Received: 1 July 2018 | Accepted: 2 February 2019
DOI: 10.1111/mafi.12223

MATHEMATICAL FINANCE An International Journal of Mathematics, Statistics and Financial Economics
WILEY

ORIGINAL ARTICLE

Optimal dividend policies with random profitability

A. Max Reppen[1] | **Jean-Charles Rochet**[2,3,4] | **H. Mete Soner**[5,6]

[1]ORFE Department, Princeton University, Princeton, New Jersey
[2]Swiss Finance Institute, University of Geneva, Geneva, Switzerland
[3]Department of Banking and Finance, University of Zürich, Zürich, Switzerland
[4]Toulouse School of Economics/Institut d'Économie Industrielle, Toulouse, France
[5]Department of Mathematics, ETH Zürich, Zürich, Switzerland
[6]Swiss Finance Institute, Zürich, Switzerland

Correspondence
H. Mete Soner, Department of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland.
Email: hmsoner@ethz.ch

Funding information
The first and third authors were partly supported by the ETH Foundation, the Swiss Finance Institute, and Swiss National Foundation grant SNF 200020_172815. The second author acknowledges support by the SCOR.

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 comparison between discontinuous sub- and supersolutions of the Hamilton–Jacobi–Bellman equation, and we provide an efficient and convergent numerical scheme for finding the solution. The value function is given by a nonlinear partial differential equation (PDE) with a gradient constraint from below in one direction. 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.

KEYWORDS
barrier strategy, dividend problem, singular control, viscosity solutions

1 | INTRODUCTION

The problem of optimizing dividend flows has its origins in the actuarial field of ruin theory, which was first treated theoretically by Lundberg (1903). The theory typically models an insurance firm, and initially revolved around minimizing the probability of ruin. However, in many situations in practice, there is an emphasis on maximizing the shareholder value—an idea that fits well into de Finetti (1957) proposal to 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 modeling 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 by Myers (1977). Although the dividend problem seeks the value of a cash flow after 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. Our direct focus is not an insurance firm, and we instead study cash reserves described by a diffusion process. Although this is not the natural insurance perspective, it is studied also there as the limiting case of the jump processes, as initiated by Iglehart (1965).

Formulated as a problem of "storage or inventory type," the general diffusion problem with singular dividend policies was solved by Shreve, Lehoczky, and Gaver (1984). In the case of constant coefficients in the cash reserves dynamics, Jeanblanc-Picqué and Shiryaev (1995) 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* that frequently appear in jump models and were first identified by Gerber (1969). 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 the 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 (1991) and more recently by Hackbarth, Miao, and Morellec (2006). Such macroeconomic effects have been studied in various forms. In particular, Anderson and Carverhill (2012) as well as Barth, Moreno-Bromberg, and Reichmann (2016) numerically study continuously changing stochastic parameters, whereas Akyildirim, Güney, Rochet, and Soner (2014) consider stochastic interest rates following a Markov chain, and Jiang and Pistorius (2012) consider model coefficients and interest rate both governed by Markov chains. Bolton, Chen, and Wang (2013) similarly study the macroeconomic impact on both financial and investment opportunities. In contrast to coefficients influenced by macroeconomic factors, Radner and Shepp (1996) already in 1996 modeled a firm that alternates between different operating strategies, thereby effectively controlling the model coefficients. Other extensions include transaction costs of dividend payments or the possibility of equity issuance, cf. Akyildirim et al. (2014); Bolton et al. (2013); Décamps, Mariotti, Rochet, and Villeneuve (2011). The papers by Cai, Gerber, and Yang (2006), and Cadenillas, Sarkar, and Zapatero (2007) both treat different models with mean-reverting cash reserves, contrasting the model here, where we instead consider mean-reverting profitability. Finally, Avanzi and Wong (2012) and Albrecher and Cani (2017) propose dividend processes proportional and affine in the cash reserves, respectively, as a means to capture the more stable dividend streams seen in practice. For further references, we refer the reader to Asmussen and Albrecher (2010); Albrecher and Thonhauser (2009) and the references therein.

Our choice of diffusion model has a continuous, stochastic drift generated by a separate profitability process. This structure yields a two-dimensional problem in which the dividend strategy depends on the current profitability. 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 profitability. 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 (DPE). The latter supports the convergence of the numerical scheme.

Mathematically, the problem is an example of a *multi-dimensional singular* optimal control. Although the abstract formulation is given generally Fleming & Soner (2006), a vast part of the literature and applications of singular control have been in one space dimension. Indeed, in multi-dimensions, several difficult technical problems arise as the geometry of the free boundary separating the two regimes is nontrivial. Consequently, the problem is studied only under structural conditions. In two space dimensions, the regularity of the free boundary is studied by Soner and Shreve (1989) using the rotational symmetry and by Chiarolla and Haussmann (1994) through monotonicity in both variables. When the so-called push direction is unique, Shreve and Soner (1991a, 1991b) exploit the connection to an obstacle problem.

