Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 9(12): 34674–34704.
DOI: 10.3934/math.20241651
Received: 14 October 2024
Revised: 04 December 2024
Accepted: 05 December 2024
Published: 12 December 2024

Research article

Optimal investment game for two regulated players with regime switching

Lin Xu[1,*], **Linlin Wang**[1], **Hao Wang**[1] **and Liming Zhang**[2]

[1] School of Mathematics and Statistics, Anhui Normal University, Wuhu, Anhui 241002, China
[2] School of Big Data and Statistics, Anhui University, Hefei, Anhui 230601, China

* **Correspondence:** Email: xlwgl11@ahnu.edu.cn; Tel: +86-553-5910663.

Abstract: This paper investigated a zero-sum stochastic investment game for two investors in a regime-switching market with common random time solvency regulations. We considered two types of intensities for the inter-arrival time of regulations: one was modeled as a function of a time-homogeneous Markov chain, while the other was treated as a deterministic function of time t. In the first case, the associated Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation was an elliptic partial differential equation (PDE). By solving an auxiliary problem, we demonstrated the existence and regularity of the value function. In the regime-switching model, players' optimal strategies resembled those in a non-regime-switching model but required dynamic adjustments based on the Markov chain state. In the second case, the associated HJBI equation was a parabolic PDE. We provided a numerical method using a Markov chain approximation scheme and presented several numerical examples to illustrate the impact of regime switching and random time solvency on optimal policies.

Keywords: stochastic investment game; solvency regulations; fixed point method; Markov chain approximation
Mathematics Subject Classification: 91-10, 90-10, 90C39

1. Motivations and literature review

1.1. Motivations

Stochastic differential games (SDGs) are a sophisticated and rewarding branch of game theory. In SDG problem, decisions are made in interactive environments, and the players of the game try to find optimal policies and balance the trade-off with their opponents. A key feature of SDGs is the use of stochastic differential equations with control variables to define the state dynamics of the system. For example, see the work of [7, 30] and the references therein. In most cases, the parameters of the controlled system in previous SDG problems are assumed to be constants or the function of the controlled system itself. Since the empirical study has witnessed the abrupt change in the return of the financial market (c.f. [11]), it is natural to a construct controlled system which can capture the effects of structural shifts in macroeconomic conditions and business cycles on price dynamics. One typical stochastic system with regime switching has its roots in early work by [24]. This inspires us to study the SDG problem in a system with regime switching. [4] provided a presentative work on investigating nonzero sum SDG problem within the jump diffusion model with regime switching.

Fund managers are often incentive to invest in high-volatility, risky assets in pursuit of higher returns or to outperform market benchmarks, commonly known as "beating the market". Such incentives can elevate the risk of future losses for investors, making these aggressive strategies unpopular with shareholders and detrimental to the stable development of the financial market. Consequently, both shareholders and regulators must closely monitor the investment behavior of financial institutions.

Among the most critical regulatory aspects is solvency, the 2008 financial crisis underscored significant gaps in capital and risk management within financial institutions. In response, global regulatory bodies implemented comprehensive reforms to enhance solvency standards and safeguard against systemic risk. One intriguing question arises: if solvency regulations were applied to players in a SDG, what changes would ensue, and how could these changes be quantified? How might one model these "regulations" in a meaningful way? These considerations prompt us to investigate SDG within a regime-switching model. Unlike the work of [4], this paper exclusively examines nonzero sum SDG with regime switching.

1.2. Literature review

In past two decades, SDGs have garnered increasing interest in finance and actuarial science. For example, [7] studied investment games within the Black-Scholes model, [4] extended this work to a jump diffusion model. [17] explored SDGs with relative performance metrics and control constraints, and [2] examined SDGs for fund managers. In actuarial research, [49] and [42] investigated nonzero sum SDGs between insurance and reinsurance companies, [23] analyzed SDGs between two defined contribution pension plans, [36] studied robust SDGs under model uncertainty, and [10] explored optimal SDGs using the mean-variance premium principle. The common feature of the models used in these works is that the parameters of the controlled system are constants. Recently, more sophisticated model were used in the SDGs (c.f. [31, 43]), or more potential risks were incorporated in the system, such as default risk or asymmetry information for the players (c.f. [13, 51]). Compared to previous works, research on SDGs under regime-switching models is relatively less prevalent. This is primarily because games under such models are often not amenable to solutions in closed-form, thereby posing challenges for study.

The first application of the Markov regime switching models in economics was proposed by [24] and consisted of the analysis of business cycles. The business cycle interpretation of the model relied on the combined analysis of the signs of the regime-specific intercept terms and the historical narrative about the periods with high values of the smoothed state probabilities for each of the regimes. Accordingly, a negative value of the intercept term coincided with the periods of economic recessions, whereas its positive value was associated with economic expansions. The regime-switching framework is particularly useful for understanding the behavior of financial markets and insurance surplus processes under varying economic conditions, making it a powerful tool for both theoretical analysis and practical applications. For detailed topics in this model, we refer to [16, 34, 48].

The focus of investigations has long been on solvency regulations pertaining to optimal investment and reinsurance strategies. [12] studied the impact of regulations on fair premium setting. There is an increasing attention on this topic recently. For example, [18] studied optimal investment and premium setting while there are solvency regulations; [9] researched optimal investment under VaR-regulations; [5] studied Pareto-optimal polices with solvency regulations; [1] derived optimal reinsurance design with solvency constraints. From a mathematical perspective, these optimal control problems are primarily modeled using single-period static frameworks, where various solvency requirements are incorporated as constraints into the optimization problems. The reason for not considering dynamic multi-period models is that quantifying solvency conditions based on control processes is often difficult to characterize or solve in dynamic models. This paper will draw on the ideas of [12] by using randomly arriving monitoring times to describe solvency constraints, with the aim of optimizing the decision-maker's performance before the arrival of these monitoring events.

1.3. The contribution of this paper

This paper addresses this issue within a competitive framework by formulating the problem as an SDG within random time horizons. The issues in [7] closely relate to the topic discussed in this paper. Compared to [7], we incorporate a regime-switching structure into the dynamic control model and focus on the impact of random time regulation. In [7], the HJBI (Hamilton-Jacobi-Bellman-Isaac) equation is an elliptic PDE (partial differential equation). However, in this paper, the HJBI equation takes two forms: when the intensity process of the regulation time is a function of an external Markov chain, it is a coupled elliptic PDE; when the intensity process is a deterministic function, the HJBI equation is a parabolic PDE. The explicit solution methods applied in [7] are invalid in this paper. Therefore, we explored two different methods for solving the aforementioned intensity processes.

In many cases, it is assumed that the random regulation times follow an exponential distribution. In this paper, for practical relevance, we model the intensity of the random solvency regulation time using two different approaches. The first model assumes that the intensity process is a function of an external Markov chain, resulting in a time-homogeneous Markov chain itself. The motivation behind this approach is that a natural understanding is: external regulations are influenced by the external macroeconomic environment. When the environment is good, there is a possibility of lower default risk and thus less regulation intensity. A proper way to model such dependence is to assume that the arrival intensity is a function of the Markov chain. Both the constant arrival intensity and the Markov chain-modeled arrival intensity are time-homogeneous. Our other interest is to treat the time-inhomogeneous intensity process. For ease of exploration, we consider the intensity process as a deterministic function of time t.

While [32, 37] have explored SDGs with random durations, their models do not incorporate regime switching or an insurance context. When the intensity process of regulation time follows a Markov chain, we address the SDG by employing an auxiliary problem approach combined with a fixed-point method. We establish the expressions for optimal policies by resolving the auxiliary problem. In the regime-switching model context, we find that the optimal strategies for both players are akin to those in models without regime switching; however, players must dynamically adjust their strategies in response to the state transitions of the Markov chain.

Through this approach, our study provides a theoretical foundation for investment games in stochastic environments and explores strategy formulation under the uncertainties of regulatory intensity and market state transitions. In the case where the intensity process is deterministic, the associated HJBI equation takes the form of a time-dependent parabolic PDE. To solve this equation, we propose a numerical method based on a Markov chain approximation scheme. Additionally, we present several numerical examples to demonstrate the effects of regime switching and random time solvency on the optimal policies. These examples illustrate how the systems dynamics influence the decision-making processes of both investors and highlight the significance of incorporating these factors into the investment strategies.

The remainder of this paper is organized as follows: Section 2 introduces the model and outlines the key issues to be addressed. It also presents the HJBI equation that the value function of the SDG must satisfy. Section 3 discusses the scenario in which the intensity process is modeled as a Markov chain, while Section 4 examines the case where the intensity is treated as a deterministic function of time t. The paper concludes with a discussion of numerical methods, algorithms, and illustrative examples to highlight our findings. At last, for reading convenience, we put the list of important notations in the paper in the following Table 1.

Table 1. Summary of notations used in this paper.

Notations	**Description**
$\{\mathbf{X}_t,\ t \geq 0\}$	External Markov chain modulating the dynamic of the market
$S_t(i), i = 1, 2$	Price of the financial market
$\theta_i, i = 1, 2$	Sharpe ratio of the two financial market
$\{f_t, g_t,\ t \geq 0\}$	Investment policies adopted by Player A and B respectively
$\{Z_t^{f,g},\ t \geq 0\}$	Ratio process of the two players under control f, g
τ	Inter-arrival random time of regulations
$\tau_x^{f,g}$	The first time that controlled process $Z_t^{f,g}$ reaches x
$\tau^{f,g}$	The first exit time of $Z_t^{f,g}$ with $Z_0 = z \in [l, u]$
$v^{f,g}(z, \alpha_i)$	Performance function of the SDG with initial state $(Z_0, X_0) = (z, \alpha_i))$ when external regulation time is time homogeneous
$v^{f,g}(t, z, \alpha_i)$	Performance function of the SDG with initial state $(Z_t, X_t) = (z, \alpha_i))$ when external regulation time is time inhomogeneous
$v^{Au,f,g}(z, \alpha_i)$	Performance function of the SDG with initial state $(Z_0, X_0) = (z, \alpha_i))$ and stopped at the change the external Markov chain state
$J^{f^h,g^h}(t, z, \alpha_i)$	performance function of the approximating Makov chain
$V^{Au}(z, \alpha_i)$	Value function of auxiliary SDG
$\mathbb{P}((z, \alpha_i), (z + h, \alpha_i) \| f^h, g^h)$	Transition probability of approximating Makov chain

2. Model and HJBI equations

2.1. Financial model

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space endowed with right-continuous, $\mathbb{P}$-completed filtration $\{\mathcal{F}_t, t \geq 0\}$. Assume that there are two correlated risky assets $S_t^{(1)}$ and $S_t^{(2)}$, a risk-free bond B_t and an external environment evolution process $X := \{X_t,\ t \geq 0\}$. While we allow both investors to invest in risk-free market, A chooses $S^{(1)} = \{S_t^{(1)},\ t \geq 0\}$ and investor B chooses $S^{(2)} = \{S_t^{(2)},\ t \geq 0\}$. Assume that $X := \{X_t,\ t \geq 0\}$ is a continuous-time, finite-state, observable Markov chain taking values in state space $\mathfrak{X} := \{\alpha_1, \alpha_2, \cdots, \alpha_d\},\ d \geq 2$. $W_t^{(i)},\ i = 1, 2$ are two correlated Brownian motions with coefficient $\rho_t \hat{=} \rho(X_t)$. Let $\mathbf{Q} := [q_{ij}]_{i,j=1,2,\ldots,d}$ be the generator of X. For each $i, j = 1, 2, \cdots, d$, q_{ij} means the constant intensity that the Markov chain X changes from state α_i to state α_j. Assume that $q_{ij} > 0,\ \sum_{j=1}^{d} q_{ij} = 0$, so $q_{ii} < 0$. Denote by $q_i = -q_{ii} > 0$. Let $\mathbf{Q}^{\mathsf{T}}$ be the transpose of a matrix, or a vector $\mathbf{Q}$. [15] presented the semi-martingale dynamics of $\mathbf{X}$ as

$$X_t = X_0 + \int_0^t \mathbf{Q}^{\mathsf{T}} X_u \mathrm{d}u + M_t,$$

where $\{M_t, t \geq 0\}$ is a martingale with respect to $\{\mathcal{F}_t, t \geq 0\}$. Denote by τ_i the i_{th} jumping time of X_t, then we have following Lemma 2.1 (c.f. [22]).

Lemma 2.1.

$$\mathbb{P}(\tau_1 > t | X_0 = \alpha_i) = \mathrm{e}^{-q_i t}; \tag{2.1}$$

$$\mathbb{P}\left(\tau_1 \leq t, X_{\tau_1} = \alpha_j | X_0 = \alpha_i\right) = (1 - \mathrm{e}^{-q_i t}) \frac{q_{ij}}{q_i}; \tag{2.2}$$

$$\mathbb{P}\left(X_{\tau_1} = \alpha_j | X_0 = \alpha_i\right) = \frac{q_{ij}}{q_i}. \tag{2.3}$$

Assume that the risky assets are evolved as

$$dS_t^{(k)} = S_t^{(k)} \left(\mu_k(X_t)dt + \sigma_k(X_t)dW_t^{(i)}\right), \quad k = 1, 2,$$

where $\mu_k(X_t) > 0, \sigma_k(X_t),\ k = 1, 2$ are return rates and volatilities of the two risky assets respectively. The dynamic of the risk-free asset is

$$dB_t = r(X_t)B_t dt,$$

where $r(\cdot) \geq 0$ for all $\alpha_i, i = 1, 2, \cdots, d$. Denote by

$$\theta_{kt}(X_t) := \frac{\mu_k(X_t) - r(X_t)}{\sigma_k(X_t)},\ k = 1, 2$$

the Sharpe ratio or the market price of risk associated to asset $S^{(k)}$ at time t.

