A Numerical Approach to Optimal Dividend Policies with Capital Injections and Transaction Costs

Zhuo Jin,*   Hailiang Yang,†   G. Yin‡

Abstract

This work focuses on numerical methods for finding optimal dividend payment and capital injection policies to maximize the present value of the difference between the cumulative dividend payment and the possible capital injections. Using dynamic programming principle, the value function obeys a quasi-variational inequality (QVI). The state constraint of the impulsive control gives rise to a capital injection region with free boundary. Since the closed-form solutions are virtually impossible to obtain, we use Markov chain approximation techniques to construct a discrete-time controlled Markov chain to approximate the value function and optimal controls. Convergence of the approximation algorithms is proved.

Key Words. Control, singular control, dividend policy, capital injection, free boundary, Markov chain approximation.

AMS Subject Classification. 65C30, 60H35, 65C05, 91B28, 91B70.

---

*Centre for Actuarial Studies, Department of Economics, The University of Melbourne, VIC 3010, Australia, zjin@unimelb.edu.au.

†Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong. Email: hlyang@hkusua.hku.hk.

‡Department of Mathematics, Wayne State University, Detroit, Michigan 48202, gyin@math.wayne.edu.

1 Introduction

Designing dividend payment policies has long been an important issue in actuarial science and finance literature. The dividend payment decision is crucial because not only does it represent an important signal about a firm's future growth opportunities and profitability but also may influence the investment and financing decisions of firms and the wealth of the policyholders. For insurance companies, because of the nature of their products, insurers tend to accumulate relatively large amounts of cash, cash equivalents, and investments in order to pay future claims and avoid financial ruin. The study of insurance companies' dividend decisions is thus desirable because the payment of dividends to shareholders may reduce an insurer's ability to survive adverse investment and underwriting experience. Recently, the financial crisis has led to the controversial discussion on the dividend policy of European insurance industry; see Reddemann et al. (2010).

Stochastic optimal control problems on dividend strategies for an insurance corporation have drawn increasing attention since the introduction of the optimal dividend payment model proposed by De Finetti (1957). Recently there have been increasing efforts on applying advanced methods of stochastic control to study the optimal dividend policy; see Asmussen and Taksar (1997), Gerber and Shiu (2004), and Schmidli (2008). As an extension of the previous work, the dividend is assumed to be paid out with the constraint that a transaction cost must be paid. The transaction cost includes two parts: a proportional cost and a fixed cost. The studies related to optimal dividend problems with transaction costs can be referred to Thonhauser and Albrecher (2011), He and Liang (2009), and Yao et al. (2010). On the other hand, to maximize the expected total discounted dividend payments, the company will bankrupt almost surly if the dividend payment is paid out as a barrier strategy. In practice, Dickson and Waters (2004) suggested that capital injections can be taken into account to maintain the business when cashflow is insufficient. That is, whenever the company is on the verge of bankruptcy, the company can raise sufficient funds to survive. Furthermore, penalty will be paid at the time of ruin, which can be considered as the transaction cost (proportional and fixed) of capital injection; see Sethi and Taksar (2002), Kulenko and Schimidli (2008), and Yao et al. (2011).

In this work, we model the surplus as a jump process. Capital injections and dividend payment policies with transaction costs are introduced as impulse and singular stochastic controls. Whenever the company is on the verge of ruin, the company can raise sufficient funds to survive. A natural payoff function is to maximize the difference between the expected total discounted dividend payment and the capital injections under the optimal controls. Scheer and Schmidli (2011) consider the capital injections with administration costs when surplus hits 0 and provides an algorithm to compute the size of constant capital injections. Comparing to the work in Scheer and Schmidli (2011), we consider a more general capital injection policy, where the capital injections are exerted when surplus hits not only 0 but also a sufficiently low threshold. To maximize the performance, the impulse controls of capital injections depend on the surplus processes and can be very large, which results in a free boundary problem and adds more difficulty to analyze the optimal policies. The formulation of our model is general and versatile. Due to the inclusion of both impulse and singular controls, closed-form solutions are virtually impossible to obtain. Thus we focus on

developing numerical solutions.

To find the optimal capital injection and dividend payment strategies, one usually solves a so-called Hamilton-Jacobi-Bellman (HJB) equation. However, in our work, because of the impulse and singular control formulation, the HJB equation is in fact a quasi-variational inequality (QVI) with free boundary. A closed-form solution is virtually impossible to obtain. A viable alternative is to employ numerical approximations. In this work, we adapt the Markov chain approximation methodology developed by Kushner and Dupuis (2001). To the best of our knowledge, the related results are relatively scarce; Budhiraja and Ross (2007) and Kushner and Martins (1991) are the only papers that carry out a convergence analysis using weak convergence and relaxed control formulation of numerical schemes for singular control problems. In a recent work, Jin et al. (2012) developed numerical algorithms for approximating optimal reinsurance and dividend payment policies. In that paper, singular controls of dividend payments are considered. This paper further treats models with jump claims and impulse controls. Moreover, the impulse controls of capital injections depend on the surplus and can be very large. These state-dependent capital injections lead to the formulation of a free boundary problem. Roughly, as demonstrated in the numerical experiments, the state-dependent "threshold" curve separates the capital injection region and continuation region. Thus, the problem becomes more complex and difficult to handle. We should add that the Markov chain approximation method requires little regularity of the value function and analytic properties of the QVI. The numerical implementation can be done using either value iterations or policy iterations.

The rest of the paper is organized as follows. A general formulation of optimal capital injection strategies and dividend policies are presented in Section 2. Section 3 deals with the numerical algorithm of Markov chain approximation method. The Poisson jumps, impulse control and the singular control are well approximated by the approximating Markov chain and the dynamic programming equation are presented. Section 4 deals with the convergence of the approximation scheme. The technique of "rescaling time" is introduced and the convergence theorems are proved. Two numerical examples are provided in Section 5 to illustrate the performance of the approximation method in different claim severity distributions. Finally, some additional remarks are provided in Section 6.

2   Formulation

We assume that the surplus $X(t)$ of an insurance company in the absence of dividend payments and capital injections satisfies the classical Cramér Lundberg process,

$$X(t) = x + ct - S(t), \quad t \geq 0. \tag{2.1}$$

where x is the initial surplus, the constant c is the rate of premium. Let ϖ_n be the inter-arrival time of the nth claim,

$$\nu_n = \sum_{j=1}^{n} \varpi_j, \quad \text{and} \quad N(t) = \max\{n \in \mathbb{N} : \nu_n \leq t\}$$

is the number of claims up to time t, which is a Poisson counting process.

$$S(t) = \sum_{i=1}^{N(t)} Y_i$$

is a compound Poisson process representing claims with arriving rate λ. Y is assumed to be the magnitude of the claim size and has the distribution $\Pi(\cdot)$. Let $f(y)$ (for $y \geq 0$) and μ denote the probability density and the mathematical expectation of Y respectively. Then the Poisson measure $N(\cdot)$ has intensity $\lambda dt \times \Pi(dy)$ where $\Pi(dy) = f(y)dy$. We are now working on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, where $\mathcal{F}_t$ is the σ-algebra generated by $\{N(s) : 0 \leq s \leq t\}$.

A dividend strategy $Z(\cdot)$ is an $\mathcal{F}_t$-adapted process $\{Z(t) : t \geq 0\}$ corresponding to the accumulated amount of dividends paid up to time t such that $Z(t)$ is a nonnegative and nondecreasing stochastic process that is right continuous with left limits. Throughout the paper, we use the convention that $Z(0^-) = 0$. The jump size of Z at time $t \geq 0$ is denoted by $\Delta Z(t) := Z(t) - Z(t^-)$, and $Z^c(t) := Z(t) - \sum_{0 \leq s \leq t} \Delta Z(s)$ denotes the continuous part of $Z(t)$.

The capital injection process $L(t) = \sum_{n=1}^{\infty} I_{\{\tau_n \leq t\}} \zeta_n$ is described by a sequence of increasing stopping times $\{\tau_n, n = 1, 2, \ldots\}$ and a sequence of random variables $\{\zeta_n, n = 1, 2, \ldots\}$, which represent the times and the sizes of capital injections. A control policy π is described by $\pi = \{Z; L\} = \{Z; \tau_1, \ldots, \tau_n, \ldots; \zeta_1, \ldots, \zeta_n, \ldots\}$. Assume the evolution of $X(t)$, subject to capital injections and dividend payments, follows a one-dimensional process on an unbounded domain $G' = (0, \infty)$. The controlled asset process is

$$X(t) = x + ct - S(t) - Z(t) + \sum_{n=1}^{\infty} I_{\{\tau_n \leq t\}} \zeta_n \tag{2.2}$$

Remark 2.1. Note that the drift c describes the premium magnitude collected by the insurance company. It is used to determine the amount charged by the insurer and mainly depends on the insurance coverage, not surplus. However, from a numerical approximation point of view, making c X-dependent will not introduce any essential difficulty.

