Optimal dividend policies with random profitability*

Max Reppen† Jean-Charles Rochet‡ H. Mete Soner§

May 31, 2017

Abstract

We study an optimal dividend problem under a bankruptcy constraint. Firms face a trade-off between potential bankruptcy and extraction of profits. In contrast to previous works, general cash flow drifts, including Ornstein–Uhlenbeck and CIR processes, are considered. We provide rigorous proofs of continuity of the value function, whence dynamic programming, as well as uniqueness of the solution to the Hamilton–Jacobi–Bellman equation, and study its qualitative properties both analytically and numerically. The value function is thus given by a nonlinear PDE with a gradient constraint from below in one dimension. We find that the optimal strategy is both a barrier and a band strategy and that it includes voluntary liquidation in parts of the state space. Finally, we present and numerically study extensions of the model, including equity issuance and gambling for resurrection.

1 Introduction

This problem of optimizing dividend flows has its origins in the actuarial field of ruin theory, which was first theoretically treated by Lundberg [27] in 1903. The theory typically models an insurance firm, and initially revolved around minimizing the probability of ruin. However, in practice, the objective function used by most firms is rather to maximize their shareholder values. This is the reason for de Finetti's [12] 1957 proposal to instead optimize the net present value of dividends paid out until the time of ruin. With a positive discount rate of the dividends, de Finetti solved the problem for cash reserves described by a random walk. Since then, this new class of dividend problems has been extensively studied, especially in the context of insurance firms.

Although a dividend problem can be seen as assigning a value to a given cash flow, the problem formulation nevertheless retains an emphasis on the ruin time. This is contrasted by cash flow valuation principles such as real option valuation, first introduced

*The first and third authors were partly supported by the ETH Foundation, the Swiss Finance Institute, and Swiss National Foundation grant SNF 200020_172815.

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

‡SFI@UZH and UNIGE, and TSE/IDEI, email jeancharles.rochet@gmail.com.

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

by Myers [29] in 1977. Whereas the dividend problem seeks the value of a cash flow after it passing through a *buffer* (the cash reserves), the real option approach evaluates a cash flow without such a buffer. In other words, the latter is a valuation of a cash flow without any liquidity constraint, as opposed to the optimal dividend problem where the firm can reach ruin. The real option value thus provides a natural bound for the optimal dividend value, which turns out to be helpful in our analysis.

In the actuarial literature, the cash reserves are commonly described by a spectrally negative Lévy process with a positive drift of premiums and negative jumps of claims. We instead study cash reserves described by a diffusion process. Although this is not the natural insurance perspective, it is, as initiated by Iglehart [21], nevertheless studied as the limiting case of the jump processes.

Formulated as a problem of 'storage or inventory type', the general diffusion problem with singular dividend policies was solved by Shreve et. al [33] in 1984. In the case of constant coefficients in the cash reserves dynamics, Jeanblanc-Picqué & Shiryaev [22] found the solution by considering limits of solutions to problems with absolutely continuous dividend strategies. The optimal solution to this singular problem formulation is described by a so-called *barrier strategy*, which yields a reflected cash reserves process by paying any excess reserves as dividends. This divides the state space into two regions: dividends are paid above the barrier (*dividend region*), but not in the region between zero and the barrier (*no-dividend region*). This is contrasted by dividend *band strategies* which frequently appear in jump models and were first identified by Gerber [17]. Instead of the two spatial regions, there then exists at least one no-dividend region in which the origin is not contained. It thus creates a band-shaped no-dividend region in-between two dividend regions.

In financial and economics literature, the main focus is on diffusion models, and extensions often involve nonconstant interest rate, drift and/or diffusion coefficients. Indeed, external, macroeconomic conditions and their effects on profitability have a substantial impact on dividend policies, as shown by Gertler and Hubbard [18] and more recently by Hackbarth et al. [19]. Such macroeconomic effects have been studied in various forms. In particular, Anderson and Carverhill [5] as well as Barth et al. [8] numerically study continuously changing stochastic parameters, whereas Akyildirim et al. [1] consider stochastic interest rate following a Markov chain, and Jiang and Pistorius [23] consider model coefficients and interest rate both governed by Markov chains. Bolton et al. [9] similarly study the macroeconomic impact on both financial and investment opportunities. In contrast to coefficients influenced by macroeconomic factors, Radner and Shepp [32] already in 1996 modelled a firm which alternates between different operating strategies, thereby effectively controlling the model coefficients. Finally, other extensions include transaction costs of dividend payments or the possibility of equity issuance, cf. [1, 9, 13]. Another diffusion model with an element of mean-reversion can be found in [11], where the authors consider mean-reverting cash reserves, in contrast to the mean-reverting profitability in this paper. For further references, we refer the reader to [2, 6] and the references therein.

Our choice of diffusion model has a continuous, stochastic drift generated by a separate cash flow rate process. This structure yields a two-dimensional problem in which the dividend strategy depends on the current cash flow rate. In particular, for low (negative) rates, a band strategy is optimal, but at higher rates, dividends are optimally paid according to a barrier strategy, with a barrier level depending on the cash flow rate. Additionally, for very low rates, we prove that it is optimal to perform a voluntary liquidation, meaning that all cash reserves are paid instantaneously. Band structures are common for jump models, but here appear in a continuous model.[1] Finally, in addition to qualitative and numerical results, we provide proofs for continuity of the value function as well as a comparison principle for the dynamic programming equation.

2 Problem formulation

Consider a *cash flow* on the form

$$\mathrm{d}C_t^\mu = \mu_t^\mu \,\mathrm{d}t + \sigma \,\mathrm{d}W_t, \quad C_0^\mu = 0,$$

where W is a Brownian motion and the *cash flow rate* μ_t is described by

$$\mathrm{d}\mu_t^\mu = \kappa(\mu_t)\,\mathrm{d}t + \tilde{\sigma}(\mu_t)\,\mathrm{d}\tilde{W}_t, \quad \mu_0^\mu = \mu,$$

for some functions κ and $\tilde{\sigma}$, as well as another Brownian motion $\tilde{W}$ with correlation $\rho \in [-1, 1]$ to W. Despite the formulation of μ_t as a continuous process, most of the results extend naturally to the Markov chains studied in the literature.

The precise assumptions on the diffusion, given in Assumption 6.1, include Ornstein–Uhlenbeck processes

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

for constants $k > 0$, $\bar{\mu}$, and $\tilde{\sigma}$ as well as another commonly considered process, the Cox–Ingersoll–Ross (CIR) process:

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

for constants $k > 0$, $\bar{\mu}$, $\tilde{\sigma}$, and a. In fact, the assumption only impose asymptotic conditions as $|\mu| \to \infty$. This means that on any given bounded domain, κ and $\tilde{\sigma}$ can be general, provided certain growth conditions are satisfied outside the bounded domain, and provided the SDE has a well-defined solution. This is naturally satisfied by bounded processes. Interpreting $-\kappa$ as the derivative of some potential, it also includes the possibility of potentials with multiple wells (local minima), thus having several points of attraction.

We model a firm whose cash flow is given by the process $C^\mu = (C_t^\mu)_{t\geq 0}$. The firm pays dividends to its shareholders using cash accumulated from the cash flow. Let L_t denote the *cumulative dividends* paid out until time t. Then the cash reserves $X = (X_t)_{t\geq 0}$ of a firm with initial cash level x can be written as

$$\mathrm{d}X_t^{(x,\mu),L} = \mathrm{d}C_t^\mu - \mathrm{d}L_t, \quad X_0^{(x,\mu),L} = x.$$

[1] Similar properties have been observed by Anderson and Carverhill [5] and Murto and Terviö [28].

The objective of the firm is to maximize its shareholders' value, defined as the expected present value of future dividends, computed under the risk-adjusted measure.[2] Denote by $\mathcal{M}$ the domain on which μ_t resides. This domain is typically the whole real line, as for a Ornstein–Uhlenbeck process, or a half-line, for a CIR process. The value function is then defined as

$$V(x,\mu) := \sup_L \mathbb{E}\left[\int_0^{\theta^{(x,\mu)}(L)} e^{-rt}\,\mathrm{d}L_t\right], \quad (x,\mu) \in \mathcal{O} := [0,\infty) \times \mathcal{M}, \tag{1}$$

where $L = (L_t)_{t\geq 0}$ is required to be càdlàg and nondecreasing[3] with $\Delta L_t \leq X_{t-}$,[4][5]

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

is the time of bankruptcy. In particular, we interpret a payout $\Delta L_t = X_{t-}$ as a decision to liquidate the firm.

In this paper we discuss and characterize the solutions to three different benchmark problems before we present the main results: We prove that liquidation is optimal when the profitability falls below a certain level; that, by stochastic methods, the value function is continuous and is a viscosity solution of the dynamic programming (DPE) equation; that the DPE satisfies the comparison principle; and finally we provide a numerical scheme and extensive numerical results.

When the starting points (x,μ) of the cash reserves and the cash flow are clear from context, the superscripts will be dropped in order to simplify the notation. Similar omissions of superscripts will be done for the bankruptcy times and strategies L when it is clear what dividend policy is followed.

Remark 2.1. The problem formulation bears resemblance to the Merton consumption problem, which has been extensively studied in the mathematical finance literature. However, the crucial difference here is that the firm is always exposed to the risk of its own operations. In other words, there is no entirely safe asset. As a result, the problem lacks desired concavity properties. More specifically, this happens due to the possible quasi-convexity of $L \mapsto \theta(L)$. Figure 1 shows a case where $\theta\left(\frac{L^1+L^2}{2}\right) < \max\{\theta(L^1), \theta(L^2)\}$ for two strategies L^1 and L^2, which means that the convex combination of strategies in some scenario would pay out dividends after bankruptcy. Note that this loss of

[2] We assume that shareholders can diversify their portfolios and that the firm under study is small, so that its decisions do not alter the risk-adjusted measure.

[3] The process L must be nondecreasing because the limited liability of shareholders implies that $\mathrm{d}L_t$ cannot be negative. Section 8 considers the case when new shares can be issued at a cost, allowing to inject new cash into the firm.

[4] Shareholders cannot distribute more dividends than available cash reserves. Otherwise this would constitute fraudulent bankruptcy.

[5] Although the model allows infinite dividend payments of very general type, we argue that it is not less realistic than the absolutely continuous case where the 'frequency' is also infinite, but interpreted as a rate. Indeed, as suggested by (3) in Section 6.3 and exploited in Section 7, this model can be considered the limit when there is no bound on the dividend rate.

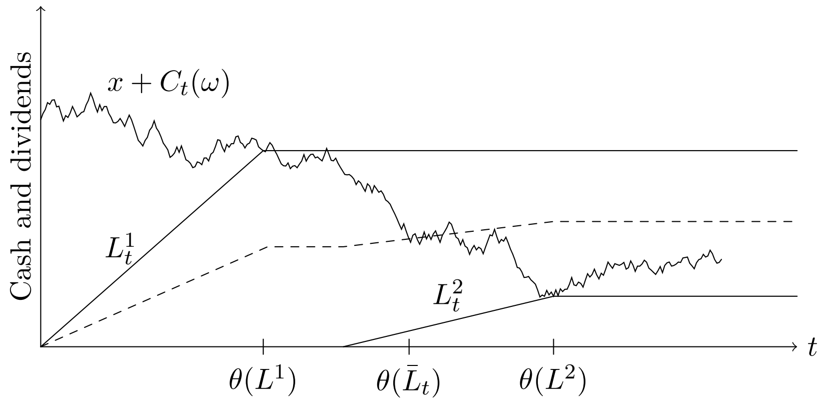

Figure 1: Recall that ruin occurs when the dividends reaches the total cash accumulated, i.e. $x + C_t$. The figure illustrates two dividend policies L^1 and L^2 (solid) as well as their convex combination $\bar{L} = (L^1 + L^2)/2$ (dashed), showing that for some path $\theta(\bar{L}) < \theta(L^2) \vee \theta(L^2)$, which corresponds to the possibility of dividend payments after ruin. Equivalently, the set of dividend strategies which are constant after ruin is nonconvex.

concavity corresponds to nonconvexity of the set of dividend processes that are constant after ruin.

3 A first benchmark: dividends of arbitrary sign

For the moment we consider the case where L is not restricted to be nondecreasing, which means that shareholders may inject new cash into the firm at no cost. In that case, cash reserves are useless and x must be distributed right away:

$$\tilde{V}(x, \mu) = x + \tilde{V}(\mu),$$

where

$$\tilde{V}(\mu) = \sup_{\tau} \mathbb{E}\left[\int_0^{\tau} e^{-rt} \mu_t \, \mathrm{d}t\right].$$

The liquidation time τ is chosen freely by the shareholders of the firm. This is a real option problem (see for example [14]). Intuitively, the owners of the firm exert the liquidation option when the profitability μ falls below a (negative) threshold μ^*, provided $(\mu_t)_{t\geq 0}$ can reach such a point. In particular, $\tilde{V}$ satisfies the following boundary value problem:

$$r\tilde{V}'(\mu) - \kappa(\mu)\tilde{V}'(\mu) - \frac{\tilde{\sigma}(\mu)^2}{2}\tilde{V}''(\mu) = \mu,$$

subject to

$$\tilde{V}(\mu^*) = \tilde{V}'(\mu^*) = 0.$$

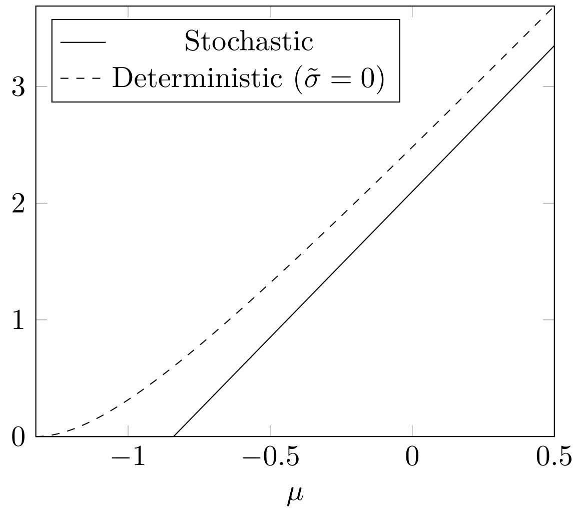

Figure 2: The value of the real option for the Ornstein–Uhlenbeck cash flow process $(\mu_t)_{t\geq 0}$. The stochastic curve is $\tilde{V}(\mu)$ and the deterministic comes from (2) ($\tilde{\sigma} = 0$), with the option to exit if the present value is negative. The parameters are $r = 0.05$, $\kappa(\mu) = 0.35(0.12 - \mu)$, and $\tilde{\sigma}(\mu) = 0.3$.

