \lim_{x\to-\infty} U_2'(x) = +\infty, \;\; U_2'(+\infty) = \lim_{x\to+\infty} U_2'(x) = 0.
\end{aligned} \tag{3}
$$

The insurer's expected total utility for a policy $(a, q, P)$ is

$$
\mathcal{J}(q, P, a) := E\left[\int_0^T U_2(q_t)dt + \sum_{i=1}^{N^a(T)} U_2(-P_i)\right]. \tag{4}
$$

The cost function of action is denoted by $V_1$, and is assumed to be positive, increasing, differentiable, strictly convex and satisfying $V_1(0) = V_1'(0) = 0$. Next, we present the participation constraint. We denote the reservation utility by $R \in \mathbb{R}$. $R$ is the expected total utility that the insured can obtain from outside options. The insurer wants to offer a contract that gives an expected total utility greater than or equal to $R$ to the insured. Otherwise, the insured will prefer outside options, and will not accept the contract offer.

The income rate of the insured is represented by $\{w_t, t \ge 0\}$. We assume that $w_t > 0$ is deterministic for every $t \ge 0$.

We denote by $\mathcal{A}$ the class of admissible controls. These are the controls $(a, q, P)$ that are adapted to the filtration $\mathbb{F}$.

**Problem 1.** The insurer wants to select the policy $(\hat{a}, \hat{q}, \hat{P}) \in \mathcal{A}$ that solves the problem

$$
\max_{(q,P,a)\in\mathcal{A}} \;\; \mathcal{J}(q, P, a)
$$

$$
s.t. \;\; E\left[\int_0^T U_1(w_t - q_t)dt + \sum_{i=1}^{N^a(T)} U_1(P_i - L_i) - \int_0^T V_1(a_t)dt\right] \ge R, \tag{5}
$$

$$
0 \le a_t \le K, \text{ for all } t \in [0, T]. \tag{6}
$$

In (3), we assume the utility functions are negative when the variables are negative. The insurer loses some amount of utility if a compensation is made and the insured loses some amount of utility if he encounters the loss from an accident. From the terms $\sum_{i=1}^{N^a(T)} U_2(-P_i)$ in (4) and $\sum_{i=1}^{N^a(T)} U_1(P_i - L_i)$ in (5), we observe that the total loss of utility due to the claims can be reduced by taking actions.

## 4. The optimal insurance contract

An extended generator on Markov processes consisting of random jumps is explicitly calculated in Theorem 5.5 in Davis (1984). Following this theorem, we will present a generator of the process $\{(I(t), t), t \ge 0\}$. The generator helps with our calculation of the expectation of $N^a(T)$. We denote the cumulative distribution function of the jump $Y$ by $F_Y$. We assume that $F_Y$ and the intensity $\rho$ defined in Section 3 are Riemann integrable.

Suppose a function $f(\cdot,\cdot)$ belongs to the domain of the generator denoted by $\mathbb{A}$. Then $\mathbb{A}$ acting on $f(I, t)$ is defined by

$$
\mathbb{A} f(I, t) := \frac{\partial f}{\partial t} - \delta I \frac{\partial f}{\partial I} + \rho(t) \int_0^\infty f(I + \theta y, t) dF_Y(y) - \rho(t) f(I, t). \tag{7}
$$

Theorem 5.5 of Davis (1984) describes the domain of the generator, and Dassios and Jang (2003) give sufficient conditions under which $f$ is in the domain of $\mathbb{A}$. In our case, $f : [0, \infty) \times [0, T] \mapsto \mathbb{R}$ belongs to the domain of $\mathbb{A}$ if $f \in C^1([0, \infty) \times [0, T]; \mathbb{R})$ and

$$
\left| \int_0^\infty f(I + \theta y, t) dF_Y(y) - f(I, t) \right| < \infty.
$$

As stated by Proposition 1 in Dassios and Embrechts (1989), $\{f(I_t, t), t \ge 0\}$ is a martingale if $\mathbb{A} f(I, t) = 0$. See also Davis (1984). Therefore, we have the following result.

**Lemma 1.** *The stochastic process*

$$
\sum_{i=0}^{M(t)} Y_i e^{\delta \tau_i} - \mu \int_0^t e^{\delta u} \rho(u) du
$$

*is a martingale.*

**Proof.** See Appendix A. □

Now we can obtain the expected number of claims.

**Proposition 1.** *The expected number of claims corresponding to actions* $a = \{a_s, s \in [0, T]\}$ *is*

$$\begin{aligned} E[N^a(T)] =& \theta \int_0^T (1 - e^{-t} A_0) e^{-\delta t} \left( Y_0 + \mu \int_0^t \rho(u) e^{\delta u} du \right) dt \\ &- \theta \int_0^T e^{-(1+\delta)t} E\left[ \int_0^t a_s r_s e^s \left( \mu \int_s^t \rho(u) e^{\delta u} du + \sum_{i=0}^{M(s)} Y_i e^{\delta \tau_i} \right) ds \right] dt. \end{aligned} \tag{8}$$

**Proof.** See Appendix A. □

Changing the order of integration, we can obtain another way to express (8).

$$\begin{aligned} E[N^a(T)] =& \theta \int_0^T (1 - e^{-t} A_0) e^{-\delta t} \left( Y_0 + \mu \int_0^t \rho(u) e^{\delta u} du \right) dt \\ &- \theta E\left[ \int_0^T a_s r_s e^s \int_s^T e^{-(1+\delta)t} \left( \mu \int_s^t \rho(u) e^{\delta u} du + \sum_{i=0}^{M(s)} Y_i e^{\delta \tau_i} \right) dt ds \right]. \end{aligned}$$

The role actions $a = \{a_s, s \in [0, T]\}$ play can also be observed through the expression above. The integration following $a_s$ is from time $s$ to $T$. It indicates that the effect of $a_s$ lasts in the time period $[s, T]$. The action exerted at different moments makes different contributions in the remaining period.

We denote

$$\begin{aligned} \bar{B} &:= \int_0^T (1 - e^{-t} A_0) e^{-\delta t} \left( Y_0 + \mu \int_0^t \rho(u) e^{\delta u} du \right) dt, \\ B_t &:= r_t e^t \int_t^T e^{-(1+\delta)s} \left( \mu \int_t^s \rho(u) e^{\delta u} du + \sum_{i=0}^{M(t)} Y_i e^{\delta \tau_i} \right) ds. \end{aligned}$$

Now, we can write $E[N^a(T)]$ as

$$E[N^a(T)] = \theta \bar{B} - \theta E\left[ \int_0^T a_t B_t dt \right]. \tag{9}$$

Since $r_t \geq 0$, $\rho(t) \geq 0$ for $t \in [0, T]$, and $Y_i > 0$ for $i = 0, 1, 2, \cdots$, it immediately follows that $B_t \geq 0$ for each $\omega \in \Omega$ and $t \in [0, T]$. Recalling that $\lambda(t)$ is nonnegative for $t \in [0, T]$, we derive that $E[N^a(T)] \geq 0$. Let $a_s = 1$ almost surely for $s \in [0, T]$, we can see $E[N^a(T)] = \theta \left( \bar{B} - E\left[ \int_0^T B_t dt \right] \right)$ from (9). Further, let $A_0 = 1$ and $r_t = 1$ almost surely for $t \in [0, T]$, then $\lambda(t) = 0$ almost surely for $t \in [0, T]$ and it results in $E[N^a(T)] = 0$. It follows that $\bar{B} = E\left[ \int_0^T B_t dt \right]$. Otherwise, $\bar{B} > E\left[ \int_0^T a_t B_t dt \right]$. $\bar{B}$ can be understood as the expected number of claims if actions are not involved. $B_t$ is the intensity rate of accidents that can be removed by one unit of action at time t.

To find the solution of the model, we use the Lagrangian method and define the functional $\mathcal{L}_1$ by

$$\begin{aligned} \mathcal{L}_1(q, P, a; \Lambda_1, \Lambda_2) :=& E\left[ \int_0^T U_2(q_t) dt + \sum_{i=1}^{N^a(T)} U_2(-P_i) \right] \\ &+ \Lambda_1 E\left[ \int_0^T U_1(w_t - q_t) dt + \sum_{i=1}^{N^a(T)} U_1(P_i - L_i) - \int_0^T V_1(a_t) dt \right] + E\left[ \int_0^T \Lambda_2^t a_t dt \right], \end{aligned} \tag{10}$$

where $\Lambda_1$ and $\Lambda_2^t$, adapted to $\mathbb{F}$, $t \in [0, T]$ are Lagrangian multipliers. The first order conditions for $q$ and $P$ are

$$U_2'(-P_i) - \Lambda_1 U_1'(P_i - L_i) = 0 \text{ and } U_2'(q_t) - \Lambda_1 U_1'(w_t - q_t) = 0. \tag{11}$$

Since $\Lambda_1$ is constant, the solution of $P_i$ from the equations above is dependent of $L_i$ only. Hence, the sequences $\{U_2(P_i)\}_{i=1,2,\cdots}$ and $\{U_1(P_i - L_i)\}_{i=1,2,\cdots}$ are i.i.d. and independent of the process $N^a$. Thus, the Lagrangian (10) can be rewritten as

$$\mathcal{L}_1(q,P,a;\Lambda_1,\Lambda_2)=E\left[\int_0^T U_2(q_t)dt\right]+E[N^a(T)]E[U_2(-P)+\Lambda_1U_1(P-L)] \\ +\Lambda_1E\left[\int_0^T U_1(w_t-q_t)dt\right]-\Lambda_1E\left[\int_0^T V_1(a_t)dt\right]+E\left[\int_0^T \Lambda_2^t a_t dt\right]. \tag{12}$$

Derive the first order condition from (12) for $a_t$ to obtain

$$\Lambda_1 V_1'(a_t)-\Lambda_2^t=-\theta E[U_2(-P)+\Lambda_1U_1(P-L)]B_t \tag{13}$$

for each $t\in[0,T]$ and $\omega\in\Omega$. The values of the Lagrangian multipliers can show important information of the solutions. Consider $\Lambda_1$ first. To ensure the first order condition (11) valid, $\Lambda_1$ must be positive. If $\Lambda_1=0$, we can get $P_i=-\infty$ and $q_t=\infty$ from (11). However, this causes a contradiction to constraint (5). From (3), we have $\lim_{P_i\to-\infty}U_1(P_i-L_i)=-\infty$ for $i=1,2,\cdots$ and $\lim_{q_t\to\infty}U_1(w_t-q_t)=-\infty$ for $t\in[0,T]$. Then, the left-hand-side of (5) is going to $-\infty$. Since $R$ is finite, (5) can not be satisfied. Hence, $\Lambda_1>0$. Consider $\Lambda_2^t$ now. If $\Lambda_2^t=0$ for some $t\in[0,T]$ and some $\omega\in\Omega$, it means the constraint (6) is not binding. The action we obtain from (13),

$$a_t=V_1'^{-1}\left(-\frac{\theta}{\Lambda_1}E[U_2(-P)+\Lambda_1U_1(P-L)]B_t\right),$$

satisfies (6). If $\Lambda_2^t<0$ for some $t\in[0,T]$ and some $\omega\in\Omega$, it means the RHS of (13) is big enough such that

$$\Lambda_1V_1'(K)<-\theta E[U_2(-P)+\Lambda_1U_1(P-L)]B_t,$$

which shows the marginal cost of action is always smaller than the marginal benefit. Inserting actions more than $K$ will bring the company more utility, but this preference is prevented by the upper bound of $a_t$. The constraint $a_t\le K$ binds and the optimal action is just $K$. If $\Lambda_2^t>0$ for some $t\in[0,T]$ and some $\omega\in\Omega$, it means the RHS of (13) is negative such that

$$\Lambda_1V_1'(0)>-\theta E[U_2(-P)+\Lambda_1U_1(P-L)]B_t,$$

which shows the marginal cost of action is always bigger than the marginal benefit. Less action is required but the constraint $0\le a_t$ binds. The optimal action is just 0.

Recalling that the utility functions are increasing functions, we have $U_2'(-x_1)>0$ and $U_1'(x_1-x_2)>0$. Recalling that the utility functions are concave functions, we have $U_2'(-x_1)$ is an increasing function of $x_1$ and $U_1'(x_1-x_2)$ is a decreasing function of $x_1$. Hence, the function $g$ defined by

$$g(x_1,x_2):=\frac{U_2'(-x_1)}{U_1'(x_1-x_2)},\qquad x_1,x_2\in\mathbb{R}$$

is a positive, increasing function of $x_1$, meaning that $g(x_1,x_2)$ is invertible for any fixed $x_2$. The inverse function is denoted by $g^{-1}(\cdot,x_2)$. Consider the function $\mathbb{U}_1$ defined by

$$\mathbb{U}_1(\Lambda_1):=\int_0^T U_1(w_t-q_t^{\Lambda_1})dt+E\left[U_1(P^{\Lambda_1}-L)\right]\theta\left(\bar{B}-E[\int_0^T a_t^{\Lambda_1}B_tdt]\right)-E\left[\int_0^T V_1(a_t^{\Lambda_1})dt\right], \tag{14}$$

where

$$\begin{aligned} &P^{\Lambda_1}=g^{-1}(\Lambda_1,L),\\ &q_t^{\Lambda_1}=-g^{-1}(\Lambda_1,-w_t),\\ &a_t^{\Lambda_1}=V_1'^{-1}\left(-\theta E\left[\frac{1}{\Lambda_1}U_2(-g^{-1}(\Lambda_1,L))+U_1(g^{-1}(\Lambda_1,L)-L)\right]B_t\right)\\ &\qquad \text{if } 0\le-\theta E\left[\frac{1}{\Lambda_1}U_2(-g^{-1}(\Lambda_1,L))+U_1(g^{-1}(\Lambda_1,L)-L)\right]B_t\le V_1'(K),\\ &a_t^{\Lambda_1}=K \text{ if } V_1'(K)<-\theta E\left[\frac{1}{\Lambda_1}U_2(-g^{-1}(\Lambda_1,L))+U_1(g^{-1}(\Lambda_1,L)-L)\right]B_t,\\ &a_t^{\Lambda_1}=0 \text{ if } -\theta E\left[\frac{1}{\Lambda_1}U_2(-g^{-1}(\Lambda_1,L))+U_1(g^{-1}(\Lambda_1,L)-L)\right]B_t<0. \end{aligned} \tag{15}$$

The controls in (15) are the solution of equations (11) and (13). $\mathbb{U}_1(\Lambda_1)$ is the customer's expected total utility corresponding to the controls $(q^{\Lambda_1},P^{\Lambda_1},a^{\Lambda_1})$. We know that $g(x_1,x_2)$ is an increasing function of $x_1$, so the inverse function is also an increasing function. Thus, $P_i^{\Lambda_1}$ increases and $q_t^{\Lambda_1}$ decreases when $\Lambda_1$ increases. That is, the customer can get more compensation and pay less premium at the same time. The customer's utility from the contract may also increase. It inspires us to think that $\mathbb{U}_1(\Lambda_1)$ may be an increasing function of $\Lambda_1$. The obstacle is we are not sure how $a_t^{\Lambda_1}$ moves according to $\Lambda_1$. From (15), we can see $a_t^{\Lambda_1}$ is closely related to

$$u(\Lambda_1) := -\frac{1}{\Lambda_1}U_2(-g^{-1}(\Lambda_1, L)) - U_1(g^{-1}(\Lambda_1, L) - L). \tag{16}$$