In this paper, we assume that the shareholders need to pay $K + \beta_2 \zeta, \beta_2 > 1$, to meet the capital injection of ζ. $K > 0$ is the fixed transaction costs, $(\beta_2 - 1)\zeta$ is the proportional transaction costs. We assume that if the company pay l as dividends, the shareholders can get $\beta_1, 0 < \beta_1 < 1$, and we omit the fixed transaction costs in the dividends payout process. Denote by $r > 0$ the discounting factor. For an arbitrary admissible pair $\pi = (Z, L)$, the performance function is

$$J(x, \pi) = E_x \left[\int_0^{\infty} e^{-rt} \beta_1 dZ - \sum_{n=1}^{\infty} e^{-r\tau_n} (K + \beta_2 \zeta_n) I_{\{\tau_n < \infty\}} \right]. \tag{2.3}$$

The pair $\pi = (Z, L)$ is said to be *admissible* if Z and L satisfy

(i) $Z(t)$ and $L(t)$ are nonnegative for any $t \geq 0$,

(ii) Z is right continuous with left limit (càdlàg), nondecreasing and adapted to $\mathcal{F}_t$,

(iii) τ_n is a stopping time w.r.t. $\mathcal{F}_t$, and $0 \le \tau_1 < \ldots < \tau_n < \ldots$, a.s.

(iv) ζ_n is measurable w.r.t. $\mathcal{F}_t$,

(v) $P(\lim_{n \to \infty} \tau_n < T) = 0, \forall T > 0$, and

(vi) $J(x, \pi) < \infty$ for any x and admissible pair $\pi = (Z, L)$, where J is the functional defined in (2.3).

Suppose that $\mathcal{A}$ is the collection of all admissible pairs. Define the value function as

$$V(x) := \sup_{\pi \in \mathcal{A}} J(x, \pi). \tag{2.4}$$

For an arbitrary $\pi \in \mathcal{A}$, and $V(\cdot) \in C^2(\mathbb{R})$, define an operator $\mathcal{L}^\pi$ by

$$\mathcal{L}^\pi V(x) = cV_x(x) + \lambda \int_0^x [V(x-y) - V(x)]f(y)dy, \tag{2.5}$$

where V_x denotes the first derivatives with respect to x. Define another capital injection operator $\mathcal{M}$ by

$$\mathcal{M}V(x) = \sup_{\widetilde{y} \ge 0} \{V(x + \widetilde{y}) - \beta_2 \widetilde{y} - K\} \tag{2.6}$$

If the value function V defined in (2.4) is sufficiently smooth, by applying the dynamic programming principle (Fleming and Soner (2006)), we formally conclude that V satisfies the following quasi variational inequality (QVI):

$$\max \left\{ \mathcal{L}^\pi V(x) - rV(x), \beta_1 - V_x(x), \mathcal{M}V(x) - V(x) \right\} = 0. \tag{2.7}$$

Similar to Yao et al. (2011), we divide the set of the surplus to three regions

(i) Continuation region:

$$\mathcal{C} := \{\mathcal{L}^\pi V(x) - rV(x) = 0, \beta_1 < V_x(x), \mathcal{M}V(x) < V(x)\}$$

(ii) Dividend payout region:

$$\mathcal{D} := \{\mathcal{L}^\pi V(x) - rV(x) < 0, \beta_1 = V_x(x), \mathcal{M}V(x) < V(x)\}$$

(iii) Capital injection region:

$$\mathcal{I} := \{\mathcal{L}^\pi V(x) - rV(x) < 0, \beta_1 < V_x(x), \mathcal{M}V(x) = V(x)\}.$$

Boundary Conditions. The capital injection will be taken into account when there is not enough solvency capital to maintain the business. Intuitively, on the boundary of the capital injection region, the value function obeys

$$V(x) = \sup_{\widetilde{y} \ge 0} \{V(x + \widetilde{y}) - \beta_2 \widetilde{y} - K\} \tag{2.8}$$

To make the company run continuously, the capital injections will definitely occur at the moments when $x = 0$. In addition, the capital injections also occur whenever the surplus is sufficiently low. The impulse control of capital injections is dependent on the surplus states and leads to a free boundary of the capital injection region. Furthermore, we also need boundary conditions at ∞ for x in the computation. Since the surplus cannot reach infinity, we only need to choose B large enough and compute the value function in the finite interval $G = [0, B]$. To make it computationally feasible, we truncate x at some large value B. When B is large enough, it follows

$$V_x(B) = \beta_1. \tag{2.9}$$

That is, the dividend payout strategy is a barrier strategy. Whenever the surplus exceed some barrier, the excess is paid out immediately as dividend.

We consider the dividend payout strategy with the capital injection as a band strategy. The decision maker will take no action until the surplus reaches the lower barrier, where a impulse control of capital injection will be taken. The dividend will be paid out immediately when the surplus reaches the upper barrier. Combing(2.7), (2.8) and (2.9), the quasi variational inequality (QVI) with the boundary conditions follows

$$\begin{cases} \max\left\{\mathcal{L}^\pi V(x) - rV(x), \beta_1 - V_x(x), \mathcal{M}V(x) - V(x)\right\} = 0, \\ V_x(B) = \beta_1, \\ V(0) = \sup_{0 \leq \widetilde{y} \leq B} \{V(\widetilde{y}) - \beta_2\widetilde{y} - K\}. \end{cases} \tag{2.10}$$

Remark 2.2. The value function V is not necessarily smooth. In fact, Bayraktar et al. (2011) provide an example in which the value function is not even continuous. While a classical solution of the QVI cannot be obtained, an alternative definition for a solution to the quasi-variational inequalities (2.10) is that of a viscosity solution (See Song et al. (2011)). However, we focus on the numerical solutions in this work, and the definition of viscosity solution will not lead any difficulty in numerical approximation.

3   Numerical Algorithm

In this section, we deal with the numerical algorithm of Markov chain method for solving the QVI. Following the original idea of Markov approximation techniques in Kushner and Dupuis (2001), we will construct a two-component discrete-time controlled Markov chain to approximate the process with impulse and singular control terms. The discrete-time and finite-state controlled Markov chain is so defined that it is locally consistent with (2.2). In this problem, the process contains a term of Poisson jumps. To construct the approximating Markov chain, let us recall some facts of Poisson random measure which is useful for the convergence theorems.

To proceed, we need introduce an equivalent way to define the process (2.2) by working with the claim times and values. Set $\nu_0 = 0$, and let ν_n, $n \geq 1$, denote the time of the nth claim, and Y_n is the corresponding claim intensity. Let $\{\nu_{n+1} - \nu_n, Y_n, n < \infty\}$ be mutually independent random variables with $\nu_{n+1} - \nu_n$ being exponentially distributed with mean $1/\lambda$.

Furthermore, let $\{\nu_{k+1} - \nu_k, Y_k, k \geq n\}$ be independent of $\{X(s), s < \nu_n, \nu_{k+1} - \nu_k, Y_k, k < n\}$, and the claim amount $S(t)$ can be written as

$$S(t) = \sum_{\nu_n \leq t} Y_n.$$

We note the local properties of claims for (2.2). Because $\nu_{n+1} - \nu_n$ is exponentially distributed, we can write

$$P\{\text{claim occurs on } [t, t + \Delta)|X(s), N(s), s \leq t\} = \lambda\Delta + o(\Delta). \tag{3.1}$$

By the independence and the definition of Y_n, for any $H \in \mathcal{B}(\mathbb{R}_+)$, we have

$$P\left\{X(t) - X(t^-) \in H | t = \nu_n \text{ for some } n; X(s), N(s), s < t; X(t^-) = x, \right\} = \Pi(Y_n \in H). \tag{3.2}$$

It is implied by the above discussion that $X(\cdot)$ satisfying (2.2) can be viewed as a process that involves impulse control and jump terms according to the claim rate defined by (3.1). Given that the nth claim occurs at time ν_n, we construct the values according to the conditional probability law (3.2) or, equivalently, write it as Y_n. Then the process given in (2.2) is a deterministic process until the time of the next claim.