Denote by f_t the proportion of A 's wealth in risky asset $S^{(1)}$ and by g_t the proportion of B's wealth in risky asset $S^{(2)}$. We made the following assumption on the control policies:

Assumption 1. *(1)* f_t(or g_t) is an anticipated, measurable function with respect to $\mathcal{F}_t$ and satisfies

$$\mathbb{E}\left[\int_0^T f_t^2 dt\right] < \infty \quad (or \quad \mathbb{E}\left[\int_0^T g_t^2 dt\right] < \infty), \forall T < \infty. \tag{2.4}$$

(2) Both short selling and borrowing are allowed in trading. Specifically, we allow that $f_t \geq 1$ (borrow) or $f_t < 0$ (short selling) and so does g_t.

Denote by Y_t^f (or Y_t^g) the wealth of investor A (or B) under policy $\{f_t, t \geq 0\}$ (or $\{g_t, t \geq 0$) with $Y_0^f = x_0$ (or $Y_0^g = y_0$), then the dynamic of Y_t^f (or Y_t^g) is given by

$$dY_t^f = Y_t^f \left(\left[f_t\sigma_1(X_t)\theta_{1t}(X_t) + r(X_t)\right] dt + f_t\sigma_1(X_t)dW_t^{(1)}\right) \tag{2.5}$$

or

$$dY_t^g = Y_t^g \left(\left[g_t\sigma_2(X_t)\theta_{2t}(X_t) + r(X_t)\right] dt + g_t\sigma_2(X_t)dW_t^{(2)}\right). \tag{2.6}$$

2.2. Relative performance and ratio process

While there are many competition objectives, we just focus on the games with payoffs related to the achievements of relative performance goals and shortfalls. For two numbers $l < u$ with $l \leq \frac{x_0}{y_0} \leq u$, we say that investor A attains its upper performance u if $Y_t^f = uY_t^g$, for some $t > 0$, and that lower shortfall occurs if $Y_t^f = lY_t^g$, for some $t > 0$. In general, A wins the game if A attains its upper performance before it reaches the lower shortfall, while B wins the game when the converse happens. In this paper, we further consider the regulation time impact on the decisions of both investors, where regulation time is specified by Definition 2.2. Similar to the discussion in [7], some specific games we consider here, starting from the perspective view of investor A, are (within regulation time)

- maximizing the probability that performance goal u is attained before the shortfall l occurred;
- minimizing the expected time of the performance u attained;
- maximizing the expected total discounted reward upon performance u reached.

Similar to the framework in [7], we investigate the *ratio* of two wealth processes. Denote the *ratio process* by $Z_t^{f,g}$ with $Z_t^{f,g} = \frac{Y_t^f}{Y_t^g}$, then the dynamic of $Z_t^{f,g}$ is given by

$$dZ_t^{f,g} = Z_t^{f,g}\left[m(f_t, g_t, X_t)dt + f_t\sigma_1(X_t)dW_t^{(1)} - g_t\sigma_2(X_t)dW_t^{(2)}\right], \tag{2.7}$$

where

$$m(f_t, g_t, X_t) \ \hat{=} \ f_t\sigma_1(X_t)\theta_{1t}(X_t) - g_t\sigma_2(X_t)\theta_{2t}(X_t) - f_tg_t\sigma_1(X_t)\sigma_2(X_t)\rho(X_t) + g_t^2\sigma_2^2(X_t). \tag{2.8}$$

Note that $\{(Z_t^{f,g}, X_t)\}$, $t \geq 0$ is a vector valued Markov process with $(Z_0, X_0) = (x, \alpha_i)$, then the infinitesimal operator of process $\{Z_t^{f,g}, X_t\}$, $t \geq 0$ is given by (suppose that function F belongs to the domain of operator $\mathcal{A}$)

$$\mathcal{A}^{f,g}F(t,z,\alpha_i) = F_t(t,z,\alpha_i) + m(f,g,\alpha_i)zF_z(t,z,\alpha_i) + \frac{1}{2}\nu^2(f,g,\alpha_i)z^2F_{zz}(t,z,\alpha_i) + \sum_{j=1}^{d} q_{ij}F(t,z,\alpha_j),$$

where F_t, F_z, F_{zz} are the first partial derivative of $F(\cdot,\cdot)$ w.r.t. t and the first partial derivative and the second partial derivative of $F(\cdot,\cdot)$ w.r.t. z,

$$\begin{aligned}
&\mu_{ki} = \mu_k(\alpha_i), \sigma_{ki} = \sigma_k(\alpha_i),\ r_i = r(\alpha_i), \rho_i = \rho(\alpha_i),\\
&\theta_{ki} = \frac{\mu_{ki} - r_i}{\sigma_{ki}},\\
&m(f,g,\alpha_i) = f\sigma_{1i}\theta_{1i} - g\sigma_{2i}\theta_{2i} - fg\sigma_{1i}\sigma_{2i}\rho_i + g^2\sigma_{2i}^2,\\
&\nu^2(f,g,\alpha_i) = f^2\sigma_{1i}^2 + g^2\sigma_{2i}^2 - 2fg\sigma_{1i}\sigma_{2i}\rho_i,\ k = 1,2, \quad i = 1,2,\cdots,d.
\end{aligned} \tag{2.9}$$

Let

$$\kappa_i = \frac{\theta_{1i}}{\theta_{2i}} \tag{2.10}$$

denote the *ratio of the market prices* of two risk assets in finance. We will see later that the parameter κ_i is a measure of the degree of advantage one player has over the other. A is said to have the advantage if $\kappa_i > 1$ and B is said to have the advantage if $\kappa_i < 1$.

2.3. Exit time and regulation time

Define by $\tau_x^{f,g}$ the first time that controlled process $Z_t^{f,g}$ reaches $x \in [l, u]$ and by $\tau^{f,g} = \min\{\tau_l^{f,g}, \tau_u^{f,g}\}$ the first exit time of $Z_t^{f,g}$ with $Z_0 = z \in [l, u]$.

Definition 2.2. (Regulation time) *Assume that τ is the inter-arrival random time of regulations for two investors. We assume that there exists a nonnegative stochastic process λ_s, $s \geq 0$, such that*

$$\mathbb{P}\left(\int_0^\infty \lambda_s \mathrm{d}s = +\infty\right) = 1; \tag{2.11}$$

$$\mathbb{P}(\tau > t) = \mathbb{E}\left[\exp\left(-\int_0^t \lambda_s \mathrm{d}s\right)\right]. \tag{2.12}$$

Remark 1. *Usually, mortality function λ_s, $s \geq 0$ is constant or a deterministic function (c.f. [33]). In this paper, we additionally consider the case that λ_s, $s \geq 0$ is a function of $\{X_t, t \geq 0\}$ and thus is a Markov chain. There is a natural explanation for this model: the external environment not only affects the performance of the financial market, but the regulation frequency from the administrator is variable to the current state of the environment. For notation ease, denote by $\lambda_i = \lambda(\alpha_i)$.*

2.4. Competition and saddle points when $\lambda_t = \lambda(X_t)$

In this subsection, we consider the case that the intensity process is the function of Markov chain X_t, i.e. $\lambda_t = \lambda(X_t) > 0$. Due to the fact that exponential distribution is "memoryless" and λ_s is a Markov process, the performance function of SDG is of the form

$$v^{f,g}(z, \alpha_i) = \mathbb{E}_{z,\alpha_i}\left[\int_0^{\tau^{f,g}\wedge\tau} \mathrm{e}^{-\delta s} c(Z_s^{f,g}) \mathrm{d}s + \mathrm{e}^{-\delta(\tau^{f,g}\wedge\tau)} h(Z^{f,g}_{\tau^{f,g}\wedge\tau})\right], \tag{2.13}$$

where $\mathbb{E}_{z,\alpha_i}$ means the condition expectation operator $\mathbb{E}_{z,\alpha_i} = \mathbb{E}[\cdot|Z_0 = z, X_0 = \alpha_i]$, $c(\cdot)$ is the reward(cost) function, and $h(\cdot)$ is the terminal reward (terminal punishment) of the game.

Remark 2. *We assume that $c(\cdot)$ and $h(\cdot)$ satisfies the polynomial growth condition, say,*

$$|c(z)| \leq C(1 + |z|^p), |h(z)| \leq C(1 + |z|^p)$$

for suitable C, p. The coefficients in (2.5) (or (2.6)) satisfy condition (5.2) and (5.3) in IV 5 of [19]. With results of Appendix D in [19], it follows that (2.5) (or (2.6)) admits a path-wise unique solution Y_t^f (or Y_t^g), which is $\mathcal{F}_t$-progressively measurable and has continuous sample paths. With similar discussion, the existence of solution of the stochastic differential equation (SDE) specified by Eq (2.7) is guaranteed. With the help of aforementioned assumptions, just as it was claimed in IV 5 of [19], the performance function (2.13) is well-defined.

The two investors compete in the following form: A wants to maximize *payoff* function $v^{f,g}(z, \alpha_i)$ while B wants to minimize $v^{f,g}(z, \alpha_i)$. We consider here only perfect observed competition, that is to say, the policy adopted by one investor at any time could be directly observed by the opponent investor instantaneously. Let

$$\underline{V}(z, \alpha_i) = \sup_f \inf_g v^{f,g}(z, \alpha_i), \bar{V}(z, \alpha_i) = \inf_g \sup_f v^{f,g}(z, \alpha_i) \tag{2.14}$$

be the lower value function and upper value function of the game respectively.

Definition 2.3. *If* $\underline{V}(z,\alpha_i) = \bar{V}(z,\alpha_i)$*, we call that the* value function *of the game exists, and naturally is given by*

$$V(z,\alpha_i) = \underline{V}(z,\alpha_i) = \bar{V}(z,\alpha_i). \tag{2.15}$$

This value can be attained if a saddle point for the payoff $v^{f,g}(z,\alpha_i), i = 1,2,\cdots,d, x \in [l,u]$ *exists, i.e. there exist* $f^* = \{f_t^*,\ t \geq 0\}$ *and* $g^* = \{g_t^*,\ t \geq 0\}$ *such that for all* $(z,\alpha) \in [l,u] \times \mathfrak{X}$ *and all admissible* f *and* g*, the following relations hold:*

$$v^{f,g^*}(z,\alpha_i) \leq v^{f^*,g^*}(z,\alpha_i) \leq v^{f^*,g}(z,\alpha_i). \tag{2.16}$$

Then, $v(z,\alpha_i) = v^{f^*,g*}(z,\alpha_i)$ *and, thus, the saddle points exist and are given by* $f^*, g*$.

2.5. Competition and saddle points when λ_t is a deterministic function of t

The second case assumes that λ_t is no longer a function of time homogeneous Markov chain X_t, but a deterministic function of t. In this case, we note that the performance function of the game not only relies on the current state z of the controlled system, but also the current time t. For notation ease, introduce

$$\begin{aligned}
\mathbb{E}_{t,z,\alpha_i} &= \mathbb{E}\left[\,\cdot\,\middle|(Z_t^{f,g}, X_t) = (z,\alpha_i)\right],\\
\mathbb{P}_{t,z,\alpha_i} &= \mathbb{P}\left[\,\cdot\,\middle|(Z_t^{f,g}, X_t) = (z,\alpha_i)\right].
\end{aligned} \tag{2.17}$$

Let $v^{f,g}(t,z,\alpha_i)$ be the payoff performance function under the policies f and g with initial value (t,z,α_i) and regulation time τ, which is defined by

$$v^{f,g}(t,z,\alpha_i) \quad = \quad \mathbb{E}_{t,z,\alpha_i}\left[\int_t^{\tau^{f,g}\wedge\tau} \mathrm{e}^{-\delta s} c(Z_s^{f,g})ds + \mathrm{e}^{-\delta(\tau^{f,g}\wedge\tau)} h(Z^{f,g}_{\tau^{f,g}\wedge\tau})\right]. \tag{2.18}$$

We similarly define the value function and saddle of the game in this case as follows.

Definition 2.4. *Let*

$$\underline{V}(t,z,\alpha_i) = \sup_f \inf_g v^{f,g}(t,z,\alpha_i), \bar{V}(t,z,\alpha_i) = \inf_g \sup_f v^{f,g}(t,z,\alpha_i) \tag{2.19}$$

be the upper value and lower value of the SDG (2.18), respectively. If

$$\underline{V}(t,z,\alpha_i) = \bar{V}(t,z,\alpha_i) \tag{2.20}$$

we call that the value function of value of the SDG (2.18) exists and is given by

$$\underline{V}(t,z,\alpha_i) = V(t,z,\alpha_i) = \bar{V}(t,z,\alpha_i). \tag{2.21}$$

If there exist $f^* = \{f_t^*,\ t \geq 0\}$ *and* $g^* = \{g_t^*,\ t \geq 0\}$ *such that for all* $(t,z,\alpha_i) \in [l,u] \times \mathfrak{X}$ *and all admissible* f *and* g*,*

$$v^{f,g^*}(t,z,\alpha_i) \leq v^{f^*,g^*}(t,z,\alpha_i) \leq v^{f^*,g}(t,z,\alpha_i) \tag{2.22}$$

then

$$V(t,z,\alpha_i) = v^{f^*,g*}(t,z,\alpha_i) \tag{2.23}$$

and thus the saddle points exist and are given by $f^*, g*$.

