IET Control Theory & Applications

Research Article

Approximation of optimal ergodic dividend strategies using controlled Markov chains

ISSN 1751-8644
Received on 26th November 2017
Revised 17th May 2018
Accepted on 30th July 2018
E-First on 5th September 2018
doi: 10.1049/iet-cta.2018.5394
www.ietdl.org

Zhuo Jin[1] ✉, *Hailiang Yang*[2], *George Yin*[3]
[1]*Centre for Actuarial Studies, Department of Economics, The University of Melbourne, VIC 3010, Australia*
[2]*Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong*
[3]*Department of Mathematics, Wayne State University, Detroit, Michigan 48202, USA*
✉ *E-mail: zjin@unimelb.edu.au*

Abstract: This study develops a numerical method to find optimal ergodic (long-run average) dividend strategies in a regime-switching model. The surplus process is modelled by a regime-switching process subject to liability constraints. The regime-switching process is modelled by a finite-time continuous-time Markov chain. Using the dynamic programming principle, the optimal long-term average dividend payment is a solution to the coupled system of Hamilton–Jacobi–Bellman equations. Under suitable conditions, the optimal value of the long-term average dividend payment can be determined by using an invariant measure. However, due to the regime switching, getting the invariant measure is very difficult. The objective is to design a numerical algorithm to approximate the optimal ergodic dividend payment strategy. By using the Markov chain approximation techniques, the authors construct a discrete-time controlled Markov chain for the approximation, and prove the convergence of the approximating sequences. A numerical example is presented to demonstrate the applicability of the algorithm.

1 Introduction

Surplus and dividend management has long been an important issue in asset and liability management of public companies. Due to the wave of demutualisation of insurance companies in the last few decades, the dividend payment plan released in the financial report for public insurance companies is an informative signal to represent the potential risks and profitability opportunities. The decision of the dividend payment is so important for a public company's financial strength because the company's share price is very sensitive to the information about dividend plans and dividend payment strategies also influence the investment and capital raising decisions of firms.

Insurance companies generally accumulate large amounts of cash in the form of premiums to pay future claims and benefits of policyholders. Hence, dividend payments, which are appealing returns to shareholders, may reduce insurance companies' ability to survive under disasters in the experience of underwriting and investment. However, due to the undergone pressure of managing balance sheets and distributing the surplus for public insurance companies, one natural objective of insurers is to optimise the management of surplus and sustained stream of dividend payments. Practitioners in insurance firms manage the surplus and dividend payment stream against various financial risks so that the firms can avoid financial ruin in the long run.

Since the optimal dividend payment model was proposed in [1], increasing efforts have been made to study the optimal dividend policy by using stochastic control theory. The majority of research is conducted aiming to maximise the present value of the cumulative dividend payment (with or without costs) in targeted time horizon, and find corresponding optimal strategies. Regular controls, singular controls, and impulse controls are involved in various scenarios. Guo *et al.* [2] studied the dividend and risk control with a diffusion where the drift is quadratic in the risk control variable. He and Liang [3] studied the mixed control of dividend, proportional reinsurance and financing for the model. Lokka and Zervos [4] solved the optimal dividend and issuance of equity policies in the presence of bankruptcy risk. Wei *et al.* [5] studied the combined control of dividend, financing and risk for the Brownian motion model. The authors of [6, 7] studied the impulse dividend control problem for a rather general linear diffusion model in which some growth and smoothness conditions are imposed on. Avram *et al*. [8] addressed the dividend and reinvestment control in a spectrally negative Lévy process. Azcue and Muler [9] analysed the strategies to maximise the accumulated discounted dividend payments of an insurance company. Jin *et al*. [10] considered the credit risk of an insurance company and derived the company's optimal dividend payment strategies and debt level.

Previous work is mainly based on the classical Cramér–Lundberg risk model to maximise the accumulated discounted dividend payments; the company will be insolvent almost surely when optimal dividend strategies are adopted. In [10], the authors propose an asset and liability model with liability constraint. The insurance company manages the surplus and designs the dividend payment strategies taking into account the liability capacity. Then, the insurance company will be in the absence of insolvency. On the other hand, since insurance companies generally have assets and liabilities with long maturity, in particular, for life insurance companies, it is very important for insurance companies to consider long-term objectives and build the dividend payment strategies with a long time horizon in mind. The long-term objectives are widely studied in investment and risk management in a variety of cases. Bielecki and Pliska [11] analysed the dynamic asset management in an infinite time horizon. Fleming and Sheu [12] studied the optimal investment strategies with a long-term objective. Pham [13] proposed a large deviation approach for finding optimal strategies of long-term investment. See also [14, 15] for related works. In this work, we extend the asset and liability model in [10] and set a new objective functions to consider the long-term impact of the dividend payment strategies, which is applicable when financial ruin is completely avoided. Instead of adopting the discounted present value, we aim to maximise the average dividend payment in the long term.

Recently, stochastic hybrid surplus models have been widely used to capture discrete movements (such as market cycles, economic environment, business trends etc.). The hybrid system investigates the coexistence of discrete events and continuous dynamics in the stochastic systems. To reflect the hybrid feature, one option is to use a finite-state Markov process to describe the transitions among different regimes. The Markov-modulated switching systems are therefore known as regime-switching

© The Institution of Engineering and Technology 2018

systems. Thus, the formulation of regime-switching models is a more versatile and general framework to describe the complex financial and insurance markets and their inherent randomness and uncertainty. In [16], optimal reinsurance and dividend payment strategies are studied in a compound Poisson regime-switching model. Sotomayor and Cadenillas [17] studied the optimal dividend problem in the regime-switching model when dividend rates are bounded, unbounded, and when there are fixed costs and taxes corresponding to the dividend payments. Zhu [18] studied the dividend optimisation for a regime-switching diffusion model with restricted dividend rates. See also [19–21]. A comprehensive introduction of regime-switching diffusions is presented in [22].

Using the dynamic programming principle to treat optimal dividend payment strategies, one usually solves a Hamilton–Jacobi–Bellman (HJB) equation. Due to the Markov regime-switching diffusion process, the HJB equation is formed as a coupled system of HJB equations. To represent the maximal average dividend payment in the long term, the unique invariant measure is constructed. Due to the complexity of the Markov switches, obtaining the explicit formula of the long-term average is virtually impossible. Employing numerical approximations is an alternative. We adopt the Markov chain approximation methodology developed in [23] to solve for the optimal performance function and the corresponding dividend payment strategies. A numerical method for finding optimal investment and dividend payment policies with capital injections under regime-switching diffusion models was developed in the work of [24]. In this work, we carry out a convergence analysis using weak convergence methods for the regime-switching models, in which case one needs to deal with a system of HJB equations with reflecting boundaries. Comparing with the work in [24], we choose a different objective function, which performs in an infinite time horizon and requires the ergodicity of the diffusion process. The long-run average objective function adds many difficulties to design the numerical schemes and to conduct the convergence analysis of the algorithm.

The remainder of the paper is organised as follows. A general formulation of the surplus process, performance functions, and assumptions are presented in Section 2. The existence and uniqueness of the invariant measure are presented. The dynamic programming equation is derived in Section 4. Section 5 deals with the Markov chain approximation method and designs the numerical algorithm. Section 6 works on the convergence of the approximation scheme. A numerical example is provided in Section 7 to illustrate the performance of the approximation method together with some further remarks.

2 Formulation

Following the framework of asset and liability management, the surplus process $X(t)$ equals the difference between the asset value $A(t)$ and liabilities $L(t)$. Then

$$X(t) = A(t) - L(t). \tag{1}$$

To describe the abrupt changes of the insurance market, we use a continuous-time irreducible Markov chain denoted by $\alpha(t)$ that takes values in a finite space. The finite space of market states is denoted by $\mathcal{M} = \{1, \dots, m\}$. Assume that the continuous-time finite-state Markov chain $\alpha(t)$ is generated by $Q = (q_{ij}) \in \mathbb{R}^{m \times m}$ $\mathbb{P}\{\alpha(t+\Delta)=j|\alpha(t)=i\} =\{ q_{ij}\Delta+o(\Delta),\ if\ j\neq i,\ 1+q_{ii}\Delta+o(\Delta),\ if\ j=i$, where $q_{ij} \geq 0$ for $i, j = 1, 2, \dots, m$ with $j \neq i$ and $q_{ii} = -\sum_{j \neq i} q_{ij} < 0$ for each $i = 1, 2, \dots, m$. Q is assumed to be irreducible.

When an insurance policy is sold at time t, the insurer takes the liability of the amount insured $L(t)$ and receives an on-going premium for it. Let $\beta(i)$ be the cost of protection for per dollar of the insured amount, where for each $i \in \mathcal{M}$. Then $\beta(\cdot)$ is the premium rate, and the asset value increases during the time period $[t, t + \mathrm{d}t]$ equals $\beta(\alpha(t))L(t)\,\mathrm{d}t$ from the insurance liability $L(t)$.

Reinsurance is a standard tool to eliminate the risks for insurance companies. Primary insurers buy insurance products from reinsurance companies to reduce the claim volatilities and pay a certain part of the premiums collected from the original contracts to reinsure. In return, the reinsurance companies are obliged to share part of the risks. Proportional reinsurance and excess-of-loss reinsurance are two major types of reinsurance schemes. In our work, we adopt the proportional reinsurance, where a fixed percentage of loss is retained for the primary insurer.

Denote λ as the retention level dictated in the reinsurance contract, where $\lambda \in [0, 1]$. Then λ will be covered by the primary insurance company per dollar of liability. Let $h(\lambda)$ be the reinsurance charge rate, which is the cost of protection for the remaining $1 - \lambda$ part per dollar of liability. Hence, the reinsurance charge for liability $L(t)$ during the time period $[t, t + \mathrm{d}t]$ is $h(\lambda)L(t)\,\mathrm{d}t$, and the out of pocket claim expense is $\lambda L(t)\,\mathrm{d}t$ accordingly.

The primary insurer collects premium with a rate β and transfers part of the risk to the reinsurer with a retention level λ.

The primary insurer collects premium with a rate β and transfers part of the risk to the reinsurer with a retention level of λ. Insurers take into account the demand of the insurance contracts and decide how much liability to take. Let $\pi(t) = L(t)/X(t)$ be the liability ratio of the insurance company. The concept of liability ratio is closely related to the leverage, which is written as the ratio between the asset and surplus. That is, $A(t)/X(t) = 1 + \pi(t)$. By using the reinsurance tools, the insurance company decides the optimal insured amount sold in insurance contracts and avoid overtaking the liabilities.