Theorem 3.1. When dividends can be of arbitrary sign, the optimal policy for shareholders is to immediately distribute the initial cash reserves at $t = 0$, and to maintain them at zero forever by choosing $\mathrm{d}L_t = \mu_t \, \mathrm{d}t + \sigma \, \mathrm{d}W_t$. If $\mathcal{M}$ has no lower bound, there exists a μ^* such that the firm is liquidated when profitability falls below the threshold μ^*: $\tau = \inf\{t > 0 : \mu_t \leq \mu^*\}$ is a maximizer.

The proof of this result is given in Section 9.1. An example of the function $\tilde{V}$ is plotted in Figure 2, and, as stated in the theorem, the real option attains the value 0 for small enough μ.

4 A second benchmark: the deterministic problem

If the two parameters σ and $\tilde{\sigma}$ determining randomness are set to zero, the problem can be solved explicitly for a mean-reverting $\kappa(\mu) = k(\bar{\mu} - \mu)$, for $k, \bar{\mu} > 0$, and the solution provides intuition for the solution of the stochastic problem.

Since $(\mu_t)_{t\geq 0}$ is mean-reverting to the positive value $\bar{\mu}$, it will never attain negative values once it has been positive. In particular, solving the ODE describing the dynamics of $(\mu_t)_{t\geq 0}$ yields

$$\mu_t = \bar{\mu} + (\mu - \bar{\mu})e^{-kt}.$$

We will treat the cases $\mu \geq 0$ and $\mu < 0$ separately.

If $\mu \geq 0$, the firm is always profitable, and there is therefore no default at $x = 0$ unless $\mu < 0$. Therefore, there is no need for a cash buffer, so it is optimal to pay all the initial cash reserves immediately. Thereafter, cash from μ is paid out as it flows in. The

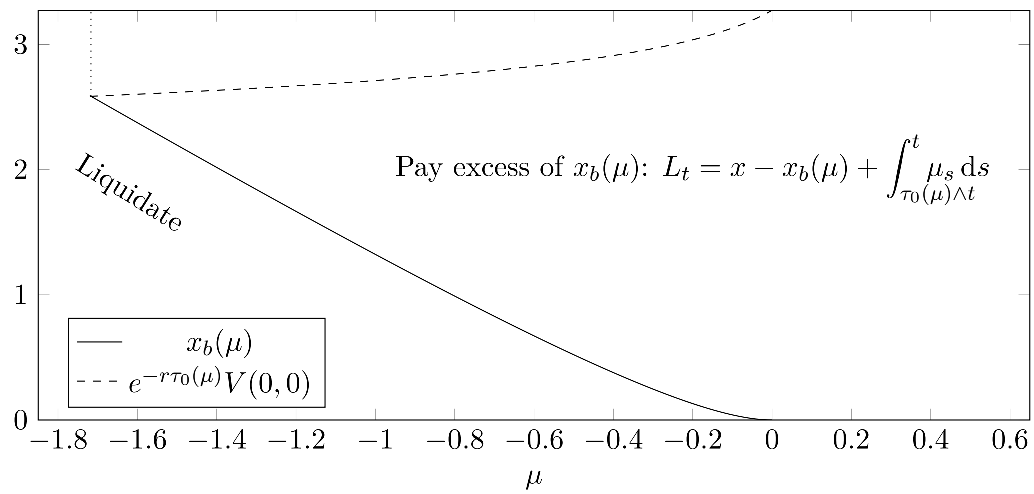

Figure 3: The value $x_b(\mu)$ is the cost of waiting for positive cash flow, whereas $e^{-r\tau_0(\mu)}V(0,0)$ is the present value of the future positive cash flow. For μ below the level at which these coincide, liquidation is thus optimal. Liquidation is also optimal when $x < x_b(\mu)$.

value of paying the incoming earnings as dividends is

$$\int_0^\infty e^{-rt}\mu_t \, \mathrm{d}t = \int_0^\infty \bar{\mu}e^{-rt} + (\mu - \bar{\mu})e^{-(k+r)t} \, \mathrm{d}t = \frac{\bar{\mu}}{r} + \frac{\mu - \bar{\mu}}{r+k}. \tag{2}$$

Hence, for $\mu \geq 0$, the value is given by

$$V(x,\mu) = x + \frac{\bar{\mu}}{r} + \frac{\mu - \bar{\mu}}{r+k}.$$

On the other hand, if the cash flow starts at a negative level, it will eventually reach a positive state, but the question is whether the firm is able to absorb the cumulated losses before. If it can, are those losses larger than future earnings? More precisely, the company could face ruin before it sees positive earnings, but even if it does not, the losses incurred could offset the value of the future positive cash flows. To address the first possibility, we calculate the minimum amount of cash needed to reach positive cash flow before the time of ruin. Denote by τ_0 the time such that $\mu_{\tau_0} = 0$. This time can be found explicitly:

$$\tau_0 = \tau_0(\mu) = \frac{\ln\left(\frac{\bar{\mu}}{\bar{\mu}-\mu}\right)}{-k}.$$

The cumulative losses until a positive cash flow is reached are:

$$\int_0^{\tau_0(\mu)} \mu_t \, \mathrm{d}t = \bar{\mu}\tau_0(\mu) + \frac{\mu - \bar{\mu}}{k}(1 - e^{-k\tau_0(\mu)}) = \bar{\mu}\tau_0(\mu) + \frac{\mu}{k}.$$

Hence, the initial cash level needs to be at least this high to survive until $\mu \geq 0$, i.e.,

$$V(x,\mu) = x, \quad \text{if} \quad x < -\bar{\mu}\tau_0(\mu) - \frac{\mu}{k} =: x_b(\mu).$$

At an initial cash level x above $x_b(\mu)$, we identify two possible strategies: Either pay out dividends of size $x - x_b(\mu)$ and wait for $(\mu_t)_{t\geq 0}$ to reach 0, or perform a liquidation by paying out x. Which strategy is optimal depends on the cost of waiting and the value of future cash flows. Hence, for $x \geq x_b(\mu)$,

$$V(x,\mu) = \max\{x, x - x_b(\mu) + e^{-r\tau_0(\mu)}V(0,0)\} = x + \max\left\{0, e^{-r\tau_0(\mu)}V(0,0) - x_b(\mu)\right\}.$$

Since x_b and τ_0 are both decreasing in μ, there exists a μ^* such that $e^{-r\tau_0(\mu^*)}V(0,0) = x_b(\mu)$, so from the last term we see that if $\mu \leq \mu^*$, it is optimal to liquidate regardless of cash level. In the model, this corresponds to paying all cash reserves as dividends at time $t = 0$, yielding the value $V(x,\mu) = x$.

With $x_b(\mu) = 0$ for $\mu \geq 0$, we have proved the following result:

Theorem 4.1. There exist thresholds $x_b(\mu)$ and μ^* such that

- it is optimal to liquidate immediately if $\mu \leq \mu^*$;
- it is optimal to liquidate immediately if $x < x_b(\mu)$;
- if $x \geq x_b(\mu)$ and $\mu \geq \mu^*$, it is optimal to immediately pay the excess $x - x_b(\mu)$ and thereafter all earnings as they arrive.

5 A third benchmark: a semi-deterministic problem

In this section we consider a special case of the dynamics described in Section 2. We assume that $\sigma > 0$, but that $(\mu_t)_{t\geq 0}$ is deterministic and mean-reverting, i.e., $\mathrm{d}\mu_t = k(\bar{\mu} - \mu_t)\,\mathrm{d}t$, for some $k > 0$ and $\bar{\mu} > 0$. By Remark 9.2, we may assume that the continuity property (and therefore dynamic programming) as well as the comparison principle hold also in this setting. The numerical results are qualitatively the same as when $(\mu_t)_{t\geq 0}$ is an Ornstein–Uhlenbeck process, but lends a bit more tractability. In Lemma 5.1, we show that the value function is concave in x for $\mu > 0$, and, by doing so, we can conclude that there cannot be more than one target wealth level for each μ in this region.

In this section we use results from Section 6. In particular, we use the continuity of V and the dynamic programming principle justified by it. Note that the proofs in Section 9 do not rely on the results presented here.

We begin by can defining the following boundary. Fix $\mu > 0$ and let $\partial_x V(x,\mu)$ be the set of super-differentials of the function $V(\cdot,\mu)$ at the point x, i.e.,

$$\partial_x V(x,\mu) = \{p \in \mathbb{R} \ : \ V(y,\mu) \leq V(x,\mu) + p(y-x), \quad \forall\ y \geq 0\}.$$

Since $V(y) - V(x) \geq y - x$ for all $x \leq y$, it holds that for all $x > 0$ and $\mu > 0$,

$$p \geq 1, \quad \forall\ p \in \partial_x V(x,\mu).$$

Furthermore, concavity implies that for $0 \leq x < y$,

$$p \in \partial_x V(x,\mu),\ \hat{p} \in \partial_x V(y,\mu) \quad \Longrightarrow \quad p \geq \hat{p}.$$

For $\mu > 0$ set

$$x^*(\mu) := \inf \{x \geq 0 \ : \ 1 \in \partial_x V(x,\mu)\}.$$

We set $x^*(\mu) = +\infty$ if the above set is empty. Then, from the above facts, it is clear that

$$\partial_x V(x,\mu) = \{1\},\ \forall\ x > x^*(\mu) \quad \text{and} \quad \partial_x V(x,\mu) \cap \{1\} = \emptyset,\ \forall\ x \in [0, x^*(\mu)).$$

It is possible that $x^*(\mu) = 0$. By the continuity properties of the sub-differentials, the map μ to $x^*(\mu)$ is lower semi-continuous, i.e.,

$$\liminf_{\hat{\mu} \to \mu}\ x^*(\hat{\mu}) \geq x^*(\mu).$$

In other words, if the value function is concave in x, the boundary separating states where dividends are paid from states where no dividends are paid is well-defined. Note that this construction is independent of the structure of $(\mu_t)_{t \geq 0}$.

The proof of the following theorem is given in Section 9.2.

Theorem 5.1. The value function is concave in x for $\mu > 0$.

6 Main results

The optimal strategy is characterized by three main regions: the dividend region, retain earnings region, and the liquidation region. The region of retained earnings is bounded by two curves and is characterized by $\mathrm{d}L_t = 0$. The dividend region and liquidation region are both characterized by $\mathrm{d}L \neq 0$, but correspond to different interpretations, and are separated by the threshold μ^*. More precisely, in the liquidation region, all available cash reserves are 'paid', leading to a liquidation. This is in contrast to the dividend region, where only the excess of the *dividend target* is paid to the shareholders. An illustration of the regions is presented in Figure 4.

The reimainder of this section is devoted to the statements of the main results. The proof of these results are given in Section 9. For the proofs we need the following set of assumptions, satisfied by for example Ornstein–Uhlenbeck and CIR processes:

Assumption 6.1. Throughout, we assume that the domain of $(\mu_t)_{t \geq 0}$ is some possibly unbounded interval $\mathcal{M}$. Moreover, we assume that κ and $\tilde{\sigma}^2$ are locally Lipschitz continuous on the interior $\mathcal{M}^{\mathrm{o}}$, that $-\mu/\kappa$ is non-negative and bounded for large (positive) μ, that $-\kappa/\mu$ is non-negative and bounded for large $-\mu$, as well as that $\tilde{\sigma}^2 \in \mathcal{O}(\mu)$ and never vanishes in $\mathcal{M}^{\mathrm{o}}$.[6] Finally, we also assume either of the following:

1. The function $\tilde{\sigma}^2$ is also locally Lipschitz on the boundary $\partial\mathcal{M}$.

[6]With small modifications, our results naturally extend to vanishing $\tilde{\sigma}$, provided it is either identically zero or not vanishing as $\mu \to \pm\infty$. However, in such cases, assumptions on existence of solutions to the SDE is necessary.

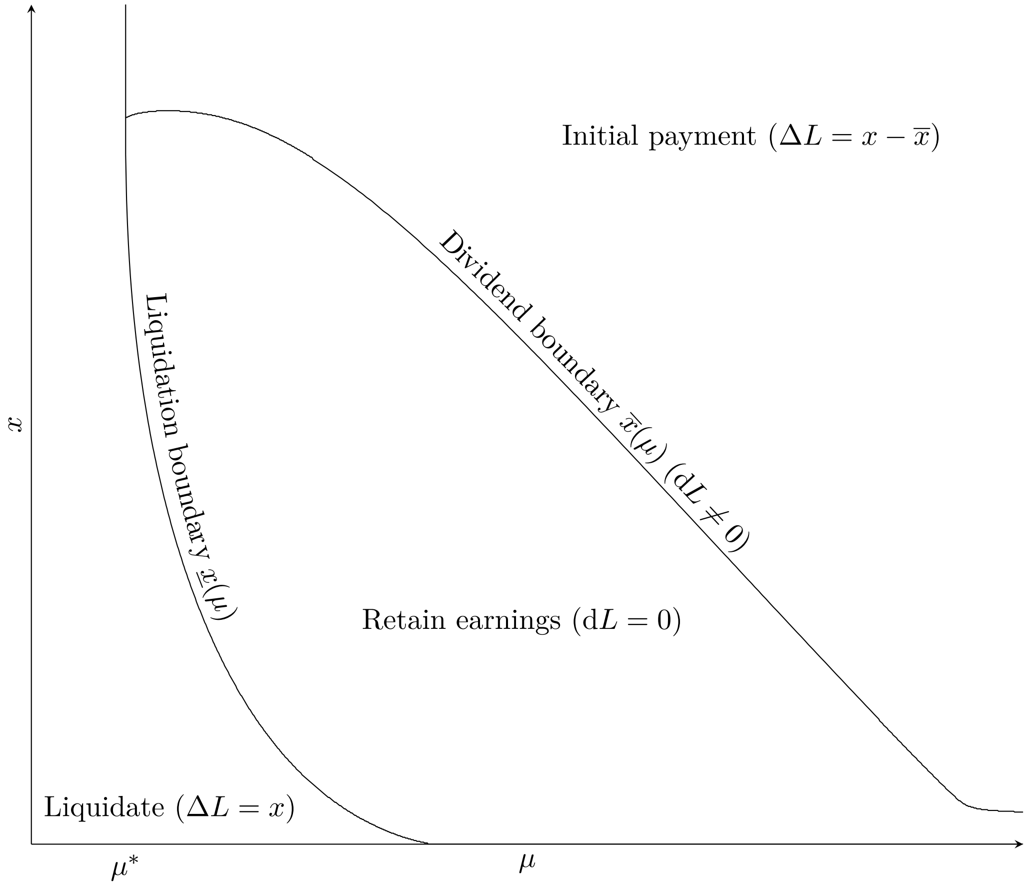

Figure 4: The figure shows the three strategy regions. In the region between the lines $\underline{x}$ and $\overline{x}$ the all incoming profits are retained. When the cash reserves fall to $\underline{x}$, the firm liquidates, whereas when it increases to $\overline{x}$, dividends are paid out according to local time at the boundary, thus reflecting the cash reserves process. In the region above $\overline{x}$, a lump sum of the excess of $\overline{x}$ is paid immediately. Finally, when $\mu \leq \mu^*$, liquidation is optimal at all cash levels.