Here, $\theta E[u(\Lambda_1)]B_t$ can be recognized as the marginal benefit of the action. $V_1'(a_t)$ can be recognized as the marginal cost of the action. When $\Lambda_2^t = 0$ for some $t$, (13) becomes $V_1'(a_t) = -\theta E[u(\Lambda_1)]B_t$. It illustrates that the optimal action is reached when its marginal benefit equals its marginal cost. To explore more connections between $a_t^{\Lambda_1}$ and $\Lambda_1$, we consider the derivative

$$u'(\Lambda_1) = \frac{1}{\Lambda_1^2}U_2(-g^{-1}(\Lambda_1, L)) + \frac{1}{\Lambda_1}U_2'(-g^{-1}(\Lambda_1, L))g^{-1'}(\Lambda_1, L) - U_1'(g^{-1}(\Lambda_1, L) - L)g^{-1'}(\Lambda_1, L). \tag{17}$$

From (11), we have $\frac{U_2'(-P_i)}{\Lambda_1} = U_1'(P_i - L_i)$. Here, $P_i^{\Lambda_1} = g^{-1}(\Lambda_1, L_i)$, so we obtain $\frac{1}{\Lambda_1}U_2'(-g^{-1}(\Lambda_1, L)) = U_1'(g^{-1}(\Lambda_1, L) - L)$. Now, we rewrite (17) to get

$$u'(\Lambda_1) = \frac{1}{\Lambda_1^2}U_2(-g^{-1}(\Lambda_1, L)). \tag{18}$$

**Theorem 1.** *$\mathbb{U}_1(\Lambda_1)$ is an increasing function of $\Lambda_1$ for $\Lambda_1 \in (0, \infty)$.*

**Proof.** See Appendix A. □

We define $\hat{\Lambda}_1$ by the following equation,

$$\mathbb{U}_1(\hat{\Lambda}_1) = R. \tag{19}$$

Then we have

**Theorem 2.** *If there exists $\hat{\Lambda}_1 > 0$ such that (19) holds, then the optimal insurance contract $(\hat{q}, \hat{P}, \hat{a}) = (q^{\hat{\Lambda}_1}, P^{\hat{\Lambda}_1}, a^{\hat{\Lambda}_1})$ is given by*

$$\hat{q}_t = -g^{-1}(\hat{\Lambda}_1, -w_t), \tag{20}$$

$$\hat{P}_i = g^{-1}(\hat{\Lambda}_1, L_i), \tag{21}$$

$$\hat{a}_t = \begin{cases} 0 & \text{if } -\theta E\left[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1, L)) + U_1(g^{-1}(\hat{\Lambda}_1, L) - L)\right]B_t < 0, \\ V_1'^{-1}\left(-\theta E\left[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1, L)) + U_1(g^{-1}(\hat{\Lambda}_1, L) - L)\right]B_t\right) & \\ \qquad \text{if } 0 \le -\theta E\left[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1, L)) + U_1(g^{-1}(\hat{\Lambda}_1, L) - L)\right]B_t \le V_1'(K), & \\ K & \text{if } V_1'(K) < -\theta E\left[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1, L)) + U_1(g^{-1}(\hat{\Lambda}_1, L) - L)\right]B_t. \end{cases} \tag{22}$$

**Proof.** See Appendix A. □

**Remark 1.** There is $\Lambda_1$ such that $\mathbb{U}_1(\Lambda_1) < R$ whatever $R$ is.[1] It must be smaller than $\hat{\Lambda}_1$ according to Theorem 1 if $\hat{\Lambda}_1$ exists. However, the existence of $\hat{\Lambda}_1$ depends on the value of $R$. In Theorem 4 of the next section, we will show the existence and uniqueness of $\hat{\Lambda}_1$ with an appropriate value of $R$.

**Remark 2.** The optimal action $\hat{a}_t$ is an increasing function of $B_t$. We can explain it in three ways. First, if $r_t$ is high, actions at this moment are more effective. The insured wants to take this opportunity to insert more actions. Second, the insured prefers to insert more actions earlier if we neglect the uncertainty elements $r_t$, $Y_i$, and $\tau_i$. For example, if $r_t = r_0$ for every $t \in [0, T]$ and $Y = 0$ almost surely, then

$$B_t = r_0 e^t \int_t^T e^{-(1+\delta)s}\left(Y_0 e^{\delta\tau_0}\right) ds = \frac{r_0 Y_0}{1+\delta}(e^{-\delta t} - e^{t-(1+\delta)T}),$$

which is a decreasing function of $t$. Thus, $\hat{a}_t$ is also a decreasing function of $t$. Especially, $B_t = 0$ when $t = T$, resulting in $a_T = 0$. The action taken at an earlier time is effective for a longer period. It can reduce the intensity of the accidents throughout the whole period. The insured is motivated to act as much as possible at the beginning. The action taken at maturity is only effective at the moment $T$. It makes almost no contribution to lowering the intensity. The insured does not want to waste his action, thus takes zero action at time $T$. Third, the bigger $\sum_{i=0}^{M(t)} Y_i e^{\delta\tau_i}$ is, the bigger $B_t$ is. Thus more actions should be inserted when the accumulated external exposure is more. Note that the same amount of action deducts the same proportion of the intensity of claims. When the exposure is high, the same amount of action can remove more intensity. The actions are therefore more valuable and the insured will choose to execute more actions at these moments.

[1] We showed $\lim_{\Lambda_1 \to 0^+} \mathbb{U}_1(\Lambda_1) = -\infty$ when we discussed Lagrangian multipliers in (13).

## 5. The reservation utility

In this section, we calculate the reservation utility $R$ of (5), which is the utility of the potential insured if he does not purchase insurance. The participation constraint (5) means that the expected total utility from purchasing insurance is greater than or equal to the expected total utility from not purchasing insurance. In this section, we will (i) calculate the reservation utility $R$ when the potential insured does not purchase insurance, (ii) compare the actions taken when the potential insured does and does not enter the insurance market, and (iii) show that $\hat{\Lambda}_1$ of Theorem 2 exists uniquely.

If the potential insured does not enter the insurance market, then he will not pay a premium and, as a consequence, will not receive any compensation. However, he will select the action to maximize his expected total utility.

We denote by $\mathcal{A}_R$ the class of stochastic processes $a : [0, T] \times \Omega \mapsto \mathbb{R}$ that are adapted to the filtration $\mathbb{F}$.

**Problem 2.** If the potential insured does not purchase insurance, he wants to obtain the control $a^* \in \mathcal{A}_R$ that solves the problem

$$\max_{a \in \mathcal{A}_R} E\left[\int_0^T U_1(w_t)dt + \sum_{i=1}^{N^a(T)} U_1(-L_i) - \int_0^T V_1(a_t)dt\right]$$
$$s.t.\ \ 0 \le a_t \le K,\ \text{for all } t \in [0, T].$$

According to (9), $E\left[\sum_{i=1}^{N^a(T)} U_1(-L_i)\right]$ can be rewritten as

$$E\left[\sum_{i=1}^{N^a(T)} U_1(-L_i)\right] = E[U_1(-L)]\left(\theta\bar{B} - \theta E\left[\int_0^T a_t B_t dt\right]\right). \tag{23}$$

We define the Lagrangian function

$$\mathcal{L}_2(a; \Lambda_3) := \int_0^T U_1(w_t)dt + E[U_1(-L)]\left(\theta\bar{B} - \theta E\left[\int_0^T a_t B_t dt\right]\right) - E\left[\int_0^T V_1(a_t)dt\right] + E\left[\int_0^T \Lambda_3^t a_t dt\right],$$

where $\Lambda_3^t$, $t \in [0, T]$, adapted to $\mathbb{F}$, are Lagrangian multipliers. We take the differentiation of the Lagrangian function with respect to $a_t$ and obtain the first order conditions

$$V_1'(a_t) - \Lambda_3^t = -\theta B_t E[U_1(-L)] \tag{24}$$

for $t \in [0, T]$ and $\omega \in \Omega$. $U_1(-L) < 0$ for $L \in R_L$ from (3), then $-\theta B_t E[U_1(-L)] \ge 0$ for each $t \in [0, T]$ and $\omega \in \Omega$. If $0 \le -\theta B_t E[U_1(-L)] \le V_1'(K)$ for some $t \in [0, T]$ and $\omega \in \Omega$, $\Lambda_3^t(\omega) = 0$. The solution of (24) for $a_t$ satisfies the constraint, so the constraint does not bind. If $-\theta B_t E[U_1(-L)] \ge V_1'(K)$ for some $t \in [0, T]$ and $\omega \in \Omega$, $\Lambda_3^t(\omega) < 0$. In this case, the marginal benefit of the action is always bigger than its marginal cost. However, the constraint $a_t \le K$ binds, so the optimal action is just $K$.

**Proposition 2.** *The optimal control of Problem 2 is given by*

$$a_t^* = \begin{cases} V_1'^{-1}(-\theta B_t E[U_1(-L)]) & \text{if } V_1'(K) \ge -\theta B_t E[U_1(-L)] \\ K & \text{if } V_1'(K) < -\theta B_t E[U_1(-L)]. \end{cases} \tag{25}$$

**Proof.** See Appendix A. □

We recall $a^{\Lambda_1}$ defined in (15). Comparing the two action processes $a^{\Lambda_1}$ and $a^*$, we have the following relation.

**Theorem 3.** *For every $t \in [0, T]$:*

$$V_1'(a_t^{\Lambda_1}) \le V_1'(a_t^*) - \frac{1}{\Lambda_1} U_2(0) B_t \theta.$$

**Proof.** See Appendix A. □

We observe that if $U_2(0) = 0$, then $V_1'(a_t^{\Lambda_1}) \le V_1'(a_t^*)$ as a consequence of Theorem 3. Since $V_1'(\cdot)$ is an increasing function, we have the following relation between the two action processes.

**Corollary 1.** *If $U_2(0) = 0$, then for every $t \in [0, T]$:*

$$a_t^{\Lambda_1} \le a_t^*.$$

Theorem 3 shows that $a^{\Lambda_1}$ is constrained by $a^*$. This constraint is more evident when $U_2(0)=0$.

Taking $a^*$ into the objective function of Problem 2, we obtain the reservation utility

$$R=\int_0^T U_1(w_t)dt+\theta E[U_1(-L)]\left(\bar{B}-E\left[\int_0^T a_t^* B_t dt\right]\right)-E\left[\int_0^T V_1(a_t^*)dt\right]. \tag{26}$$

We define $\underline{\Lambda_1}$ by the equation

$$E[U_1(g^{-1}(\underline{\Lambda_1},L)-L)]=E[U_1(-L)]. \tag{27}$$

**Lemma 2.** *$\underline{\Lambda_1}$ exists uniquely. Furthermore, $\mathbb{U}_1(\underline{\Lambda_1})<R$, where $R$ is the reservation utility defined by (26).*

**Proof.** See Appendix A. □

**Theorem 4.** *There exists a unique $\hat{\Lambda}_1$ such that (19) holds and $\hat{\Lambda}_1\in(\underline{\Lambda_1},\infty)$.*

**Proof.** See Appendix A. □

Thus, Theorem 4 completes the solution of Problem 1.

We define the highest income rate by $w_{sup}:=\sup\{w_t:t\in[0,T]\}$. We also define $\bar{\Lambda}_1:=\frac{U_2'(0)}{U_1'(w_{sup})}$. Then we have the following constraint for $\hat{\Lambda}_1$.

**Corollary 2.** *If $U_2(0)=0$, then there exists a unique $\hat{\Lambda}_1$ such that (19) holds and $\hat{\Lambda}_1\in(\underline{\Lambda_1},\bar{\Lambda}_1)$.*

**Proof.** See Appendix A. □

## 6. The exponential utility and the quadratic cost

In this section, we apply the theory developed in Sections 4 and 5 to the case

$$U_1(x)=-e^{-\gamma_1 x},\ \ U_2(x)=-e^{-\gamma_2 x},\ \ V_1(x)=mx^2,\ \ K=1,\ \ w_t=0,$$

where $\gamma_1>\gamma_2>0$ and $m>0$ are constant parameters. Then, $g$ is given by

$$g(x_1,x_2)=\frac{U_2'(-x_1)}{U_1'(x_1-x_2)}=\frac{\gamma_2 e^{\gamma_2 x_1}}{\gamma_1 e^{-\gamma_1(x_1-x_2)}}.$$

For a fixed $x_2$, the inverse function $g^{-1}(\cdot,x_2)$ is given by

$$g^{-1}(y,x_2)=\frac{\ln(y)+\ln(\frac{\gamma_1}{\gamma_2})+\gamma_1 x_2}{\gamma_1+\gamma_2}.$$

From (20) and (21), we obtain

$$\hat{q}_t=-g^{-1}(\hat{\Lambda}_1,-w_t)=-g^{-1}(\hat{\Lambda}_1,0)=-\frac{\ln(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})}{\gamma_1+\gamma_2}\text{ for }t\in[0,T]; \tag{28}$$

$$\hat{P}_i=g^{-1}(\hat{\Lambda}_1,L_i)=\frac{\gamma_1 L_i+\ln(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})}{\gamma_1+\gamma_2}\text{ for }i=1,2,3,\cdots. \tag{29}$$

We have

$$\begin{aligned}
&-\theta E[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1,L))+U_1(g^{-1}(\hat{\Lambda}_1,L)-L)]B_t\\
&=\theta E\left[\frac{1}{\hat{\Lambda}_1}e^{\frac{\gamma_2}{\gamma_1+\gamma_2}(\gamma_1 L+\ln(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2}))}+e^{-\frac{\gamma_1}{\gamma_1+\gamma_2}(\ln(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})-\gamma_2 L)}\right]B_t\\
&=\theta B_t(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{\gamma_1}{\gamma_1+\gamma_2}}E[e^{\frac{\gamma_1\gamma_2 L}{\gamma_1+\gamma_2}}](1+\frac{\gamma_1}{\gamma_2}),
\end{aligned}$$

which is positive for every $t\in[0,T]$. In this example, $V_1'(x)=2mx$, so $V_1'(K)=V_1'(1)=2m$ and $V_1'^{-1}(y)=\frac{y}{2m}$. Hence,

$$\hat{a}_t=\begin{cases}\frac{\theta}{2m}B_t(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{\gamma_1}{\gamma_1+\gamma_2}}E[e^{\frac{\gamma_1\gamma_2 L}{\gamma_1+\gamma_2}}](1+\frac{\gamma_1}{\gamma_2}) & \text{if }\theta B_t(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{\gamma_1}{\gamma_1+\gamma_2}}E[e^{\frac{\gamma_1\gamma_2 L}{\gamma_1+\gamma_2}}](1+\frac{\gamma_1}{\gamma_2})\le 2m\\ 1 & \text{if }\theta B_t(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{\gamma_1}{\gamma_1+\gamma_2}}E[e^{\frac{\gamma_1\gamma_2 L}{\gamma_1+\gamma_2}}](1+\frac{\gamma_1}{\gamma_2})>2m.\end{cases} \tag{30}$$

Since $-\theta B_t E[U_1(-L)] = \theta B_t E\left[e^{\gamma_1 L}\right]$, applying (25), we obtain

$$a_t^* = \begin{cases} \frac{\theta}{2m} B_t E\left[e^{\gamma_1 L}\right] & \text{if } \theta B_t E\left[e^{\gamma_1 L}\right] \le 2m \\ 1 & \text{if } \theta B_t E\left[e^{\gamma_1 L}\right] > 2m. \end{cases}$$

