Optimal ratcheting of dividends in insurance

Hansjörg Albrecher*, Pablo Azcue† and Nora Muler†

Abstract

We address a long-standing open problem in risk theory, namely finding the optimal strategy to pay out dividends from an insurance surplus process, if the dividends are paid according to a dividend rate that is not allowed to decrease. The optimality criterion here is to maximize the expected value of the aggregate discounted dividend payments up to the time of ruin. In the framework of the classical Cramér-Lundberg risk model, we solve the corresponding two-dimensional optimal control problem and show that the value function is the unique viscosity solution of the corresponding Hamilton-Jacobi-Bellman equation. We also show that the value function can be approximated arbitrarily closely by ratcheting strategies with only a finite number of possible dividend rates and identify the free boundary and the optimal strategies in several concrete examples. These implementations illustrate that the restriction of ratcheting does not lead to a large efficiency loss when compared to the classical un-constrained optimal dividend strategy.

1 Introduction

How to optimally pay out dividends from an insurance surplus process is a classical research question starting with the papers of de Finetti [11] and Gerber [13]. When the criterion is to maximize the expected aggregate discounted dividend payments up to the time of ruin, the challenge is to find the right compromise between paying early in view of the discounting and paying late in order not to have ruin too early and profit from the typically positive safety loading for a longer time. The problem turns out to be very challenging from a mathematical point of view, and many variants have been studied over the last decades, using various different techniques, see e.g. Schmidli [18] and Albrecher & Thonhauser [3] for an overview. In recent years, the problem became well understood within the framework of modern stochastic control theory and the concept of viscosity solutions for corresponding Hamilton-Jacobi-Bellman equations, cf. Azcue & Muler [6].

In terms of the practical insight from the resulting optimal payout strategies, one aspect often raised critically in discussions by practitioners was the following: dividend strategies implemented in practice often are designed in a way as to not decrease over time, since a decrease would send unfavorable signals to the market. Such a monotonicity of dividend rates over time (also referred to as *ratcheting*) is, however, not automatically present in the optimal strategies without this ratcheting constraint, as the optimal strategies are of band type (and often of simpler threshold form: pay no dividends below a certain threshold, and at maximal rate above the threshold). Hence it is an interesting question to (a) look for the optimal strategies when such a ratcheting constraint is imposed and (b) see whether this additional constraint comes at the cost of losing a lot of efficiency

*Department of Actuarial Science, Faculty of Business and Economics, University of Lausanne, CH-1015 Lausanne and Swiss Finance Institute. Supported by the Swiss National Science Foundation Project 200021_168993.

†Departamento de Matematicas, Universidad Torcuato Di Tella. Av. Figueroa Alcorta 7350 (C1428BIJ) Ciudad de Buenos Aires, Argentina.

when compared to the un-constrained value function.[1] A first step towards answering these questions was recently obtained in [2], where explicit calculations were performed for a restricted form of a ratcheting strategy, namely that once during the lifetime of the process the dividend rate can be increased. It was then studied, both in the Cramér-Lundberg model and its diffusion approximation, to what extent and at which surplus level such an increase should optimally be implemented, leading to some surprising relations of the optimal ratcheting level with the threshold level of unconstrained dividend strategies. However, finding the optimal solution to the general ratcheting problem for a continuum of available ratcheting levels of the dividend rate was still open. From a technical point of view, it becomes clear that in a stochastic control formulation one is faced with a (Markovian) two-dimensional problem, keeping track of both the current surplus level and the currently implemented dividend rate. The analysis of two-dimensional control problems in risk theory can be quite intricate, see e.g. Albrecher et al. [1], Azcue et al. [8] and Gu et al. [15] in the context of other dividend problem formulations.

In this paper, we solve the two-dimensional ratcheting problem and establish the value function as the unique viscosity solution of the respective Hamilton-Jacobi-Bellman equation. It will turn out that allowing the maximal dividend rate to exceed the rate of incoming premiums leads to some additional analytical challenges, but one can derive the respective results in that case as well. Note that the concept of ratcheting has been studied in the context of lifetime consumption in the corporate finance community, see Dybvig [12] and the very recent nice extension of Angoshtari et al. [4]. In those papers the focus is on a geometric Brownian motion as an underlying and a logarithmic or power utility function applied to the consumption rate, which together with interest rate considerations renders this model setup within the framework of Merton-type consumption problems. Despite some apparent analogies, the present risk theory setup does not fall within the class of models studied there and the techniques used for its study are quite different.

The rest of the paper is structured as follows. Section 2 describes the model setup in more detail and some basic results are derived in Section 3. Some of the respective proofs are, however, quite technical and hence delegated to an appendix.[2] In Section 4 it is then proved that the value function of the general ratcheting problem is the unique viscosity solution of the respective Hamilton-Jacobi-Bellman equation. Section 5 studies properties of ratcheting strategies when only finitely many different dividend rates are possible, and in Section 6 it is shown that these strategies converge uniformly to the general value function, when the number of possible dividend rates tends to infinity. Section 7 identifies the resulting optimal ratcheting strategies, the free boundaries and the corresponding value functions for a number of concrete examples with exponentially and Gamma distributed claims. We also compare the optimal solutions to their counterparts in the un-constrained case (without ratcheting) and in the case when only one switch of dividend rate is allowed, as studied in [2]. It turns out that the efficiency loss due to ratcheting is remarkably small, and that a one-switch strategy already performs very similarly to the optimal general ratcheting solution. Finally, Section 8 concludes.

2 Model

Consider the free surplus X_t of an insurance portfolio according to the Cramér-Lundberg model

$$X_t = x + pt - \sum_{i=1}^{N_t} U_i, \tag{1}$$

where x is the initial surplus, p is the premium rate and U_i is the size of the i-th claim. All claims are assumed to be i.i.d. random variables with continuous distribution function F. N_t is the number of claims up to time t and assumed to follow a Poisson process with intensity β. Let us denote by

[1] The question was for instance posed in an academic environment by Elias Shiu at the First Int. Workshop on Gerber-Shiu functions in Montreal back in 2006, see also Avanzi et al. [9] for a more recent motivation.

[2] Due to the page limit of the journal, this 9-page appendix is made available in a separate file which can be downloaded from the webpage of the corresponding author.

τ_i the arrival time of claim i. The process N_t and the random variables U_i are independent of each other, and we have the safety loading condition $p > \beta \mathbb{E}(U_i)$. Let Ω be the set of paths with left and right limits and let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$ be the complete probability space generated by the process X_t.

The company uses part of the surplus to pay dividends to the shareholders with a finite rate less than or equal to a fixed rate $\overline{c} > 0$. Let us denote by C_t the rate at which the company pays dividends at time t. Given an initial surplus $X_0 = x$ and a minimum dividend rate c at the beginning, a dividend ratcheting strategy $C = (C_t)_{t\geq 0}$ is admissible if it is càdlàg, adapted with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$, non-decreasing and if it satisfies $c \leq C_t \leq \overline{c}$ for all t. Moreover, the controlled surplus process can be written as

$$X_t^C = X_t - \int_0^t C_s ds.$$

Let us define $\Pi_{x,c,\overline{c}}$ as the set of all the admissible dividend ratcheting strategies. Given $x \geq 0$, $c \in [0, \overline{c}]$ and an admissible dividend ratcheting strategy $C \in \Pi_{x,c,\overline{c}}$, the value function of this strategy is given by

$$J(x; C) = \mathbb{E}\left[\int_0^\tau e^{-qs} C_s ds\right],$$

where $\tau = \inf\left\{t \geq 0 : X_t^C < 0\right\}$ is the ruin time. Hence, for any initial surplus $x \geq 0$ and initial dividend rate $c \in [0, \overline{c}]$, our aim is to maximize

$$V(x, c) = \sup_{C \in \Pi_{x,c,\overline{c}}} J(x; C). \tag{2}$$

Remark 2.1 In the case that $c \leq p$, any ratcheting strategy C in $\Pi_{0,c,\overline{c}}$ with $C_0 > p$ can not be optimal because the corresponding ruin time is 0. Also, in the case that $C_{t-} \leq p$ and $X_t^C = 0$ for some $t > 0$, any dividend ratcheting strategy with $C_s > p$ for $s > t$ can not be optimal because it implies immediate ruin as well. So, without loss of generality, we only consider admissible strategies that satisfy the following property: if $C_{t-} = p$ with $X_t^C = 0$ for $t > 0$, then $C_s = p$ for $s \geq t$ until ruin time. Also, the only admissible strategy in $\Pi_{0,p,\overline{c}}$ that we consider is to pay dividends at rate p up to the arrival of the first claim, which is the ruin time.

Remark 2.2 The dividend optimization problem without the ratcheting constraint, that is where the dividend strategy $C = (C_t)_{t\geq 0}$ is not necessarily non-decreasing, was studied intensively in the literature (see e.g. Gerber and Shiu [14], Schmidli [18, Sec.2.4] and Azcue and Muler [5]). Unlike the ratcheting optimization problem, this non-ratcheting problem is one dimensional. If $V^{NR}(x)$ denotes the optimal value function of this non-ratcheting problem, then clearly $V(x, c) \leq V^{NR}(x)$ for all $x \geq 0$ and $c \in [0, \overline{c}]$. It is known that V^{NR} is non-decreasing with $\lim_{x\to\infty} V^{NR}(x) = \overline{c}/q$. Moreover, in [5], it was proved that there exists an optimal strategy and it has a band structure. It is characterized by three sets which partition the state space of the surplus process $[0, \infty)$. Each set is associated with a certain dividend payment action: define $\mathcal{O}^{NR}$ as the set of values where no dividends are paid, $\mathcal{B}^{NR}$ as the set of values where dividends are paid at the maximum possible rate $\overline{c}$ and $\mathcal{A}^{NR}$ as the set of values where dividends are paid at rate p. The topological properties of these sets depend on whether the premium rate p is larger than the dividend-rate ceiling $\overline{c}$; for example $\mathcal{A}^{NR}$ is empty if $\overline{c} < p$, and $\mathcal{B}^{NR}$ is empty in the case $\overline{c} = p$. The band strategies are *stationary* in the sense that they only depend on the current surplus. The simplest band strategies are the so-called *threshold* strategies, according to which dividends are paid at the maximal admissible rate $\overline{c}$ as soon as the surplus exceeds a certain threshold level $x_{NR} \geq 0$ and no dividends are paid when the surplus is less than x_{NR}. More precisely, the threshold strategy is characterized by the sets $\mathcal{O}^{NR} = [0, x_{NR})$ and $\mathcal{B}^{NR} = [x_{NR}, \infty)$ in the case $\overline{c} < p$, by the sets $\mathcal{O}^{NR} = [0, x_{NR})$ and $\mathcal{A}^{NR} = [x_{NR}, \infty)$ in the case $\overline{c} = p$, and by the sets $\mathcal{O}^{NR} = [0, x_{NR})$, $\mathcal{A}^{NR} = \{x_{NR}\}$ and $\mathcal{B}^{NR} = (x_{NR}, \infty)$ in the case $\overline{c} > p$.

Remark 2.3 If we denote by $V_{\overline{c}}(x, c)$ the optimal function defined in (2) with maximum rate $\overline{c} > 0$, it is straightforward that $V_{\overline{c}_1}(x, c) \leq V_{\overline{c}_2}(x, c)$ for $0 \leq c \leq \overline{c}_1 < \overline{c}_2$. Moreover, it is easy to see that for the limit $V_\infty(x, c) := \lim_{\overline{c}\to\infty} V_{\overline{c}}(x, c)$ one gets $x \leq V_\infty(x, c) \leq x + p/q$. Indeed, $V_\infty(x, c)$ is bounded from above by the optimal value function without ratcheting or dividend rate constraints defined in [6, Eq.1.10], and the latter can not exceed $x + p/q$ (see [6, Prop.1.2]). At the same time, for every $\overline{c}$

the function $V_{\overline{c}}(x,c)$ is bounded from below by the expected value of the dividend ratcheting strategy of paying the maximum rate $\overline{c}$ up to ruin, so that

$$V_{\infty}(x,c)=\lim_{\overline{c}\to\infty}V_{\overline{c}}(x,c)\geq\lim_{\overline{c}\to\infty}\mathbb{E}\left[I_{\tau_1\geq x/\overline{c}}\int_0^{x/\overline{c}}e^{-qs}\,\overline{c}\,ds\right]=\lim_{\overline{c}\to\infty}\frac{e^{-\beta x/\overline{c}}\,\overline{c}\,(1-e^{-qx/\overline{c}})}{q}=x.$$

3 Basic Results

Let us first derive some basic properties of the optimal value function (2). In the case $\overline{c}\leq p$ we will show that the optimal value function V is globally Lipschitz. In contrast, for $\overline{c}>p$ there are some issues with the regularity and the proofs are more involved. It is clear that in the case of $\overline{c}>p$, the optimal value function V is not continuous at the point $(0,p)$. Indeed, by Remark 2.1, we have that

$$V(0,p)=\mathbb{E}\left[\int_0^{\tau_1}e^{-qs}p\,ds\right]=\frac{p}{q+\beta};$$

but $V(0,c)=0$ for all $c\in(p,\overline{c}]$ because all the admissible strategies lead to immediate ruin. As a consequence,

$$\lim_{c\to p^+}V(0,c)=0<V(0,p)=p/(q+\beta).$$

In the case $\overline{c}>p$, we prove the following results depending on the value of c: (1) V is Lipschitz in $[0,\infty)\times[0,p]$, (2) V is continuous with respect to c in $(0,\infty)\times\{p\}$. (3) V is locally Lipschitz in $[0,\infty)\times(p,\overline{c}]$ with a Lipschitz bound that goes to infinity as $c\searrow p$. In particular, we conclude that V is continuous at any point except $(0,p)$.