We assume that the process of asset value $A(t)$ follows a geometric Brownian motion process

$$\frac{\mathrm{d}A(t)}{A(t)} = \mu(\alpha(t))\,\mathrm{d}t + \sigma(\alpha(t))\,\mathrm{d}w(t), \tag{2}$$

where $\mu(i)$ is the return rate of the asset, $\sigma(i)$ is the corresponding volatility for $i \in \mathcal{M}$ and $w(t)$ is a standard Brownian motion. Hence, combining (1) and (2), the surplus process $\widetilde{X}(t)$ in the absence of claims and dividend payments follows:

$$\begin{aligned} \mathrm{d}\widetilde{X}(t) = \quad & (\beta(\alpha(t)) - h(\lambda))L(t)\,\mathrm{d}t \\ & + A(t)(\mu(\alpha(t))\,\mathrm{d}t + \sigma(\alpha(t))\,\mathrm{d}w(t)). \end{aligned} \tag{3}$$

To proceed, we consider the claims. Denoted by $R(t)$ the accumulated claims up to time t. Let $c(t)$ be a claim rate against insured liabilities, which means that an amount of $c(t)$ is claimed per dollar of the liability up to time t. Hence, the accumulated claims up to time T is denoted as

$$R(T) = \int_0^T c(t)L(t)\,\mathrm{d}t, \tag{4}$$

Practically, the claim rate $c(t)$ is risky and is not predictable. The claim rates of different types of insurance products are very different and volatile in different market states. For example, the CDS, which can be considered as an insurance contract to protect against credit events, is affected by various factors such as credit ratings of firms, the demand for CDOs in the market, and government regulation etc. Furthermore, it is largely influenced by the uncertainty and randomness of the economic environment. In [10], the claim rate $c(t)$ is formulated as a diffusion process to describe its randomness. However, one of the main drawbacks of the diffusion process in the work is that the claim rate can be negative, which is very difficult to calibrate and explain with real market data. To guarantee the positivity of the claim, we assume that the claim rate follows a continuous-time Markov process, taking values in a set of positive values. That is, the claim rate depends on $\alpha(t)$, so $c(\alpha(t))$ in lieu of $c(t)$ is used. Hence, the accumulated claims follow

$$R(T) = \int_0^T c(\alpha(t))L(t)\,\mathrm{d}t. \tag{5}$$

© The Institution of Engineering and Technology 2018

The cumulative dividend payments $D(\cdot)$ is an $\mathscr{F}_t$-adapted process $\{D(t): t \ge 0\}$, which represents the accumulated dividend payments paid up to time t. Hence, $D(t)$ is a non-decreasing and non-negative process, and is right continuous with left limits. In our model, we assume that dividend payments are proportional to the surplus. Hence, a restricted dividend payment rate $u(t) \in [0, 1]$ is introduced. Let $U = [0, 1]$. Then the accumulated dividend $D(t)$ follows:

$$dD(t) = u(t)X(t)\,\mathrm{d}t, \tag{6}$$

where $u \in U$ is an $\mathscr{F}_t$-adapted process.

Thus, considering the impact of reinsurance, the claims and dividend payments, the insurer's surplus process follows:

$$\mathrm{d}X(t) = \mathrm{d}\widetilde{X}(t) - \lambda\,\mathrm{d}R(t) - \mathrm{d}D(t). \tag{7}$$

Together with the initial condition, (7) follows:

$$\begin{cases} \mathrm{d}X(t) = [(\beta(\alpha(t)) - h(\lambda) - \lambda c(\alpha(t)))L(t) + \mu(\alpha(t))A(t) \\ \qquad - u(t)X(t)]\,\mathrm{d}t + A(t)\sigma(\alpha(t))\,\mathrm{d}w(t), \\ X(0) = x \ge 0 \end{cases} \tag{8}$$

for all $t < \tau$, where $\tau = \inf\{t \ge 0: X(t) < 0\}$ represents the time of financial ruin. We impose $X(t) = 0$ for all $t > \tau$. No dividend will be paid after the financial ruin. Incorporating the liability ratio into (8), (8) can be rewritten as

$$\begin{cases} \dfrac{\mathrm{d}X(t)}{X(t)} = [\pi(t)(\beta(\alpha(t)) - h(\lambda) - \lambda c(\alpha(t)) + \mu(\alpha(t))) \\ \qquad + \mu(\alpha(t)) - u(t)]\,\mathrm{d}t + (\pi(t) + 1)\sigma(\alpha(t))\,\mathrm{d}w(t), \\ X(0) = x. \end{cases} \tag{9}$$

For a dividend payment rate, $u(t)$ is non-negative and restricted to a bounded region. An admissible strategy $u(\cdot)$ is progressively measurable with respect to $\{w(s), \alpha(s): 0 \le s \le t\}$. Denote the set of all admissible strategies by $\mathscr{U}$. Then the admissible strategy set $\mathscr{U}$ can be defined as

$$\mathscr{U} = \{u \in \mathbb{R}: 0 \le u \le 1\}. \tag{10}$$

A Borel measurable function $u(x, \alpha)$ is an admissible feedback strategy if (9) has a unique solution.

The objective of the representative financial institute is to maximise the average dividend payment in the long term. For an arbitrary admissible feedback control $u(\cdot, \cdot)$, the performance function is the average dividend payment in the long term given by

$$J(x, u, i) = \limsup_{T \to \infty} \frac{1}{T}\mathbb{E}_{x,i}\left[\int_0^T u(X(t), \alpha(t))X(t)\,\mathrm{d}t\right], \quad \forall i \in \mathscr{M}, \tag{11}$$

where $\mathbb{E}_{x,i}$ denote the expectation conditioned on $X(0) = x$ and $\alpha(0) = i$, and let $\mathbb{P}_{x,i}$ denote the conditional probability on $X(0) = x$ and $\alpha(0) = i$.

Denote by $\gamma(u) = J(x, u, i)$. Define the optimal value as

$$\bar{\gamma} := \sup_{u \in \mathscr{U}} \gamma(u). \tag{12}$$

For $i \in \mathscr{M}$, an arbitrary strategy $u \in \mathscr{U}$ and $V(\cdot, i) \in C^2(\mathbb{R})$, we define an operator $\mathscr{L}^u$ by

$$\begin{aligned} \mathscr{L}^u V(x, i) = \; & V_x(x, i)x(\pi(\beta(i) - h(\lambda) - \lambda c(i) + \mu(i)) + \mu(i) - u) \\ & + \frac{1}{2}(u + 1)^2\sigma^2(i)x^2V_{xx}(x, i) + QV(x, \cdot)(i) \end{aligned} \tag{13}$$

where V_x and V_{xx} present the first- and second-order derivatives with respect to x, and

$$QV(x, \cdot)(i) = \sum_{j \ne i} q_{ij}(V(x, j) - V(x, i)).$$

If $\bar{\gamma}$ exists, by using the dynamic programming principle [25], there exists a sufficiently smooth function V that normally satisfies the following coupled system of HJB equations:

$$\bar{\gamma} = \max_{u \in \mathscr{U}} \{\mathscr{L}^u V(x, i) + ux\}, \quad \text{for each } i \in \mathscr{M}. \tag{14}$$

In view of (9), the surplus is always non-negative in the infinite time horizon. The insurance company will run the business with probability one in the long run. It is worthwhile to consider the ergodic control of dividend payment when the operation period $T \to \infty$. On the other hand, in reality, the surplus of an insurance company cannot reach infinity. Once the surplus is substantially high, the decision maker will undergo pressure from the shareholders to pay dividend. Hence, we only need to choose B large enough and set the surplus in the finite interval $G = [0, B]$. To make $J(x, u, i)$ computationally feasible, we truncate x at some large value B.

3 Invariant measure

To obtain the expected average dividend payment in the infinite time horizon, one approach is to replace the instantaneous measures with invariant measures. Note that the state of the process in our formulation has two components: one component is the diffusion process $X(t)$; the other component is the Markov regime switching process $\alpha(t)$. We denote by $Z(t) = (X(t), \alpha(t))$ the state of the process.

To proceed, we need the following assumption.

(A) $Z(t)$ is positive recurrent with respect to some bounded domain $E \times \{i\}$, where $E \subset G \subset \mathbb{R}$, i is fixed and $i \in \mathscr{M}$.

Lemma 1 (assume (A)): $Z(t)$ is the positive recurrent with respect to $G \times \mathscr{M}$.

Proof: The result is immediately obtained by applying Theorem 3.12 in [22]. □

We proceed to define a sequence of stopping times $\{\eta_k\}, k = 0, 1, 2, \ldots$. Let $\eta_0 = 0$, η_{2k+1} be the first time after η_{2k} when $Z(t)$ reaches the boundary $\partial E \times \{i\}$, and η_{2k+2} be the first time after η_{2k+1} when $Z(t)$ reaches the boundary $\partial G \times \{i\}$. Then, the sample path of $Z(t)$ can be divided into the cycles as

$$[\eta_0, \eta_2), [\eta_2, \eta_4), \ldots, [\eta_{2k}, \eta_{2k+2}), \ldots \tag{15}$$

By Lemma 1, $Z(t)$ is the positive recurrent with respect to $G \times \mathscr{M}$. Hence the stopping times $\{\eta_k\}, k = 0, 1, 2, \ldots$ are finite almost surely. Without loss of generality, we assume $x = 0$. It follows that the sequence $Z_n = (X_n, i) = Z(\eta_{2n}), n = 0, 1, \ldots,$ is a Markov chain on $\partial G \times \{i\}$. Denote by $\mathscr{B}(\partial G)$ the collection of Borel measurable sets on ∂G. Starting from (x, i), $Z(t)$ may jump many times before it reaches the set (H, i) where $H \in \mathscr{B}(\partial G)$. The one-step transition probability of the Markov chain Z_n is defined as

$$\tilde{p}^{(1)}(x, H) = \mathbb{P}(Z_1 \in (H \times \{i\}) | Z_0 = (x, i)). \tag{16}$$

Analogously, the n-step transition probability of the Markov chain Z_n is denoted by $\tilde{p}^{(n)}(x, H)$. Now we will construct the stationary distribution of $Z(t)$.