2. For any sufficiently small $\eta > 0$, we assume that for $\nu = \inf \mathcal{M}$,

$$\limsup_{\mu \to \nu} \left(\frac{1}{\mu - \nu} - \frac{2\kappa(\mu) + \eta\rho\sigma\tilde{\sigma}(\mu)}{\tilde{\sigma}(\mu)^2} \right) < \infty,$$

and for $\nu = \sup \mathcal{M}$

$$\liminf_{\mu \to \nu} \left(\frac{1}{\mu - \nu} - \frac{2\kappa(\mu) + \eta\rho\sigma\tilde{\sigma}(\mu)}{\tilde{\sigma}(\mu)^2} \right) > -\infty.$$

The economic interpretation of the growth conditions on κ is that even if the profitability is very large, it eventually returns to a more reasonable level. The growth condition on $\tilde{\sigma}$ simply ensures that the diffusion does not overpower this effect. Finally, the lim sup and lim inf conditions at the boundary points are needed to ensure that the profitability process behaves well enough close to the boundary.

6.1 Liquidation threshold

Theorem 6.2. If $\mathcal{M}$ has no lower bound, there exists a value μ^* such that it is optimal to liquidate immediately whenever $\mu \le \mu^*$, i.e. $V(x, \mu) \equiv x$.

6.2 Continuity

Theorem 6.3. The value function is continuous everywhere.

6.3 Dynamic programming equation

Following the continuity of Theorem 6.3, we refer to [15] for proving the dynamic programming principle. For a general proof of dynamic programming, we refer to [24, 25]. Writing

$$\mathcal{L}V = \mu V_x + \kappa(\mu)V_\mu + \operatorname{Tr}\Sigma(\mu)D^2V,$$

where

$$\Sigma = \begin{bmatrix} \sigma^2 & \rho\sigma\tilde{\sigma} \\ \rho\sigma\tilde{\sigma} & \tilde{\sigma}^2 \end{bmatrix}$$

is the covariance matrix, the dynamic programming equation corresponding to (1) is given by

$$\min\{rV - \mathcal{L}V, V_x - 1\} = 0, \qquad \text{in } \mathbb{R}_{>0} \times \mathcal{M}, \tag{3}$$

with $V(0, \cdot) \equiv 0$.

Theorem 6.4 (Comparison)**.** Let u and v be upper and lower semicontinuous, polynomially growing viscosity sub- and supersolutions of (3). Then $u \le v$ for $x = 0$ implies that $u \le v$ everywhere in $\mathcal{O} := \mathbb{R}_{\ge 0} \times \mathcal{M}$.

Corollary 6.5 (Uniqueness)**.** The value function is the unique subexponentially growing *viscosity solution* to the dynamic programming equation (3).

Proof. By the dynamic programming principle, the value function V is a solution to (3). To obtain uniqueness, observe that, by Theorem 6.4, V, being both a sub- and a supersolution, both dominates and is dominated by any other solution. In other words, it is equal to any other solution, and thus unique. $\square$

Note that the importance of the comparison principle goes beyond the uniqueness of the solution; the principle underpins the stability property of viscosity solutions. The stability property, in turn, leads to convergence of numerical schemes to the (unique) solution [7].

7 Numerical results

The numerical results presented in this section are all obtained through policy iteration. Policy iteration is an iterative technique where one chooses a policy/control, calculates the corresponding payoff function, then updates the policy where the payoff function suggests it is profitable, and finally iterates this procedure until convergence. That the scheme does indeed converge to the value function is supported by the comparison principle in Theorem 6.4 and the uniqueness result in Corollary 6.5.

The idea implemented here is to approximate the singularity with increasingly large controls which are absolutely continuous with respect to time. In particular, let $K > 0$ be any large constant and consider control variables $L_t = \int_0^t \ell(X_t^L)\,\mathrm{d}t$, where $\ell(x) \in [0, K]$ is measurable. Then, the problem of optimizing over functions ℓ amounts to a penalization of the DPE (3) with penalization factor K.

To see that the limit of these problems gives the solution to the original problem, we begin by writing out the PDE for the approximation:

$$\min_{\ell \in [0,K]} \left(rV^K - \mathcal{L}V^K \right) + \ell \left(V_x^K - 1 \right) = 0. \tag{4}$$

By dividing by K and subtracting and adding equal terms, we reach

$$\min_{\lambda \in [0,1]} (1-\lambda) \frac{rV^K - \mathcal{L}V^K}{K} + \lambda \left(V_x^K - 1 + \frac{rV^K - \mathcal{L}V^K}{K} \right) = 0,$$

which is equivalent to

$$\min \left\{ rV^K - \mathcal{L}V^K, V_x^K - 1 + \frac{rV^K - \mathcal{L}V^K}{K} \right\} = 0.$$

Finally, letting $K \to \infty$ and using the stability property of viscosity solutions guaranteed by the comparison principle, we find that $V^K \to V$.

Due to this approximation of the state dynamics, it holds that in any given space discretization, the transition rate between states is bounded away from zero. This means that the continuous time Markov chain on the discretized space can be reduced to a discrete time Markov chain (cf., e.g., [31]). Thus, after a suitable space discretization, the problem is solved using standard methods of policy iteration that are known to converge (to the penalized problem). Let $\mathcal{D}$ be a discretization of $\mathcal{O}$ consisting of N points, and let $\mathcal{L}^{\ell}_{\mathcal{D}}$ be a corresponding discretization of $rV^K - \mathcal{L}V^K - \ell V^K_x$ from (4). Then, starting with any control ℓ_0, the policy iteration scheme with tolerance $\tau \geq 0$[7] is given by the following steps:[8]

Policy iteration algorithm

1. Compute $V_i \in \mathbb{R}^N$ such that

$$\sum_{(x',\mu')\in\mathcal{D}} \mathcal{L}^{\ell_i}_{\mathcal{D}}(x,\mu,x',\mu')V_i(x',\mu') + \ell_i = 0, \quad \forall (x,\mu) \in \mathcal{D}.$$

   Halt if $|V_i - V_{i-1}| \leq \tau$.

2. For each $(x,\mu) \in \mathcal{D}$, compute $\ell_{i+1}(x,\mu)$ according to

$$\ell_{i+1}(x,\mu) \in \operatorname*{arg\,min}_{\hat{\ell}\in[0,K]} \left(\sum_{(x',\mu')\in\mathcal{D}} \mathcal{L}^{\hat{\ell}}_{\mathcal{D}}(x,\mu,x',\mu')V_i(x',\mu') + \hat{\ell} \right).$$

3. Return to step (i).

7.1 Results and comparative statics

The scheme was implemented for the Ornstein–Uhlenbeck model

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

The resulting optimal strategies presented in Figure 5, with $k = 0.5$, $\bar{\mu} = 0.15$, and $\tilde{\sigma} = 0.1$ (left) as well as $\tilde{\sigma} = 0.3$ (right). The other parameter choices are $\sigma = 0.1$, $\rho = 0$, and $r = 0.05$. The white regions indicates dividend payments or liquidation, i.e., $V_x = 1$ and $\mathrm{d}L > 0$, whereas the black region indicates that the firm retains all its earnings, i.e., $rV - \mathcal{L}V = 0$ and $\mathrm{d}L = 0$. The figures show the optimal policy, from which the value function can be obtain by solving a linear system of equations.

Figure 6 shows the effect of changing one parameter at a time. Varying the parameters does not seem to change the qualitative properties significantly. The parameter $\tilde{\sigma}$ primarily changes the width of the band region; k and $\bar{\mu}$ affect the size and extension into the region of negative μ; σ changes the height; and finally ρ influences the shape. Note that although the free boundary is nonmonotone in $\tilde{\sigma}$ for x right below 2, it is monotone for smaller x.

[7] The policy iteration scheme halts even for $\tau = 0$.

[8] Since we solve in a finite domain, some care has to be taken at the boundaries. However, thanks to the condition given on κ, it is natural to impose a reflecting boundary along the μ-directions, provided the domain is large enough. Moreover, the optimal policy will naturally reflect at the upper x-boundary (at a cost), provided the domain is large enough to contain the no-dividend region. For these reasons, the precise choice of boundary condition is of relatively small importance, if $\mathcal{D}$ is chosen appropriately.

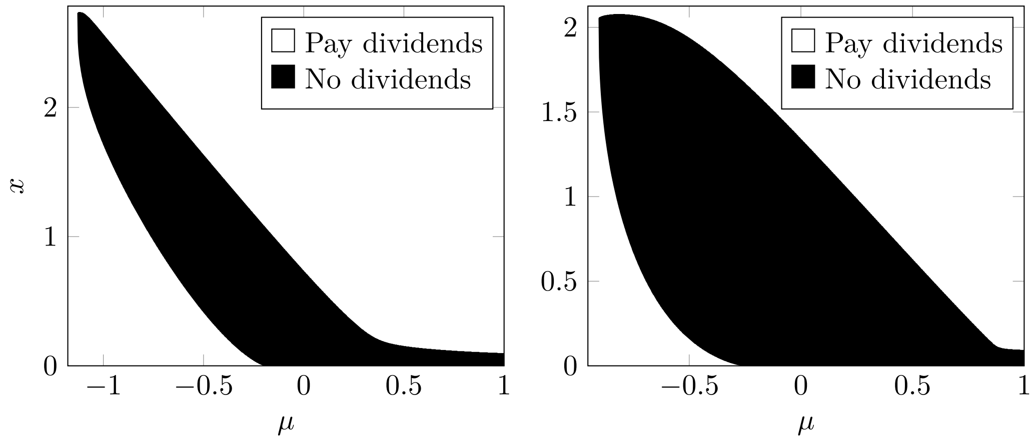

Figure 5: The black region corresponds to $\mathrm{d}L = 0$, whereas the white region corresponds to $\mathrm{d}L > 0$. The latter case is interpreted as either dividend payments or liquidation, depending on the position in the state space, see Figure 4.

8 Model extensions

8.1 Equity issuance

A firm in need of liquidity could see itself issuing equity to outside investors. In the sequel, we assume that this happens whenever desired, but at a cost. We consider two costs: one cost λ_p *proportional* to the capital received and one *fixed cost* λ_f which is independent on the amount of equity issued. Mathematically, we follow the model in [13] and write

$$\mathrm{d}X_t = \mu_t \,\mathrm{d}t + \sigma \,\mathrm{d}W_t - \mathrm{d}L_t + \mathrm{d}I_t,$$

where $I = (I_t)_{t \geq 0}$, just like L, is an adapted, increasing, RCLL control process. We allow for the costs to be μ-dependent,[9] and write $\lambda_p(\mu_t)$ and $\lambda_f(\mu_t)$. For emphasis, we keep this dependence explicit.

The figures presented in this section are generated with the Ornstein–Uhlenbeck model

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

for $k = 0.5$, $\bar{\mu} = 0.15$, and $\tilde{\sigma} = 0.3$. The other parameter choices are $\sigma = 0.1$, $\rho = 0$, and $r = 0.05$.

[9] This reflects the fact that a more profitable company (higher μ) typically has better access to financial markets.

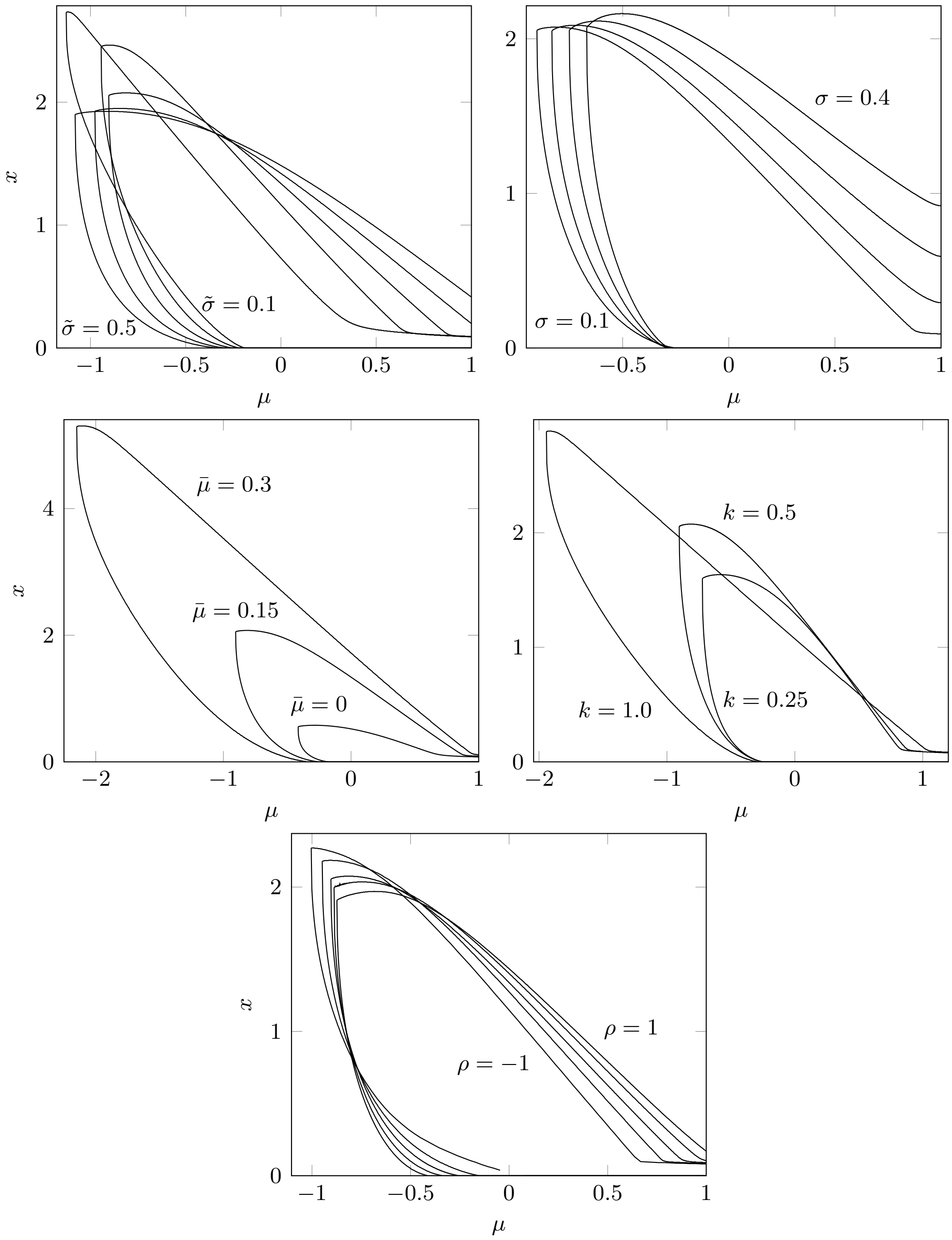

Figure 6: Comparative statics. Apart from for the parameter being varied, the chosen values were $r = 0.05$, $k = 0.5$, $\bar{\mu} = 0.15$, $\tilde{\sigma} = 0.3$, $\sigma = 0.1$, and $\rho = 0$. The parameter varied is indicated in the respective figure. The values considered were $k = 0.25, 0.5, 1.0$, $\bar{\mu} = 0.0, 0.15, 0.3$, $\tilde{\sigma} = 0.1, 0.2, 0.3, 0.4, 0.5$, $\sigma = 0.1, 0.2, 0.3, 0.4$, and $\rho = -1.0, -0.5, 0.0, 0.5, 1.0$. To address the effect of the boundary conditions, most calculations were run on a larger domain than plotted here. The lower boundary for $\rho = -1.0$ displayed signs of numerical instability around the origin and has therefore not been plotted in this region.