Let us start with a straightforward result regarding the boundedness and monotonicity of the optimal value function.

Proposition 3.1 *The optimal value function $V(x,c)$ is bounded by $\overline{c}/q$, non-decreasing in x and non-increasing in c.*

Proof. Since the discounted value of paying the maximum rate $\overline{c}$ up to infinity is $\overline{c}/q$, we conclude the boundedness result.

On the one hand $V(x,c)$ is non-increasing in c because given $c_1<c_2$ we have $\Pi_{x,c_2,\overline{c}}\subset\Pi_{x,c_1,\overline{c}}$ for any $x\geq 0$. On the other hand, given $x_1<x_2$ and an admissible ratcheting strategy $C^1\in\Pi_{x_1,c,\overline{c}}$ for any $c\in[0,\overline{c}]$, let us define $C^2\in\Pi_{x_2,c,\overline{c}}$ as $C_t^2=C_t^1$ until the ruin time of the controlled process $X_t^{C^1}$ with $X_0^{C^1}=x_1$, and pay the maximum rate $\overline{c}$ afterwards. Thus, $J(x;C_1)\leq J(x;C_2)$ and we have the result. ■

Note that the previous proposition implies

$$0\leq V(x_2,c_1)-V(x_1,c_2)$$

for all $0\leq x_1\leq x_2$ and $c_1\leq c_2$.

In order to obtain the Lipschitz results we add the following assumption for technical reasons:

A1. The claim size distribution F is (globally) Lipschitz, that is if $x<y$, $0\leq F(y)-F(x)\leq K(y-x)$ for some $K>0$.

Note that for instance exponential, mixture of exponentials and Pareto claim size distributions satisfy the condition A.1; Gamma and Weibull distributions satisfy the condition A.1 for a shape parameter greater than or equal to one.

The following proposition establishes that V is Lipschitz in the case $\overline{c}\leq p$ and also in the case $\overline{c}>p$ for $(x,c)\in[0,\infty)\times[0,p]$.

Proposition 3.2 *There exists a constant $K_1>0$ such that*

$$0\leq V(x_2,c_1)-V(x_1,c_2)\leq K_1\left[(x_2-x_1)+(c_2-c_1)\right]$$

for all $0\leq x_1\leq x_2$ and $c_1\leq c_2\leq\min\{\overline{c},p\}$.

The proof of this proposition is given in the separate Appendix file.

Lipschitz bounds for the case $\overline{c} > p$ with $(x, c) \in [0, \infty) \times (p, \overline{c}]$ are as follows:

Proposition 3.3 *Assume that $\overline{c} > p$, then there exist constants $K_2 > 0$ and $K_3 > 0$ such that*

$$0 \leq V(x_2, c_1) - V(x_1, c_2) \leq \left[K_2 + \frac{K_3}{c_1 - p}\right](x_2 - x_1) + \left[K_2 + \frac{K_3 x_2}{(c_1 - p)^2}\right](c_2 - c_1)$$

for all $0 \leq x_1 \leq x_2$ and $p < c_1 \leq c_2 \leq \overline{c}$.

The proof of this proposition is given in the separate Appendix file.

Note that in the case $\overline{c} > p$, from Proposition 3.2 we have a global Lipschitz condition in $[0, \infty) \times [0, p]$ for V and Proposition 3.3 guarantees a local Lipschitz condition in $[0, \infty) \times (p, \overline{c}]$. The next proposition deals with the continuity from above of V in the set $(0, \infty) \times \{p\}$.

Proposition 3.4 *Assume $\overline{c} > p$, then we have that $\lim_{c \to p^+} V(x, c) = V(x, p)$ for $x > 0$.*

The proof of this proposition is given in the separate Appendix file.

4 Viscosity Solutions

In this section we introduce the Hamilton-Jacobi-Bellman (HJB) equation of the ratcheting problem and show that, in some sense, the optimal value function V defined in (2) is the unique viscosity solution of the HJB equation with boundary condition $\overline{c}/q$ when x goes to infinity. In the case that $\overline{c} \leq p$, we will prove that the optimal value function V is the unique viscosity solution in $(0, \infty) \times (0, \overline{c}]$ satisfying $\lim_{x \to \infty} V(x, c) = \overline{c}/q$. For $\overline{c} > p$, the scenario is more complex: we first prove that V is the unique viscosity solution in $(0, \infty) \times (p, \overline{c}]$ satisfying $\lim_{x \to \infty} V(x, c) = \overline{c}/q$ and afterwards that V is the unique viscosity solution in $(0, \infty) \times [0, p]$ satisfying $V(x, p) = \lim_{c \to p^+} V(x, c)$ for $x > 0$ (here, we use the continuity result of Proposition 3.4).

Let us define the operator

$$\mathcal{L}(u)(x, c) = c + (p - c)u_x(x, c) - (q + \beta)u(x, c) + \beta \int_0^x u(x - \alpha, c) dF(\alpha). \quad (3)$$

The Hamilton-Jacobi-Bellman equation associated to (2) is given by

$$\max\{\mathcal{L}(u)(x, c), u_c(x, c)\} = 0 \text{ for } x \geq 0 \text{ and } 0 \leq c \leq \overline{c}. \quad (4)$$

Definition 4.1 *(a) A locally Lipschitz function $\overline{u} : [0, \infty) \times [c_1, c_2) \to \mathbb{R}$, where $0 \leq c_1 < c_2 \leq \overline{c}$, is a viscosity supersolution of (4) at $(x, c) \in (0, \infty) \times [c_1, c_2)$ if any continuously differentiable function $\varphi : [0, \infty) \times [c_1, c_2) \to \mathbb{R}$ with $\varphi(x, c) = \overline{u}(x, c)$ such that $\overline{u} - \varphi$ reaches the minimum at (x, c) satisfies*

$$\max\{\mathcal{L}(\varphi)(x, c), \varphi_c(x, y)\} \leq 0.$$

The function φ is called a ***test function for supersolution*** *at (x, c).*

(b) A function $\underline{u} : [0, \infty) \times [c_1, c_2) \to \mathbb{R}$ is a viscosity subsolution of (4) at $(x, c) \in (0, \infty) \times [c_1, c_2)$ if any continuously differentiable function $\psi : [0, \infty) \times [c_1, c_2) \to \mathbb{R}$ with $\psi(x, c) = \underline{u}(x, c)$ such that $\underline{u} - \psi$ reaches the maximum at (x, c) satisfies

$$\max\{\mathcal{L}(\psi)(x, c), \psi_c(x, c)\} \geq 0.$$

The function ψ is called a ***test function for subsolution*** *at (x, c).*

(c) A function $u : [0, \infty) \times [c_1, c_2)$ which is both a supersolution and subsolution at $(x, c) \in [0, \infty) \times [c_1, c_2)$ is called a viscosity solution of (4) at (x, c).

We first prove that V is a viscosity solution of the HJB equation except at the points of the set $(0, \infty) \times \{p\}$ where V might not be locally Lipschitz. Let us first state the dynamic programming principle. The proof is similar to the one of Lemma 1.2 in [6].

Lemma 4.1 *Given any stopping time $\widetilde{\tau}$, we can write*

$$V(x,c)=\sup_{C\in\Pi_{x,c,\overline{c}}}\mathbb{E}\left[\int_0^{\tau\wedge\widetilde{\tau}}e^{-qs}C_s ds+e^{-q(\tau\wedge\widetilde{\tau})}V(X^C_{\tau\wedge\widetilde{\tau}},C_{\widetilde{\tau}})\right].$$

Proposition 4.2 *(i) If $\overline{c}\leq p$, V is a viscosity solution of (4) in $(0,\infty)\times[0,\overline{c})$. (ii) If $\overline{c}>p$, V is a viscosity solution of (4) in $(0,\infty)\times[0,p)$ and also in $(0,\infty)\times[c_1,\overline{c})$ for any $c_1>p$.*

Proof. We prove here part (i). The proof of part (ii) is similar.

Let us first show that V is a viscosity supersolution in $(0,\infty)\times[0,\overline{c})$ for $\overline{c}\leq p$. By Proposition 3.1, $V_c\leq 0$ in $(0,\infty)\times[0,\overline{c})$ in the viscosity sense.

Consider now $(x,c)\in(0,\infty)\times[0,\overline{c})$ and the admissible strategy $C\in\Pi_{x,c,\overline{c}}$ which pays dividends at constant rate c up to the ruin time τ. Let us denote the corresponding controlled surplus process as $X_t^C=X_t-ct$ and suppose that there exists a test function φ for supersolution (4) at (x,c). This means that φ is a continuously differentiable function $\varphi:[0,\infty)\times[0,\overline{c}]\to\mathbb{R}$ with $\varphi(x,c)=V(x,c)$ and such that $V-\varphi$ reaches the minimum at (x,c). We extend the definition of both V and φ as $\varphi=0$ for $x<0$. Using Lemma 4.1 we get for $h>0$,

$$\begin{array}{rcl}\varphi(x,c) & = & V(x,c)\\ & \geq & \mathbb{E}\left[\int_0^{\tau_1\wedge h}e^{-q\,s}\,cds\right]+\mathbb{E}\left[e^{-q(\tau_1\wedge h)}\varphi(X^C_{\tau_1\wedge h},c))\right].\end{array}$$

Hence,

$$\begin{array}{rcl}0 & \geq & \mathbb{E}\left[\int_0^{\tau_1\wedge h}e^{-q\,s}\,c\,ds\right]+\mathbb{E}\left[I_{\tau_1>h}e^{-q\,(\tau_1\wedge h)}\varphi(x+(p-c)h,c)\right]\\ & & +\mathbb{E}\left[I_{\tau_1\leq h}e^{-q\,(\tau_1\wedge h)}\varphi(x+(p-c)h-U_1,c)\right]-\varphi(x,c).\end{array}$$

So, dividing by h and taking $h\to 0^+$, we get

$$\mathcal{L}(\varphi)(x,c)\leq 0$$

and so it is a viscosity supersolution at (x,c).

Let us prove now that V is a viscosity subsolution in $(0,\infty)\times[0,\overline{c})$. Arguing by contradiction, we assume that V is not a subsolution of (4) at $(x,c)\in(0,\infty)\times[0,\overline{c})$, then there exist $\varepsilon>0$, $0<h<\min\{x/2,\overline{c}-c\}$ and a continuously differentiable function ψ with $\psi(x,c)=V(x,c)$ such that $\psi\geq V$,

$$\max\{\mathcal{L}(\psi)(y,d),\psi_c(y,d)\}\leq -q\varepsilon<0 \tag{5}$$

for $(y,d)\in[x-h,x+h]\times[c,c+h]$ and

$$V(y,d)\leq\psi(y,d)-\varepsilon \tag{6}$$

for $(y,d)\notin[x-h,x+h]\times[c,c+h]$.

Consider the controlled risk process X_t corresponding to an admissible strategy $C\in\Pi_{x,c,\overline{c}}$ and define

$$\tau^*=\inf\{t>0:\ \ (X_t,C_t)\notin[x-h,x+h]\times[c,c+h]\}.$$

Since C_t is non-decreasing and right-continuous, it can be written as

$$C_t=c+\int_0^t dC_s^{co}+\sum_{\substack{C_s\neq C_{s^-}\\ 0\leq s\leq t}}(C_s-C_{s^-}) \tag{7}$$

where C_s^{co} is a continuous and non-decreasing function.

