




OPTIMAL PROPORTIONAL REINSURANCE TO MAXIMIZE AN INSURER'S EXPONENTIAL UTILITY UNDER UNOBSERVABLE DRIFT

XIAOQING LIANG ,* *Hebei University of Technology*
VIRGINIA R. YOUNG,** *University of Michigan*

Abstract

We study an optimal reinsurance problem for a diffusion model, in which the drift of the claim follows an Ornstein–Uhlenbeck process. The aim of the insurer is to maximize the expected exponential utility of its terminal wealth. We consider two cases: full information and partial information. Full information occurs when the insurer directly observes the drift; partial information occurs when the insurer observes only its claims. By applying stochastic control and by solving the corresponding Hamilton–Jacobi–Bellman equations, we find the value function and the optimal reinsurance strategy under both full and partial information. We determine a relationship between the value function and reinsurance strategy under full information with the value function and reinsurance strategy under partial information.

Keywords: Exponential utility; mean-reverting drift; incomplete information; optimal reinsurance

2020 Mathematics Subject Classification: Primary 91G05; 93E20; 93E11
Secondary 49J40; 60G15

1. Introduction

Research on applying stochastic control theory to analyze insurance problems has long attracted a great deal of interest among actuaries because optimal control provides both theoretical and practical solutions to optimization problems in insurance. There are many research papers studying optimal insurance problems involving reinsurance and investment, which help the insurer to increase profits and to reduce the claim risk. For example, [8] considered two optimal investment problems for an insurer under a diffusion risk model, namely, maximizing expected exponential utility of terminal wealth and minimizing the probability of ruin. The probability of ruin for a diffusion risk model was minimized via reinsurance and investment in [26]; [28] studied a similar problem for a compound Poisson risk model; [17] incorporated proportional and excess-of-loss reinsurance into a jump-diffusion model with investment to maximize the expected utility of terminal wealth; [22] explored an optimal proportional reinsurance and investment model in a stock market driven by an Ornstein–Uhlenbeck process; and [20] studied an optimal investment and reinsurance problem under the mean-variance criterion. See also [3, 15, 29, 32], to name just a few.

Received 12 October 2021; revision received 29 August 2022.

* Postal address: Department of Statistics, School of Sciences, Hebei University of Technology, Tianjin 300401, P. R. China. Email: liangxiaoqing115@hotmail.com

** Postal address: Department of Mathematics, University of Michigan, Ann Arbor, Michigan, 48109. Email: vryoung@umich.edu

© The Author(s), 2023. Published by Cambridge University Press on behalf of Applied Probability Trust.

In most of above-referenced research, the insurer has complete knowledge of the model and of the values of the processes in that model, i.e. the problems are considered under full information. However, in reality, the insurance company generally only has partial information, which generally assumes knowledge of the model but *not* of the values of all the processes in that model. Portfolio optimization problems with unobservable information has been an active topic in mathematical finance. However, results related to insurance models are relatively few. Research on portfolio optimization generally assumes the drift of the traded stock unobservable; see, for example, [6, 7]. Also, [18, 19] investigated a similar problem by using the martingale-duality approach; [2] assumed that the dynamic of the stock price follows a geometric Brownian motion with Poisson jumps, in which the jump intensity is unobservable; and [21] considered an optimal reinsurance and investment problem by maximizing the expected exponential utility of the insurer's terminal wealth in a Black–Scholes financial market. The claim process is a compound Poisson process in which the claim intensity and the jump-size distribution depend on the state of a non-observable Markov chain. See also [1, 4, 5, 14, 16, 25, 30].

In this paper we find the optimal reinsurance strategy for an insurer to maximize the expected exponential utility of its terminal wealth. We use a diffusion model to describe the dynamics of the claim, in which the drift of the claims follows a mean-reverting Ornstein–Uhlenbeck process. We consider two cases: full information and partial information. Full information occurs when the insurer directly observes the drift; partial information occurs when the insurer observes only its claims. We use the filtering technique to transfer the unobservable problem into an observable one; then, by applying the dynamic programming approach, we derive explicit expressions for the value function and the corresponding optimal reinsurance strategy. We also determine a relationship between the value function and reinsurance strategy under full information with the value function and reinsurance strategy under partial information. Finally, we present numerical examples to illustrate possible outcomes of the model.

The rest of this paper is organized as follows. In Section 2, we describe the insurance model in both full and partial information frameworks. In Section 3, we use a dynamic Hamilton–Jacobi–Bellman (HJB) equation approach to find the value function and optimal reinsurance strategy under complete information. In Section 4, we consider the problem with unobservable information. Moreover, we discuss how to handle a constraint on the proportion reinsured, and we provide numerical simulation results to compute the probability that the reinsurance proportion lies outside of the interval $[0, 1]$ for some values of the parameters. Finally, in Section 5 we compare the value function and the optimal reinsurance strategy under partial observation with the ones under full observation. We also provide numerical examples to show the difference between the two optimal reinsurance proportions at the end of this section.

2. Problem formulation

In this section we describe the claim process of the insurer, and we formulate the problem of maximizing the insurer's expected exponential utility of terminal wealth. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space that supports two correlated standard Brownian motions W_1 and W_2 , with constant coefficient of correlation $\rho \in [-1, 1]$. As in [26], assume that the claim process $S = \{S_t\}_{0 \leq t \leq T}$ follows Brownian motion with drift, in which S_t equals the cumulative claims paid by the insurer during the time interval $[0, t]$. However, unlike [26], the drift follows a random process. Specifically, $dS_t = \mu_t dt - \sigma dW_{1,t}$, in which $S_0 = 0$ and σ is a positive constant, and

$$d\mu_t = -a(\mu_t - \bar{\mu})dt + b dW_{2,t}, \quad (2.1)$$

in which $a, b,$ and $\bar{\mu}$ are positive constants, i.e. $\{\mu_t\}_{0 \leq t \leq T}$ follows a mean-reverting Ornstein–Uhlenbeck process. We loosely interpret $\bar{\mu}$ as a long-run value of μ_t , and a measures the ‘speed’ at which μ_t moves towards $\bar{\mu}$. This model was explored in [12] as the diffusion approximation of a Cox process with a shot-noise process as the claim intensity.

The insurer collects premium at the constant rate $c = (1 + \theta)\bar{\mu}$, in which $\theta \geq 0$ is the constant proportional risk loading. The insurer is able to purchase proportional reinsurance with constant proportional risk loading $\eta \geq \theta$. Let q_t denote the proportional amount of business retained at time $t \in [0, T]$; thus, the controlled surplus $X = \{X_t\}_{0 \leq t \leq T}$ follows the stochastic differential equation (SDE)

$$\begin{aligned} dX_t &= (cdt - dS_t) - (1 - q_t)(1 + \eta)\bar{\mu} dt + (1 - q_t)dS_t \\ &= (q_t((1 + \eta)\bar{\mu} - \mu_t) - (\eta - \theta)\bar{\mu})dt + q_t\sigma dW_{1,t}, \end{aligned} \tag{2.2}$$

with $X_0 = x$. If $0 \leq q_t \leq 1$, then q_t is the usual proportional reinsurance. In Examples 4.1 and 4.2 we calculate the probability that the optimal q_t lies outside the interval $[0, 1]$.

The insurer chooses a retention strategy $q = \{q_t\}_{0 \leq t \leq T}$ based on the available information, and we consider two cases in this paper:

Full information In this case, the insurer observes both $\{S_t\}$ and $\{\mu_t\}$. A retention strategy q is *admissible* in this case if (i) q is adapted to the filtration $\mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T}$, in which $\mathcal{F}_t = \sigma(S_s, \mu_s : 0 \leq s \leq t)$ for all $t \in [0, T]$; (ii) q is conditional L^2 -integrable, i.e. $\mathbb{E}[\int_t^T q_u^2 du | \mathcal{F}_t] < \infty$ for any $0 \leq t \leq T$; and (iii) the SDE (2.2) has a pathwise unique solution $\{X_t^q\}_{t \in [0, T]}$. Let \mathcal{A}^f denote the set of admissible strategies in the full information case.

Partial information In this case, the insurer observes only $\{S_t\}$ and does not know the drift of its claim process, although the insurer knows the conditional expectation and variance of μ_0 . A retention strategy q is *admissible* in this case if (i) q is adapted to the filtration $\mathbb{G} = \{\mathcal{G}_t\}_{0 \leq t \leq T}$, in which $\mathcal{G}_t = \sigma(S_s : 0 \leq s \leq t)$ for all $t \in [0, T]$; (ii) if $\mathbb{E}[\int_t^T q_u^2 du | \mathcal{G}_t] < \infty$ for any $0 \leq t \leq T$; and (iii) the SDE of the controlled surplus under \mathcal{G}_t has a pathwise unique solution. Let \mathcal{A}^p denote the set of admissible strategies in the partial information case. Note that $\mathcal{G}_t \subset \mathcal{F}_t$ for all $t \in [0, T]$, i.e. \mathbb{G} is a subfiltration of \mathbb{F} ; we also assume that \mathbb{F} and \mathbb{G} are augmented to satisfy the usual conditions of completeness and right continuity.

In both cases, the insurer chooses q to maximize the expectation of exponential utility of wealth at time T . Let V^f denote the maximum expected exponential utility of terminal wealth under full information, i.e.

$$V^f(t, x, \mu) = \sup_{q \in \mathcal{A}^f} \mathbb{E}(-e^{-\gamma X_T} | X_t = x, \mu_t = \mu), \tag{2.3}$$

in which $\gamma > 0$ is the (constant) coefficient of absolute risk aversion.