To begin with, we construct a discrete-time, finite-state Markov chain to approximate the process without claims, with the dynamic system

$$X(t) = x + ct - Z(t) + \sum_{n=1}^{\infty} I_{\{\tau_n \leq t\}}\zeta_n \tag{3.3}$$

Let $h > 0$ be a discretization parameter. Define $S_h' = \{x : x = kh, k = 0, \pm 1, \pm 2, \dots\}$ and $S_h = S_h' \cap G$. Moreover, assume without loss of generality that the boundary point B is an integer multiple of h. Let $\{\xi_n^h, n < \infty\}$ be a controlled discrete-time Markov chain on S_h and denote by $p_D^h((x, y)|\pi^h)$ the transition probability from a state x to another state y under the control π^h. We need to define p_D^h so that the chain's evolution well approximates the local behavior of the controlled process (2.2). At any discrete time n, we can either exercise a capital injection, a dividend payment or no action. That is, if we put $\Delta\xi_n^h = \xi_{n+1}^h - \xi_n^h$, then

$$\Delta\xi_n^h = \Delta\xi_n^h I_{\{\text{capital injection at } n\}} + \Delta\xi_n^h I_{\{\text{dividend payment at } n\}} + \Delta\xi_n^h I_{\{\text{no action at } n\}}. \tag{3.4}$$

The chain and the control will be chosen so that there is exactly one term in (3.4) is nonzero. Denote by $\{I_n^h : n = 0, 1, \dots\}$ a sequence of control actions, where $I_n^h = 0, 1$ or 2, if we exercise a capital injection, a dividend payment or no action at time n, respectively.

If $I_n^h = 0$, then we denote by Δl_n^h the impulse control that is the capital injection for the chain at time n. Note that $\Delta\xi_n^h = \Delta l_n^h$. If $I_n^h = 1$, or $\xi_n^h = B + h$, dividend payment is exerted definitely as a singular control. Dividend is paid out to lower the surplus level. Moreover, we require this dividend payment takes the state from $B + h$ to B. That is, if we denote by Δz_n^h the random variable that is the reflection action for the chain at time n, then $\Delta\xi_n^h = -\Delta z_n^h = -h$.

If $I_n^h = 2$, the decision makers will take no action. Let $\tilde{\Delta}t^h(\cdot) > 0$ be the *interpolation interval* on S_h. Assume $\inf_x \tilde{\Delta}t^h(x) > 0$ for each $h > 0$ and $\lim_{h \to 0} \sup_x \tilde{\Delta}t^h(x) \to 0$. Let $E_{x,n}^{h,2}$, $\mathrm{Var}_{x,n}^{h,2}$ and $P_{x,n}^{h,2}$ denote the conditional expectation, variance, and marginal probability given $\{\xi_k^h, I_k^h, k \le n, \xi_n^h = x, I_n^h = 2\}$, respectively. The sequence $\{\xi_n^h\}$ is said to be *locally consistent*, if it satisfies

$$
E_{x,n}^{h,2}[\Delta\xi_n^h] = c\tilde{\Delta}t^h(x) + o(\tilde{\Delta}t^h(x)) = O(h),
$$

$$
\mathrm{Var}_{x,n}^{h,2}(\Delta\xi_n^h) = O(h^2), \tag{3.5}
$$

$$
\sup_{n,\omega\in\Omega} |\Delta\xi_n^h| \to 0 \ \text{ as } \ h \to 0.
$$

Also we require the singular control and impulse control to be "instantaneous." In other words, the interpolation interval on $S_h \times \{0, 1, 2\}$ is

$$
\Delta t^h(x, i) = \tilde{\Delta}t^h(x)I_{\{i=2\}}, \ \text{ for any } (x, i) \in S_h \times \{0, 1, 2\}. \tag{3.6}
$$

Denote by $\pi^h := \{\pi_n^h, n \ge 0\}$ the sequence of control actions and $\mathcal{A}^h$ all piecewise constant admissible controls with

$$
\pi_n^h := \Delta l_n^h I_{\{I_n^h=0\}} + \Delta z_n^h I_{\{I_n^h=1\}}.
$$

The sequence π^h is said to be *admissible* if π_n^h is $\sigma\{\xi_0^h, \dots, \xi_n^h, \pi_0^h, \dots, \pi_{n-1}^h\}$-adapted and for any $E \in \mathcal{B}(S_h)$, we have

$$
P\{\xi_{n+1}^h \in E | \sigma\{\xi_0^h, \dots, \xi_n^h, \pi_0^h, \dots, \pi_n^h\}\} = p^h(\xi_n^h, E | \pi_n^h),
$$

and

$$
P\{\xi_{n+1}^h = B | \xi_n^h = B + h, \sigma\{\xi_0^h, \dots, \xi_n^h, \pi_0^h, \dots, \pi_n^h\}\} = 1.
$$

Put

$$
t_0^h := 0, \ t_n^h := \sum_{k=0}^{n-1} \Delta t^h(\xi_k^h, I_k^h), \ \text{ and } \ n^h(t) := \max\{n : t_n^h \le t\}.
$$

Then the piecewise constant interpolations, denoted by $\xi^h(\cdot)$, $l^h(\cdot)$, and $z^h(\cdot)$, are naturally defined as

$$
\xi^h(t) = \xi_n^h, \ \ l^h(t) = \sum_{k \le n^h(t)} \Delta l_k^h I_{\{I_k^h=0\}}, \ \ z^h(t) = \sum_{k \le n^h(t)} \Delta z_k^h I_{\{I_k^h=1\}}, \tag{3.7}
$$

for $t \in [t_n^h, t_{n+1}^h)$. Let $\xi_0^h = x \in S_h$ and $\pi^h = (z^h, l^h) \in \mathcal{A}^h$ be an admissible control, where $\mathcal{A}^h$ denotes the collection of corresponding admissible controls. The cost function for the controlled Markov chain is defined as

$$
J_B^h(x, \pi^h) = E \sum_{k=1}^{\infty} e^{-rt_k^h}(\beta_1 \Delta z_k^h - \beta_2 \Delta l_k^h - K), \tag{3.8}
$$

which is analogous to (2.3). Thanks to the definition of interpolation intervals in (3.6). The value function of the controlled Markov chain is

$$
V_B^h(x) = \sup_{\pi^h \text{ admissible}} J_B^h(x, \pi^h). \tag{3.9}
$$

We shall show that $V_B^h(x)$ satisfies the dynamic programming equation:

$$V_B^h(x) = \max \left\{ \sum_y e^{-r\Delta t^h(x,i)} p^h((x,y)|\pi) V_B^h(y), \sum_y p^h((x,y)|\pi) V_B^h(y) + \beta_1 h, \right.$$
$$\left. \sup_{0 \le \widetilde{y} \le B-x} p^h((x, x+\widetilde{y})|\pi) V_B^h(x+\widetilde{y}) - \beta_2 \widetilde{y} - K \right\}, \quad \text{for } x \in S_h. \tag{3.10}$$

For simplicity of notation, we use $V^h(x)$ for $V_B^h(x)$ henceforth. Note that discount does not appear in the second and third term above because singular control and impulse control are instantaneous. In the actual computing, we use iterations on value space or policy space to solve V^h. Moreover, the impulse control of capital injection makes the capital injection region have a free boundary, where an optimal control is exerted to determine the capital injection size. The computations turns out to be very involved.

Define the approximation to the derivative of $V(\cdot)$ by finite difference method in the first part of QVI (2.10) using stepsize $h > 0$ as:

$$V(x) \to V^h(x),$$
$$V_x(x) \to \frac{V^h(x+h) - V^h(x)}{h} \quad \text{since} \quad c > 0. \tag{3.11}$$

For the second part of the QVI, we choose

$$V_x(x) \to \frac{V^h(x) - V^h(x-h)}{h}.$$

It leads to

$$\max_{\pi \in \mathcal{A}} \left\{ c \frac{V^h(x+h) - V^h(x)}{h} - rV^h(x), \beta_1 - \frac{V^h(x) - V^h(x-h)}{h}, \right.$$
$$\left. \sup_{0 \le \widetilde{y} \le B-x} V^h(x+\widetilde{y}) - \beta_2 \widetilde{y} - K - V^h(x) \right\} = 0. \tag{3.12}$$

Simplifying (3.12) and comparing the result with (3.10), we achieve the transition probabilities of the first part of the right side of (3.10) as follows:

$$p_D^h((x, x+h)|\pi) = 1, \quad p_D^h(\cdot) = 0, \quad \text{otherwise,}$$
$$\Delta t^h(x) = \frac{h}{c + rh}, \tag{3.13}$$

We also find the transition probability for the second part of the right side of (3.10). That is,

$$p_D^h((x, x-h)|\pi) = 1.$$

Suppose that the current state is $\xi_n^h = x$. The next interpolation interval $\Delta t^h(x)$ is determined by (3.13). To present the claim terms, we determine the next state ξ_{n+1}^h by noting:

1. No claims occur in $[t_n^h, t_{n+1}^h)$ with probability $1 - \lambda \Delta t^h(x) + o(\Delta t^h(x))$; we determine ξ_{n+1}^h by transition probability $p_D^h(\cdot)$ as in (3.13).