Take a non-negative continuously differentiable function $\psi(x,c)$ in $(0,\infty)\times[0,\overline{c}]$. Since the function $e^{-qt}\psi(x,c)$ is continuously differentiable, using the expression (7) and the change of variables formula for finite variation processes (see for instance [16]), we can write

$$
\begin{array}{rl}
& \psi(X^C_{\tau^*}, C_{\tau^*})e^{-q\tau^*} - \psi(x,c) \\
= & \int_0^{\tau^*} e^{-qs}\psi_x(X^C_{s^-}, C_{s^-})(p - C_{s^-})ds + \int_0^{\tau^*} e^{-qs}\psi_c(X^C_{s^-}, C_{s^-})dC^{co}_s \\
& + \sum_{\substack{C_s \neq C_{s^-} \\ 0\le s\le \tau^*}} e^{-qs}(C_s - C_{s^-})\psi_c(X^C_{s^-}, C_{s^-}) \\
& + \sum_{\tau_i \le \tau^*} \left(\psi(X^C_{s^-} - U_i, C_{s^-}) - \psi(X^C_{s^-}, C_{s^-})\right) e^{-qs} - q\int_0^{\tau^*} \psi(X^C_{s^-}, C_{s^-})e^{-qs}ds.
\end{array}
\tag{8}
$$

We have that

$$
\begin{array}{rl}
M_t = & \sum_{\tau_i \le \tau} \left(\psi(X^C_{s^-} - U_i, C_{s^-}) - \psi(X^C_{s^-}, C_{s^-})\right) e^{-qs} \\
& -\beta\int_0^t e^{-qs}\int_0^\infty \left(\psi(X^C_{s^-} - \alpha, C_{s^-}) - \psi(X^C_{s^-}, C_{s^-})\right) dF(\alpha)ds
\end{array}
$$

is a martingale with zero expectation. Hence, from (5), we can write

$$
\begin{array}{rl}
& e^{-q\tau^*}\psi(X_{\tau^*}, C_{\tau^*}) - \psi(x,c) \\
= & \int_0^{\tau^*} e^{-qs}\mathcal{L}_{C_{s^-}}(\psi)(X^C_{s^-}, C_{s^-})ds + M_{\tau^*} - \int_0^{\tau^*} e^{-qs}C_{s^-}ds \\
& + \int_0^{\tau^*} e^{-qs}\psi_c(X^C_{s^-}, C_{s^-})dC^c_s + \sum_{\substack{C_s \neq C_{s^-} \\ 0\le s\le \tau^*}} e^{-qs}(C_s - C_{s^-})\psi_c(X^C_{s^-}, C_{s^-}) \\
\le & \int_0^{\tau^*} e^{-qs}(-q\varepsilon)ds + M_{\tau^*} - \int_0^{\tau^*} e^{-qs}C_{s^-}ds \\
& + \int_0^{\tau^*} e^{-qs}(-q\varepsilon)dC^c_s + \sum_{\substack{C_s \neq C_{s^-} \\ 0\le s\le \tau^*}} e^{-qs}(C_s - C_{s^-})(-q\varepsilon) \\
= & \varepsilon\left(e^{-q\tau^*} - 1\right) + M_{\tau^*} - \int_0^{\tau^*} e^{-qs}C_{s^-}ds - q\varepsilon\left(\int_0^{\tau^*} e^{-qs}dC_s\right).
\end{array}
$$

So, taking expectation we obtain from (6)

$$
\begin{array}{rl}
& \mathbb{E}\left[e^{-q\tau^*}V(X_{\tau^*}, C_{\tau^*})\right] \\
\le & \mathbb{E}\left[\psi(x,c) - e^{-q\tau^*}\varepsilon\right] + \mathbb{E}\left[\psi(X_{\tau^*}, C_{\tau^*})e^{-q\tau^*} - \psi(x,c)\right] \\
\le & \psi(x,c) - \varepsilon\mathbb{E}\left[e^{-q\tau^*}\right] - \left(\varepsilon\mathbb{E}\left[1 - e^{-q\tau^*}\right] + q\varepsilon\mathbb{E}\left[\int_0^{\tau^*} e^{-qs}dC_s\right]\right) - \mathbb{E}\left[\int_0^{\tau^*} e^{-qs}C_{s^-}ds\right] \\
\le & \psi(x,c) - \varepsilon\mathbb{E}\left[e^{-q\tau^*}\right] - \varepsilon\mathbb{E}\left[1 - e^{-q\tau^*}\right] - \mathbb{E}\left[\int_0^{\tau^*} e^{-qs}C_{s^-}ds\right] \\
= & \psi(x,c) - \varepsilon - \mathbb{E}(\int_0^{\tau^*} e^{-qs}C_{s^-}ds).
\end{array}
$$

Hence, using the dynamic programming principle of Lemma 4.1, we have that

$$
V(x,c) = \sup_{C\in\Pi_{x,c,\overline{c}}} \mathbb{E}\left(\int_0^{\tau^*} e^{-qs}C_{s^-}ds + e^{-c\tau^*}V(X^C_{\tau^*}, C_{\tau^*})\right) \le \psi(x,c) - \varepsilon.
$$

but this is a contradiction in view of the assumption $V(x,c) = \psi(x,c)$. ■

$V(x,\overline{c})$ corresponds to the value function of the strategy that pays dividends at constant rate $\overline{c}$, so the following lemma is a standard one-dimensional result.

Lemma 4.3 *The optimal value function $V(x,\overline{c})$ is a solution of $\mathcal{L}(V)(x,\overline{c}) = 0$ for $x > 0$, where $\mathcal{L}$ is the operator defined in (3).*

We now show that V satisfies a boundary condition as x goes to infinity.

Proposition 4.4 *The optimal value function V satisfies $\lim_{x\to\infty} V(x,c) = \overline{c}/q$ for all $c \in [0,\overline{c}]$.*

Proof. We first prove the result for $c = \overline{c}$. Since $V(\cdot,\overline{c}) : [0,\infty) \to \mathbb{R}$ is bounded, non-decreasing and Lipschitz, the set $\widetilde{\mathcal{D}}$ of points where $V(\cdot,\overline{c})$ is differentiable has full measure in $[0,\infty)$. Let us prove that there exists a sequence $x_n \to \infty$ such that $x_n \in \widetilde{\mathcal{D}}$, $V_x(x_n,\overline{c}) \to 0$. Given any $\varepsilon > 0$, if $V_x(x,\overline{c}) \geq \varepsilon$ for all $x \in \widetilde{\mathcal{D}} \cap [n,\infty)$ with $n \in \mathbb{N}$, then $V(x,\overline{c})$ cannot be bounded and this is a contradiction, so such a sequence exists. We obtain

$$\begin{aligned} 0 &= \lim_{n\to\infty} \mathcal{L}(V)(x_n,\overline{c}) \\ &= \overline{c} + (p-\overline{c}) \lim_{n\to\infty} V_x(x_n,\overline{c}) - (q+\beta)\lim_{x\to\infty} V(x_n,\overline{c}) + \beta \int_0^\infty \lim_{x\to\infty} V(x,\overline{c}) dF(\alpha) \\ &= \overline{c} - q \lim_{x\to\infty} V(x,\overline{c}). \end{aligned}$$

Finally, for $c \in [0,\overline{c})$, since $V(x,c) \leq \overline{c}/q$ but also $V(x,c) \geq V(x,\overline{c})$ for all $c \in [0,\overline{c}]$, so that we obtain the result. ■

We now give the comparison result for viscosity solutions.

Lemma 4.5 *In the case $\overline{c} \leq p$ consider any interval $[c_1,c_2] \subset [0,\overline{c}]$ and in the case $\overline{c} > p$ consider any interval $[c_1,c_2] \subset [0,p]$ or any interval $[c_1,c_2] \subset (p,\overline{c}]$. Let us assume that (i) $\underline{u}$ is a viscosity subsolution and $\overline{u}$ is a viscosity supersolution of the HJB equation (4) for all $x > 0$ and for all $c \in (c_1,c_2)$, (ii) $\underline{u}$ and $\overline{u}$ are non-decreasing in the variable x and Lipschitz in $[0,\infty) \times [c_1,c_2]$, (iii) $\lim_{x\to\infty} \underline{u}(x,c) = \lim_{x\to\infty} \overline{u}(x,c) = L > 0$ and (iv) $\underline{u}(x,c_2) \leq \overline{u}(x,c_2)$ for all $x \geq 0$; then $\underline{u} \leq \overline{u}$ in $[0,\infty) \times [c_1,c_2]$.*

Proof. The proof of this lemma is a two-dimensional generalization of Proposition 4.2 of Azcue and Muler [5]. We use in this proof an equivalent formulation of viscosity solution, see for example Sayah [17] and Benth, Karlsen and Reikvam [10]: Let us define the operators

$$\begin{aligned} &\mathcal{L}(u,\psi)(x,c) = c + (p-c)\psi_x(x,c) - (q+\beta)u(x,c) + \beta \int_0^x u(x-\alpha,c) dF(\alpha) \text{ and} \\ &\overline{\mathcal{L}}(\psi)(x,c) = \psi_c(x,c). \end{aligned}$$

A locally Lipschitz function $\overline{u} : [0,\infty) \times [c_1,c_2] \to \mathbb{R}$ is a viscosity supersolution of (4) at $(x,c) \in (0,\infty) \times (c_1,c_2)$ if any test function φ for supersolution at (x,c) satisfies

$$\max\{\mathcal{L}(\overline{u},\varphi)(x,c), \overline{\mathcal{L}}(\varphi)(x,c)\} \leq 0,$$

and a locally Lipschitz function $\underline{u} : [0,\infty) \times [c_1,c_2] \to \mathbb{R}$ is a viscosity subsolution of (4) at $(x,c) \in (0,\infty) \times (c_1,c_2)$ if any test function ψ for subsolution at (x,c) satisfies

$$\max\{\mathcal{L}(\underline{u},\psi)(x,c), \overline{\mathcal{L}}(\psi)(x,c)\} \geq 0.$$

Suppose that there is a point $(x_0,c_0) \in [0,\infty) \times (c_1,c_2)$ such that $\underline{u}(x_0,c_0) - \overline{u}(x_0,c_0) > 0$. Let us define $h(c) = 1 + e^{-\frac{c}{c_2}}$ and

$$\overline{u}^s(x,c) = sh(c)\overline{u}(x,c)$$

for any $s > 1$. We have that φ is a test function for supersolution of $\overline{u}$ at (x,c) if and only if $\varphi^s = sh(c)\varphi$ is a test function for supersolution of $\overline{u}^s$ at (x,c). We have

$$\mathcal{L}(\overline{u}^s,\ \varphi^s)(x,c) = sh(c)\mathcal{L}(\overline{u},\varphi)(x,c) + c(1-sh(c)) < 0, \tag{9}$$

and

$$\overline{\mathcal{L}}(\varphi^s)(x,c) \leq -\frac{s}{c_2}\varphi(x,c)e^{-\frac{c}{c_2}} < 0 \tag{10}$$

for $\varphi(x,c) > 0$. Take $s_0 > 1$ such that $\underline{u}(x_0,c_0) - \overline{u}^{s_0}(x_0,c_0) > 0$. We define

$$M = \sup_{x\geq 0, c_1\leq c\leq c_2} \left(\underline{u}(x,c) - \overline{u}^{s_0}(x,c)\right). \tag{11}$$

Since $\lim_{x\to\infty} \underline{u}(x,c) = \lim_{x\to\infty} \overline{u}(x,c) = L$, we have that there exists a $b > x_0$ such that

$$\sup_{x\geq 0, c_1\leq c,d\leq c_2} \underline{u}(x,c) - \overline{u}^{s_0}(x,d) < 0 \text{ for } x \geq b. \tag{12}$$

We obtain from (12) that

$$0 < \underline{u}(x_0,c_0) - \overline{u}^{s_0}(x_0,c_0) \leq M := \max_{x\in[0,b], c_1\leq c\leq c_2} \left(\underline{u}(x,c) - \overline{u}^{s_0}(x,c)\right). \tag{13}$$

Call $(x^*, c^*) := \arg \max_{x \in [0,b], c_1 \leq c \leq c_2} (\underline{u}(x,c) - \overline{u}^{s_0}(x,c))$. Since $\underline{u}$ and $\overline{u}^{s_0}$ are locally Lipschitz, there exists a constant $m > 0$ such that

$$|\underline{u}(x,c) - \underline{u}(y,d)| \leq m \left\|(x-y, c-d)\right\|_2 \text{ and } |\overline{u}^{s_0}(x,c) - \overline{u}^{s_0}(y,d)| \leq m \left\|(x-y, c-d)\right\|_2 \qquad (14)$$

for $0 \leq x_2 \leq x_1 \leq b$. Consider the set

$$A = \{(x,y,c,d) : 0 \leq x \leq y \leq b, c_1 \leq \ c \leq c_2, c_1 \leq d \leq c_2\}$$

and, for all $\lambda > 0$, the functions

$$\begin{array}{l} \Phi^{\lambda}(x,y,c,d) = \dfrac{\lambda}{2}(x-y)^2 + \dfrac{\lambda}{2}(c-d)^2 + \frac{2m}{\lambda^2(y-x)+\lambda}, \\ \Sigma^{\lambda}(x,y,c,d) = \underline{u}(x,c) - \overline{u}^{s_0}(y,d) - \Phi^{\lambda}(x,y,c,d). \end{array} \qquad (15)$$