For the partial information case, as in [4, 6, 7], we first project the drift process μ_t onto the observable filtration \mathcal{G} , in order to reduce the partially observable problem to an equivalent problem with full information. Define $m_t = \mathbb{E}(\mu_t | \mathcal{G}_t)$, $0 \leq t \leq T$; then, [23, Theorem 10.3] shows us that $\{m_t\}_{0 \leq t \leq T}$ follows the SDE

$$dm_t = -a(m_t - \bar{\mu})dt + \left(\rho b - \frac{\text{Var}(\mu_t | \mathcal{G}_t)}{\sigma} \right) d\bar{W}_{1,t}, \tag{2.4}$$

in which $\bar{W}_1 = \{\bar{W}_{1,t}\}_{0 \leq t \leq T}$ is the so-called *innovations process* given by

$$\bar{W}_{1,t} = W_{1,t} + \frac{1}{\sigma} \int_0^t (m_s - \mu_s) ds, \tag{2.5}$$

and \bar{W}_1 is a (\mathbb{P}, \mathbb{G}) -standard Brownian motion. In other words, $\{m_t\}_{0 \leq t \leq T}$ follows a \mathbb{G} -Ornstein–Uhlenbeck process with non-constant volatility.

If we define $v(t) = \text{Var}(\mu_t | \mathcal{G}_t)$ for $t \in [0, T]$, then $v = v(t)$ satisfies the Riccati equation

$$\frac{dv(t)}{dt} = b^2 - 2av(t) - \left(\frac{\sigma \rho b - v(t)}{\sigma} \right)^2, \tag{2.6}$$

with initial value $v(0) = \text{Var}(\mu_0 | \mathcal{G}_0)$. See Appendix A for a derivation of (2.6) and the following solution:

$$v(t) = \sigma^2 \Delta \cdot \frac{Re^{2\Delta \cdot t} - 1}{Re^{2\Delta \cdot t} + 1} - \sigma(\sigma a - \rho b),$$

in which

$$\Delta = \frac{1}{\sigma} \sqrt{(\sigma a - \rho b)^2 + b^2(1 - \rho^2)}, \tag{2.7}$$

and

$$R = \frac{\sigma^2 \Delta + \sigma(\sigma a - \rho b) + v(0)}{\sigma^2 \Delta - \sigma(\sigma a - \rho b) - v(0)}. \tag{2.8}$$

Moreover, by substituting for W_1 in terms of \bar{W}_1 in (2.2), we obtain that X follows the dynamics

$$dX_t = (q_t((1 + \eta)\bar{\mu} - m_t) - (\eta - \theta)\bar{\mu})dt + q_t \sigma d\bar{W}_{1,t},$$

which is \mathbb{G} -adapted in the partial information case because q is \mathbb{G} -adapted in that case. Let V^p denote the maximum expected exponential utility of terminal wealth under partial information, i.e.

$$V^p(t, x, m) = \sup_{q \in \mathcal{A}^p} \mathbb{E}(-e^{-\gamma X_T} | X_t = x, m_t = m). \tag{2.9}$$

3. Full information case

We begin by stating a relevant verification theorem without proof because the proof is standard in the actuarial and financial mathematics literature; see, for example, [26, Theorem 2.1].

Theorem 3.1. *Suppose $v^f \in C^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$ takes values in \mathbb{R}^- , is non-decreasing and concave in x , and satisfies the HJB equation*

$$0 = v_t^f - (\eta - \theta)\bar{\mu}v_x^f - a(\mu - \bar{\mu})v_\mu^f + \frac{1}{2}b^2v_{\mu\mu}^f + \sup_q [q((1 + \eta)\bar{\mu} - \mu)v_x^f + q\sigma\rho bv_{x\mu}^f + \frac{1}{2}q^2\sigma^2v_{xx}^f], \tag{3.1}$$

with terminal condition $v^f(T, x, \mu) = -e^{-\gamma x}$. The maximizer of (3.1) is

$$q^*(t, x, \mu) = - \frac{((1 + \eta)\bar{\mu} - \mu)v_x^f(t, x, \mu) + \sigma\rho bv_{x\mu}^f(t, x, \mu)}{\sigma^2v_{xx}^f(t, x, \mu)}.$$

If the retention strategy q^f given in feedback form by $q_t^f = q^*(t, X_t^*, \mu_t)$ for all $0 \leq t \leq T$ is admissible, then the value function V^f defined by (2.3) equals v^f . Here, X_t^* is the optimally controlled surplus at time t .

In the following theorem, we solve the HJB equation in (3.1) with boundary condition $v^f(T, x, \mu) = -e^{-\gamma x}$. Theorem 3.1 then allows us to deduce that the solution equals the value function V^f in (2.3).

Theorem 3.2. *The maximum expected exponential utility of terminal wealth under full information is $V^f(t, x, \mu) = -e^{-\gamma x} \exp\{A(t)\mu^2 + B(t)\mu + C(t)\}$, in which*

$$A(t) = -\frac{e^{2\Delta \cdot (T-t)} - 1}{\alpha_1 e^{2\Delta \cdot (T-t)} + \alpha_2}, \tag{3.2}$$

$$B(t) = \frac{2\bar{\mu}}{\Delta} \cdot \frac{e^{\Delta \cdot (T-t)} - 1}{\alpha_1 e^{2\Delta \cdot (T-t)} + \alpha_2} \left\{ (1 + \eta)\Delta(e^{\Delta \cdot (T-t)} + 1) + \eta a(e^{\Delta \cdot (T-t)} - 1) \right\}, \tag{3.3}$$

$$C(t) = \int_t^T \left\{ \frac{b^2(1 - \rho^2)}{2} B^2(s) + \left(a - \frac{(1 + \eta)b\rho}{\sigma} \right) \bar{\mu} B(s) + b^2 A(s) \right\} ds + \left\{ (\eta - \theta)\bar{\mu}\gamma - \frac{(1 + \eta)^2 \bar{\mu}^2}{2\sigma^2} \right\} (T - t), \tag{3.4}$$

with Δ given in (2.7), and we define $\alpha_1 > 0$ and $\alpha_2 > 0$ as $\alpha_1 = 2\sigma(\sigma\Delta + (\sigma a - \rho b))$, $\alpha_2 = 2\sigma(\sigma\Delta - (\sigma a - \rho b))$. Moreover, the optimal retention strategy q^f is given in feedback form by

$$q_t^f = \frac{1}{\sigma^2 \gamma} \{ (2\sigma\rho b A(t) - 1)\mu_t + \sigma\rho b B(t) + (1 + \eta)\bar{\mu} \}. \tag{3.5}$$

Proof. See Appendix B for a proof of this theorem. □

Remark 3.1. We derive explicit expressions for A and B in (3.2) and (3.3), respectively, because q^f in (3.5) relies on A and B . However, for the sake of space, we do not present an explicit expression for C .

Remark 3.2. Because $A(t) \leq 0$ for all $0 \leq t \leq T$, as μ_t increases the proportion retained, namely q_t^f , decreases. This monotonicity makes sense because, as the drift of claims increases, we expect the insurance company to retain less of its risk, especially given a fixed premium rate $(1 + \eta)\bar{\mu}$. Also, note that, when $\mu_t = 0$, $q_t^f > 0$ because $B(t) \geq 0$ for all $0 \leq t \leq T$.

In the following corollary we show how A and B change with time. We omit the calculations because they are straightforward from (3.2) and (3.3).

Corollary 3.1. *The derivative of $A(t)$ is*

$$A'(t) = \frac{8\sigma^2 \Delta^2 e^{2\Delta \cdot (T-t)}}{(\alpha_1 e^{2\Delta \cdot (T-t)} + \alpha_2)^2},$$

which is positive for $0 \leq t \leq T$, and the derivative of $B(t)$ is

$$B'(t) = -\frac{4\bar{\mu} e^{\Delta \cdot (T-t)}}{(\alpha_1 e^{2\Delta \cdot (T-t)} + \alpha_2)^2} \times \left\{ (1 + \eta)4\sigma^2 \Delta^2 e^{\Delta \cdot (T-t)} + \eta a(e^{\Delta \cdot (T-t)} - 1)(\alpha_1 e^{\Delta \cdot (T-t)} + \alpha_2) \right\},$$

which is negative for $0 \leq t \leq T$. Thus, the slope of q_t^f as a linear function of μ_t , namely $(1/(\sigma^2\gamma))(2\sigma\rho bA(t) - 1)$, becomes less negative over time (specifically, it increases to $-1/(\sigma^2\gamma)$ as t increases to T), and the intercept, namely $(1/(\sigma^2\gamma))(\sigma\rho bB(t) + (1 + \eta)\bar{\mu})$, becomes less positive over time (specifically, it decreases to $(1 + \eta)\bar{\mu}/(\sigma^2\gamma)$ as t increases to T).

Remark 3.3. Corollary 3.1 shows us that, over time, the proportion of retained risk, as a linear function of μ_t , flattens. This flattening is consistent with the risk aversion of the insurer. As time approaches the horizon T , the insurer will not wish to change its retention as much (as a function of μ_t) as when further from the horizon. Intuitively, the closer time is to the horizon, the less time the insurer has to maximize its expected utility and, therefore, the insurer reacts less strongly to changes in the drift of the surplus. We see a similar phenomenon in the optimal investment strategy of [6]; namely, the closer time is to T , the less the investor changes their investment in reaction to changes in the drift of the risky asset.

4. Partial information case

In this section we analyze the problem under partial information. The corresponding verification theorem and its solution parallel the results in Section 3.