Another approach to singular control is to characterize the free boundary as a solution of an integral equation. In the context of the classical Stephan problem of melting ice, this was first observed by Kolodner (1956). A similar characterization through a very different method is given in Ferrari (2015); De Angelis, Federico, and Ferrari (2017). These papers use stochastic singular control to model problems in irreversible investment. More recently, a public debt management problem studied by Ferrari (2018) and an integral equation as well as the regularity of the free boundary is obtained. Mathematically, this model is the closest to the one studied in this paper. However, there are substantial differences as well. In our model, as opposed to Ferrari (2018), bankruptcy happens and is central to the difficulties in our analysis. In particular, there are at least two free boundaries in our problem, and the derivation of integral equations seems not possible. Consequently, we follow the dynamic programming approach and prove a general comparison result, Theorem 3.7. In Section 4, we then use it to prove the convergence of our numerical method. In turn, this allows us to numerically study the problem extensively.

The paper is organized as follows. We begin by describing the problem in Section 2, followed by our assumptions and analytical results in Section 3. In Section 4, we present a numerical algorithm as well as its results. Intuition for these results is then provided by studying a related, simpler problem in Section 5. Finally, we suggest and numerically study some possible extensions of the model in Section 6. At the end of the paper, after the concluding comments in Section 7, the proofs of the statements in Section 3 are given in Section 8.

2 | PROBLEM FORMULATION

Consider a *cash flow* of the form

$$dC_t^{\mu} = \mu_t dt + \sigma dW_t, \quad C_0^{\mu} = 0,$$

[1] Similar properties have been observed by Anderson and Carverhill (2012) and Murto and Terviö (2014).

where W is a Brownian motion and the *profitability* (cash flow rate) $(\mu_t)_{t\geq 0}$ is described by

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

for some functions κ and $\tilde{\sigma}$, as well as another Brownian motion $\tilde{W}$ with correlation $\rho \in [-1, 1]$ to W. When it is necessary to emphasize the starting point of $(\mu_t)_{t\geq 0}$, a superscript will be added. In other words, $(\mu_t^\nu)_{t\geq 0}$ is the process with $\mu_0 = \nu$ that solves the above stochastic differential equation (SDE). 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 Assumptions 3.2 and 3.3, 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 3.2 only imposes 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

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

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, a half-line, for a CIR process, or a bounded interval, for a bounded process. The value function is then defined as

$$\begin{aligned} V(x,\mu) &:= \sup_{L\in\mathcal{A}(x,\mu)} J(x,\mu;L) \\ &:= \sup_{L\in\mathcal{A}(x,\mu)} \mathbb{E}\left[\int_0^{\theta^{(x,\mu)}(L)} e^{-rt}\mathrm{d}L_t \,\middle|\, (X_0,\mu_0) = (x,\mu)\right], \quad (x,\mu)\in\mathcal{O} := [0,\infty)\times\mathcal{M}, \quad (1) \end{aligned}$$

[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.

where $\mathcal{A}(x, \mu)$ is the set of processes $L = (L_t)_{t\geq 0}$ that are cádlàg (right-continuous with left limits) adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by C^μ and μ, as well as nondecreasing[3] with $\Delta L_t \leq X_{t-}^{(x,\mu),L}$,[4,5] and

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

is the time of bankruptcy. We model a limited liability firm, and therefore do not impose a penalty at ruin.[6] In particular, we interpret a payout $\Delta L_t = X_{t-}$ as a decision to liquidate the firm. In the context of insurance, one may imagine a ruin penalty is also possible, see for instance Liang and Young (2012). Mathematically, this would change the boundary conditions, but we believe that most of the analysis and in particular the computational method still apply.

Note that as μ satisfies comparison (see Revuz & Yor, 1999, theorem 3.7, p. 394), $X_t^{(x,\mu),L} \leq X_t^{(y,\nu),L}$ P-a.s. for any $t \leq \theta^{(x,\mu)}$ if $x \leq y$ and $\mu \leq \nu$. Hence, the value function is monotone in each argument. We use this comparison property repeatedly in the proofs.

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.

3 | MAIN RESULTS

Before presenting the main results, we give a brief account of the structure of the state space. The state space can be divided into a number of regions, each characterized by its role in the DPE—interpreted as a control action—as illustrated in Figure 1. The value function is characterized by three main regions: the dividend region, retained 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 intersect at the line at the threshold value 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 boundary* (or *dividend target*) is paid to the shareholders.

The remainder of this section is devoted to the statements of the main results that are all proved in Section 8. For the proofs, we need the following set of assumptions, satisfied by for example Ornstein–Uhlenbeck and CIR processes. This assumption essentially says that the profitability process is mean-reverting and well behaved in the sense of a Feller condition. Mean-reversion of profitability is empirically well established by Fama and French (2000) and others.

We will use the so-called big O notation.

[3] The process L must be nondecreasing because the limited liability of shareholders implies that $\mathrm{d}L_t$ cannot be negative. Section 6 considers the case when new shares can be issued at a cost, allowing to inject new cash into the firm. Also any monotone process such as L has left limits and we could dispose with the requirement.