$\hat{\Lambda}_1$ in (28)-(30) is the solution of $\mathbb{U}_1(\hat{\Lambda}_1) = R$. We denote $C_t := \theta B_t (\frac{\hat{\Lambda}_1 \gamma_1}{\gamma_2})^{-\frac{\gamma_1}{\gamma_1+\gamma_2}} E[e^{\frac{\gamma_1\gamma_2 L}{\gamma_1+\gamma_2}}](1+\frac{\gamma_1}{\gamma_2})$. Recalling (14), we have

$$\begin{aligned}
\mathbb{U}_1(\hat{\Lambda}_1) = & \int_0^T U_1\left(0 + \frac{\ln(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})}{\gamma_1+\gamma_2}\right)dt + E\left[U_1\left(\frac{\gamma_1 L_i + \ln(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})}{\gamma_1+\gamma_2} - L\right)\right]\theta\left(\bar{B} - E[\int_0^T \hat{a}_t B_t dt]\right) \\
& - E\left[\int_0^T m\hat{a}_t^2 dt\right] \\
= & -(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{\gamma_1}{\gamma_1+\gamma_2}} T - E\left[e^{\frac{\gamma_1\gamma_2 L}{\gamma_1+\gamma_2}}\right](\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{\gamma_1}{\gamma_1+\gamma_2}}\theta \\
& \left(\bar{B} - E\left[\int_0^T \left(\frac{\theta}{2m} B_t^2 (\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{\gamma_1}{\gamma_1+\gamma_2}} E[e^{\frac{\gamma_1\gamma_2 L}{\gamma_1+\gamma_2}}](1+\frac{\gamma_1}{\gamma_2})\mathbb{I}_{\{C_t \le 2m\}} + B_t \mathbb{I}_{\{C_t > 2m\}}\right)dt\right]\right) \\
& - E\left[\int_0^T m\left(\frac{\theta^2}{4m^2} B_t^2 (\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{2\gamma_1}{\gamma_1+\gamma_2}}\left(E[e^{\frac{\gamma_1\gamma_2 L}{\gamma_1+\gamma_2}}]\right)^2 (1+\frac{\gamma_1}{\gamma_2})^2 \mathbb{I}_{\{C_t \le 2m\}} + \mathbb{I}_{\{C_t > 2m\}}\right)dt\right] \\
= & -(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{\gamma_1}{\gamma_1+\gamma_2}} T - E\left[e^{\frac{\gamma_1\gamma_2 L}{\gamma_1+\gamma_2}}\right](\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{\gamma_1}{\gamma_1+\gamma_2}}\theta\left(\bar{B} - E\left[\int_0^T B_t \mathbb{I}_{\{C_t > 2m\}} dt\right]\right) \\
& + \frac{\theta^2}{2m}\left(E[e^{\frac{\gamma_1\gamma_2 L}{\gamma_1+\gamma_2}}]\right)^2 \left(1 + \frac{\gamma_1}{\gamma_2} - \frac{1}{2}(1+\frac{\gamma_1}{\gamma_2})^2\right) E\left[\int_0^T B_t^2 \mathbb{I}_{\{C_t \le 2m\}} dt\right](\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{2\gamma_1}{\gamma_1+\gamma_2}} \\
& - mE\left[\int_0^T \mathbb{I}_{\{C_t > 2m\}} dt\right] \\
= & \frac{\theta^2}{4m}\left(E[e^{\frac{\gamma_1\gamma_2 L}{\gamma_1+\gamma_2}}]\right)^2 (1 - \frac{\gamma_1^2}{\gamma_2^2}) E\left[\int_0^T B_t^2 \mathbb{I}_{\{C_t \le 2m\}} dt\right](\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{2\gamma_1}{\gamma_1+\gamma_2}} \\
& - \left(T + \theta E[e^{\frac{\gamma_1\gamma_2 L}{\gamma_1+\gamma_2}}]\left(\bar{B} - E\left[\int_0^T B_t \mathbb{I}_{\{C_t > 2m\}} dt\right]\right)\right)(\frac{\hat{\Lambda}_1\gamma_1}{\gamma_2})^{-\frac{\gamma_1}{\gamma_1+\gamma_2}} - mE\left[\int_0^T \mathbb{I}_{\{C_t > 2m\}} dt\right].
\end{aligned}$$

According to (26), we have

$$\begin{aligned}
R = & \int_0^T U_1(0)dt + \theta E[U_1(-L)]\left(\bar{B} - E\left[\int_0^T a_t^* B_t dt\right]\right) - E\left[\int_0^T m(a_t^*)^2 dt\right] \\
= & -(T + \theta \bar{B} E[e^{\gamma_1 L}]) + \theta E[e^{\gamma_1 L}] E\left[\int_0^T \left(\frac{\theta}{2m} B_t E[e^{\gamma_1 L}] \mathbb{I}_{\{\theta B_t E[e^{\gamma_1 L}] \le 2m\}} + \mathbb{I}_{\{\theta B_t E[e^{\gamma_1 L}] > 2m\}}\right) B_t dt\right] \\
& - mE\left[\int_0^T \left(\frac{\theta^2}{4m^2}(E[e^{\gamma_1 L}])^2 B_t^2 \mathbb{I}_{\{\theta B_t E[e^{\gamma_1 L}] \le 2m\}} + \mathbb{I}_{\{\theta B_t E[e^{\gamma_1 L}] > 2m\}}\right) dt\right] \\
= & \frac{\theta^2}{4m}(E[e^{\gamma_1 L}])^2 E\left[\int_0^T B_t^2 \mathbb{I}_{\{\theta B_t E[e^{\gamma_1 L}] \le 2m\}} dt\right] \\
& + E\left[\int_0^T (\theta B_t E[e^{\gamma_1 L}] - m)\mathbb{I}_{\{\theta B_t E[e^{\gamma_1 L}] > 2m\}} dt\right] - (T + \theta \bar{B} E[e^{\gamma_1 L}]).
\end{aligned}$$

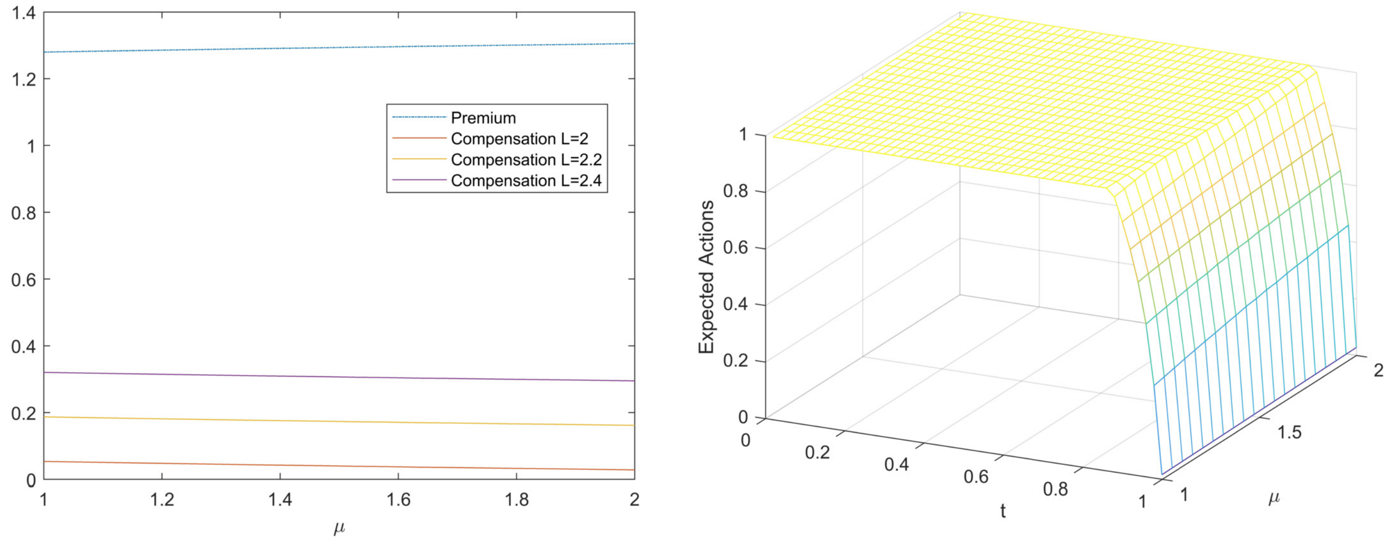


**Fig. 1.** The other parameters are $\theta = 1$, $\rho = 1$, $\gamma_1 = 2$, and $\gamma_2 = 1$. Furthermore, $P\{L = 2.0\} = 0.5$, $P\{L = 2.2\} = 0.3$, $P\{L = 2.4\} = 0.2$.

Even though the equation $\mathbb{U}_1(\hat{\Lambda}_1) = R$ looks complicated, the monotonicity of $\mathbb{U}_1(\Lambda_1)$ and the uniqueness of $\hat{\Lambda}_1$ allow us to use the bisection method to find $\hat{\Lambda}_1$ in the following numerical analysis.

Equation (29) shows that $\hat{P}_i \neq L_i$, so full compensation is not optimal.

**Example 1.** To consider a numerical example, assume that the magnitude $Y$ of the external risky events has exponential distribution, and the intensity $\rho$ is constant: $\rho(t) \equiv \rho \in [0, \infty)$.

We will investigate how the solution depends on the parameters $\theta$, $\rho$, $E[L]$, $\mu$, $\gamma_1$, $\gamma_2$, and the variance of $I(t)$ for $t \in [0, T]$. We fix the other parameters as $T = 1$, $Y_0 = 1$, $m = 5$, $\delta = 1$, $r_t = 1$, and $A_0 = 0$.

The benchmark parameter values are $\theta = 1$, $\rho = 1$, $\mu = 1$, $\gamma_1 = 2$, $\gamma_2 = 1$, and $L$ has probability distribution $P\{L = 2\} = 0.5$, $P\{L = 2.2\} = 0.3$, $P\{L = 2.4\} = 0.2$. Then, $\hat{\Lambda}_1 = 0.0108$ and the optimal insurance contract (for these parameter values) is given by

$$\hat{q}_t = 1.2794;$$

$$\hat{P}_i = \begin{cases} 0.0539 & \text{if } L_i = 2.0 \\ 0.1873 & \text{if } L_i = 2.2 \\ 0.3206 & \text{if } L_i = 2.4; \end{cases}$$

$$\hat{a}_t = \min\{16.2321 B_t, 1\}.$$

Since $\hat{a}$ is a stochastic process, we will consider $E[\hat{a}_t]$. Figs. 1 to 4 show that $\hat{q}_t$, $\hat{P}_i$, and $E[\hat{a}_t]$ increase when the parameters $\mu$, $\theta$, $\rho$, and $E[L]$ increase. These four parameters reflect the risk in different aspects. Thus, when the risk increases, the insurer requires a higher premium, pays less compensation, and requires the insured to increase his expected action.

Figs. 1 to 4 also show that the expected insured's action decreases when time passes, and that the insured is required to take no action when maturity approaches. This is consistent with Remark 2 of Theorem 2.

Fig. 5 shows that when the insured's risk aversion $\gamma_1$ increases, the premium increases, the compensation decreases, and the expected action increases.

Fig. 6 shows how the solution depends on the insurer's risk aversion $\gamma_2$. We recall that the insured's reservation utility presented in Section 5 is not affected by the insurer's risk aversion $\gamma_2$. Fig. 6 shows that the premium and the compensation decrease when the insurer's risk aversion increases. This makes sense because, as the risk aversion increases, the insurer avoids risk by paying less compensation in exchange for receiving less premium.

We have also studied the situation in which the mean remains the same but the variance changes. Fig. 7 shows that the variance does not affect much the optimal premium $q$ or compensation $P$ when the mean is fixed. However, the optimal expected action $E[\hat{a}]$ decreases when the variance of $I(t)$ increases. Since it is impossible to list the variances of $I(t)$ for all $t \in [0, T]$ in the figure, we use the variance of $I(T)$ as a representation.

## 7. Conclusions

We have studied the optimal insurance contract that an insurer should propose to a potential insured. Motivated by climate change and catastrophic events, we have assumed that the number of claims process is a shot-noise Cox process. However, this model for the number of claims can be applied to many other risk management problems. This is the first paper on optimal insurance contracts that allows the number of claims to be a shot-noise Cox process. It is also a model in which persistent actions affect a Cox process.

To the best of our knowledge, we have obtained the first analytical solution for the optimal premium, the optimal compensation, and the optimal actions of the insured when the number of claims process is a Cox process. The solution shows that the optimal expected action decreases over time. It also shows that the amount of action decided by the insurer is restricted by the amount of action the potential insured selects when he is not in the insurance market.

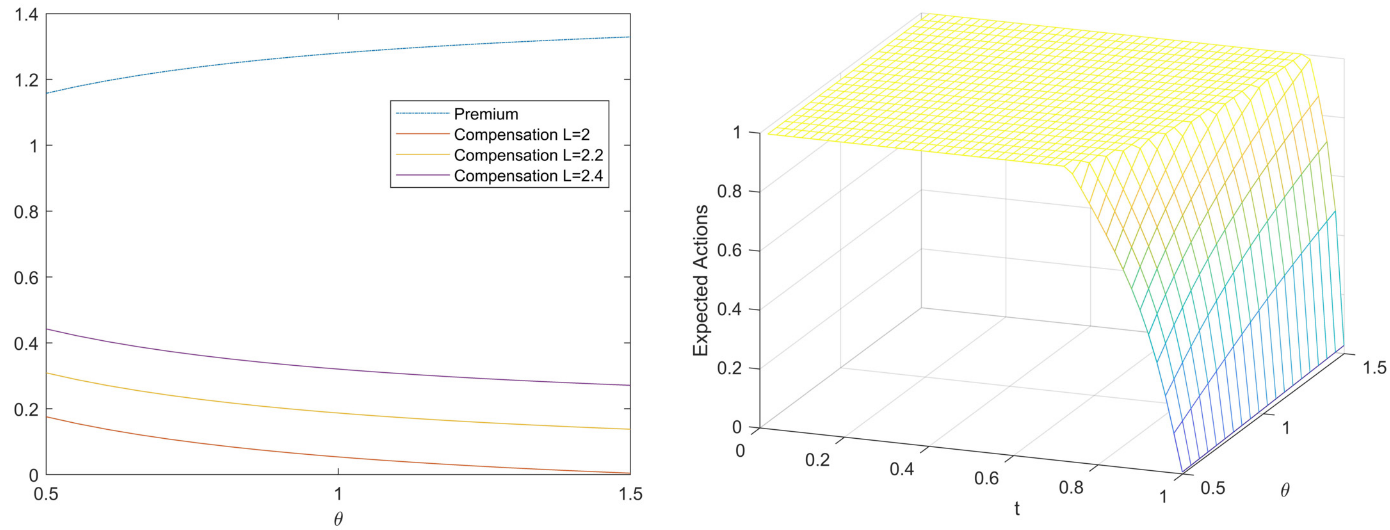


**Fig. 2.** The other parameters are $\rho = 1$, $\mu = 1$, $\gamma_1 = 2$, and $\gamma_2 = 1$. Furthermore, $P\{L = 2.0\} = 0.5$, $P\{L = 2.2\} = 0.3$, $P\{L = 2.4\} = 0.2$.