We have that the partial derivatives are

$$\begin{array}{l} \Phi^{\lambda}_x(x,y,c,d) = \lambda(x-y) + \frac{2m}{(\lambda(y-x)+1)^2}, \ \Phi^{\lambda}_y(x,y,c,d) = -\lambda(x-y) - \frac{2m}{(\lambda(y-x)+1)^2}; \\ \Phi^{\lambda}_c(x,y,c,d) = \lambda(c-d) \ \text{ and } \Phi^{\lambda}_d(x,y,c,d) = -\lambda(c-d). \end{array}$$

Calling $M^{\lambda} = \max_A \Sigma^{\lambda}$ and $(x_{\lambda}, y_{\lambda}, c_{\lambda}, d_{\lambda}) = \arg\max_A \Sigma^{\lambda}$, we obtain that $M^{\lambda} \geq \Sigma^{\lambda}(x^*, x^*, c^*, c^*) = M$ and so

$$\liminf_{\lambda \to \infty} M^{\lambda} \geq M. \qquad (16)$$

We show that there exists λ_0 large enough such that if $\lambda \geq \lambda_0$, then $(x_{\lambda}, y_{\lambda}, c_{\lambda}, d_{\lambda}) \notin \partial A$. The maximum is not achieved on the boundary $y = x$ because

$$\liminf_{h \searrow 0} \frac{\Sigma^{\lambda}(x,x+h,c,d) - \Sigma^{\lambda}(x,x,c,d)}{h} = \liminf_{h \searrow 0} \frac{\overline{u}^{s_0}(x,d) - \overline{u}^{s_0}(x+h,d)}{h} - \Phi^{\lambda}_y(x,x,c,d) \geq m > 0. \qquad (17)$$

Let us see now that the maximum is also not achieved on the boundary $x = 0$. Since Σ^{λ} is continuous and locally Lipschitz, by (17) there exists an open set

$$O_1 \supset \{(0,0)\} \times [c_1, c_2] \times [c_1, c_2],$$

where Σ^{λ} does not achieve the maximum. Correspondingly, there exists an $\varepsilon > 0$ such that the maximum of Σ^{λ} is not achieved at the points $(0,y,c,d)$ with $0 \leq y < \varepsilon$. Moreover, since $\underline{u}$ is a non-decreasing function in x, we have from (12) and (15) that

$$\begin{array}{rcl} \liminf_{h \searrow 0} \frac{\Sigma^{\lambda}(h,y,c,d) - \Sigma^{\lambda}(0,y,c,d)}{h} & = & \liminf_{h \searrow 0} \frac{\underline{u}(h,c) - \underline{u}(0,c)}{h} - \Phi^{\lambda}_x(0,y,c,d) \\ & \geq & \lambda y - \frac{2m}{(\lambda y+1)^2} > 0 \end{array} \qquad (18)$$

for λ large enough if $y > \varepsilon > 0$; so the maximum is not achieved on the boundary $x = 0$. With similar arguments, it can be proved that the maximum is not achieved on the boundaries $y = b$, $c = c_1,\ c = c_2,\ d = c_1$ and $d = c_2$.

Since $\Sigma^{\lambda}(x,y,c,d) = \underline{u}(x,c) - \overline{u}^{s_0}(y,d) - \Phi^{\lambda}(x,y,c,d)$ reaches the maximum in $(x_{\lambda}, y_{\lambda}, c_{\lambda}, d_{\lambda})$ in the interior of the set A, the function

$$\psi(x,c) = \Phi^{\lambda}(x, y_{\lambda}, c, d_{\lambda}) - \Phi^{\lambda}(x_{\lambda}, y_{\lambda}, c_{\lambda}, d_{\lambda}) + \underline{u}(x_{\lambda}, c_{\lambda})$$

is a test for subsolution for $\underline{u}$ at $(x_{\lambda}, c_{\lambda})$, and so

$$\max\{\mathcal{L}(\underline{u}, \psi)(x_{\lambda}, c_{\lambda}), \overline{\mathcal{L}}(\psi)(x_{\lambda}, c_{\lambda})\} \geq 0. \qquad (19)$$

Furthermore,

$$\varphi^{s_0}(y,d) = -\Phi^{\lambda}(x_{\lambda}, y, c_{\lambda}, d) + \Phi^{\lambda}(x_{\lambda}, y_{\lambda}, c_{\lambda}, d_{\lambda}) + \overline{u}^{s_0}(y_{\lambda}, d_{\lambda})$$

is a test for supersolution for $\overline{u}^{s_0}$ at (y_λ, d_λ) and so

$$\max\{\mathcal{L}(\overline{u}^{s_0}, \varphi^{s_0})\,(y_\lambda, d_\lambda)\,, \overline{\mathcal{L}}(\varphi^{s_0})\,(y_\lambda, d_\lambda)\} \leq 0. \tag{20}$$

Since

$$\overline{\mathcal{L}}(\varphi^s)(y_\lambda, d_\lambda) \leq -\frac{s_0}{c_2}\varphi(y_\lambda, d_\lambda)e^{-\frac{c}{c_2}} < 0$$

(because $y_\lambda > 0$) and $\overline{\mathcal{L}}(\psi)(x_\lambda, c_\lambda) = \overline{\mathcal{L}}(\varphi^s)(y_\lambda, d_\lambda) < 0$, we have from (19) that

$$\mathcal{L}(\underline{u}, \psi)(x_\lambda, c_\lambda) \geq 0. \tag{21}$$

Therefore, from (20), (19) and $\psi_x(x_\lambda, c_\lambda) = \varphi^s_x(y_\lambda, d_\lambda)$, we get

$$\begin{array}{l}\mathcal{L}(\overline{u}^{s_0}, \varphi^{s_0})\,(y_\lambda, d_\lambda) - \mathcal{L}(\underline{u}, \psi)(x_\lambda, c_\lambda) \\ = (q+\beta)\,(\underline{u}(x_\lambda, c_\lambda) - \overline{u}^{s_0}(y_\lambda, d_\lambda)) \\ \leq \beta\left(\int_0^{x_\lambda} \underline{u}(x_\lambda - \alpha, c_\lambda)dF(\alpha) - \int_0^{y_\lambda} \overline{u}^{s_0}(y_\lambda - \alpha, d_\lambda)dF(\alpha)\right).\end{array} \tag{22}$$

Using the inequality

$$\Sigma^\lambda\,(x_\lambda, x_\lambda, c_\lambda, c_\lambda) + \Sigma^\lambda\,(y_\lambda, y_\lambda, d_\lambda, d_\lambda) \leq 2\Sigma^\lambda\,(x_\lambda, y_\lambda, c_\lambda, d_\lambda)$$

we obtain that

$$\lambda\,(x_\lambda - y_\lambda)^2 + \lambda\,(c_\lambda - d_\lambda)^2 \leq \underline{u}(x_\lambda, c_\lambda) - \underline{u}(y_\lambda, d_\lambda) + \overline{u}^{s_0}(x_\lambda, c_\lambda) - \overline{u}^{s_0}(y_\lambda, d_\lambda) + 4m(y_\lambda - x_\lambda),$$

which together with (14) gives

$$\lambda\,\|(x_\lambda - y_\lambda, c_\lambda - d_\lambda)\|_2^2 \leq 6m\,\|(x_\lambda - y_\lambda, c_\lambda - d_\lambda)\|_2\,. \tag{23}$$

We can find a sequence $\lambda_n \to \infty$ such that $(x_{\lambda_n}, y_{\lambda_n}, c_{\lambda_n}, d_{\lambda_n}) \to \left(\widehat{x}, \widehat{y}, \widehat{c}, \widehat{d}\right) \in A$. From (23), we get that $\|(x_{\lambda_n} - y_{\lambda_n}, c_{\lambda_n} - d_{\lambda_n})\|_2 \leq \dfrac{6m}{\lambda_n}$, which gives $\widehat{x} = \widehat{y}$ and $\widehat{c} = \widehat{d}$. Using that $y_{\lambda_n} \geq x_{\lambda_n}$ for all n, we obtain from (22) that

$$(q+\beta)\,(\underline{u}(\widehat{x}, \widehat{c}) - \overline{u}^{s_0}(\widehat{x}, \widehat{c})) \leq \beta\left(\int_0^{\widehat{x}} (\underline{u}(\widehat{x} - \alpha, \widehat{c}) - \overline{u}^{s_0}(\widehat{x} - \alpha, \widehat{c}))\,dF(\alpha)\right) \leq \beta M. \tag{24}$$

From (23) we get that $\lim_{n\to\infty} \lambda_n\,\|(x_{\lambda_n} - y_{\lambda_n}, c_{\lambda_n} - d_{\lambda_n})\|_2^2 = 0$, hence from (16) and (24) we obtain

$$\begin{array}{rl} M \leq \liminf\limits_{\lambda\to\infty} M_\lambda \leq \lim\limits_{n\to\infty} M_{\lambda_n} &= \lim\limits_{n\to\infty} \Sigma^{\lambda_x}(x_{\lambda_n}, y_{\lambda_n}, c_{\lambda_n}, d_{\lambda_n}) = \underline{u}(\widehat{x}, \widehat{c}) - \overline{u}^{s_0}(\widehat{x}, \widehat{c}) \\ &\leq \dfrac{\beta}{q+\beta}\displaystyle\int_0^{\widehat{x}} (\underline{u}(\widehat{x} - \alpha, \widehat{c}) - \overline{u}^{s_0}(\widehat{x} - \alpha, \widehat{c}))\,dF(\alpha) \leq \dfrac{\beta}{q+\beta}M. \end{array}$$

This is a contradiction, which establishes the result. ■

As a consequence of the previous lemma, we have the following two propositions concerning uniqueness for the cases $\overline{c} \leq p$ and $\overline{c} > p$. In the case $\overline{c} \leq p$ the uniqueness is a direct consequence of Lemmas 4.5 and 4.4 together with Proposition 4.2.

Proposition 4.6 *If $\overline{c} \leq p$, the optimal value function V is the unique function non-decreasing in x that is a viscosity solution of (4) in $(0,\infty) \times [0,\overline{c})$ with limit $\overline{c}/q$ as $x \to \infty$.*

In the case $\overline{c} > p$ the uniqueness is a direct consequence of Lemmas 4.5 and 4.4 together with Propositions 4.2 and 3.4.

Proposition 4.7 *If $\overline{c} > p$, the optimal value function V is the unique continuous function non-decreasing in x that has limit $\overline{c}/q$ as $x \to \infty$ for all $c \in [0,\overline{c}]$ and that is a viscosity solution of (4) both in $[0,\infty) \times [0,p]$ and in $[0,\infty) \times [c_1,\overline{c})$ for any $c_1 \in (p,\overline{c})$.*

From Definition 2, Lemma 4.5, and 4.4 together with Proposition 4.2, we also get the following verification theorem that will be used in the next section.

Theorem 4.8 *In the case $\overline{c} \leq p$ consider any interval $[c_1, c_2] \subset [0, \overline{c}]$ and in the case $\overline{c} > p$ consider any interval $[c_1, c_2] \subset [0, p]$ or any interval $[c_1, c_2] \subset (p, \overline{c}]$. Consider a family of strategies $\{C_{x,c} \in \Pi_{x,c,\overline{c}} : (x, c) \in [0, \infty) \times [c_1, c_2]\}$. If the function $W(x, c) := J(x; C_{x,c})$ is a viscosity supersolution of the HJB equation (4) in $(0, \infty) \times (c_1, c_2)$ with $\lim_{x\to\infty} W(x, c) = \overline{c}/q$, then W is the optimal value function V. Also, if for each $k \geq 1$ there exists a family of strategies $\left\{C^k_{x,c} \in \Pi_{x,c,\overline{c}} : (x, c) \in [0, \infty) \times [c_1, c_2]\right\}$ such that $W(x, c) := \lim_{k\to\infty} J(x; C^k_{x,c})$ is a viscosity supersolution of the HJB equation (4) in $(0, \infty) \times (c_1, c_2)$ with $\lim_{x\to\infty} W(x, c) = \overline{c}/q$, then W is the optimal value function V.*

5 Finite ratcheting strategies

In this section we introduce ratcheting strategies when only a finite number N of dividend rates are possible and find the optimal value function in this restricted setting. This optimization problem is no longer two-dimensional and can be reduced to N one-dimensional obstacle problems. We also show that there exists an optimal finite ratcheting strategy and we construct it recursively. In Section 6 we will then use the optimal value function of this restricted setting to approximate the optimal value function V for the general case.