Theorem 1: The positive recurrent process $Z(t)$ has a unique stationary distribution $\nu(\cdot, \cdot)$. Let $\theta(\cdot, \cdot)$ be the stationary

© The Institution of Engineering and Technology 2018

density associated with the stationary distribution. Then for any $(x, i) \in G \times \mathcal{M}$

$$\mathbb{P}_{x,i}\left(\lim_{T \to \infty} \frac{1}{T}\int_0^T u(X(t), \alpha(t))X(t)\,\mathrm{d}t = \bar{\gamma}\right) = 1, \tag{17}$$

where

$$\bar{\gamma} = \sum_{i=1}^{m} \int_{\mathbb{R}} u(x,i)x\theta(x,i)\,\mathrm{d}x\,. \tag{18}$$

Proof: In view of Lemma 4.1 in [22], Z_n has a unique stationary distribution $\phi(\cdot)$. For any $H \in \mathscr{B}(\mathbb{R})$, $\phi(H) = \lim_{n\to\infty} \tilde{p}^{(n)}(x, H)$. Recall that the cycles are defined in (15). Denote by $\tau^{H\times\{i\}}$ the time spent by the path of $Z(t)$ in the set $(H \times \{i\})$ during the first cycle. Set

$$\tilde{\nu}(H, i) := \int_{\partial G} \phi(\mathrm{d}x)\mathbb{E}_x \tau^{H\times\{i\}}\,. \tag{19}$$

Using Theorem 4.3 in [22], we have

$$\begin{aligned} &\sum_{i=1}^{m}\int_{\mathbb{R}} u(x,i)x\tilde{\nu}(\mathrm{d}x,i) \\ &= \int_{\partial G}\phi(\mathrm{d}x)\mathbb{E}_x\int_0^{\eta_2} u(X(t),\alpha(t))X(t)\,\mathrm{d}t \\ &= \sum_{i=1}^{m}\int_{\mathbb{R}} \mathbb{E}_{x,i}u(X(t),\alpha(t))X(t)\tilde{\nu}(\mathrm{d}x,i)\,. \end{aligned} \tag{20}$$

Hence, the desired stationary distribution is defined by the normalised measure as

$$\nu(H,i) = \frac{\tilde{\nu}(H,i)}{\sum_{j=1}^{m}\tilde{\nu}(\mathbb{R},j)}, \quad \forall i \in \mathcal{M}\,. \tag{21}$$

Now we will prove (17). Regarding the stationary distribution, we know that starting from an arbitrary point (x, i) with arbitrary initial distribution is asymptotically equivalent to starting with an initial distribution that is the stationary distribution. Then we will only need to verify the case when the initial distribution is the stationary distribution of the Markov chain Z_n. Hence, for any $H \in \mathscr{B}(\partial G)$

$$\mathbb{P}\{(X(0), \alpha(0)) \in (H \times \{i\})\} = \phi(H)\,.$$

Take a look at the sequence of random variables

$$\rho_n = \int_{\eta_{2n}}^{\eta_{2n+2}} u(X(t),\alpha(t))X(t)\,\mathrm{d}t\,.$$

From (19) and (20), we have

$$\mathbb{E}\rho_n = \sum_{i=1}^{m}\int_{\mathbb{R}} u(x,i)x\tilde{\nu}(\mathrm{d}x,i), \tag{22}$$

for all $n = 0, 1, 2, \ldots$. Let $\varphi(T)$ denote the number of cycles completed up to time T. Then

$$\varphi(T) := \max\left\{n \in \mathbb{N}\colon \sum_{k=1}^{n}(\eta_{2k} - \eta_{2k-2}) \le T\right\}.$$

Hence, $\int_0^T u(X(t),\alpha(t))X(t)\,\mathrm{d}t$ can be decomposed as

$$\begin{aligned} &\int_0^T u(X(t),\alpha(t))X(t)\,\mathrm{d}t \\ &= \sum_{n=0}^{\varphi(T)}\rho_n + \int_{\eta_{2\varphi(T)}}^T u(X(t),\alpha(t))X(t)\,\mathrm{d}t\,. \end{aligned}$$

Note that both $u(\cdot)$ and $X(\cdot)$ are non-negative, we have

$$\sum_{n=0}^{\varphi(T)}\rho_n \le \int_0^T u(X(t),\alpha(t))X(t)\,\mathrm{d}t \le \sum_{n=0}^{\varphi(T)+1}\rho_n\,.$$

Then

$$\begin{aligned} \frac{1}{\varphi(T)}\sum_{n=0}^{\varphi(T)}\rho_n &\le \frac{1}{\varphi(T)}\int_0^T u(X(t),\alpha(t))X(t)\,\mathrm{d}t \\ &\le \frac{1}{\varphi(T)}\sum_{n=0}^{\phi(T)+1}\rho_n\,. \end{aligned}$$

As $T \to \infty$, $\varphi(T) \to \infty$. Combining with (22), we have

$$\frac{1}{\varphi(T)}\int_0^T u(X(t),\alpha(t))X(t)\,\mathrm{d}t \to \sum_{i=1}^{m}\int_{\mathbb{R}} u(x,i)x\tilde{\nu}(\mathrm{d}x,i)\,. \tag{23}$$

On the other hand, the law of large numbers implies

$$\mathbb{P}\left\{\frac{1}{n}\sum_{k=0}^{n}\rho_k \to \sum_{i=1}^{m}\int_{\mathbb{R}} u(x,i)x\tilde{\nu}(\mathrm{d}x,i),\ \text{as}\, n \to \infty\right\} = 1. \tag{24}$$

Particularly, when $u(x, i) = 1/x$, (24) implies

$$\mathbb{P}\left\{\frac{\eta_{2n+2}}{n} \to \sum_{i=1}^{m}\tilde{\nu}(\mathrm{d}x,i),\ \text{as}\, n \to \infty\right\} = 1. \tag{25}$$

Since $\eta_{2n} \le T \le \eta_{2n+2}$ and $\eta_{2n}/\eta_{2n+2} \to 1$ almost surely as $T \to \infty$, we have

$$\mathbb{P}\left\{\frac{T}{\varphi(T)} \to \sum_{i=1}^{m}\tilde{\nu}(\mathrm{d}x,i),\ \text{as}\, T \to \infty\right\} = 1. \tag{26}$$

Now, using (23) and (26), we have as $T \to \infty$

$$\begin{aligned} &\frac{1}{T}\int_0^T u(X(t),\alpha(t))X(t)\,\mathrm{d}t \\ &= \frac{\int_0^T u(X(t),\alpha(t))X(t)\,\mathrm{d}t}{\varphi(T)}\frac{\varphi(T)}{T} \\ &\to \sum_{i=1}^{m}\int_{\mathbb{R}} u(x,i)x\tilde{\nu}(\mathrm{d}x,i) \times \frac{1}{\sum_{i=1}^{m}\tilde{\nu}(\mathrm{d}x,i)} \\ &= \sum_{i=1}^{m}\int_{\mathbb{R}} u(x,i)x\nu(\mathrm{d}x,i)\ \text{almost surely.} \end{aligned} \tag{27}$$

Hence

$$\begin{aligned} \mathbb{P}\Bigg(&\lim_{T\to\infty}\frac{1}{T}\int_0^T u(X(t),\alpha(t))X(t)\,\mathrm{d}t \\ &= \sum_{i=1}^{m}\int_{\mathbb{R}} u(x,i)x\nu(\mathrm{d}x,i),\ \text{as}\, T \to \infty\Bigg) = 1. \end{aligned} \tag{28}$$

Since (28) holds for any $(x, i) \in G \times \mathcal{M}$, then

© The Institution of Engineering and Technology 2018

$$\mathbb{P}_{x,i}\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T u(X(t),\alpha(t))X(t)\,\mathrm{d}t = \sum_{i=1}^{m}\int_{\mathbb{R}} u(x,i)x\nu(\mathrm{d}x,i),\ \text{as}\ T\to\infty\right) = 1. \tag{29}$$

Note that $\theta(\cdot,\cdot)$ is the stationary density associated with the stationary distribution $\nu(\cdot,\cdot)$, (29) can be written as

$$\mathbb{P}_{x,i}\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T u(X(t),\alpha(t))X(t)\,\mathrm{d}t = \sum_{i=1}^{m}\int_{\mathbb{R}} u(x,i)x\theta(x,i)\,\mathrm{d}x\right) = 1. \tag{30}$$

Thus, (17) and (18) hold. □

4 Dynamic programming equation

We have constructed the stationary distribution $\nu(\cdot,\cdot)$. However, it is generally not easy to approximate the invariant measure. To obtain the optimal ergodic control of dividend payment, we will refer to the dynamic programming equation in (14). To solve for (14), we will design a two-component Markov chain to approximate the state process. Then we will rewrite (14) by a dynamic programming equation with a Markov chain with transition probabilities.

Before we write the dynamic programming equations, let us recall some results of Markov chains. By using the ergodic theorem for Markov chains in [26, 27], we can find an auxiliary function $W(x,i,u)$ such that the pair $(W(x,i,u),\gamma(u))$ satisfies

$$W(x,i,u) = \sum_y p((x,i),(y,j)|u)W(y,i,u) + u(x,i)x - \gamma(u), \tag{31}$$

for each feedback control $u(\cdot)$, where $p((x,i),(y,j)|u)$ is the transition probability from a state (x,i) to another state (y,j) under the control $u(\cdot)$.

Define $\bar{\gamma} = \max_u \gamma(u)$, where $u(\cdot)\in\mathscr{U}$. Then there is an auxiliary function $V(x,i)$ such that the pair $(V(x,i),\bar{\gamma})$ satisfies the dynamic programming equation:

$$V(x,i) = \max_{u\in\mathscr{U}}\left\{\sum_y p((x,i),(y,j)|u)V(y,i) + u(x,i)x - \bar{\gamma}\right\}. \tag{32}$$

To keep $V(x,i)$ from blowing up, (32) can be written in a centred form as follows:

$$V(x,i) = \max_{u\in\mathscr{U}}\left\{\sum_y p((x,i),(y,j)|u)\widetilde{V}(y,i) + u(x,i)x\right\}, \tag{33}$$

where

$$\widetilde{V}(y,i) = V(y,i) - V(x_0,i).$$

x_0 is determined such that $\bar{\gamma} = V(x_0,i)$.

4.1 Boundary conditions

For the purpose of numerical analysis, it is always necessary to consider a compact state space. In our problem, the surplus could potentially grow to an arbitrary high level. Our control variable is the dividend payment strategies. When surplus is too high, it is optimal to pay the dividend according to our objective function. Furthermore, the domain of the surplus process is compactified for the computation purpose where a large enough right boundary B was imposed. To be consistent with the reality, it is natural to set a reflecting boundary on the right side. For the left side, the surplus follows a log-normal distribution and is always positive. We will also choose a reflecting boundary on the left side for the computation purpose. Hence, for the boundaries, $V(x,i)$ follows:

$$V_x(x,i) = 0. \tag{34}$$

5 Numerical algorithm

Our objective is to develop a numerical scheme to approximate $\bar{\gamma}$ in (12). In what follows, Section 5.1 will construct an approximating Markov chain in the state space. The discretion of the dynamic programming equation is presented in Section 5.2 and the transition probability of the approximating Markov chain is derived.