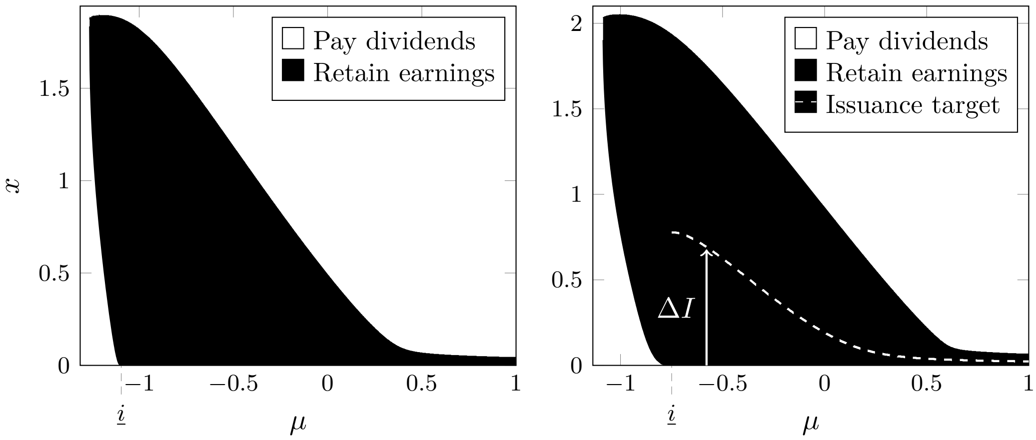

Figure 7: In the left figure, proportional issuance costs start at 34 % for low μ and decaying to 25 %. Equity is issued according to local time at the boundary $x = 0$ where otherwise ruin would occur, and no equity is issued below $\underline{i} = -1.0953$. In the right figure also fixed costs are present, starting at 0.14 decaying to 0.06. Again, equity is only issued at the boundary, and no equity is issued below $\underline{i} = -0.7493$. The dashed white line indicates the cash reserve level after issuance, i.e., how much equity was issued.

8.1.1 Proportional issuance costs

If the costs are purely proportional, i.e., $\lambda_f = 0$, the payoff corresponding to any two controls L and I is

$$J(x, \mu; L, I) = \mathbb{E}\left[\int_0^{\theta(L,I)} e^{-rt}\, \mathrm{d}\big(L - (1 + \lambda_p(\mu_t))I\big)_t\right],$$

where $\theta(L, I)$ is the first time the process X becomes negative. In this case, the DPE bears great resemblance to that of the original model, since the issuance simply has the opposite effect of dividend payments:

$$\min\{rV - \mathcal{L}V, \quad V_x - 1, \quad 1 + \lambda_p(\mu) - V_x\} = 0.$$

The interpretation is that the state space consists of three different regions defined by the optimal action: pay dividends, issue equity, or doing neither. Equity is thus issued whenever $V_x(x, \mu) = 1 + \lambda_p(\mu)$. This means that issuance occurs whenever the marginal value is equal to the marginal cost.

Since issuance is costly and can be done at any time, it is optimal to only issue equity at points where ruin would otherwise be reached, i.e., where $x = 0$. However, whether to do so at the boundary depends on the current cash flow rate. Indeed, as seen in Figure 7, equity is only issued when the cash flow rate is above a certain level, below which we still see the band structure of the original problem.

8.1.2 Fixed issuance costs

On the other hand, if the fixed cost is nonzero, we assume, without loss of generality, that

$$I_t = \sum_{k=1}^{\infty} i_k 1_{\{t \geq \tau_k\}},$$

for some strictly increasing sequence of stopping times τ_k and positive $\mathcal{F}_{\tau_k}$-measurable random variables i_k. The stopping times are interpreted as issuance dates, and the random variables as the issued equity. If I_t can not be written on this form, it means it at some point has infinite issuance frequency, which would, because of the fixed cost, come at infinite cost. This form is therefore a natural restriction, and the corresponding payoff functional is given by

$$J(x,\mu;L,I) = \mathbb{E}\left[\int_0^{\theta(L,I)} e^{-rt}\,\mathrm{d}L_t - \sum_{k=1}^{\infty} e^{-r\tau_k}(\lambda_f(\mu_t) + \lambda_p(\mu_t)i_k)1_{\{\tau_k < \theta(L,I)\}}\right].$$

The value function given by this problem is then the solution to the following nonlocal DPE:

$$\min\left\{rV - \mathcal{L}V, \quad V_x - 1, \quad V(x,\mu) - \sup_{i\geq 0}\Big(V(x+i,\mu) - \lambda_p i - \lambda_f\Big)\right\}.$$

The last conditions states that the value at any given point is at least equal to the value in any point after issuance less the issuance costs.

Just like for proportional costs, issuance optimally only occurs at the boundary. However, with fixed costs, the amount of equity issued is now larger in order to avoid incurring another fixed cost soon in the future. The magnitude is presented as the issuance target in Figure 7. Note that, the numerical method employed in the fixed cost case can be interpreted as the limit of the issuance structure in [20] when the investor arrival rate tends to infinity. The precise scheme is presented for another problem in [3].

By letting the fixed cost grow sufficiently fast in $-\mu$, one can obtain substantially different issuance policies. As shown in Figure 8, such structure can have a wave-like shape, not dissimilar to the shape of the continuity region. This seems to indicate two factors at play: Either one issues equity as a last resort at $x = 0$, or at an earlier time in fear of higher issuance costs in the future. Moreover, in this regime, the target points no longer constitute a continuous like, but instead has a jump discontinuity even to the side of the *wave*.

8.2 Gambling for resurrection

As seen in Figure 5, the value function is not necessarily concave in the x-direction. It is sometimes argued that concavity is desirable, because of the possibility to enter a (fair) speculative strategies and thus receiving the average of surrounding points.[10] Such

[10] One example of such behavior by FedEx is described by Frock [16, Chapter 18], one of the firm's co-founders.

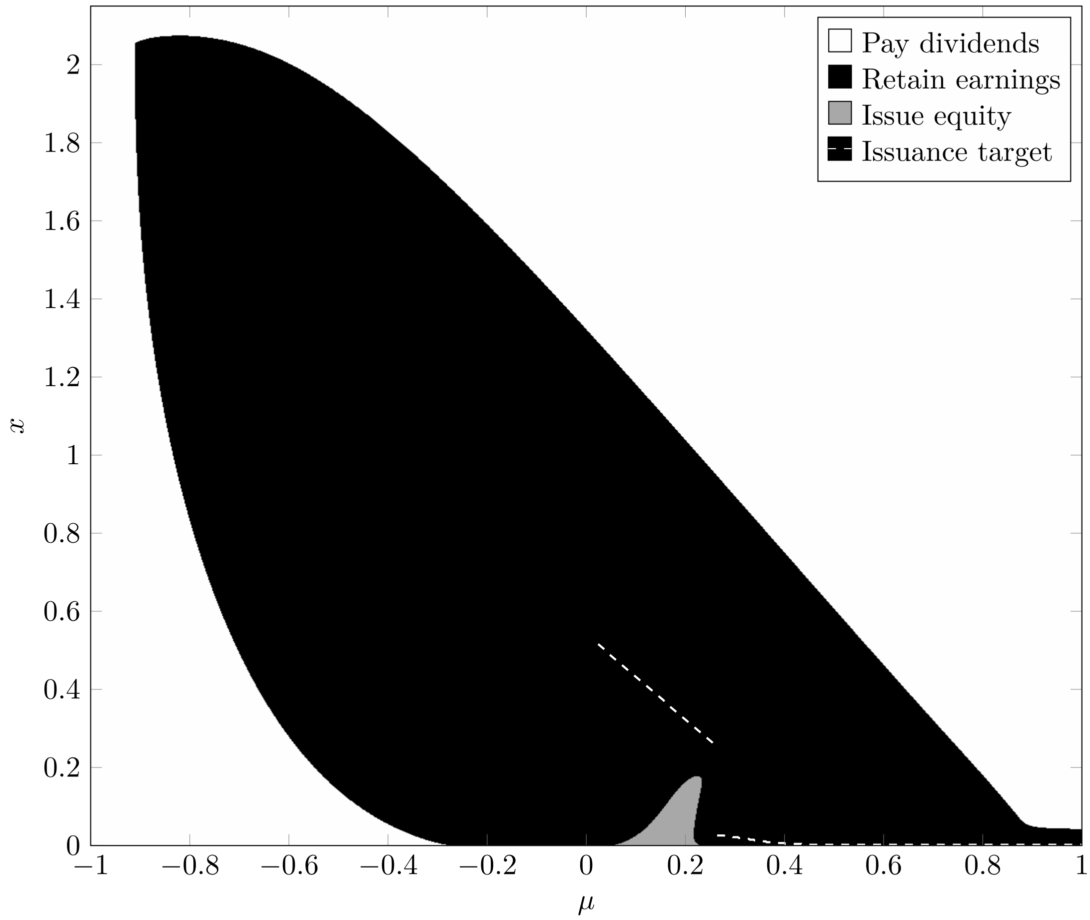

Figure 8: When issuance costs are sufficiently high for negative μ, it is optimal to issue equity away from the boundary $x = 0$.

possibilities can be incorporated into the model, by considering another control process $G = (G_t)_{t\geq 0}$ and cash reserves given by

$$X_t = x + \int_0^t \mu_t \,\mathrm{d}t + \sigma W_t - L_t + G_t,$$

for processes of the form

$$G_t = \sum_{k=1}^{\infty} g_k 1_{\{t\geq \tau_k\}},$$

for some sequence of predictable stopping times and $\mathcal{F}_{\tau_k}$-measurable random variables g_k[11] satisfying $\mathbb{E}[g_k] = 0$ and $X_{\tau_k -} + g_k \geq 0$ P-a.s. This leads to the DPE

$$\min\{rV - \mathcal{L}V, \quad V_x - 1, \quad -V_{xx}\} = 0,$$

[11]One possible interpretation of g_k is to think of it as a forward contract. More precisely, it should be interpreted as the limit when the forward contract can be entered with arbitrarily short maturity.

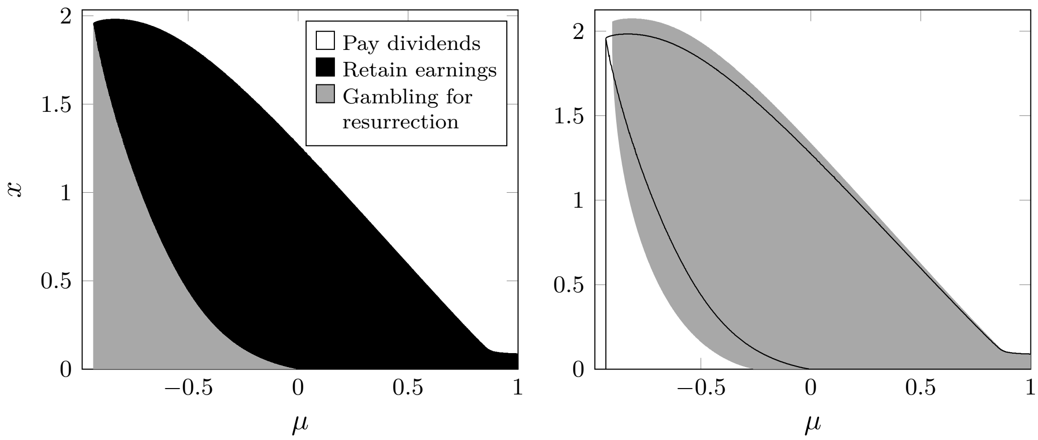

Figure 9: The figure on the left shows the model which allows gambling for resurrection. The figure on the right shows the model without gambling in gray, overlaid by the free boundaries of the gambling model in solid lines. When gambling is allowed, voluntary liquidation is no longer optimal in the 'band region'. This is because entering a 'lottery' gives a chance of reaching a higher point in the no activity region, thus concavifying the problem.

from which the last conditions makes it directly clear that the value function is now concave. However, as seen in Figure 9, this is not the concave envelope in the x-direction, since the *retain earnings region* shifts in the μ-direction. The cause of this is that gambling occurs in what otherwise would have been the no-dividend region, thus affecting the solution in the μ-direction through the elliptic operator $\mathcal{L}$.

8.3 Credit lines

If the firm is sufficiently profitable (μ large enough), it could be granted a credit line by a bank, whereby the cash reserves $X = (X_t)_{t\geq 0}$ are allowed to be negative up to a certain threshold $\underline{x}(\mu)$, below which bankruptcy occurs. Suppose the interest on these are $\rho_- \geq 0$ and define

$$\rho(x) = \begin{cases} 0, & x \geq 0 \\ \rho_-, & x < 0 \end{cases}.$$

The cash balance (reserves) process can then be written as

$$\mathrm{d}X_t = (\rho(X_t)X_t + \mu_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t - \mathrm{d}L_t.$$

Although the credit line has the effect of shifting the dividend region downwards, closer to the new bankruptcy boundary, as seen in Figure 10, the effect of the interest rates ρ seems to be minimal. Indeed, with interest rates of order 1 % and optimal cash levels of order 1, the effecting increase in drift is then also of order 1 %. This effect is therefore generally small in comparison to the magnitude of the cash flow rate.