2. There is a claim in $[t_n^h, t_{n+1}^h)$ with probability $\lambda \Delta t^h(x) + o(\Delta t^h(x))$, we determine ξ_{n+1}^h by

$$\xi_{n+1}^h = \xi_n^h - y_h,$$

where $y_h \sim \Pi(\cdot)$, and $y_h \in S_h \subseteq \mathbb{R}_+$ such that y_h is the nearest value of y so that $\xi_{n+1}^h \in S_h$. Then $|y_h - y| \to 0$ as $h \to 0$.

Let H_n^h denote the event that ξ_{n+1}^h is determined by the first alternative above, T_n^h denote the event of the second case. Let $I_{H_n^h}$ and $I_{T_n^h}$ be corresponding indicator functions, respectively. Then $I_{H_n^h} + I_{T_n^h} = 1$. Then we need a new definition of the local consistency for Markov chain approximation of compound Poisson process.

Definition 3.1. A controlled Markov chain $\{\xi_n^h, n < \infty\}$ is said to be locally consistent with (2.2), if there is an interpolation interval $\Delta t^h(x) \to 0$ as $h \to 0$ uniformly in x such that

1. there is a transition probability $p_D^h(\cdot)$ that is locally consistent with (3.3) in the sense that (3.5) holds.

2. there is a $\delta^h(x) = o(\Delta t^h(x))$ such that the one-step transition probability $\{p^h(x, y)|\pi\}$ is given by

$$p^h((x, y)|\pi) = (1 - \lambda \Delta t^h(x) + \delta^h(x))p_D^h((x, y)|\pi) \\ + (\lambda \Delta t^h(x) + \delta^h(x))\Pi(x - y)\}. \tag{3.14}$$

Furthermore, the system of dynamic programming equations is a modification of (3.10). That is,

$$V^h(x) = \begin{cases} \max_{\pi \in \mathcal{A}} \Big[(1 - \lambda \Delta t^h(x) + \delta^h(x))e^{-r\Delta t^h(x)} \sum_{(y, \iota)} p_D^h((x, y)|\pi)V^h(y) \\ \quad + (\lambda \Delta t^h(x) + \delta^h(x))e^{-r\Delta t^h(x)} \int_0^x V^h(x - y)\Pi(dy), \\ \quad V^h(x - h) + \beta_1 h, \sup_{0 \le \tilde{y} \le B - x} V^h(x + \tilde{y}) - \beta_2 \tilde{y} - K \Big], \text{ for } x \in S_h, \\ \sup_{0 \le \tilde{y} \le B} V^h(\tilde{y}) - \beta_2 \tilde{y} - K, \qquad\qquad\qquad\qquad \text{for } x = 0, \\ V^h(B - h) + \beta_1 h, \qquad\qquad\qquad\qquad\qquad\qquad \text{for } x = B. \end{cases} \tag{3.15}$$

4   Convergence of Numerical Approximation

This section focuses on the asymptotic properties of the approximating Markov chain proposed in the last section. The main techniques are methods of weak convergence. As mentioned above, we will compute the value function in the finite interval $[0, B]$. Our ultimate goal is to show V^h converges to V in a large enough interval $[0, B]$ as $h \to 0$. A common approach (Kushner and Dupuis (2001)) is to show that the collection $\{\xi^h(\cdot), S^h(\cdot), l^h(\cdot), z^h(\cdot)\}$ is tight and then appropriately characterize the subsequential weak limit. However, the above scheme is problematic since in general, the processes $\{l^h(\cdot), z^h(\cdot)\}$ may fail to be tight. Owing to the time scaling, it is possible that the "limit" has a "jump" resulting in no convergence taking place under the Skorokhod topology. To overcome this difficulty, we will

suitably re-scale the time so that the processes involved in the convergence analysis are tight in the new time scale and carry out weak convergence analysis with the rescaled processes. Finally the rescaled processes will be reverted back to the original time scale to obtain the convergence of V^h to V.

The technique of time rescaling and the interpolation of the approximation sequences is introduced in Section 4.1. Section 4.2 deals with weak convergence of $\{\widehat{\xi}^h(\cdot), \widehat{N}^h(\cdot), \widehat{S}^h(\cdot), \widehat{z}^h(\cdot), \widehat{l}^h(\cdot), \widehat{T}^h(\cdot)\}$, a sequence of rescaled process. As a result, a sequence of controlled surplus processes converges to a limit surplus process. By using the techniques of inversion, convergences of the surplus process and the value function are established.

4.1 Interpolation and Rescaling

Based on the approximating Markov chain constructed above, the piecewise constant interpolation is obtained and the appropriate interpolation interval is chosen. Recalling (3.7), the continuous-time interpolations $\xi^h(\cdot)$, $l^h(\cdot)$, and $z^h(\cdot)$ are defined. Let the discrete times at which claims occur be denoted by ν_j^h, $j = 1, 2, \ldots$ Then we have

$$\xi_{\nu_j^h - 1}^h - \xi_{\nu_j^h}^h = y_j^h.$$

Define $\mathcal{D}_n^h$ as the smallest σ-algebra of $\{\xi_k^h, H_k^h, l_k^h, z_k^h, k \leq n; \nu_k^h, y_k^h : \nu_k^h \leq t_n\}$.

Using the representations of impulse control, singular control, and the interpolations defined above, (3.4) yields

$$\begin{aligned}
\xi_n &= x + \sum_{k=0}^{n-1}[\Delta\xi_k^h I_{H_k^h} + \Delta\xi_k^h(1 - I_{H_k^h})] - \sum_{k=0}^{n-1} z_k^h + \sum_{k=0}^{n-1} l_k^h \\
&= x + \sum_{k=0}^{n-1} E_k^h \Delta\xi_k^h I_{H_k^h} + \sum_{k=0}^{n-1}(\Delta\xi_k^h - E_k^h \Delta\xi_k^h)I_{H_k^h} + \sum_{k=0}^{n-1}\Delta\xi_k^h(1 - I_{H_k^h}) - \sum_{k=0}^{n-1} z_k^h + \sum_{k=0}^{n-1} l_k^h.
\end{aligned}$$
$$(4.1)$$

The local consistency leads to

$$\begin{aligned}
\sum_{k=0}^{n-1} E_k^h \Delta\xi_k^h I_{H_k^h} &= \sum_{k=0}^{n-1}(c\Delta t_k^h + o(\Delta t_k^h))I_{H_k^h} \\
&= \sum_{k=0}^{n-1}(c\Delta t_k^h + o(\Delta t_k^h)) - (\max_{k' \leq n} \Delta t_{k'}^h)O(\sum_{k=0}^{n-1} I_{T_k^h}).
\end{aligned}$$
$$(4.2)$$

and

$$\sum_{k=0}^{n-1}(\Delta\xi_k^h - E_k^h \Delta\xi_k^h)I_{H_k^h} = O(h).$$
$$(4.3)$$

Denote

$$S_n^h = -\sum_{k=0}^{n-1}\Delta\xi_k^h(1 - I_{H_k^h}) = \sum_{k:\nu_k < n} y_k^h.$$
$$(4.4)$$

Note that

$$E \sum_{k=0}^{n-1} I_{T_k^h} = E[\text{number of } n : \nu_n^h \le t] \to \lambda t \quad \text{as} \quad h \to 0.$$

This implies

$$(\max_{k' \le n} \Delta t_{k'}^h) O(\sum_{k=0}^{n-1} I_{T_k^h}) \to 0 \quad \text{in probability as} \quad h \to 0.$$

Hence we can drop the term involving $I_{T_k^h}$ without affecting the limit in (4.2). Combining (4.2) and (2.2), we rewrite (4.1) as

$$\xi^h(t) = x + ct - S^h(t) - z^h(t) + l^h(t) + \varepsilon^h(t), \tag{4.5}$$

where $S^h(t) = \sum_{\nu_n^h \le t} y_n^h$, and $\varepsilon^h(t)$ is a negligible error satisfying

$$\lim_{h \to \infty} \sup_{0 \le t \le T} E|\varepsilon^h(t)| \to 0 \quad \text{for any} \quad 0 < T < \infty. \tag{4.6}$$