Consider a finite set $\mathcal{G} = \{c_1, c_2, \ldots c_N\}$ in the interval $[0, \overline{c}]$ with $c_k < c_{k+1}$ and $c_N = \overline{c}$. The task is then to find the optimal value function among the ratcheting strategies with dividend rates in $\mathcal{G}$. To that end, let us define the family of admissible strategies $\Pi^{\mathcal{G}}_{x,c,\overline{c}} \subset \Pi_{x,c,\overline{c}}$ as

$$\Pi^{\mathcal{G}}_{x,c,\overline{c}} = \{C \in \Pi_{x,c,\overline{c}} \text{ such that } \operatorname{Im}(C) \subset \mathcal{G}\}.$$

and the optimal value function within the restricted class as

$$V^{\mathcal{G}}(x, c) = \sup_{C\in\Pi^{\mathcal{G}}_{x,c,\overline{c}}} J(x; C). \tag{25}$$

By definition, $V^{\mathcal{G}}(x, c) = V^{\mathcal{G}}(x, \widetilde{c})$ where $\widetilde{c} = \min\{c_k \in \mathcal{G} : c_k \geq c\}$ and $V^{\mathcal{G}}(x, c) \leq V(x, c)$ for all finite sets $\mathcal{G}$.

Let us first state some basic properties of $V^{\mathcal{G}}$. These properties mirror the properties of the optimal value function V. The proofs are the same as those of Propositions 3.1, 3.2 and 4.4 but considering admissible strategies in the set $\Pi^{\mathcal{G}}_{x,c,\overline{c}}$ instead of $\Pi_{x,c,\overline{c}}$.

Proposition 5.1

1. *$V^{\mathcal{G}}(x, c)$ is non-increasing in c with $V^{\mathcal{G}}(x, \overline{c}) = V(x, \overline{c})$, and non-decreasing in x with $\lim_{x\to\infty} V^{\mathcal{G}}(x, c) = \overline{c}/q$.*
2. *There exists a constant $K_1 > 0$ such that*

$$0 \leq V^{\mathcal{G}}(x_2, c_k) - V^{\mathcal{G}}(x_1, c_l) \leq K_1 \left[(x_2 - x_1) + (c_l - c_k)\right]$$

for all $0 \leq x_1 \leq x_2$, and $c_k \in \mathcal{G}, c_l \in \mathcal{G}$ with $c_k \leq c_l \leq \min\{\overline{c}, p\}$.

3. *Assume that $\overline{c} > p$, then there exist constants $K_2 > 0$ and $K_3 > 0$ such that*

$$0 \leq V^{\mathcal{G}}(x_2, c_k) - V^{\mathcal{G}}(x_1, c_l) \leq \left[K_2 + \frac{K_3}{c_k - p}\right](x_2 - x_1) + \left[K_2 + \frac{K_3 x_2}{(c_k - p)^2}\right](c_l - c_k)$$

for all $0 \leq x_1 \leq x_2$, and $c_k \in \mathcal{G}, c_l \in \mathcal{G}$ with $p < c_k \leq c_l \leq \overline{c}$.

From the problem definition, for any given finite set $\mathcal{G} = \{c_1, c_2, \ldots c_N\}$ in the interval $[0, \overline{c}]$ with $c_k < c_{k+1}$ and $c_N = \overline{c}$, we have that

$$V^{\mathcal{G}}(x, c_N) = V^{\mathcal{G}}(x, \overline{c}) = \mathbb{E}\left[\int_0^\tau e^{-qs}\, \overline{c}\, ds\right].$$

We can now describe $V^{\mathcal{G}}(x, c_k)$ for $k = N-1, N-2, ..., 1$ recursively as a problem of optimal irreversible switching times as follows. Analogously to Section 2 in Azcue and Muler [7], consider the decision-time problem with obstacle function $V^{\mathcal{G}}(x, c_{k+1})$. Given any initial surplus $x \geq 0$ and $c_k \in \mathcal{G}$, take the strategy that pays dividends at constant rate c_k up to the stopping time $T_k \geq 0$. Define

$$V^{\mathcal{G}}(x, c_k) = \sup_{T_k} \left(\mathbb{E}\left[\int_0^{T_k \wedge \tau} e^{-qs} c_k ds \right] + \mathbb{E}[e^{-q(T_k \wedge \tau)} V^{\mathcal{G}}(X_{T_k \wedge \tau}, c_{k+1})] \right). \tag{26}$$

The value $V^{\mathcal{G}}(x, c_k)$ can be interpreted as the expected discounted dividend payment at rate c_k up to the optimal stopping time $T_k \wedge \tau$ plus an exit dividend payment of $V^{\mathcal{G}}(X_{T_k \wedge \tau}, c_{k+1})$ at this time. With this recursive construction, the decision time T_k corresponds to the time at which the admissible strategy $C = (C_t)_{t \geq 0} \in \Pi^{\mathcal{G}}_{x,c,\overline{c}}$ changes from c_k to some c_l with $l > k$. We define $T_N = \infty$ because $c_N = \overline{c}$ is the maximum possible dividend rate.

Let us define the operator

$$\mathcal{L}_c(v)(x) := c + (p - c)v'(x) - (q + \beta)v(x) + \beta \int_0^x v(x - \alpha) dF(\alpha), \tag{27}$$

then by Lemma 4.3 and Proposition 4.4, $V^{\mathcal{G}}(\cdot, c_N)$ is the unique solution of the integro-differential equation $\mathcal{L}_{\overline{c}}(v)(x) = 0$ in $[0, \infty)$ with $\lim_{x \to \infty} u(x) = \overline{c}/q$. Analogously to the proofs of Section 3 in Azcue and Muler [7], it can be proved that $V^{\mathcal{G}}(\cdot, c_k)$ is a viscosity solution of the obstacle problem

$$\max\{\mathcal{L}_{c_k}(v)(x), V^{\mathcal{G}}(x, c_{k+1}) - v(x)\} = 0 \tag{28}$$

in $(0, \infty)$ and that $V^{\mathcal{G}}(\cdot, c_k)$ is the smallest viscosity supersolution of (28) with $\lim_{x \to \infty} u(x) = \overline{c}/q$.

Remark 5.1 From the previous result, we conclude that if the value function of any admissible strategy in $\Pi^{\mathcal{G}}_{x,c_k,\overline{c}}$ is a viscosity supersolution of (28), then it is $V^{\mathcal{G}}(x, c_k)$.

If we define the closed sets

$$\mathcal{D}^*_k := \{x : V^{\mathcal{G}}(x, c_{k+1}) - V^{\mathcal{G}}(x, c_k) = 0\} \text{ for } k \leq N-1, \tag{29}$$

we have that $\mathcal{D}^*_k$ is non-empty, because otherwise $\lim_{x \to \infty} V^{\mathcal{G}}(x, c_k) = c_k/q < \overline{c}/q$. Given any initial surplus $x \geq 0$ and $c_k \in \mathcal{G}$, the optimal decision time T^*_k is the first time at which the surplus process X_t hits $\mathcal{D}^*_k$. We define $\mathcal{D}^*_N = \varnothing$ because $c_N = \overline{c}$ is the maximum possible dividend rate.

We have that there exist optimal strategies $C^{\mathcal{G}}_{x,c} \in \Pi^{\mathcal{G}}_{x,c,\overline{c}}$ of the (restricted) optimization problem (25) for any $(x, c) \in [0, \infty) \times [0, \overline{c}]$ and these strategies are described by the *optimal change region*

$$\mathcal{D}^{\mathcal{G}} = \textstyle\bigcup_{k=1}^{N-1} \mathcal{D}^*_k \times \{c_k\} \subset [0, \infty) \times \mathcal{G} \tag{30}$$

in the following way:

- Given $(x, c) \in [0, \infty) \times [0, \overline{c}]$, take $k_1 = \min\{k : \text{with } 1 \leq k \leq N,\ c_k \geq c \text{ and } x \notin \mathcal{D}^*_k\}$, and pay dividends at constant rate c_{k_1} up to the first time T_{k_1} that the controlled surplus process X_t hits $\mathcal{D}^*_{k_1}$.
- Take $k_2 = \min\{k : \text{with } k_1 < k \leq N \text{ and } X_{T_{k_1}} \notin \mathcal{D}^*_k\}$, and pay dividends at constant rate c_{k_2} up to the first time T_{k_2} that the controlled surplus process X_t hits $\mathcal{D}^*_{k_2}$, etc. Note, that since $1 \leq k_1 < k_2 < \cdots \leq N$, the number of $k_i's$ is at most N.

More precisely, the optimal strategy $C^{\mathcal{G}}_{x,c} = (C_t)_{t \geq 0} \in \Pi^{\mathcal{G}}_{x,c,\overline{c}}$ is given by

$$C_t = \sum_{i \geq 1} c_{k_i} I_{t \in [T_{k_i} \wedge \tau, T_{k_{i+1}} \wedge \tau)}.$$

Note that the optimal strategies $C^{\mathcal{G}}_{x,c} \in \Pi^{\mathcal{G}}_{x,c,\overline{c}}$ are *stationary* in the state space $[0, \infty) \times [0, \overline{c}]$, in the sense that the current dividend rate depends only on $(X_t, C_{t-}) \in [0, \infty) \times [0, \overline{c}]$ (using the notation $C_{0^-} = c$).

Remark 5.2 Given a finite set $\mathcal{G}$, $\mathcal{D}^*_k$ with $k = 1, ..., N$ are called the *optimal change sets* and the complements $\mathcal{U}^*_k = [0, \infty) - \mathcal{D}^*_k$ are called the *optimal non-change sets.* In Section 7, we will find examples in which the closed sets $\mathcal{D}^*_k$ defined in (29) are of the form $[d^*_k, \infty)$ and satisfy $\mathcal{D}^*_{k+1} \subset \mathcal{D}^*_k$. In this case d^*_k is called the *threshold* of the set $\mathcal{D}^*_k$. We will also find examples where the closed sets $\mathcal{D}^*_k$ have two connected components and $\mathcal{D}^*_{k+1} \not\subseteq \mathcal{D}^*_k$.

6 Approximation with value functions of finite ratcheting strategies

Let us now use the value functions $V^{\mathcal{G}}$ of finite ratcheting strategies to approximate the optimal value function V as the mesh size of $\mathcal{G}$ goes to zero. Consider for $n \geq 0$, a sequence of sets $\mathcal{G}^n$ (with $k_n + 1$ elements) of the form

$$\mathcal{G}^n = \{c_0^n = 0 < c_1^n < c_2^n < \cdots < c_{k_n}^n = \overline{c}\}.$$

satisfying $\mathcal{G}^0 = \{0, \overline{c}\}$, $\mathcal{G}^n \subset \mathcal{G}^{n+1}$ and mesh-size $\delta(\mathcal{G}^n) := \max_{k=1,..,k_n} (c_k^n - c_{k-1}^n) \searrow 0$ as $n \to \infty$. For convenience, we assume that $p \in \mathcal{G}^n$ for $n \geq 1$ in the case that $\overline{c} > p$. We use the abbreviation

$$V^n(x, c) := V^{\mathcal{G}^n}(x, c);$$

we will prove in this section that $\lim_{n\to\infty} V^n(x, c) = V(x, c)$ and we will study the uniform convergence of this limit.

Remark 6.1 The value function of the one-step ratcheting problem considered by Albrecher et al. in [2], which increases the dividend payment from 0 to $\overline{c}$ only once and for all, corresponds to V^0 with $\mathcal{G}^0 = \{0, \overline{c}\}$ in our setting.

Since $V^n \leq V^{n+1} \leq V$, we can define

$$\overline{V}(x, c) = \lim_{n\to\infty} V^n(x, c). \tag{31}$$

Remark 6.2 From Proposition 5.1, we obtain immediately similar results for $\overline{V}$:

(1) $\overline{V}(x, c)$ is non-increasing in c with $\overline{V}(x, \overline{c}) = V(x, \overline{c})$, and non-decreasing in x with $\lim_{x\to\infty} \overline{V}(x, c) = \overline{c}/q$.

(2) There exists a constant $K_1 > 0$ such that

$$0 \leq \overline{V}(x_2, c_1) - \overline{V}(x_1, c_2) \leq K_1 \left[(x_2 - x_1) + (c_2 - c_1)\right]$$

for all $0 \leq x_1 \leq x_2$ and $0 \leq c_1 \leq c_2 \leq \min\{\overline{c}, p\}$.

