Discrete dividend payments in continuous time

Jussi Keppo* Max Reppen†§ H. Mete Soner‡§

May 14, 2018

Abstract

We propose a model in which dividend payments occur at regular intervals in an otherwise continuous model. This contrasts traditional models where either the payment of continuous dividends is controlled or the dynamics are given by discrete time processes. Moreover, between two dividend payments, the structure allows for other types of control; we consider the possibility of equity issuance at any point in time. We prove the convergence of an efficient numerical algorithm that we use to study the problem. The model enables us to find the loss caused by infrequent dividend payments. We show that under realistic parameter values this loss varies from around 1% to 24% depending on the state of the system, and that using the optimal policy from the continuous problem further increases the loss.

1 Introduction

Continuous time decision making is prevalent and of great importance, but some phenomena really only occur at discrete intervals. In models for asset trading, these intervals are typically sufficiently short to justify a continuous time model. However, other types of events take place on larger time scales and thus weakens the basis for a continuous time approximation. One such example is dividend payments that we study in this paper. Variation of the dividend frequency typically ranges from monthly to annually, far less frequently than, for instance, trading on a large exchange.

Our approach for tackling this discrepancy between the models and practise is to consider a continuous time model in which dividends may only be paid out at predetermined, discrete time points. The distinction between such a model and a

*NUS Business School and Institute of Operations Research and Analytics, National University of Singapore, email: keppo@nus.edu.sg. Partly supported by Institute of Operations Research and Analytics (National University of Singapore) grant WBS-R-726-000-009-646.

†ETH Zürich, Department of Mathematics, Rämistrasse 101, 8092 Zürich, Switzerland, email: max.reppen@math.ethz.ch.

‡ETH Zürich, Department of Mathematics, Rämistrasse 101, 8092 Zürich, Switzerland, and Swiss Finance Institute, email: mete.soner@math.ethz.ch.

§Partly supported by the ETH Foundation, the Swiss Finance Institute, and Swiss National Foundation grant SNF 200020_172815.

traditional discrete time model lies in the possibility to model other continuous decision-making problems in-between the dividend payments. In particular, we allow for the possibility to issue equity. Since we do not wish to restrict equity issuance to predetermined time points—as would be the case in a discrete time model—we model this issuance as a continuous time control problem. In other words, equity can be issued at any point in time. Although our choice of continuous time control is equity issuance, the same methodology can also be used for other types of decisions and models, for instance capital investments. Notwithstanding, the model at hand is not only a good example for the method, but our numerical results facilitates an interesting comparison to the continuous time counterpart. More specifically, we show that under realistic parameter values the difference between the value function of discrete time dividends and the corresponding value function with continuous time dividends is often around 1–3% relative to the continuous dividend case, but increases to 24% depending on the state of the system.

Moreover, this type of structure does not only appear in problems with control decisions at discrete times. In fact, problems in which monitoring occurs at discrete time points readily fit into the same framework. One example of such a model is that of leveraged exchange traded funds (leveraged ETFs or LETFs). The goal of an LETF is to track the returns of some index on some time scale—typically daily—by a predetermined multiple. In [9], the authors model this tracking problem by imposing a 'monitoring condition' at the end of each trading day; at the end of the day, it incurs a penalty depending how closely it tracks the underlying index. From a structural point of view, the trading and transaction cost payment, happening in continuous time, is akin to the equity issuance, whereas the 'monitoring' takes the role of the dividend payment (despite not being actively controlled).

Our main focus is to characterize the value function as the solution to a parabolic PDE with a fixed point structure. To numerically find a solution, one needs to iteratively solve a related control problem without the discrete time element, i.e., without the dividends/monitoring. In our model below, we do not otherwise make any specific assumptions on the cash flow process other than that the cash flow process cannot be too large. In particular, the results hold for both the commonly studied jump models, cf. [12], as well as their diffusion model counterparts, cf. e.g. [11, 13].

In the context of dividend problems specifically, one common point of criticism of many optimal dividend problems is their irregular dividend payments when following the optimal policy. One way to alleviate this is to consider dividend policies that are proportional or affine as a function of the current reserves, cf. [2, 5]. Although we do not explicitly consider any such models, they still fit naturally into the framework presented in this paper. For further references on optimal dividend problems, we refer the reader to [3, 4].

The structure of the paper is given as follows: For the purpose of showcasing the main idea, we begin with a nonrigorous description of the general structure in Section 2. We thereafter give an account of how an optimal dividend problem with continuous issuance and discrete dividends fit into this framework in Section 3. In this context we present the structure of the main equation and describe a numerical scheme for finding its solution. In Section 4, we extend the optimal dividend model and the convergence results to a multidimensional model in the spirit of [15]. Numerical studies on the value and policy impact of dividend discretization is conducted in both settings. Finally, Section 5 provides a summary of our findings along with our interpretations and conclusions.

2 General structure

Although the focus of the paper is on optimal dividend problems, the core idea extends to a wider class of problems and is best showcased in a general setting. What follows in this section is a formal discussion of the ideas that will later be made rigorous for two dividend problems.

We consider a specific type of infinite horizon (possibly singular) stochastic optimal control problem with discounting. What distinguishes these problems is that they, at regular, equidistant intervals, involve a singular action and/or monitoring or a singular jump in the payoff function. In this sense, the structure can be considered as a mix of continuous and discrete time control problems.

The structure of the problem can thus be separated into two components: one for what happens at the discrete time points and one for what happens in-between. For simplicity, we will represent these components by incremental operators that represent the equations determining the solution across these time regions.

In particular, we allow for two controls α and β that represent the continuous control and the discrete control respectively. For a given choice of α and β, we denote by $X^{\alpha,\beta} = (X_t^{\alpha,\beta})_{t\geq 0}$ the state process corresponding to this choice. It is here implicit that the process does not depend on β in-between the discrete time point. Similarly, we consider two types of cost structures: $F_t^\alpha = F(\alpha_t, X_t^{\alpha,\beta})$ for the cumulative (undiscounted) continuous cost and $G_t^\beta = G(\beta_t, X_{t-}^{\alpha,\beta})$ for the cost incurred at the discrete points. Note that α may be a singular control process. Finally, let T be the time between two of the discrete time points.

With this structure and with $T\mathbb{N} = \{0, T, 2T, \dots\}$, we can write the control problem as

$$V(x) = \sup_{\alpha,\beta} \mathbb{E}_x \left[\int_0^\infty e^{-\rho t}\, \mathrm{d}F_t^\alpha + \sum_{n\in T\mathbb{N}} e^{-\rho n} G_n^\beta \right],$$

where $\mathbb{E}_x$ denotes expectation with respect to a measure under which the state process starts at x (before any control is activated), the supremum is over some set of admissible controls, and ρ is the discounting rate. To proceed, we require the *discrete time dynamic programming principle* (DTDPP) to hold, i.e., dynamic programming at the discrete time points $t \in T\mathbb{N}$. This and that the value function is universally measurable are established by Bertsekas and Shreve [8].