Theorem 4.1. Suppose $v^p \in C^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$ takes values in \mathbb{R}^- , is non-decreasing and concave in x , and satisfies the HJB equation

$$0 = v_t^p - (\eta - \theta)\bar{\mu}v_x^p - a(m - \bar{\mu})v_m^p + \frac{1}{2}\left(\rho b - \frac{v(t)}{\sigma}\right)^2 v_{mm}^p + \sup_q \left[q((1 + \eta)\bar{\mu} - m)v_x^p + q\sigma\left(\rho b - \frac{v(t)}{\sigma}\right)v_{xm}^p + \frac{1}{2}q^2\sigma^2v_{xx}^p \right], \quad (4.1)$$

with terminal condition $v^p(T, x, \mu) = -e^{-\gamma x}$. The maximizer of (4.1) is

$$q^*(t, x, m) = -\frac{((1 + \eta)\bar{\mu} - m)v_x^p(t, x, m) + \sigma\rho b v_{xm}^p(t, x, m)}{\sigma^2 v_{xx}^p(t, x, m)}.$$

If the retention strategy q^p given in feedback form by $q_t^p = q^*(t, X_t^*, m_t)$ for all $0 \leq t \leq T$ is admissible, then the value function V^p defined by (2.9) equals v^p . Here, X_t^* is the optimally controlled surplus at time t .

In the following theorem we solve the HJB equation in (4.1) with boundary condition $v^p(T, x, \mu) = -e^{-\gamma x}$. Theorem 4.1 then allows us to deduce that the solution equals the value function V^p in (2.9).

Theorem 4.2. The maximum expected exponential utility of terminal wealth under partial information is $V^p(t, x, m) = -e^{-\gamma x} \exp\{\widehat{A}(t)m^2 + \widehat{B}(t)m + \widehat{C}(t)\}$, in which

$$\widehat{A}(t) = -\frac{1}{4\sigma^2\Delta} \cdot \frac{(\text{Re}^{2\Delta \cdot t} + 1)(e^{2\Delta \cdot (T-t)} - 1)}{\text{Re}^{2\Delta \cdot T} + 1}, \quad (4.2)$$

$$\begin{aligned} \widehat{B}(t) &= \frac{\bar{\mu}}{2\sigma^2\Delta^2} \cdot \frac{(\text{Re}^{2\Delta \cdot t} + 1)(e^{\Delta \cdot (T-t)} - 1)}{\text{Re}^{2\Delta \cdot T} + 1} \\ &\quad \times \left\{ (1 + \eta)\Delta(e^{\Delta \cdot (T-t)} + 1) + \eta a(e^{\Delta \cdot (T-t)} - 1) \right\}, \end{aligned} \quad (4.3)$$

$$\widehat{C}(t) = \int_t^T \left\{ \left(a - (1 + \eta) \frac{\sigma \rho b - v(t)}{\sigma^2} \right) \bar{\mu} \widehat{B}(s) + \left(\rho b - \frac{v(s)}{\sigma} \right)^2 \widehat{A}(s) \right\} ds + \left\{ (\eta - \theta) \bar{\mu} \gamma - \frac{(1 + \eta)^2 \bar{\mu}^2}{2\sigma^2} \right\} (T - t), \tag{4.4}$$

in which R is given in (2.8).

Moreover, the optimal retention strategy q^p is given in feedback form by

$$q_t^p = \frac{1}{\sigma^2 \gamma} \{ (2(\sigma \rho b - v(t)) \widehat{A}(t) - 1) m_t + (\sigma \rho b - v(t)) \widehat{B}(t) + (1 + \eta) \bar{\mu} \}. \tag{4.5}$$

Proof. See Appendix C for a proof of this theorem. □

Remark 4.1. As in Section 3, we derive explicit expressions for \widehat{A} and \widehat{B} in (4.2) and (4.3), respectively, because q^p in (4.5) relies on \widehat{A} and \widehat{B} . However, for the sake of space, we do not present an explicit expression for \widehat{C} .

Because $\sigma \rho b - v(t)$ can be negative, it is not clear whether the slope of q_t^p as a function of m_t is negative, as in the case for q_t^p as a function of μ_t . The following corollary tells us that the slope of q_t^p is, indeed, negative.

Corollary 4.1. *The slope of q_t^p as a linear function of m_t is negative, i.e.*

$$\frac{1}{\sigma^2 \gamma} (2(\sigma \rho b - v(t)) \widehat{A}(t) - 1) < 0, \tag{4.6}$$

for all $0 \leq t \leq T$.

Proof. By substituting for \widehat{A} and v from (4.2) and (2.6), respectively, and by simplifying the result, we can show that inequality (4.6) is equivalent to

$$\frac{1}{2\Delta} \cdot \frac{((\Delta + a)e^{2\Delta \cdot (T-t)} + (\Delta - a)(Re^{2\Delta \cdot t} + 1))}{Re^{2\Delta \cdot T} + 1} > 0,$$

or

$$\frac{Re^{2\Delta \cdot t} + 1}{Re^{2\Delta \cdot T} + 1} > 0. \tag{4.7}$$

By substituting for R from (2.8), we find that inequality (4.7) is equivalent to

$$\frac{\sigma^2 \Delta (e^{2\Delta \cdot t} + 1) + \sigma(\sigma a - \rho b)(e^{2\Delta \cdot t} - 1) + v(0)(e^{2\Delta \cdot t} - 1)}{\sigma^2 \Delta (e^{2\Delta \cdot T} + 1) + \sigma(\sigma a - \rho b)(e^{2\Delta \cdot T} - 1) + v(0)(e^{2\Delta \cdot T} - 1)} > 0,$$

which is true because $\sigma \Delta > \pm(\sigma a - \rho b)$. □

As in Corollary 3.1, we show how \widehat{A} and \widehat{B} change with time in the following corollary.

Corollary 4.2. *The derivative of $\widehat{A}(t)$ is*

$$\widehat{A}'(t) = \frac{1}{2\sigma^2} \cdot \frac{Re^{2\Delta \cdot t} + e^{2\Delta \cdot (T-t)}}{Re^{2\Delta \cdot T} + 1}$$

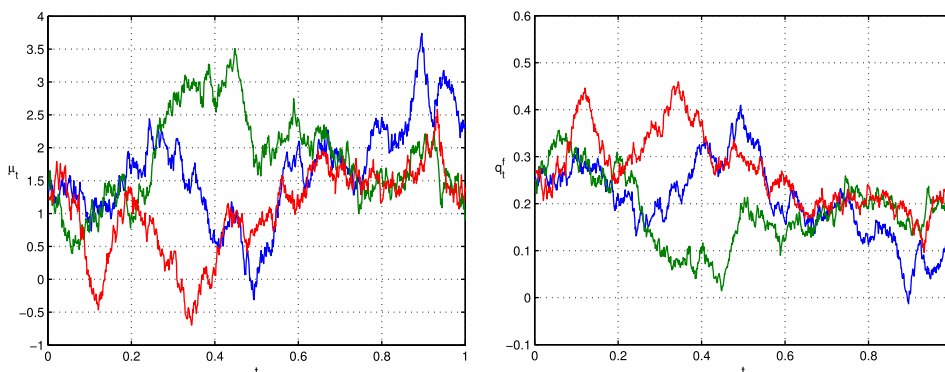


FIGURE 1. Left: Sample paths of μ_t when T is small. Right: Sample paths of q_t^f when T is small.

for $0 \leq t \leq T$, and the derivative of $\widehat{B}(t)$ is

$$\widehat{B}'(t) = -\frac{\bar{\mu}}{\sigma^2 \Delta} \cdot \frac{(1 + \eta)\Delta(\text{Re}^{2\Delta \cdot t} + e^{2\Delta \cdot (T-t)}) + \eta a(\text{Re}^{2\Delta \cdot t} + e^{\Delta \cdot (T-t)})(e^{\Delta \cdot (T-t)} - 1)}{\text{Re}^{2\Delta \cdot T} + 1},$$

for $0 \leq t \leq T$.

Remark 4.2. Unlike Corollary 3.1, we cannot assert that $\widehat{A}'(t) \geq 0$ and $\widehat{B}'(t) \leq 0$ because these inequalities might not hold if R is negative enough, which occurs, for example, when $v(0)$ is relatively large. On the other hand, if $R > 0$, then it is clear that $\widehat{A}'(t) \geq 0$ and $\widehat{B}'(t) \leq 0$.

In the following, we present two numerical examples to further explore the reinsurance strategies in both full and partial information cases. For each example, we set $a = 1$, $b = 2.5$, $\theta = 0.4$, $\eta = 0.8$, $\rho = 0.4$, $\sigma = 3$, $\gamma = 1.2$, $\bar{\mu} = 2$, and $\mu_0 \sim N(0, 1)$.

Example 4.1. In Figure 1 we set $T = 1$ and used a Monte Carlo approach to simulate the sample paths of μ_t and q_t^f under full information. We present three sample paths in these two figures.

To compute the probability that $q_t^f < 0$, $q_t^f > 1$, or $\mu_t < 0$, we worked with 3000 sample paths, each discretized into 2000 time intervals. Let i denote a sample path, and let j denote a time instance. For $i = 1, 2, \dots, 3000$ and $j = 1, 2, \dots, 2000$, we counted the number of points for which $q^f(i, j) < 0$, and computed the proportion of that number divided by the total number of observations, namely, 3000×2000 . That proportion is our estimate of the probability of $q_t^f < 0$. Similarly, we estimated the probabilities of $q_t^f > 1$ and $\mu_t < 0$. For our parameter values, we computed that the probability of $q_t^f < 0$ equals 0.0682, the probability of $q_t^f > 1$ equals 0, and the probability of $\mu_t < 0$ equals 0.1081.