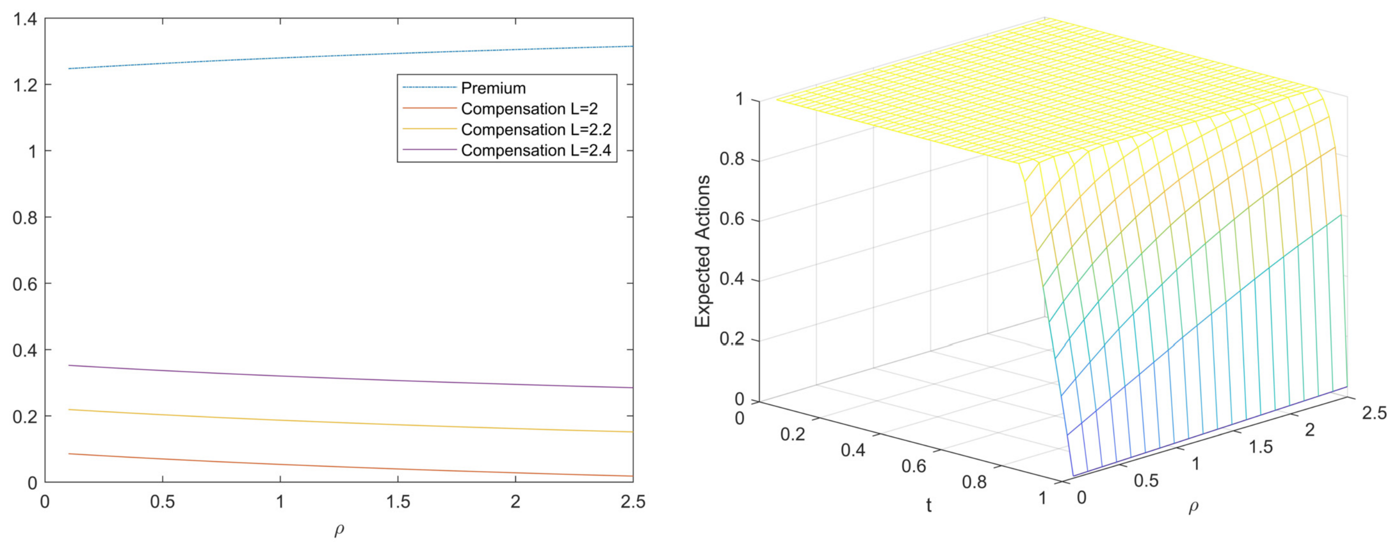


**Fig. 3.** The other parameters are $\theta = 1$, $\mu = 1$, $\gamma_1 = 2$, and $\gamma_2 = 1$. Furthermore, $P\{L = 2.0\} = 0.5$, $P\{L = 2.2\} = 0.3$, $P\{L = 2.4\} = 0.2$.

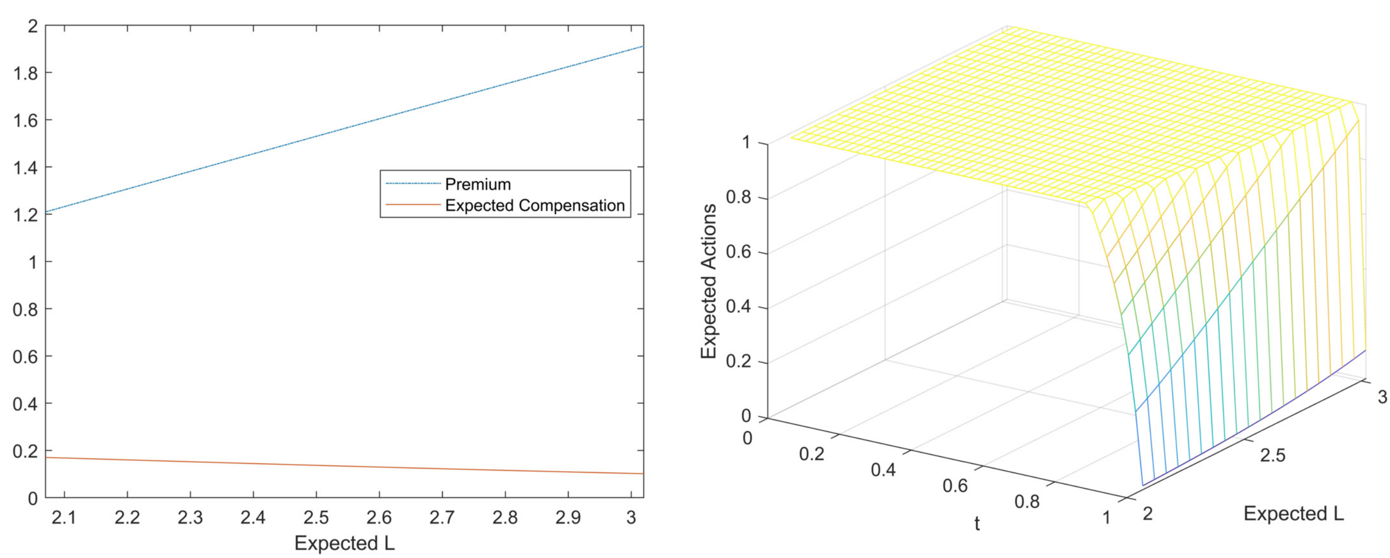


**Fig. 4.** The other parameters are $\theta = 1$, $\rho = 1$, $\mu = 1$, $\gamma_1 = 2$, and $\gamma_2 = 1$.

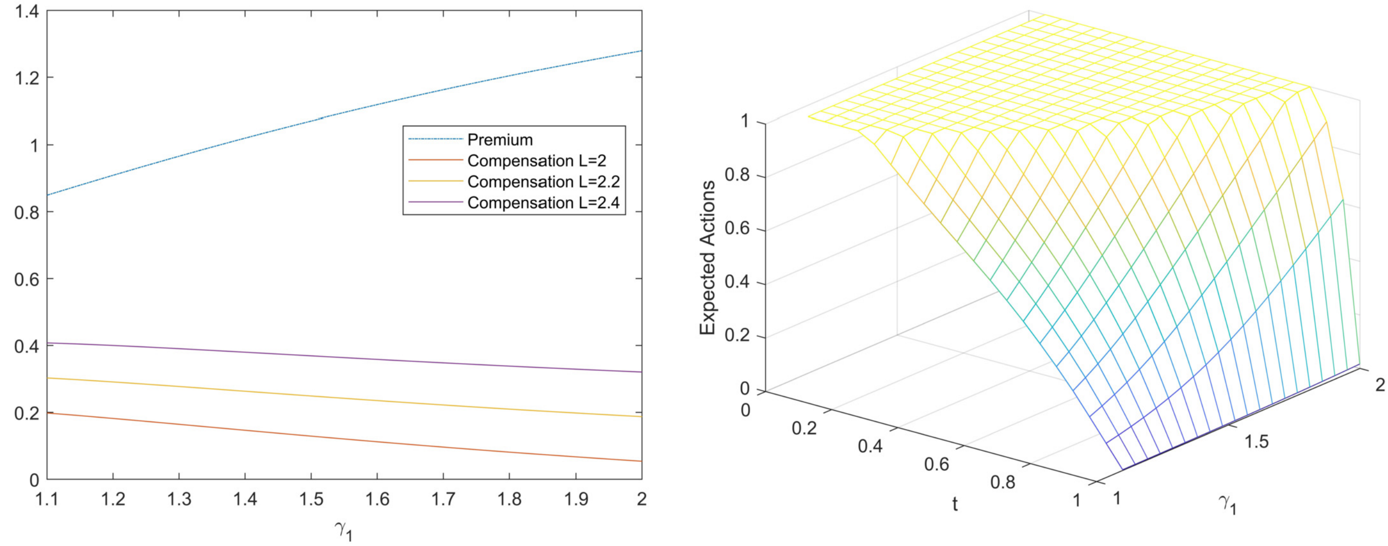


**Fig. 5.** The other parameters are $\theta = 1$, $\mu = 1$, $\rho = 1$, and $\gamma_2 = 1$. Furthermore, $P\{L = 2.0\} = 0.5$, $P\{L = 2.2\} = 0.3$, $P\{L = 2.4\} = 0.2$.

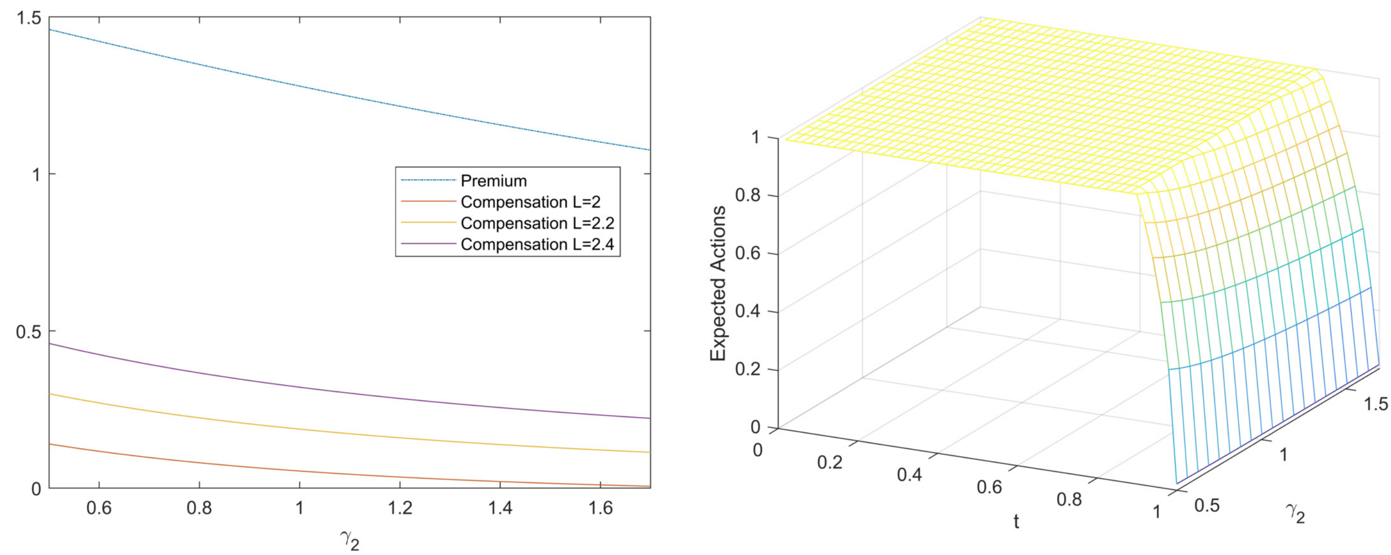


**Fig. 6.** The other parameters are $\theta = 1$, $\mu = 1$, $\rho = 1$, and $\gamma_1 = 2$. Furthermore, $P\{L = 2.0\} = 0.5$, $P\{L = 2.2\} = 0.3$, $P\{L = 2.4\} = 0.2$.

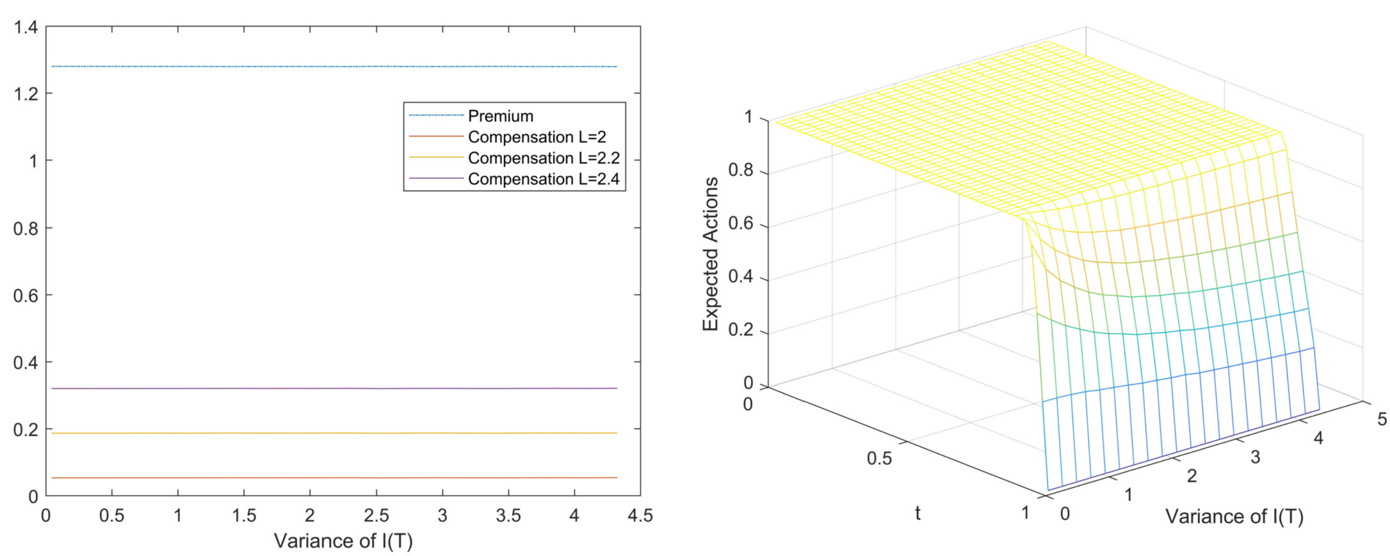


**Fig. 7.** The other parameters are $\theta = 1$, $\gamma_1 = 2$, and $\gamma_2 = 1$. Furthermore, $P\{L = 2.0\} = 0.5$, $P\{L = 2.2\} = 0.3$, $P\{L = 2.4\} = 0.2$.

An example with exponential utilities allows us to see how the solution depends on the parameters of the model.

### Declaration of competing interest

None.

### Data availability

No data was used for the research described in the article.

### Acknowledgements

The research of A. Cadenillas and W. Liu was funded by the Social Sciences and Humanities Research Council of Canada (SSHRC) grant 435-2017-0511. The authors would like to thank Chenzhe Diao for his help in coding. Preliminary versions of this paper have been presented at the 2022 CORS/INFORMS International Conference, Vancouver, June 5–8, 2022, and at the Advances in Stochastic Control and Optimal Stopping with Applications in Finance and Economics Conference, *Centre International de Rencontres Mathématiques*, Marseille, September 12-16, 2022. We are grateful for comments from participants at these conferences. The authors thank the referees for their constructive remarks.