Now suppose there exists a space $\mathcal{X}$ such that the composition of the following two 'incremental' operators are well-defined. First, the continuous operator is given by

$$\mathcal{L}\phi(x) = \sup_{\alpha} \mathbb{E}_x \left[\int_0^T e^{-\rho t}\, \mathrm{d}F_t^\alpha + e^{-\rho T} \phi(X_{T-}^{\alpha,0}) \right].$$

Second, the discrete operator is given by

$$\mathcal{D}\phi(x) = \sup_{\beta} \Big(\phi(x+\beta) + G(\beta, x) \Big).$$

In other words, there exists a space $\mathcal{X}$ such that $\mathcal{D} \circ \mathcal{L} : \mathcal{X} \to \mathcal{X}$ or $\mathcal{L} \circ \mathcal{D} : \mathcal{X} \to \mathcal{X}$.[1] Without loss of generality, assume that it holds for $\mathcal{T} := \mathcal{D} \circ \mathcal{L}$. Note that the universal measurability of the value function V makes the operators well-defined on V, and the DTDPP states precisely that $V = \mathcal{T}V$.

Our goal is to show that the value function can be found by iteratively applying $\mathcal{T}$. To do so, we seek a complete metric space $(\mathcal{X}, d)$ such that $\mathcal{T}$ is a strict contraction and $V \in \mathcal{X}$. If such a space exists, $\mathcal{T}$ will have a (unique) fixed point, provided the space is not empty. Then, by the DTDPP, the value function is the fixed point, and $\lim_{n\to\infty}(\mathcal{D} \circ \mathcal{L})^n \phi = V$ for every $\phi \in \mathcal{X}$.[2]

In the rest of this paper, we will show how $\mathcal{X}$ and d can be chosen for the two optimal dividend problems.

3 Discrete dividend payments with capital injections

This section is devoted to the optimal dividend problem for which the fixed point idea from Section 2 can be applied. An equity capital constrained firm pays dividends to its shareholders at discrete, predetermined time intervals. The firm may also choose to issue equity at any point in time.

More precisely, the firm is endowed with some cash flow $C = (C_t)_{t\geq 0}$ that are placed in the firm's *cash reserves*. Dividends may be paid from these reserves at any point until the time of ruin/bankruptcy, θ. Let x denote the initial capital, $L = (L_t)_{t\geq 0}$ the cumulative dividends, and $I = (I_t)_{t\geq 0}$ the cumulative equity issued. Then the net cash reserves $X = (X_t)_{t\geq 0}$ are given by[3]

$$\mathrm{d}X_t^{L,I} = \mathrm{d}C_t - \mathrm{d}L_t + \mathrm{d}I_t, \quad X_0^{L,I} = x.$$

Let $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$ be the filtration generated by C. We restrict dividends to be fully covered by reserves of the firm, i.e., $\Delta L_t \leq X_{t-}^{L,I}$. The ruin time is then given by

$$\theta(L, I) = \inf\{t > 0 : X_t^{L,I} < 0\}.$$

The aim of the firm is to maximize the discounted value of dividends, net of capital injections. In modelling the equity issuance, we follow [10]. This type of model has since also been studied in [1, 15]. With a discounting rate of $\rho > 0$ as well as fixed and proportional issuance costs $\lambda_f \geq 0$ and $\lambda_p \geq 0$ respectively, the value of the firm is then

$$V(x) = \sup_{L,I} \mathbb{E}_x \left[\sum_{t\in T\mathbb{N}\cap[0,\theta(L,I)]} e^{-\rho t}\Delta L_t - \sum_{t\geq 0} e^{-\rho t}(\lambda_f + (1+\lambda_p)\Delta I_t) 1_{\{\Delta I_t > 0\}} \right],$$

[1] Note that the first operator in these compositions may map to an intermediate space.
[2] In fact, this is nothing else than so-called value iteration [7].
[3] We impose the condition that all three processes are RCLL.

where the optimization is over increasing, RCLL, adapted processes and $\mathbb{E}_x$ denotes expectation under a measure such that $X^{L,I}_{0-} = x$.

Assumption 3.1. For the cash flow process $C = (C_t)_{t\geq 0}$, it holds that $\mathbb{E}_x[(C_T)^+] < \infty$.

3.1 Periodization

To put this problem into a periodic structure, we define the function

$$v(t,x) \;=\; \sup_{L,I} \mathbb{E}_x\left[\sum_{s\in T\mathbb{N}\cap[t,\theta(L,I)]} e^{-\rho(s-t)}\Delta L_t - \sum_{s\geq t} e^{-\rho(s-t)}(\lambda_f + (1+\lambda_p)\Delta I_s)1_{\{\Delta I_s>0\}}\right],$$

for $t \geq 0$. Note that since dividends are paid at $t \in T\mathbb{N}$, $v(\cdot,x)$ is LCRL for all x, in contrast to the stochastic processes introduced in the beginning of this section. In particular, $v(nT,x) = v(nT-,x)$ for $n \in \mathbb{N}$. Like in Section 2, for $t \in (0,T]$,

$$v(t,x) = \sup_{L,I} \mathbb{E}_x\Bigg[-\sum_{T>s\geq t} e^{-\rho(s-t)}(\lambda_f + (1+\lambda_p)\Delta I_s)1_{\{\Delta I_s>0\}} \\ + e^{-\rho(T-t)} v(T-, X^{L,I}_{T-})1_{\{T<\theta(L,I)\}}\Bigg],$$

which suggests that, for $t \in (0,T)$, v satisfies

$$\min\left\{-(\partial_t + \mathcal{A} - \rho)v(t,x), v(t,x) - \sup_{i\geq 0}\Big(v(t,x+i) - (1+\lambda_p)i - \lambda_f\Big)\right\} = 0, \qquad (1)$$

where $\mathcal{A}$ is the generator of $C = (C_t)_{t\geq 0}$, with boundary condition $v(t,0) = 0$ in the viscosity sense, i.e.,

$$v(t,0) = \max\left\{0, (\partial_t + \mathcal{A} + 1 - \rho)v(t,0+), \sup_{i\geq 0}\Big(v(t,i) - (1+\lambda_p)i - \lambda_f\Big)\right\}. \qquad (2)$$

At the time of a dividend payout ΔL, it holds that

$$v(T,x) \geq v(T+, x - \Delta L) + \Delta L.$$

Since $\theta(L,I)$ does not depend on the starting time, $v(0+,x) = v(T+,x)$ and $v(0,x) = v(T,x)$. By exploiting this recurrence structure, we may consider only one 'period': Optimizing over dividend policies, we obtain the periodic condition

$$v(T,x) = \sup_{\ell\leq x}\Big(v(x-\ell, 0+) + \ell\Big). \qquad (3)$$

This is the periodic initial-terminal condition for (1).