In Figure 2 we set $T = 50$, and we present one sample path of the mean-reverting process μ_t and the reinsurance proportion q_t^f . In this case, we estimated, via 3000 discretized sample paths, that the probability of $q_t^f < 0$ equals 0.1461, and the probability of $\mu_t < 0$ equals 0.1323. Thus, as the terminal time T increases, the probabilities of $q_t^f < 0$ and $\mu_t < 0$ also increase.

As an aside, if b is relatively small and if T is large, say, 50, then the value of μ_t is close to $\bar{\mu}$ for a good portion of the interval $[0, T]$; thus, the probability of $\mu_t < 0$ is small. If $T = 1$, then because μ_0 might be very different from $\bar{\mu}$, we cannot say that the probability of $\mu_t < 0$ is small.

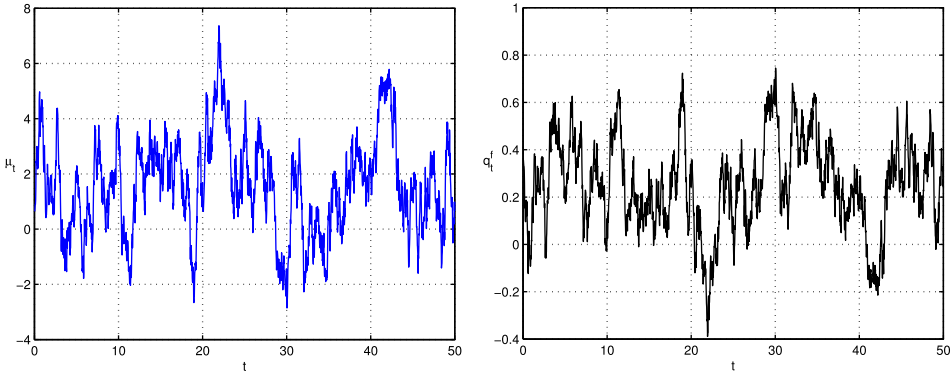


FIGURE 2. Left: Sample path of μ_t when T is large. Right: Sample path of q_t^f when T is large.

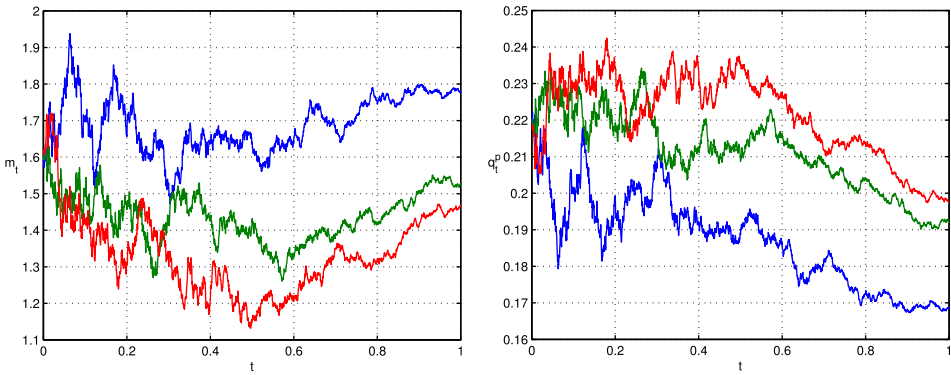


FIGURE 3. Left: Sample paths of m_t when T is small. Right: Sample paths of q_t^p when T is small.

We also observed (in work not shown here) that, as $\bar{\mu}$ decreases, the probability of $q_t^f < 0$ increases, and as $\bar{\mu}$ increases, the probability of $q_t^f > 1$ increases. Moreover, as $\bar{\mu}$ increases, then the probability of $\mu_t < 0$ decreases.

Example 4.2. In Figure 3 we simulate the sample paths of m_t and q_t^p under partial information. We set $T = 1$, $m_0 = 1.6$, and $v(0) = 0.5$. As in Example 4.1, we plot three sample paths of q_t^p and m_t . In this example, among 3000 discretized sample paths, we find that the optimal reinsurance proportion q_t^p always lies in $[0, 1]$, and the filtered drift process m_t is always positive.

In Figure 4 we set $T = 50$ and plot one sample path. As in the case for $T = 1$, among 3000 discretized sample paths, the optimal reinsurance proportion q_t^p always lies in $[0, 1]$, and the filtered drift process m_t is always positive.

For the optimization problem with the reinsurance constraint, that is, $0 \leq q_t \leq 1$ for all $0 \leq t \leq T$, as in the discussion before, the value function under full information (or partial information) satisfies the HJB equation in (3.1) (or (4.1)) for V^f (or V^p), except that q is constrained to lie in $[0, 1]$. Indeed, from [10, 31], we know that a value function for a constrained problem is the unique viscosity solution of its HJB equation subject to the constraint. However, due to

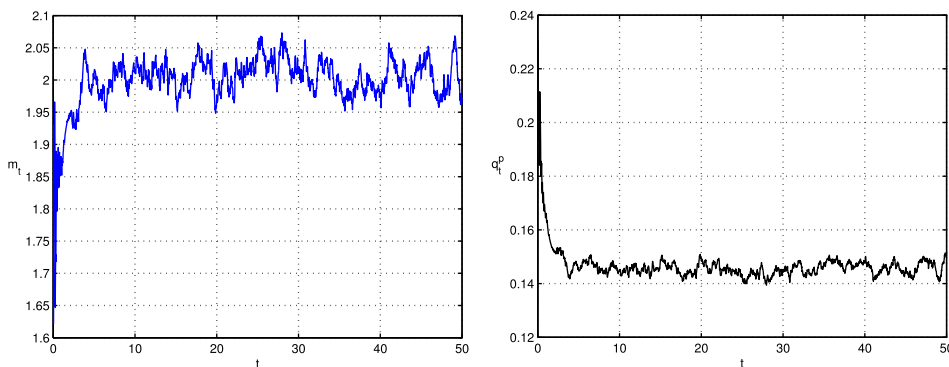


FIGURE 4. Left: Sample path of m_t when T is large. Right: Sample path of q_t^p when T is large.

the reinsurance constraint, we cannot derive further explicit results by following the classical HJB equation approach. Motivated by the convex-duality approach, as in [11, 27], we need to introduce an auxiliary unconstrained optimization problem by modifying the original problem with an auxiliary stochastic parameter, then find the relationship between the value function of the auxiliary problem and that of the original problem. Due to the randomness of the auxiliary parameter process, the corresponding value function satisfies a stochastic HJB equation, which is a special backward stochastic partial differential equation or infinite-dimensional BSDE. The solvability of the stochastic HJB equation was studied in [24], which provided an existence and uniqueness theorem for the case in which the volatility coefficient of the state process does not contain the control variable. In the finance and insurance literature, the BSDE approach is becoming an efficient technique to solve utility maximization problems with stochastic coefficients. A stochastic Stackelberg differential game between an insurer and a reinsurer was considered in [9] by applying the BSDE approach. See also [13] for more details about the applications in actuarial and financial models with the BSDE approach. Because the convex-duality-plus-BSDE approach differs greatly from the HJB equation approach in this paper, we will leave that work for future research.

5. The relationship between V^f and V^p

In this section we adapt the technique in [7] to show the link between the value function V^f under full observation and the value function V^p under partial observation.

Theorem 5.1. *The following relationships hold among A, B, C and $\widehat{A}, \widehat{B}, \widehat{C}$:*

$$\widehat{A}(t) = \frac{A(t)}{1 - 2v(t)A(t)}, \tag{5.1}$$

$$\widehat{B}(t) = \frac{B(t)}{1 - 2v(t)A(t)}, \tag{5.2}$$

$$\widehat{C}(t) = C(t) + \frac{B^2(t)v(t)}{2(1 - 2v(t)A(t))} - \frac{1}{2} \ln(1 - 2v(t)A(t)) + \frac{N(t)v(t)}{2(1 - 2v(t)A(t))} + Q(t), \tag{5.3}$$

in which N and Q satisfy the following differential equations:

$$\begin{aligned} \frac{dN(t)}{dt} &= -\frac{1}{\sigma^2}(1 - 2\sigma\rho bA(t))^2 - 4b^2(1 - \rho^2)A(t)N(t) + \frac{2}{\sigma}(\sigma a - \rho b)N(t), \\ \frac{dQ(t)}{dt} &= -\frac{1}{2}b^2(1 - \rho^2)N(t), \end{aligned}$$

with boundary conditions $N(T) = Q(T) = 0$. Moreover, the value functions V^f and V^p satisfy

$$\mathbb{E}[V^f(t, X_t, \mu_t) | \mathcal{G}_t] = V^p(t, X_t, m_t)e^{-\hat{P}(t)}, \tag{5.4}$$

in which

$$\hat{P}(t) = \frac{N(t)v(t)}{2(1 - 2v(t)A(t))} + Q(t).$$

Finally, the optimal retention strategies q^f and q^p satisfy

$$\mathbb{E}[q_t^f | \mathcal{G}_t] = (1 - 2v(t)A(t))q_t^p + \frac{v(t)}{\sigma^2\gamma} \{2(1 + \eta)\bar{\mu}A(t) + B(t)\}.$$

Proof. The validity of (5.1)–(5.3) can be checked directly by differentiating terms on the right-hand sides and comparing them with the existing differential equations satisfied by the left-hand sides.