5.1 Approximating Markov chain

In this section, we will construct a locally consistent discrete-time Markov chain to approximate the controlled regime-switching diffusion system. The constructed Markov chain will be locally consistent with (9). In (9), the process state has two components x and α. Applying the methodology in [23], our approximating Markov chain will include two components: one component is to approximate the diffusive part; the other component delineates the market regimes.

Let $h>0$ be the step size. Define $S_h' = \{x: x = kh, k = 0, \pm 1, \pm 2, \ldots\}$ and $S_h = S_h' \cap G_h$, where $G_h = (0, B+h)$. B is a large value representing the upper bound for computation purpose. Furthermore, without loss of generality, assume that the upper bound B is an integer multiple of h. Let $\{(\xi_n^h,\alpha_n^h), n<\infty\}$ be a controlled discrete-time Markov chain on $S_h\times\mathscr{M}$. The transition probability from one state (x,i) to another state (y,j) under the control u^h is denoted as $p^h((x,i),(y,j)|u^h)$. p^h should be well defined later so that the constructed Markov chain's evolution is able to approximate the local behaviour of the controlled regime-switching diffusion process (9).

We proceed as follows. At a discrete time n, we can either pay a dividend payment as a regular control or have a reflection on the boundary. Thus, if we let $\Delta\xi_n^h = \xi_{n+1}^h - \xi_n^h$, then

$$\Delta\xi_n^h = \Delta\xi_n^h I_{\{\text{dividend payment at } n\}} + \Delta\xi_n^h I_{\{\text{reflection step on the left at } n\}} + \Delta\xi_n^h I_{\{\text{reflection step on the right at } n\}}. \tag{35}$$

The constructed Markov chain and control will be carefully chosen so that only one term in (35) is non-zero. Let $\left\{I_n^h: n = 0,1,\ldots\right\}$ be a sequence of control actions, where $I_n^h = 0$, 1, or 2, if we exercise a dividend payment, or reflect on the left or right boundaries at time n.

If $I_n^h = 0$, then we denote by $u_n^h \subset U$ the random variable that is the dividend payment action for the chain at time n. Let $\tilde{\Delta}t^h(\cdot,\cdot,\cdot)>0$ be the interpolation interval on $S_h\times\mathscr{M}\times U$. Assume $\inf_{x,i,u}\tilde{\Delta}t^h(x,i,u)>0$ for each $h>0$ and $\lim_{h\to 0}\sup_{x,i,u}\tilde{\Delta}t^h(x,i,u)\to 0$. If $I_n^h = 1$, or $\xi_n^h = 0$, reflection step on the left boundary is exerted definitely. We assume that the reflection step on the left side is from 0 to h. Hence, denote by Δz_n^h the left reflection size at time n, then we have $\Delta\xi_n^h = \Delta z_n^h = h$. If $I_n^h = 2$, or $\xi_n^h = B+h$, reflection step on the right boundary is exerted definitely. We assume that the reflection step on the right side is from $B+h$ to B. Hence, denote by Δg_n^h the right reflection size at time n, then we have $\Delta\xi_n^h = -\Delta g_n^h = -h$.

Let $\mathbb{E}_{x,i,n}^{u,h,0}$, $\mathrm{Var}_{x,i,n}^{u,h,0}$ and $\mathbb{P}_{x,i,n}^{u,h,0}$ denote the conditional expectation, variance and marginal probability given $\{\xi_k^h,\alpha_k^h,u_k^h,I_k^h,k\le n,\xi_n^h = x,\alpha_n^h = i,I_n^h = 0,u_n^h = u\}$, respectively. The sequence $\{(\xi_n^h,\alpha_n^h)\}$ is said to be *locally consistent*, if it satisfies

© The Institution of Engineering and Technology 2018

$$\begin{aligned}
&\mathbb{E}^{u,h,0}_{x,i,n}[\Delta\xi^h_n] = x[\pi(\beta(i) - h(\lambda) - \lambda c(i) + \mu(i)) + \mu(i) \\
&\qquad -u]\tilde{\Delta}t^h(x,i,u) + o(\tilde{\Delta}t^h(x,i,u)), \\
&\mathrm{Var}^{u,h,0}_{x,i,n}(\Delta\xi^h_n) = (\pi+1)^2\sigma^2(i)x^2\tilde{\Delta}t^h(x,i,u) \\
&\qquad +o(\tilde{\Delta}t^h(x,i,u)), \\
&\mathbb{P}^{u,h,0}_{x,i,n}\{\alpha^h_{n+1} = j\} = q_{ij}\tilde{\Delta}t^h(x,i,u) \\
&\qquad +o(\tilde{\Delta}t^h(x,i,u)),\ \text{for } j \neq i, \\
&\mathbb{P}^{u,h,0}_{x,i,n}\{\alpha^h_{n+1} = i\} = 1 + q_{ii}\tilde{\Delta}t^h(x,i,u) + o(\tilde{\Delta}t^h(x,i,u)). \\
&\sup_{n,\omega\in\Omega} |\Delta\xi^h_n| \to 0 \text{ as } h \to 0.
\end{aligned} \tag{36}$$

We require the reflections to be 'impulsive' or 'instantaneous' when $I^h_n = 1$ and $I^h_n = 2$. That is, the interpolation interval on $S_h \times \mathscr{M} \times U \times \{0,1,2\}$ is

$$\begin{aligned}
&\Delta t^h(x,i,u,\bar{i}) = \tilde{\Delta}t^h(x,i,u)I_{\{\bar{i}=0\}}, \\
&\quad \text{for any } (x,i,u,\bar{i}) \in S_h \times \mathscr{M} \times U \times \{0,1,2\}.
\end{aligned} \tag{37}$$

The sequence u^h is said to be admissible if u^h_n is $\sigma\left\{(\xi^h_0,\alpha^h_0),\dots,(\xi^h_n,\alpha^h_n),u^h_0,\dots,u^h_{n-1}\right\}$-adapted and for any $\mathscr{E} \in \mathscr{B}(S_h \times \mathscr{M})$, we have

$$\begin{aligned}
&\mathbb{P}\left\{(\xi^h_{n+1},\alpha^h_{n+1}) \in \mathscr{E} \,|\, \sigma\{(\xi^h_0,\alpha^h_0),\dots,(\xi^h_n,\alpha^h_n),u^h_0,\dots,u^h_n\}\right\} \\
&\quad = p^h((\xi^h_n,\alpha^h_n),\mathscr{E}\,|\,u^h_n),
\end{aligned}$$

$$\begin{aligned}
&\mathbb{P}\left\{(\xi^h_{n+1},\alpha^h_{n+1}) = (h,i) \,|\, (\xi^h_n,\alpha^h_n) = (0,i),\right. \\
&\quad \left.\sigma\{(\xi^h_0,\alpha^h_0),\dots,(\xi^h_n,\alpha^h_n),u^h_0,\dots,u^h_n\}\right\} = 1,
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{P}\left\{(\xi^h_{n+1},\alpha^h_{n+1}) = (B,i) \,|\, (\xi^h_n,\alpha^h_n) = (B+h,i),\right. \\
&\quad \left.\sigma\{(\xi^h_0,\alpha^h_0),\dots,(\xi^h_n,\alpha^h_n),u^h_0,\dots,u^h_n\}\right\} = 1.
\end{aligned}$$

Put

$$\begin{aligned}
&t^h_0 := 0,\ t^h_n := \sum_{k=0}^{n-1} \Delta t^h(\xi^h_k,\alpha^h_k,u^h_k,I^h_k), \\
&\Delta t^h_k = \Delta t^h(\xi^h_k,\alpha^h_k,u^h_k,I^h_k),\ \text{and } n^h(t) := \max\left\{n : t^h_n \le t\right\}.
\end{aligned}$$

Then the piecewise constant interpolations, denoted by $(\xi^h(\cdot),\alpha^h(\cdot))$, $u^h(\cdot)$, $z^h(\cdot)$, and $g^h(\cdot)$, are defined accordingly as

$$\begin{aligned}
&\xi^h(t) = \xi^h_n,\ \alpha^h(t) = \alpha^h_n,\ u^h(t) = u^h_n = u(\xi^h_n), \\
&z^h(t) = \sum_{k \le n^h(t)} \Delta z^h_k I_{\{I^h_k = 1\}},\ g^h(t) = \sum_{k \le n^h(t)} \Delta g^h_k I_{\{I^h_k = 2\}}
\end{aligned} \tag{38}$$

for $t \in [t^h_n, t^h_{n+1})$. Let $(\xi^h_0,\alpha^h_0) = (x,i) \in S_h \times \mathscr{M}$ and u^h be an admissible strategy. Then the cost function for the controlled Markov chain follows:

$$J^h_B(x,i,u) = \limsup_n \frac{\mathbb{E}_{x,i}\sum_{k=1}^{n-1} u^h_k \xi^h_k \Delta t^h_k}{\mathbb{E}_{x,i}\sum_{k=1}^{n-1} \Delta t^h_k}, \tag{39}$$

which is analogous to (11) regarding the definition of interpolation intervals in (37). Since $J^h_B(x,i,u)$ does not depend on the initial condition (x, i), we write it as $\gamma^h(u)$. Likewise, we denote

$$\bar{\gamma}^h = \sup_{u^h \text{ admissible}} \gamma^h(u). \tag{40}$$

Note that we are considering feedback controls $u(\cdot)$ here. Similarly to ν in (21), let $\nu^h(u) = \nu^h(x,u), x \in S_h$ denote the associate invariant measure in the approximating space. Then $\gamma^h(u)$ can be rewritten as

$$\begin{aligned}
\gamma^h(u) &= \limsup_n \frac{\mathbb{E}_{x,i}\sum_{k=1}^{n-1} u^h_k \xi^h_k \Delta t^h_k}{\mathbb{E}_{x,i}\sum_{k=1}^{n-1} \Delta t^h_k} \\
&= \frac{\sum_{x,i} u(x,i)x\Delta t^h(x,i,u(x,i),0)\nu^h(x,u)}{\sum_{x,i} \Delta t^h(x,i,u(x,i),0)\nu^h(x,u)}.
\end{aligned} \tag{41}$$

Since, the time interval of the approximating Markov chain $\Delta t^h(x,i,u(x,i),0)$ depends on x and u, the invariant measure for the approximating Markov chain needs considering the time spent on each state of the interpolated process. Then, we define a new measure $\omega^h(u) = \omega^h(x,i,u), x \in S_h$ such that

$$\omega^h(x,i,u) = \frac{\Delta t^h(x,i,u(x,i),0)\nu^h(x,i)}{\sum_{x,i} \Delta t^h(x,i,u(x,i),0)\nu^h(x,i)}. \tag{42}$$

Hence, $\gamma^h(u)$ can be written in a simple form as

$$\gamma^h(u) = \sum_x u(x,i)x\omega^h(x,i,u). \tag{43}$$

Let $\mathbb{E}^u_{\omega^h(u)}$ be the expectation for the stationary process under control $u(\cdot)$. In view of (41), (43) can also be written as

$$\gamma^h(u) = \mathbb{E}^u_{\omega^h(u)} \int_0^1 u(\xi^h(s))\xi^h(s)\,\mathrm{d}s. \tag{44}$$

Remark 1: Practically, it is much more difficult to calculate the invariant measure $\omega^h(u)$ than to calculate the summation $\sum^x u(x,i)x\omega^h(x,i,u)$. By using the iteration method, the convergence speed for computing the value of $\gamma^h(u)$ is much faster than that for computing the invariant measure $\omega^h(u)$. Hence, we focus on the converge of the state process and objective functions instead of the invariant measure itself.