To summarize, v—thus also V—is characterized by (1) and the boundary conditions (2) and (3). The periodic structure of the problem is thus captured by the boundary condition in the time dimension and the original value function is given by $V(x) = v(0,x) = \sup_{\ell\leq x}(v(x-\ell,0+)+\ell)$.

3.2 Numerical convergence

Like in Section 2, we define the operator $\mathcal{L}$ as the operator mapping a function ϕ according to

$$\mathcal{L}\phi(x) = \sup_I \mathbb{E}_x\left[-\sum_{T>s\geq 0} e^{-\rho s}(\lambda_f + (1+\lambda_p)\Delta I_s)1_{\{\Delta I_s>0\}} + e^{-\rho T}\phi(X^{0,I}_{T-})1_{\{T<\theta(L,I)\}}\right]. \quad (4)$$

We think of this as the solution of (1) and (2). The discrete operator $\mathcal{D}$ is then given by (3), i.e.,

$$\mathcal{D}\phi(x) = \sup_{\ell\leq x}(\phi(x-l)+\ell), \quad (5)$$

and $\mathcal{T} := \mathcal{D}\circ\mathcal{L}$. Next, we construct a space $\mathcal{X}$ in which $\mathcal{T}$ has a fixed point:

Theorem 3.2. There exists a space $\mathcal{X}$ on which $\mathcal{T} : \mathcal{X} \to \mathcal{X}$ is a strict contraction.

Proof. First, for $A \geq 0$, let $\mathcal{X}_A$ be the space of universally measurable functions $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $f(x) \in [x, A+x]$. We begin by showing that there exists an A^* such that $\mathcal{T} : \mathcal{X}_{A^*} \to \mathcal{X}_{A^*}$. To emphasize the upper and lower boundaries, let $\underline{f} := \mathrm{id}_\mathbb{R}$ and $\overline{f} := A + \mathrm{id}_\mathbb{R}$, where $\mathrm{id}_\mathbb{R}$ is the identity function on $\mathbb{R}$, i.e., $\mathrm{id}_\mathbb{R}(x) := x$. It is clear that for every $f \in \mathcal{X}_A$, $\mathcal{D}f \geq f(0) + \underline{f}$, by the possibility of paying out all surplus as dividends. Moreover, by monotonicity (see (4)), $\mathcal{L}f \geq \mathcal{L}\underline{f}$, for every $f \in \mathcal{X}_A$. This shows the lower bound, i.e., that $\mathcal{T}f \geq \mathcal{D}\mathcal{L}\underline{f} \geq \mathcal{L}\underline{f}(0) + \underline{f} \geq \underline{f}$, for $f \in \mathcal{X}_A$. Moreover, observe that

$$\mathcal{L}\overline{f} = \mathcal{L}(A + \mathrm{id}_\mathbb{R}) \leq e^{-\rho T}A + \mathcal{L}\mathrm{id}_\mathbb{R}.$$

Using the stochastic representation of $\mathcal{L}$,

$$\begin{aligned}
\mathcal{L}\mathrm{id}_\mathbb{R}(x) &= \sup_I \mathbb{E}_x\left[-\sum_{t\in[0,1)} e^{-\rho t}(\lambda_f + (1+\lambda_p)\Delta I_t)1_{\{\Delta I_t>0\}} + e^{-\rho T}X^{0,I}_{T-}1_{\{\inf_{t\in(0,T)} e^{-\rho T}X^{0,I}_t\geq 0\}}\right] \\
&= \sup_I \mathbb{E}_x\left[-\sum_{t\in[0,1)} e^{-\rho t}(\lambda_f + (1+\lambda_p)\Delta I_t)1_{\{\Delta I_t>0\}} + e^{-\rho T}(X^{0,0}_{T-} + I_{T-})1_{\{\inf_{t\in(0,T)} X^{0,I}_t\geq 0\}}\right] \\
&\leq \mathbb{E}_x\left[e^{-\rho T}X^{0,0}_{T-}1_{\{\inf_{t\in(0,T)} X^{0,I}_t\geq 0\}}\right] \leq \mathbb{E}_x\left[e^{-\rho T}\left(X^{0,0}_{T-}\right)^+\right] \leq e^{-\rho T}x + e^{-\rho T}\mathbb{E}_x[(C_{T-})^+].
\end{aligned}$$

Therefore,

$$\mathcal{T}\overline{f}(x) \leq x + e^{-\rho T}(A + \mathbb{E}_x[(C_{T-})^+]) \leq \overline{f} + (e^{-\rho T}-1)A + e^{-\rho T}\mathbb{E}_x[(C_{T-})^+],$$

so $\mathcal{T}\overline{f} \leq \overline{f}$ whenever

$$A \geq \frac{e^{-\rho T}\mathbb{E}_x[(C_{T-})^+]}{-(e^{-\rho T}-1)} =: A^*,$$

which is finite, by Assumption 3.1. From now on, let $\mathcal{X} = \mathcal{X}_{A^*}$.

We have yet to show that for some metric, $\mathcal{T}$ is a strict contraction. Define the metric $d(f,g) = \sup_{\mathbb{R}_{\geq 0}} |f-g|$. Then, for any $f, g \in \mathcal{X}$,[4]

$$
\begin{aligned}
(\mathcal{T}f - \mathcal{T}g)(x) &\leq \sup_{\ell \leq x} \sup_{I} \mathbb{E}_x \Big[e^{-\rho T} (f-g)(x - \ell + C_{T-} + I_{T-}) 1_{\{\inf_{t\in(0,T)} X_t^{0,I} - \ell \geq 0\}} \Big] \\
&\leq e^{-\rho T} \sup_{\mathbb{R}_{\geq 0}} (f-g),
\end{aligned}
$$

which, after taking the supremum over the left hand side shows that

$$
d(\mathcal{T}f, \mathcal{T}g) \leq e^{-\rho T} d(f,g),
$$

i.e., that $\mathcal{T}$ is a strict contraction. □

3.3 Numerical results

To compute the value function, the two operators $\mathcal{D}$ in (5) and $\mathcal{L}$ in (4) need to be implemented. The former is straight-forward to implement, but the latter requires a bit more work.

For a model without equity issuance, $\mathcal{L}$ can for instance easily be implemented by means of Monte Carlo simulations. This is particularly convenient for cash flow processes without diffusion, like Cramér–Lundberg model. The reason for this is that to evaluate the indicator function in (4), a test for ruin only has to be made at the time of a jump. On the other hand, in a diffusion model, this has to be estimated by making increasingly smaller time steps.