(3) In the case $\overline{c} > p$, there exist constants $K_2 > 0$ and $K_3 > 0$ such that

$$0 \leq \overline{V}(x_2, c_1) - \overline{V}(x_1, c_2) \leq \left[K_2 + \frac{K_3}{c_1 - p}\right](x_2 - x_1) + \left[K_2 + \frac{K_3 x_2}{(c_1 - p)^2}\right](c_2 - c_1)$$

for all $0 \leq x_1 \leq x_2$, and $p < c_1 \leq c_2 \leq \overline{c}$.

We have the following results about uniform convergence in (31).

Proposition 6.1 *In the case $\overline{c} \leq p$, the sequence V^n converges uniformly to $\overline{V}$. In the case $\overline{c} > p$, the sequence V^n converges uniformly to $\overline{V}$ in $[0, \infty) \times [c_1, \overline{c}]$ for any $c_1 > p$.*

Proof. Let us show that V^n converges uniformly to $\overline{V}$ in the case $\overline{c} \leq p$, the case $\overline{c} > p$ is similar. Since $\overline{V}$ is non-decreasing in x and $\lim_{x\to\infty} \overline{V}(x, \overline{c}) = \overline{c}/q$, for any $\varepsilon > 0$ there exists x_0 such that

$$0 \leq \overline{c}/q - \overline{V}(x_0, \overline{c}) \leq \varepsilon/2.$$

Take n_0 large enough such that $\overline{V}(x_0, \overline{c}) - V^n(x_0, \overline{c}) < \varepsilon/2$ for all $n \geq n_0$ and so

$$0 \leq \overline{c}/q - V^n(x_0, \overline{c}) \leq \varepsilon.$$

Since $\overline{V}$ and V^n are non-decreasing in x and non-increasing in c, we have

$$0 \leq \overline{V}(x, c) - V^n(x, c) \leq \overline{c}/q - V^n(x_0, \overline{c}) \leq \varepsilon \tag{32}$$

for $(x, c) \in [x_0, \infty) \times [0, \overline{c}]$ and $n \geq n_0$. Consider the compact set $K = [0, x_0] \times [0, \overline{c}]$. For any point $(x_1, c_1) \in K$, we show first that there exists k large enough and $\eta > 0$ small enough such that if $\|(x, c) - (x_1, c_1)\| < \eta$ and $n \geq k$, then

$$\overline{V}(x, c) - V^n(x, c) < \varepsilon. \tag{33}$$

Indeed, by pointwise convergence at (x_1, c_1), there exists a k such that

$$\overline{V}(x_1, c_1) - V^n(x_1, c_1) < \varepsilon/3 \text{ for } n \geq k.$$

By Proposition 5.1 and Remark 6.2, there exists an $\eta > 0$ such that if $\|(x,c) - (x_1, c_1)\| < \eta$, then

$$|V^n(x,c) - V^n(x_1, c_1)| < \varepsilon/3 \text{ and } \left|\overline{V}(x,c) - \overline{V}(x_1, c_1)\right| < \varepsilon/3.$$

Therefore, we obtain (33). Taking a finite covering of the compact set K by balls of radius η we conclude that there exists an n_1 such that $\overline{V}(x,c) - V^n(x,c) < \varepsilon$ for any $(x,c) \in K$ and $n \geq n_1$. The result follows from (32). ■

Remark 6.3 In the case $\overline{c} > p$, we cannot prove the uniform convergence in the set $(0,\infty) \times [p, p+\varepsilon]$ with $\varepsilon > 0$ because of the lack of the Lipschitz condition of V^n, $\overline{V}$ and V at the points in the line $(0,\infty) \times \{p\}$. However, there is pointwise convergence.

Theorem 6.2 *The function $\overline{V}$ defined in (31) is the optimal value function V.*

Proof. Since $\overline{V}(x,c)$ is a limit of the value functions $V^n(x,c)$ of admissible strategies and $\lim_{x\to\infty} \overline{V}(x,\overline{c}) = \overline{c}/q$, by virtue of Theorem 4.8 it is enough to prove that $\overline{V}$ is a viscosity supersolution of (4) at any point (x_0, c_0) with $x_0 > 0$ and $c_0 \neq p$. Since $\overline{V}$ is non-increasing in c, $\overline{V}_c(x_0, c_0) \leq 0$ in the viscosity sense; so it is sufficient to show that $\mathcal{L}(\overline{V})(x_0, c_0) \leq 0$ in the viscosity sense. Take a test function φ for viscosity supersolution of (4) at (x_0, c_0), i.e. a continuously differentiable function φ with

$$\overline{V}(x,c) \geq \varphi(x,c) \text{ and } \overline{V}(x_0,c_0) = \varphi(x_0,c_0). \tag{34}$$

In order to prove that $\mathcal{L}(\varphi)(x_0, c_0) \leq 0$, consider now, for $\gamma > 0$ small enough,

$$\varphi_\gamma(x,c) = \varphi(x,c) - \gamma(x - x_0)^2.$$

Given $n > 0$, let us define

$$\begin{aligned} c_n &:= \min\{c \in \mathcal{G}^n : c \geq c_0\}, \\ a_n^\gamma &:= \min\{V^n(x,c_n) - \varphi_\gamma(x,c_n) : x \in [0, x_0+1]\}, \\ x_n^\gamma &:= \arg\min\{V^n(x,c_n) - \varphi_\gamma(x,c_n) : x \in [0, x_0+1]\}, \end{aligned}$$

and

$$b_n^\gamma := \max\{\overline{V}(x,c_n) - V^n(x,c_n) : x \in [0, x_0+1]\}.$$

We have that $c_n \searrow c_0$ and, from Proposition 6.1, $\lim_{n\to\infty} a_n^\gamma = 0$ and $\lim_{n\to\infty} b_n^\gamma = 0$. We also have that $\lim_{n\to\infty} x_n^\gamma = x_0$ because

$$\begin{aligned} 0 &= V^n(x_n^\gamma, c_n) - (\varphi_\gamma(x_n^\gamma, c_n) + a_n^\gamma) \\ &= \left(V^n(x_n^\gamma, c_n) - \overline{V}(x_n^\gamma, c_n)\right) + \left(\overline{V}(x_n^\gamma, c_n) - \varphi_\gamma(x_n^\gamma, c_n)\right) - a_n^\gamma \\ &\geq -b_n^\gamma + 0 - a_n^\gamma + \gamma(x_n^\gamma - x_0)^2 \end{aligned}$$

and then

$$(x_n^\gamma - x_0)^2 \leq \frac{a_n^\gamma + b_n^\gamma}{\gamma} \to 0 \text{ as } n \to \infty.$$

Note that $\overline{\varphi}^n(\cdot) = \varphi_\gamma(\cdot, c_n) + a_n^\gamma$ is a test function for viscosity supersolution of $V^n(\cdot, c_n)$ in equation (28) at the point x_n^γ because

$$\varphi_\gamma(x_n^\gamma, c_n) + a_n^\gamma = V^n(x_n^\gamma, c_n) \text{ and } \varphi_\gamma(x, c_n) + a_n^\gamma \leq V^n(x, c_n) \text{ for } x \in [0, x_0+1].$$

Hence, we obtain $\mathcal{L}_{c_n}(\overline{\varphi}^n)(x_n^\gamma) \leq 0$. Since $(x_n^\gamma, c_n) \to (x_0, c_0)$, $\overline{\varphi}^n(\cdot) = \varphi_\gamma(\cdot, c_n) + a_n^\gamma \to \varphi_\gamma(\cdot, c_0)$ as $n \to \infty$ and φ_γ is continuously differentiable, one gets

$$\mathcal{L}(\varphi_\gamma)(x_0, c_0) = \lim_{n\to\infty} \mathcal{L}_{c_n}(\overline{\varphi}^n)(x_n^\gamma) \leq 0.$$

Finally, as

$$\partial_x \varphi_\gamma(x_0, c_0) = \partial_x \varphi(x_0, c_0)$$

and $\varphi_\gamma \nearrow \varphi$ as $\gamma \searrow 0$, we obtain that $\mathcal{L}(\varphi)(x_0, c_0) \leq 0$ and the result follows. ■

7 Numerical illustrations

In this section we consider the sequence of sets

$$\mathcal{G}^n := \left\{ c_k^n = \frac{k}{2^n}\overline{c} \ : k = 0, ..., 2^n \right\} \tag{35}$$

with $k_n = 2^n + 1$ elements. We present some examples in which we approximate the optimal ratcheting value V by the (optimal) finite ratcheting function $\widehat{V} := V^8$ corresponding to the finite set $\mathcal{G} = \mathcal{G}^8$ with $2^8 + 1 = 257$ possible positive values for the dividend rate as defined in (35). In the examples, we will compare the value functions $V^0(x, 0)$, $\widehat{V}(x, 0)$ and $V^{NR}(x)$; clearly, by Remarks 2.2 and 6.1, we have that $V^0(x, 0) \leq \widehat{V}(x, 0) \leq V^{NR}(x)$ for all $x \geq 0$. It also follows from Proposition 5.1 that

$$\lim_{x\to\infty} V^0(x, 0) = \lim_{x\to\infty} \widehat{V}(x, 0) = \lim_{x\to\infty} V(x, 0) = \lim_{x\to\infty} V^{NR}(x) = \overline{c}/q.$$

Let us describe first how we obtain the optimal finite ratcheting function for any finite set $\mathcal{G}$. Since an optimal finite ratcheting strategy exists for any finite set $\mathcal{G}$ and it is associated to the optimal change region $\mathcal{D}^{\mathcal{G}}$ given in (30), we define, for any family of change sets $\mathcal{D} = (\mathcal{D}_k)_{k=1,...,N-1}$ with $\mathcal{D}_k$ closed in $[0, \infty)$ the associated finite ratcheting strategy as follows:

- Given $(x, c) \in [0, \infty) \times [0, \overline{c}]$, take $k_1 = \min\{k : \text{with } 1 \leq k \leq N,\ c_k \geq c \text{ and } x \notin \mathcal{D}_k\}$, and pay dividends at constant rate c_{k_1} up to the first time T_{k_1} that the controlled surplus process X_t hits $\mathcal{D}_{k_1}$.
- Take $k_2 = \min\{k : \text{with } k_1 < k \leq N \text{ and } X_{T_{k_1}} \notin \mathcal{D}_k\}$, and pay dividends at constant rate c_{k_2} up to the first time T_{k_2} that the controlled surplus process X_t hits $\mathcal{D}_{k_2}$, etc.

Denote the non-change sets as $\mathcal{U}_k := [0, \infty) \setminus \mathcal{D}_k$. For the value function $W^{\mathcal{D}}$ associated to the family $\mathcal{D} = (\mathcal{D}_k)_{k=1,...,N-1}$ we have the following:

- If $c_N = \overline{c}$, then $W^{\mathcal{D}}(\cdot, \overline{c})$ is the unique solution of $\mathcal{L}_{\overline{c}}(v) = 0$ with boundary condition $\lim_{x\to\infty} v(x) = \overline{c}/q$.
- If $0 \leq c_k < \overline{c}$, then $W^{\mathcal{D}}(\cdot, c_k) = W^{\mathcal{D}}(\cdot, c_{k+1})$ in $\mathcal{D}_k$.
- If $p < c_k < \overline{c}$ and U is a connected component of $\mathcal{U}_k$, then $W^{\mathcal{D}}(\cdot, c_k)$ is the unique solution of $\mathcal{L}_{c_k}(v) = 0$ in U with boundary condition either $v(0) = 0$ if $0 \in U$ or $v(u_0) = W^{\mathcal{D}}(u_0, c_{k+1})$ if $u_0 = \inf U \in \mathcal{D}_k$.
- If $p = c_k$ and U is a connected component of $\mathcal{U}_k$, then $W^{\mathcal{D}}(\cdot, p)$ is the unique solution of $\mathcal{L}_p(v) = 0$ in U.
- If $0 \leq c_k < p$ and U is a connected component of $\mathcal{U}_k$, then $W^{\mathcal{D}}(\cdot, c_k)$ is the unique solution of $\mathcal{L}_{c_k}(v) = 0$ in U with boundary condition $v(u_0) = W^{\mathcal{D}}(u_0, c_{k+1})$ where $u_0 = \sup U \in \mathcal{D}_k$.
  Note that $W^{\mathcal{D}}(\cdot, c_k)$ only depends on the sets $D_k, ..., D_{N-1}$.