Next we introduce the rescaling process. The basic idea of rescaling time is to "stretch out" the control and state processes so that they are "smoother", so the tightness of $l^h(\cdot)$ and $z^h(\cdot)$ can be proved. Define $\Delta \widehat{t}_n^h$ by

$$\Delta \widehat{t}_n^h = \begin{cases} \Delta t^h & \text{for no action on step } n, \\ |\Delta z_n^h| & \text{for a dividend payment on step } n, \\ |\Delta l_n^h| & \text{for a capital injection on step } n. \end{cases} \tag{4.7}$$

Define $\widehat{T}^h(\cdot)$ by $\widehat{T}^h(t) = \sum_{i=0}^{n-1} \Delta t_i^h = t_n^h$ for $t \in [\widehat{t}_n^h, \widehat{t}_{n+1}^h]$. Thus, $\widehat{T}^h(\cdot)$ will increase with the slope of unity if an only if a regular control is exerted. In addition, define the rescaled and interpolated process $\widehat{\xi}^h(t) = \xi^h(\widehat{T}^h(t))$, likewise define $\widehat{S}^h(t)$, $\widehat{l}^h(t)$ and $\widehat{z}^h(t)$ similarly. The time scale is stretched out by h at the impulse and singular control steps. We can now write

$$\widehat{\xi}^h(t) = x + ct - \widehat{S}^h(t) - \widehat{z}^h(t) + \widehat{l}^h(t) + \varepsilon^h(t). \tag{4.8}$$

4.2 Convergence

Theorem 4.1. *Let the approximating chain $\{\xi_n^h, n < \infty\}$ constructed with transition probabilities defined in (3.13) be locally consistent with (2.2), $\xi^h(\cdot)$ be the continuous-time interpolation defined in (3.7), and $\{\widehat{\xi}^h(\cdot), \widehat{N}^h(\cdot), \widehat{S}^h(\cdot), \widehat{z}^h(\cdot), \widehat{l}^h(\cdot), \widehat{T}^h(\cdot)\}$ be the corresponding rescaled processes. Then $\{\widehat{\xi}^h(\cdot), \widehat{N}^h(\cdot), \widehat{S}^h(\cdot), \widehat{z}^h(\cdot), \widehat{l}^h(\cdot), \widehat{T}^h(\cdot)\}$ is tight. Let $\{\widehat{X}(\cdot), \widehat{N}^h(\cdot), \widehat{S}(\cdot), \widehat{z}(\cdot), \widehat{l}(\cdot), \widehat{T}(\cdot)\}$ be the limit of the weakly convergent subsequence of $\{\widehat{\xi}^h(\cdot), \widehat{N}^h(\cdot), \widehat{S}^h(\cdot), \widehat{z}^h(\cdot), \widehat{l}^h(\cdot), \widehat{T}^h(\cdot)\}$. The limit processes satisfy*

$$\widehat{X}(t) = x + c \int_0^t d\widehat{T}(s) - \widehat{S}(t) - \widehat{z}(t) + \widehat{l}(t). \tag{4.9}$$

Proof. In view of (Kushner and Dupuis, 2001, Theorem 9.2.1), the sequence $\{N^h(\cdot)\}$ is tight because the mean number of claims on any bounded interval $[t, t+s]$ is bounded by $\lambda s + \delta_1^h(s)$, where $\delta_1^h(s)$ goes to zero as $h \to 0$, and

$$\liminf_{\delta \to 0 \; h,n} P\{\nu_{n+1}^h - \nu_n^h > \delta | \text{data up to } \nu_n^h\} = 1.$$

This also implies the tightness of $\{\widehat{N}^h(\cdot)\}$ and $\{S^h(\cdot)\}$. On the other hand, due to the definition of $\widehat{S}^h(\cdot)$, we have

$$\left| E[(\widehat{S}^h(t+s) - \widehat{S}^h(t))^2 | \mathcal{F}_t^h] \right| \leq \left| E[(S^h(t+s) - S^h(t))^2 | \mathcal{F}_t^h] \right| \leq \gamma(s),$$

where

$$\lim_{s \to 0} \limsup_{h \to 0} E\gamma(s) = 0.$$

Then we obtain $\widehat{S}^h(\cdot)$ is tight. Furthermore, following the definition of "stretched out" timescale,

$$|\widehat{z}^h(\tau_h + \delta) - \widehat{z}^h(\tau_h)| \leq |\delta| + O(h),$$
$$|\widehat{l}^h(\tau_h + \delta) - \widehat{l}^h(\tau_h)| \leq |\delta| + O(h).$$

Thus $\{\widehat{z}^h(\cdot), \widehat{l}^h(\cdot)\}$ is tight. These results and the boundedness of c implies the tightness of $\{\xi^h(\cdot)\}$. Therefore it follows that $\{\widehat{\xi}^h(\cdot), \widehat{N}^h(\cdot), \widehat{S}^h(\cdot), \widehat{z}^h(\cdot), \widehat{l}^h(\cdot), \widehat{T}^h(\cdot)\}$ is tight.

Since $\{\widehat{\xi}^h(\cdot), \widehat{N}^h(\cdot), \widehat{S}^h(\cdot), \widehat{z}^h(\cdot), \widehat{l}^h(\cdot), \widehat{T}^h(\cdot)\}$ is tight, we can extract a weakly convergent subsequence denoted by $\{\widehat{X}(\cdot), \widehat{N}^h(\cdot), \widehat{S}(\cdot), \widehat{z}(\cdot), \widehat{l}(\cdot), \widehat{T}(\cdot)\}$. Also, the paths of $\{\widehat{X}(\cdot), \widehat{N}^h(\cdot), \widehat{S}(\cdot), \widehat{z}(\cdot), \widehat{l}(\cdot), \widehat{T}(\cdot)\}$ are continuous w.p.1.

Let $h \to 0$, by using the Skorohod representation, we obtain

$$E \left| \int_0^t c\, d\widehat{T}^h(s) - \int_0^t c\, d\widehat{T}(s) \right| \to 0 \tag{4.10}$$

uniformly in t. Then we have

$$\widehat{x}(t) = x + c \int_0^t d\widehat{T}(s) - \widehat{S}(t) - \widehat{z}(t) + \widehat{l}(t) + \varepsilon^\delta(t), \tag{4.11}$$

where $\lim_{\delta \to 0} E|\varepsilon^\delta(t)| = 0$. Finally, taking limits in the above equation as $\delta \to 0$, (4.9) is obtained. $\square$

Theorem 4.2. *For $t < \infty$, define the inverse $T(t) = \inf\{s : \widehat{T}(s) > t\}$. Then $T(t)$ is right continuous and $T(t) \to \infty$ as $t \to \infty$ w.p.1. For any process $\widehat{\psi}(\cdot)$, define the rescaled process $\psi(\cdot)$ by $\psi(t) = \widehat{\psi}(T(t))$. Then, (2.2) holds.*

Proof. Since $\widehat{T}(t) \to \infty$ w.p.1 as $t \to \infty$, $T(t)$ exists for all t and $T(t) \to \infty$ as $t \to \infty$ w.p.1. A rescaling of (4.9) yields

$$X(t) = x + ct - S(t) - z(t) + l(t). \tag{4.12}$$

In other words, (2.2) holds. $\square$

Theorem 4.3. *Let h index the weak convergent subsequence of $\{\widehat{\xi}^h(\cdot), \widehat{N}^h(\cdot), \widehat{S}^h(\cdot), \widehat{z}^h(\cdot),$*
$\widehat{l}^h(\cdot), \widehat{T}^h(\cdot)\}$ with the limit $\{\widehat{X}(\cdot), \widehat{N}^h(\cdot), \widehat{S}(\cdot), \widehat{z}(\cdot), \widehat{l}(\cdot), \widehat{T}(\cdot)\}$. Then,

$$J^h(x, \pi^h) \to E_x^\pi \int_0^\infty e^{-rt} \beta_1 dZ - \sum_{n=1}^\infty e^{-r\tau_n}(K + \beta_2 \zeta_n) I_{\{\tau_n < \infty\}} = J(x, \pi). \tag{4.13}$$

Proof. Note that $\Delta z^h = h$, the uniform integrability of dZ can be easily verified. Due to the tightness and the uniform integrability properties, for any t, $\int_0^t d\widehat{Z}$ can be well approximated by a Reimann sum uniformly in h. By the weak convergence and the Skorohod representation,

$$E \sum_{k=1}^\infty e^{-rt_k^h} \beta_1 \Delta z_k^h \to E_x^\pi \int_0^\infty e^{-r\widehat{T}(t)} \beta_1 d\widehat{Z}.$$