## Appendix A. Proofs

### *A.1. Proof of Lemma 1*

**Proof.** We denote

$$f(I,t) := \frac{1}{\theta} I e^{\delta t} - \mu \int_0^t e^{\delta u} \rho(u) du.$$

It is obvious that $f(I,t)$ is differentiable with respect to each $I$ and $t$.

$$\begin{aligned}
&\Big| \int_0^\infty f(I+\theta y, t) dF_Y(y) - f(I,t) \Big| \\
&= \Big| \int_0^\infty \left( \frac{1}{\theta}(I+\theta y) e^{\delta t} - \mu \int_0^t e^{\delta u} \rho(u) du \right) dF_Y(y) - \left( \frac{1}{\theta} I e^{\delta t} - \mu \int_0^t e^{\delta u} \rho(u) du \right) \Big| \\
&= \left| e^{\delta t} \mu \right| < \infty
\end{aligned}$$

Applying (7), we obtain

$$\begin{aligned}
\mathbb{A} f(I,t) = \frac{1}{\theta} I \delta e^{\delta t} - \mu e^{\delta t} \rho(t) - \frac{1}{\theta} I \delta e^{\delta t} + \rho(t) \int_0^\infty \left( \frac{1}{\theta}(I+\theta y) e^{\delta t} - \mu \int_0^t e^{\delta u} \rho(u) du \right) dF_Y(y) \\
- \rho(t) \left( \frac{1}{\theta} I e^{\delta t} - \mu \int_0^t e^{\delta u} \rho(u) du \right) \\
= -\mu e^{\delta t} \rho(t) + \rho(t) \frac{1}{\theta} I e^{\delta t} + \rho(t) \mu e^{\delta t} - \rho(t) \mu \int_0^t e^{\delta u} \rho(u) du \\
- \rho(t) \frac{1}{\theta} I e^{\delta t} + \rho(t) \mu \int_0^t e^{\delta u} \rho(u) du \\
= 0.
\end{aligned}$$

According to Proposition 1 in Dassios and Embrechts (1989), we obtain that the stochastic process defined by

$$f(I(t),t) = \frac{1}{\theta} I(t) e^{\delta t} - \mu \int_0^t e^{\delta u} \rho(u) du$$

is a martingale. From (1), we can get the required statement. □

*A.2. Proof of Proposition 1*

**Proof.** $N^a = \{N^a(t); t \geq 0\}$ is a Cox process with intensity process $\lambda(\cdot)$. From Lemma 3a of Grandell (1976) or Theorem 2.7 of Dassios and Jang (2003), we have

$$E[N^a(T)] = \int_0^T E[\lambda(t)]dt. \tag{31}$$

According to equation (2),

$$E[\lambda(t)] = \theta\left((1-e^{-t}A_0)E\Big[\sum_{i=0}^{M(t)} Y_i e^{\delta(\tau_i - t)}\Big] - e^{-(1+\delta)t}E\Big[\Big(\sum_{i=0}^{M(t)} Y_i e^{\delta\tau_i}\Big)\int_0^t a_s r_s e^s ds\Big]\right).$$

According to Lemma 1,

$$E\Big[\sum_{i=0}^{M(t)} Y_i e^{\delta(\tau_i - t)}\Big] = e^{-\delta t}E\Big[\sum_{i=0}^{M(t)} Y_i e^{\delta\tau_i}\Big] = e^{-\delta t}\left(Y_0 + \mu\int_0^t \rho(u)e^{\delta u}du\right)$$

and

$$\begin{aligned}
E\Big[\Big(\sum_{i=0}^{M(t)} Y_i e^{\delta\tau_i}\Big)\int_0^t a_s r_s e^s ds\Big] &= \int_0^t E\left[a_s r_s e^s\Big(\sum_{i=0}^{M(t)} Y_i e^{\delta\tau_i}\Big)\right]ds \\
&= E\left[\int_0^t a_s r_s e^s E\Big[\sum_{i=0}^{M(t)} Y_i e^{\delta\tau_i}|\mathcal{F}_s\Big]ds\right] \\
&= E\left[\int_0^t a_s r_s e^s\left(\mu\int_s^t \rho(u)e^{\delta u}du + \sum_{i=0}^{M(s)} Y_i e^{\delta\tau_i}\right)ds\right].
\end{aligned}$$

Therefore,

$$\begin{aligned}
E[\lambda(t)] &= \theta(1-e^{-t}A_0)e^{-\delta t}\left(Y_0 + \mu\int_0^t \rho(u)e^{\delta u}du\right) \\
&\quad -\theta e^{-(1+\delta)t}E\left[\int_0^t a_s r_s e^s\left(\mu\int_s^t \rho(u)e^{\delta u}du + \sum_{i=0}^{M(s)} Y_i e^{\delta\tau_i}\right)ds\right].
\end{aligned} \tag{32}$$

We replace $E[\lambda(t)]$ in (31) by (32) to obtain (8). □

*A.3. Proof of Theorem 1*

**Proof.** We split $\bar{B}$ as $\bar{B} = E\left[\int_0^T KB_t dt\right] + \bar{B} - E\left[\int_0^T KB_t dt\right]$. Then, we rewrite (14) to obtain

$$\mathbb{U}_1(\Lambda_1) = \int_0^T U_1(w_t - q_t^{\Lambda_1})dt \tag{33}$$

$$+E\left[U_1(P^{\Lambda_1} - L)\right]\theta\left(\bar{B} - E\left[\int_0^T KB_t dt\right]\right) \tag{34}$$

$$+E\left[U_1(P^{\Lambda_1} - L)\right]\theta\left(E\left[\int_0^T KB_t dt\right] - E\left[\int_0^T a_t^{\Lambda_1}B_t dt\right]\right) - E\left[\int_0^T V_1(a_t^{\Lambda_1})dt\right] \tag{35}$$

From (15), $q_t^{\Lambda_1}$ is a decreasing function of $\Lambda_1$ for every $t \in [0, T]$. Thus, the term (33) is an increasing function of $\Lambda_1$. Recalling from (9) that $\bar{B} - E\left[\int_0^T a_t B_t dt\right] \geq 0$, we obtain $\bar{B} - E\left[\int_0^T KB_t dt\right] \geq 0$. Also recalling that $P^{\Lambda_1}$ is an increasing function of $\Lambda_1$ for every $L \in R_L$, we see that the term (34) is an increasing function of $\Lambda_1$. Next, we will analyze the remaining terms in (35). For each $\omega \in \Omega$ and each $t \in [0, T]$, consider

$$\varphi(\Lambda_1) := \theta E\left[U_1(P^{\Lambda_1} - L)\right]\left(K - a_t^{\Lambda_1}\right) B_t - V_1(a_t^{\Lambda_1}).$$

We will show $\varphi(\Lambda_1)$ is an increasing function of $\Lambda_1$. $a_t^{\Lambda_1}$ takes different values for different $\Lambda_1$, so we will discuss the following three cases.

(i) If $\Lambda_1$ is such that $a_t^{\Lambda_1} = 0$, then we have

$$\varphi(\Lambda_1) = \theta E\left[U_1(P^{\Lambda_1} - L)\right] K B_t - V_1(0) = E\left[U_1(P^{\Lambda_1} - L)\right] K B_t. \tag{36}$$

Recalling $K \geq 0$, $B_t \geq 0$, and $P^{\Lambda_1}$ is an increasing function of $\Lambda_1$ for every $L \in R_L$, we get (36) is an increasing function of $\Lambda_1$.

(ii) If $\Lambda_1$ is such that $a_t^{\Lambda_1} = K$, then we have

$$\varphi(\Lambda_1) = \theta E\left[U_1(P^{\Lambda_1} - L)\right](K - K) B_t - V_1(K) = -V_1(K). \tag{37}$$

(37) is constant.

(iii) If $\Lambda_1$ is such that $a_t^{\Lambda_1} = V_1'^{-1}\left(-\theta E[\frac{1}{\Lambda_1}U_2(-g^{-1}(\Lambda_1, L)) + U_1(g^{-1}(\Lambda_1, L) - L)]B_t\right)$, then we have

$$\varphi'(\Lambda_1) = \theta E\left[U_1'(P^{\Lambda_1} - L)\frac{\partial P^{\Lambda_1}}{\partial \Lambda_1}\right](K - a_t^{\Lambda_1})B_t + \theta E\left[U_1(P^{\Lambda_1} - L)\right] B_t(-\frac{\partial a_t^{\Lambda_1}}{\partial \Lambda_1}) - V_1'(a_t^{\Lambda_1})\frac{\partial a_t^{\Lambda_1}}{\partial \Lambda_1}.$$

Here, $P^{\Lambda_1} = g^{-1}(\Lambda_1, L)$ and $V_1'(a_t^{\Lambda_1}) = -\theta E[\frac{1}{\Lambda_1}U_2(-g^{-1}(\Lambda_1, L)) + U_1(g^{-1}(\Lambda_1, L) - L)]B_t$. Now we have

$$\begin{aligned}
\varphi'(\Lambda_1) &= \theta E\left[U_1'\left(g^{-1}(\Lambda_1, L) - L\right)\frac{\partial g^{-1}(\Lambda_1, L)}{\partial \Lambda_1}\right](K - a_t^{\Lambda_1})B_t + \theta E\left[U_1\left(g^{-1}(\Lambda_1, L) - L\right)\right] B_t(-\frac{\partial a_t^{\Lambda_1}}{\partial \Lambda_1}) \\
&\quad + \theta E[\frac{1}{\Lambda_1}U_2(-g^{-1}(\Lambda_1, L)) + U_1(g^{-1}(\Lambda_1, L) - L)]B_t\frac{\partial a_t^{\Lambda_1}}{\partial \Lambda_1} \\
&= \theta E\left[U_1'\left(g^{-1}(\Lambda_1, L) - L\right)\frac{\partial g^{-1}(\Lambda_1, L)}{\partial \Lambda_1}\right](K - a_t^{\Lambda_1})B_t + \theta E[\frac{1}{\Lambda_1}U_2(-g^{-1}(\Lambda_1, L))]B_t\frac{\partial a_t^{\Lambda_1}}{\partial \Lambda_1}.
\end{aligned} \tag{38}$$

Recalling the definition of $u(\Lambda_1)$ in (16), we can see $a_t^{\Lambda_1} = V_1'^{-1}(\theta E[u(\Lambda_1)]B_t)$. From (18), we obtain

$$\begin{aligned}
\frac{\partial a_t^{\Lambda_1}}{\partial \Lambda_1} &= V_1'^{-1'}(\theta E[u(\Lambda_1)]B_t)\theta B_t E[u'(\Lambda_1)] \\
&= V_1'^{-1'}(\theta E[u(\Lambda_1)]B_t)\theta B_t E[\frac{1}{\Lambda_1^2}U_2(-g^{-1}(\Lambda_1, L))].
\end{aligned}$$

We rewrite (38) to get

$$\varphi'(\Lambda_1) = \theta E\left[U_1'\left(g^{-1}(\Lambda_1, L) - L\right)\frac{\partial g^{-1}(\Lambda_1, L)}{\partial \Lambda_1}\right](K - a_t^{\Lambda_1})B_t \tag{39}$$

$$+\theta^2\frac{1}{\Lambda_1^3}\left(E[U_2(-g^{-1}(\Lambda_1, L))]\right)^2 B_t^2 V_1'^{-1'}(\theta E[u(\Lambda_1)]B_t). \tag{40}$$

$U_1$ is an increasing function and $g^{-1}(\Lambda_1, L)$ is an increasing function of $\Lambda_1$, meaning

$$U_1'\left(g^{-1}(\Lambda_1, L) - L\right)\frac{\partial g^{-1}(\Lambda_1, L)}{\partial \Lambda_1} > 0.$$

We also know that $K - a_t^{\Lambda_1} > 0$ and $B_t \geq 0$ for every $t \in [0, T]$ and $\omega \in \Omega$. Therefore, (39) is non-negative. $V_1'$ is an increasing function, so its inverse $V_1'^{-1}$ must also be an increasing function. We can state that $V_1'^{-1'}(\theta E[u(\Lambda_1)]B_t) \geq 0$ and therefore (40) is non-negative.

To summarize, we have shown that $\varphi(\Lambda_1)$ is an increasing function of $\Lambda_1$ in each case. It is obvious that $\varphi(\Lambda_1)$ is continuous, so we state that $\varphi(\Lambda_1)$ is an increasing function of $\Lambda_1$ in the interval $\Lambda_1 \in (0, \infty)$.

Taking the integration of $\varphi(\Lambda_1)$ from 0 to $T$ and then taking the expectation on the integration, we obtain (35). So (35) increases when $\Lambda_1$ increases. Recalling that (33) and (34) also increase when $\Lambda_1$ increases, we conclude that $\mathbb{U}_1(\Lambda_1)$ is an increasing function of $\Lambda_1 \in (0, \infty)$. $\square$

*A.4. Proof of Theorem 2*

**Proof.** First, we want to verify that the process $a$ defined by (22) satisfies the constraint (6).

We consider three possibilities for $-\theta E\left[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1,L))+U_1(g^{-1}(\hat{\Lambda}_1,L)-L)\right]B_t$. If

$$-\theta E\left[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1,L))+U_1(g^{-1}(\hat{\Lambda}_1,L)-L)\right]B_t<0,$$

then $\hat{a}_t=0$ and the constraint (6) is trivially satisfied. If

$$-\theta E\left[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1,L))+U_1(g^{-1}(\hat{\Lambda}_1,L)-L)\right]B_t>V_1'(K),$$

then $\hat{a}_t=K$ and the constraint (6) is trivially satisfied. If

$$0\le-\theta E\left[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1,L))+U_1(g^{-1}(\hat{\Lambda}_1,L)-L)\right]B_t\le V_1'(K),$$

then the strict convexity of $V_1$ and the condition $V_1'(0)=0$ imply that

$$0\le V_1'^{-1}\left(-\theta E\left[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1,L))+U_1(g^{-1}(\hat{\Lambda}_1,L)-L)\right]B_t\right)\le K,$$

which is equivalent to $0\le\hat{a}_t\le K$. Hence, $\hat{a}_t\in[0,K]$ for each $t\in[0,T]$ and (22) satisfies the condition (6).

Let $a$ be a fixed admissible action process that satisfies (6). Then we find the first order conditions similar to (11) for $P_i$ and $q_t$,

$$U_2'(-P_i)-\Lambda^a U_1'(P_i-L_i)=0,\quad U_2'(q_t)-\Lambda^a U_1'(w_t-q_t)=0, \tag{41}$$

where $\Lambda^a$ is the Lagrangian multiplier. Since $U_1$ and $U_2$ are increasing functions, $\Lambda^a$ must be positive to make the equations above meaningful. The solution of the first order conditions is

$$P_i^a=g^{-1}(\Lambda^a,L_i),\quad q_t^a=-g^{-1}(\Lambda^a,-w_t).$$

We define

$$\mathbb{U}_a(\Lambda^a):=E\left[\int_0^T U_1(w_t+g^{-1}(\Lambda^a,-w_t))dt+\sum_{i=1}^{N^a(T)}U_1(g^{-1}(\Lambda^a,L_i)-L_i)-\int_0^T V_1(a_t)dt\right].$$

We denote the root of $\mathbb{U}_a(\Lambda^a)=R$ by $\hat{\Lambda}^a$ and correspondingly we define $\hat{P}_i^a:=g^{-1}(\hat{\Lambda}^a,L_i)$ and $\hat{q}_t^a:=-g^{-1}(\hat{\Lambda}^a,-w_t)$. Next, we discuss the existence of $\hat{\Lambda}^a$ for a fixed process $a$. We will show that $\hat{\Lambda}^a$ exists if for the fixed process $a$, there are compensation and premium processes such that (5) holds. For the fixed process $a$, let $P=\{P_i;i=1,2,\cdots\}$ and $q=\{q_t;t\in[0,T]\}$ be any adapted compensation sequence and premium process that satisfy (5). When $\Lambda^a\to\infty$,

$$g^{-1}(\Lambda^a,L_i)\to\infty,\quad g^{-1}(\Lambda^a,-w_t)\to\infty,$$

which yields $g^{-1}(\Lambda^a,L_i)\ge P_i$ for $i=1,2,\cdots$ and $g^{-1}(\Lambda^a,-w_t)\ge -q_t$ for $t\in[0,T]$. Recalling that $U_1$ is an increasing function, we have

$$\lim_{\Lambda^a\to\infty}\mathbb{U}_a(\Lambda^a)\ge E\left[\int_0^T U_1(w_t-q_t)dt+\sum_{i=1}^{N^a(T)}U_1(P_i-L_i)-\int_0^T V_1(a_t)dt\right]\ge R$$

from (5). When $\Lambda^a\to 0^+$,

$$g^{-1}(\Lambda^a,L_i)\to-\infty,\quad g^{-1}(\Lambda^a,-w_t)\to-\infty,$$

resulting in $\mathbb{U}_a(\Lambda^a)\to-\infty$ and consequently $\mathbb{U}_a(\Lambda^a)<R$. Due to the continuity of $\mathbb{U}_a(\Lambda^a)$, we see that there exists $\hat{\Lambda}^a\in(0,\infty)$ such that $\mathbb{U}_a(\hat{\Lambda}^a)=R$ holds.