Remark 3. *The existence of the value function and the saddle point of SDG plays a fundamental role in the study of SDG. For instance, see the works of [8, 14, 20, 41]. However, there are various challenges in proving the existence of value functions, depending on the framework of the current SDG. The characteristic of the SDG in this paper is that, in addition to the control terms of the two players, it accommodates a Markov modulated structure in the drifts and diffusions, as well as an external random "stoping" time. The focus of this paper is to find optimal policies for the players. Motivated by the results of [40] and [26], we find that it suffices to verify the conditions A1)–A5) and A7) from [40]; our framework meets these conditions. Consequently, Theorem 5.3 of [40], which establishes the existence of the value function in SDG with a Markov regime switching structure over a stochastic time horizon, is applicable to our context.*

3. Optimal policies for players when $\lambda_t = \lambda(X_t)$

We first introduce some notations and definitions:

(1) For any function $V(z, \alpha_i), i = 1, 2, \cdots, d$ with continuously second order partial derivative w.r.t. z, let's denote by Θ the differential operator specified by

$$\Theta V(z, \alpha_i) = (1 - \rho_i^2)\left[V_z(z, \alpha_i) + zV_{zz}(z, \alpha_i)\right]^2 - V_z(z, \alpha_i)^2. \tag{3.1}$$

(2) $V(z, \alpha_i), i = 1, 2, \cdots, d$ is said to be *sufficiently fast-increasing* on an interval (a, b) if the following condition holds:

$$2V_z(z, \alpha_i) + zV_{zz}(z, \alpha_i) > 0 \tag{3.2}$$

for $i = 1, 2, \cdots, d$ and $z \in [l, u]$.

We note that in our model, the advantage of the two investors is variable with respect to the economic environment, which differs significantly from the case presented by [7], making our problem more complex and realistic in practice. The following Theorem 3.1 presents the HJBI equation associated with problem (2.15). The proof of this theorem is similar to that of Theorem 4.1, so we only provide the proof for Theorem 4.1.

Theorem 3.1. *Suppose that the value function $V(z, \alpha_i) : [l, u] \times \mathfrak{X} \mapsto \mathbf{R}, i = 1, 2, ..., d$ has continuously second order partial derivatives w.r.t. z, strictly concave, fulfilling condition (3.2), then $V(z, \alpha_i), i = 1, 2, \cdots, d$ solve the following equations for all $z \in [l, u]$:*

$$\begin{aligned} &\frac{zV_z(z, \alpha_i)^2}{2\Theta V(z, \alpha_i)}\theta_{2i}^2 \left[(2\kappa_i(\rho_i - \kappa_i)V_z(z, \alpha_i) - (1 + \kappa_i^2 - 2\rho_i\kappa_i)zV_{zz}(z, \alpha_i)\right] \\ &+c(z) - (\lambda_i + \delta)V(z, \alpha_i) + \sum_{j=1}^{d} q_{ij}V(z, \alpha_j) = 0, \quad i = 1, 2, \cdots, d, \end{aligned} \tag{3.3}$$

with

$$V(l, \alpha_i) = h(l) \quad \text{and} \quad V(u, \alpha_i) = h(u) \quad \text{for} \quad i = 1, 2, \cdots, d.$$

If $w(z, \alpha_i), i = 1, 2, \cdots, d$ solve coupled HJBI equations (3.3) and satisfy

(1) *for all admissible policies f and g and for all $t \geq 0$,*

$$\int_0^t \mathbb{E}\left[Z_s^{f,g} w_z(Z_s^{f,g}, X_s)\right]^2 [f_s^2 + g_s^2]ds < \infty; \tag{3.4}$$

(2) *function*

$$z w_z(z, \alpha_i) \frac{w_z(z, \alpha_i) + |z w_{zz}(z, \alpha_i)|}{|\Theta w(z, \alpha_i)|} \tag{3.5}$$

are uniformly bounded for all $i = 1, 2, \cdots, d$.

Then we have

$$w(z, \alpha_i) = V(z, \alpha_i) \tag{3.6}$$

and the feedback optimal controls are given by

$$f^*(z, \alpha_i) = \frac{\theta_{1i}}{\sigma_{1i}} \left(\frac{w_z(z, \alpha_i)}{\Theta w(z, \alpha_i)}\right) \left[\left(\frac{\rho_i}{\kappa_i} - 1\right)(w_z(z, \alpha_i) + z w_{zz}(z, \alpha_i))\right], \tag{3.7}$$

$$g^*(z, \alpha_i) = \frac{\theta_{2i}}{\sigma_{2i}} \left(\frac{w_z(z, \alpha_i)}{\Theta w(z, \alpha_i)}\right) \Big[(1 - \rho_i \kappa_i)(w_z(z, \alpha_i) + z w_{zz}(z, \alpha_i))\Big]. \tag{3.8}$$

Moreover,

$$\frac{f^*}{g^*} = \frac{\sigma_{2i}(\rho_i - \kappa_i)}{\sigma_{1i}(1 - \rho_i \kappa_i)}. \tag{3.9}$$

3.1. An auxiliary game problem and optimal policies

Deriving explicit expressions for the coupled HJBI equations (3.3) is generally not straightforward. In [45], a stochastic differential game was considered, yet explicit solutions were derived only under specific constraints on the system's coefficients. In this paper, we adopt the "fixed point method" from [25] to investigate optimal dividends within a Markov regime-switching model. This approach has been applied by [50] for singular optimal dividend control in a regime-switching Cramér-Lundberg model with interest on credit and debit, by [21] for portfolio optimization in a regime-switching market with derivatives, and by [46] for optimal investment and dividend strategies involving tax payments. Here, we re-examine a game problem subject to random time regulation constraints, where the process halts if the current regime switches. Specifically, let τ_1 , denote the first instance the environment shifts. We then define an auxiliary game problem as follows:

- *Auxiliary performance function*:

$$v^{Au,f,g}(z, \alpha_i) = \mathbb{E}_{z,\alpha_i}\left[\int_0^{\tau^{f,g} \wedge \tau \wedge \tau_1} \mathrm{e}^{-\delta s} c(Z_s^{f,g}) \mathrm{d}s + \mathrm{e}^{-\delta(\tau^{f,g} \wedge \tau \wedge \tau_1)} h(Z^{f,g}_{\tau^{f,g} \wedge \tau \wedge \tau_1})\right]. \tag{3.10}$$

- *Auxiliary value function*:

$$V^{Au}(z, \alpha_i) = \sup_f \inf_g v^{Au,f,g}(z, \alpha_i) = \inf_g \sup_f v^{Au,f,g}(t, z, \alpha_i). \tag{3.11}$$

With similar discussion to Theorem 3.1, the HJBI equation associated with the auxiliary problem is given by Corollary 1.

Corollary 1. *Suppose that the current state of external environment is α_i, $V^{Au}(z,\alpha_i)$ is a function with continuously second order partial derivatives w.r.t. z, strictly concave, fulfilling condition (3.2), then $V^{Au}(z,\alpha)$ solves the following equation for all $z \in [l,u]$:*

$$\begin{aligned}&\frac{zV_z^{Au}(z,\alpha_i)^2}{2\Theta V^{Au}(z,\alpha_i)}\theta_{2i}^2\left[2k_i(\rho_i-k_i)V_z^{Au}(z,\alpha_i)-(1+k_i^2-2\rho_ik_i)zV_{zz}^{Au}(z,\alpha_i)\right]\\&+c(z)-(\lambda_i+\delta+q_i)V^{Au}(z,\alpha_i)=0,\end{aligned}\tag{3.12}$$

with

$$V^{Au}(l,\alpha_i)=h(l)\quad \text{and} \quad V^{Au}(u,\alpha_i)=h(u).$$

For any give α_i, if there exists a regular solution $w(z,\alpha_i)$ to (3.12) that satisfies analogue conditions to (3.4) and (3.5), then

$$w(z,\alpha_i)=V^{Au}(z,\alpha_i)\tag{3.13}$$

and "feedback optimal control" have the same form as it were in (3.7) and (3.8).

Proof. The proof is very similar to the one for Theorem 3.1 of [7] and we omit here. □

We note that the HJBI equation in the auxiliary problem is not coupled; it is valid only until the current state changes. Assuming the current time is zero, the effective time interval for this policy is given by $[0,\tau\wedge\tau_1)$. For the remainder of this section, we will proceed under the assumptions outlined in Corollary 1.

3.2. Optimal policies for SDG (2.13)

Inspired by the Markov property of $\{X_t, t\geq 0\}$, we introduce a candidate control process $\{f_t, g_t, t\geq 0\}$ for the original problem over the entire control time interval as

$$f_t=f^*_{Au,X(t)}=f^*_{Au,X(\tau_k)},\quad \text{if } \tau_k\leq t<\tau_{k+1},\tag{3.14}$$

$$g_t=g^*_{Au,X(t)}=g^*_{Au,X(\tau_k)},\quad \text{if } \tau_k\leq t<\tau_{k+1}.\tag{3.15}$$

We observe that the candidate control process is piecewise deterministic, contingent solely on the current environment state. Consequently, under this policy, investors A and B each adopt environment-specific strategies and adjust their policies only upon state changes. Theorem 3.2 below establishes that the policies derived from Eqs (3.14) and (3.15) are indeed optimal for both investors. The proof, for brevity, is provided in Appendix B.

Theorem 3.2. *Suppose that $\lambda_t=\lambda(X_t)$, then the controlled process defined by Eqs (3.14) and (3.15) are optimal for both investors.*

Proof. For reading convenience, we put the proof in Appendix A. □

3.3. Minimizing the expected time of A winning the game

In this subsection, we analyze the auxiliary game problem aimed at maximizing or minimizing a player's expected time to outperform their opponent. Specifically, we focus on Investor A's objective to minimize the expected duration of victory, as represented in the value function:

$$N(h,\alpha_i)=\inf_f\sup_g \mathbb{E}_{z,\alpha_i}[\tau_u^{f,g}\wedge\tau\wedge\tau_1]=\sup_g\inf_f \mathbb{E}_{z,\alpha_i}[\tau_u^{f,g}\wedge\tau\wedge\tau_1].\tag{3.16}$$

Similarly, let $\tilde{N}(z,\alpha_i) = \inf_g \sup_f \mathbb{E}_{z,\alpha_i}[\tau_u^{f,g} \wedge \tau \wedge \tau_1] = \sup_f \inf_g \mathbb{E}_{z,\alpha_i}[\tau_u^{f,g} \wedge \tau \wedge \tau_1]$, so $N(z,\alpha_i) = -\tilde{N}(z,\alpha_i)$. Note that in this case, $c(\cdot) \equiv 1,\ \delta \equiv 0,\ d(\cdot) = 0$; thus, by Corollary 1, $\tilde{N}(z,\alpha_i)$ is the solution to equation

$$\frac{z\tilde{N}_z(z,\alpha_i)^2}{2\Theta\tilde{N}(z,\alpha_i)}\theta_{2i}^2\left[2k_i(\rho_i - k_i)\tilde{N}_z(z,\alpha_i) - (1 + k_i^2 - 2\rho_i k_i)z\tilde{N}_{zz}(z,\alpha_i)\right] + 1 - (\lambda_i + q_i)\tilde{N}(z,\alpha_i) = 0, \quad (3.17)$$

with boundary condition $\tilde{N}(u,\alpha_i) = 0$. [7] solved Eq (3.17) when $\lambda_i + q_i = 0$, which motivated us to find an explicit expression for $\tilde{N}(z,\alpha_i)$ in a special case. Assume that $\tilde{N}(z,\alpha_i)$ is of the form of $\tilde{N}(z,\alpha_i) = \frac{1}{\lambda_i+q_i}\left[\left(\frac{z}{u}\right)^{\zeta} + 1\right]$. By [7], we can get the solution of the problem, and the final result is

$$\tilde{N}(z,\alpha_i) = \frac{1}{\lambda_i + q_i}\left[\left(\frac{z}{u}\right)^{\zeta^+} + 1\right],$$

where the form of ζ^+ is

$$\zeta^+ = \frac{\theta_{2i}^2(1-k_i^2) + \sqrt{\Delta}}{2\theta_{2i}^2\left(1 + k_i^2 - 2\rho_i k_i\right) + 4(\lambda_i + q_i)(1-\rho_i^2)},$$

$$\Delta = \left[\theta_{2i}^2(1-k_i^2)\right]^2 + 8(\lambda + \lambda_1)\theta_2^2\left(1 + k_i^2 - 2\rho_i k_i\right) + 16(\lambda_i + q_i)^2(1-\rho_i^2).$$

Finally, the value of (3.16) is given by

$$N(h,\alpha_i) = -\frac{1}{\lambda_i + q_i}\left[\left(\frac{z}{u}\right)^{\zeta^+} + 1\right]. \quad (3.18)$$

Then,

$$N_z = -\frac{1}{\lambda_i + q_i}\frac{1}{u}\zeta^+\left(\frac{z}{u}\right)^{\zeta^+-1},$$

$$N_{zz} = -\frac{1}{\lambda_i + q_i}\frac{1}{u}\zeta^+(\zeta^+-1)\left(\frac{z}{u}\right)^{\zeta^+-2}.$$

By calculation, the associated saddle point is given by

$$f^*(z) = \frac{\theta_{1i}}{\sigma_{1i}}\left[\frac{(\rho_i/k_i - 1)\zeta^+ - 1}{\left(1-\rho_i^2\right)(\zeta^+)^2 - 1}\right] \quad \text{and} \quad g^*(z) = \frac{\theta_{2i}}{\sigma_{2i}}\left[\frac{(1-\rho_i k_i)\zeta^+ - 1}{\left(1-\rho_i^2\right)(\zeta^+)^2 - 1}\right]. \quad (3.19)$$