In order to find the optimal finite ratcheting strategy, we assume that the optimal change sets $\mathcal{D}_k^*$ have finitely many connected components and so $\mathcal{U}_k^*$ are bounded and have finitely many connected components. We construct the optimal finite ratcheting function $V^{\mathcal{G}}(\cdot, c_k)$ defined in (25) and the optimal change sets $\mathcal{D}_k^*$ for $k \leq N$ by going backward recursively as follows: $V^{\mathcal{G}}(\cdot, c_N)$ is the unique solution of $\mathcal{L}_{\overline{c}}(v) = 0$ with boundary condition $\lim_{\ x\to\infty} v(x) = \overline{c}/q$. Having constructed $V^{\mathcal{G}}(\cdot, c_{k+1})$ and $\mathcal{D}_{k+1}^*, .., \mathcal{D}_{N-1}^*$, we consider first the case in which the sets $\mathcal{D}_k$ and $\mathcal{U}_k$ have one connected component, i.e. $\mathcal{D}_k = [d_k, \infty)$, and maximize $W^{\mathcal{D}}(\cdot, c_k)$ with the one parameter d_k. If there exists a maximized value function $W^{\mathcal{D}}(\cdot, c_k)$ and it is a viscosity supersolution of (28) then, by Remark 5.1, $V^{\mathcal{G}}(\cdot, c_k) = W^{\mathcal{D}}(\cdot, c_k)$ and $\mathcal{D}_k^* = [d_k^*, \infty)$ where d_k^* is the value for which the maximum is attained. If this is not the case, we consider change sets $\mathcal{D}_k$ with two connected components and non-change sets $\mathcal{U}_k$ with one, that is $\mathcal{D}_k = [0, d_k^1] \cup [d_k^2, \infty)$. We then maximize $W^{\mathcal{D}}(\cdot, c_k)$ with two parameters $d_k^1 < d_k^2$, ; if there exists a maximized value function $W^{\mathcal{D}}(\cdot, c_k)$ and it is a viscosity supersolution of (28), $V^{\mathcal{G}}(\cdot, c_k) = W^{\mathcal{D}}(\cdot, c_k)$ and $\mathcal{D}_k = [0, d_k^{1*}] \cup [d_k^{2*}, \infty)$, with $d_k^{1*} < d_k^{2*}$ being the values for which this maximum is attained. If this not the case we proceed to three parameters, and so on.

We first consider two examples with exponentially distributed claim sizes: one for $\overline{c} < p$ and one with $\overline{c} > p$.

Example 7.1 We take $\beta = 4$, $p = 2.3$, $F(x) = 1 - e^{-2x}$, $q = 0.1$ and $\overline{c} = 1.72 = (3/4)\,p$. Figure 1 shows the approximation $\widehat{V}(x,0)$ of the optimal value function $V(x,0)$ as a function of x. Figure 2 depicts in gray the change region $\mathcal{D}^{\mathcal{G}^8} = \bigcup_{k=1}^{2^8-1} \mathcal{D}_k^* \times \{\frac{k}{2^8}\overline{c}\} \subset [0,\infty) \times \mathcal{G}^8$ (note that all the sets $\mathcal{D}_k^*$ are of the form $[d_k^*, \infty)$ and satisfy $\mathcal{D}_{k+1}^* \subset \mathcal{D}_k^*$).

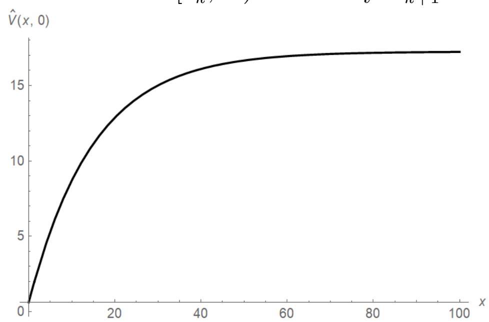

Figure 1: Value function under ratcheting as a function of initial capital x

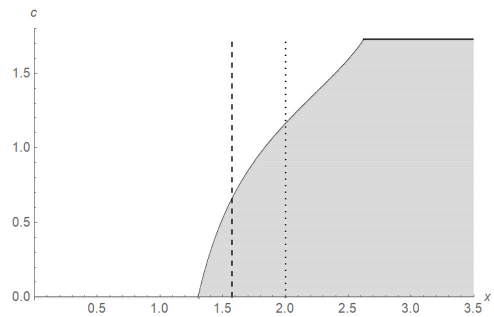

Figure 2: Change region (in gray) for the optimal ratcheting strategy

In Figure 3, we compare $\widehat{V}(x,0)$ with the value function $V^0(x,0)$ of only one possible switch during the lifetime of the process. We find that the optimal strategy for V^0 is given by the set $\mathcal{D}^{\mathcal{G}^0} = \mathcal{D}_1^* \times \{0\} \subset [0,\infty) \times \mathcal{G}^0$ with $\mathcal{D}_1^* = [2.00, \infty)$, i.e. to switch to the maximal dividend rate 1.72 as soon as the surplus reaches the value 2 (cf. the dotted line in Figure 2). One sees that the difference of the value functions is surprisingly small relative to their absolute values. That is, the improvement from being allowed to only raise your dividend rate once rather than ratcheting continuously is quite minor. Note also the spread of the optimal boundary to 'enter' a dividend rate level around the one-time jump level at $x = 2$. In order to see how the number of possible switches changes the value function, Table 1 gives the maximum error (w.r.t. the approximation $V^8(x,0)$) across the considered x-range for the respective refinements $V^n(x,0)$ as n increases from 0 to 7. Depending on the desired accuracy one can then decide which value of n is advised. In particular, $n = 8$ (e.g. 256 possible switches) already seems to be a very satisfactory approximation of $V(x,0)$, given that the improvement over the case with 128 possible switches ($n = 7$) is already quite minor.
On the other hand, the efficiency loss of the expected discounted dividend payments by ratcheting compared to the general un-constrained dividend problem is relatively larger (yet still surprisingly small in relative terms). Figure 4 depicts the corresponding difference as a function of initial capital x. Note that the optimal strategy in the un-constrained problem is a threshold strategy with threshold $x_{NR} = 1.57$, i.e. whenever the surplus is above $x_{NR} = 1.57$, dividends are paid at the maximal rate 1.72 and no dividends are paid when the surplus is below that level (see e.g. [14]). For illustration, x_{NR} is plotted as a dashed line in Figure 2.

n	0	1	2	3	4	5	6	7
	$8.05{\cdot}10^{-3}$	$2.73{\cdot}10^{-3}$	$4.96{\cdot}10^{-4}$	$1.17{\cdot}10^{-4}$	$2.79{\cdot}10^{-4}$	$6.72{\cdot}10^{-6}$	$1.62{\cdot}10^{-6}$	$3.23{\cdot}10^{-7}$

Table 1: $\max_x\{V^8(x,0) - V^n(x,0)\}$ for various values of n

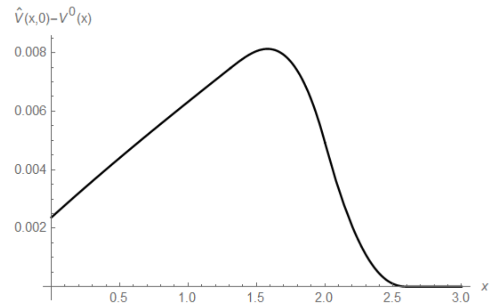

Figure 3: Improvement from switching once to general ratcheting as a function of x

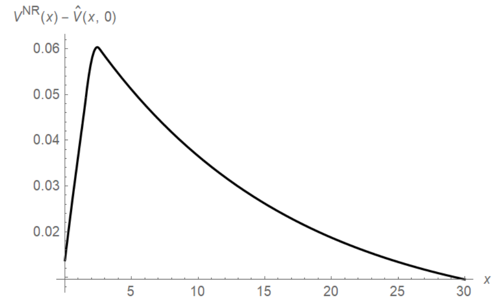

Figure 4: Improvement from the optimal ratcheting strategy to the un-constrained optimal dividend strategy as a function of x

Example 7.2 The geometry of the change region can become more complex, if the maximal dividend rate is allowed to exceed the premium rate p. As an illustration, consider $\beta = 4$, $p = 2.3$, $F(x) = 1 - e^{-2x}$, $q = 0.1$ and $\bar{c} = 4.6 = 2p$. Figure 5 gives the approximation $\widehat{V}(x,0)$ of the optimal value function $V(x,0)$ in this case, and the change region $\mathcal{D}^{\mathcal{G}^8}$ is plotted (in gray) in Figure 6 (again, all the sets $\mathcal{D}_k^*$ are of the form $[d_k^*,\infty)$ and satisfy $\mathcal{D}_{k+1}^* \subset \mathcal{D}_k^*$).

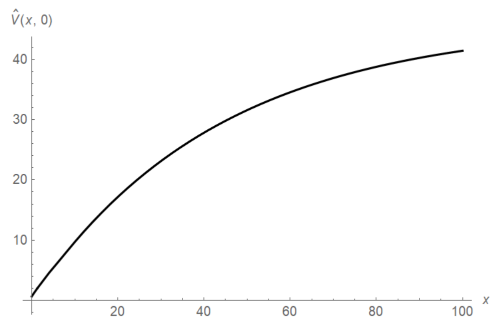

Figure 5: Value function under ratcheting as a function of initial capital x

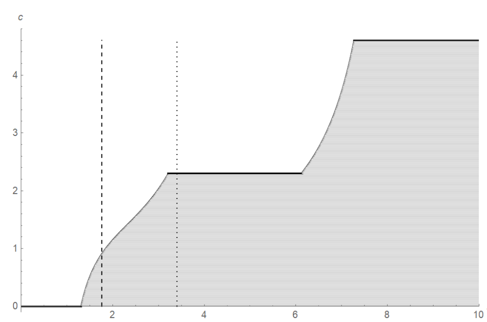

Figure 6: Change region (in gray) for the optimal ratcheting strategy

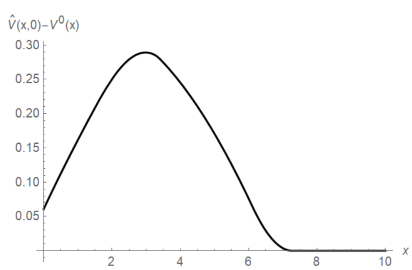

Figure 7: Improvement from switching once to general ratcheting as a function of x

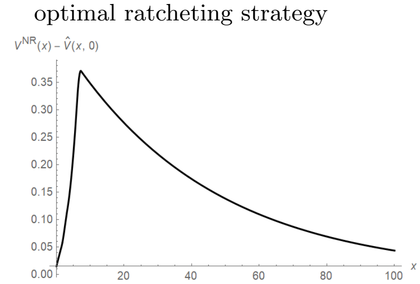

Figure 8: Improvement from the optimal ratcheting strategy to the un-constrained optimal dividend strategy as a function of x

n	0	1	2	3	4	5	6	7
	$2.89{\cdot}10^{-1}$	$2.22{\cdot}10^{-2}$	$7.51{\cdot}10^{-3}$	$1.18{\cdot}10^{-3}$	$7.61{\cdot}10^{-4}$	$5.38{\cdot}10^{-5}$	$1.44{\cdot}10^{-5}$	$2.55{\cdot}10^{-6}$

Table 2: $\max_x\{V^8(x,0) - V^n(x,0)\}$ for various values of n

One observes that now we have a region of the free boundary where the dividend rate c is not increased for growing x. Concretely, it takes a much larger surplus value until it is optimal to start making use of the possibility to pay out dividends at a higher rate than the one of incoming premiums. Figure 7 again shows the efficiency gain from the optimal one-switch strategy (which in this case is characterized by the set $\mathcal{D}^{\mathcal{G}^0} = \mathcal{D}_1^* \times \{0\} \subset [0,\infty) \times \mathcal{G}^0$ with $\mathcal{D}_1^* = [3.40,\infty)$) to the general ratcheting strategy, and Table 2 shows again the quality improvement of the approximation when allowing more and more switches. Figure 8, on the other hand, depicts the difference of the value functions of the optimal dividend problem without and with ratcheting. In this case, the optimal threshold for the non-constrained case is $x_{NR} = 1.76$; since $\bar{c} > p$, this means that the optimal strategy is to pay no dividends if the current surplus is less than x_{NR}, to pay dividends at maximum possible rate $\bar{c}$ if the current surplus is greater than x_{NR} and to pay dividends at rate p if the current surplus coincides with x_{NR}.

Remark 7.1 The numerical approximation suggests that, in the examples above, the optimal ratcheting strategy is given by the optimal free boundary

$$G = \{(x, C^*(x)) : x \geq 0\},$$

which separates the non-change region with the change region; here C^* is a non-decreasing function with $C^*(x) = 0$ for x small and $C^*(x) = \bar{c}$ for x large enough. More precisely,

- If the initial values are (x, c) with $c < C^*(x)$, the optimal strategy is to pay dividends at rate $C^*(x)$.
- If the initial values are (x, c) with $c > C^*(x)$, the optimal strategy is to pay dividends at rate c until the controlled trajectory $(X_t^C - ct, c)$ in the state space reaches the free boundary G.
- If the initial values (x, c) are on the free boundary G with $c < p$, the optimal strategy is to pay dividends at rate $C^*(x)$. In this case the trajectory in the state space $\left(X_t^C, C^*(X_t^C)\right)$ remains on the free boundary G until either ruin occurs or a next claim arrives.
- If the initial values (x, c) are on the free boundary G with $c = p$, the optimal strategy is to pay dividends at rate p. In this case the trajectory in the state space $\left(X_s^C, C^*(X_s^C)\right)$ is constant (x, p) until either ruin occurs or a next claim arrives.
- If the initial values (x, c) are on the free boundary G with $c > p$, the optimal strategy is to pay dividends at rate c. In this case the trajectory in the state space falls immediately into the non-change region.