We next prove (5.4). Because the distribution of μ_t conditional on \mathcal{G}_t is Gaussian with mean m_t and variance $v(t)$, we have

$$\begin{aligned} \mathbb{E}[\exp(A(t)\mu_t^2 + B(t)\mu_t) | \mathcal{G}_t] &= \frac{1}{\sqrt{2\pi v(t)}} \int_{\mathbb{R}} e^{A(t)x^2 + B(t)x} \exp\left\{-\frac{(x - m_t)^2}{2v(t)}\right\} dx \\ &= \frac{1}{\sqrt{2\pi v(t)}} \int_{\mathbb{R}} \exp\left\{-\frac{1 - 2v(t)A(t)}{2v(t)}\left(x - \frac{m_t + v(t)B(t)}{1 - 2v(t)A(t)}\right)^2\right. \\ &\quad \left. + \frac{(B(t)v(t) + m_t)^2 - m_t^2(1 - 2v(t)A(t))}{2(1 - 2v(t)A(t))v(t)}\right\} dx \\ &= \frac{1}{\sqrt{1 - 2v(t)A(t)}} \exp\left\{\frac{B^2(t)v(t) + 2B(t)m_t}{2(1 - 2v(t)A(t))} + \frac{A(t)m_t^2}{1 - 2v(t)A(t)}\right\} \\ &= \exp\left\{\frac{A(t)}{1 - 2v(t)A(t)}m_t^2 + \frac{B(t)}{1 - 2v(t)A(t)}m_t + \frac{B^2(t)v(t)}{2(1 - 2v(t)A(t))}\right. \\ &\quad \left. - \frac{1}{2}\ln(1 - 2v(t)A(t))\right\}. \end{aligned}$$

Hence, we obtain

$$\mathbb{E}[\exp(A(t)\mu_t^2 + B(t)\mu_t + C(t)) | \mathcal{G}_t] = \exp(\hat{A}(t)m_t^2 + \hat{B}(t)m_t + \hat{C}(t))e^{-\hat{P}(t)}.$$

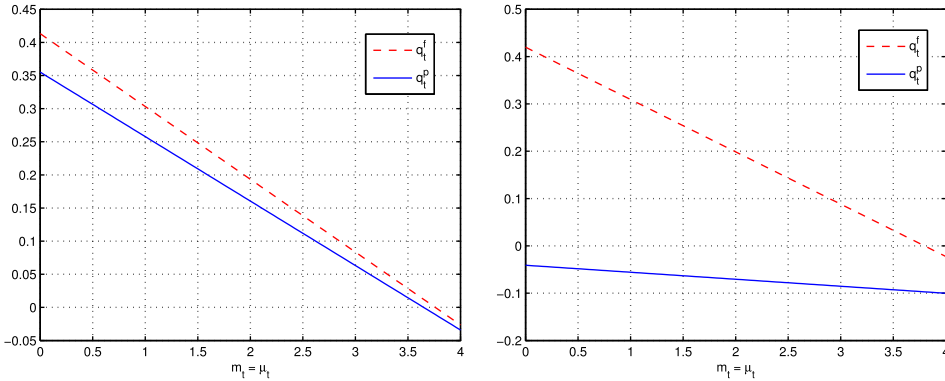


FIGURE 5. Left: Optimal proportional retention under full and partial information when $v(0)$ is small. Right: Optimal proportional retention under full and partial information when $v(0)$ is large.

Finally, from the expressions of q_t^f and q_t^p in (3.5) and (4.5), respectively, we derive

$$\begin{aligned} & (1 - 2v(t)A(t))q_t^p + \frac{v(t)}{\sigma^2\gamma} \{2(1 + \eta)\bar{\mu}A(t) + B(t)\} \\ &= \frac{1}{\sigma^2\gamma} \{2(\sigma\rho b - v(t))A(t) - (1 - 2v(t)A(t))m_t + (\sigma\rho b - v(t))B(t)\} \\ & \quad + \frac{1}{\sigma^2\gamma} (1 + \eta)(1 - 2v(t)A(t))\bar{\mu} + \frac{v(t)}{\sigma^2\gamma} \{2(1 + \eta)\bar{\mu}A(t) + B(t)\} \\ &= \frac{1}{\sigma^2\gamma} \{(2\sigma\rho bA(t) - 1)m_t + \sigma\rho bB(t) + (1 + \eta)\bar{\mu}\} = \mathbb{E}[q_t^f | \mathcal{G}_t], \end{aligned}$$

which completes our proof. □

Note that the slopes of both q_t^f and q_t^p as functions of μ_t and m_t , respectively, are negative for $0 \leq t \leq T$. Also, the vertical intercept of q_t^f thought of as a function of μ_t is positive, although the corresponding statement for the vertical intercept of q_t^p is not necessarily true; see, for example, the right panel of Figure 5. A natural question is how these slopes and intercepts compare to each other, and the following corollary answers this query.

Corollary 5.1. For all $0 \leq t \leq T$,

$$\frac{1}{\sigma^2\gamma} \{2\sigma\rho bA(t) - 1\} \leq \frac{1}{\sigma^2\gamma} \{2(\sigma\rho b - v(t))\widehat{A}(t) - 1\} < 0 \tag{5.5}$$

and

$$\frac{1}{\sigma^2\gamma} \{\sigma\rho bB(t) + (1 + \eta)\bar{\mu}\} > \frac{1}{\sigma^2\gamma} \{(\sigma\rho b - v(t))\widehat{B}(t) + (1 + \eta)\bar{\mu}\}, \tag{5.6}$$

with strict inequalities when $0 \leq t < T$.

Proof. If $t = T$, then the first inequality in (5.5) is an equality because $A(T) = \widehat{A}(T) = 0$. For $0 \leq t < T$, the first inequality in (5.5) holds strictly if and only if the following string of

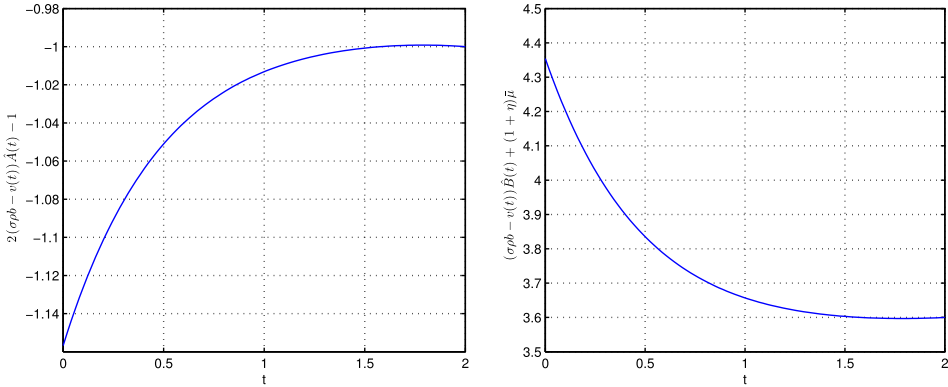


FIGURE 6. Left: The slope of q_t^p divided by the ratio $1/(\sigma^2\gamma)$ when $v(0)$ is small. Right: The intercept of q_t^p divided by the ratio $1/(\sigma^2\gamma)$ when $v(0)$ is small.

implications holds:

$$2\sigma\rho bA(t) < 2(\sigma\rho b - v(t))\widehat{A}(t) \iff \sigma\rho bA(t) < (\sigma\rho b - v(t)) \frac{A(t)}{1 - 2v(t)A(t)}$$

$$\iff^{A(t)<0} \sigma\rho b > \frac{\sigma\rho b - v(t)}{1 - 2v(t)A(t)},$$

which is true because $A(t) < 0$ and $v(t) > 0$ for all $0 \leq t < T$. The proof of (5.6) is similar (using $B(t) > 0$ when $0 \leq t < T$), so we omit it. □

It is intuitively pleasing that q_t^p reacts less strongly to changes in m_t than q_t^f reacts to μ_t . Indeed, in the partial information case, the risk-averse insurer has less information and, therefore, is more cautious in changing the proportion of retained risk. Similarly, inequality (5.6) implies that the insurer retains less risk when $m_t = 0$ in the partial information case than when $\mu_t = 0$ in the full information case. See Figure 5 for an illustration of Corollary 5.1.

In the following, we present three numerical examples to further explore the difference between q_t^f and q_t^p . For each example, we choose $b = 2.5$, $\theta = 0.4$, $\eta = 0.8$, $\rho = 0.4$, $\sigma = 3$, $\gamma = 1.2$, $\bar{\mu} = 2$, and $T = 2$.

Example 5.1. In the left panel of Figure 5, we set $a = 1$ and $v(0) = 0.5$. We plot the graphs of q_t^f and q_t^p at time $t = 0.5$ as functions of μ_t and m_t by assuming that $m_t = \mu_t$ at this time. We observe that both q_t^f and q_t^p are linear functions of $\mu_t = m_t$, as we expect from Theorems 3.2 and 4.2, but the slope of q_t^f is steeper than that of q_t^p , as we expect from Corollary 5.1. From our algebraic work, we also note that $\lim_{t \rightarrow T^-} q_t^f = q_t^p$ when $\mu_t = m_t$, which our numerical work (not shown here) confirms.

Next, in the right panel of Figure 5 we enlarge the value of $v(0)$ by setting $v(0) = 100$, and we plot the graphs of q_t^f and q_t^p at time $t = 0$. We observe that, unlike the full information case, for which the vertical intercept of q_t^f is positive (see Remark 3.2), the intercept of q_t^p as a function of m_t can be negative.