3.4. Maximizing the probability of reaching upper level game

In this game, player A aims to maximize the probability of reaching a higher level u, while player B aims to minimize it. When the game involves a single player and as u approaches infinity with $l = 0$, this problem simplifies to minimizing the ruin probability in the presence of investment opportunities, as discussed in [6]. According to Theorem 3.2, for a given current external state α_i, it is necessary to first solve a single-state optimization problem. Now, let $\tilde{R}(z,\alpha_i)$ be the value function of the auxiliary game, then,

$$\begin{aligned}\tilde{R}(z,\alpha_i) &= \sup_f \inf_g \mathbb{P}_{z,\alpha_i}\left(Z_{\tau^{f,g}\wedge\tau\wedge\tau_1} = u\right) \\ &= \sup_f \inf_g \mathbb{P}_{z,\alpha_i}\left(\tau^{f,g} \wedge \tau \wedge \tau_1 = \tau_u^{f,g}\right).\end{aligned}$$

Note that in this case, $c(\cdot) \equiv 0$, $\delta \equiv 0$, $h(\cdot) = 1_{\{Z_{\tau}^{f,g}{}_{\wedge\tau\wedge\tau_1} = u\}}$; thus, by Corollary 1, $\tilde{R}(z, \alpha_i)$ is the solution to equation

$$\frac{z\tilde{R}_z(z, \alpha_i)^2}{2\Theta\tilde{R}(z, \alpha_i)}\theta_{2i}^2\left[2k_i(\rho_i - k_i)\tilde{R}_z(z, \alpha_i) - (1 + k_i^2 - 2\rho_i k_i)z\tilde{R}_{zz}(z, \alpha_i)\right] - (\lambda_i + a_i)\tilde{R}(z, \alpha_i) = 0, \tag{3.20}$$

with boundary condition $\tilde{R}(u, \alpha_i) = 1$, $\tilde{R}(l, \alpha_i) = 0$. Substituting the expression of $\Theta\tilde{R}(z, \alpha_i)$ into (3.20) yields

$$\begin{aligned}
&2k_i(\rho_i - k_i)z\tilde{R}_z^2(z, \alpha_i)\theta_{2i}^2 - (1 + k_i^2 - 2\rho_i k_i)z\tilde{R}_z(z, \alpha_i)\theta_{2i}^2\\
&-2(\lambda_i + q_i)z^2\tilde{R}(z, \alpha_i)\tilde{R}_{zz}(z, \alpha_i)^2 - 2(\lambda_i + q_i)z\tilde{R}(z, \alpha_i)\tilde{R}_z(z, \alpha_i)\tilde{R}_{zz}(z, \alpha_i)\\
&= \rho_i(\lambda_i + q_i)\tilde{R}_z(z, \alpha_i)^2 + 2\rho_i(\lambda_i + a_i)z\tilde{R}_z(z, \alpha_i)\tilde{R}_{zz}(z, \alpha_i)\\
&+\rho_i(\lambda_i + q_i)z^2\tilde{R}_{zz}(z, \alpha_i)^2 = 0.
\end{aligned} \tag{3.21}$$

This equation can be tracked by numerical method.

4. Optimal policies for players when λ_t is a deterministic function

The following Theorem 4.1 gives the HJBI equation associated with the SDG problem when λ_t is a deterministic function. For convenience, the proof of Theorem 4.1 is provided in the Appendix.

Theorem 4.1. *Suppose that λ_t is a positive deterministic function of t, $w(t, z, \alpha_i) : [l, u] \times \mathfrak{X} \mapsto \mathbf{R}$ is a function with continuous second-order partial derivatives w.r.t. z, strictly concave, fulfilling condition (3.2), and solves the following equation for all $z \in [l, u]$:*

$$\begin{aligned}
&w_t(t, z, \alpha_i) + \frac{zw_z^2(t, z, \alpha_i)}{2\Theta w(t, z, \alpha_i)}\theta_{2i}^2\left[(1 - k_i^2)w_z(t, z, \alpha_i) - (1 + k_i^2 - 2\rho_i k_i)\left(w_z(t, z, \alpha_i) + zw_{zz}(t, z, \alpha_i)\right)\right]\\
&+c(z) - (\delta + \lambda_t)w(t, z, \alpha_i) + \sum_{j=1}^{d} q_{ij}w(t, z, \alpha_j) = 0, \quad i = 1, 2, \cdots, d,
\end{aligned} \tag{4.1}$$

with

$$w(t, l, \alpha_i) = h(l) \quad \text{and} \quad w(t, u, \alpha_i) = h(u) \quad \text{for} \quad i = 1, 2, \cdots, d.$$

We further suppose that

(1) for all admissible policies f and g and for all $t \geq 0$,

$$\int_0^t \mathbb{E}\left[Z_s^{f,g}w_z(s, Z_s^{f,g}, X_s)\right]^2 [f_s^2 + g_s^2]ds < \infty; \tag{4.2}$$

(2) function

$$zw_z(t, z, \alpha_i)\frac{w_z(t, z, \alpha_i) + |zw_{zz}(t, z, \alpha_i)|}{|\Theta w(t, z, \alpha_i)|} \tag{4.3}$$

is uniformly bounded for all $i = 1, 2, \cdots, d$.

Then, $w(t, z, \alpha_i)$ is the value function of SDG, i.e.,

$$\begin{aligned} w(t, z, \alpha_i) &= v^{f^*, g^*}(t, z, \alpha_i) \\ &= \sup_f \inf_g v^{f,g}(t, z, \alpha_i) = \inf_g \sup_f v^{f,g}(t, z, \alpha_i). \end{aligned}$$

The "feedback" saddle points of this SDG are specified by

$$f^* = \frac{\theta_{1i}}{\sigma_{1i}} \left(\frac{w_z(t, z, \alpha_i)}{\Theta w(t, z, \alpha_i)} \right) \left[\left(\frac{\rho_i}{k_i} - 1 \right) (w_z(t, z, \alpha_i) + z w_{zz}(t, z, \alpha_i)) - w_z(t, z, \alpha_i) \right], \tag{4.4}$$

$$g^* = \frac{\theta_{2i}}{\sigma_{2i}} \left(\frac{w_z(t, z, \alpha_i)}{\Theta w(t, z, \alpha_i)} \right) \left[(1 - \rho_i k_i) (w_z(t, z, \alpha_i) + z w_{zz}(t, z, \alpha_i)) - w_z(t, z, \alpha_i) \right]. \tag{4.5}$$

4.1. Numerical method

Section 3 presents two examples of the SDG in which the arrival intensity of regulation is piece-wise constant. However, in more general cases, deriving explicit solutions can be challenging, necessitating a shift to numerical methods. Since a Markovian SDG problem can be treated as a Markovian control problem, the approach to constructing numerical schemes for SDG can leverage numerical methods for stochastic control (see [26, 28, 38, 39]). Note that the controlled wealth process is a map $[0, \infty) \mapsto [l, u]$ and stopped at $\tau^{f,g} \wedge \tau$.

Let $h > 0$ and define $L_h = \{z : z = l + kh, k = 0, 1, 2, \cdots, [\frac{u-l}{h}] + 1\}$, where $[\cdot]$ is the integer function. L_h is a discrete segmentation of interval $[l, u]$, where $\{(\xi_k^h, e_k^h), k < \infty\}$ is a controlled Markov chain on L_h, where $\{\xi_k^h, k < \infty\}$ is used to approximate the underlying controlled wealth process $\{Z_t^{f,g}, t \geq 0\}$, and $\{e_k^h, k < \infty\}$ is the discrete time observation of the external environment process $\{X_t, t \geq 0\}$. Hence, for any chosen h, the domain of our numerical schemes with step h is

$$\mathfrak{D}^h = \{(z, \alpha_i) : \ i = 1, 2, ..., d, \ z \in L_h, i = 1, 2, \cdots, d\}. \tag{4.6}$$

The design of the approximate Markov chain within the domain D_h is analogous to that presented in [40]. This controlled Markov chain is constructed to be both discrete-time and finite-state for computational efficiency, while adhering to the local consistency properties of the controlled state system. Therefore, a crucial step in designing this Markov chain is establishing the transition probabilities. Denote the transition probability from state (x, α_i) to (y, α_j) under control (f^h, g^h) by $\mathbb{P}((x, i), (y, j) | f^h, g^h)$. To determine $\mathbb{P}((x, i), (y, j) | f^h, g^h)$, we have to meet the following three conditions:

(1) (Local moment consistent) It is crucial to determine $\mathbb{P}((x, i), (y, j) | f^h, g^h)$ such that the Markov chain $\{\xi_k^h, k \geq 1\}$ has the same first and second moment with the $Z_t^{f,g}$ in a very small time interval.

(2) (Continuous time Markov chain and value function) To approximate the continuous time controlled state process $Z_t^{f,g}$, we should choose an appropriate continuous time interpolation in any small time epoch. Suppose an interpolation epoch $\Delta t_k^h = \Delta t_k^h(\xi_k^h, e_k^h) > 0, k \geq 1$ is given, define interpolated time $t_k^h = \sum_{s=1}^{k-1} \Delta t_s^h$. Then, a piece-wise constant interpolation ξ_t^h is specified by

$$\xi_t^h = \xi_k^h \text{ for } t \in [t_k^h, t_{k+1}^h). \tag{4.7}$$

The interpolation interval satisfies

$$\inf_{z\in L_h} \Delta t_k^h(z) > 0 \text{ and } \lim_{h\to 0}\sup_{z\in L_h} \Delta t_k^h(z) = 0. \tag{4.8}$$

For the continuous time interpolated process $\{\xi_t^h, e_t^h,\ t \geq 0\}$, define by

$$\tau_h^{f^h,g^h} = \inf\{t : \xi_t^h \bar{\in} [l,u]\} \tag{4.9}$$

the first exit time of Markov chain $\{\xi_t, t \geq 0\}$ and by

$$N^h - 1 \hat{=} \tau_h^{f^h,g^h} \wedge \left[\frac{\tau}{h}\right]. \tag{4.10}$$

An approximating performance function is then defined as

$$J^{f^h,g^h}(t,z,\alpha_i) = \begin{cases} \mathbb{E}_{x,\alpha_i}\left[\sum_{n=0}^{N^h-1} \mathrm{e}^{-\delta nh} c(\xi_n^h) + \mathrm{e}^{-\delta N^h} h(\xi_{N^h}^h)\right], & \text{if } t \geq \tau; \\ \mathbb{E}_{x,\alpha_i}\left[\sum_{n=[\frac{t}{h}]}^{N^h-1} \mathrm{e}^{-\delta nh} c(\xi_n^h) + \mathrm{e}^{-\delta(N^h-\frac{t}{h})} h(\xi_{N^h}^h)\right], & \text{if } t < \tau. \end{cases} \tag{4.11}$$

The upper value function and lower value function of the controlled Markov chain is then given by

$$\overline{V}^h(t,x,\alpha_i) = \inf_{g^h\in\mathcal{A}} \sup_{f^h\in\mathcal{A}} J^{f^h,g^h}(t,z,\alpha_i),$$
$$\underline{V}^h(t,x,\alpha_i) = \sup_{f^h\in\mathcal{A}} \inf_{g^h\in\mathcal{A}} J^{f^h,g^h}(t,z,\alpha_i).$$

Notably, the control problem in this paper consists of an external regulation time, with involves in stopping time of the controlled system. For implementing the computation steps, we need to discretize the integration $\int_0^\infty \lambda_s ds$ as follows. Let

$$\lambda_j^h = \int_{t_{j-1}^h}^{t_j^h} \lambda_s ds \text{ and } \Lambda_k^h = \sum_{j=1}^{k} \lambda_j^h,$$
$$p_j^h = e^{-\lambda_j^h} \text{ and } \bar{F}_k^h = e^{-\Lambda_k^h},\ j = 1,2,\cdots. \tag{4.12}$$

Specifically, the discretized value function $\overline{V}^h(t,z,\alpha_i)$ satisfies the following dynamic equation

$$\begin{aligned} \overline{V}^h(t_k^h,z,\alpha_i) \quad &= \min_{g^h}\max_{f^h}\bigg\{e^{-q_ih}\bigg[\bar{F}_k^h\Big[\mathbb{P}((z,\alpha_i),(z+h,\alpha_i)|f^h,g^h)\overline{V}^h(t_k^h+\Delta t_k^h, z+h) \\ &+\mathbb{P}((z,\alpha_i),(z-h,\alpha_i)|f^h,g^h)\overline{V}^h(t_k^h+\Delta t_k^h, z-h,\alpha_i)\Big] \\ &+p_k^h\mathbb{P}((z,\alpha_i),(z,\alpha_i)|f^h,g^h)\overline{V}^h(t_k^h+\Delta t_k^h, z,\alpha_i)\bigg] \\ &+(1-e^{-q_ih})\frac{q_{ij}}{q_i}\overline{V}^h(t_k^h+\Delta t_k^h, z,\alpha_j) + c(z)\Delta t_k^h\bigg\}. \end{aligned} \tag{4.13}$$

Similarly, we have

$$
\begin{aligned}
\underline{V}^h(t_k^h, z, \alpha_i) \quad &= \max_{f^h} \min_{g^h} \Big\{ e^{-q_i h} \Big[\bar{F}_k^h \Big[\mathbb{P}((z, \alpha_i), (z+h, \alpha_i) | f^h, g^h) \underline{V}^h(t_k^h + \Delta t_k^h, z+h) \\
&+ \mathbb{P}((z, \alpha_i), (z-h, \alpha_i) | f^h, g^h) \underline{V}^h(t_k^h + \Delta t_k^h, z-h, \alpha_i) \Big] \\
&+ p_k^h \mathbb{P}((z, \alpha_i), (z, \alpha_i) | f^h, g^h) \underline{V}^h(t_k^h + \Delta t_k^h, z, \alpha_i) \Big] \\
&+ (1 - e^{-q_i h}) \frac{q_{ij}}{q_i} \underline{V}^h(t_k^h + \Delta t_k^h, z, \alpha_j) + c(z) \Delta t_k^h \Big\}. \qquad (4.14)
\end{aligned}
$$