By an inverse transformation,

$$E_x^\pi \int_0^\infty e^{-r\widehat{T}(t)} \beta_1 d\widehat{Z} = E_x^\pi \int_0^\infty e^{-rt} \beta_1 dZ.$$

Also,

$$E \sum_{k=1}^\infty e^{-rt_k^h}(\beta_2 \Delta l_k^h + K) \to E \sum_{n=1}^\infty e^{-r\tau_n}(\beta_2 \zeta_n + K) I_{\{\tau_n < \infty\}}.$$

Thus, as $h \to 0$,

$$J^h(x, \pi^h) \to J(x, \pi).$$

$\square$

Theorem 4.4. *$V^h(x)$ and $V(x)$ are value functions defined in (3.9) and (2.4), respectively. Then $V^h(x) \to V(x)$ as $h \to 0$.*

Proof. First, to prove

$$\limsup_h V^h(x) \le V(x). \tag{4.14}$$

Since $V(x)$ is the maximizing cost function, for any admissible control $\pi(\cdot)$,

$$J(x, \pi) \le V(x).$$

Let $\widetilde{\pi}^h(\cdot) = (\widetilde{z}^h(\cdot), \widetilde{l}^h(\cdot))$. That is,

$$V^h(x) = J^h(x, \widetilde{\pi}^h) = \sup_{\pi^h} J^h(x, \pi^h).$$

Choose a subsequence $\{\widetilde{h}\}$ of $\{h\}$ such that

$$\lim_{\widetilde{h} \to 0} V^{\widetilde{h}}(x) = \limsup_{\widetilde{h} \to 0} V^{\widetilde{h}}(x) = \lim_{\widetilde{h} \to 0} J^{\widetilde{h}}(x, \widetilde{\pi}^{\widetilde{h}}).$$

Without loss of generality (passing to an additional subsequence if needed), we may assume that $\{\xi^{\widetilde{h}}(\cdot), N^{\widetilde{h}}(\cdot), S^{\widetilde{h}}(\cdot), z^{\widetilde{h}}(\cdot), l^{\widetilde{h}}(\cdot)\}$ converges weakly to $\{X(\cdot), N(\cdot), S(\cdot), z(\cdot), l(\cdot)\}$, where

$\pi(\cdot) = (z(\cdot), l(\cdot))$ is an admissible control. Then the weak convergence and the Skorohod representation yield that

$$\limsup_h V^h(x) = J(x, \pi) \leq V(x). \tag{4.15}$$

We proceed to prove the reverse inequality.

We claim that

$$\liminf_h V^h(x) \geq V(x). \tag{4.16}$$

Suppose that $\overline{\pi}$ is an optimal control such that $\overline{X}(\cdot)$ is the associated trajectory, and $J(x, \overline{\pi}) = V(x)$. Given any $\varepsilon > 0$, there are an $\varepsilon > 0$ and an ordinary control $\overline{\pi}^\varepsilon(\cdot)$ that takes only finite many values, that $\overline{\pi}^\varepsilon(\cdot)$ is a constant on $[k\varepsilon, k\varepsilon + \varepsilon)$, that $\overline{\pi}^\varepsilon(\cdot)$ is its corresponding optimal control representation, and let $\overline{X}^\varepsilon(\cdot)$ be the associated solution. Then if $\{\overline{\pi}^\varepsilon(\cdot), N(\cdot), S(\cdot)\}$ converges weakly to $\{\overline{\pi}(\cdot), N(\cdot), S(\cdot)\}$, we also have $\{\overline{X}^\varepsilon(\cdot), \overline{\pi}^\varepsilon(\cdot), N(\cdot), S(\cdot)\}$ converges weakly to $\{X(\cdot), \overline{\pi}(\cdot), N(\cdot), S(\cdot)\}$. Then $J^h(x, \overline{\pi}^\varepsilon) \to J(x, \overline{\pi})$, and that $J^h(x, \overline{\pi}^\varepsilon) \geq V(x) - h$. Thus,

$$\liminf_h V^h(x) \geq J^h(x, \overline{\pi}^h) \geq V(x) - h.$$

The arbitrariness of h then implies that $\liminf_h V^h(x) \geq V(x)$. Using (4.15) and (4.16) together with the weak convergence and the Skorohod representation, we obtain the desired result. The proof of the theorem is concluded. □

5 Numerical Examples

This section is devoted to a couple of examples. We approximate the value functions and optimal controls in two different claim severity distributions. The claim sizes in the first example follows exponential distribution in which the tail of distribution is considered light. Exponential distribution is applicable for automobile losses. In the seconde example, we consider the Pareto distribution, which provides the best fit to heavy-tailed business such as liability insurance. The corresponding capital injection sizes and barriers for regions are also obtained in the numerical examples.

Based on the algorithm constructed above, we carry out the computation by value iterations. For $n \in Z^+$, define the vectors

$$V_n^h = \{V_n^h(h), V_n^h(2h), \ldots, V_n^h(B)\}$$
$$V^h = \{V^h(h), V^h(2h), \ldots, V^h(B)\}.$$

Using the method of value iteration, we obtain $V_n^h \to V^h$ as $n \to \infty$.

1. Set $n = 0$. $\forall x \in S_h$, we set the initial value $V^h(x) = 1$.

2. Find improved values $V_{n+1}^h(x)$ by (3.15) and record the corresponding optimal control.

$$V_{n+1}^h(x) = \max_{\pi \in \mathcal{A}} \Big[(1 - \lambda \Delta t^h(x) + \delta^h(x)) e^{-r\Delta t^h(x)} \sum_{(y,\iota)} p_D^h((x,y)|\pi) V_n^h(y)$$

$$+ (\lambda \Delta t^h(x) + \delta^h(x)) e^{-r\Delta t^h(x)} \int_0^x V_n^h(x-y) \Pi(dy),$$

$$V_n^h(x-h) + \beta_1 h, \sup_{0 \leq \widetilde{y} \leq B-x} V_n^h(x+\widetilde{y}) - \beta_2 \widetilde{y} - K \Big]$$

$$\pi_{n+1}^h = \underset{\pi \in \mathcal{A}}{\operatorname{argmax}}\Big[(1 - \lambda \Delta t^h(x) + \delta^h(x))e^{-r\Delta t^h(x)} \sum_{(y,\iota)} p_D^h((x,y)|\pi)V_n^h(y)$$

$$+(\lambda \Delta t^h(x) + \delta^h(x))e^{-r\Delta t^h(x)} \int_0^x V_n^h(x-y)\Pi(dy),$$

$$V_n^h(x-h) + \beta_1 h, \sup_{0 \leq \widetilde{y} \leq B-x} V_n^h(x+\widetilde{y}) - \beta_2\widetilde{y} - K\Big].$$

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

5.1 Exponential Distribution

Example 5.1. We assume $S(t)$ is the compound process. The claim severity distribution follows exponential distribution with density function $f(y) = \alpha e^{-\alpha y}$ where $\alpha = 0.1$. Furthermore, $\{\nu_{n+1} - \nu_n\}$ is a sequence of exponentially distributed random variables with mean 10. Then $\lambda = 0.1$. Let the upper bound of the computation interval $B = 100$, the premium collection rate $c = 10$, the interest rate $r = 0.05$, the fixed capital injection cost $K = 0.1$, the parameters for the proportion costs of dividend payments and capital injections $\beta_1 = 0.9$, $\beta_2 = 1.1$. Set $h = 0.2506$ and tolerance $= 10^{-7}$. We obtain the computation results depicted in Figure 5.1 as follows.
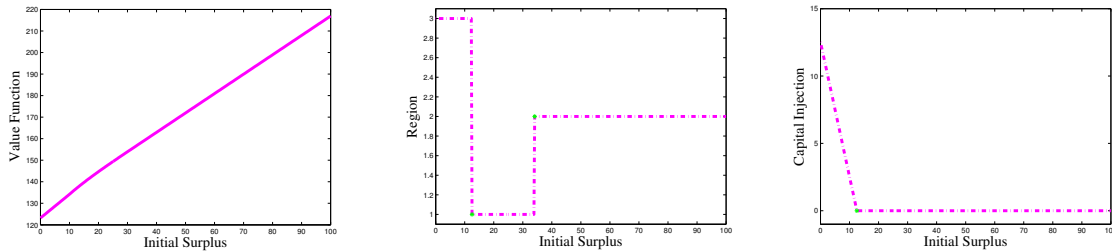
5.1.1 Performance function versus initial surplus

5.1.2 Optimal dividend payment and capital injection strategies versus initial surplus

5.1.3 Optimal capital injection sizes versus initial surplus

Figure 5.1: Exponential claim size distribution with $B = 100$, $\alpha = 0.1$, $\lambda = 0.1$, $K = 0.1$, $c = 10$, $\beta_1 = 0.9$, $\beta_2 = 1.1$.