We will prove Theorem 2 in two steps. First, we will show $\mathcal{J}(\hat{q}^a,\hat{P}^a,a)\ge\mathcal{J}(q,P,a)$ for any fixed action process $a$ that satisfies (6). Afterwards, we will show $\mathcal{J}(\hat{q},\hat{P},\hat{a})\ge\mathcal{J}(\hat{q}^a,\hat{P}^a,a)$. We need some preparation before starting the steps.

**Lemma 3.**

$$\theta E\left[\frac{1}{\hat{\Lambda}_1}U_2(-\hat{P})+U_1(\hat{P}-L)\right]B_t(a_t-\hat{a}_t)\ \ge\ V_1(\hat{a}_t)-V_1(a_t).$$

**Proof.** According to (20) and (21), it is sufficient to prove that

$$\theta E\left[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1,L))+U_1(g^{-1}(\hat{\Lambda}_1,L)-L)\right]B_t(a_t-\hat{a}_t)\ \geq\ V_1(\hat{a}_t)-V_1(a_t).$$

If $0<\hat{a}_t<K$,

$$V_1'(\hat{a}_t)(\hat{a}_t-a_t)=\theta E[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1,L))+U_1(g^{-1}(\hat{\Lambda}_1,L)-L)]B_t(a_t-\hat{a}_t).$$

If $\hat{a}_t=K$, then

$$\hat{a}_t-a_t=K-a_t\geq 0\ \text{ and }\ V_1'(\hat{a}_t)=V_1'(K)<-\theta E[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1,L))+U_1(g^{-1}(\hat{\Lambda}_1,L)-L)]B_t$$

which yields

$$V_1'(\hat{a}_t)(\hat{a}_t-a_t)\leq\theta E[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1,L))+U_1(g^{-1}(\hat{\Lambda}_1,L)-L)]B_t(a_t-\hat{a}_t).$$

If $\hat{a}_t=0$, then

$$\hat{a}_t-a_t=0-a_t\leq 0\ \text{ and }\ V_1'(\hat{a}_t)=V_1'(0)=0>-\theta E[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1,L))+U_1(g^{-1}(\hat{\Lambda}_1,L)-L)]B_t$$

which yields

$$V_1'(\hat{a}_t)(\hat{a}_t-a_t)\leq\theta E[\frac{1}{\hat{\Lambda}_1}U_2(-g^{-1}(\hat{\Lambda}_1,L))+U_1(g^{-1}(\hat{\Lambda}_1,L)-L)]B_t(a_t-\hat{a}_t).$$

Due to the convexity of $V_1$, we have $V_1'(\hat{a}_t)(\hat{a}_t-a_t)\geq V_1(\hat{a}_t)-V_1(a_t)$. The required statement follows. □

**Step 1.** Since $U_1$ and $U_2$ are both concave functions, we obtain the inequality

$$\begin{aligned}&\int_0^T U_1(w_t-q_t)dt+\sum_{i=1}^{N^a(T)}U_1(P_i-L_i)-\left(\int_0^T U_1(w_t-\hat{q}_t^a)dt+\sum_{i=1}^{N^a(T)}U_1(\hat{P}_i^a-L_i)\right)\\&\leq\int_0^T U_1'(w_t-\hat{q}_t^a)(\hat{q}_t^a-q_t)dt+\sum_{i=1}^{N^a(T)}\Big(U_1'(\hat{P}_i^a-L_i)(P_i-\hat{P}_i^a)\Big).\end{aligned}\tag{42}$$

Furthermore, (4) implies

$$\mathcal{J}(\hat{q}^a,\hat{P}^a,a)-\mathcal{J}(q,P,a)=E\left[\int_0^T\big(U_2(\hat{q}_t^a)-U_2(q_t)\big)dt+\sum_{i=1}^{N^a(T)}\Big(U_2(-\hat{P}_i^a)-U_2(-P_i)\Big)\right],$$

which yields

$$\mathcal{J}(\hat{q}^a,\hat{P}^a,a)-\mathcal{J}(q,P,a)\geq E\left[\int_0^T U_2'(\hat{q}_t^a)(\hat{q}_t^a-q_t)dt+\sum_{i=1}^{N^a(T)}\Big(U_2'(-\hat{P}_i^a)(P_i-\hat{P}_i^a)\Big)\right].\tag{43}$$

According to (41), we can replace $U_2'(\hat{q}_t^a)$ by $\hat{\Lambda}^aU_1'(w_t-\hat{q}_t^a)$ and replace $U_2'(-\hat{P}_i^a)$ by $\hat{\Lambda}^aU_1'(\hat{P}_i^a-L_i)$ in (43). Comparing (42) and (43), we obtain

$$\begin{aligned}&\mathcal{J}(\hat{q}^a,\hat{P}^a,a)-\mathcal{J}(q,P,a)\\&\geq E\left[\int_0^T\hat{\Lambda}^aU_1'(w_t-\hat{q}_t^a)(\hat{q}_t^a-q_t)dt+\hat{\Lambda}^a\sum_{i=1}^{N^a(T)}\Big(U_1'(\hat{P}_i^a-L_i)(P_i-\hat{P}_i^a)\Big)\right]\\&\geq\hat{\Lambda}^aE\left[\int_0^T U_1(w_t-q_t)dt+\sum_{i=1}^{N^a(T)}U_1(P_i-L_i)-\left(\int_0^T U_1(w_t-\hat{q}_t^a)dt+\sum_{i=1}^{N^a(T)}U_1(\hat{P}_i^a-L_i)\right)\right].\end{aligned}$$

According to (5), we obtain

$$\mathcal{J}(\hat{q}^a,\hat{P}^a,a)-\mathcal{J}(q,P,a)\geq\hat{\Lambda}^a\left((R+E\left[\int_0^T V_1(a_t)dt\right])-(R+E\left[\int_0^T V_1(a_t)dt\right])\right)=0.$$

Therefore, $\hat{q}^a$ and $\hat{P}^a$ are the optimal controls when $a$ is the fixed action process.

**Step 2.** As a Lagrangian multiplier, $\hat{\Lambda}^a$ is a constant. The randomness of $\hat{P}_i^a$ depends on $L_i$ only, so $\hat{P}_i^a$ is independent of $N^a(t)$ for $i = 1, 2, \cdots$, and we get the following equations for any $a$ satisfying (5) and (6).

$$E\left[\sum_{i=1}^{N^a(T)} U_1(\hat{P}_i^a - L_i)\right] = E[N^a(T)]E\left[U_1(\hat{P}^a - L)\right],\ E\left[\sum_{i=1}^{N^a(T)} U_2(-\hat{P}_i^a)\right] = E[N^a(T)]E\left[U_2(-\hat{P}^a)\right], \tag{44}$$

where $E[N^a(T)] = \theta E\left[\bar{B} - \int_0^T a_t B_t dt\right]$ from (9). Similarly, we obtain

$$E\left[\sum_{i=1}^{N^{\hat{a}}(T)} U_1(\hat{P}_i - L_i)\right] = E[N^{\hat{a}}(T)]E\left[U_1(\hat{P} - L)\right],\ E\left[\sum_{i=1}^{N^{\hat{a}}(T)} U_2(-\hat{P}_i)\right] = E[N^{\hat{a}}(T)]E\left[U_2(-\hat{P})\right], \tag{45}$$

where $E[N^{\hat{a}}(T)] = \theta E\left[\bar{B} - \int_0^T \hat{a}_t B_t dt\right]$. Hence, the difference between $\mathcal{J}(\hat{q}, \hat{P}, \hat{a})$ and $\mathcal{J}(\hat{q}^a, \hat{P}^a, a)$ is

$$\begin{aligned}
&\mathcal{J}(\hat{q}, \hat{P}, \hat{a}) - \mathcal{J}(\hat{q}^a, \hat{P}^a, a)\\
&= \int_0^T \left(U_2(\hat{q}_t) - U_2(\hat{q}_t^a)\right) dt + E\left[\sum_{i=1}^{N^{\hat{a}}(T)} U_2(-\hat{P}_i) - \sum_{i=1}^{N^a(T)} U_2(-\hat{P}_i^a)\right]\\
&= \int_0^T \left(U_2(\hat{q}_t) - U_2(\hat{q}_t^a)\right) dt\\
&\quad + \theta E\left[\bar{B} - \int_0^T \hat{a}_t B_t dt\right] E\left[U_2(-\hat{P})\right] - \theta E\left[\bar{B} - \int_0^T a_t B_t dt\right] E\left[U_2(-\hat{P}^a)\right]\\
&= \int_0^T \left(U_2(\hat{q}_t) - U_2(\hat{q}_t^a)\right) dt + \theta E\left[\int_0^T (a_t - \hat{a}_t) B_t dt\right] E\left[U_2(-\hat{P})\right]\\
&\quad + \theta E\left[\bar{B} - \int_0^T a_t B_t dt\right] E\left[U_2(-\hat{P}) - U_2(-\hat{P}^a)\right].
\end{aligned}$$

Recalling $E\left[\bar{B} - \int_0^T a_t B_t dt\right] \geq 0$ and the concavity of the utility function $U_2$, we obtain

$$\begin{aligned}
\mathcal{J}(\hat{q}, \hat{P}, \hat{a}) - \mathcal{J}(\hat{q}^a, \hat{P}^a, a) \geq & \int_0^T U_2'(\hat{q}_t)(\hat{q}_t - \hat{q}_t^a) dt + \theta E\left[\int_0^T (a_t - \hat{a}_t) B_t dt\right] E\left[U_2(-\hat{P})\right]\\
& + \theta E\left[\bar{B} - \int_0^T a_t B_t dt\right] E\left[U_2'(-\hat{P})(\hat{P}^a - \hat{P})\right].
\end{aligned}$$

According to (11), this inequality can be rewritten as

$$\begin{aligned}
\mathcal{J}(\hat{q}, \hat{P}, \hat{a}) - \mathcal{J}(\hat{q}^a, \hat{P}^a, a) \geq & \int_0^T \hat{\Lambda}_1 U_1'(w_t - \hat{q}_t)(\hat{q}_t - \hat{q}_t^a) dt + \theta E\left[\int_0^T (a_t - \hat{a}_t) B_t dt\right] E\left[U_2(-\hat{P})\right]\\
& + \theta E\left[\bar{B} - \int_0^T a_t B_t dt\right] E\left[\hat{\Lambda}_1 U_1'(\hat{P} - L)(\hat{P}^a - \hat{P})\right].
\end{aligned}$$

Due to the concavity of the utility function $U_1$, we have

$$\begin{aligned}
&\mathcal{J}(\hat{q}, \hat{P}, \hat{a}) - \mathcal{J}(\hat{q}^a, \hat{P}^a, a)\\
&\geq \hat{\Lambda}_1 \int_0^T \left(U_1(w_t - \hat{q}_t^a) - U_1(w_t - \hat{q}_t)\right) dt
\end{aligned}$$

$$+\hat{\Lambda}_1\theta E\left[\int_0^T (a_t-\hat{a}_t)B_t dt\right] E\left[\frac{1}{\hat{\Lambda}_1}U_2(-\hat{P})+U_1(\hat{P}-L)-U_1(\hat{P}-L)\right]$$

$$+\hat{\Lambda}_1\theta E\left[\bar{B}-\int_0^T a_t B_t dt\right] E\Big[U_1(\hat{P}^a-L)-U_1(\hat{P}-L)\Big].$$

Applying Lemma 3, we obtain

$$\begin{aligned}
&\frac{1}{\hat{\Lambda}_1}\Big(\mathcal{J}(\hat{q},\hat{P},\hat{a})-\mathcal{J}(\hat{q}^a,\hat{P}^a,a)\Big)\\
&\geq \int_0^T \Big(U_1(w_t-\hat{q}_t^a)-U_1(w_t-\hat{q}_t)\Big)dt + E\left[\int_0^T \Big(V_1(\hat{a}_t)-V_1(a_t)\Big)dt\right]\\
&\quad +\theta E\left[\int_0^T(\hat{a}_t-a_t)B_t dt\right]E\Big[U_1(\hat{P}-L)\Big]+\theta E\left[\bar{B}-\int_0^T a_t B_t dt\right]E\Big[U_1(\hat{P}^a-L)-U_1(\hat{P}-L)\Big]\\
&= \int_0^T \Big(U_1(w_t-\hat{q}_t^a)-U_1(w_t-\hat{q}_t)\Big)dt + E\left[\int_0^T \Big(V_1(\hat{a}_t)-V_1(a_t)\Big)dt\right]\\
&\quad +\theta E\left[\bar{B}-\int_0^T a_t B_t dt\right]E\Big[U_1(\hat{P}^a-L)\Big]-\theta E\left[\bar{B}-\int_0^T \hat{a}_t B_t dt\right]E\Big[U_1(\hat{P}-L)\Big].
\end{aligned}$$

Applying (44) and (45)to the expression above, we obtain

$$\begin{aligned}
\frac{1}{\hat{\Lambda}_1}\Big(\mathcal{J}(\hat{q},\hat{P},\hat{a})-\mathcal{J}(\hat{q}^a,\hat{P}^a,a)\Big) &\geq E\left[\int_0^T U_1(w_t-\hat{q}_t^a)dt+\sum_{i=1}^{N^a(T)}U_1(\hat{P}_i^a-L_i)-\int_0^T V_1(a_t)dt\right]\\
&\quad -E\left[\int_0^T U_1(w_t-\hat{q}_t)dt+\sum_{i=1}^{N^{\hat{a}}(T)}U_1(\hat{P}_i-L_i)-\int_0^T V_1(\hat{a}_t)dt\right]\\
&= R-R=0.
\end{aligned}$$

Therefore, $\mathcal{J}(\hat{q},\hat{P},\hat{a})\geq\mathcal{J}(q,P,a)$ for every admissible control $(q,P,a)$ that satisfies the constraints of Problem 1. If $\hat{\Lambda}_1>0$ satisfies (19), we conclude that $(\hat{q},\hat{P},\hat{a})$ is the optimal solution. □

### A.5. Proof of Proposition 2

**Proof.** Let $\{a_t\}_{t\in[0,T]}$ be any action process that satisfies the constraints of Problem 2. We will compare the utilities from implementing the two action processes $a^*$ and $a$. We denote by $D(a^*,a)$ the difference of the expected total utilities associated with $a^*$ and $a$. That is,

$$\begin{aligned}
D(a^*,a):=&E\left[\int_0^T U_1(w_t)dt+\sum_{i=1}^{N^{a^*}(T)}U_1(-L_i)-\int_0^T V_1(a_t^*)dt\right]\\
&-E\left[\int_0^T U_1(w_t)dt+\sum_{i=1}^{N^{a}(T)}U_1(-L_i)-\int_0^T V_1(a_t)dt\right].
\end{aligned}$$