We will show that $V^h(x,i)$ satisfies the dynamic programming equation as follows:

$$V^h(x,i) = \begin{cases} p^h((x,i),(y,j)|u)V^h(y,j) \\ +(ux - \gamma^h)\Delta t^h(x,i,u,0),\ \text{for } x \in S_h, \\ p^h((x,i),(y,j)|u)V^h(y,j),\ \text{for } x \in \partial S_h. \end{cases} \tag{45}$$

In actual computing, we will use value iteration or policy iteration in an enlarged state space due to the presence of regime-switching comparing the work in [23].

5.2 Discretisation

Define the approximation to the first and the second derivatives of $V(\cdot,i)$ by the finite difference method in [24] using stepsize $h > 0$ as:

$$\begin{aligned}
&V(x,i) \to V^h(x,i) \\
&V_x(x,i) \to \frac{V^h(x+h,i) - V^h(x,i)}{h} \\
&\qquad \text{for } x(\pi(\beta(i) - h(\lambda) - \lambda c(i) + \mu(i)) + \mu(i) - u) > 0, \\
&V_x(x,i) \to \frac{V^h(x,i) - V^h(x-h,i)}{h} \\
&\qquad \text{for } x(\pi(\beta(i) - h(\lambda) - \lambda c(i) + \mu(i)) + \mu(i) - u) < 0, \\
&V_{xx}(x,i) \to \frac{V^h(x+h,i) - 2V^h(x,i) + V^h(x-h,i)}{h^2}.
\end{aligned} \tag{46}$$

© The Institution of Engineering and Technology 2018

It leads to, $\forall x \in S_h, i \in \mathscr{M}$,

$$\max_{u \in U} \left\{ \frac{V^h(x+h,i) - V^h(x,i)}{h} \left[x(\pi(\beta(i) - h(\lambda) - \lambda c(i) + \mu(i)) + \mu(i) - u) \right]^+ - \frac{V^h(x,i) - V^h(x-h,i)}{h} \left[x(\pi(\beta(i) - h(\lambda) - \lambda c(i) + \mu(i)) + \mu(i) - u) \right]^- + \frac{(\pi+1)^2 \sigma^2(i) x^2}{2} \frac{V^h(x+h,i) - 2V^h(x,i) + V^h(x-h,i)}{h^2} + \sum_j V^h(x,\cdot) q_{ij} + ux - \bar{\gamma} \right\} = 0, \tag{47}$$

where $[x(\pi(\beta(i) - h(\lambda) - \lambda c(i) + \mu(i)) + \mu(i) - u)]^+$ and $[x(\pi(\beta(i) - h(\lambda) - \lambda c(i) + \mu(i)) + \mu(i) - u)]^-$ are the positive and negative parts of $[x(\pi(\beta(i) - h(\lambda) - \lambda c(i) + \mu(i)) + \mu(i) - u)]$, respectively.

For the reflecting boundaries, we choose

$$V_x(x,i) \to \frac{V^h(x,i) - V^h(x-h,i)}{h}. \tag{48}$$

That is, the process will be reflected left into the domain of the x. For simplicity, let

$$b(x,i,u) = x(\pi(\beta(i) - h(\lambda) - \lambda c(i) + \mu(i)) + \mu(i) - u),$$
$$\sigma(x,i,u) = (\pi+1)\sigma(i)x.$$

Comparing (47) and (48) with (45), we achieve the transition probabilities of $V^h(x,i)$ in the interior of the domain as the following:

$$\begin{aligned} &p^h((x,i),(x+h,i)|u) = \frac{\sigma(x,i,u)^2/2 + h[b(x,i,u)]^+}{D}, \\ &p^h((x,i),(x-h,i)|u) = \frac{\sigma(x,i,u)^2/2 + h[b(x,i,u)]^-}{D}, \\ &p^h((x,i),(x,j)|u) = \frac{q_{ij}h^2}{D}, \text{ for } i \neq j, \\ &p^h(\cdot) = 0, \text{ otherwise}, \\ &\Delta t^h(x,i,u,2) = \frac{h^2}{D}, \end{aligned} \tag{49}$$

with

$$D = \sigma(x,i,u)^2 + h|b(x,i,u)| - h^2 q_{ii}$$

being well defined. We also find the transition probability of $V^h(x,i)$ on the boundaries comparing with (45) as follows:

$$p^h((x,i),(x+h,i)|u) = 1, \quad \text{for } x = 0, \tag{50}$$

and

$$p^h((x,i),(x-h,i)|u) = 1, \quad \text{for } x = B. \tag{51}$$

6 Convergence of numerical approximation

This section deals with the asymptotic properties of the approximating Markov chain by using the techniques of weak convergence. Section 6.1 introduces the technique of time rescaling and the interpolation of the approximation sequences. Relax controls are defined in Section 6.2. Section 6.3 analyses the weak convergence of the sequence of the rescaled processes $\{\hat{\xi}^h(\cdot), \hat{\alpha}^h(\cdot), \hat{m}^h(\cdot), \hat{w}^h(\cdot), \hat{z}^h(\cdot), \hat{g}^h(\cdot), \hat{T}^h(\cdot)\}$. Section 6.4 establishes the convergence of the optimal value.

6.1 Interpolation and rescaling

According to the approximating Markov chain defined in the last section, we obtain the piecewise constant interpolation and choose an appropriate interpolation interval level. In view of (38), the continuous-time interpolations $(\xi^h(\cdot), \alpha^h(\cdot))$, $u^h(\cdot)$, $g^h(\cdot)$, and $z^h(\cdot)$ are defined. In addition, let $\mathscr{U}^h$ be the collection of strategies. $\mathscr{U}^h$ is determined by a sequence of measurable functions $F_n^h(\cdot)$. For $u_n^h \in \mathscr{U}^h$,

$$u_n^h = F_n^h(\xi_k^h, \alpha_k^h, k \le n; u_k^h, k \le n). \tag{52}$$

Define $\mathscr{D}_t^h$ as the smallest σ-algebra generated by $\{\xi^h(s), \alpha^h(s), u^h(s), g^h(s), z^h(s), s \le t\}$. $\mathscr{U}^h$ defined by (52) is equivalent to the set of all piecewise constant admissible strategies with respect to $\mathscr{D}_t^h$.

Using the notations of the regular controls, interpolations and reflection steps defined above, (35) yields

$$\begin{aligned} \xi^h(t) &= x + \sum_{k=0}^{n-1} [\mathbb{E}_k^h \Delta\xi_k^h + (\Delta\xi_k^h - \mathbb{E}_k^h \Delta\xi_k^h)] \\ &= x + \sum_{k=0}^{n-1} b(\xi_k^h, \alpha_k^h, u_k^h) \Delta t^h(\xi_k^h, \alpha_k^h, u_k^h, 0) \\ &\quad + \sum_{k=0}^{n-1} (\Delta\xi_k^h - \mathbb{E}_k^h \Delta\xi_k^h) + \varepsilon^h(t) \\ &= x + B^h(t) + M^h(t) + \varepsilon^h(t), \end{aligned} \tag{53}$$

where

$$B^h(t) = b(\xi_k^h, \alpha_k^h, u_k^h) \Delta t^h(\xi_k^h, \alpha_k^h, u_k^h, 0),$$
$$M^h(t) = \sum_{k=0}^{n-1} (\Delta\xi_k^h - \mathbb{E}_k^h \Delta\xi_k^h),$$

and $\varepsilon^h(t)$ is a negligible error satisfying

$$\lim_{h \to \infty} \sup_{0 \le t \le T} E|\varepsilon^h(t)|^2 \to 0 \text{ for any } 0 < T < \infty. \tag{54}$$

Also, $M^h(t)$ is a martingale with respect to $\mathscr{D}_t^h$, and its discontinuity goes to zero as $h \to 0$. We attempt to represent $M^h(t)$ similar to the diffusion term in (9). Define $w^h(\cdot)$ as

$$\begin{aligned} w^h(t) &= \sum_{k=0}^{n-1} (\Delta\xi_k^h - \mathbb{E}_k^h \Delta\xi_k^h)/\sigma(\xi_k^h, \alpha_k^h, u_k^h), \\ &= \int_0^t \sigma^{-1}(\xi^h(s), \alpha^h(s), u^h(s))\, \mathrm{d}M^h(s). \end{aligned} \tag{55}$$

We can now rewrite (53) as

$$\begin{aligned} \xi^h(t) = x &+ \int_0^t b(\xi^h(s), \alpha^h(s), u^h(s))\, \mathrm{d}s \\ &+ \int_0^t \sigma(\xi^h(s), \alpha^h(s), u^h(s))\, \mathrm{d}w^h(s) + \varepsilon^h(t). \end{aligned} \tag{56}$$

Now we introduce the procedures of rescaling. The technique of time-rescaling is to 'stretch out' the control and state processes to make them smoother. Then the tightness of $g^h(\cdot)$ and $z^h(\cdot)$ can be proved. Define $\Delta\hat{t}_n^h$ by

$$\Delta\hat{t}_n^h = \begin{cases} \Delta t^h & \text{for a diffusion on step } n, \\ |\Delta z_n^h| = h & \text{for a left reflection on step } n, \\ |\Delta g_n^h| = h & \text{for a right reflection on step } n. \end{cases} \tag{57}$$