Figure 5.1.1 shows that the value function is concave and monotone increasing. The initial value $V(0)$ is nonzero because the capital injections can always guarantee the continuity of the business even with the zero initial surplus. Furthermore, starting as a curve, the value function increases linearly after some barrier level, which means the extra surplus will all be paid out as the dividend after reaching certain barrier. With the cost of dividend payment, the slope should be β_1 instead of unity.

Figure 5.1.2 describes the regions of the QVI. We use "1" to denote the continuation region when no action is taken, "2" to denote the dividend payout region and "3" to denote the capital injection region. It shows that the capital injection region is dominant when

surplus is lower than a bound. The dividend payout strategy is a barrier strategy, where the dividend payout region is dominant beyond some barrier level. In addition, we see that the region switches to continuation region when the surplus reaches $x_1 = 10.2551$. That is, if the surplus reaches x_1, the company can self finance its business without the exogenous capital injections. The continuation switches to the dividend payout region at the barrier level $x_2 = 33.8346$. Then the surplus more than x_2 will be paid out immediately.

Figure 5.1.3 provides the relationship between the optimal capital injection size and the initial surplus. It shows that the capital injection size decrease with the increase of the surplus. That is, healthier financial condition needs less capital injections, which is in accordance with the intuitive thinking. The optimal capital injection size reaches 0 at $x_3 = 12.5313$.

A natural guess is that if $x_1 = x_3$. That is, when it is optimal to take no action at x_1, should the optimal capital injection size be 0 at the same time (at x_3)? However, due to the fixed capital injection cost K, the difference between the two switching points x_1 and x_3 exists. Between $[x_1, x_3]$, the maximal term in (2.7) is the first term, which means no action will be taken in this continuation region. At the meantime, the third term will reach its maximum by choosing a nonzero $\widetilde{y}$, even though it is not the maximal value of the QVI. We can show that the fixed capital injection cost K significantly influence the gap between this two switching points, and the difference vanishes when there is no fixed capital injection cost. Referring to Table 5.1, we can see that the difference $x_3 - x_1 \to 0$ as $K \to 0$.

K	4	2	0.5	0.1	0.05	0.001	0
x_3	12.5313	12.5313	12.5313	12.5313	12.5313	12.5313	12.5313
x_1	0	2.5063	7.0175	10.0251	10.7769	12.2807	12.5313
$x_3 - x_1$	12.5313	10.0250	5.5138	2.5062	1.7544	0.2506	0

Table 5.1: Difference between x_1 and x_3 respect to K

From Tables 5.2 and 5.3, we can analyze the effects of changes in β_1, β_2. From the observation that x_1 and x_3 are decreasing with smaller β_1, it is shown that smaller β_1 will result in smaller sizes of capital injections and postponed capital injection times; and slightly increase barrier level x_2. Similarly, larger β_2 will result in smaller sizes of capital injections and postponed capital injection times; and slightly increase barrier level x_2. The differences of x_2 with respect to different values of β_1 and β_2 are no more than 0.2506, which is just one step size h and can be caused by computation errors. We can conclude that the barrier levels x_2 are independent of β_1 and β_2. Referring to x_1 and x_3, the natural interpretation is that smaller β_1 and larger β_2 mean more cost of transaction. To maximize the performance, the capital injection time will be postponed and the size of the capital injection will be reduced. This result is consistent with result in Yao et al. (2011).

From Tables 5.4 and 5.5, we study the effects of λ and α. It is shown that the barrier level x_2 increases with larger λ. It can be interpreted that the dividend payment should be postponed to avoid surplus short if more frequent claims come. Similarly, smaller α means larger expectation of claim size. The barrier level x_2 decreases as α increases with the same reason. In addition, since larger λ and smaller α will increase the cost for the insurance

β_1	0.99	0.9	0.8	0.7	0.6	0.5	0.4
x_1	14.0351	10.0251	6.2657	3.0075	0	0	0
x_2	33.5840	33.8346	33.8346	34.0852	34.0852	34.0852	34.0852
x_3	17.0426	12.5313	8.5213	4.7619	1.5038	0.2506	0.2506

Table 5.2: Switching points with respect to β_1

β_2	1.01	1.1	1.2	1.3	1.4	1.7	1.9
x_1	13.5338	10.0251	7.2682	5.2632	3.5088	0	0
x_2	33.8346	33.8346	34.0852	34.0852	34.0852	34.0852	34.0852
x_3	16.5414	12.5313	9.2732	7.0175	5.0125	0.7519	0.2506

Table 5.3: Switching points with respect to β_2

company to pay for the claims, the sizes of possible capital injection also increase.

λ	0.01	0.05	0.1	0.2	0.5
x_1	0	4.0100	10.0251	15.2882	17.0426
x_2	12.2807	27.3183	33.8346	40.3509	46.1153
x_3	0.2506	6.5163	12.5313	18.0451	20.0501

Table 5.4: Switching points with respect to λ

α	0.08	0.1	0.2	0.5	0.7
x_1	10.2757	10.0251	6.7669	3.0075	2.2556
x_2	37.8446	33.8346	21.8045	11.2782	8.7719
x_3	13.0326	12.5313	8.5213	4.2607	3.2582

Table 5.5: Switching points with respect to α

Figures 5.1 and 5.2 indicate the effect of β_2 on capital injection controls. Let $\beta_2 = 1.1$ in Figure 5.1 and $\beta_2 = 1.9$ in Figure 5.2. From Figures 5.1.2 and 5.2.2, it is shown that the capital injection region is much bigger when β_2 is smaller. Then the capital injection strategy can be seen as a singular control in the interval $[0, x_1]$. Whilst for the bigger β_2, the capital injection region is shrank to one point "0", and the capital injection strategy becomes the impulse control on the left boundary. In addition, comparing Figures 5.1.3 and 5.2.3, a relatively smaller capital injection size was suggested with larger β_2. It can be interpreted that larger β_2 means larger capital injection costs, and the decision maker will postpone the capital injection time and reduce the capital injection size as much as possible.
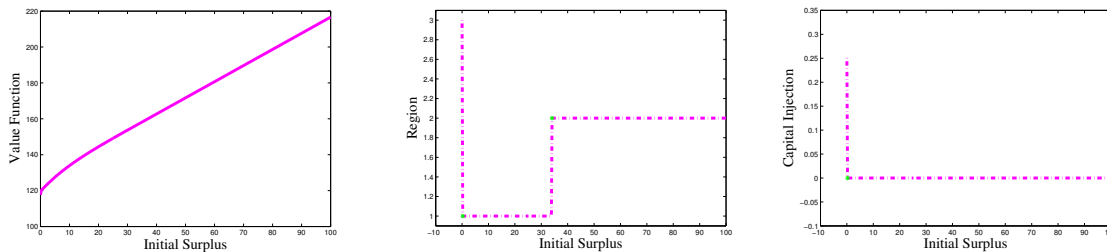
5.2.1 Performance function versus initial surplus

5.2.2 Optimal dividend payment and capital injection strategies versus initial surplus

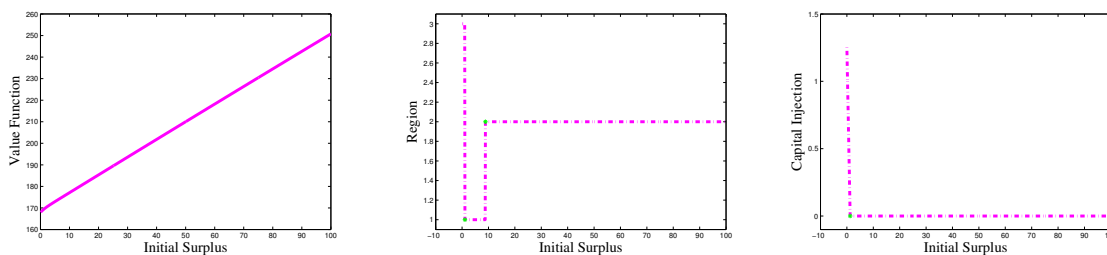5.2.3 Optimal capital injection sizes versus initial surplus

Figure 5.2: Exponential claim size distribution with $B = 100$, $\alpha = 0.1$, $\lambda = 0.1$, $K = 0.1$, $c = 10$, $\beta_1 = 0.9$, $\beta_2 = 1.9$.