Fortunately, also with equity issuance, alternative methods can be employed to solve the problem. We assume that the PDE representation (1)–(2) holds and opt for the semi-Lagrangian method presented in [6]. Since we wish to compute the solution on a bounded domain, an artificial boundary condition has to be specified. At any time point, any additional inflow of cash at the upper boundary is paid out as dividends at the next opportunity, provided the reserves do not fall below the dividend barrier. As the computational domain is chosen larger, it is therefore increasingly unlikely that additional cash is not paid out. Hence, if the domain is chosen sufficiently large, the present value of Δx at the boundary is its discounted value $e^{-\rho(T-t)}\Delta x$. The boundary condition $v_x(x,t) = e^{-\rho(T-t)}$ is therefore a good approximation.

With means for calculating both operators $\mathcal{D}$ and $\mathcal{L}$, we may proceed to iteratively apply $\mathcal{T}$ to any arbitrary initial function. We do so for the C being a Brownian motion with drift, as in [13], i.e., $\mathcal{A} = \mu\partial_x + \frac{1}{2}\sigma^2\partial_{xx}$. Our choice of parameters for the computations in this section comes from [14] and are listen in Figure 1. For our purposes, we consider all parameters to be in fractions of their so-called regulatory risk-weighted assets.

The results without equity issuance is presented in Figure 1. The loss due to discretization of dividend payments quickly falls to a level below 1.4%. The primary impression of the result is that the loss is relatively small. Note that for larger values of x,

[4] If the point of evaluation is negative for some outcomes, the indicator function is zero, so we consider the whole expression to be zero.

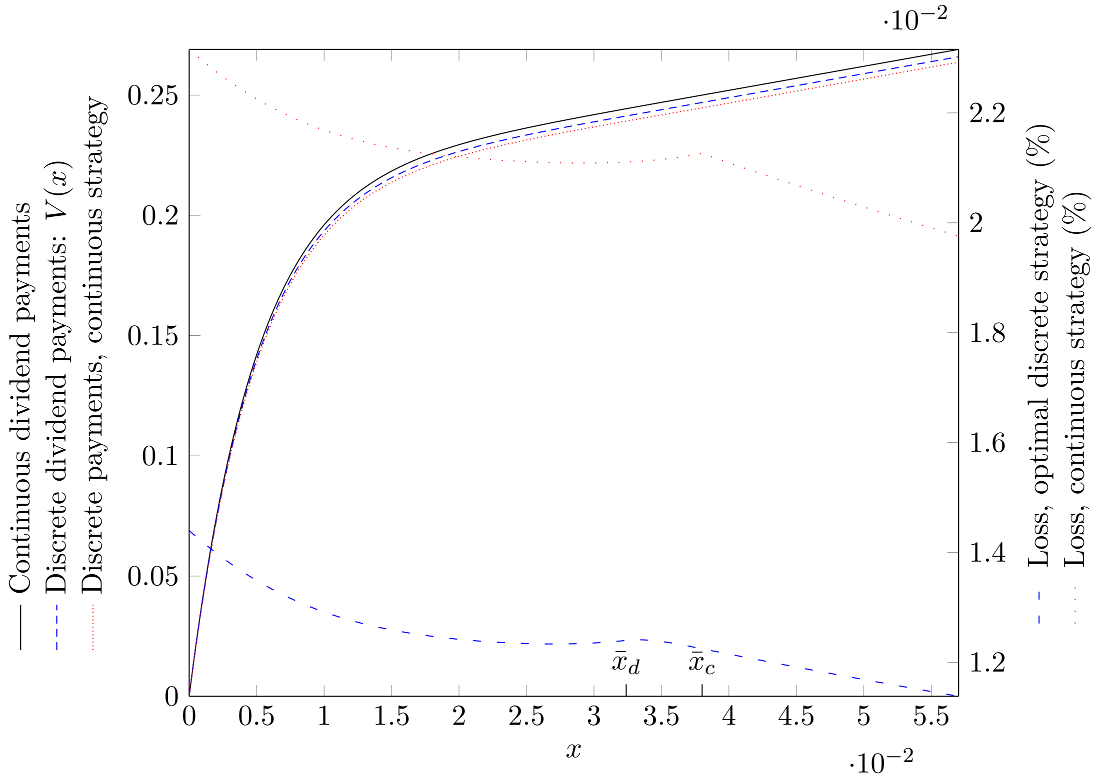

Figure 1: Value functions and dividend policies without equity issuance. On the left axis are plots of the value function V in the discrete model (blue, dashed), the value function for the problem with continuous (singular) dividend payments (black, solid), and the value obtained from (suboptimally) using the optimal continuous strategy in the discrete model (red, dotted). On the right axis are the losses in percent due to discretization of dividend payments, relative to the continuous model. The two lines denote the losses using the optimal discrete strategy as well as the (suboptimal) continuous strategy. The cash flow is given by $C_t = \mu t + \sigma W_t$ and the parameters values are $\rho = 0.04$, $T = 1$, $\sigma = 0.01$, and $\mu = 0.01$. The values $\bar{x}_d$ and $\bar{x}_c$ at the bottom are the dividend barriers in the discrete and continuous models respectively. The difference corresponds to 14% lower reserves in the discrete model.