© The Institution of Engineering and Technology 2018

Define $\hat{T}^h(\cdot)$ by

$$\hat{T}^h(t) = \sum_{i=0}^{n-1} \Delta t^h = t_n^h, \quad \text{for } t \in [\hat{T}_n^h, \hat{t}_{n+1}^h] .$$

Then $\hat{T}^h(\cdot)$ will increase with unit slop when a regular control is exerted. Moreover, we define the rescaled and interpolated process $\hat{\xi}^h(t) = \xi^h(\hat{T}^h(t))$. $\hat{\alpha}^h(t)$, $\hat{u}^h(t)$, $\hat{g}^h(t)$ are defined similarly. The time scale is stretched out by h at the left and right reflection steps. Then the stretched process can be written as follows:

$$\begin{aligned}\hat{\xi}^h(t) = x &+ \int_0^t b(\hat{\xi}^h(s), \hat{\alpha}^h(s), \hat{u}^h(s))\,\mathrm{d}s \\ &+ \int_0^t \sigma(\hat{\xi}^h(s), \hat{\alpha}^h(s), \hat{u}^h(s))\mathrm{d}w^h(s) + \varepsilon^h(t) .\end{aligned} \tag{58}$$

6.2 Relaxed controls

Let $\mathscr{B}(U \times [0, \infty))$ be the σ-algebra of Borel subsets of $U \times [0, \infty)$. An *admissible relaxed control* $m(\cdot)$ is a measure on $\mathscr{B}(U \times [0, \infty))$ such that $m(U \times [0, t]) = t$ for each $t \geq 0$. Given a relaxed control $m(\cdot)$, there is an $m_t(\cdot)$ such that $m(\mathrm{d}\phi\, \mathrm{d}t) = m_t(\mathrm{d}\phi)\, \mathrm{d}t$. Given a relaxed control $m(\cdot)$ of $u^h(\cdot)$, we define the derivative $m_t(\cdot)$ such that

$$m^h(H) = \int_{U \times [0, \infty)} I_{\{(u^h, t) \in H\}} m_t(\mathrm{d}\phi)\, \mathrm{d}t \tag{59}$$

for all $H \in \mathscr{B}(U \times [0, \infty))$. For each t, $m_t(\cdot)$ is a measure on $\mathscr{B}(U)$ satisfying $m_t(U) = 1$. Then $m_t(\cdot)$ can be defined as the left-hand derivative for $t > 0$,

$$m_t(O) = \lim_{\delta \to 0} \frac{m(O \times [t - \delta, t])}{\delta}, \quad \forall O \in \mathscr{B}(U) . \tag{60}$$

The relaxed control representation $m^h(\cdot)$ of $u^h(\cdot)$ can be defined by

$$m_t^h(O) = I_{\{u^h(t) \in O\}}, \ \forall O \in \mathscr{B}(U) . \tag{61}$$

Denote by $\mathscr{F}_t^h$ a filtration. $\mathscr{F}_t^h$ is the minimal σ-algebra that measures

$$\{\xi^h(s), \alpha^h(\cdot), m_s^h(\cdot), w^h(s), z^h(s), g^h(s), s \leq t\} . \tag{62}$$

Let Γ^h be the set of admissible relaxed controls $m^h(\cdot)$ with respect to $(\alpha^h(\cdot), w^h(\cdot))$ such that $m_t^h(\cdot)$ is a fixed probability measure in the interval $[t_n^h, t_{n+1}^h)$ given $\mathscr{F}_t^h$. Then Γ^h is a larger control space containing $\mathscr{U}^h$. Referring to the stretched out time scale, we denote the rescaled relax control as $m_{\hat{T}^h(t)}(\mathrm{d}\psi)$. Define $M_t(O)$ and $M_t^h(\mathrm{d}\psi)$ by

$$\begin{aligned}M_t(O)\, \mathrm{d}t &= \mathrm{d}w(t) I_{u(t) \in O}, \ \forall O \in \mathscr{B}(U), \\ M_t^h(\mathrm{d}\psi)\, \mathrm{d}t &= \mathrm{d}w^h(t) I_{u^h(t) \in \mathscr{U}} .\end{aligned}$$

Similarly, we let

$$\hat{M}_{\hat{T}^h}^h(t)(\mathrm{d}\psi)\, \mathrm{d}\hat{T}^h(t) = \mathrm{d}\hat{w}^h(\hat{T}^h(t)) I_{u^h(\hat{T}^h(t)) \in \mathscr{U}} .$$

Taking into account the relaxed controls, we rewrite (56), (58), and (40) as

$$\begin{aligned}\xi^h(t) = \quad & x + \int_0^t \int_{\mathscr{U}} b(\xi^h(s), \alpha^h(s), \psi) m_s^h(\mathrm{d}\psi)\, \mathrm{d}s \\ &+ \int_0^t \int_{\mathscr{U}} \sigma(\xi^h(s), \alpha^h(s), \psi) M_s^h(\mathrm{d}\psi)\, \mathrm{d}s + \varepsilon^h(t),\end{aligned} \tag{63}$$

$$\begin{aligned}\hat{\xi}^h(t) = x &+ \int_0^t \int_{\mathscr{U}} b(\hat{\xi}^h(s), \hat{\alpha}^h(s), \psi) \hat{m}_{\hat{T}^h(s)}^h(\mathrm{d}\psi)\, \mathrm{d}\hat{T}^h(s) \\ &+ \int_0^t \int_{\mathscr{U}} \sigma(\hat{\xi}^h(s), \hat{\alpha}^h(s), \psi) \hat{M}_{\hat{T}^h(s)}(\mathrm{d}\psi)\, \mathrm{d}\hat{T}^h(s) + \varepsilon^h(t),\end{aligned} \tag{64}$$

and

$$\bar{\gamma}^h = \inf_{m^h \in \Gamma^h} \gamma^h(m^h) . \tag{65}$$

Now we define the existence and uniqueness of the solution in the weak sense.

Definition 1: By a weak solution of (63), we mean that there exists a probability space $(\Omega, \mathscr{F}, \{\mathscr{F}^t\}, P)$, and the sequence of processes $(x(\cdot), \alpha(\cdot), m(\cdot), w(\cdot))$ such that $w(\cdot)$ is a standard $\mathscr{F}^t$-Wiener process, $\alpha(\cdot)$ is a continuous-time Markov chain, $m(\cdot)$ is admissible with respect to $x(\cdot)$ is $\mathscr{F}^t$-adapted, and (63) is satisfied. For an initial condition (x, i), we say that the probability law of the admissible process $(\alpha(\cdot), m(\cdot), w(\cdot))$ determines the probability law of solution $(x(\cdot), \alpha(\cdot), m(\cdot), w(\cdot))$ to (63) by the weak sense uniqueness, irrespective of probability space.

In addition, we have one more assumption.

(A1) Let $u(\cdot)$ be an admissible ordinary control with respect to $w(\cdot)$ and $\alpha(\cdot)$. Assume that $u(\cdot)$ is a piecewise constant and takes values in a finite set. For each initial condition, there exists a solution, which is unique in the weak sense, to (63) where $m(\cdot)$ is the relaxed control representation of $u(\cdot)$.

6.3 Convergence of a sequence of surplus processes

In this section, we deal with the convergence of the approximation sequence to the regime-switching process and the surplus process. We will derive one lemma and three theorems, whose proof is provided in the Appendix.

Lemma 2: Using the transition probabilities $\{p^h(\cdot)\}$ defined in (49), the interpolated process of the constructed Markov chain $\{\hat{\alpha}^h(\cdot)\}$ converges weakly to $\hat{\alpha}(\cdot)$, the Markov chain with generator $Q = (q_{\ell\iota})$.

Theorem 2: Let the approximating chain $\{\xi_n^h, \alpha_n^h, n < \infty\}$ constructed with transition probabilities defined in (49) be locally consistent with (9), $m^h(\cdot)$ be the relaxed control representation of $\{u_n^h, n < \infty\}$, $(\xi^h(\cdot), \alpha^h(\cdot))$ be the continuous-time interpolation defined in (38), and $\{\hat{\xi}^h(\cdot), \hat{\alpha}^h(\cdot), \hat{m}^h(\cdot), \hat{w}^h(\cdot), \hat{z}^h(\cdot), \hat{g}^h(\cdot), \hat{T}^h(\cdot)\}$ be the corresponding rescaled processes. Then $\{\hat{\xi}^h(\cdot), \hat{\alpha}^h(\cdot), \hat{m}^h(\cdot), \hat{w}^h(\cdot), \hat{z}^h(\cdot), \hat{g}^h(\cdot), \hat{T}^h(\cdot)\}$ is tight.

Theorem 3: Let $\{\hat{x}(\cdot), \hat{\alpha}(\cdot), \hat{m}(\cdot), \hat{w}(\cdot), \hat{z}(\cdot), \hat{g}(\cdot), \hat{T}(\cdot)\}$ be the limit of weakly convergent subsequence of $\{\hat{\xi}^h(\cdot), \hat{\alpha}^h(\cdot), \hat{m}^h(\cdot), \hat{w}^h(\cdot), \hat{z}^h(\cdot), \hat{g}^h(\cdot), \hat{T}^h(\cdot)\}$. $w(\cdot)$ is a standard $\mathscr{F}_t$-Wiener process and $m(\cdot)$ is admissible. Let $\hat{\mathscr{F}}_t$ be the σ-algebra generated by $\{\hat{\xi}^h(\cdot), \hat{\alpha}^h(\cdot), \hat{m}^h(\cdot), \hat{w}^h(\cdot), \hat{z}^h(\cdot), \hat{g}^h(\cdot), \hat{T}^h(\cdot)\}$. Then $\hat{w}(t) = w(\hat{T}(t))$ is an $\hat{\mathscr{F}}_t$-martingale with quadratic variation $\hat{T}(t)$. The limit process follows:

© The Institution of Engineering and Technology 2018

$$\hat{x}(t) = x + \int_0^t \int_{\mathcal{U}} b(\hat{x}(s), \hat{\alpha}(s), \psi)\hat{m}^h_{\hat{T}(s)}(\mathrm{d}\psi)\,\mathrm{d}\hat{T}(s) + \int_0^t \int_{\mathcal{U}} \sigma(\hat{x}(s), \hat{\alpha}(s), \psi)\hat{M}_{\hat{T}(s)}(\mathrm{d}\psi)\,\mathrm{d}\hat{T}(s)\,. \tag{66}$$

Theorem 4: For $t < \infty$, the inverse can be defined as

$$\mathcal{T}(t) = \inf\{s: \hat{T}(s) > t\}\,.$$

Then $\mathcal{T}(t)$ is right continuous. $\mathcal{T}(t) \to \infty$ as $t \to \infty$ w. p. 1. Define the rescaled process $\varphi(\cdot)$ by $\varphi(t) = \hat{\varphi}(T(t))$ for any process $\hat{\varphi}(\cdot)$. Then, $w(\cdot)$ is a standard $\mathcal{F}_t$-Wiener process and (9) holds.