We know that if the value function of the game exists, then

$$
\lim_{h \to 0} \overline{V}^h(t_k^h, z, \alpha_i) = \lim_{h \to 0} \underline{V}^h(t_k^h, z, \alpha_i) = V(t, z, \alpha_i), i = 1, 2, \cdots. \qquad (4.15)
$$

(3) (Approximating to HJBI equations) Suppose that $V^h(t, z, \alpha_i)$ is given by (4.15). The finite difference method indicates that we need to approximate the first and second derivatives of $V(t, z)$ using step size $h > 0$ as

$$
\overline{V}^h(t, z, \alpha_i) \to \overline{V}(t, z, \alpha_i), \qquad (4.16)
$$

$$
\frac{\overline{V}^h(t+h, z, \alpha_i) - \overline{V}^h(t, z, \alpha_i)}{h} \to \frac{\partial V(t, z, \alpha_i)}{\partial t}, \qquad (4.17)
$$

$$
\frac{\overline{V}^h(t, z+h, \alpha_i) - \overline{V}^h(t, z, \alpha_i)}{h} \to \frac{\partial V(t, z, \alpha_i)}{\partial z}, \text{ if } m_i^h > 0, \qquad (4.18)
$$

$$
\frac{\overline{V}^h(t, z, \alpha_i) - \overline{V}^h(t, z-h, \alpha_i)}{h} \to \frac{\partial V(t, z, \alpha_i)}{\partial z}, \text{ if } m_i^h \leq 0, \qquad (4.19)
$$

$$
\frac{\overline{V}^h(t, z+h, \alpha_i) + \overline{V}^h(t, z-h, \alpha_i) - 2\overline{V}^h(t, z, \alpha_i)}{h^2} \to \frac{\partial^2 \overline{V}(t, z, \alpha_i)}{\partial z^2}. \qquad (4.20)
$$

Now, we turn to determine transition probabilities. For notation convenience, let

$$
\begin{aligned}
&\nu_i^h \hat{=} \nu^2(f^h, g^h, \alpha_i) = f^{h2} \sigma_{1i}^2 + g^{h2} \sigma_{2\alpha_i}^2 - 2 f^h g^h \sigma_{1i} \sigma_{2i} \rho_i \\
&m_i^h \hat{=} f^h \sigma_{1i} \theta_{1i} - g^h \sigma_{2i} \theta_{2i} - f^h g^h \sigma_{1i} \sigma_{2i} \rho_i + g_h^2 \sigma_{2i}^2, \\
&m_i^{h+} = \max\{m_i^h, 0\}, m_i^{h-} = \min\{m_i^h, 0\}, \\
&m_i^h = m_i^{h+} + m_i^{h-}, |m_i^h| = m_i^{h+} - m_i^{h-}, i = 1, 2, \cdots, d. \qquad (4.21)
\end{aligned}
$$

Substituting (4.16)–(4.20) into (4.1) yields

$$
\begin{aligned}
&\Big[h(1 + m_i^h z) + \nu_i^h z^2 - (\lambda_j^h + \delta) h^2 \Big] \overline{V}^h(t_k^h, z, \alpha_i) \\
= \quad & \sup_{f^h} \min_{g^h} \Big\{ \overline{V}^h(t_k^h + \Delta_{t_k}^h, z) h + [m_i^{h+} z h + \frac{1}{2} \nu_i^h z^2] \overline{V}^h(t_k^h + \Delta_{t_k}^h, z+h) \\
& + (m_i^{h-} h z + \frac{1}{2} \nu_i^h z^2) \overline{V}^h(t_k^h + \Delta_{t_k}^h, z-h) \Big\} + c(z) + \Delta_{t_k}^h + \sum_{j=1}^{d} q_{ij} \overline{V}(t_k^h + \Delta_{t_k}^h, z, \alpha_j). \qquad (4.22)
\end{aligned}
$$

By comparing coefficients of Eqs (4.13), (4.14) and (4.22), we can determine the transition probabilities of the constructed Markov chain, which are specified as

$$\begin{aligned}
&\mathbb{P}((z,\alpha_i),(z+h,\alpha_i)|f^h,g^h)\\
=\;&\left(\frac{\bar{F}_k^h}{1-e^{-q_i h}p_k^h}\right)\left(\frac{m_i^{h-}zh+\frac{1}{2}\nu_i^h z^2}{h(1+m_i^h z)+\nu_i^h z^2-(\lambda_j^h+\delta)h^2}\right), \quad (4.23)\\
&\mathbb{P}((z,\alpha_i),(z-h,\alpha_i)|f^h,g^h)\\
=\;&\left(\frac{\bar{F}_k^h}{1-e^{-q_i h}p_k^h}\right)\left(\frac{m_i^{h+}zh+\frac{1}{2}\nu_i^h z^2}{h(1+m_i^h z)+\nu_i^h z^2-(\lambda_j^h+\delta)h^2}\right), \quad (4.24)\\
&\mathbb{P}(z,z,\alpha_i|f^h,g^h)\\
=\;&1-\mathbb{P}((z,\alpha_i),(z+h,\alpha_i)|f^h,g^h)-\mathbb{P}((z,\alpha_i),(z-h,\alpha_i)|f^h,g^h), \quad (4.25)\\
\Delta t_k^h =\;&\left(\frac{\bar{F}_k^h}{1-e^{-q_i h}p_k^h}\right)\left(\frac{h^2}{h(1+m_i^h z)+\nu_i^h z^2-(\lambda_j^h+\delta)h^2}\right). \quad (4.26)
\end{aligned}$$

It is easy to verify that

$$\begin{aligned}
&0<\mathbb{P}((z,\alpha_i),(z+h,\alpha_i)|f^h,g^h),\quad \mathbb{P}((z,\alpha_i),(z-h,\alpha_i)|f^h,g^h)<1,\\
&\mathbb{P}((z,\alpha_i),(z+h,\alpha_i)|f^h,g^h)+\mathbb{P}((z,\alpha_i),(z-h,\alpha_i)|f^h,g^h)<1, \quad (4.27)
\end{aligned}$$

thus, the transition probabilities are well-defined.

It is straightforward to verify that the constructed Markov chain, with transition probabilities specified by Eqs (4.23) and (4.24), meets the local consistency conditions indicated by Conditions (1)–(3). Specifically, under the aforementioned transition probabilities, the constructed Markov chain satisfies the following local consistency properties:

$$\mathbb{E}\Delta\xi_k^h = m_i^h\Delta t_k^h + o(\Delta t_k^h), \quad (4.28)$$

$$Var\Delta\xi_k^h = \nu_i^h\Delta t_k^h + o(\Delta t_h^k). \quad (4.29)$$

Thus far, we have established the transition probabilities for the approximate Markov chain, as defined by Eqs (4.23) and (4.24). By substituting these transition probabilities and the interpolated time epochs into Eqs (4.13) and (4.14), we construct the iteration process for approximating the discrete-time value function with the prescribed boundary conditions

$$V^h(t,l,\alpha_i)=h(l),\quad V^h(t,u,\alpha)=h(u). \quad (4.30)$$

By letting $h\to 0$ and Eqs (4.16)–(4.20), we can then approximate the value function and optimal investment and reinsurance policies numerically.

Remark 4. *Comparing with the algorithms in [26], one may observe that the primary distinction between the two algorithms lies in the modification of the Markov chains transition probabilities by the external regulation time. Consequently, these transition probabilities are no longer time-homogeneous, as they now depend on both the current time and the residual distribution of the regulation time.*

4.2. An illustrative example: goal reaching game

In a typical game problem, Agent A aims to maximize the probability of reaching the upper level u before regulation arrives or reaching the lower level l. The boundary condition for this scenario is given by

$$V^h(t, l, \alpha_i) = 0, \quad V^h(t, u, \alpha_i) = 1. \tag{4.31}$$

We note that this goal-reaching problem is well-known in finance. Optimal control problems on this subject have yielded extensive results. For example, [3] studied the optimization of a bequest goal problem at a random time, specifically the death time of the insured individual. More recent research in this area includes [29]. The parameter settings for this example are provided as follows.

(1) Parameters of environment For the dynamic of environment, we only consider a two-state Markov chain, i.e., $X_t \in (\alpha_1, \alpha_2)$. State $\mathbf{e}_1$ means that the macroeconomic environment is "bad" versus "good". Suppose that the Q-matrix of the Markov chain is given by

$$\begin{pmatrix} -0.1 & 0.1 \\ 0.2 & -0.2 \end{pmatrix}.$$

(2) Parameters of financial market For the financial market, we assume that the parameters are provided in Table 2. The Sharpe ratios for the two players operating in distinct environments are detailed in Table 3. Due to the setup of these parameters, it is apparent that the stock selected by Player A carries higher risk compared to the stock chosen by Player B. Nevertheless, the Sharpe ratio of the stock chosen by Player A exceeds that of Player B. Concurrently, the Sharpe ratio in a bull market surpasses that in a bear market. This parameter configuration is designed to closely mimic real-world conditions in our model.

Table 2. Parameters of financial market for two players.

Player	parameter	bear market	bull market
A	$\mu_\mathbf{1}$	0.08	0.12
A	$\sigma_\mathbf{1}$	0.025	0.03
B	$\mu_\mathbf{2}$	0.06	0.09
B	$\sigma_\mathbf{2}$	0.025	0.03
risk free interest rate	$\mathbf{r}$	0.03	0.05

Table 3. Sharpe ratio for two investors.

Player	Sharp ratio	bear market	bull market
A	$\theta_\mathbf{1}$	2	2.3
B	$\theta_\mathbf{2}$	1.5	1.6

$\mathbf{r} = (r_1, r_2) = (0.01, 0.013), \mu = (\mu_1, \mu_2) = (0.012, 0.018), \sigma_\mathbf{S} = (\sigma_{S1}, \sigma_{S2}) = (0.02, 0.025)$. It can be observed that the Sharpe ratio in a "bad" environment is 0.1, while in a "good" environment it is 0.2. This indicates that the market price in a good environment is higher than in a bad environment.

(3) Parameters of regulation tensity Assume that the force of mortality follows the famous Gompertz-Makeham law of mortality (c.f. [27]), i.e., the hazard rate λ_s is given by

$$\lambda_s = \left(Ae^{Bs} + C\right)\exp\left[-Cs - \frac{A}{B}\left(e^{Bs} - 1\right)\right] \tag{4.32}$$

and

$$\bar{F}(s) = \exp(-\int_0^t \lambda_s \mathrm{d}s) = \exp\left[-Ct - \frac{A}{B}\left(e^{Bt} - 1\right)\right]. \tag{4.33}$$

One can find that the exponential distribution is a special case of the Gompertz-Makeham law. Here we adopted the parameter estimation results in [27] as an example, which are specified by

$$A = 0.0007, B = 0.0006, c = 0.0831. \tag{4.34}$$

The conditional expectation of the residual regulation time, given as current time t, is given by

$$\frac{\frac{1}{B}e^{\frac{A}{B}}(\frac{A}{B} - \ln(\frac{A}{B}) - C)}{\bar{F}(t)}. \tag{4.35}$$

Numerical results show that with the parameters given by (4.34), the expected lifetime is about 79.04. However, for solvency regulation, this time epoch is too long. Based on (4.35), we revise the parameter as

$$A = 0.07, B = 0.06, c = 0.0831. \tag{4.36}$$

Then, the expected regulatory time is 1.1715.

Figure 1 presents the value function with respect to the residual regulation time t and current wealth z. The range of the residual regulation time t is $[0, 2]$, and the wealth interval is $[2, 10]$. To enhance visualization, we standardize the scales of the horizontal and vertical axes, transforming the range of values to $[0, 100]$. From Figure 1, it is evident that the value function of the game problem is smooth and convex over its domain. Figures 2 and 3 illustrate the investment amounts chosen by the two players.

It can be observed that, at the onset of the game, each player tends to invest a significant amount in risky assets. However, as the game nears the "end of regulation time", the investment amounts become more stationary and conservative.

Figure 4 provides a comparison of the value function for the goal-reaching game across different regime scenarios. $\Psi_1(z)$ represents the value function in a "bad" macroeconomic environment, while $\Psi_1(z)$ represents the value function in a "good" environment. One may observe that when current wealth is relatively low, the environment significantly impacts the value of the game. Conversely, when current wealth in the ratio process is relatively high, the value of the game converges. This phenomenon can be interpreted as follows: since both players operate under the same environment, when current wealth is very high, the environment has a diminished impact on the games winning probability.

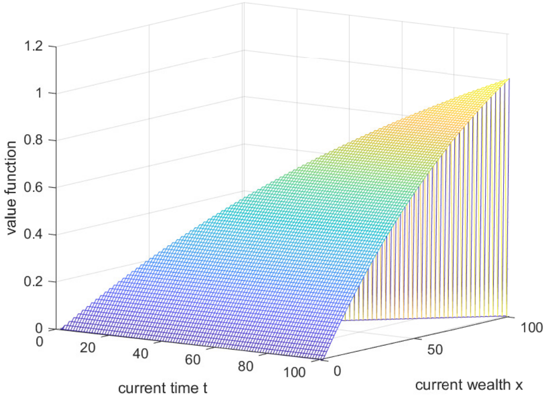

Figure 1. Value function of goal reaching game.