5.2   Pareto Distribution

Example 5.2. Comparing to Example 5.1, we only change the assumption of the claim size distribution and assume it follows the two parameter Pareto distribution with density

$$f(x) = \frac{\rho \theta^\rho}{(x + \theta)^{\rho+1}},$$

where $\rho = 2$, $\theta = 1$. Set $h = 0.08304$. Similar to Figures 5.1, we obtain the computation results depicted in Figure 5.3 as follows. The value function is given. Three regions for different control strategies and the optimal capital injections sizes are obtained.

5.3.1 Performance function versus initial surplus

5.3.2 Optimal dividend payment and capital injection strategies versus initial surplus

5.3.3 Optimal capital injection sizes versus initial surplus

Figure 5.3: Pareto claim size distribution with $B = 100$, $\alpha = 0.1$, $\lambda = 0.1$, $K = 0.1$, $c = 10$, $\beta_1 = 0.9$, $\beta_2 = 1.1$, $\rho = 2$, $\theta = 1$.

From Tables 5.6 and 5.7, we can see that β_1 and β_2 have little or no effect on the barrier level of dividend payment x_2. In addition, it is shown that smaller β_1 will result in smaller sizes of capital injections and postponed capital injection times and larger β_2 will result in smaller sizes of capital injections and postponed capital injection times. These phenomena are consistent with the results in Example 5.1.

β_1	0.99	0.9	0.8	0.7	0.6	0.5	0.3
x_1	0.9174	0.5838	0.3336	0.1668	0.0834	0.0834	0.0834
x_2	8.9241	9.0075	8.8407	9.0075	8.9241	9.0909	9.3411
x_3	1.7515	1.2510	0.9174	0.5838	0.4170	0.1668	0.0834

Table 5.6: Switching points with respect to β_1

β_2	1.01	1.1	1.2	1.3	1.4	1.7	1.9
x_1	0.8340	0.5838	0.4170	0.3336	0.2502	0.0834	0.0834
x_2	8.8407	9.0075	9.0075	8.8407	8.9241	8.8407	8.9241
x_3	1.6681	1.2510	1.0008	0.7506	0.6672	0.3336	0.2502

Table 5.7: Switching points with respect to β_2

λ	0.01	0.05	0.1	0.2	0.5
x_1	0.0834	0.0834	0.5838	1.4179	3.3361
x_2	3.2527	6.5888	9.0075	13.1777	37.7815
x_3	0.0834	0.6672	1.2510	2.2519	4.3369

Table 5.8: Switching points with respect to λ

K	0.5	0.1	0.02	0.001	0
x_3	1.2510	1.2510	1.2510	1.2510	1.2510
x_1	0.0834	0.5838	0.9174	1.0842	1.0842
$x_3 - x_1$	1.1676	0.6672	0.3336	0.1668	0.1668

Table 5.9: Difference between x_1 and x_3 respect to K

From Tables 5.8 and 5.9, we study the effects of λ and K. It is shown that the barrier level x_2 increases as λ increases and larger λ will lead to bigger x_1 and x_3. Moreover, the capital injection K has similar effect on the two switching points x_1 and x_3. The interpretations to above results are also similar to those in Example 5.1.

ρ	1	2	3	5	10	20
x_1	1.3344	0.5838	0.3336	0.1668	0.0834	0
x_2	17.0142	9.0075	5.6714	3.0859	2.6689	0.0834
x_3	2.5021	1.2510	0.8340	0.4170	0.1668	0.0834

Table 5.10: Switching points with respect to ρ

θ	0.5	1	2	3	5	10
x_1	0.2502	0.5838	1.2510	1.8349	2.8357	4.1701
x_2	6.8390	9.0075	12.5938	15.1793	19.6831	26.5221
x_3	0.7506	1.2510	2.2519	3.0859	4.3369	6.1718

Table 5.11: Switching points with respect to θ

From Tables 5.10 and 5.11, we analyze the effects of ρ and θ. The expectation of the claim sizes is $\frac{\theta}{\rho-1}$. It is shown in Table 5.10 that x_1 decreases and x_2 decreases as ρ increases. Smaller ρ means larger expectation of claim size. The barrier level x_2 decrease as ρ increase because the dividend payment should be postponed with more frequent claims to avoid surplus short. Also smaller ρ will increase the cost for the insurance company to pay for the claims, the sizes of possible capital injection x_1 also increase. Similarly in Table 5.11, x_1 increases and x_2 increases as θ increases. Larger θ has the same effect on expectation of the claim sizes as smaller ρ, thus the results can be interpreted with the same reasons.

6   Further Remarks

In this work, we have developed numerical approximation schemes to find the optimal capital injection and dividend payment policy to maximize the total discounted dividend payment excluding the possible capital injections. Although one could derive the QVI by using the usual dynamic programming approach, with the free boundary of the capital injection region, solving the impulse and singular control problem analytically is very difficult. Taking into account the jump process, the claim size varies with different claim severity distributions. Thus, to obtain a closed-form solution is virtually impossible. As an alternative, we presented a Markov chain approximation method using mainly probabilistic methods. A technique of time rescaling is used. In the actual computation, the optimal value function can be obtained by using the value or policy iteration methods. Two examples of exponential and Pareto distributions are provided and the corresponding properties of each parameter are analyzed. This method can also be extended to more complicated claim severity models and performance functions.

Acknowledgments

References

Asmussen, S. and Taksar, M. Controlled diffusion models for optimal dividend pay-out. *Insurance: Math. and Economics*, 20: 1–15 (1997)

Bayraktar, E., Song, Q.S., and Yang, J. On the continuity of stochastic exit time control problems. *Stoch. Anal. Appl.*, 29(1): 48–60 (2011)

Budhiraja, A. and Ross, K. Convergent numerical scheme for singular stochastic control with state constraints in a portfolio selection problem. *SIAM J. Control Optim.*, 45(6): 2169–2206 (2007)

Dickson, DCM and Waters, HR Some optimal dividends problems. *ASTIN Bulletin* 34(1):49–74 (2004).

De Finetti, B. Su unimpostazione alternativa della teoria collettiva del rischio. *Transactions of the XVth International Congress of Actuaries* 2: 433–443 (1957).

Fleming, W. and Soner, H. *Controlled Markov Processes and Viscosity Solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, New York, NY, second edition, 2006

Gerber, H. and Shiu, E. Optimal dividends: analysis with Brownian motion. *North American Actuarial Journal*, 8: 1–20 (2004).

He, L. and Liang, Z.X. Optimal financing and dividend control of the insurance company with fixed and proportional transaction costs. *Insurance: Mathematics and Economics*, 42: 88–94 (2009).

Kulenko, N. and Schimidli, H. An optimal dividend strategy in a Cráme-Lundberg model with capital injections. Insurance: Mathmatics and Economics, 43:270–278 (2008).

Kushner, H. and Dupuis, P. *Numerical Methods for Stochstic Control Problems in Continuous Time*, volume 24 of *Stochastic Modelling and Applied Probability*. Springer, New York, second edition, 2001

Kushner, H. J. and Martins, L. F. Numerical methods for stochastic singular control problems. *SIAM J. Control Optim.*, 29:1443–1475 (1991).

Jin, Z., Yin, G., and Zhu, C. Numerical solutions of optimal risk control and dividend optimization policies under a generalized singular control formulation. *Automatica*, 48(8): 1489–1501 (2012)

Reddemann, S., Basse, T., von der Schulenburg, and Johann-Matthias G. On the impact of the financial crisis on the dividend policy of the European insurance industry. *Geneva Papers on Risk and Insurance: Issues and Practice*, 35(1):53–62 (2010)

Schmidli, H. *Stochastic Control in Insurance*, Springer Verlag, 2008

Scheer, Natalie and Schmidli, Hanspeter. Optimal dividend strategies in a Cráme-Lundberg model with capital injections and administration costs. *European Actuarial Journal*, 1(1):57–92 (2011)

Sethi, S.P. and Taksar, M. Optimal financing of a corporation subject to random returns. *Mathematical Finance*, 12:155–172 (2002)

Song, Q., Stockbridge, R. and Zhu, C (2011). On optimal harvesting problems in random environments. *SIAM J. Control Optim.*, 49(2):859–89.

Thonhauser, S. and Albrecher, H Optimal dividend strategies for a compound Poisson process under transaction costs and power utility. *Stochastic Models*, 27(1):120–140 (2011)

Yao, D., Yang, H. and Wang, R. Optimal financing and dividend strategies in a dual model with proportional costs. *Journal of Industrial and Management Optimization*, 6(4): 761–777 (2010)

Yao, D., Yang, H. and Wang, R. Optimal dividend and capital injection problem in the dual model with proportional and fixed transaction costs. *European Journal of Operational Research*, 211:568–576 (2011)