6.4 Convergence of the optimal value

To prove the convergence of the optimal value of the objective function, we proceed to find a comparison ε-optimal control.

Lemma 3: For each $\varepsilon > 0$, there exists a continuous feedback control $u^{\varepsilon}(\cdot)$ that is ε-optimal to all admissible controls. The solution to (9) is unique in a weak sense and has a unique invariant measure under this ε-optimal control.

Proof: The existence of a smooth ε-optimal can be guaranteed by modifying the method in [28] for our formulation. □

Theorem 5: Assume the conditions of Theorems 3 and 4 are satisfied. Then as $h \to 0$

$$\bar{\gamma}^h(x, i) \to \bar{\gamma}\,. \tag{67}$$

Proof: First, to prove

$$\bar{\gamma}^h(x, i) \le \bar{\gamma}\,. \tag{68}$$

Let $\tilde{u}(\cdot)$ be the optimal control and $\tilde{m}^h(\cdot)$ be the relaxed control representation of $\tilde{u}^h(\cdot)$. Then, $\bar{\gamma}^h(x, i) = \gamma^h(\tilde{u}^h)$. Hence, in view of (44)

$$\begin{aligned}\bar{\gamma}^h(x, i) &= \mathbb{E}^{\tilde{u}^h}\int_0^1 u(\xi^h(s))\xi^h(s)\,\mathrm{d}s \\ &= \mathbb{E}^{\tilde{u}^h}\int_0^1 \int_{\mathcal{U}} \xi^h(s)\psi\tilde{m}^h_s(\mathrm{d}\psi)\,\mathrm{d}s \\ &\to \mathbb{E}^{\tilde{m}}\int_0^1 \int_{\mathcal{U}} x(s)\psi\tilde{m}_s(\mathrm{d}\psi)\,\mathrm{d}s \\ &= \lim_T \frac{1}{T}\mathbb{E}^{\tilde{m}}\int_0^T \int_{\mathcal{U}} \psi\tilde{m}_s(\mathrm{d}\psi)\,\mathrm{d}s \\ &= \gamma(\tilde{m}) \\ &\le \bar{\gamma},\end{aligned} \tag{69}$$

where $\gamma(\tilde{m})$ is the optimal value of the performance function for the limit stationary process.

On the other hand, from Lemma 3, we have ε-optimal control u^{ε} such that

$$\begin{aligned}\bar{\gamma}^h(x, i) &\ge \gamma^h(u^{\varepsilon}, i) \\ &= \mathbb{E}^{u^{\varepsilon}}\int_0^1 u^{\varepsilon}(\xi^h(s))\xi^h(s)\,\mathrm{d}s \\ &\to \mathbb{E}^{u^{\varepsilon}}\int_0^1 u^{\varepsilon}(x(s))x(s)\,\mathrm{d}s \\ &= \lim_T \frac{1}{T}\mathbb{E}^{u^{\varepsilon}}\int_0^T u^{\varepsilon}(x(s))x(s)\,\mathrm{d}s \\ &= \gamma(u^{\varepsilon}, i) \\ &\ge \bar{\gamma} - \varepsilon\,.\end{aligned} \tag{70}$$

Combining (69) and (70) yields (67). □

7 Numerical examples and further remarks

7.1 Numerical examples

We present a numerical example in this section. The regime-switching process is set to have two states for simplicity. By using the value iteration method, we solve the optimal strategies numerically.

According to the algorithm constructed in previous sections, we conduct the computation by using the method of value iteration as the following. For $n \in Z^+$ and $i \in \mathcal{M}$, define the vectors

$$\begin{aligned}V_n^h &= \{V_n^h(h, 1), V_n^h(2h, 1), \ldots, V_n^h(B, 1), \ldots V_n^h(h, n_0), \\ &\quad V_n^h(2h, m), \ldots, V_n^h(B, m)\}, \\ V^h &= \{V_n(h, 1), V_n(2h, 1), \ldots, V_n(B, 1), \ldots V_n(h, m), \\ &\quad V_n(2h, m), \ldots, V_n(B, m)\}\,.\end{aligned}$$

We obtain $V_n(x_0, i)$ at last. The numerical experiments demonstrate that $V_n(x_0, i) \to \bar{\gamma}$ as $n \to \infty$. The procedure is as follows.

1. Set $n = 0$. $\forall x \in S_h$ and $i \in \mathcal{M}$, we set the initial value $V_0^h(x, i) = \tilde{V}_0^h(x) = 0$.
2. Choose a x_0. Find improved values $V_{n+1}^h(x, i)$ by iteration and record the corresponding optimal control

$$V_{n+1}^h(x, i) = \max_{u \in \mathcal{U}} \Big[\sum_{(y, j)} (p^h((x, i), (y, j)) \mid u)\tilde{V}_n^h(y, i) + ux\Big],$$

$$\tilde{V}_n^h(x, i) = V_n^h(x, i) - V_n^h(x_0, i)\,.$$

3. If $|V_{n+1}^h - V_n^h| >$ tolerance, then $n \to n + 1$ and go to step 2; else the iteration stops.

The continuous-time finite-state Markov chain $\alpha(t)$ has the generator

$$Q = \begin{pmatrix}-10 & 10 \\ 800 & -800\end{pmatrix}$$

and $\mathcal{M} = \{1, 2\}$. The claim severity distribution follows an exponential distribution with density function $f(y) = ae^{-ay}$ where $a = 0.1$. The premium rate depends on the regimes with $\beta(1) = 0.02$ and $\beta(2) = 0.06$. The dividend rate $u(t)$ taking values in $[0, 1]$ is the control. Based on different market states, the yield rate of the asset is $\mu(1) = 0.1$ and $\mu(2) = 0.02$. The volatility of the financial market $\sigma(\alpha(t))$ is valued as $\sigma(1) = 0.05$ and $\sigma(2) = 0.1$. The claim rates in different regimes are set as $c(1) = 0.01$ and $c(2) = 0.1$. Hence, we are considering two insurance market modes to represent the insurance cycle. Market mode 1 represents a 'soft' market, where the investment return is high and the premium rate is low. While market mode 2 represents a 'hard' market, where the investment return is low and the premium rate is high. Obviously, it is much easier for insurance companies when the market is in mode 1. The insurance company is more likely to expand its business and write more policies. Then the liability ratio is higher. Hence, we set $\pi(1)$=0.8. For market mode 2, the insurance and financial market are much harder. The insurance company will preserve sufficient surplus and write less policies to pay for the future claims. Then the liability ratio is lower. Then, we set $\pi(1)$=0.2. To compute the optimal average dividend payment, we choose the value iteration and impose the upper bound of the computation interval of surplus as $B = 50$.

Furthermore, note that the x_0 is arbitrarily chosen to initiate the algorithm. Theoretically, different x_0 is supposed to lead to the same $\bar{\gamma}$. In practical computation, there are inevitable computational errors in calculating the optimal value of γ. The

© The Institution of Engineering and Technology 2018

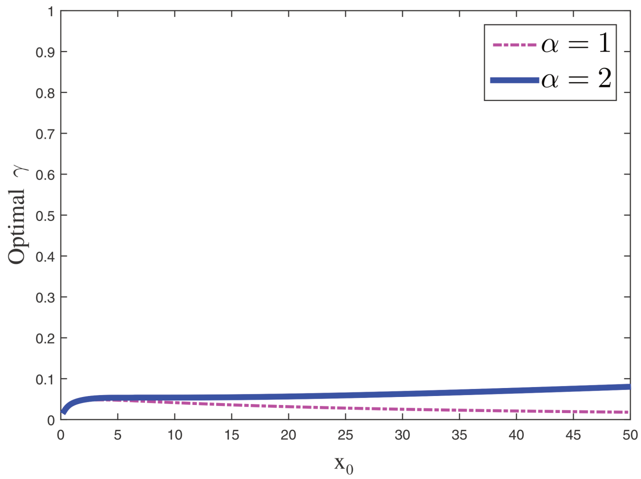

Fig. 1 *Optimal γ versus initial status*

optimal values are the average of convergent values of $V_n(x_0)$ in all available x_0's. To show the stability of the convergence with different x_0's, we plot the values $V_n(x_0)$ with respect to x_0 when the iteration stops in Fig. 1.

From Fig. 1, it is shown that $V_n(x_0)$ is fluctuating within the range $[0, 0.1]$. After a small hike when x_0 is small, the convergence value of $V_n(x_0)$ is flat and stable when x_0 is bigger. According to the stationary of the process, the average of the convergence value of $V_n(x_0)$ is a good approximation to the optimal long-term average dividend payment. The average of the values of blue dots is 0.06, and the average of the values of red dots is 0.03. The optimal ergodic control of dividend payment can be approximated by the mean as 0.045.

7.2 Further remarks

This work focused on finding the optimal ergodic dividend payment strategies of an insurance company with a long-term goal. The parameters in the model including premium rate, return rate of the assets, and claim rate, depend on the state of the economy, which is described by a finite-state continuous-time Markov chain. Considering the impact of reinsurance strategies on the surpluses of insurance companies, we aimed to maximise the long-run average dividend payment in an infinite time horizon. A generalised stationary diffusion process of surplus is presented. The invariant measure is constructed and the optimal value is obtained accordingly. By using the dynamic programming approach, we derive the associated system of HJB equations. Due to the regime-switching, approximating the invariant measure is very difficult. Then, we develop a numerical method to approximate the optimal ergodic dividend payment strategy directly. A two-component discrete-time controlled Markov chain is constructed to approximate the controlled regime-switching diffusion process. Convergence of the approximation algorithm is presented, and the optimal ergodic control is obtained in the numerical example under given parameter settings.