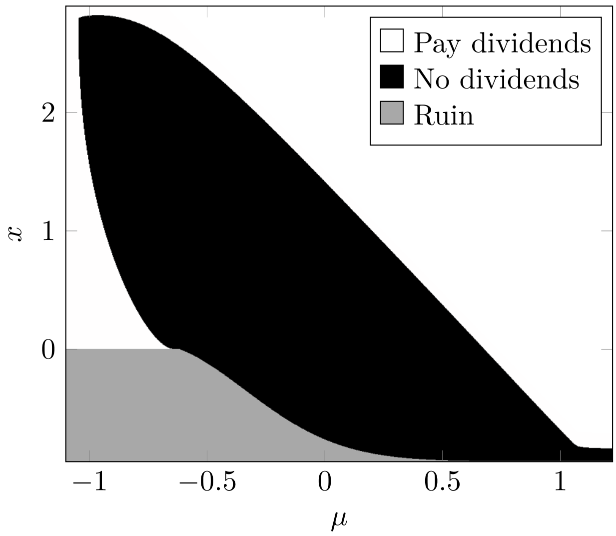

Figure 10: The credit line shifts the upper boundary downwards, but the main characteristics of the model are otherwise retained.

8.4 $\sigma = 0$

Arguably, the case $\sigma = 0$ is also of interest, and the main difference appears in the proof of continuity, where bankruptcy does not happen at the boundary for all values of μ. Continuity at the ruin points is what is necessary to bootstrap the proof of continuity elsewhere in Theorem 6.3. Therefore, continuity in points $(x, \mu) \in \{0\} \times (\mathcal{M} \cap \mathbb{R}_{\leq 0})$ is sufficient. For $\mu < 0$, the proof of continuity at the boundary follows the same steps. Continuity thus hinges on continuity at the origin, which only holds if the ruin is reached instantaneously, i.e., $\theta^{(0,0)}(0) = 0$. In the general case, proving this turns out to be a nontrivial task without additional assumptions.

9 Proofs

This section is dedicated to the proofs of of previous sections. We begin with a result that is needed in multiple proofs. It is proven under slightly stronger assumptions, which turn out to satisfied without loss of generality where the lemma is needed.

Lemma 9.1. If, in addition to Assumption 6.1, $\inf \mathcal{M} > -\infty$ or $\mu \in \mathcal{O}(\kappa(\mu))$ as $\mu \to -\infty$, there exists a sublinearly growing function H and a constant C such that, for any stopping time τ,

$$\mathbb{E}\left[\max_{0 \leq t \leq \tau} |\mu_t|\right] \leq C\mathbb{E}[H(\tau)].$$

Remark 9.2. For $\tilde{\sigma}$ identically zero for large $|\mu|$, this condition is satisfied thanks to the assumptions on κ pushing the process 'inwards'. Moreover, the theorem assumes that $\tilde{\sigma} > 0$, but this can be partially dropped by choosing the c sufficiently large in the proof.

Proof. We will use the result by Peskir [30] to obtain a function H which is sublinearly growing. For some $c \in \mathcal{M}^{\mathrm{o}}$, let

$$S'(\mu) = \exp\left(-2\int_c^\mu \frac{\kappa(\nu)}{\tilde{\sigma}(\nu)^2}\,\mathrm{d}\nu\right)$$

and

$$m(\mathrm{d}\nu) = \frac{2\,\mathrm{d}\nu}{S'(\nu)\tilde{\sigma}(\nu)^2}.$$

Finally, define

$$F(\mu) = \int_c^\mu m((c,\nu])S'(\nu)\,\mathrm{d}\nu.$$

Note that by the assumptions of the statement, κ and $\tilde{\sigma}$ behave analogously for large positive and negative μ, so, without loss of generality, we may consider only positive values of μ. In particular, for large μ, $S'(\mu)$ grows as $\exp(a\mu^\gamma)$ for some $a > 0$ and $\gamma \geq 1$, and we are done if we can verify the following condition:

$$\sup_{\mu > c}\left(\frac{F(\mu)}{\mu}\int_\mu^\infty \frac{\mathrm{d}\nu}{F(\nu)}\right) < \infty.$$

All involved functions are continuous, so we are done if it has a finite limit (or is negative) as $\mu \to \infty$. L'Hôpital's rule yields the fraction

$$\frac{\frac{\mathrm{d}}{\mathrm{d}\mu}\int_\mu^\infty \frac{\mathrm{d}\nu}{F(\nu)}}{\frac{\mathrm{d}}{\mathrm{d}\mu}\frac{\mu}{F}} = \frac{F(\mu)}{\mu F'(\mu) - F(\mu)}.$$

If the denominator were bounded from above, Grönwall's inequality would imply linear growth of F, which contradicts the growth of S'. Hence, the expression is either negative (and we are done), or we may use l'Hôpital's rule again:

$$\frac{F'(\mu)}{\mu F''(\mu)} = \frac{S'(\mu)m((c,\mu])}{\frac{2\mu}{\tilde{\sigma}(\mu)^2} + \mu S''(\mu)m((c,\mu])} \xrightarrow{\mu\to\infty} 0,$$

since $S'(\mu)/\mu S''(\mu) = \tilde{\sigma}(\mu)^2/(-2\kappa(\mu)\mu) \to 0$. Thus, with $H = F^{-1}$, [30] allows us to conclude that

$$\mathbb{E}\left[\max_{0\leq s\leq\tau}|\mu_s|\right] \leq C\mathbb{E}[H(\tau)], \tag{5}$$

for some constant C and any stopping time τ. In particular, for $\tau = t$, the expression is finite and sublinearly growing in t. $\square$

9.1 Dividends of arbitrary sign

Lemma 9.3. When dividends can be of arbitrary sign, the optimal policy for shareholders is to immediately distribute the initial cash reserves at $t = 0$, and to maintain them at zero forever.

Proof. Suppose L be any strategy for which $X_t^L \geq 0$ until some (liquidation) time τ. Then define

$$L'_t = L_t + X_t^L.$$

Since X_t^L is nonnegative until τ, it is clear that $X_t^{L'} = 0$ and that $L'_t \geq L_t$ for $t \leq \tau$. Hence, L' is admissible whenever L is, and it also produces a higher payoff. □

Theorem 3.1. When dividends can be of arbitrary sign, the optimal policy for shareholders is to immediately distribute the initial cash reserves at $t = 0$, and to maintain them at zero forever by choosing $\mathrm{d}L_t = \mu_t \,\mathrm{d}t + \sigma \,\mathrm{d}W_t$. If $\mathcal{M}$ has no lower bound, there exists a μ^* such that the firm is liquidated when profitability falls below the threshold μ^*: $\tau = \inf\{t > 0 : \mu_t \leq \mu^*\}$ is a maximizer.

Proof. If $\mathcal{M}$ has no lower bound, but κ is not growing linearly as $\mu \to -\infty$, consider instead of $(\mu_t)_{t\geq 0}$ another process with the same $\tilde{\sigma}$, but which also fulfills this growth condition. The corresponding value function dominates our original one, so it is enough to prove it in this case.

Setting $\mathrm{d}L_t = \mu_t \,\mathrm{d}t + \sigma \,\mathrm{d}W_t$ until a stopping time τ yields

$$J(x, \mu; L) = x + \mathbb{E}\int_0^\tau e^{-rt}\mu_t \,\mathrm{d}t + \mathbb{E}\int_0^\tau e^{-rt}\sigma \,\mathrm{d}W_t.$$

Since the last term is zero, the value function is obtained by maximizing over τ:

$$V(x, \mu) = x + \sup_\tau \mathbb{E}\int_0^\tau e^{-rt}\mu_t \,\mathrm{d}t = x + \tilde{V}(\mu).$$

We now try to find a point in which $\tilde{V}$ is 0. Consider the equation

$$\min\left\{-\nu + r\phi - \kappa(\nu)\phi' - \frac{1}{2}\tilde{\sigma}(\nu)^2\phi'', \phi\right\} = 0, \tag{6}$$

and suppose it has a solution which never touches 0, i.e, $\phi > 0$. Then,

$$\phi(\mu) = \mathbb{E}\int_0^\infty e^{-rt}\mu_t \,\mathrm{d}t = \int_0^\infty e^{-rt}\mathbb{E}[\mu_t] \,\mathrm{d}t.$$

Using Itô's formula, the preceding martingale result, the growth bounds on κ, and (5), we obtain

$$\mathbb{E}[\mu_t] \leq \mu + tC(1 + F^{-1}(t)),$$

for some new constant C. Hence,

$$\phi(\mu) \leq \int_0^\infty e^{-rt}\left(\mu + tC(1 + F^{-1}(t))\right) \mathrm{d}t \leq \frac{\mu}{r} + C',$$

for yet another constant C'. Thus, $\phi(\mu) \to -\infty$ as $\mu \to -\infty$, which contradicts that $\phi \geq 0$. We conclude that a solution ϕ must indeed touch 0.

Finally, we are done if $\tilde{V}$ satisfies the dynamic programming equation (6). By Lemma 9.1 and [26], the optimal stopping time is the hitting time of $A_0 = \{\mu : \tilde{V}(\mu) = 0\} \neq 0$. Hence, the function is smooth everywhere, except possibly at $\mu^* := \sup A_0$. However, since $\tilde{\sigma}$ never vanishes,[12] an argument analogous to in the proof of Theorem 9.9 yields continuity also at μ^*, from which (6) can be derived. □

9.2 Semi-deterministic problem

We first need some lemmata, and separate the result into two parts depending on if $\mu \geq \bar{\mu}$ or if $\mu \in (0, \bar{\mu})$.

Lemma 9.4. If $V(2x, \mu) \leq 2V(x, \mu)$ for every $x > 0$, then $x \mapsto V(x, \mu)$ is concave.

Proof. Consider the three strategies L^x, $\hat{L}$, and L^{x+2h} for the starting points (x, μ), $(x + h, \mu)$, and $(x + 2h, \mu)$, respectively. Assume L^x and L^{x+2h} are ε-optimal for some $\varepsilon > 0$, and define $\hat{L}$ as the control equal to $\frac{L_t^x + L_t^{x+2h}}{2}$ until $\theta^x(L^x)$, and thereafter ε-optimal. Note that

$$X_{\theta^x(L^x)}^{(x+h,\mu),\hat{L}} = X_{\theta^x(L^x)}^{(x+2h,\mu),L^{x+2h}}/2.$$

Therefore, by the dynamic programming principle, the definition of $\hat{L}$, and the assumption,

$$\begin{aligned}
V(x+2h,\mu) - \varepsilon &\leq \mathbb{E}\left[\int_0^{\theta^x(L^x)} e^{-rt}\,\mathrm{d}L_t^{x+2h} + e^{-r\theta^x(L^x)} V\left(X_{\theta^x(L^x)}^{(x+2h,\mu),L^{x+2h}}, \mu_{\theta^x(L^x)}\right)\right] \\
&\leq 2\mathbb{E}\left[\int_0^{\theta^x(L^x)} e^{-rt}\,\mathrm{d}\hat{L}_t + e^{-r\theta^x(L^x)} V\left(\frac{X_{\theta^x(L^x)}^{(x+2h,\mu),L^{x+2h}}}{2}, \mu_{\theta^x(L^x)}\right)\right] \\
&\quad - V(x,\mu) \\
&\leq 2\mathbb{E}\left[\int_0^{\theta^{x+h}(\hat{L})} e^{-rt}\,\mathrm{d}\hat{L}_t + \varepsilon\right] - V(x,\mu) \\
&\leq 2V(x+h,\mu) - V(x,\mu) + 2\varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this shows that V is midpoint concave in x, and hence also concave. □

Lemma 9.5. The value function is concave in x for $\mu \geq \bar{\mu}$.

Proof. Fix x and $\mu \geq \bar{\mu}$. Denote by L^{2x} an ε-optimal strategy from starting at $(2x, \mu)$, i.e.,

$$V(2x,\mu) - \varepsilon \leq J(2x,\mu; L^{2x}),$$

[12] If it does vanish somewhere, the same result could be obtained by considering the limit of $\tilde{\sigma} + \varepsilon$ for $\varepsilon \searrow 0$ instead.

and let $\theta^x(L^{2x})$ be the time of ruin when starting in (x, μ) and following the dividend strategy L^{2x}.

Then,

$$V(x,\mu) \geq J(x,\mu; L^{2x}),$$

and therefore

$$V(2x,\mu) - \varepsilon \leq V(x,\mu) + \mathbb{E}\left[e^{-r\theta^x(L^{2x})} V\left(x, \mu_{\theta^x(L^{2x})}\right)\right] \leq 2V(x,\mu),$$

since $\mu \geq \bar{\mu}$. The statement follows from Lemma 9.4, since $\varepsilon > 0$ is arbitrary. □

The second case is more technical so in order to prove concavity for $\mu \in (0, \bar{\mu})$, we need the following lemmata.

Lemma 9.6. For any x, μ, and δ,

$$V(x, \mu+\delta) \leq V(x+\delta/k, \mu).$$

Proof. Denote by $(\mu_t^\mu)_{t\geq 0}$ the process starting in μ. Then $\mu_t^{\mu+\delta} = \mu_t^\mu + \delta e^{-kt}$. Hence, for any dividend policy L,

$$\begin{aligned} x + \int_0^t \mu_s^{\mu+\delta}\,\mathrm{d}s + \sigma W_t - L_t &= x + \int_0^t \mu_s^\mu\,\mathrm{d}s + \frac{\delta}{k}(1-e^{-kt}) + \sigma W_t - L_t \\ &\leq (x+\delta/k) + \int_0^t \mu_s^\mu\,\mathrm{d}s + \sigma W_t - L_t, \end{aligned}$$

which means that any dividend strategy admissible for $(x, \mu+\delta)$ is also admissible for $(x+\delta/k, \mu)$, from which the statement follows. □

For a continuous function

$$f : [0,\infty) \to \mathbb{R}$$

as well as $x \geq 0$, $h \geq 0$, and $y \geq x+h$, define

$$I_f(x,y,h) := f(x) - f(x+h) - f(y-h) + f(y).$$

Lemma 9.7. The function f is concave if and only if there is a $h_0 > 0$ so that $I_f(x,y,h) \leq 0$ for all $x \geq 0$, $y \geq x+h$ and $h \in [0, h_0]$.

Lemma 9.8. For any continuous function f, suppose there exist $x^* > 0$ and $\alpha \geq 0$ so that

1. f is concave on $[0, 2x^*]$;
2. $I_f(x,y,h) \leq \alpha$ for all $x^* \leq x$, $y \geq x+h$, and $h \in [0, x^*]$.

Then $I_f(x,y,h) \leq \alpha$ for all $0 \leq x$, $y \geq x+h$, and $h \in [0, x^*]$.

Proof. Consider the case $x \in [0, x^*]$ and $y > 2x^*$. In all other cases, by hypothesis, we have $I_f(x, y, h) \leq \alpha$. Since f is concave on $[0, 2x^*]$ and $h \leq x^*$, it holds that $x + h \leq 2x^*$ and $I_f(x, x^* + h, h) \leq 0$. Therefore,

$$f(x) - f(x + h) \leq f(x^*) - f(x^* + h).$$

Using this, we obtain

$$I_f(x, y, h) \leq f(x^*) - f(x^* + h) - f(y - h) + f(y) = I_f(x^*, y, h),$$

for any $y > 2x^*$ and $h \leq x^*$, since $x + h \leq 2x^* < y$. Hence, by hypothesis, $I_f(x^*, y, h) \leq \alpha$. ☐

Theorem 5.1. The value function is concave in x for $\mu > 0$.