According to (23), we have

$$\begin{aligned}
D(a^*,a)=&E\left[U_1(-L)\right]\left(\theta\bar{B}-\theta E\left[\int_0^T a_t^* B_t dt\right]\right)-E\left[U_1(-L)\right]\left(\theta\bar{B}-\theta E\left[\int_0^T a_t B_t dt\right]\right)\\
&+E\left[\int_0^T\big(V_1(a_t)-V_1(a_t^*)\big)dt\right]\\
=&\theta E[U_1(-L)]\left(E\left[\int_0^T(a_t-a_t^*)B_t dt\right]\right)+E\left[\int_0^T\big(V_1(a_t)-V_1(a_t^*)\big)dt\right].
\end{aligned}$$

The convexity of $V_1$ implies

$$D(a^*, a) \geq \theta E[U_1(-L)] \left( E\left[ \int_0^T (a_t - a_t^*) B_t dt \right] \right) + E\left[ \int_0^T V_1'(a_t^*)(a_t - a_t^*) dt \right]$$
$$= E\left[ \int_0^T \left( V_1'(a_t^*) + \theta E[U_1(-L)] B_t \right) (a_t - a_t^*) dt \right].$$

Next, we consider the two cases described in equation (25). If $a_t^* = K$, from (25), we have

$$a_t - a_t^* = a_t - K \leq 0 \text{ and } V_1'(a_t^*) = V_1'(K) \leq -\theta E[U_1(-L)] B_t$$

which yields

$$\left( V_1'(a_t^*) + \theta E[U_1(-L)] B_t \right) (a_t - a_t^*) \geq 0.$$

Otherwise, if $a_t^* = V_1'^{-1}\left(-\theta B_t E[U_1(-L)]\right)$, we have $V_1'(a_t^*) = -\theta E[U_1(-L)] B_t$, which yields

$$\left( V_1'(a_t^*) + \theta E[U_1(-L)] B_t \right) (a_t - a_t^*) = 0.$$

Now we can obtain $D(a^*, a) \geq 0$ and conclude that the action process $a^*$ is the optimal control of Problem 2. □

*A.6. Proof of Theorem 3*

**Proof.** Since $U_2(0) \leq 0$ and $\Lambda_1 > 0$, we have $-\dfrac{1}{\Lambda_1} U_2(0) \theta B_t \geq 0$. We will consider three cases for $a_t^{\Lambda_1}$.
(i) Consider $a_t^{\Lambda_1} = 0$. Then, $V_1'(a_t^{\Lambda_1}) = V_1'(0) = 0$. Noting $a_t^* > 0$, we know $V_1'(a_t^*) > 0$. It follows that

$$V_1'(a_t^{\Lambda_1}) \leq V_1'(a_t^*) \leq V_1'(a_t^*) - \frac{1}{\Lambda_1} U_2(0) \theta B_t.$$

(ii) Consider $a_t^{\Lambda_1} = K$. From (15), we have

$$V_1'(a_t^{\Lambda_1}) = V_1'(K) < -\theta E\left[ \frac{1}{\Lambda_1} U_2(-g^{-1}(\Lambda_1, L)) + U_1(g^{-1}(\Lambda_1, L) - L) \right] B_t.$$

If $a_t^* = V_1'^{-1}\left(-\theta B_t E[U_1(-L)]\right)$, we have

$$V_1'(a_t^*) = -\theta B_t E[U_1(-L)].$$

It follows that

$$V_1'(a_t^{\Lambda_1}) - V_1'(a_t^*) \leq -\theta B_t E\left[ \frac{1}{\Lambda_1} U_2(-g^{-1}(\Lambda_1, L)) + U_1(g^{-1}(\Lambda_1, L) - L) - U_1(-L) \right].$$

The concavity of the utility functions implies

$$U_1(g^{-1}(\Lambda_1, L) - L) - U_1(-L) \geq g^{-1}(\Lambda_1, L) U_1'(g^{-1}(\Lambda_1, L) - L)$$

and

$$U_2(0) - U_2(-g^{-1}(\Lambda_1, L)) \leq g^{-1}(\Lambda_1, L) U_2'(-g^{-1}(\Lambda_1, L))$$

for every $L \in R_L$, so we have

$$\begin{aligned} V_1'(a_t^{\Lambda_1}) - V_1'(a_t^*) &\leq -\theta B_t E\left[ \frac{1}{\Lambda_1} U_2(-g^{-1}(\Lambda_1, L)) + g^{-1}(\Lambda_1, L) U_1'(g^{-1}(\Lambda_1, L) - L) \right] \\ &= -\theta B_t E\left[ \frac{1}{\Lambda_1} U_2(-g^{-1}(\Lambda_1, L)) + g^{-1}(\Lambda_1, L) \frac{1}{\Lambda_1} U_2'(-g^{-1}(\Lambda_1, L)) \right] \\ &\leq -\frac{1}{\Lambda_1} \theta U_2(0) B_t. \end{aligned} \tag{46}$$

If $a_t^* = K$, then

$$V_1'(a_t^{\Lambda_1}) - V_1'(a_t^*) = V_1'(K) - V_1'(K) = 0 \leq -\frac{1}{\Lambda_1} \theta U_2(0) B_t.$$

(iii) Consider $a_t^{\Lambda_1} = V_1'^{-1}\left( -\theta E[\dfrac{1}{\Lambda_1} U_2(-g^{-1}(\Lambda_1, L)) + U_1(g^{-1}(\Lambda_1, L) - L)] B_t \right)$.

If $a_t^* = V_1'^{-1}(-\theta B_t E[U_1(-L)])$, we have

$$V_1'(a_t^{\Lambda_1}) - V_1'(a_t^*) = -\theta B_t E\left[\frac{1}{\Lambda_1}U_2(-g^{-1}(\Lambda_1, L)) + U_1(g^{-1}(\Lambda_1, L) - L) - U_1(-L)\right].$$

Now we can repeat (46) to get $V_1'(a_t^{\Lambda_1}) - V_1'(a_t^*) \le -\frac{1}{\Lambda_1}\theta U_2(0)B_t$.

If $a_t^* = K$, it is obvious that

$$V_1'(a_t^{\Lambda_1}) - V_1'(a_t^*) = V_1'(a_t^{\Lambda_1}) - V_1'(K) < 0 \le -\frac{1}{\Lambda_1}\theta U_2(0)B_t.$$

As a summary of all the cases discussed above, the required statement is proved. □

### A.7. Proof of Lemma 2

**Proof.** We consider $\phi(\underline{\lambda}) := E[U_1(g^{-1}(\underline{\lambda}, L) - L)]$ as a function of $\underline{\lambda}$. From the definition of the function $g$, we have $g(0, L) = \frac{U_2'(0)}{U_1'(0-L)}$. It follows that $g^{-1}\left(\frac{U_2'(0)}{U_1'(-L)}, L\right) = 0$. When $\underline{\lambda} = \frac{U_2'(0)}{U_1'(-\inf R_L)}$, $\underline{\lambda} \ge \frac{U_2'(0)}{U_1'(-L)}$ for every $L \in R_L$ due to the concavity of $U_1$. Since $g^{-1}(\cdot, x_2)$ is an increasing function, $g^{-1}(\underline{\lambda}, L) \ge 0$ for every $L \in R_L$. It results in $\phi(\underline{\lambda}) \ge E[U_1(-L)]$. When $\underline{\lambda} = \frac{U_2'(0)}{U_1'(-\sup R_L)}$, $\underline{\lambda} \le \frac{U_2'(0)}{U_1'(-L)}$ for every $L \in R_L$. Then we have $g^{-1}(\underline{\lambda}, L) \le 0$ for every $L \in R_L$ and $\phi(\underline{\lambda}) \le E[U_1(-L)]$. $\phi(\underline{\lambda})$ is continuous and monotone because $g^{-1}$ and $U_1$ are continuous and monotone functions. Using the Mean Value Theorem, we can conclude there is a unique $\underline{\Lambda_1}$ such that $\phi(\underline{\Lambda_1}) = E[U_1(-L)]$ and $\underline{\Lambda_1} \in \left[\frac{U_2'(0)}{U_1'(-\sup R_L)}, \frac{U_2'(0)}{U_1'(-\inf R_L)}\right]$.

Noting $P^{\underline{\Lambda_1}} = g^{-1}(\underline{\Lambda_1}, L)$, we have $E[U_1(P^{\underline{\Lambda_1}} - L)] = E[U_1(-L)]$ according to (27). From (14),

$$\mathbb{U}_1(\underline{\Lambda_1}) = \int_0^T U_1(w_t - q_t^{\underline{\Lambda_1}})dt + E[U_1(-L)]\theta\left(\bar{B} - E[\int_0^T a_t^{\underline{\Lambda_1}} B_t dt]\right) - E\left[\int_0^T V_1(a_t^{\underline{\Lambda_1}})dt\right].$$

Comparing (26) and the expression above, we obtain

$$R - \mathbb{U}_1(\underline{\Lambda_1}) = \int_0^T \left(U_1(w_t) - U_1(w_t - q_t^{\underline{\Lambda_1}})\right)dt + \theta E[U_1(-L)]E\left[\int_0^T (a_t^{\underline{\Lambda_1}} - a_t^*)B_t dt\right]$$
$$+ E\left[\int_0^T \left(V_1(a_t^{\underline{\Lambda_1}}) - V_1(a_t^*)\right)dt\right].$$

The range of $\underline{\Lambda_1}$ indicates that $\underline{\Lambda_1} < \frac{U_2'(0)}{U_1'(0)} < \frac{U_2'(0)}{U_1'(w_t)}$. It yields $q_t^{\underline{\Lambda_1}} > 0$ and $U_1(w_t) - U_1(w_t - q_t^{\underline{\Lambda_1}}) > 0$ for $t \in [0, T]$. Thus, the equation above implies

$$R - \mathbb{U}_1(\underline{\Lambda_1}) > \theta E[U_1(-L)]E\left[\int_0^T (a_t^{\underline{\Lambda_1}} - a_t^*)B_t dt\right] + E\left[\int_0^T \left(V_1(a_t^{\underline{\Lambda_1}}) - V_1(a_t^*)\right)dt\right].$$

Since $V_1(\cdot)$ is a convex function, $V_1(a_t^{\underline{\Lambda_1}}) - V_1(a_t^*) \ge V_1'(a_t^*)\left(a_t^{\underline{\Lambda_1}} - a_t^*\right)$. Hence,

$$R - \mathbb{U}_1(\underline{\Lambda_1}) > E\left[\int_0^T \left(\theta E[U_1(-L)]B_t + V_1'(a_t^*)\right)(a_t^{\underline{\Lambda_1}} - a_t^*)dt\right]. \tag{47}$$

Next, we consider the two cases described in (25). If $a_t^* = K$, then from (25),

$$V_1'(a_t^*) < -\theta E[U_1(-L)]B_t \quad \text{and} \quad a_t^{\underline{\Lambda_1}} \le a_t^*,$$

which yield $\left(\theta E[U_1(-L)]B_t + V_1'(a_t^*)\right)(a_t^{\underline{\Lambda_1}} - a_t^*) \ge 0$. If $a_t^* = V_1'^{-1}(-\theta E[U_1(-L)]B_t)$, then

$$V_1'(a_t^*) = -\theta E[U_1(-L)]B_t$$

which yields $\left(\theta E[U_1(-L)]B_t + V_1'(a_t^*)\right)(a_t^{\underline{\Lambda_1}} - a_t^*) = 0$. Then, from (47), we obtain $R - \mathbb{U}_1(\underline{\Lambda_1}) > 0$. □

*A.8. Proof of Theorem 4*

**Proof.** Our first objective is to show that $\mathbb{U}_1(\Lambda_1) \geq R$ when $\Lambda_1 \to \infty$. Here $R$ is presented in (26). Since $\lim_{\Lambda_1\to\infty} \frac{1}{\Lambda_1} U_2(0)\theta B_t = 0$ almost surely for each $t \in [0, T]$, we have $\lim_{\Lambda_1\to\infty} a_t^{\Lambda_1} \leq a_t^*$ almost surely for $t \in [0, T]$ according to Theorem 3. From the definition of $P^{\Lambda_1}$ and $q_t^{\Lambda_1}$ in (15), we have

$$\frac{U_2'(-P^{\Lambda_1})}{U_1'(P^{\Lambda_1} - L)} = \Lambda_1 \quad \text{and} \quad \frac{U_2'(q_t^{\Lambda_1})}{U_1'(w_t - q_t^{\Lambda_1})} = \Lambda_1.$$

When $\Lambda_1 \to \infty$, we obtain $P^{\Lambda_1} \to \infty$ and $q_t^{\Lambda_1} \to -\infty$, which means $P^{\Lambda_1} > 0$ for every $L \in R_L$ and $q_t^{\Lambda_1} < 0$ for every $t \in [0, T]$. To simplify the notation, we rewrite $\bar{B}$ as $\bar{B} = \int_0^T b_t dt$, where

$$b_t := (1 - e^{-t} A_0)e^{-\delta t} \left( Y_0 + \mu \int_0^t \rho(u)e^{\delta u} du \right).$$

If $\lim_{\Lambda_1\to\infty} a_t^{\Lambda_1} = K$, then $a_t^* = K$ and

$$\begin{aligned}
&\left(\theta E\left[U_1(P^{\Lambda_1} - L)\right]\left(b_t - a_t^{\Lambda_1} B_t\right) - V_1(a_t^{\Lambda_1})\right) - \left(\theta E[U_1(-L)]\left(b_t - a_t^* B_t\right) - V_1(a_t^*)\right) \\
&= \left(\theta E\left[U_1(P^{\Lambda_1} - L)\right](b_t - K B_t) - V_1(K)\right) - \left(\theta E[U_1(-L)](b_t - K B_t) - V_1(K)\right) \\
&= \theta E\left[U_1(P^{\Lambda_1} - L) - U_1(-L)\right]\left(b_t - a_t^* B_t\right)
\end{aligned} \tag{48}$$

almost surely when $\Lambda_1 \to \infty$. If $\lim_{\Lambda_1\to\infty} a_t^{\Lambda_1} < K$, then from (15), we have

$$\begin{aligned}
\lim_{\Lambda_1\to\infty} V_1'(a_t^{\Lambda_1}) &\geq \lim_{\Lambda_1\to\infty} -\theta E\left[\frac{1}{\Lambda_1} U_2(-g^{-1}(\Lambda_1, L)) + U_1(g^{-1}(\Lambda_1, L) - L)\right] B_t \\
&= \lim_{\Lambda_1\to\infty} -\theta E\left[\frac{1}{\Lambda_1} U_2(-P^{\Lambda_1}) + U_1(P^{\Lambda_1} - L)\right] B_t.
\end{aligned}$$

Noting $\lim_{\Lambda_1\to\infty} P^{\Lambda_1} > 0$ and the negativity property of $U_2$ in (3), we get

$$\lim_{\Lambda_1\to\infty} V_1'(a_t^{\Lambda_1}) \geq \lim_{\Lambda_1\to\infty} -\theta E\left[U_1(P^{\Lambda_1} - L)\right] B_t.$$