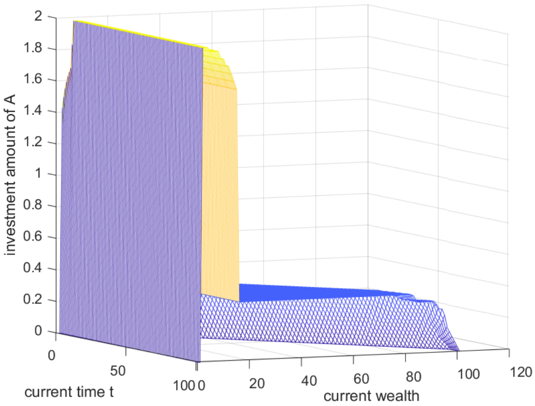

Figure 2. Investment amount of *A*.

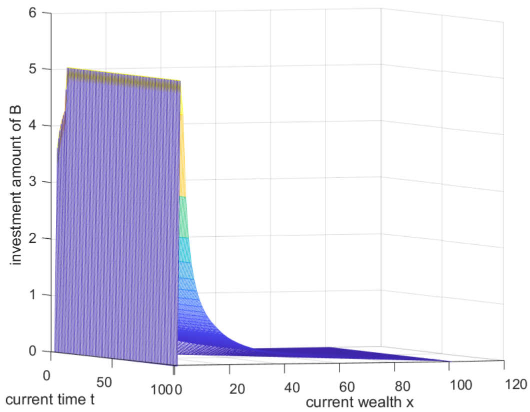

Figure 3. Investment amount of *B*.

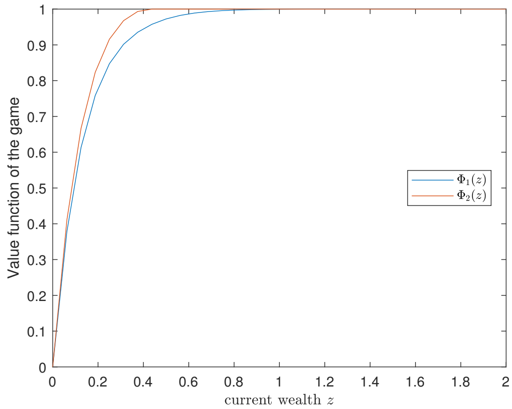

Figure 4. Comparing of value function with different environment.

To assess the sensitivity of numerical results to the parameters in this paper, we proceeded to compare the value function under varying parameters. Table 4 presents the value function for different values of μ_1, with $\mu_2 = 0.08$, while keeping the other parameters constant as listed in Table 3. Similarly, Table 5 shows the value function for different values of μ_2, with $\mu_1 = 0.08$ and the remaining parameters unchanged from Table 3. Upon examining these tables, it is evident that the numerical results exhibit stability in response to changes in parameters. Specifically, regarding the variations in μ_1 and μ_2, Table 4 reveals that an increase in μ_1 leads to an increase in the probability of Player A winning the game, and this marginal effect increases with higher values of μ_1 . However, it is also noteworthy that as the current ratio z increases, the marginal effect due to increases in μ_1 ? diminishes. In contrast, Table 5 indicates that an increase in μ_2 results in a decrease in the probability of Player A winning, yet the marginal effect does not decrease, which differs from the findings in Table 4. Similar to the previous observation, the marginal effect in Table 5 also decreases with higher values of the current ratio z. This suggests that when Player A has a significant advantage over Player B, the benefits gained from selecting risky assets become less crucial for winning the game.

Table 4. Value function with various μ_1.

Current ratio z	$\mu_1 = 0.11$	$\mu_1 = 0.13$	$\mu_1 = 0.15$
0.000000	0.000000	0.000000	0.000000
0.101266	0.265215	0.265609	0.266390
0.202532	0.355853	0.356389	0.357510
0.303797	0.444018	0.444608	0.445943
0.405063	0.531364	0.532000	0.533563
0.506329	0.618786	0.619368	0.620788
0.607595	0.704511	0.705023	0.706283
0.708861	0.788952	0.789406	0.790530
0.810127	0.872271	0.872670	0.873683
0.911392	0.954548	0.954900	0.955817
0.974684	0.995346	1.005855	1.037951

Table 5. Value function with various μ_2

current ratio z	$\mu_2 = 0.04$	$\mu_2 = 0.05$	$\mu_2 = 0.06$
0.000000	0.000000	0.000000	0.000000
0.012658	0.154382	0.154380	0.154378
0.202532	0.347252	0.347247	0.347241
0.303797	0.435608	0.435601	0.435593
0.405063	0.522875	0.522864	0.522855
0.506329	0.610320	0.610307	0.610294
0.607595	0.696173	0.696161	0.696148
0.708861	0.780713	0.780700	0.780688
0.810127	0.864116	0.864103	0.864090
0.911392	0.946462	0.946450	0.946437
0.974684	0.997462	0.997450	0.997438

5. Conclusions

This paper investigates optimal investment games for two investors subject to random time solvency regulations. We first introduce administrative random time regulations into the stochastic investment game problem, providing a practical framework to understand the impact of regulation on fund managers risk management strategies. Additionally, we incorporate regime switching coefficients into the SDEs, enhancing the models applicability across various economic scenarios.

Methodologically, we prove the regularity of the value function when the intensity of regulatory time is a Markov chain, enabling optimal feedback control. By approximating the derivatives of the value function using difference methods, we simplify the numerical computation process. Furthermore, we develop a numerical scheme for the value function when the intensity of regulatory time is a deterministic function of time, utilizing a Markov chain approximation approach to solve PDEs with time-dependent parameters. These methods ensure robust and efficient numerical solutions for complex control problems.

On the other hand, the practical relevance of this paper can be enhanced in several ways. For instance, incorporating scenarios where the two players have distinct regulation intensities and exploring potential dependencies between their intensity processes would enrich the analysis. Additionally, a rigorous proof of the existence of solutions to the HJBI equation using viscosity solution theory could provide stronger theoretical support. Lastly, adopting appropriate statistical methods for parameter calibration is paramount for the practical application of the regime-switching model.

Author contributions

Lin Xu: Validation, Methodology; Linlin Wang: Software, Investigation; Hao Wang: Resources, Writing-review & editing; Liming Zhang: Formal analysis, Software. All authors have read and approved the final version of the manuscript for publication.

Acknowledgments

The authors express their sincere gratitude for the constructive comments provided by the reviewers and editors. Lin Xu was supported by National Natural Science Foundation of Anhui Province [grant number 2408085MA019], LinLin Wang was supported by National Natural Science Foundation of China [grant number 11971034], Hao Wang was supported by National Natural Science Foundation of China [grant number 12301597], Liming Zhang was supported by National Natural Science Foundation of China [grant number 12201006].

Conflict of interest

The authors declare that they have no conflicts of interest.

References

1. B. Avanzi, H. Lau, M. Steffensen, Optimal reinsurance design under solvency constraints, *Scand. Actuar. J.*, **2024** (2024), 383–416. https://doi.org/10.1080/03461238.2023.2257405
2. S. Basak, D. Makarov, Strategic asset allocation in money management, *J. Finance*, **69** (2014), 179–217. https://doi.org/10.1111/jofi.12106
3. E. Bayraktar, V. R. Young, Optimally investing to reach a bequest goal, *Insurance: Math. Econom.*, **70** (2016), 1–10. https://doi.org/10.1016/j.insmatheco.2016.05.015
4. A. Bensoussan, C. C. Siu, S. C. P. Yam, H. Yang, A class of non-zero-sum stochastic differential investment and reinsurance games, *Automatica*, **50** (2014), 2025–2037. https://doi.org/10.1016/j.automatica.2014.05.033
5. T. J. Boonen, W. Jiang, Pareto-optimal reinsurance with default risk and solvency regulation, *Prob. Eng. Inform. Sci.*, **37** (2023), 518–545. https://doi.org/10.1017/S0269964822000079
6. S. Browne, Optimal investment policies for a firm with a random risk process: expo nential utility and minimizing the probability of ruin, *Math. Oper. Res.*, **20** (1995), 937–958. https://doi.org/10.1287/moor.20.4.937
7. S. Browne, Stochastic differential portfolio games, *J. Appl. Prob.*, **37** (2000), 126–147. https://doi.org/10.1239/jap/1014842273
8. R. Buckdahn, J. Li, N. Zhao, Representation of limit values for nonexpansive stochastic differential games, *J. Differ. Equ.*, **276** (2021), 187–227. https://doi.org/10.1016/j.jde.2020.12.009
9. A. Chen, T. Nguyen, M. Stadje, Optimal investment under var-regulation and minimum insurance, *Insurance: Math. Econom.*, **79** (2018), 194–209. https://doi.org/10.1016/j.insmatheco.2018.01.008
10. S. Chen, H. Yang, Y. Zeng, Stochastic differential games between two insurers with generalized mean-variance premium principle, *ASTIN Bull.: J. Int. Actuar. Assoc.*, **48** (2018), 413–434. https://doi.org/10.1017/asb.2017.35

11. M. P. Clements, H. M. Krolzig, Can regime-switching models reproduce the busi ness cycle features of us aggregate consumption, investment and output? *Int. J. Finance Econom.*, **9** (2004), 1–14. https://doi.org/10.1002/ijfe.231

12. J. D. Cummins, Risk-based premiums for insurance guaranty funds, *J. Finance*, **43** (1988), 823–839. https://doi.org/10.1111/j.1540-6261.1988.tb02607.x

13. C. Deng, X. Zeng, H. Zhu, Non-zero-sum stochastic differential reinsurance and investment games with default risk, *Eur. J. Oper. Res.*, **264** (2018), 1144–1158. https://doi.org/10.1016/j.ejor.2017.06.065

14. R. Elliott, The existence of value in stochastic differential games, *SIAM J. Control Optim.*, **14** (1976), 85–94. https://doi.org/10.1137/0314006

15. R. J. Elliott, L. Aggoun, J. B. Moore, *Hidden Markov models*, Berlin: Springer, 1994.

16. R. J. Elliott, J. Van der Hoek, An application of hidden markov models to asset allocation problems, *Finance Stochast.*, **1** (1997), 229–238. https://doi.org/10.1007/s007800050022

17. G. E. Espinosa, N. Touzi, Optimal investment under relative performance concerns, *Math. Finance*, **25** (2013), 221–257. https://doi.org/10.1111/mafi.12034

18. D. Filipovic, R. Kremslehner, A. Muermann, Optimal investment and premium policies under risk shifting and solvency regulation, *J. Risk Insur.*, **82** (2015), 261–288. https://doi.org/10.1111/jori.12021

19. W. H. Fleming, H. M. Soner, *Controlled Markov processes and viscosity solutions*, Berlin: Springer, 2006.

20. W. H. Fleming, P. E. Souganidis, On the existence of value-functions of 2-player, zero-sum stochastic differential-games, *Ind. Uni. Math. J.*, **38** (1989), 293–314.

21. J. Fu, J. Wei, H. Yang, Portfolio optimization in a regime-switching market with derivatives, *Eur. J. Oper. Res.*, **233** (2014), 184–192. https://doi.org/10.1016/j.ejor.2013.08.033

22. J. Grandell, *Aspects of risk theory*, Berlin: Springer, 1991.

23. G. Guan, Z. Liang, A stochastic nash equilibrium portfolio game between two DC pension funds, *Insurance: Math. Econom.*, **70** (2016), 237–244. https://doi.org/10.1016/j.insmatheco.2016.06.015

24. J. D. Hamilton, A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, **57** (1989), 357–384. https://doi.org/10.2307/1912559

25. Z. Jiang, M. Pistorius, Optimal dividend distribution under markov regime switching, *Finance Stochast.*, **16** (2012), 449–476. https://doi.org/10.1007/s00780-012-0174-3

26. Z. Jin, G. Yin, F. Wu, Optimal reinsurance strategies in regime-switching jump diffusion models: stochastic differential game formulation and numerical methods, *Insurance: Math. Econom.*, **53** (2013), 733–746. https://doi.org/10.1016/j.insmatheco.2013.09.015

27. T. Kenan, M. Sucu, Gompertz-makeham parameter estimations and valuation approaches: Turkish life insurance sector, *Eur. Actuar. J.*, **5** (2015), 447–468.. https://doi.org/10.1007/s13385-015-0110-y

28. H. J. Kushner, P. Dupuis, *Numerical methods for stochastic control problems in continuous time*, Berlin: Springer, 2001.