Proof. By Lemma 9.5, it remains to consider only the case $\mu \in (0, \bar{\mu})$. We know that V is a super-solution to the HJB equation. If there exists a viscosity test function φ in some point (x, μ), the HJB equation implies the inequalities

$$\begin{aligned}
\frac{\sigma^2}{2}\varphi_{xx} &\leq rV - \mu\phi_x - k(\bar{\mu} - \mu)\phi_\mu \\
&\leq rV - \mu\varphi_x \\
&\leq rV - \mu.
\end{aligned}$$

We conclude that for every $\delta^\mu \in (0, \bar{\mu})$ there exists a $\delta^x > 0$ such that for any point in $(0, \delta^x) \times (\delta^\mu, \bar{\mu})$ and any test function ϕ, the second derivative $\phi_{xx} < 0$.

By proof by contradiction, we show that this condition on the test functions ensures that also V is concave in x in $(0, \delta^x) \times (\delta^\mu, \bar{\mu})$. Assume that V is not concave. Then there exists two points (x_1, ν) and (x_2, ν) such that if

$$f(x) = \frac{x_2 - x}{x_2 - x_1} V(x_1, \nu) + \frac{x - x_1}{x_2 - x_1} V(x_2, \nu),$$

then $x \mapsto V(x, \nu) - f(x)$ has a local minimum strictly smaller than 0, attained in (x_1, x_2). Denote by x_0 such a minimizer and let

$$g(x, \mu) = f(x) - f(x_0) + V(x_0, \nu) - a|\mu - \nu|,$$

where $a > f'/k$ and f' is the slope of the straight line given by f. We construct a set on which the minimum of $V - g$ is attained on the interior. Let

$$\begin{aligned}
N = \{(x_0 + \delta_1, \nu + \delta_2) : \delta_1 &\in [x_1 - x_0, x_2 - x_0], \\
\nu + \delta_2 &\in [\delta^\mu, \bar{\mu}], \\
-\delta_1 + \frac{|\delta_2|}{k} &\leq x_0 - x_1, \\
\delta_1 + \frac{|\delta_2|}{k} &\leq x_2 - x_0\}.
\end{aligned}$$

Note that $\min_{[x_1,x_2]}(V(\cdot,\nu)-g(\cdot,\nu))=(V-g)(x_0,\nu)=0$. Thus $V-g$ on N attains it's minimum in (x_0,ν), since

$$\begin{aligned}V(x_0\pm\delta_1,\nu-\delta_2)-&g(x_0\pm\delta_1,\nu-\delta_2)\\ &\geq V(x_0\pm\delta_1-\delta_2/k,\nu)-g(x_0\pm\delta_1-\delta_2/k,\nu)\\ &\quad-(f(x_0\pm\delta_1)-f(x\pm\delta_1-\delta_2/k)-a\delta_2)\\ &\geq -f'\delta_2/k+a\delta_2+(V-g)(x_0,\nu)\\ &\geq (V-g)(x_0,\nu).\end{aligned}$$

The case $\mu>\nu$ is trivial as V is increasing in μ, and g is symmetric. Furthermore, the calculation above shows that the minimum can only be attained on the line where $\mu=\nu$. Due to the fact that $[x_1,x_2]\times\{\nu\}$ only intersects ∂N in (x_1,ν) and (x_2,ν), the minimum is not attained at the boundary.

Define

$$\zeta=\inf_{\partial N}(V-g)>0.$$

The function g does not have the regularity necessary to use it as a viscosity test function. Instead we consider the mollifications $\{g^\epsilon\}_{\epsilon>0}$. Let

$$\xi(\epsilon)=\sup_N|g-g^\epsilon|\xrightarrow{\epsilon\to 0}0.$$

There must exist an ϵ such that $V-g^\epsilon$ attains its minimum on the interior of N, for otherwise

$$\zeta=\inf_{\partial N}(V-g)\leq\inf_{\partial N}(V-g^\epsilon)+\xi(\epsilon)\leq(V-g^\epsilon)(x_0,\nu)+\xi(\epsilon)\xrightarrow{\epsilon\to 0}0,$$

contradicting that $\zeta>0$.

Finally, we observe that the linearity in x is preserved under mollification. Therefore $\partial_{xx}g^\epsilon=0$ so the choice $\varphi=g$ leads to a contradiction. Hence, V must be concave in $(0,\delta^x)\times(\delta^\mu,\bar\mu)$.

To show that V is also concave for larger x, we apply Lemma 9.8 with $x^*=\delta^x/2$. Let

$$\alpha=\alpha_\epsilon:=\sup\{I_{V(\cdot,\mu)}(x,y,h)\ :\ \mu\in(\delta^\mu,\bar\mu]\ 0<h\leq x^*\leq x,y\geq x+h\}.$$

Since the process $\mu^{\bar\mu}$ is constant, the function $V(x,\bar\mu)$ coincides with the value function found in [22, 33], hence bounded by $x+b$ for some b. Therefore, since V is growing in μ, we have that for any $z>0$, $z<V(z,\mu)<z+b$, from which it follows that α is finite.

By the argument above, $V(\cdot,\mu)$ is concave on $[0,2x^*]$ for $\mu\in(\delta^\mu,\bar\mu]$. Thus, by the previous lemma,

$$I_{V(\cdot,\mu)}(x,y,h)\leq\alpha,\quad\forall\ 0<h\leq x^*,0\leq x,y\geq x+h.$$

Fix $\mu\in(\delta^\mu,\bar\mu]$, $0<h\leq x^*\leq x$ as well as $y\geq x+h$ and choose ε-optimal controls L^x, L^y so that

$$V(x,\mu)-\varepsilon=J(x,\mu;L^x),\quad V(y,\mu)-\varepsilon=J(y,\mu;L^y).$$

We assume that the optimal trajectory starting from (x, μ) always stays below the one from (y, μ). Otherwise we set them equal to each other if they ever become equal. It is standard that the resulting strategy is also ε-optimal.

Now construct the stopping time T according to

$$T = \inf\{t > 0 : x^* + \sigma W_t \notin (0, 2x^*)\}.$$

Then $T > 0$ P-a.s., and since no dividends are paid below $2x^*$, it is smaller than the time of ruin for any optimal process starting above x^*.

Let $\theta = T \wedge \tau$, where τ is the stopping time at which the difference between the trajectory of the controlled processes starting (y, μ) and (x, μ) is equal to h, (note that it starts larger than h). Set

$$X_t^* := x + \int_0^t \mu_u du + \sigma W_t - L_t^x, \quad Y_t^* := y + \int_0^t \mu_u du + \sigma W_t - L_t^y.$$

Then, on $[0, \theta]$, $0 \leq X_t^* \leq Y_t^* + h$. This implies that L^y is admissible on the interval $[0, \theta]$ starting from $y - h$. Hence, by dynamic programming,

$$V(y - h, \mu) \geq \mathbb{E}\left[\int_0^\theta e^{-rt}\, \mathrm{d}L_t^y + e^{-r\theta} V(Y_\theta^* - h, \mu_\theta)\right].$$

We also have

$$V(x + h, \mu) \geq \mathbb{E}\left[\int_0^\theta e^{-rt}\, \mathrm{d}L_t^x + e^{-r\theta} V(X_\theta^* + h, \mu_\theta)\right].$$

Hence, the ε-optimality of L^x and L^y implies that

$$\begin{aligned} I_{V(\cdot,\mu)}(x, y, h) - 2\varepsilon &\leq \mathbb{E}\left[e^{-r\theta}\left(V(X_\theta^*, \mu_\theta) - V(X_\theta^* + h, \mu_\theta) - V(Y_\theta^* - h, \mu_\theta) + V(Y_\theta^*, \mu_\theta)\right)\right] \\ &= \mathbb{E}\left[e^{-r\theta}\left(I_{V(\cdot,\mu_\theta)}(X_\theta^*, Y_\theta^*, h)\right)\right]. \end{aligned}$$

Notice that, $Y_\tau^* = X_\tau^* + h$ (recall that τ is defined above as the first time this equality holds). Hence,

$$I_{V(\cdot,\mu_\tau)}(X_\tau^*, Y_\tau^*, h) = 0.$$

Therefore,

$$\begin{aligned} \mathbb{E}\left[e^{-r\theta}\left(I_{V(\cdot,\mu_\theta)}(X_\theta^*, Y_\theta^*, h)\right)\right] &\leq \mathbb{E}\left[e^{-rT}\left(I_{V(\cdot,\mu_T)}(X_T^*, Y_T^*, h)\right)\right] \\ &\leq \mathbb{E}\left[e^{-rT}\ \alpha\right] \leq \Lambda\alpha, \end{aligned}$$

where $\Lambda = \mathbb{E}[e^{-rT}] < 1$.

Letting $\varepsilon \searrow 0$, we have shown that for any $\mu \in (\delta^\mu, \bar{\mu}]$, $0 < h \leq x^* \leq x$ as well as $y \geq x + h$,

$$I_{V(\cdot,\mu)}(x, y, h) \leq \Lambda\alpha.$$

Since α is defined as the sup over the left hand side, and since $\Lambda < 1$, we conclude that $\alpha = 0$.

In view of the previous lemma, this proves that $I_{V(\cdot,\mu)}(x, y, h) \leq 0$ for all $x \geq 0$, $h \in [0, x^*]$, $y \geq x + h$, and $\mu \in (\delta^\mu, \bar{\mu}]$. Since δ^μ can be chosen arbitrarily close to 0, we conclude that V is concave in x for all $\mu \in (0, \bar{\mu}]$. □

9.3 Liquidation threshold

Theorem 6.2. If $\mathcal{M}$ has no lower bound, there exists a value μ^* such that it is optimal to liquidate immediately whenever $\mu \leq \mu^*$, i.e. $V(x, \mu) \equiv x$.

Proof. Until the time of ruin $\theta(L)$, $L_t \leq x + \int_0^t \mu_s \, \mathrm{d}s + \int_0^t \sigma \, \mathrm{d}W_s$. Hence, by stochastic integration by parts,

$$V(x, \mu) = \sup_L \mathbb{E} \int_0^{\theta(L)} e^{-rt} \, \mathrm{d}L_t \leq x + \sup_L \mathbb{E} \int_0^{\theta(L)} e^{-rt} \mu_t \, \mathrm{d}t + \mathbb{E} \int_0^{\theta(L)} e^{-rt} \sigma \, \mathrm{d}W_t.$$

First observe that the last term is equal to 0. Then, since the second term is smaller than or equal to $\tilde{V}(\mu)$, the result is a direct consequence of Theorem 3.1. □

9.4 Continuity

When proving the continuity of the value function, we need a weak form of a dynamic programming *inequality.* More precisely, for any control L and stopping time τ with values between 0 and $\theta(L)$,

$$E \left[\int_\tau^{\theta(L)} e^{-rt} \, \mathrm{d}L_t \middle| \mathcal{F}_\tau \right] \leq e^{-r\tau} \bar{V}(X_\tau, \mu_\tau), \quad P\text{-a.s.}, \tag{7}$$

where $\bar{V}$ denotes the upper-semicontinuous envelope of V. The reason we need to use $\bar{V}$ and not V is that we do not know a priori whether V is measurable. This inequality very much related to the weak dynamic programming principle [10] which also establishes a similar inequality in the other direction. However, (7) is more primitive than the inequality in the other direction.

These measurability issues are arguably the most notable obstacles in establishing the dynamic programming principle. However, for continuous value functions, proofs of the dynamic programming principle are well known [15]. For the general case, we once again refer to [24, 25].

In this section we establish the continuity of the value function, from which the dynamic programming principle then follows.

Theorem 9.9. The value function is continuous at $x = 0$.

Proof. Let $\{(x^n, \mu^n)\}_{n \geq 1}$ be a sequence converging to $(0, \mu^\infty)$. Without loss of generality, assume $x^n > 0$ converges monotonically to 0. Since $V \equiv 0$ at $x = 0$, it is sufficient to consider monotonically decreasing sequences in μ, by monotonicity in μ. For simplicity, also assume that $x^1 < 1$ and $|\mu^1 - \mu^\infty| < 1$.

Let τ be any strictly positive, bounded stopping time such that for $t \leq \tau$, $\mu_t^{\mu^1} \leq |\mu^\infty|+1$ and $X_t^{(x^1,\mu^1),0} \leq 1$, P-a.s for the uncontrolled process corresponding to $L = 0$. Hence, we also have $X^{(x^n,\mu^n),L} \leq X^{x^1,\mu^1,0} \leq 1$ for all $n \in \mathbb{N}$ and dividend processes L.

Starting with the definition of the value function, one obtains

$$\begin{aligned}
V(x^n, \mu^n) &= \sup_L E\left[\int_0^{\tau\wedge\theta_n(L)} e^{-rt}\,\mathrm{d}L_t + 1_{\{\tau<\theta_n(L)\}}\int_{\tau\wedge\theta_n(L)}^{\theta_n(L)} e^{-rt}\,\mathrm{d}L_t\right] \\
&\leq \sup_L E\left[\int_0^{\tau\wedge\theta_n(L)} e^{-rt}\,\mathrm{d}L_t + 1_{\{\tau<\theta_n(L)\}}\int_0^{\theta_n(L)} e^{-rt}\,\mathrm{d}L_t\right] \\
&\leq \sup_L E\left[x^n + (|\mu^\infty|+1)(\tau\wedge\theta_n(L)) + V(1, |\mu^\infty|+1)1_{\{\tau<\theta_n(L)\}}\right] \\
&\leq x^n + (|\mu^\infty|+1)E[\tau\wedge\theta_n(0)] + V(1, |\mu^\infty|+1)P[\tau<\theta_n(0)].
\end{aligned}$$

Now make the observation that,

$$\{\tau < \theta_{n+1}(0)\} \subseteq \{\tau < \theta_n(0)\}.$$

Since $\sigma > 0$, $\theta_n(0) \to 0$ P-a.s., and therefore

$$\lim_{n\to\infty} P[\tau < \theta_n(0)] = P\left(\lim_{n\to\infty}\{\tau < \theta_n(0)\}\right) = P[\tau \leq 0] = 0.$$

By letting $n \to \infty$, we obtain $\lim_n V(x^n, \mu^n) \leq (|\mu^\infty|+1)E[\tau]$, but since τ can be chosen arbitrarily small, we conclude that

$$\lim_{n\to\infty} V(x^n, \mu^n) \leq 0.$$

Since V is non-negative and zero where $x = 0$,

$$\lim_{n\to\infty} V(x^n, \mu^n) = 0 = V(0, \mu^\infty).$$

$\square$

Lemma 9.10. For each starting point (x, μ), the payoffs $\int_0^{\theta(L)} e^{-rt}\,\mathrm{d}L_t$ for strategies L are uniformly bounded L^1.