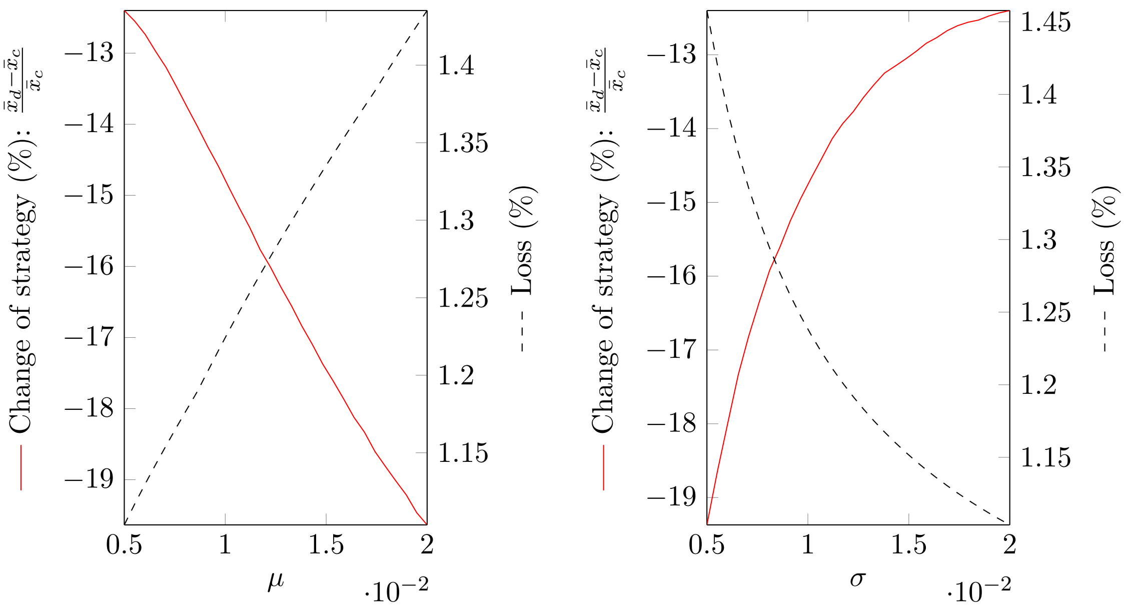

Figure 2: Effect of the parameters μ and σ without issuance. The change in strategy in terms of the relative distance between the continuous dividend strategy and the discrete one. The loss is evaluated at the optimal dividend barrier for the continuous problem. The fixed parameters are the same as in Figure 1.

the absolute loss stays constant, so the relative loss decays as the value function increases linearly. Although the change in the value function is not very large, the dividend barrier moves considerably, decreasing by 14% (slightly more than 0.005 units) in Figure 1. We attribute this mainly to paying out some of the expected income during the next period. Note, however, that due to the need of keeping a buffer, only a bit more than half of the expected cash flow is paid out in advance.

One important aspect of discretization of dividends is that the use of the continuous time optimal dividend threshold in the discrete problem induces further losses, since it is no longer optimal in the discrete model. To shed some light on the effect of using the wrong policy in this way, Figure 1 also shows the value function resulting from using the continuous time dividend barrier of the discrete dividend payments (the smallest of the three functions). We also plot the relative losses in comparison to the continuous dividend model. For the parameters in the figure, we observe that using the wrong policy adds a bit more than 0.8 percentage points to the losses.

Figures 2 and 3 both show the effect of varying some of the parameters. The loss comparisons are all made at the optimal barrier of the continuous continuous model, $\bar{x}_c$. The rationale for this is that it is the level of reserves of a healthy firm. Changes to μ and σ that increase the value of the continuous model also raises the *relative* loss from discretizing the dividend strategy. We also observe that for all parameter values

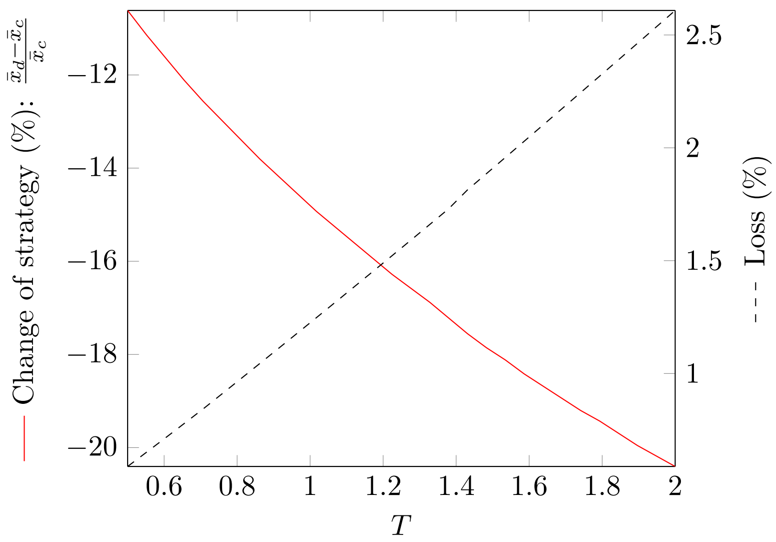

Figure 3: Effect of the parameter T without issuance. The change in strategy measures the relative distance between the continuous dividend strategy and the discrete one. The loss is evaluated at the optimal dividend barrier for the continuous problem. The fixed parameters are the same as in Figure 1.

in the ranges considered, there is a significant shift in the optimal strategy. In Figure 3 we solve the problem for different values of T. This is the only parameter that does not affect the continuous time problem. As expected, the size of T has a strong impact on the losses. The figure suggests that the loss in the value function is low for quarterly dividend payments, but still the dividend strategy is quite different.

Figure 4 shows the value function for the model with issuance. As expected for issuance costs independent of t and x, issuance only occurs at the boundary. Note that since $\lambda_p = 0$, the optimal issuance target[5] coincides with the dividend barrier, in this case roughly 0.0125. We observe that the size of issued equity grows as time passes, with the exception of the period right before the time of dividend payment where it drops to its initial value.

4 Discrete dividends with random profitability

Instead of the constant drift considered in Section 3.3, one could consider the drift—the *profitability*—to be described by another random process. Suppose that the cash flow $C = C^\mu$ depends on some profitability process $(\mu_t)_{t\geq 0}$.[6] We then write the net cash

[5]Because $\lambda_p = 0$, the target is not unique at time points coinciding with dividend payments, since excessive issuance can be offset by dividend payments at no cost. At these points we consider the optimal issuance target to be the smallest optimizer.

[6]The typical dependency of C^μ on μ is given by $\mathrm{d}C_t = \mu_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$. However, we do not restrict ourselves to this structure in our theoretical analysis.

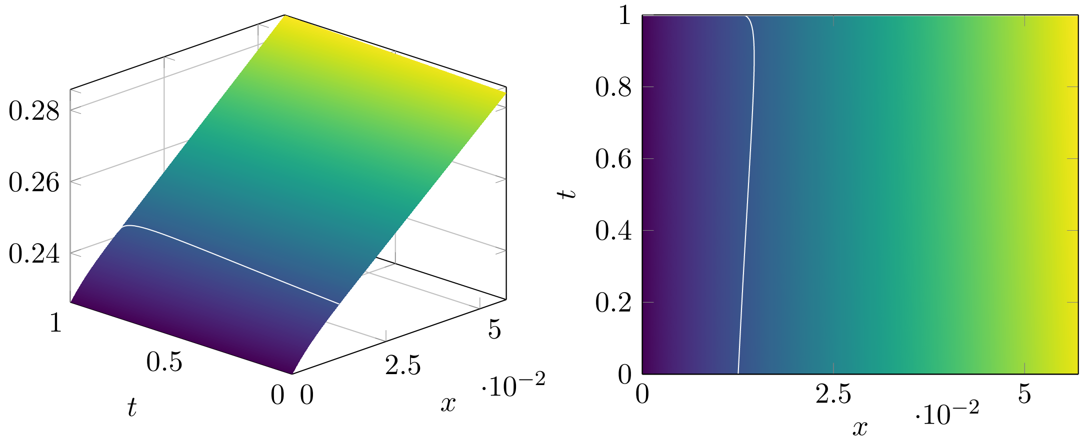

Figure 4: Surface plots of the value function $v(x,t)$ with issuance, and the optimal issuance target for $t \in [0,1)$. Issuance only occurs at the boundary $x = 0$, and the issuance target is presented as the white line on the surface. The issuance costs are $\lambda_p = 0$, $\lambda_f = 0.0025$, and the remaining parameters are the same as in Figure 1.