Hence,

$$V_1(a_t^*) - \lim_{\Lambda_1\to\infty} V_1(a_t^{\Lambda_1}) \geq \lim_{\Lambda_1\to\infty} V_1'(a_t^{\Lambda_1})(a_t^* - a_t^{\Lambda_1}) \geq \lim_{\Lambda_1\to\infty} -\theta E\left[U_1(P^{\Lambda_1} - L)\right] B_t(a_t^* - a_t^{\Lambda_1})$$

almost surely, and consequently

$$\begin{aligned}
&\left(\theta E\left[U_1(P^{\Lambda_1} - L)\right]\left(b_t - a_t^{\Lambda_1} B_t\right) - V_1(a_t^{\Lambda_1})\right) - \left(\theta E[U_1(-L)]\left(b_t - a_t^* B_t\right) - V_1(a_t^*)\right) \\
&\geq \theta E\left[U_1(P^{\Lambda_1} - L)\right]\left(b_t - a_t^{\Lambda_1} B_t\right) - \theta E[U_1(-L)]\left(b_t - a_t^* B_t\right) - \theta E\left[U_1(P^{\Lambda_1} - L)\right] B_t(a_t^* - a_t^{\Lambda_1}) \\
&= \theta E\left[U_1(P^{\Lambda_1} - L) - U_1(-L)\right]\left(b_t - a_t^* B_t\right)
\end{aligned} \tag{49}$$

almost surely when $\Lambda_1 \to \infty$. From (48) and (49), we see that it is almost surely that

$$\begin{aligned}
&\left(\theta E\left[U_1(P^{\Lambda_1} - L)\right]\left(b_t - a_t^{\Lambda_1} B_t\right) - V_1(a_t^{\Lambda_1})\right) - \left(\theta E[U_1(-L)]\left(b_t - a_t^* B_t\right) - V_1(a_t^*)\right) \\
&\quad \geq \theta E\left[U_1(P^{\Lambda_1} - L) - U_1(-L)\right]\left(b_t - a_t^* B_t\right)
\end{aligned}$$

for each case when $\Lambda_1 \to \infty$. Integrating and taking expectation on both sides of the above inequality, we obtain

$$\begin{aligned}
&\left(\theta E\left[U_1(P^{\Lambda_1} - L)\right] E\left[\int_0^T b_t dt - \int_0^T a_t^{\Lambda_1} B_t dt\right] - E\left[\int_0^T V_1(a_t^{\Lambda_1}) dt\right]\right) \\
&\quad - \left(\theta E\left[U_1(-L)\right] E\left[\int_0^T b_t dt - \int_0^T a_t^* B_t dt\right] - E\left[\int_0^T V_1(a_t^*) dt\right]\right)
\end{aligned}$$

$$\geq \theta E\left[U_1(P^{\Lambda_1}-L)-U_1(-L)\right] E\left[\int_0^T b_t dt-\int_0^T a_t^* B_t dt\right],$$

which is equivalent to

$$\begin{aligned}&\left(\theta E\left[U_1(P^{\Lambda_1}-L)\right] E\left[\bar{B}-\int_0^T a_t^{\Lambda_1} B_t dt\right]-E\left[\int_0^T V_1(a_t^{\Lambda_1}) dt\right]\right)\\&\quad-\left(\theta E\left[U_1(-L)\right] E\left[\bar{B}-\int_0^T a_t^* B_t dt\right]-E\left[\int_0^T V_1(a_t^*) dt\right]\right)\\&\geq \theta E\left[U_1(P^{\Lambda_1}-L)-U_1(-L)\right] E\left[\bar{B}-\int_0^T a_t^* B_t dt\right].\end{aligned} \tag{50}$$

Recalling $\bar{B}-E\left[\int_0^T a_t^* B_t dt\right]\geq 0$ and $\lim_{\Lambda_1\to\infty} P^{\Lambda_1}>0$ for every $L\in R_L$, we obtain that the right-hand-side of (50) is non-negative. Thus,

$$\begin{aligned}&\theta E\left[U_1(P^{\Lambda_1}-L)\right] E\left[\bar{B}-\int_0^T a_t^{\Lambda_1} B_t dt\right]-E\left[\int_0^T V_1(a_t^{\Lambda_1}) dt\right]\\&\geq \theta E\left[U_1(-L)\right] E\left[\bar{B}-\int_0^T a_t^* B_t dt\right]-E\left[\int_0^T V_1(a_t^*) dt\right]\end{aligned} \tag{51}$$

when $\Lambda_1\to\infty$. Recalling that $\lim_{\Lambda_1\to\infty} q_t^{\Lambda_1}<0$ for $t\in[0,T]$, we have

$$\lim_{\Lambda_1\to\infty} U_1(w_t-q_t^{\Lambda_1})>U_1(w_t) \tag{52}$$

for $t\in[0,T]$. Combining (51) and (52), we obtain

$$\begin{aligned}&\int_0^T U_1(w_t-q_t^{\Lambda_1})dt+\theta E\left[U_1(P^{\Lambda_1}-L)\right]\left(\bar{B}-E\left[\int_0^T a_t^{\Lambda_1} B_t dt\right]\right)-E\left[\int_0^T V_1(a_t^{\Lambda_1}) dt\right]\\&>\int_0^T U_1(w_t)dt+\theta E[U_1(-L)]\left(\bar{B}-E\left[\int_0^T a_t^* B_t dt\right]\right)-E\left[\int_0^T V_1(a_t^*) dt\right]\end{aligned}$$

when $\Lambda_1\to\infty$. This is equivalent to $\lim_{\Lambda_1\to\infty}\mathbb{U}_1(\Lambda_1)>R$. Lemma 2 states that $\mathbb{U}_1(\underline{\Lambda}_1)<R$. $\mathbb{U}_1(\Lambda_1)$ is a continuous function of $\Lambda_1$. From Theorem 1, we also know that $\mathbb{U}_1(\Lambda_1)$ is an increasing function of $\Lambda_1$. Therefore, there is a unique $\hat{\Lambda}_1$ such that (19) holds and $\hat{\Lambda}_1\in(\underline{\Lambda}_1,\infty)$. □

*A.9. Proof of Corollary 2*

**Proof.** From (15), we see that $q_t^{\Lambda_1}=-g^{-1}(\Lambda_1,-w_t)=0$ when $\Lambda_1=g(0,-w_t)=\dfrac{U_2'(0)}{U_1'(w_t)}$ for each $t\in[0,T]$. Noting that $\bar{\Lambda}_1=\dfrac{U_2'(0)}{U_1'(w_{sup})}\geq\dfrac{U_2'(0)}{U_1'(w_t)}$ and that $q_t^{\Lambda_1}$ is a decreasing function of $\Lambda_1$ for $t\in[0,T]$, we have $q_t^{\bar{\Lambda}_1}\leq 0$ for $t\in[0,T]$.

From (15), we see that $P^{\Lambda_1}=g^{-1}(\Lambda_1,L)=0$ when $\Lambda_1=g(0,L)=\dfrac{U_2'(0)}{U_1'(-L)}$ for each $L\in R_L$. Noting that $\bar{\Lambda}_1=\dfrac{U_2'(0)}{U_1'(w_{sup})}>\dfrac{U_2'(0)}{U_1'(-L)}$ and that $P^{\Lambda_1}$ is an increasing function of $\Lambda_1$ for $L\in R_L$, we have $P^{\bar{\Lambda}_1}>0$ for $L\in R_L$.

From (14) and (26), we obtain

$$\begin{aligned}\mathbb{U}_1(\bar{\Lambda}_1)-R=&\int_0^T\left(U_1(w_t-q_t^{\bar{\Lambda}_1})-U_1(w_t)\right)dt\\&+\theta E\left[U_1(P^{\bar{\Lambda}_1}-L)\right]\left(\bar{B}-E\left[\int_0^T a_t^{\bar{\Lambda}_1} B_t dt\right]\right)-\theta E[U_1(-L)]\left(\bar{B}-E\left[\int_0^T a_t^* B_t dt\right]\right)\end{aligned}$$

$$+E\left[\int_0^T \left(V_1(a_t^*)-V_1(a_t^{\bar{\Lambda}_1})\right)dt\right].$$

In the above equation, we have $U_1(w_t-q_t^{\bar{\Lambda}_1})-U_1(w_t)\geq 0$ for $t\in[0,T]$ because $q_t^{\bar{\Lambda}_1}\leq 0$ for $t\in[0,T]$. Since $P^{\bar{\Lambda}_1}>0$ for $L\in R_L$, we have

$$-\theta E[U_1(-L)]\left(\bar{B}-E\left[\int_0^T a_t^* B_t dt\right]\right)\geq -\theta E[U_1(P^{\bar{\Lambda}_1}-L)]\left(\bar{B}-E\left[\int_0^T a_t^* B_t dt\right]\right).$$

From (15), we also have

$$\begin{aligned} V_1(a_t^*)-V_1(a_t^{\bar{\Lambda}_1}) &\geq V_1'(a_t^{\bar{\Lambda}_1})(a_t^*-a_t^{\bar{\Lambda}_1}) \\ &=-\theta E\left[\frac{1}{\bar{\Lambda}_1}U_2(-P^{\bar{\Lambda}_1})+U_1(P^{\bar{\Lambda}_1}-L)\right]B_t(a_t^*-a_t^{\bar{\Lambda}_1}). \end{aligned}$$

Hence, we obtain

$$\begin{aligned} \mathbb{U}_1(\bar{\Lambda}_1)-R \geq\ & \theta E\left[U_1(P^{\bar{\Lambda}_1}-L)\right]\left(\bar{B}-E\left[\int_0^T a_t^{\bar{\Lambda}_1}B_t dt\right]\right)-\theta E[U_1(P^{\bar{\Lambda}_1}-L)]\left(\bar{B}-E\left[\int_0^T a_t^* B_t dt\right]\right) \\ & -\theta E\left[\frac{1}{\bar{\Lambda}_1}U_2(-P^{\bar{\Lambda}_1})+U_1(P^{\bar{\Lambda}_1}-L)\right]E\left[\int_0^T B_t(a_t^*-a_t^{\bar{\Lambda}_1})dt\right] \\ =\ & -\theta E\left[\frac{1}{\bar{\Lambda}_1}U_2(-P^{\bar{\Lambda}_1})\right]E\left[\int_0^T B_t(a_t^*-a_t^{\bar{\Lambda}_1})dt\right]. \end{aligned} \tag{53}$$

Here, $E\left[U_2(-P^{\bar{\Lambda}_1})\right]\leq 0$ because $P^{\bar{\Lambda}_1}\geq 0$ for each $L\in R_L$. Corollary 1 shows that $a_t^*-a_t^{\bar{\Lambda}_1}\geq 0$ for every $t\in[0,T]$ when $U_2(0)=0$. Now we can get $\mathbb{U}_1(\bar{\Lambda}_1)-R\geq 0$ from (53). Because $\mathbb{U}_1(\Lambda_1)$ is an increasing function of $\Lambda_1$, $\hat{\Lambda}_1<\bar{\Lambda}_1$. Theorem 4 shows that $\hat{\Lambda}_1>\underline{\Lambda_1}$, so we can conclude the unique $\hat{\Lambda}_1$ is located in the interval $(\underline{\Lambda_1},\bar{\Lambda}_1)$. □

## References


Federal Emergency Management Agency, 2022. National Flood Insurance Program, Summary of Coverage. https://agents.floodsmart.gov/sites/default/files/fema_nfip-summary-of-coverage-brochure-11-2022.pdf.

Flex Insurance Company Residential Strata Plan Insurance Product Disclosure Statement and Policy Wording. https://www.flexinsurance.com.au/assets/Documents/FlexInsurance_ResiStrataPDS_Jan2022_online.pdf.

Albrecher, H., Asmussen, S., 2006. Ruin probabilities and aggregate claims distributions for shot noise Cox processes. Scandinavian Actuarial Journal 2006, 86–110.

Beard, R.E., Pentikäinen, T., Pesonen, E., 1984. Risk Theory. Chapman and Hall Ltd., London.

Bolton, P., Dewatripont, M., 2005. Contract Theory. The MIT Press.

Bühlmann, H., 1970. Mathematical Methods in Risk Theory. Springer-Verlag, Berlin, Heidelberg.

Cadenillas, A., Cvitanić, J., Zapatero, F., 2007. Optimal risk-sharing with effort and project choice. Journal of Economic Theory 133, 403–440.

Cvitanić, J., Zhang, J., 2013. Contract Theory in Continuous-Time Model. Springer.

Dassios, A., Embrechts, P., 1989. Martingales and insurance risk. Communications in Statistics. Stochastic Models 5, 181–217.

Dassios, A., Jang, J., 2003. Pricing of catastrophe reinsurance and derivatives using the Cox process with shot noise intensity. Finance and Stochastics 7, 73–95.

Dassios, A., Jang, J., Zhao, H., 2015. A risk model with renewal shot-noise Cox process. Insurance. Mathematics & Economics 65, 55–65.

Davis, M.H.A., 1984. Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models. Journal of the Royal Statistical Society, Series B, Methodological 46, 353–388.

Demarzo, P.M., Sannikov, Y., 2017. Learning, termination, and payout policy in dynamic incentive contracts. The Review of Economic Studies 84, 182–236.

Grandell, J., 1976. Doubly Stochastic Poisson Processes. Springer-Verlag, Berlin, Heidelberg.

Hoffmann, F., Inderst, R., Opp, M., 2021. Only time will tell: a theory of deferred compensation. The Review of Economic Studies 88, 1253–1278.

Hopenhayn, H., Jarque, A., 2010. Unobservable persistent productivity and long term contracts. Review of Economic Dynamics 13, 333–349.

Jarque, A., 2010. Repeated moral hazard with effort persistence. Journal of Economic Theory 145, 2412–2423.

Lindskog, F., McNeil, A.J., 2003. Common Poisson shock models: applications to insurance and credit risk modelling. ASTIN Bulletin 33, 209–238.

Macci, C., Torrisi, G.L., 2011. Risk processes with shot noise Cox claim number process and reserve dependent premium rate. Insurance. Mathematics & Economics 48, 134–145.

Medhi, J., 1982. Stochastic Processes. New Age International(P) Limited, Publishers.

Moore, K.S., Young, V.R., 2006. Optimal insurance in a continuous-time model. Insurance. Mathematics & Economics 39, 47–68.

Mukoyama, T., Şahin, A., 2005. Repeated moral hazard with persistence. Economic Theory 25, 831–854.

Sannikov, Y., 2008. A continuous - time version of the principal: agent problem. The Review of Economic Studies 75, 957–984.

Schmidt, T., 2014. Catastrophe insurance modeled by shot-noise processes. Risks 2, 3–24.

Williams, N., 2009. On Dynamic Principal-Agent Problems in Continuous Time. Working paper. University of Wisconsin-Madison.

Williams, N., 2015. A solvable continuous time dynamic principal–agent model. Journal of Economic Theory 159, 989–1015.

Zhu, L., 2013. Ruin probabilities for risk processes with non-stationary arrivals and subexponential claims. Insurance. Mathematics & Economics 53, 544–550.

Zou, B., Cadenillas, A., 2014. Optimal investment and risk control policies for an insurer: expected utility maximization. Insurance. Mathematics & Economics 58, 57–67.

Zou, B., Cadenillas, A., 2017. Optimal investment and liability ratio policies in a multidimensional regime switching model. Risks 5 (6).