Example 5.2. From Corollary 3.1, we know that the slope of q_t^f increases with t , and the intercept of q_t^f decreases with t . Motivated by this corollary, in this example, we investigate the

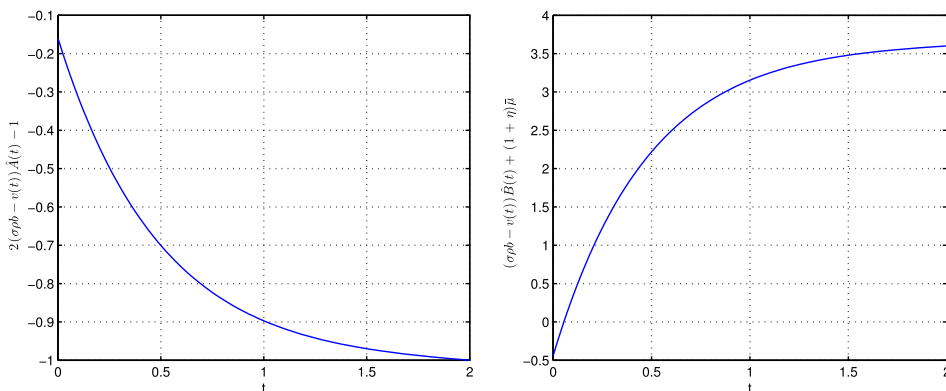


FIGURE 7. Left: The slope of q_t^p divided by the ratio $1/(\sigma^2\gamma)$ when $v(0)$ is large. Right: The intercept of q_t^p divided by the ratio $1/(\sigma^2\gamma)$ when $v(0)$ is large.

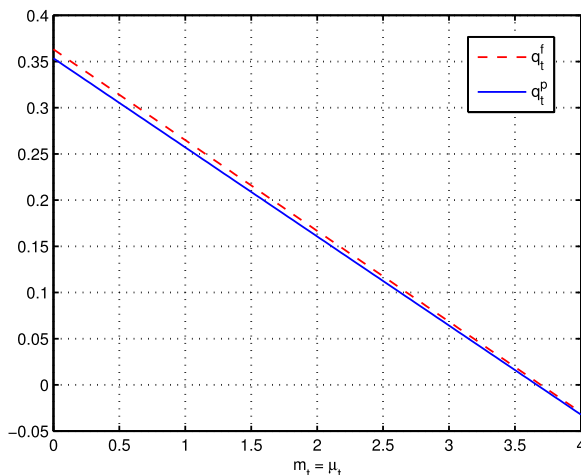


FIGURE 8. Optimal proportional retention under full and partial information when a is large.

changes of the slope and the intercept of q_t^p with t . In the left panel of Figure 6 we set $a = 1$ and $v(0) = 0.5$. We plot the graph of the slope of q_t^p divided by the ratio $1/(\sigma^2\gamma)$ as a function of $t \in [0, T]$. In the right panel of Figure 6 we plot the graph of the intercept of q_t^p divided by the ratio $1/(\sigma^2\gamma)$ as a function of $t \in [0, T]$. In these graphs, when $v(0)$ is relatively small, we see that the slope and intercept of q_t^p increase and decrease with t , respectively, as is true for q_t^f .

In Figure 7 we plot the slope and intercept of q_t^p divided by the ratio $1/(\sigma^2\gamma)$ when $v(0)$ is relatively large, that is, $v(0) = 100$. We find that both the slopes and intercepts of q_t^p are monotonic with time t , but the monotonicity is the opposite of that when $v(0)$ is relatively small, that is, $v(0) = 0.5$.

Example 5.3. In Figure 8 we plot the graph of the optimal proportional retention q_t^f and q_t^p when the parameter a is large. We set $a = 3$ and $v(0) = 0.5$. This graph shows that the values of q_t^f and q_t^p are close to each other when the parameter a is large.

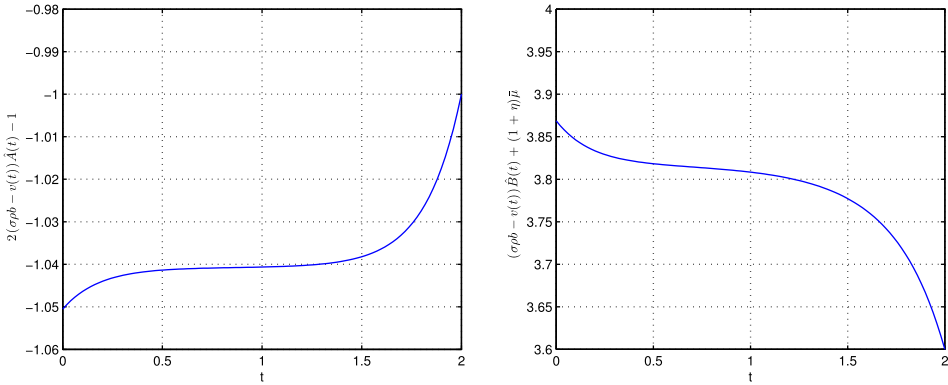


FIGURE 9. Left: The slope of q_t^p divided by the ratio $1/(\sigma^2\gamma)$ when a is large. Right: The intercept of q_t^p divided by the ratio $1/(\sigma^2\gamma)$ when a is large.

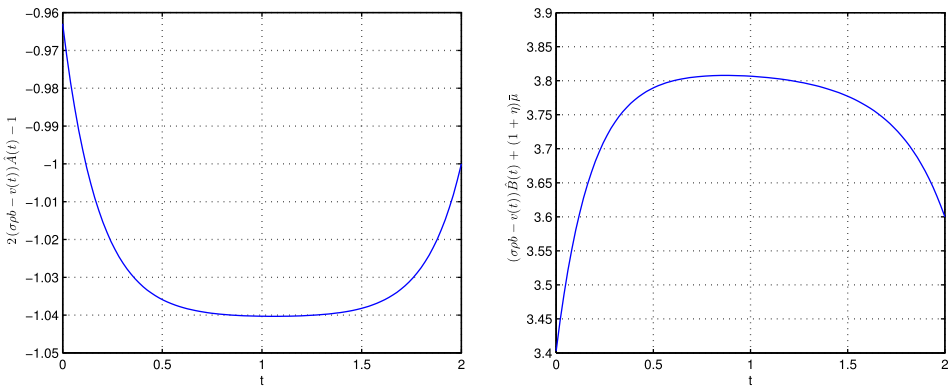


FIGURE 10. Left: The slope of q_t^p divided by the ratio $1/(\sigma^2\gamma)$ when both a and $v(0)$ are large. Right: The intercept of q_t^p divided by the ratio $1/(\sigma^2\gamma)$ when both a and $v(0)$ are large.

In Figure 9 we plot the slope and intercept of q_t^p divided by the ratio $1/(\sigma^2\gamma)$ in this case; we see that they are monotonic with respect to t , with the same monotonicity that the slope and intercept of q_t^f possess. Furthermore, when we take a larger value of $v(0)$, such as $v(0) = 5$, Figure 10 shows that they both lose monotonicity with t .

Appendix A. Derivation and solution of the Riccati equation (2.6)

First, [23, Section 10.2.1] proved that $v(t) = \mathbb{E}((\mu_t - m_t)^2)$, i.e. the conditional variance of μ_t equals the unconditional variance.

Define the process $\{\delta_t\}_{0 \leq t \leq T}$ by $\delta_t = \mu_t - m_t$. Then, by (2.1) and (2.4), we have

$$d\delta_t = -a\delta_t dt + b dW_{2,t} - \frac{\sigma \rho b - v(t)}{\sigma} d\bar{W}_{1,t}.$$

Use (2.5) to replace $d\bar{W}_{1,t}$ with $dW_{1,t} - (\delta_t/\sigma)dt$,

$$d\delta_t = \delta_t \left(-a + \frac{\sigma\rho b - v(t)}{\sigma^2} \right) dt + b dW_{2,t} - \frac{\sigma\rho b - v(t)}{\sigma} dW_{1,t},$$

and Itô's formula gives us

$$\begin{aligned} \delta_t^2 &= \delta_0^2 + 2 \int_0^t \delta_s \left(\left(-a + \frac{\sigma\rho b - v(s)}{\sigma^2} \right) \delta_s ds + b dW_{2,s} - \frac{\sigma\rho b - v(s)}{\sigma} dW_{1,s} \right) \\ &+ \int_0^t \left(b^2 - 2\rho b \frac{\sigma\rho b - v(s)}{\sigma} + \left(\frac{\sigma\rho b - v(s)}{\sigma} \right)^2 \right) ds. \end{aligned} \tag{A.1}$$

Because W_1 and W_2 are \mathbb{F} -standard Brownian motions, and because $v(t)$ and $\mathbb{E}(\delta_t^2)$ are continuous with respect to t and, thus, are bounded on $[0, T]$, we have

$$\mathbb{E} \left(\int_0^t \delta_s \frac{\sigma\rho b - v(s)}{\sigma} dW_{1,s} \right) = 0 = \mathbb{E} \left(\int_0^t \delta_s b dW_{2,s} \right).$$

By taking (unconditional) expectation of both sides of (A.1), we obtain

$$\begin{aligned} v(t) &= v(0) + \int_0^t \left(2 \left(-a + \frac{\sigma\rho b - v(s)}{\sigma^2} \right) v(s) + b^2 - 2\rho b \frac{\sigma\rho b - v(s)}{\sigma} \right. \\ &\quad \left. + \left(\frac{\sigma\rho b - v(s)}{\sigma} \right)^2 \right) ds \\ &= v(0) + \int_0^t \left(b^2 - 2av(s) - \left(\frac{\sigma\rho b - v(s)}{\sigma} \right)^2 \right) ds, \end{aligned}$$

which gives us the Riccati equation (2.6).