Proof. Without loss of generality, we may assume that $\mu \in \mathcal{O}(\kappa(\mu))$ as $\mu \to -\infty$, since this yields a larger or equally large process μ_t, and therefore also $\int_0^{\theta(L)} e^{-rt}\,\mathrm{d}L_t$. Integration by parts yields

$$\begin{aligned}
\int_0^{\theta(L)} e^{-rt}\,\mathrm{d}L_t &\leq x + \int_0^{\theta(L)} e^{-rt}\mu_t\,\mathrm{d}t + \int_0^{\theta(L)} e^{-rt}\sigma(\mu_t)\,\mathrm{d}W_t \\
&\leq x + \int_0^{\infty} e^{-rt}|\mu_t|\,\mathrm{d}t + \sup_T \int_0^{T} e^{-rt}\sigma(\mu_t)\,\mathrm{d}W_t,
\end{aligned}$$

where the sup is taken over stopping times T. This provides the L-independent bound if it has finite expectation.

The expectation of the first integral is finite, since, by Lemma 9.1,

$$E\left[\int_0^\infty e^{-rt}|\mu_t|\,\mathrm{d}t\right] \le \int_0^\infty e^{-rt}H(t)\,\mathrm{d}t < \infty.$$

Similarly, by Doob's inequality, Itô isometry and Lemma 9.1, we obtain, for some C,

$$E\left[\left(\sup_T \int_0^T e^{-rt}\sigma(\mu_t)\,\mathrm{d}W_t\right)^2\right] \le 2\int_0^\infty e^{-2rt}C\left(1+H(t)\right)\mathrm{d}t < \infty.$$

□

Theorem 6.3. The value function is continuous everywhere.

Proof. Denote by $\theta^n(L)$ the bankruptcy time of starting in (x^n, μ^n) and using the dividend policy L. Denote by L^n ε-optimal stratgies starting in (x^n, μ^n). Similarly, let X_t^n be the process associated with the starting point (x^n, μ^n) and dividend process L^n.

First consider a sequence (x^n, μ^n) which is non-decreasing in both coordinates and converges to (x^∞, μ^∞). Then

$$\begin{aligned} V(x^\infty, \mu^\infty) - \varepsilon &\le E\left[\int_0^{\theta^n(L^\infty)} e^{-rt}\,\mathrm{d}L_t^\infty + \int_{\theta^n(L^\infty)}^{\theta^\infty(L^\infty)} e^{-rt}\,\mathrm{d}L_t^\infty\right] \\ &\le V(x^n, \mu^n) + E\left[\int_{\theta^n(L^\infty)}^{\theta^\infty(L^\infty)} e^{-rt}\,\mathrm{d}L_t^\infty\right]. \end{aligned} \tag{8}$$

Therefore, if we can show that the second term tends to zero as n tends to infinity, we are done.

By Lebesgue's dominated convergence theorem (see Lemma 9.10), and then by the strong Markov property (7),

$$\lim_{n\to\infty} E\left[\int_{\theta^n(L^\infty)}^{\theta^\infty(L^\infty)} e^{-rt}\,\mathrm{d}L_t^\infty\right] = E\left[\lim_{n\to\infty}\int_{\theta^n(L^\infty)}^{\theta^\infty(L^\infty)} e^{-rt}\,\mathrm{d}L_t^\infty\right] \tag{9}$$

where $\bar{V}$ denotes the upper semicontinuous envelope of V.

We will prove that the limit inside the expectation is 0 on the following set:

$$\begin{aligned} \Omega' = &\left\{\sup_L \int_0^{\theta(L)} e^{-rt}\,\mathrm{d}L_t < \infty\right\} \\ &\cap \left\{\mathbb{E}\left[\int_{\theta^n(L^\infty)}^{\theta^\infty(L^\infty)} e^{-rt}\,\mathrm{d}L_t^\infty \middle| \mathcal{F}_{\theta^n(L^\infty)}\right] \le e^{-r\theta^n(L^\infty)}\bar{V}\left(X^\infty_{\theta^n(L^\infty)}, \mu^\infty_{\theta^n(L^\infty)}\right), \forall n \in \mathbb{N}\right\}. \end{aligned}$$

Note that by Lemma 9.1 and (7), $P(\Omega') = 1$. For any $\omega \in \Omega'$, consider the following two cases:

Let $\theta^k = \theta^{n(k)}(L^\infty)(\omega)$ be any subsequence which converges to ∞. Then,

$$\left(\int_{\theta^k}^{\theta^\infty(L^\infty)} e^{-rt}\,\mathrm{d}L_t^\infty\right)(\omega) \le e^{-r\theta^k/2}\left(\int_{\theta^k}^{\theta^\infty(L^\infty)} e^{-rt/2}\,\mathrm{d}L_t^\infty\right)(\omega) \xrightarrow{k\to\infty} 0,$$

because $\omega \in \Omega'$ ensures that the integral factor is bounded, and the exponential factor converges to 0.

On the other hand, let $\theta^k = \theta^{n(k)}(L^\infty)(\omega)$ be a bounded subsequence. We then wish to use the continuity of V, and therefore also of $\bar{V}$, at 0 to argue that

$$\lim_{k\to\infty} \bar{V}\left(X_{\theta^k}^\infty(\omega), \mu_{\theta^k}^\infty(\omega)\right) = 0.$$

Since $\sup_k \theta^k(\omega) < \infty$, $X_{\theta^k}^\infty = X_{\theta^k}^\infty - X_{\theta^k}^k = x^\infty - x^k + \int_0^{\theta^k} \mu_t^\infty - \mu_t^k\,\mathrm{d}t \xrightarrow{k\to\infty} 0$, because of continuity with respect to initial points. Therefore, since V is increasing, it is sufficient to find a bound to $\mu_{\theta^k}^\infty(\omega)$.

Begin by considering the process $M_t^\infty = \sup_{0\le s\le t} \mu_s^\infty$. Then, since θ^k is a bounded sequence and M_t^∞ is continuous, $M_{\theta^k}^\infty(\omega)$ is bounded by some constant C. Therefore, by Theorem 9.9,

$$\lim_{k\to\infty} \bar{V}\left(X_{\theta^k}^\infty(\omega), \mu_{\theta^k}^\infty(\omega)\right) \le \lim_{k\to\infty} \bar{V}\left(X_{\theta^k}^\infty(\omega), C\right) = \bar{V}(0, C) = 0.$$

Hence, for almost every ω it holds that

$$\left(\int_{\theta^n(L^\infty)}^{\theta^\infty(L^\infty)} e^{-rt}\,\mathrm{d}L_t^\infty\right)(\omega)$$

converges to 0 along any bounded subsequence and any subsequence converging to ∞. As a consequence of this, the whole sequence converges to 0, for every $\omega \in \Omega'$, i.e., P-a.s. Returning to (8), this yields

$$V(x^\infty, \mu^\infty) - \varepsilon \le \lim_{n\to\infty} V(x^n, \mu^n) \le V(x^\infty, \mu^\infty),$$

by monotonicity. Since this holds for any choice of $\varepsilon > 0$, equality is obtained.

Now let (x^n, μ^n) be non-increasing in both coordinates and converge to (x^∞, μ^∞). Then, in the same manner as above,

$$\lim_{n\to\infty} V(x^n, \mu^n) - \varepsilon \le V(x^\infty, \mu^\infty) + E\left[\lim_{n\to\infty} \int_{\theta^\infty(L^n)}^{\theta^n(L^n)} e^{-rt}\,\mathrm{d}L_t^n\right],$$

and by analogous arguments,

$$\lim_{n\to\infty} V(x^n, \mu^n) = V(x^\infty, \mu^\infty).$$

As a final step, consider an arbitrary convergent sequence (x^n, μ^n). By monotonicity,

$$V\left(\inf_{m\ge n} x_m, \inf_{m\ge n} \mu_m\right) \le V(x^n, \mu^n) \le V\left(\sup_{m\ge n} x_m, \sup_{m\ge n} \mu_m\right).$$

Since the sequences $(\inf_{m\geq n} x_m, \inf_{m\geq n} \mu_m)$ and $(\sup_{m\geq n} x_m, \sup_{m\geq n} \mu_m)$ are non-decreasing and non-increasing, respectively, it follows that

$$\lim_{n\to\infty} V(x^n, \mu^n) = V(x^\infty, \mu^\infty),$$

and V is continuous everywhere. □

9.5 Comparison principle

Lemma 9.11. If a function u is a viscosity subsolution to (3), then

$$\tilde{u}(x,\mu) := e^{-\eta x - \eta g(\mu)} u(x,\mu) \tag{10}$$

is a viscosity subsolution to

$$\begin{aligned} \min\Bigg\{ \Big(& r - \eta\mu - \eta g'(\mu)\kappa(\mu) - \eta^2\Sigma_{11} \\ & - \eta^2 g'(\mu)^2\Sigma_{22} - \eta g''(\mu)\Sigma_{22} - 2\eta^2 g'(\mu)\Sigma_{12}\Big)\tilde{V} \\ & - (\mu + \eta\Sigma_{11} + 2\eta g'(\mu)\Sigma_{12})\tilde{V}_x \\ & - (\kappa(\mu) + \eta g'(\mu)\Sigma_{22} + 2\eta\Sigma_{12})\tilde{V}_\mu \\ & - \operatorname{Tr}\Sigma D^2\tilde{V}, \\ & \qquad \eta\tilde{V} + \tilde{V}_x - e^{-\eta x - \eta g(\mu)}\Bigg\} = 0, \text{ in } \mathbb{R}_{>0}\times\mathcal{M}, \end{aligned} \tag{11}$$

for any given η and $g \in C^2(\mathbb{R})$. A corresponding statement is true for supersolutions.

Proof. Let u be a viscosity subsolution to (3). Let $\tilde{\varphi} \in C^2$ and (x_0, μ_0) be a local maximum of $\tilde{u} - \tilde{\varphi}$ where $(\tilde{u} - \tilde{\varphi})(x_0, \mu_0) = 0$. Finally, define

$$\varphi(x,\mu) := e^{\eta x + \eta g(\mu)}\tilde{\varphi}(x,\mu).$$

Then (x_0, μ_0) is also a local maximum of $u - \varphi$. Using the viscosity property of u and plugging in φ yields the viscosity property for $\tilde{u}$ and the transformed equation (11). □

Lemma 9.12. The parameter η and function g can be chosen in such a way that the coefficient of $\tilde{V}$ in (11) is strictly positive.

Proof. Fix any large $y > 0$ and let g be a twice differentible function with

$$g'(\mu) = \begin{cases} -\eta_-, & \text{if } \mu < -y, \\ \eta_+, & \text{if } \mu > y, \end{cases}$$

for strictly postive constants η_- and η_+ as well as $\mu \in [-y, y]^c$.[13] For any such choice, the coefficient is strictly positive in $[-y, y]$, provided η is small enough. Moreover, with our choice of g, the condition reduces to

$$r - \eta(\mu + \eta_\pm \kappa(\mu)) - \eta^2(\Sigma_{11} + \eta_\pm^2 \Sigma_{22} + 2\eta_\pm \Sigma_{12}) > 0, \quad \text{in } [-y, y]^c.$$

Note that the last two terms both grow at most linearly in μ, by the growth conditions on κ, $\tilde{\sigma}$, and σ. Furthermore, since κ is negative and linearly growing for large (positive) μ, we can choose η_+ such that $-(\mu + \eta_+ \kappa(\mu))$ is linearly increasing. Then, for sufficiently small η, the whole expression is increasing, for large μ.

Similarly, if both η_- and η are chosen sufficiently small, the same is true also for large, negative μ. Hence, for such a choice of η and g, the coefficient is strictly positive. □

Remark 9.13. Assumption 6.1 could possibly be generalized by finding strict supersolutions to the equation

$$r - \eta\mu - \eta g'(\mu)\kappa(\mu) - \eta^2\Sigma_{11} - \eta^2 g'(\mu)^2 \Sigma_{22} - \eta g''(\mu)\Sigma_{22} - 2\eta^2 g'(\mu)\Sigma_{12} = 0,$$

since this is sufficient for the transformed equation to be proper.

Theorem 6.4 (Comparison)**.** Let u and v be upper and lower semicontinuous, polynomially growing viscosity sub- and supersolutions of (3). Then $u \le v$ for $x = 0$ implies that $u \le v$ everywhere in $\mathcal{O} := \mathbb{R}_{\ge 0} \times \mathcal{M}$.

Proof of Theorem 6.4. Comparison is shown for the transformed DPE (11) with η chosen as in Lemma 9.12. Since the transformation (10) is sign-preserving, this is sufficient to establish comparison for (3) thanks to Lemma 9.11. For the sake of simpler presentation later on, (11) is shortened to

$$\begin{aligned} \min\{f\tilde{V} - f^x \tilde{V}_x - f^\mu \tilde{V}_\mu - \operatorname{Tr}\Sigma D^2 \tilde{V}, \\ \eta\tilde{V} + \tilde{V}_x - e^{-\eta x - \eta g(\mu)}\} = 0. \end{aligned}$$

Note that the coefficients f^x, and f^μ are locally Lipschitz continuous on the interior of the domain, and the coefficient f is continuous.

Denote by $\tilde{u}$ and $\tilde{v}$ the functions transformed as in (10). We note that these tend to 0 at infinity and that $\tilde{u}$ as well as $\tilde{v}$ are bounded. We distinguish between two cases:

1. The function $\tilde{u} - \tilde{v}$ attains a maximum in $[0, \infty) \times \mathcal{M}^o$. If the maximum is at $x = 0$, we are done. Otherwise, the maximum is attained in the interior $\mathcal{O}^o$.

2. There exists a maximizing sequence converging to a point $(\hat{x}, \hat{\mu})$ in $[0, \infty) \times \partial\mathcal{M}$. Without loss of generality, assume that $\hat{\mu}$ is a lower boundary point. An upper boundary point is handled analogously. The method employed here originates in [4].

[13] This can be done by a mollification argument on g'.