reserves $X = (X_t)_{t \geq 0}$

$$\mathrm{d}X_t^{L,I} = \mathrm{d}C_t^{\mu} - \mathrm{d}L_t + \mathrm{d}I_t, \quad X_0^{L,I} = x.$$

Let $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$ be the filtration generated by (C^{μ}, μ). Again, we restrict dividends to be fully covered by reserves of the firm, i.e., $\Delta L_t \leq X_{t-}^{L,I}$, and the ruin time is given by

$$\theta(L, I) = \inf\{t > 0 : X_t^{L,I} < 0\}.$$

Just as before, the aim of the firm is to maximize the discounted value of dividends net of equity issuance, i.e.,

$$V(x,\mu) = \sup_{L,I} \mathbb{E}_{x,\mu}\left[\sum_{t \in T\mathbb{N} \cap [0,\theta(L,I)]} e^{-\rho t}\Delta L_t - \sum_{t\geq 0} e^{-\rho t}(\lambda_f + (1+\lambda_p)\Delta I_t)1_{\{\Delta I_t > 0\}}\right],$$

where the optimization again is over increasing, RCLL, adapted processes, and where $\mathbb{E}_{x,\mu}$ denotes expectation under a measure such that $(X_{0-}^{L,I}, \mu_0) = (x, \mu)$.

4.1 Periodization and numerical convergence

Note that the periodization in Section 3.1 does not rely on the dynamics for X or C. We can thus take the same steps to arrive at

$$\begin{aligned}\mathcal{L}\phi(x,\mu) := \sup_I \mathbb{E}_{x,\mu}\Bigg[&-\sum_{T>s\geq 0} e^{-\rho s}(\lambda_f + (1+\lambda_p)\Delta I_s)1_{\{\Delta I_s>0\}} \\ &+ e^{-\rho T}\phi(X^{0,I}_{T-},\mu_{T-})1_{\{T<\theta(L,I)\}}\Bigg].\end{aligned} \tag{6}$$

This is the solution of (1) and (2). The discrete operator $\mathcal{D}$ is given by (3), i.e.,

$$\mathcal{D}\phi(x,\mu) := \sup_{\ell\leq x}(\phi(x-l,\mu)+\ell).$$

As seen in the proof of Theorem 3.2, the equity issuance poses no extra obstacle. For the sake of simplifying the involved expressions, we assume here that equity issuance is not allowed.

We make the following assumption on C^μ and $(\mu_t)_{t\geq 0}$. The assumption plays the same role as Assumption 3.1, but ensures that the effect of random profitability is sufficiently well behaved. In particular, it restricts the profitability process from having too strong growth.

Assumption 4.1. There exists an $\alpha:\mathbb{R}\to[1,\infty)$ so that for all μ

1. $\mathbb{E}_{x,\mu}[(x+C^\mu_{T-})^+] \leq x + A\alpha(\mu)$, for some $A\geq 0$.
2. $\mathbb{E}_{x,\mu}[\alpha(\mu_{T-})] \leq e^{\rho T/2}\alpha(\mu)$;

Theorem 4.2. There exists a metric space $(\mathcal{X}_\alpha, d_\alpha)$ such that the operator $\mathcal{T}$ maps $\mathcal{X}_\alpha$ into itself and is a strict contraction.

Proof. We prove the statements for the following subspace of universally measurable functions:

$$\mathcal{X}_\alpha := \Big\{x\leq\phi(x,\mu)\leq x + A_\phi\alpha(\mu) \text{ for some } A_\phi\Big\}$$

with metric

$$d_\alpha(\phi,\psi) := \sup_{x\geq 0,\mu\in\mathbb{R}} \frac{|\phi(x,\mu)-\psi(x,\mu)|}{\alpha(\mu)}.$$

Note that this implies that $|\phi(x,\mu)-\psi(x,\mu)|\leq d_\alpha(\phi,\psi)\alpha(\mu)$.

Then, for $\phi\in\mathcal{X}_\alpha$,

$$\begin{aligned}e^{\rho T}\mathcal{L}\phi(x,\mu) &\leq \mathbb{E}_{x,\mu}[(X_{T-}+A_\phi\alpha(\mu_{T-}))1_{\{\theta\geq T\}}] \\ &\leq \mathbb{E}_{x,\mu}[(x+C^\mu_{T-})^+] + A_\phi\mathbb{E}_{x,\mu}[\alpha(\mu_{T-})] \\ &\leq x + A\alpha(\mu) + e^{\rho T/2}A_\phi\alpha(\mu) \\ &\leq x + A'\alpha(\mu).\end{aligned}$$

Hence, $\mathcal{T}\phi(x,\mu) \leq x + e^{-\rho T} A' \alpha(\mu)$, so $\mathcal{T}\phi \in \mathcal{X}_\alpha$.

It is left to show that $\mathcal{T}$ is a strict contraction. By the properties of $\mathcal{T}$ and the construction of d_α,

$$\begin{aligned} |\mathcal{T}\phi(x,\mu) - \mathcal{T}\psi(x,\mu)| &\leq e^{-\rho T}\mathbb{E}_{x,\mu}[d_\alpha(\phi,\psi)\alpha(\mu_{T-})] \\ &\leq e^{-\rho T} e^{\rho T/2} d_\alpha(\psi,\phi)\alpha(\mu) \\ &\leq e^{-\rho T/2} d_\alpha(\phi,\psi)\alpha(\mu). \end{aligned}$$

This implies that $d_\alpha(\mathcal{T}\phi, \mathcal{T}\psi) \leq e^{-\rho T/2} d_\alpha(\phi,\psi)$, showing that $\mathcal{T}$ is indeed a strict contraction. □