29. D. Li, D. Li, V. R. Young, Optimality of excess-loss reinsurance under a mean variance criterion, *Insurance: Math. Econom.*, **75** (2017), 82–89. https://doi.org/10.1016/j.insmatheco.2017.05.001

30. X. Lin, C. Zhang, T. K. Siu, Stochastic differential portfolio games for an insurer in a jump-diffusion risk process, *Math. Meth. Oper. Res.*, **75** (2012), 83–100. https://doi.org/10.1007/s00186-011-0376-z

31. S. Luo, M. Wang, W. Zhu, Stochastic differential reinsurance games in diffusion approximation models, *J. Comput. Appl. Math.*, **386** (2021), 113252. https://doi.org/10.1016/j.cam.2020.113252

32. J. Marin-Solano, E. V. Shevkoplyas, Non-constant discounting and differential games with random time horizon, *Automatica*, **47** (2011), 2626–2638. https://doi.org/10.1016/j.automatica.2011.09.010

33. A. Melnikov, Y. Romaniuk, Evaluating the performance of gompertz, makeham and leeccarter mortality models for risk management with unit-linked contracts, *Insurance: Math. Econom.*, **39** (2006), 310–329. https://doi.org/10.1016/j.insmatheco.2006.02.012

34. S. Pliska, *Introduction to mathematical finance*, Oxford UK: Blackwell publishers, 1997.

35. P. E. Protter, *Stochastic integration and differential equations*, 2 Eds., Berlin: Springer, 2004. https://doi.org/10.1007/978-3-662-10061-5_6

36. C. S. Pun, H. Y. Wong, Robust non-zero-sum stochastic differential reinsurance game, *Insurance: Math. Econom.*, **68** (2016), 169–177. https://doi.org/10.1016/j.insmatheco.2016.02.007

37. E. V. Shevkoplyas, The hamilton-jacobi-bellman equation for a class of differential games with random duration, *Autom. Remote Control*, **75** (2014), 959–970. https://doi.org/10.1134/S0005117914050142

38. Q. S. Song, Convergence of markov chain approximation on generalized HJB equation and its applications, *Automatica*, **44** (2008), 761–766. https://doi.org/10.1016/j.automatica.2007.07.014

39. Q. S. Song, G. Yin, Z. Zhang, Numerical methods for controlled regime-switching diffusions and regime-switching jump diffusions, *Automatica*, **42** (2006), 1147–1157. https://doi.org/10.1016/j.automatica.2006.03.016

40. Q. Song, G. G. Yin, Z. Zhang, Numerical solutions for stochastic differential games with regime switching, *IEEE Trans. Autom. Control*, **53** (2008), 509–521. https://doi.org/10.1109/TAC.2007.915169

41. A. Swiech, Another approach to the existence of value functions of stochastic differential games, *J. Math. Anal. Appl.*, **204** (1996), 884–897. https://doi.org/10.1006/jmaa.1996.0474

42. M. Taksar, X. Zeng, Optimal non-proportional reinsurance control and stochastic differential games, *Insurance: Math. Econom.*, **48** (2011), 64–71. https://doi.org/10.1016/j.insmatheco.2010.09.006

43. N. Wang, N. Zhang, Z. Jin, L. Qian, Stochastic differential investment and reinsurance games with nonlinear risk processes and var constraints, *Insurance: Math. Econom.*, **96** (2021), 168–184. https://doi.org/10.1016/j.insmatheco.2020.11.004

44. S. Frank, A course in the theory of stochastic processes, *SIAM Rev.*, **24** (1982), 795–815. https://doi.org/10.1137/1024085

45. L. Xu, R. Wang, D. Yao, Optimal stochastic investment games under markov regime switching market, *J. Ind. Manag. Optim.*, **10** (2014), 795–815. http://doi.org/10.3934/jimo.2014.10.795

46. L. Xu, D. Yao, G. Cheng, Optimal investment and dividend for an insurer under a markov regime switching market with high gain tax, *J. Ind. Manag. Optim.*, **13** (2020), 325–356. http://doi.org/10.3934/jimo.2018154

47. H. Yang, L. Zhang, Optimal investment for insurer with jump-diffusion risk process, *Insurance: Math. Econom.*, **37** (2005), 615–634. https://doi.org/10.1016/j.insmatheco.2005.06.009

48. K. F. C. Yiu, J. Liu, T. K. Siu, W. K. Ching, Optimal portfolios with regime switching and value-at-risk constraint, *Automatica*, **46** (2010), 979–989. https://doi.org/10.1016/j.automatica.2010.02.027

49. X. Zeng, Stochastic differential reinsurance games, *J. Appl. Probab.*, **47** (2010), 335–349. https://doi.org/10.1239/jap/1276784895

50. J. Zhu, Singular optimal dividend control for the regime-switching Cramer-Lundberg model with credit and debit interest, *J. Comput. Appl. Math.*, **257** (2014), 212–239. https://doi.org/10.1016/j.cam.2013.08.033

51. J. Zhu, G. Guan, S. Li, Time-consistent non-zero-sum stochastic differential reinsurance and investment game under default and volatility risks, *J. Comput. Appl. Math.*, **374** (2020), 112737. https://doi.org/10.1016/j.cam.2020.112737

Appendix

Appendix A. Proof of Theorem 4.1

Proof. Note that in zero-sum game problem, policies adopted by one investor would be instantaneously obtained by his opponent, and thus, for both investors, the game problem became a min max control problem or max min control problem. For later discussion convenience, we introduce the following shift operator of the Markov process. For detailed introduction on this concept, readers are referred to [44]. Let

$$(X_\cdot, Z_\cdot^{f,g}) : (\mathfrak{E}\times\mathbb{R})^{\mathbb{R}^+} \mapsto (\mathfrak{E}\times\mathbb{R})^{\mathbb{R}^+} \tag{A.1}$$

be the controlled canonical state process. Define the shift operators $\theta_t : (\mathfrak{E}\times\mathbb{R})^{\mathbb{R}^+} \mapsto (\mathfrak{E}\times\mathbb{R})^{[t,\infty)}$ for $t > 0$ by $(\theta_t\omega)s = \omega_{s+t},\ s,t \in \mathbb{R}^+, \omega \in (\mathfrak{E}\times\mathbb{R})^{\mathbb{R}^+}$. It is clear that $\theta_t \in \mathcal{F}_t$ and $\theta_t(X_s, Z_s^{f,g}) = (X_{t+s}, Z_{t+s}^{f,g})$. Let $\tau_0 = 0$, then we have

$$\tau_{n+1} = \tau_n + \theta_{\tau_n}\tau_1,\ n = 0, 1, 2, \cdots \tag{A.2}$$

For given suitable function w and control policies f and g, define two operators $\mathfrak{F}$ and $\mathfrak{W}$ on function w as

$$\mathfrak{F}_w^{f,g}(z,\alpha_i) = \mathbb{E}_{z,\alpha_i}\left[\int_0^{\tau^{f,g}\wedge\tau\wedge\tau_1} \mathrm{e}^{-\delta s}c(Z_s^{f,g})\mathrm{d}s + \mathrm{e}^{-\delta(\tau^{f,g}\wedge\tau\wedge\tau_1)}w(Z^{f,g}_{\tau^{f,g}\wedge\tau\wedge\tau_1})\right]. \tag{A.3}$$

Let

$$\overline{\mathfrak{W}}_w(z,\alpha_i) = \sup_f \inf_g \mathfrak{F}_w^{f,g}(z,\alpha_i),$$

$$\underline{\mathfrak{W}}_w(z,\alpha_i) = \inf_g \sup_f \mathfrak{F}_w^{f,g}(z,\alpha_i) \tag{A.4}$$

and if $\underline{\mathfrak{W}}_w(z,\alpha_i) = \overline{\mathfrak{W}}_w(z,\alpha_i)$, define

$$\mathfrak{W}_w(z,\alpha_i) = \underline{\mathfrak{W}}_w(z,\alpha_i) = \overline{\mathfrak{W}}_w(z,\alpha_i). \tag{A.5}$$

By dynamic programming principle, for any policy g adopted by investor B, the value function for investor A satisfies

$$\begin{aligned}
\overline{V}(z,\alpha_i) \quad &= \sup_f \inf_g \mathbb{E}_{z,\alpha_i}\Big[\int_0^{\tau^{f,g}\wedge\tau\wedge\tau_1} \mathrm{e}^{-\delta s} c(Z_s^{f,g})\mathrm{d}s \\
&+\mathrm{e}^{-\delta\left(\tau^{f,g}\wedge\tau\wedge\tau_1\right)}\Big[\int_{\tau^{f,g}\wedge\tau-\tau^{f,g}\wedge\tau\wedge\tau_1}^{\tau^{f,g}\wedge\tau} \mathrm{e}^{-\delta s} c(Z_s^{f,g})\mathrm{d}s + \mathrm{e}^{-\delta\left(\tau^{f,g}\wedge\tau-\tau^{f,g}\wedge\tau\wedge\tau_1\right)} h(Z^{f,g}_{\tau^{f,g}\wedge\tau\wedge})\Big]\Big] \\
&= \sup_f \inf_g \mathbb{E}_{z,\alpha_i}\Big[\int_0^{\tau^{f,g}\wedge\tau\wedge\tau_1} \mathrm{e}^{-\delta s} c(Z_s^{f,g})\mathrm{d}s + \mathrm{e}^{-\delta\left(\tau^{f,g}\wedge\tau\wedge\tau_1\right)}\overline{V}\left(Z^{f,g}_{\tau^{f,g}\wedge\tau\wedge\tau_1}, X_{\tau^{f,g}\wedge\tau\wedge\tau_1}\right)\Big] \quad (A.6)\\
&= \sup_f \inf_g \mathfrak{F}_V^{f,g}(z,\alpha_i) = \overline{\mathfrak{W}}_V(z,\alpha_i). \quad (A.7)
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\underline{V}\left(Z^{f,g}_{\tau^{f,g}\wedge\tau\wedge\tau_1}, X_{\tau^{f,g}\wedge\tau\wedge\tau_1}\right) = \underline{\mathfrak{W}}_V(Z^{f,g}_{\tau^{f,g}\wedge\tau\wedge\tau_1}, X_{\tau^{f,g}\wedge\tau\wedge\tau_1}) \\
&V(Z^{f,g}_{\tau^{f,g}\wedge\tau\wedge\tau_n}, X_{\tau^{f,g}\wedge\tau\wedge\tau_n}) = \mathfrak{W}_V(Z^{f,g}_{\tau^{f,g}\wedge\tau\wedge\tau_n}, X_{\tau^{f,g}\wedge\tau\wedge\tau_n}).
\end{aligned} \tag{A.8}$$

Thus, if we want to prove the regularities of the value function, we just need to prove the operator $\mathfrak{W}$ is a contractive operator and the candidate polices specified by (3.14 and 3.15) are the right optimal policies.

To see this, from the structure of $\{(f_t^*, g_t^*),\ t \ge 0\}$, we can see that, given the initial state $X_0 = \alpha_i$, the optimal strategy is $f^*(z,\alpha_i), g^*(z,\alpha_i)$ before time τ_1. Hence, if the current state of X_t is ξ_{τ_n}, the optimal strategy is $f^*(Z_{\tau_n}, X_{\tau_n}), g^*(Z_{\tau_n}, X_{\tau_n})$ before the next jump time of ξ_t. By noting that the operator $\mathfrak{F}_V^{f,g}(.,.,.)$ is defined by the path of $(Z_t^{f^*,g^*}, X_t),\ t \ge 0$ up to the first transition time, and using (A.8), we conclude that

$$\overline{V}^{f*,g}(z,\alpha_i) = \mathbb{E}_{(z,\alpha_i)}\left[\int_0^{\tau^{f,g}\wedge\tau\wedge\tau_k} \mathrm{e}^{-\delta s} c(Z_s)\mathrm{d}s + \mathrm{e}^{-\delta\tau_k}\overline{V}(Z_{\tau_k}^{f^*,g^*}, X_{\tau_k})\right], \quad \text{for } k = 1,2,3,\cdots. \tag{A.9}$$

Following the method in [25], we prove Eq (A.9) by the mathematical induction method. It is obviously true for $k = 1$ (see Eq (A.8). Suppose that Eq (A.9) holds for $k = n$. Then,

$$\begin{aligned}
\overline{V}(z,\alpha_i) &= \mathbb{E}_{z,\alpha_i}\left[\int_0^{\tau^{f,g}\wedge\tau\wedge\tau_n} \mathrm{e}^{-\delta s} c(Z_s)\mathrm{d}s + \mathrm{e}^{-\delta\tau_n}\overline{V}(Z_{\tau_n}, X_{\tau_n}) 1_{(\tau_n<\tau^{f,g}\wedge\tau)}\right] \\
&= \mathbb{E}_{z,\alpha_i}\left[\int_0^{\tau^{f,g}\wedge\tau\wedge\tau_n} \mathrm{e}^{-\delta s} c(Z_s)\mathrm{d}s\right] \\
&+ \mathbb{E}_{z,\alpha_i}\left[\mathrm{e}^{-\delta\tau_n}\overline{V}(Z_{\tau_n}, X_{\tau_n}) 1_{(\tau_n<\tau^{f,g}\wedge\tau<\tau_{n+1})}\right]
\end{aligned} \tag{A.10}$$

$$+\quad \mathbb{E}_{(z,\alpha_i)}\left[\mathrm{e}^{-\delta\tau_n}\overline{V}(Z_{\tau_n},X_{\tau_n})1_{(\tau_{n+1}<\tau^{f,g}\wedge\tau)}\right]. \tag{A.11}$$

Note that $\overline{V}(Z_{\tau_n},X_{\tau_n})=\theta_{\tau_n}\overline{V}(Z_0,X_0)$. By the induction hypothesis, we have that

$$\begin{aligned}\overline{V}(Z^{f^*,g^*}_{\tau_n},X_{\tau_n}) &= \theta_{\tau_n}\overline{V}(Z_0,X_0)\\ &= \mathbb{E}_{(z,\alpha_i)}\left[\int_{\tau_n}^{\tau_{n+1}}\mathrm{e}^{-\delta s}c(Z_{\omega_{s+\tau_n}})\mathrm{d}s\Big|\mathcal{F}_{\tau_n}\right] &(A.12)\\ &+\ \mathbb{E}_{(z,\alpha_i)}\left[\mathrm{e}^{-\delta\tau_{n+1}}1_{(\tau_{n+1}<\theta_{\tau_n}\tau^{f,g}\wedge\tau)}\overline{V}(Z^{f^*,g^*}_{\tau_{n+1}},X_{\tau_{n+1}})\Big|\mathcal{F}_{\tau_n}\right]. &(A.13)\end{aligned}$$