To solve this equation (see [19, Remark 4.2], which also provides an explicit solution for the constant-coefficient, one-dimensional Riccati equation), first define the function y by

$$y(t) = \frac{\sigma\rho b - v(t)}{\sigma^2}; \tag{A.2}$$

(2.6) gives us the following Riccati equation for y :

$$\frac{dy(t)}{dt} = \frac{2\sigma a\rho b - b^2}{\sigma^2} - 2ay(t) + y^2(t). \tag{A.3}$$

Next, define the function u by

$$y(t) = -\frac{u'(t)}{u(t)},$$

or equivalently,

$$u(t) = \exp \left\{ - \int y(t) dt \right\}; \tag{A.4}$$

then, (A.3) gives us the following second-order ordinary differential equation (ODE) for u :

$$\frac{2\sigma a\rho b - b^2}{\sigma^2} u(t) + 2au'(t) + u''(t) = 0. \tag{A.5}$$

The ODE in (A.5) has the general solution $u(t) = A_1 e^{r_1 t} + A_2 e^{r_2 t}$, in which A_1 and A_2 are constants to be determined, $r_1 = -a + \Delta$, and $r_2 = -a - \Delta$, with Δ given in (2.7). By reversing (A.2) and (A.4), we obtain the following general expression for v :

$$v(t) = \sigma^2 \Delta \cdot \frac{A_1 e^{2\Delta \cdot t} - A_2}{A_1 e^{2\Delta \cdot t} + A_2} - \sigma(\sigma a - \rho b),$$

or equivalently,

$$v(t) = \sigma^2 \Delta \cdot \frac{R e^{2\Delta \cdot t} - 1}{R e^{2\Delta \cdot t} + 1} - \sigma(\sigma a - \rho b),$$

in which $R = A_1/A_2$. By using the given initial condition $v(0)$, we determine R :

$$v(0) = \sigma^2 \Delta \cdot \frac{R - 1}{R + 1} - \sigma(\sigma a - \rho b),$$

which gives us R as in (2.8).

Appendix B. Proof of Theorem 3.2

From related work with exponential utility (e.g. [6]), we hypothesize that the value function is of the form

$$V^f(t, x, \mu) = -e^{-\gamma x} \exp\{A(t)\mu^2 + B(t)\mu + C(t)\} \tag{B.1}$$

for some functions of time A, B , and C , with $A(T) = B(T) = C(T) = 0$. The terminal conditions follow from $V^f(T, x, \mu) = -e^{-\gamma x}$.

Because V^f in (B.1) is concave with respect to x , the first-order necessary condition in (3.1) is sufficient, and we obtain the optimal retention in feedback form as

$$\begin{aligned} q^f(t, x, \mu) &= -\frac{((1 + \eta)\bar{\mu} - \mu_t)V_x^f(t, x, \mu) + \sigma \rho b V_{x\mu}^f(t, x, \mu)}{\sigma^2 V_{xx}^f(t, x, \mu)} \\ &= \frac{1}{\sigma^2 \gamma} \{(2\sigma b \rho A(t) - 1)\mu_t + (1 + \eta)\bar{\mu} + \sigma b \rho B(t)\}, \end{aligned} \tag{B.2}$$

in which we abuse notation slightly by using q^f to refer both to the optimal retention strategy (as in $q^f = \{q_t^f\}_{0 \leq t \leq T}$) and to the deterministic function q^f in (B.2). Note that q^f in (B.2) is independent of the surplus.

By substituting (B.1) and (B.2) into (3.1) and rearranging terms, we obtain

$$\begin{aligned} &\left\{ A'(t) + 2b^2(1 - \rho^2)A^2(t) - \frac{2}{\sigma}(\sigma a - \rho b)A(t) - \frac{1}{2\sigma^2} \right\} \mu^2 \\ &+ \left\{ B'(t) + \left(2b^2(1 - \rho^2)A(t) - \frac{1}{\sigma}(\sigma a - \rho b) \right) B(t) + 2\bar{\mu} \left(a - \frac{\rho b(1 + \eta)}{\sigma} \right) A(t) \right. \\ &\quad \left. + \frac{(1 + \eta)\bar{\mu}}{\sigma^2} \right\} \mu + \left\{ C'(t) + \frac{1}{2}b^2(1 - \rho^2)B^2(t) + \left(a - \frac{\rho b(1 + \eta)}{\sigma} \right) \bar{\mu} B(t) \right. \\ &\quad \left. + b^2 A(t) + (\eta - \theta)\bar{\mu}\gamma - \frac{(1 + \eta)^2 \bar{\mu}^2}{2\sigma^2} \right\} = 0. \end{aligned}$$

Thus, we have the following three differential equations for A, B , and C :

$$A'(t) + 2b^2(1 - \rho^2)A^2(t) - \frac{2}{\sigma}(\sigma a - \rho b)A(t) - \frac{1}{2\sigma^2} = 0, \tag{B.3}$$

$$B'(t) + \left(2b^2(1 - \rho^2)A(t) - \frac{1}{\sigma}(\sigma a - \rho b) \right) B(t) + 2\bar{\mu} \left(a - \frac{\rho b(1 + \eta)}{\sigma} \right) A(t) + \frac{(1 + \eta)\bar{\mu}}{\sigma^2} = 0, \tag{B.4}$$

$$C'(t) + \frac{1}{2}b^2(1 - \rho^2)B^2(t) + \left(a - \frac{\rho b(1 + \eta)}{\sigma} \right) \bar{\mu} B(t) + b^2A(t) + (\eta - \theta)\bar{\mu}\gamma - \frac{(1 + \eta)^2\bar{\mu}^2}{2\sigma^2} = 0, \tag{B.5}$$

with terminal conditions $A(T) = 0, B(T) = 0,$ and $C(T) = 0.$

Equation (B.3) is a constant-coefficient Riccati equation, which we can solve explicitly by using the same method as in Appendix A, although we are given $A(T) = 0$ instead of $v(0)$; by doing so, we obtain the expression for A in (3.2).

Equation (B.4) is a linear differential equation, which, by substituting for A from (3.2), we can rewrite as

$$0 = \frac{d}{dt} \left[\left\{ \alpha_1 e^{\Delta \cdot (T-t)} + \alpha_2 e^{-\Delta \cdot (T-t)} \right\} B(t) \right] + 2\bar{\mu} \left\{ ((1 + \eta)\Delta + \eta a) e^{\Delta \cdot (T-t)} + ((1 + \eta)\Delta - \eta a) e^{-\Delta \cdot (T-t)} \right\}.$$

By integrating this from t to T and using $B(T) = 0,$ we obtain the expression for B in (3.3).

Finally, by integrating (B.5) from t to T and by using $C(T) = 0,$ we obtain the integral representation for C in (3.4).

It remains to show that V^f and q^f satisfy the conditions of Theorem 3.1. By construction, V^f satisfies the HJB equation (3.1) with boundary condition $V^f(T, x, \mu) = -e^{-\gamma x}.$ To show that the retention strategy q^f in (3.5) is admissible, we check the three conditions in the definition of admissibility. First, note that q^f is adapted to the filtration \mathbb{F} by its definition. Second, q^f is conditional L^2 -integrable because $A(t)$ and $B(t)$ are bounded functions on $[0, T],$ and $\mu = \{\mu_t\}$ is conditional L^2 -integrable. Third, the SDE (2.2) has a pathwise unique solution, which is easy to see because q^f is independent of the surplus $X.$ Thus, q^f is admissible.

From Theorem 3.1, we deduce that V^f and q^f as stated in Theorem 3.2 equal the value function and optimal retention strategy, respectively.

Appendix C. Proof of Theorem 4.2

As in Appendix B, we hypothesize that the value function is of the form

$$V^p(t, x, m) = -e^{-\gamma x} \exp \{ \widehat{A}(t)m^2 + \widehat{B}(t)m + \widehat{C}(t) \} \tag{C.1}$$

for some functions of time $\widehat{A}, \widehat{B},$ and $\widehat{C},$ with $\widehat{A}(T) = \widehat{B}(T) = \widehat{C}(T) = 0.$ The terminal conditions follow from $V^p(T, x, \mu) = -e^{-\gamma x}.$

Because V^p in (C.1) is concave with respect to $x,$ the first-order necessary condition in (4.1) is sufficient, and we obtain the optimal retention in feedback form as

$$q^p(t, x, m) = - \frac{((1 + \eta)\bar{\mu} - m)V_x^p(t, x, m) + (\rho\sigma b - v(t))V_{xm}^p(t, x, m)}{\sigma^2 V_{xx}^p(t, x, m)} = \frac{1}{\sigma^2 \gamma} \{ (2(\sigma\rho b - v(t))\widehat{A}(t) - 1)m + (\sigma\rho b - v(t))\widehat{B}(t) + (1 + \eta)\bar{\mu} \}, \tag{C.2}$$

in which we abuse notation slightly by using q^p to refer both to the optimal retention strategy and to the deterministic function q^p in (C.2).