For some small $\gamma > 0$, let $N = \{(x, \mu) : \mu - \hat{\mu} < \gamma\} \cap ([0, \infty) \times \mathcal{M})$ and define

$$\Psi_\delta(x, \mu) = \tilde{u}(x, \mu) - \tilde{v}(x, \mu) - \delta h(\mu), \quad (x, \mu) \in N,$$

for $\delta \geq 0$ and

$$h(x, \mu) = \int_\mu^{\hat{\mu}+\gamma} e^{C(\xi - \hat{\mu})}(\xi - \hat{\mu})^{-1}\, \mathrm{d}\xi, \quad (x, \mu) \in N,$$

with

$$C = \sup\left\{\frac{1}{\mu} - \frac{2f^\mu(\mu)}{\tilde{\sigma}(\mu)^2} : 0 < \mu - \hat{\mu} < \gamma\right\}.$$

Note that by Assumption 6.1, $C < \infty$. It is easily verified that $h > 0$, $h(\mu) \to \infty$ as $\mu \to \hat{\mu}$, and that

$$f^\mu h' + \frac{\tilde{\sigma}^2}{2} h'' \leq 0, \quad \text{in } N.$$

Hence, $\tilde{w}_\delta := \tilde{u} - \delta h$ is also a subsolution in N.

Let $(x_n, \mu_n) \to (\hat{x}, \hat{\mu})$ be the maximizing sequence, and set $\delta = \delta_n = 1/nh(\mu_n)$. Then $\delta \to 0$ as $\mu \to \hat{\mu}$. Thus,

$$\sup_N(\tilde{u} - \tilde{v}) \geq \sup_N \Psi_\delta \geq \Psi_\delta(x_n, \mu_n) = (\tilde{u} - \tilde{v})(x_n, \mu_n) - 1/n,$$

so

$$\lim_{\delta \to 0} \sup_N \Psi_\delta = \sup_N\,(\tilde{u} - \tilde{v})\,.$$

Moreover, Ψ_δ attains a maximum $(x_\delta, \mu_\delta) \in \overline{N}$. For sufficiently small $\delta > 0$, a maximum is attained in the interior, or otherwise a maximum of $\tilde{u} - \tilde{v}$ is attained for $\mu = \hat{\mu} + \gamma \in \mathcal{O}^{\mathrm{o}}$. It follows that

$$\sup_N(\tilde{u} - \tilde{v}) \geq (\tilde{u} - \tilde{v})(x_\delta, \mu_\delta) = \sup_N \Psi_\delta + \delta h(\mu_\delta) \geq \sup_N \Psi_\delta,$$

so $\delta h(\mu_\delta) \to 0$. If $\max_N(\tilde{w}_\delta - \tilde{v}) \leq 0$,

$$\sup_{\mathcal{O}}(\tilde{u} - \tilde{v}) = \lim_{\delta \to 0} \max_N(\tilde{w}_\delta - \tilde{v}) \leq 0$$

follows. It therefore leads to no loss of generality to assume that $\tilde{u} - \tilde{v}$ attains a maximum in $\mathcal{O}^{\mathrm{o}}$.

By the preceeding argument, we may assume that $\tilde{u} - \tilde{v}$ attains local maximum $(x_0, \mu_0) \in \mathcal{O}^{\mathrm{o}}$. Let $N \subseteq N' \subseteq \mathcal{O}$ be two neighborhoods of (x_0, μ_0) in which this (x_0, μ_0) is a maximum and with $\overline{N} \subseteq N'$. For all $\epsilon > 0$, the function

$$\Phi_\epsilon(x, \mu, y, \nu) := \tilde{u}(x, \mu) - \tilde{v}(y, \nu) - \frac{1}{2\epsilon}\left(|x - y|^2 + |\mu - \nu|^2\right) - \|(x, \mu) - (x_0, \mu_0)\|_2^4\,.$$

has a maximum in $\overline{N} \times \overline{N}$, which we denote by $(x_\epsilon, \mu_\epsilon, y_\epsilon, \nu_\epsilon)$. In particular as $\epsilon \to 0$, the sequence converges along a subsequence.

From the construction of Φ_ϵ it follows that

$$\frac{1}{2\epsilon}\left(|x_\epsilon - y_\epsilon|^2 + |\mu_\epsilon - \nu_\epsilon|^2\right) \leq \tilde{u}(x_\epsilon, \mu_\epsilon) - \tilde{v}(y_\epsilon, \nu_\epsilon) - \max_{\bar{N}}(\tilde{u} - \tilde{v}).$$

Observing that the superior limit of the right hand side is bounded from above by 0 yields

$$\limsup_{\epsilon \to 0} \frac{1}{2\epsilon}\left(|x_\epsilon - y_\epsilon|^2 + |\mu_\epsilon - \nu_\epsilon|^2\right) \leq 0.$$

This estimate is used later in the proof. Moreover, $(x_\epsilon, \mu_\epsilon) \to (x_0, \mu_0)$, which means they are local maxima in N'.

By the Crandall–Ishii lemma, there exist matrices X_ϵ and Y_ϵ such that

$$(p_\epsilon, X_\epsilon) \in \bar{J}^{2,+}u(x_\epsilon, \mu_\epsilon), \quad (p_\epsilon, Y_\epsilon) \in \bar{J}^{2,-}v(y_\epsilon, \nu_\epsilon)$$

and

$$X_\epsilon \leq Y_\epsilon + \mathcal{O}\left(\frac{1}{\epsilon}\|(x,\mu) - (x_0, \mu_0)\|_2^2 + \|(x,\mu) - (x_0,\mu_0)\|_2^4\right),$$

for

$$p_\epsilon = \frac{1}{\epsilon}(x_\epsilon - y_\epsilon, \mu_\epsilon - \nu_\epsilon).$$

In particular, $X_\epsilon \leq Y_\epsilon + o(1)$ as $\epsilon \to 0$. Using the viscosity property of $\tilde{u}$ and $\tilde{v}$ yields

$$\min\left\{ f(\mu_\epsilon)\tilde{u}(x_\epsilon, \mu_\epsilon) - f^x(\mu_\epsilon)\frac{x_\epsilon - y_\epsilon}{\epsilon} - f^\mu(\mu_\epsilon)\frac{\mu_\epsilon - \nu_\epsilon}{\epsilon} - \operatorname{Tr}\Sigma(\mu_\epsilon)D^2\tilde{X}_\epsilon, \right.$$
$$\left. \eta\tilde{u}(x_\epsilon, \mu_\epsilon) + \frac{x_\epsilon - y_\epsilon}{\epsilon} - e^{-\eta x_\epsilon - \eta g(\mu_\epsilon)} \right\} \leq 0.$$

and

$$\begin{aligned} \min\Bigg\{ & f(\nu_\epsilon)\tilde{v}(y_\epsilon, \nu_\epsilon) - f^x(\nu_\epsilon)\frac{x_\epsilon - y_\epsilon}{\epsilon} - f^\mu(\nu_\epsilon)\frac{\mu_\epsilon - \nu_\epsilon}{\epsilon} - \operatorname{Tr}\Sigma(\nu_\epsilon)D^2\tilde{Y}_\epsilon, \\ & \eta\tilde{v}(y_\epsilon, \nu_\epsilon) + \frac{x_\epsilon - y_\epsilon}{\epsilon} - e^{-\eta y_\epsilon - \eta g(\nu_\epsilon)} \Bigg\} \geq 0. \end{aligned} \tag{12}$$

The rest of the proof is dividend into two cases, depending on whether

$$f(\mu_\epsilon)\tilde{u}(x_\epsilon, \mu_\epsilon) - f^x(\mu_\epsilon)\frac{x_\epsilon - y_\epsilon}{\epsilon} - f^\mu(\mu_\epsilon)\frac{\mu_\epsilon - \nu_\epsilon}{\epsilon} - \operatorname{Tr}\Sigma(\mu_\epsilon)D^2\tilde{X}_\epsilon \leq 0 \tag{13}$$

or

$$\eta\tilde{u}(x_\epsilon, \mu_\epsilon) + \frac{x_\epsilon - y_\epsilon}{\epsilon} - e^{-\eta x_\epsilon - \eta g(\mu_\epsilon)} \leq 0. \tag{14}$$

Case 1. Assume (13). Using (12) we arrive at

$$\begin{aligned} & f(\mu_\epsilon)(\tilde{u}(x_\epsilon, \mu_\epsilon) - \tilde{v}(y_\epsilon, \nu_\epsilon)) + (f(\mu_\epsilon) - f(\nu_\epsilon))\tilde{v}(y_\epsilon, \nu_\epsilon) \\ & \qquad - (f^x(\mu_\epsilon) - f^x(\nu_\epsilon))\frac{x_\epsilon - y_\epsilon}{\epsilon} - (f^\mu(\mu_\epsilon) - f^\mu(\nu_\epsilon))\frac{\mu_\epsilon - \nu_\epsilon}{\epsilon} \\ & \qquad\qquad - \operatorname{Tr}(\Sigma(\mu_\epsilon) - \Sigma(\nu_\epsilon))Y_\epsilon \leq \operatorname{Tr}\Sigma(\mu_\epsilon)(X_\epsilon - Y_\epsilon) \in o(1) \end{aligned}$$

Using the (local) Lipschitz continuity of the coefficients as well as the quadratic convergence rates of $x_\epsilon - y_\epsilon$ and $\mu_\epsilon - \nu_\epsilon$, we find that

$$f(\mu_0)(\tilde{u} - \tilde{v})(x_0, \mu_0) = \limsup_{\epsilon \to 0} f(\mu_\epsilon)(\tilde{u}(x_\epsilon, \mu_\epsilon) - \tilde{v}(y_\epsilon, \nu_\epsilon)) \leq 0$$

We conclude that

$$0 \geq (\tilde{u} - \tilde{v})(x_0, \mu_0).$$

Case 2. Assume (14). Using (12) yields

$$\eta(\tilde{u}(x_\epsilon, \mu_\epsilon) - \tilde{v}(y_\epsilon, \nu_\epsilon)) \leq e^{-\eta x_\epsilon - \eta g(\mu_\epsilon)} - e^{-\eta y_\epsilon - \eta g(\nu_\epsilon)}.$$

Once again we use the convergence results, and once again we conclude that

$$0 \geq (\tilde{u} - \tilde{v})(x_0, \mu_0) = \max_{\mathcal{O}}(\tilde{u} - \tilde{v}).$$

The inequality $\tilde{u} \leq \tilde{v}$ holds at any local maximum. Moreover, as mentioned in the beginning of the proof, we may assume that $\tilde{u} - \tilde{v}$ is attains its maximum on the interior. The theorem statement follows from the fact that

$$\tilde{u} \leq \tilde{v} \iff u \leq v.$$

□

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 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.

[3] Albert Altarovici, Max Reppen, and H. Mete Soner. Optimal consumption and investment with fixed and proportional transaction costs. *SIAM Journal on Control and Optimization*, 55(3):1673–1710, 2017.

[4] Anna Lisa Amadori. Uniqueness and comparison properties of the viscosity solution to some singular HJB equations. *Nonlinear Differential Equations and Applications NoDEA*, 14(3-4):391–409, 2007.

[5] Ronald W Anderson and Andrew Carverhill. Corporate liquidity and capital structure. *Review of Financial Studies*, 25(3):797–837, 2012.

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

[7] Guy Barles and Panagiotis E Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic analysis*, 4(3):271–283, 1991.

[8] Andrea Barth, Santiago Moreno-Bromberg, and Oleg Reichmann. A non-stationary model of dividend distribution in a stochastic interest-rate setting. *Computational Economics*, 47(3):447–472, 2016.

[9] Patrick Bolton, Hui Chen, and Neng Wang. Market timing, investment, and risk management. *Journal of Financial Economics*, 109(1):40–62, 2013.

[10] Bruno Bouchard and Nizar Touzi. Weak dynamic programming principle for viscosity solutions. *SIAM Journal on Control and Optimization*, 49(3):948–962, 2011.

[11] Abel Cadenillas, Sudipto Sarkar, and Fernando Zapatero. Optimal dividend policy with mean-reverting cash reservoir. *Mathematical Finance*, 17(1):81–109, 2007.

[12] Bruno de Finetti. Su un'impostazione alternativa della teoria collettiva del rischio. In *Transactions of the XVth international congress of Actuaries*, volume 2, pages 433–443. New York, 1957.

[13] 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.

[14] Avinash K Dixit and Robert S Pindyck. *Investment under uncertainty*. Princeton university press, 1994.

[15] Wendell H Fleming and Halil Mete Soner. *Controlled Markov processes and viscosity solutions*, volume 25. Springer Science & Business Media, 2006.

[16] Roger Frock. *Changing how the world does business*. Berrett-Koehler, San Francisco, 2006.

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

[18] Mark Gertler and R Glenn Hubbard. Corporate financial policy, taxation, and macroeconomic risk. Technical report, National Bureau of Economic Research, 1991.

[19] Dirk Hackbarth, Jianjun Miao, and Erwan Morellec. Capital structure, credit risk, and macroeconomic conditions. *Journal of Financial Economics*, 82(3):519–550, 2006.

[20] Julien Hugonnier, Semyon Malamud, and Erwan Morellec. Capital supply uncertainty, cash holdings, and investment. *Review of Financial Studies*, 28(2):391–445, 2015.

[21] Donald L. Iglehart. Limiting diffusion approximations for the many server queue and the repairman problem. *Journal of Applied Probability*, 2(02):429–441, 1965.

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

[23] Zhengjun Jiang and Martijn Pistorius. Optimal dividend distribution under Markov regime switching. *Finance and Stochastics*, 16(3):449–476, 2012.

[24] Nicole El Karoui and Xiaolu Tan. Capacities, measurable selection and dynamic programming part i: abstract framework. *arXiv preprint arXiv:1310.3363*, 2013.

[25] Nicole El Karoui and Xiaolu Tan. Capacities, measurable selection and dynamic programming part ii: application in stochastic control problems. *arXiv preprint arXiv:1310.3364*, 2013.

[26] Magdalena Kobylanski, Marie-Claire Quenez, et al. Optimal stopping time problem in a general framework. *Electron. J. Probab*, 17(72):1–28, 2012.

[27] Filip Lundberg. *1. Approximerad framställning af sannolikhetsfunktionen: 2. Återförsäkring af kollektivrisker. Akademisk afhandling.* Almqvist &Wiksells, 1903.

[28] Pauli Murto and Marko Terviö. Exit options and dividend policy under liquidity constraints. *International Economic Review*, 55(1):197–221, 2014.

[29] Stewart C Myers. Determinants of corporate borrowing. *Journal of Financial Economics*, 5(2):147–175, 1977.

[30] Goran Peskir. Bounding the maximal height of a diffusion by the time elapsed. *Journal of Theoretical Probability*, 14(3):845–855, 2001.

[31] Martin L Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* John Wiley & Sons, Inc., 1994.

[32] Roy Radner and Larry Shepp. Risk vs. profit potential: A model for corporate strategy. *Journal of economic dynamics and Control*, 20(8):1373–1393, 1996.

[33] Steven E Shreve, John P Lehoczky, and Donald P Gaver. Optimal consumption for general diffusions with absorbing and reflecting barriers. *SIAM Journal on Control and Optimization*, 22(1):55–75, 1984.