Substituting Eqs (A.12) and (A.13) into Eqs (A.10) and (A.11), we have

$$\begin{aligned}\overline{V}(z,\alpha_i) &= \mathbb{E}_{(z,\alpha_i)}\left[\left(\int_0^{\tau^{f,g}}\mathrm{e}^{-\delta s}c(Z_s)\mathrm{d}s\right)1_{(\tau_n<\tau^{f,g}<\tau_{n+1})}+\left(\int_0^{\tau_{n+1}}\mathrm{e}^{-\delta s}c(Z_s)\mathrm{d}s\right)1_{(\tau_{n+1}<\tau^{f,g})}\right]\\ &+\ \mathbb{E}_{(z,\alpha_i)}\left[\mathrm{e}^{-\delta\tau_{n+1}}\overline{V}(Z^{f^*,g^*}_{\tau_{n+1}},X_{\tau_{n+1}})1_{(\tau_{n+1}<\tau^{f,g}\wedge\tau)}\right]\\ &= \mathbb{E}_{(z,\alpha_i)}\left[\int_0^{\tau^{f,g}\wedge\tau\wedge\tau_{n+1}}\mathrm{e}^{-\delta s}c(Z_s)\mathrm{d}s+\mathrm{e}^{-\delta\tau_{n+1}}1_{(\tau_{n+1}<\tau^{f,g}\tau)}\overline{V}(Z^{f^*,g^*}_{\tau_{n+1}},X_{\tau_{n+1}})\right]. &(A.14)\end{aligned}$$

This indicates that (A.9) also holds for $k=n+1$. Since we have proved that $\overline{V}(z,\alpha_i)$ is bounded and we note that $\lim_{n\to\infty}\tau_n=\infty$, letting $n\to\infty$ in the above equation, we have

$$\lim_{n\to\infty}\mathrm{e}^{-\delta\tau_n}1_{(\tau_n<\tau^{f,g}\wedge\tau)}\overline{V}(Z^{f^*,g^*}_{\tau_n},X_{\tau_n})=\mathbb{E}\left[e^{-\delta\tau^{f,g}\wedge\tau}h(X_{\tau^{f,g}\wedge\tau})\right] \tag{A.15}$$

and this indicates that under the policy (f^*,g^*), the performance function is really the value function and the operator is a conductive operator. □

Appendix B. Proof of Theorem 3.2

Proof. The idea of proof is: for a certain investment strategy of B, investor A chooses the corresponding optimal investment strategy and applies the same method to B. Then, by Eq (2.16), we find the optimal differential game policies. Specifically, for any policy g adopted by investor B, the HJBI equation of investor A for maximizing $v^{f,g}(t,z,\alpha_i)$ is

$$\sup_f\left\{\mathcal{A}^{f,g}v^{f,g}(t,z,\alpha_i)+c-(\delta+\lambda_t)v^{f,g}(t,z,\alpha_i)\right\}=0. \tag{A.16}$$

Denote by $\tilde{f}(t,z,\alpha_i:g)$ the maximizer for investor A of Eq (A.16) under given policy g, and we further have

$$\mathcal{A}^{\tilde{f},g}v^{\tilde{f},g}+c-(\delta+\lambda_t)v^{\tilde{f},g}=0. \tag{A.17}$$

Assuming that $v^{\tilde{f},g}_{zz}<0$ and then differentiating (A.17) on both sides w.r.t f yields that the maximizer $\tilde{f}(t,z,\alpha_i:g)$ is of the form

$$\tilde{f}(t,z,\alpha_i:g)=\frac{g\sigma_{2i}\rho_i}{\sigma_{1i}}\left(1+\frac{v^{*,g}_z(t,z,\alpha_i)}{zv^{*,g}_{zz}(t,z,\alpha_i)}\right)-\frac{\theta_{1i}}{\sigma_{1i}}\frac{v^{*,g}_z(t,z,\alpha_i)}{zv^{*,g}_{zz}(t,z,\alpha_i)} \tag{A.18}$$

where

$$v^{*,g}(t,z,\alpha_i) = \sup_f v^{f,g}(t,z,\alpha_i) = v^{\tilde{f},g}(t,z,\alpha_i).$$

Obviously,

$$\inf_g v^{*,g}(t,z,\alpha_i) = \inf_g \sup_f v^{f,g}(t,z,\alpha_i) = \bar{V}(t,z,\alpha_i)$$

is the upper value of SDG.

Similarly, for any given policy f adopted by Investor A, the HJBI equation of investor B for minimizing $v^{f,g}$ is given by

$$\inf_g \{\mathcal{A}^{f,g} v(t,z,\alpha_i) + c - (\delta + \lambda_t) v(t,z,\alpha_i)\} = 0,$$

then the minimizer for Investor B is specified by

$$\tilde{g}(t,z,\alpha_i : f) = \frac{\theta_{2i}}{\sigma_{2i}} \frac{z v_z^{f,*}(t,z,\alpha_i)}{2z v_z^{f,*}(t,z,\alpha_i) + z^2 v_{zz}^{f,*}(t,z,\alpha_i)} + f\rho_i \frac{\sigma_{1i}}{\sigma_{2i}} \frac{z v_z^{f,*}(t,z,\alpha_i) + z^2 v_{zz}^{f,*}(t,z,\alpha_i)}{2z v_z^{f,*}(t,z,\alpha_i) + z^2 v_{zz}^{f,*}(t,z,\alpha_i)}, \quad \text{(A.19)}$$

where $v^{f,*} = \inf_g v^{f,g} = v^{f,\tilde{g}}$ and also

$$\sup_f \inf_g v^{f,g}(t,z,\alpha_i) = \sup_f v^{f,*} = \sup_f v^{f,\tilde{g}(f)} = \underline{V}(t,z,\alpha_i)$$

is the lower value function of SDG. Since we have shown that the saddle point of SDG (2.18) exists, then the game must have an achievable value with

$$v^{*,\tilde{g}} = v^{\tilde{f},*}.$$

If this is the case, then we can substitute Eq (A.19) into Eq (A.18). By some manipulations, one can find that

$$f^*(t,z,\alpha_i) = \frac{\theta_{1i}}{\sigma_{1i}} \left(\frac{v_z(t,z,\alpha_i)}{\Theta v(t,z,\alpha_i)} \right) \left[\left(\frac{\rho}{k_i} - 1 \right) (v_z(t,z,\alpha_i) + z v_{zz}(t,z,\alpha_i)) - v_z(t,z,\alpha_i) \right]. \quad \text{(A.20)}$$

With a vice versa, it results in

$$g^*(t,z,\alpha_i) = \frac{\theta_{2i}}{\sigma_{2i}} \left(\frac{v_z(t,z,\alpha_i)}{\Theta v(t,z,\alpha_i)} \right) \Big[(1 - \rho_i k_i) (v_z(t,z,\alpha_i) + z v_{zz}(t,z,\alpha_i)) - v_z(t,z,\alpha_i) \Big]. \quad \text{(A.21)}$$

Substituting (A.20) and (A.21) into (A.17) we finally find that the equations satisfied by value function $v(t,z,\alpha_i)$ are of the form

$$\begin{aligned} & v_t + \frac{z v_z(t,z,\alpha_i)^2}{2\Theta v(t,z,\alpha_i)} \theta_{2i}^2 \left[(1 - k_i^2) v_z(t,z,\alpha_i) - (1 + k_i^2 - 2\rho k_i) (v_z(t,z,\alpha_i) + z v_{zz}(t,z,\alpha_i)) \right] \\ & + c(z) - (\delta + \lambda_t) v(t,z,\alpha_i) + \sum_{j=1}^d v(t,z,\alpha_j) = 0, \quad i = 1,2,\cdots,d, \end{aligned} \quad \text{(A.22)}$$

and naturally with boundary conditions

$$v(t,l,\alpha_i) = h(l) \quad \text{and} \quad v(t,u,\alpha_i) = h(u) \quad \text{for} \quad i = 1,2,\cdots,d.$$

One may find that Eq (4.1) is just the reformulation of Eq (A.22). Thus when the value function of the game exists and smooth enough, it solves the coupled HJB equation Eq (4.1). On the other hand, we need to verify that the solutions to coupled Eq (4.1) is the value function of SDG. Although we can rely on the result of [20] to complete our proof, here we want to prove our "verification" theorem by the "Martingale optimality principle" which is widely used in many literatures. For example, see [47].

Suppose that $w(t, z, \alpha_i),\ i = 1, 2, \cdots, d$ are the solutions to couple equations (4.1). For any policy pair (f, g), define a process

$$M_h^{f,g} := \mathrm{e}^{-\delta h} w(t+h, Z_{t+h}^{f,g}, X_{t+h}) + \int_t^{t+h} \mathrm{e}^{-\delta s} c(Z_s^{f,g}) ds. \tag{A.23}$$

Note that $\{(Z_t^{f,g}, X_t), t \geq 0\}$ is a vector valued Markov process and $\{X_t, t \geq 0\}$ is a process with bounded total variation on finite time interval; thus, by Itô's lemma (see [35]), we have that for any $t \wedge \tau^{f,g}$,

$$\begin{aligned} M_h^{f,g} &= M_0^{f,g} + \int_t^{(t+h)\wedge\tau^{f,g}\wedge\tau} \mathrm{e}^{-\delta s} \left[w_s(s, Z_s^{f,g}, X_s) + \mathcal{A} w(s, Z_s^{f,g}, X_s) + c(Z_s^{f,g}) - \delta w(s, Z_s^{f,g}, X_s) \right] ds \\ &+ \int_t^{(t+h)\wedge\tau^{f,g}\wedge\tau} \mathrm{e}^{-\delta s} Z_s^{f,g} w_z(s, Z_s^{f,g}, X_s) \left[f_s \sigma_1(X_s) dW_s^{(1)} - g_s \sigma_2(X_s) dW_s^{(2)} \right]. \end{aligned}$$

If Eq (4.2) of Theorem 4.1 holds, then for any $t \geq 0$, $M_{t\wedge\tau^{f,g}\wedge\tau}^{f,g}$ is a local martingale, and further, if Eq (4.3) holds, then $M_{t\wedge\tau^{f,g}\wedge\tau}^{f,g}$ is uniformly integrable sup-martingale or sub-martingale, depending on the choice of investment policies f and g. Note that $\mathbb{P}(\tau^{f,g} \wedge \tau < \infty) = 1$ and interval $[l, u]$ is bounded, supposing that $(Z_t^{f,g}, X_t) = (z, \alpha_i)$; and letting $h \to \infty$, it is easy to find that

$$\begin{aligned} v^{f,g}(t, z, \alpha_i) &= \mathbb{E}_{t,z,\alpha_i} \left[1_{\{\tau\wedge\tau^{f,g} > t\}} \int_t^{\tau^{f,g}\wedge\tau} \mathrm{e}^{-\delta t} c(Z_t^{f,g}) ds + \mathrm{e}^{-\delta(\tau^{f,g}\wedge\tau)} h(Z_{\tau^{f,g}\wedge\tau}^{f,g}) \right] \\ &= \mathbb{E}_{t,z,\alpha_i} \left[M_{\tau^{f,g}\tau}^{f,g} \right] \\ &= \mathbb{E}_{t,z,\alpha_i} \left[M_h^{f,g} + \int_{t+h}^{\tau^{f,g}} \mathrm{e}^{-\delta s} \left[\mathcal{A}^{f,g} w(s, Z_s^{f,g}, X_s) + c(Z_s^{f,g}) - (\delta + \lambda_s) w(s, Z_s^{f,g}, X_s) \right] ds \right] \\ &= M_0^{f,g} + \mathbb{E}_{t,z,\alpha_i} \left[\int_t^{\tau^{f,g}} \mathrm{e}^{-\delta s} \left[\mathcal{A}^{f,g} w(s, Z_s^{f,g}, X_s) + c(Z_s^{f,g}) - (\delta + \lambda_s) w(s, Z_s^{f,g}, X_s) \right] ds \right] \\ &= w(t, z, \alpha_i) + \mathbb{E}_{t,z,\alpha_i} \Big[\int_t^{\tau^{f,g}\wedge\tau} \mathrm{e}^{-\delta s} \Big[\mathcal{A} w(s, Z_s^{f,g}, X_s) \\ &+ c(Z_s^{f,g}) - (\delta + \lambda_s) w(s, Z_s^{f,g}, X_s) \Big] ds \Big]. \end{aligned} \tag{A.24}$$

For given policy g^* adopted by Investor B and any policy f adopted by Investor A, Equation (A.16) implies that

$$\mathcal{A}^{f,g^*} w + c - (\delta + \lambda_s) w \leq 0$$

and

$$\mathcal{A}^{f^*,g^*} w + c - (\delta + \lambda_t) w = 0.$$

Thus, by Eq (A.24) we have

$$v^{f,g^*}(t, z, \alpha_i) \leq w(t, z, \alpha_i) \tag{A.25}$$

and $v^{f^*,g^*}(t,z,\alpha_i) = w(t,z,\alpha_i)$. By Eq (A.19), with a similar discussion, we find that for given policy f^* we have

$$\mathcal{A}^{f^*,g} w + c - (\delta + \lambda_s) w \geq 0,$$
$$\mathcal{A}^{f^*,g^*} w + c - (\delta + \lambda_s) w = 0$$

and

$$v^{f^*,g}(t,z,\alpha_i) \geq w(t,z,\alpha_i)$$

and

$$v^{f^*,g^*}(t,z,\alpha_i) = w(t,z,\alpha_i). \tag{A.26}$$

Together with Eqs (A.25) and (A.26), we have Eq (2.16), i.e.,

$$v^{f,g^*}(t,z,\alpha_i) \leq v^{f^*,g^*}(t,z,\alpha_i) = w(t,z,\alpha_i) \leq v^{f^*,g}(t,z,\alpha_i)$$

which proves that the solutions of the coupled Eq (4.1) are the value of SDG. □

AIMS Press