Example 7.3 As a final example, we would like to see how the ratcheting constraint changes the un-constrained dividend problem in the case where a band strategy is optimal for the latter. To that end, choose $\beta = 10$, $p = 21.4$, $F(x) = 1 - e^{-x}(1 + x)$, $q = 0.1$ and $\bar{c} = 16.5 = (3/4)\,p$ (cf. [**?**, Sec.10.1]). Figure 9 depicts the approximation $\widehat{V}(x,0)$ of the optimal value function $V(x,0)$ for this choice of parameters, while Figure 10 shows in gray the change region $\mathcal{D}^{\mathcal{G}^8}$, and Figures 11 and 12 compare $\widehat{V}(x,0)$ again with the one-switch value function $V^0(x,0)$ and the un-constrained value function $V^{NR}(x,0)$, respectively. Table 3 depicts the approximation improvement when increasing the number of possible switches.
From Figure 10 one sees that in this case the sets $\mathcal{D}_k^* = [0, d_k^{1*}] \cup [d_k^{2*}, \infty)$ have two components with $d_k^{1*} < d_{k+1}^{1*}$ and $d_k^{2*} < d_{k+1}^{2*}$ and so $\mathcal{D}_{k+1}^* \not\subseteq \mathcal{D}_k^*$. In Figure 13, we show the derivative of $\widehat{V}(x,0)$ for small values of x. Note that the function $\widehat{V}(x,0)$ is continuous but not differentiable at the point

n	0	1	2	3	4	5	6	7
	$5.65{\cdot}10^{-2}$	$9.45{\cdot}10^{-3}$	$2.45{\cdot}10^{-3}$	$5.97{\cdot}10^{-4}$	$1.47{\cdot}10^{-4}$	$3.63{\cdot}10^{-5}$	$8.67{\cdot}10^{-6}$	$1.72{\cdot}10^{-6}$

Table 3: $\max_x\{V^8(x,0) - V^n(x,0)\}$ for various values of n

$d_0^{1*} = 0.076$ because the optimal strategy consists of paying dividends at maximum rate $\bar{c}$ if the current surplus is less than or equal to d_0^{1*}, but paying no dividends if the current surplus is slightly above d_0^{1*}. Only for much larger values of the initial capital x it is again optimal to pay dividends.

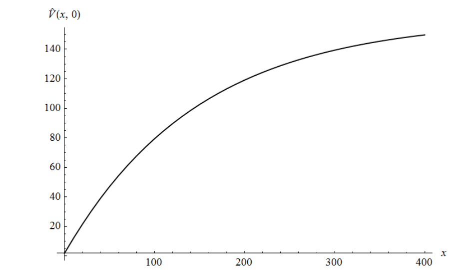

Figure 9: Value function under ratcheting as a function of initial capital x

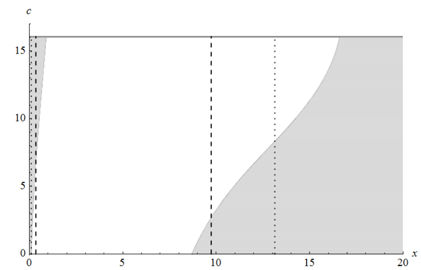

Figure 10: Change region (in gray) for the optimal ratcheting strategy

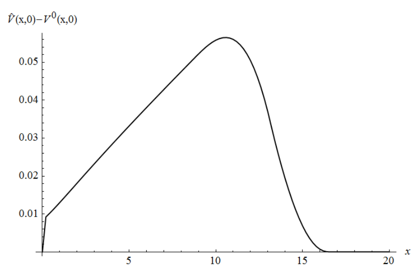

Figure 11: Improvement from switching once to general ratcheting as a function of x

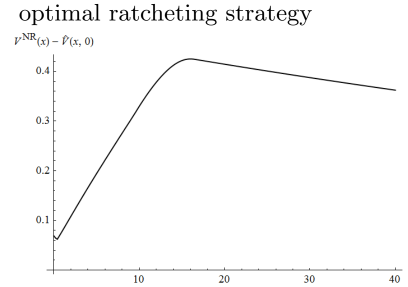

Figure 12: Improvement from the optimal ratcheting strategy to the un-constrained optimal dividend strategy as a function of x

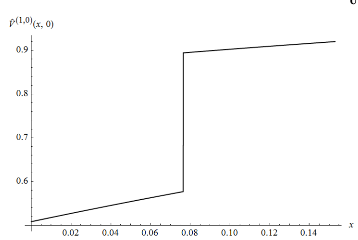

Figure 13: The derivative of $\widehat{V}(x,0)$ with respect to initial capital x

This is exactly the extension of the band strategy of the un-constrained case (which consists of paying dividends at maximum possible rate $\bar{c}$ if the current surplus is in the set $[0, 0.35] \cup [9.73, \infty)$, cf. the dashed lines in Figure 10, and paying no dividends elsewhere – so not a threshold strategy anymore) to the ratcheting case. The optimal strategy for V^0 (i.e. when only one switch of dividend rates is possible) is given by the set $\mathcal{D}^{\mathcal{G}^0} = \mathcal{D}_1^* \times \{0\} \subset [0,\infty) \times \mathcal{G}^0$, where $\mathcal{D}_1^* = [0, 0.10] \cup [13.13, \infty)$ now has two connected components (cf. dotted lines in Figure 10).

Remark 7.2 In Example 7.3 the optimal ratcheting strategy is given by two optimal free boundaries

$$G_1 = \{(x, C_1^*(x)) : x \geq 0\} \text{ and } G_2 = \{(x, C_2^*(x)) : x \geq 0\}$$

which bound the non-change region on the left and on the right respectively; here C_1^* and C_2^* are non-decreasing functions with $C_2^* \leq C_1^*$ and $C_i^*(x) = 0$ for x small and $C_i^*(x) = \overline{c}$ for x large enough for $i = 1, 2$. More precisely,

- If the initial values are (x, c) with $c \geq C_1^*(x)$, the optimal strategy is to pay dividends at maximum rate $\overline{c}$.
- If the initial values are (x, c) with $c < C_2^*(x)$, the optimal strategy is to pay dividends at rate $C_2^*(x)$.
- If the initial values (x, c) are on the free boundary G_2 but not in G_1, the optimal strategy is to pay dividends at rate $C_2^*(x)$; in this case the trajectory in the state space $\left(X_t^C, C_2^*(X_t^C)\right)$ remains on the free boundary G_2 until either ruin occurs or the next claim arrives.
- If the initial values (x, c) are in the non-change region (i.e. $C_2^*(x) < c < C_1^*(x)$), then the optimal strategy is to pay dividends at rate c until the controlled trajectory $(X_t^C - ct, c)$ in the state space exits the non-change region.

8 Conclusion

In this paper we solved the general problem of identifying optimal dividend strategies in an insurance risk model under the additional constraint that the dividend rate needs to be non-decreasing over time. We showed that the value function is the unique viscosity solution of a two-dimensional Hamilton-Jacobi-Bellman equation and can be approximated arbitrarily closely by optimal strategies for finitely many possible dividend rates. The analysis is considerably more complicated when the maximal dividend rate is allowed to exceed the incoming premium rate. We derived the free boundaries and optimal strategies numerically for a number of concrete cases with exponential and Gamma claim sizes, and the results illustrate that the value function with ratcheting is not much lower than the one without the ratcheting constraint. Also, a comparison shows that the previously studied one-switch strategy performs remarkably well, i.e. the further improvement in the case of general ratcheting is typically not substantial. We also showed that for parameter settings where a band strategy is optimal in the non-constrained case, the band-type structure remains optimal for the ratcheting solution, then with two free boundaries in the domain of initial surplus and initial dividend rate.

There are several interesting directions of future research. The numerical approximations implemented in Section 7 were based on an equi-distant grid of the k_n possible dividend rates, which was not restrictive for the purpose of the paper, as k_n was large in the illustrations. If, however, k_n is small, one may be able to achieve a higher value function by assigning the dividend rates in a different than equidistant way. Note in this context that Example 3.2 of [2] illustrated numerically for a diffusion setup that for two possible levels, the optimal strategy is to pay out nothing until the switching barrier is hit, and then pay the maximal amount (exactly corresponding to V^0 used in Section 7). It can be interesting to study the optimal dividend rate increases in the presence of a small number k_n of switching levels. Linked to the latter, one could try to identify the optimal ratcheting strategy when dividend increases are only possible at discrete random (e.g. Poissonian) times. Furthermore, a deeper investigation as to why V^n is such a good approximation of the optimal V for small values of n already would be relevant. Finally, further directions include considering the ratcheting problem for more general surplus processes and variants where the dividend rate may be decreased by a certain percentage of its current value (known as drawdown strategies in finance, cf. [4]) or interrupted (dropped down to 0) for an otherwise increasing ratcheting structure.

References

[1] Albrecher, H., Azcue, P. and Muler N. (2017). Optimal dividend strategies for two collaborating insurance companies. *Advances in Applied Probability* **49**, No.2, 515–548.

[2] Albrecher H., Bäuerle N. and Bladt M. (2018). Dividends: From refracting to ratcheting. *Insurance Math. Econom.* **83,** 47–58.

[3] Albrecher, H. and Thonhauser, S. (2009). Optimality results for dividend problems in insurance. *RACSAM-Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas* **103**, No.2, 295–320.

[4] Angoshtari, B., Bayraktar, E. and Young, V.R. (2019) Optimal dividend distribution under drawdown and ratcheting constraints on dividend rates. *SIAM Journal on Financial Mathematics* **10**, No.2, 547–577.

[5] Azcue, P. and Muler, N. (2012). Optimal dividend policies for compound Poisson processes: The case of bounded dividend rates. *Insurance Math. Econom.* **51,** 26–42.

[6] Azcue, P. and Muler, N. (2014). *Stochastic Optimization in Insurance: a Dynamic Programming Approach.* Springer Briefs in Quantitative Finance. Springer.

[7] Azcue, P. and Muler, N. (2015). Optimal dividend payment and regime switching in a compound Poisson risk model. *SIAM J. Control Optim.* **53**(5), 3270-3298.

[8] Azcue, P., Muler, N. and Z. Palmowski (2019). Optimal dividend payments for a two-dimensional insurance risk process. *Eur. Actuar. J.* **9**(15), 241-272.

[9] Avanzi, B., Tu, V., and Wong, B. (2016). A note on realistic dividends in actuarial surplus models. *Risks* **4**, 4, 37, 1–9.

[10] Benth, F.E., Karlsen, K. H. and Reikvam, K. (2002). Portfolio optimization in a Lévy market with intertemporal substitution and transaction costs. *Stochastics Stochastics Rep.* **74,** 517–569.

[11] De Finetti, B. (1957). Su un'Impostazione Alternativa della Teoria Collettiva del Rischio. *Transactions of the 15th Int. Congress of Actuaries* **2**, 433–443.

[12] Dybvig, P.H. (1995). Dusenberry's ratcheting of consumption: optimal dynamic consumption and investment given intolerance for any decline in standard of living. *The Review of Economic Studies*, **62**, No.2, 287–313.

[13] Gerber, H.U. (1969). Entscheidungskriterien fuer den zusammengesetzten Poisson-Prozess. *Schweiz. Aktuarver. Mitt.* (1969), No.1, 185–227.

[14] Gerber, H.U. and Shiu, E.S.W. (2006). On optimal dividend strategies in the compound Poisson model. *N. Am. Actuar. J.* **10**, 2, 76–93.

[15] Gu, J., Steffensen, M. and Zheng, H.(2017). Optimal dividend strategies of two collaborating businesses in the diffusion approximation model. *Mathematics of Operations Research* **43**, No.2, 377–398.

[16] Protter, P. (1992). *Stochastic integration and differential equations.* Berlin: Springer Verlag.

[17] Sayah, A. (1991). Équations d'Hamilton-Jacobi du premier ordre avec termes intégro différentiels.I. Unicité des solutions de viscosité. *Comm. Partial Differential Equations* **16,** 1057–1074.

[18] Schmidli, H. (2008). *Stochastic Control in Insurance.* Springer, New York.