Remark 4.3. Assumption 4.1 is satisfied by $C_t^\mu = \int_0^t \mu_s \,\mathrm{d}s + \sigma W_t$, where μ is the Ornstein–Uhlenbeck processes

$$\mathrm{d}\mu_t = k(\bar{\mu} - \mu_t)\,\mathrm{d}t + \tilde{\sigma}\,\mathrm{d}\tilde{W}_t,$$

where k, $\bar{\mu}$, and $\tilde{\sigma}$ are positive constants. By the time-scaled representation of Ornstein–Uhlenbeck processes, for $t \in [0,1]$ we have

$$\begin{aligned} \mathbb{E}_{x,\mu}[(\mu_t)^+] &= \mathbb{E}_{x,\mu}\left[\left(\mu e^{-kt} + \bar{\mu}(1-e^{-kt}) + \frac{\tilde{\sigma}}{\sqrt{2k}} e^{-kt}\tilde{W}_{e^{2kt}-1}\right)^+\right] \\ &\leq \mu^+ + \bar{\mu} + \frac{\tilde{\sigma}}{\sqrt{2k}}\mathbb{E}_{x,\mu}\left[\sup_{t\in[0,e^{2k}-1]} \tilde{W}_t\right] \\ &= \mu^+ + \bar{\mu} + \tilde{\sigma}\sqrt{\frac{e^{2k}-1}{k\pi}}. \end{aligned}$$

Hence, $\alpha(\mu) = \mu^+ + A$ satisfies the second condition of Assumption 4.1 for any

$$A \geq \frac{\bar{\mu} + \sigma\sqrt{\frac{e^{2k}-1}{k\pi}}}{e^{\rho T/2} - 1} \vee 1.$$

However, this estimate is also sufficient for the first condition, since

$$\mathbb{E}_{x,\mu}\left[\left(x + \int_0^1 \mu_t \,\mathrm{d}t + \sigma W_t\right)^+\right] \leq x^+ + \int_0^1 \mathbb{E}_{x,\mu}[(\mu_t)^+]\,\mathrm{d}t + \sigma\sqrt{\frac{2}{\pi}}.$$

We therefore conclude that the conditions of Assumption 4.1 are satisfied for this choice of C^μ and μ.

4.2 Numerical results

Assuming the dynamic programming principle holds, it follows that the value function solves

$$\min\left\{(\partial_t + \mathcal{A} - \rho)v(t,x,\mu), \quad v(t,x,\mu) - \sup_{i\geq 0}(v(t,x+i,\mu) - (1+\lambda_p)i - \lambda_f)\right\} = 0, \tag{7}$$

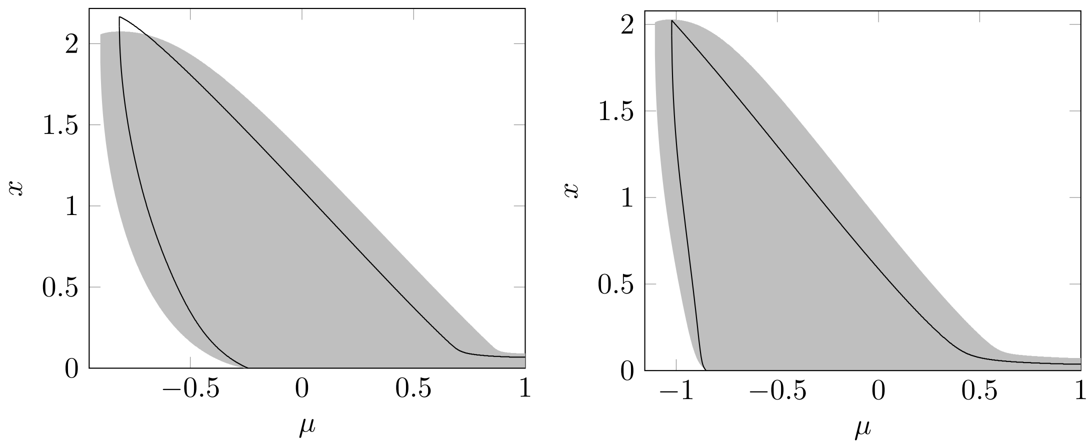

Figure 5: State space and the free boundaries of (7)–(8) **(black lines).** The left panel is without equity issuance and the right panel is with equity issuance. Between the two lines, it is optimal to not pay dividends, whereas outside it is. The gray area corresponds the same model, but allowing for continuous payments of dividends. The interpretation is the same, but the area between the lines is filled. The cash flow is given by $C_t^\mu = \int_0^t \mu_s \,\mathrm{d}s + \sigma W_t$, where $\mathrm{d}\mu_t = k(\bar{\mu} - \mu_t)\,\mathrm{d}t + \tilde{\sigma}\,\mathrm{d}\tilde{W}_t$. Parameter values are $\lambda_f = 0.1$, $\lambda_p = 0.2$, $\rho = 0.05$, $T = 1$, $\sigma = 0.1$, $k = 0.5$, $\bar{\mu} = 0.15$, $\tilde{\sigma} = 0.3$, and $\mathrm{Cov}(W_t, \tilde{W}_t) = 0$.

with the boundary condition $v(t, 0, \mu) = 0$ in the viscosity sense, i.e.,

$$v(t,0,\mu) = \max\{0,\ -(\partial_t + \mathcal{A} + 1 - \rho)v(t,0,\mu),\ \sup_{i\geq 0}(v(t,i,\mu) - (1+\lambda_p)i - \lambda_f\}. \quad (8)$$

Like in the one-dimensional case, we will employ this PDE formulation for the numerical solution of the problem.

We solve the model for $\mathrm{d}C_t^\mu = \mu_t \,\mathrm{d}t + \sigma\,\mathrm{d}W_t$, where μ is an Ornstein–Uhlenbeck process.[7] This model was explored for continuous dividend payments in [15]. Also in this case, we opt for a semi-Lagrangian scheme, and for the same reason as in the one-dimensional model, we place the same boundary condition $v_x = e^{-\rho(T-t)}$ on the upper boundary in the x-dimension. In the μ-dimension boundary conditions also have to be set. For the sake of our calculations, we mirror the process μ at the boundary. There are better choices, but we expect it to have a relatively small impact due to the Ornstein–Uhlenbeck process' strong inward drift at the boundary.[8]

The dividend boundaries can be seen for models with and without equity issuance in Figure 5. As in [15], we interpret the two lines constituting the dividend boundary in different ways: The upper line has the same interpretation as the dividend barrier in

[7] Recall from Remark 4.3 that this class of processes satisfies Assumption 4.1 required for numerical convergence.

[8] The results are consistent with disregarding the diffusion at the μ-boundary. In fact, disregarding the diffusion at the lower boundary and mirroring at the upper seems to allow the smallest domain without impacting the free boundaries, i.e., retaining stability with respect to choosing a larger domain.