By substituting (C.1) and (C.2) into (4.1) and rearranging terms, we obtain

$$\begin{aligned} & \left\{ \widehat{A}'(t) - \frac{2}{\sigma} \left((\sigma a - \rho b) + \frac{v(t)}{\sigma} \right) \widehat{A}(t) - \frac{1}{2\sigma^2} \right\} m^2 \\ & + \left\{ \widehat{B}'(t) - \frac{1}{\sigma} \left((\sigma a - \rho b) + \frac{v(t)}{\sigma} \right) \widehat{B}(t) + 2\bar{\mu} \left(a - (1 + \eta) \frac{\sigma \rho b - v(t)}{\sigma^2} \right) \widehat{A}(t) \right. \\ & \quad \left. + \frac{(1 + \eta)\bar{\mu}}{\sigma^2} \right\} m + \left\{ \widehat{C}'(t) + \left(a - (1 + \eta) \frac{\sigma \rho b - v(t)}{\sigma^2} \right) \bar{\mu} \widehat{B}(t) + \left(\rho b - \frac{v(t)}{\sigma} \right)^2 \widehat{A}(t) \right. \\ & \quad \left. + (\eta - \theta)\bar{\mu}\gamma - \frac{(1 + \eta)^2 \bar{\mu}^2}{2\sigma^2} \right\} = 0. \end{aligned}$$

Thus, we have the following three differential equations for $\widehat{A}(t)$, $\widehat{B}(t)$, and $\widehat{C}(t)$:

$$\widehat{A}'(t) - \frac{2}{\sigma} \left((\sigma a - \rho b) + \frac{v(t)}{\sigma} \right) \widehat{A}(t) - \frac{1}{2\sigma^2} = 0, \tag{C.3}$$

$$\widehat{B}'(t) - \frac{1}{\sigma} \left((\sigma a - \rho b) + \frac{v(t)}{\sigma} \right) \widehat{B}(t) + 2\bar{\mu} \left(a - (1 + \eta) \frac{\sigma \rho b - v(t)}{\sigma^2} \right) \widehat{A}(t) + \frac{(1 + \eta)\bar{\mu}}{\sigma^2} = 0, \tag{C.4}$$

$$\begin{aligned} & \widehat{C}'(t) + \left(a - (1 + \eta) \frac{\sigma \rho b - v(t)}{\sigma^2} \right) \bar{\mu} \widehat{B}(t) + \left(\rho b - \frac{v(t)}{\sigma} \right)^2 \widehat{A}(t) + (\eta - \theta)\bar{\mu}\gamma \\ & - \frac{(1 + \eta)^2 \bar{\mu}^2}{2\sigma^2} = 0, \end{aligned} \tag{C.5}$$

with terminal conditions $\widehat{A}(T) = 0$, $\widehat{B}(T) = 0$, and $\widehat{C}(T) = 0$.

We can rewrite (C.3) as

$$0 = \frac{d}{dt} \left[\frac{\widehat{A}(t)}{(\text{Re}^{\Delta \cdot t} + e^{-\Delta \cdot t})^2} \right] - \frac{1}{2\sigma^2} \cdot \frac{1}{(\text{Re}^{\Delta \cdot t} + e^{-\Delta \cdot t})^2}.$$

and by integrating this from t to T and using $\widehat{A}(T) = 0$, we obtain the expression for \widehat{A} in (4.2).

Equation (C.4) is a linear differential equation, which, by substituting for \widehat{A} from (4.2), we rewrite as

$$0 = \frac{d}{dt} \left[\frac{\widehat{B}(t)}{\text{Re}^{\Delta \cdot t} + e^{-\Delta \cdot t}} \right] + \frac{\bar{\mu}}{2\sigma^2 \Delta} \cdot \frac{e^{\Delta \cdot t}}{\text{Re}^{2\Delta \cdot T} + 1} \left\{ (1 + \eta)\Delta(e^{2\Delta \cdot (T-t)} + 1) + \eta a(e^{2\Delta \cdot (T-t)} - 1) \right\},$$

and by integrating this from t to T and using $\widehat{B}(T) = 0$, we obtain the expression for \widehat{B} in (4.3).

Finally, by integrating (C.5) from t to T and using $\widehat{C}(T) = 0$, we obtain the integral representation of \widehat{C} in (4.4).

It remains to show that V^p and q^p satisfy the conditions of Theorem 4.1. The argument is similar to the one in the proof of Theorem 3.2, so we omit it. Thus, V^p and q^p as stated in Theorem 4.2 equal the value function and optimal retention strategy, respectively.

Funding information

X. Liang thanks the Research Foundation for Returned Scholars of Hebei Province (C20200102), Natural Science Foundation of Tianjin (19JCYBJC30400), and the Natural Science Foundation of Hebei Province (A2020202033) for financial support, and Virginia R. Young thanks the Cecil J. and Ethel M. Nesbitt Professorship of Actuarial Mathematics for financial support.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

- [1] BÄUERLE, N. AND LEIMCKE, G. (2021). Robust optimal investment and reinsurance problems with learning. *Scand. Actuar. J.* **2021**, 82–109.
- [2] BÄUERLE, N. AND RIEDER, U. (2007). Portfolio optimization with jumps and unobservable intensity process. *Math. Finance* **17**, 205–224.
- [3] BI, J., LIANG, Z. AND YUEN, K. C. (2019). Optimal mean-variance investment/reinsurance with common shock in a regime-switching market. *Math. Meth. Operat. Res.* **90**, 109–135.
- [4] BJÖRK, T., DAVIS, M. H. A. AND LANDÉN, C. (2010). Optimal investment under partial information. *Math. Meth. Operat. Res.* **71**, 371–399.
- [5] BRACHETTA, M. AND CECI C. (2020). A BSDE-based approach for the optimal reinsurance problem under partial information. *Insurance Math. Econom.* **95**, 1–16.
- [6] BRENDLE, S. (2004). Portfolio selection under partial observation and constant absolute risk aversion. Available at <https://ssrn.com/abstract=590362>.
- [7] BRENDLE, S. (2006). Portfolio selection under incomplete information. *Stoch. Process. Appl.* **116**, 701–723.
- [8] BROWNE, S. (1995). Optimal investment policies for a firm with a random risk process: Exponential utility and minimizing the probability of ruin. *Math. Operat. Res.* **20**, 937–958.
- [9] CHEN, L. AND SHEN, Y. (2018). On a new paradigm of optimal reinsurance: A stochastic Stackelberg differential game between an insurer and a reinsurer. *ASTIN Bull.* **48**, 905–960.
- [10] CRANDALL, M. G., ISHII, H. AND LIONS, P. L. (1992). User's guide to viscosity solutions of second-order partial differential equations. *Bull. Amer. Math. Soc.* **27**, 1–67.
- [11] CVITANIĆ, J. AND KARATZAS, I. (1992). Convex duality in constrained portfolio optimization, *Ann. Appl. Prob.* **2**, 767–818.
- [12] DASSIOS, A. AND JANG J.-W. (2005). Kalman–Bucy filtering for linear systems driven by the Cox process with shot noise intensity and its application to the pricing of reinsurance contracts. *J. Appl. Prob.* **42**, 93–107.
- [13] DELONG, L. (2013). *Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications*. Springer, New York.
- [14] GENNOTTE, G. (1986). Optimal portfolio choice under incomplete information. *J. Finance* **41**, 733–749.
- [15] HIPPE, C. AND TAKSAR, M. (2010). Optimal non-proportional reinsurance control. *Insurance Math. Econom.* **47**, 246–254.
- [16] HONDA, T. (2003). Optimal portfolio choice for unobservable and regime-switching mean returns. *J. Econom. Dynam. Control* **28**, 45–78.
- [17] IRGENS, C. AND PAULSEN, J. (2004). Optimal control of risk exposure, reinsurance and investments for insurance portfolios. *Insurance Math. Econom.* **35**, 21–51.
- [18] LAKNER, P. (1995). Utility maximization with partial information. *Stoch. Process. Appl.* **56**, 247–273.
- [19] LAKNER, P. (1998). Optimal trading strategy for an investor: The case of partial information. *Stoch. Process. Appl.* **76**, 77–97.
- [20] LI, D., LI, D. AND YOUNG, V. R. (2017). Optimality of excess-loss reinsurance under a mean-variance criterion. *Insurance Math. Econom.* **75**, 82–89.
- [21] LIANG, Z. AND BAYRAKTAR, E. (2014). Optimal reinsurance and investment with unobservable claim size and intensity. *Insurance Math. Econom.* **55**, 156–166.
- [22] LIANG, Z., YUEN, K. C. AND GUO, J. (2011). Optimal proportional reinsurance and investment in a stock market with Ornstein–Uhlenbeck process. *Insurance Math. Econom.* **49**, 207–215.
- [23] LIPTSER, R. S. AND SHIRYAEV, A. N. (1997). *Statistics of Random Processes* Vol. I (Stoch. Model. Appl. Prob. 5). Springer, New York.

- [24] PENG, S. (1992). Stochastic Hamilton–Jacobi–Bellman equations. *SIAM J. Control Optimization* **30**, 284–304.
- [25] PENG, X. AND HU, Y. (2013). Optimal proportional reinsurance and investment under partial information. *Insurance Math. Econom.* **53**, 416–428.
- [26] PROMISLOW, S. D. AND YOUNG, V. R. (2005). Minimizing the probability of ruin when claims follow Brownian motion with drift. *N. Amer. Actuar. J.* **9**, 110–128.
- [27] PUTSCHÖGL, W. AND SASS, J. (2011). Optimal investment under dynamic risk constraints and partial information. *Quant. Finance* **11**, 1547–1564.
- [28] SCHMIDLI, H. (2002). On minimizing the ruin probability by investment and reinsurance. *Ann. Appl. Prob.* **12**, 890–907.
- [29] SUN, Z., ZHANG, X. AND YUEN, K. C. (2020). Mean-variance asset-liability management with affine diffusion factor process and a reinsurance option. *Scand. Actuar. J.* **2020**, 218–244.
- [30] XIONG, J., XU, Z. AND ZHENG, J. (2021). Mean-variance portfolio selection under partial information with drift uncertainty. *Quant. Finance* **21**, 1–13.
- [31] YONG, J. AND ZHOU, X. (1999). *Stochastic Controls: Hamiltonian Systems and HJB Equations* (Stoch. Model. Appl. Prob. **43**). Springer, New York.
- [32] ZHANG, X. AND SIU, T. K. (2009). Optimal investment and reinsurance of an insurer with model uncertainty. *Insurance Math. Econom.* **45**, 81–88.