In the future study, the techniques of constructing invariant measures and approximating Markov chain can be extended to a variety of optimisation problems of risk-sensitive controls for ergodic processes, where the objective is to maximiee/minimise various performance functions over long term. Although the specific aim of this study is devoted to developing the optimal long-term insurance policies, the methods can be readily adopted to treat other optimal control problems with a long-run average aim and regime-switching diffusion formulation. The future effort may also be devoted to another variant of related game problems with long-term goal objective functions.

8 Acknowledgments

The research of Zhuo Jin and Hailiang Yang was supported in part by Research Grants Council of the Hong Kong Special Administrative Region (project No. HKU 17330816). The research of G. Yin was supported in part by the National Science Foundation under DMS-1710827.

9 References

[1] De Finetti, B.: 'Su unimpostazione alternativa della teoria collettiva del rischio', *Trans. XVth Int. Congress Actuaries*, 1957, **2**, pp. 433–443

[2] Guo, X., Liu, J., Zhou, X.Y.: 'A constrained non-linear regular-singular stochastic control problem, with applications', *Stoch. Process. Appl.*, 2004, **109**, pp. 167–187

[3] He, L., Liang, Z.: 'Optimal financing and dividend control of the insurance company with proportional reinsurance policy', *Insur. Math. Econ.*, 2008, **42**, pp. 976–983

[4] Løkka, A., Zervos, M.: 'Optimal dividend and issuance of equity policies in the presence of proportional costs', *Insur. Math. Econ.*, 2008, **42**, pp. 954–961

[5] Wei, F., Wu, L., Zhou, D.: 'Optimal control problem for an insurance surplus model with debt liability', *Math. Methods Appl. Sci.*, 2014, **37**, pp. 1652–1667

[6] Alvarez, L.H.R., Lempa, J.: 'On the optimal stochastic impulse control of linear diffusions', *SIAM J. Control Optim.*, 2008, **47**, pp. 703–732

[7] Bai, L., Paulsen, J.: 'On non-trivial barrier solutions of the dividend problem for a diffusion under constant and proportional transaction costs', *Stoch. Process. Appl.*, 2012, **122**, pp. 4005–4027

[8] Avram, F., Palmowski, Z., Pistorius, M.R.: 'On the optimal dividend problem for a spectrally negative Lévy process', *Ann. Appl. Prob.*, 2007, **17**, pp. 156–180

[9] Azcue, P., Muler, N.: 'Optimal investment policy and dividend payment strategy in an insurance company', *Ann. Appl. Prob.*, 2010, **20**, pp. 1253–1302

[10] Jin, Z., Yang, H., Yin, G.: 'Optimal debt ratio and dividend payment strategies with reinsurance', *Insur. Math. Econ.*, 2015, **64**, pp. 351–363

[11] Bielecki, T.R., Pliska, S.R.: 'Risk sensitive dynamic asset management', *Appl. Math. Optim.*, 1999, **39**, pp. 337–360

[12] Fleming, W.H., Sheu, S.J.: 'Risk-sensitive control and an optimal investment model', *Math. Finance*, 2000, **10**, pp. 197–213

[13] Pham, H.: 'A large deviations approach to optimal long term investment', *Finance Stochast.*, 2003, **7**, pp. 169–195

[14] Bielecki, T.R., Pliska, S.R.: 'Economic properties of the risk sensitive criterion for portfolio management', *Rev. Account. Finance*, 2003, **2**, pp. 3–17

[15] Fleming, W.H., McEneaney, W.M.: 'Risk-sensitive control on an infinite time horizon', *SIAM J. Control Optim.*, 1995, **33**, pp. 1881–1915

[16] Wei, J., Yang, H., Wang, R.: 'Classical and impulse control for the optimization of dividend and proportional reinsurance policies with regime switching', *J. Optim. Theory Appl.*, 2010, **147**, pp. 358–377

[17] Sotomayor, L., Cadenillas, A.: 'Classical, singular, and impulse stochastic control for the optimal dividend policy when there is regime switching', *Insur. Math. Econ.*, 2011, **48**, pp. 344–354

[18] Zhu, J.: 'Dividend optimization for a regime-switching diffusion model with restricted dividend rates', *ASTIN Bull.*, 2014, **44**, pp. 459–494

[19] Zhu, J., Chen, F.: 'Dividend optimization for regime-switching general diffusions', *Insur. Math. Econ.*, 2013, **53**, pp. 439–456

[20] Jin, Z., Yin, G., Yang, H.: 'Numerical methods for dividend policy of regime-switching jump-diffusion models', *Math. Control Relat. Fields*, 2011, **1**, pp. 21–40

[21] Jin, Z., Yin, G.: 'Numerical methods for optimal dividend payment and investment strategy for Markov-modulated jump diffusion models with regular and singular controls', *J. Optim. Theory Appl.*, 2013, **159**, pp. 246–271

[22] Yin, G., Zhu, C.: '*Hybrid switching diffusions: properties and applications*' (Springer, New York, 2010)

[23] Kushner, H., Dupuis, P.: '*Numerical methods for stochastic control problems in continuous time*' (Springer, New York, 2001, 2nd edn.)

[24] Jin, Z., Yang, H., Yin, G.: 'Numerical methods for optimal dividend payment and investment strategies of regime-switching jump diffusion models with capital injections', *Automatica*, 2013, **49**, pp. 2317–2329

[25] Fleming, W.H., Soner, H.: '*Controlled Markov processes and viscosity solutions*' (Springer-Verlag, New York, NY, 2006, 2nd edn.)

[26] Bertsekas, D.P.: '*Dynamic programming: deterministic and stochastic models*' (Prentice-Hall, Englewood Cliffs, NJ, 1987)

[27] Kushner, H.J.: '*Introduction to stochastic control theory*' (Holt Rinehart and Winston, New York, 1972)

[28] Kushner, H.J.: 'Optimality conditions for the average cost per unit time problem with a diffusion model', *SIAM J. Control Optim.*, 1978, **16**, pp. 330–346

[29] Yin, G., Zhang, Q., Badowski, G.: 'Discrete-time singularly perturbed Markov chains: aggregation, occupation measures, and switching diffusion limit', *Adv. Appl. Probab.*, 2003, **35**, pp. 449–476

10 Appendix

10.1 Proof of Lemma 2

Proof: We see that $\alpha^h(\cdot)$ is tight. The proof is similar to Theorem 3.1 in [29]. Then so is $\hat{\alpha}^h(\cdot)$ due to the rescaled time. □

© The Institution of Engineering and Technology 2018

10.2 Proof of Theorem 2

Proof: In view of Lemma 2, $\{\hat{\alpha}^h(\cdot)\}$ is tight. Since its range space is compact, the sequence $\{\hat{m}^h(\cdot)\}$ is tight. Let $T < \infty$, and let τ_h be an $\mathscr{F}_t$-stopping time which is not larger than T. Then for $\delta > 0$

$$\mathbb{E}^{u^h}_{\tau_h}(w^h(\tau_h + \delta) - w^h(\tau_h))^2 = \delta + \varepsilon_h, \tag{71}$$

where $\varepsilon_h \to 0$ uniformly in τ_h. Taking $\limsup_{h \to 0}$ followed by $\lim_{\delta \to 0}$ yield the tightness of $\{w^h(\cdot)\}$. Similar to the argument of $\alpha^h(\cdot)$, the tightness of $\hat{w}^h(\cdot)$ is obtained. Furthermore, following the definition of 'stretched out' timescale

$$|\hat{z}^h(\tau_h + \delta) - \hat{z}^h(\tau_h)| \le |\delta| + O(h),$$
$$|\hat{g}^h(\tau_h + \delta) - \hat{g}^h(\tau_h)| \le |\delta| + O(h).$$

Thus, $\{\hat{z}^h(\cdot), \hat{g}^h(\cdot)\}$ is tight. These results and the boundedness of $b(\cdot)$ imply the tightness of $\{\xi^h(\cdot)\}$. Thus, $\{\hat{\xi}^h(\cdot), \hat{\alpha}^h(\cdot), \hat{u}^h(\cdot), \hat{w}^h(\cdot), \hat{z}^h(\cdot), \hat{g}^h(\cdot), \hat{T}^h(\cdot)\}$ is tight.□

Since $\{\hat{x}^h(\cdot), \hat{\alpha}^h(\cdot), \hat{m}^h(\cdot), \hat{w}^h(\cdot), \hat{z}^h(\cdot), \hat{g}^h(\cdot), \hat{T}^h(\cdot)\}$ is tight, A weakly convergent subsequence can be extracted and denoted by $\{\hat{\xi}(\cdot), \hat{\alpha}(\cdot), \hat{m}(\cdot), \hat{w}(\cdot), \hat{z}(\cdot), \hat{g}(\cdot), \hat{T}(\cdot)\}$. Also, the paths of $\{\hat{x}(\cdot), \hat{\alpha}(\cdot), \hat{m}(\cdot), \hat{w}(\cdot), \hat{z}(\cdot), \hat{g}(\cdot), \hat{T}(\cdot)\}$ are continuous w. p. 1. □

10.3 Proof of Theorem 3

Proof: The proof is similar to Theorem 4.4 in [24], thus we omit it here. The readers may refer to [24] or related references. □

10.4 Proof of Theorem 4

Proof: Since $\hat{T}(t) \to \infty$ w.p. 1 as $t \to \infty$, $\mathscr{T}(t)$ exists for all t and $\mathscr{T}(t) \to \infty$ as $t \to \infty$ w. p. 1

$$\mathbb{E}S(\xi^h(t_k), \alpha^h(t_k), w^h(t_k), (\psi_j, m^h)_{t_k}, z^h(t_k), g^h(t_k), j \le q,$$
$$k \le p)[w(t+s) - w(t)] = 0.$$

$$\mathbb{E}S(\xi^h(t_k), \alpha^h(t_k), w^h(t_k), (\psi_j, m^h)_{t_k}, z^h(t_k), g^h(t_k), j \le q,$$
$$k \le p)[w^2(t+\delta) - w^2(t) - (\mathscr{T}(t+s) - \mathscr{T}(t))] = 0.$$

Thus, we can verify $w(\cdot)$ is an $\mathscr{F}_t$-Wiener process. A rescaling of (66) yields

$$\begin{aligned} x(t) = \quad & x + \int_0^t \int_{\mathscr{U}} b(x(s), \alpha(s), \psi) m_s(\mathrm{d}\psi)\, \mathrm{d}s \\ & + \int_0^t \int_{\mathscr{U}} \sigma(x(s), \alpha(s), \psi) M_s(\mathrm{d}\psi)\, \mathrm{d}s . \end{aligned} \tag{72}$$

In other words, (9) holds. □

© The Institution of Engineering and Technology 2018