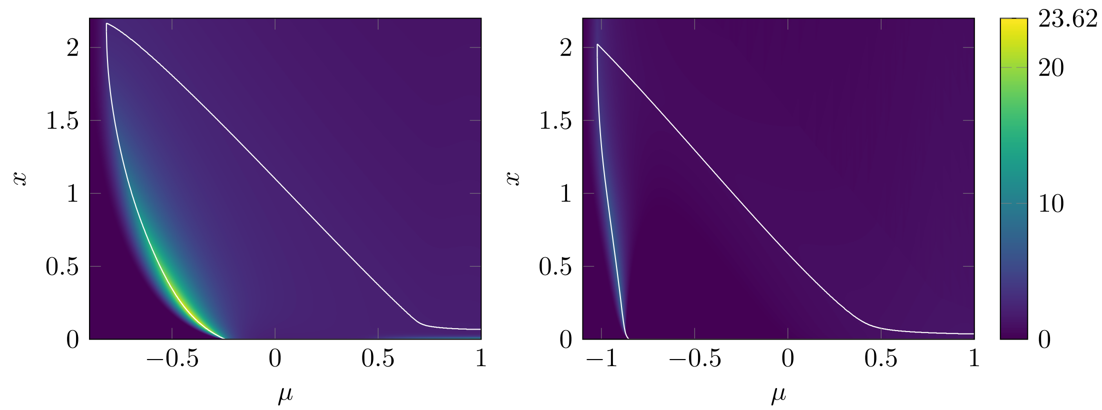

Figure 6: Heatmap of the relative loss from discrete dividends relative to continuously paid dividends. The left panel is without issuance and the right panel is with issuance. The scale is given in percent. Parameters are the same as in Figure 5. The white curves constitute the dividend/liquidation boundaries for the discrete problem.

the one-dimensional setting: whenever the reserves are above it at the time of dividend payments, dividends are paid out such that the reserves move down to the line. The lower boundary has a more subtle interpretation. Mathematically seen, dividends are paid out whenever the reserves lie below this line. Since the new state will still lie below the line, dividends must be paid until the reserves reach zero. The interpretation of this is that the firm liquidates whenever the reserves dip below the line. We will call these two lines the dividend boundary and the liquidation boundary. For points to the left of these lines, the profitability is so low that liquidation is optimal regardless of reserves.

The general effect of dividend discretization is consistent in the two figures; the dividend boundary moves downwards for most values of the profitability, with the exception of points close to where it meets the liquidation boundary. Just like with constant profitability, we ascribe the lower dividend boundary to paying out profits in advance. The liquidation boundary, on the other hand, moves upwards/inwards for all points. This is likely due to the reduction in the prospective future value in the event of higher profitability, thus reducing today's value of not liquidating. In both the continuous and discrete models, issuance only occurs at the boundary for the chosen parameter values.

Figure 6 shows the relative loss of discrete dividends to continuous dividends for the various points in the state space. The losses peak around the liquidation boundary for the discrete solution. At these points, the losses are close to 25% without issuance and a bit above 8% with issuance. For higher profitability and larger reserves, the losses soon dip below 3% without issuance and 1% with, decaying to 0 as x increases. In particular in the model with equity issuance, we see that the loss from discrete dividend payments is relatively small. The average loss for all the points of the shown domain is less than 0.8%.

5 Concluding remarks

For the one-dimensional dividend problem of Section 3, we find that the losses from dividend discretization are relatively low. We have observed the same, relatively small losses also for other parameter choices, and believe that it extends to most reasonable choices in this one-dimensional setting. In particular, for quarterly or more frequent dividends, the losses are especially small. The overall small losses provide justification for using a continuous model as a substitute, if the goal is to find the value function/value the cash flow.

On the other hand, the richer model presented in Section 4 paints another picture. For the parameters considered, we see that the total losses increase to almost 24% in some parts of the state space. This suggests that the impact of dividend discretization is by all means very model dependent.

In both models there is a pronounced shift in the optimal strategy. This implies that using the optimal continuous dividend policy would induce further losses, as is illustrated in Figure 1.

We conclude with the remark that whether the traditional continuous modelling is appropriate in a given setting is highly model-dependent. In particular, the choice of dividends modelling has to be made on a case-by-case basis.

References

[1] Erdinç Akyildirim, I. Ethem Güney, Jean-Charles Rochet, and H. Mete Soner. Optimal dividend policy with random interest rates. *Journal of Mathematical Economics*, 51:93–101, 2014.

[2] Hansjörg Albrecher and Arian Cani. Risk theory with affine dividend payment strategies. In *Number Theory–Diophantine Problems, Uniform Distribution and Applications*, pages 25–60. Springer, 2017.

[3] Hansjörg Albrecher and Stefan Thonhauser. Optimality results for dividend problems in insurance. *RACSAM-Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas*, 103(2):295–320, 2009.

[4] Søren Asmussen and Hansjörg Albrecher. *Ruin Probabilities (Advanced series on statistical science & applied probability; v. 14).* World Scientific, 2010.

[5] Benjamin Avanzi and Bernard Wong. On a mean reverting dividend strategy with brownian motion. *Insurance: Mathematics and Economics*, 51(2):229–238, 2012.

[6] Parsiad Azimzadeh, Erhan Bayraktar, and George Labahn. Convergence of approximation schemes for weakly nonlocal second order equations. *arXiv preprint arXiv:1705.02922*, 2017.

[7] Richard Bellman. A markovian decision process. *Indiana Univ. Math. J.*, 6:679–684, 1957. ISSN 0022-2518.

[8] Dimitir P Bertsekas and Steven Shreve. *Stochastic optimal control: the discrete-time case.* 1978.

[9] Min Dai, Steven Kou, Mete Soner, and Chen Yang. Daily rebalancing of leveraged ETFs. 2018.

[10] Jean-Paul Décamps, Thomas Mariotti, Jean-Charles Rochet, and Stéphane Villeneuve. Free cash flow, issuance costs, and stock prices. *The Journal of Finance*, 66(5):1501–1544, 2011.

[11] Hans U Gerber and Elias SW Shiu. Optimal dividends: analysis with brownian motion. *North American Actuarial Journal*, 8(1):1–20, 2004.

[12] Hans-Ulrich Gerber. *Entscheidungskriterien für den zusammengesetzten Poisson-Prozess*. PhD thesis, ETH Zurich, 1969.

[13] Monique Jeanblanc-Picqué and Albert N Shiryaev. Optimization of the flow of dividends. *Russian Mathematical Surveys*, 50(2):257, 1995.

[14] Samu Peura and Jussi Keppo. Optimal Bank Capital with Costly Recapitalization*. *The Journal of Business*, 79(4):2163–2201, July 2006. ISSN 0021-9398, 1537-5374. doi: 10.1086/503660. URL `https://www.jstor.org/stable/10.1086/503660`.

[15] Max Reppen, Jean-Charles Rochet, and H Mete Soner. Dividends with random profitability rate. *arXiv preprint arXiv:1706.01813*, 2017.