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

[5] Although the model allows dividend payments with infinite "frequency" 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 (2) in Section 3.4 and exploited in Section 4, this model can be considered the limit when there is no bound on the dividend rate.

[6] Limited liability implies that the bankruptcy costs are not paid by the owners of the firm but by its creditors.

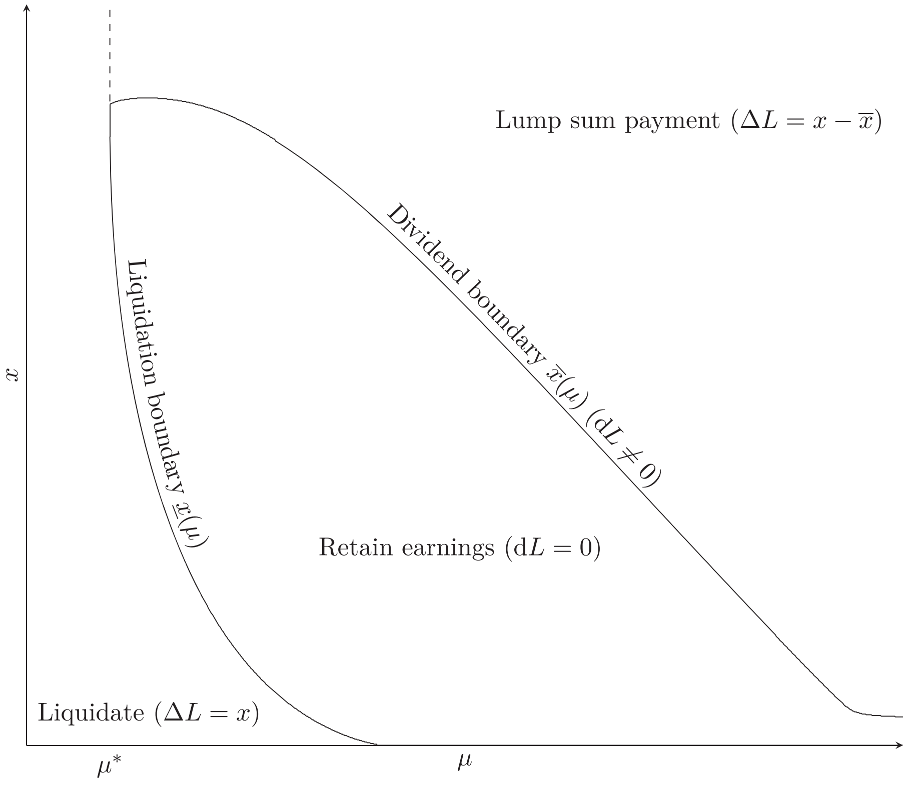

FIGURE 1 The figure shows the three regimes. In the region between the lines $\underline{x}$ and $\overline{x}$, the all incoming profits are retained. When the cash reserves fall to $\underline{x}$, the firm liquidates, whereas when it increases to $\overline{x}$, dividends are paid out according to the 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

Definition 3.1. A function f is said to be in $O(g)$ (so-called big O) as $y \to a$ if

$$\limsup_{y \to a} \left| \frac{f(y)}{g(y)} \right| < \infty.$$

We are now ready to state the following set of assumptions on the growth of the coefficients and the regularity of μ at a boundary of its domain. The following condition on the μ process, without explicitly mentioning, is assumed throughout the paper.

Assumption 3.2. Throughout, we assume that the domain of $(\mu_t)_{t \geq 0}$ is some possibly unbounded interval $\mathcal{M} \subseteq \mathbb{R}$. In other words, for any starting point $\mu_0 \in \mathcal{M}$, the process remains in $\mathcal{M}$.

Moreover, we assume that κ and $\tilde{\sigma}^2$ are locally Lipschitz continuous on the interior $\mathcal{M}^{\text{o}}$, that $-\mu/\kappa$ is nonnegative and bounded for large (positive) μ, that $-\kappa/\mu$ is nonnegative and bounded for large $-\mu$, as well as that $\tilde{\sigma}^2 \in O(\mu)$ and never vanishes in $\mathcal{M}^{\text{o}}$.

The following assumption is needed to prove comparison for non-Lipschitz coefficients κ and $\tilde{\sigma}$.

Assumption 3.3. If $\mathcal{M} \neq \mathbb{R}$, that is, it has an upper or lower boundary, we require some regularity at this boundary.

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

2. For any sufficiently small $\eta > 0$, we assume that if $-\infty < \inf \mathcal{M} = \nu$, then

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

and if $\infty > \sup \mathcal{M} = \nu$, then

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

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 origin of the lim sup and lim inf conditions at the boundary points are due to conditions needed for the comparison proof (cf. Section 8.4), while still general enough to encompass processes like the ones in Section 2.

We assume that Assumption 3.2 holds throughout. This assumption is not necessary for the comparison proof, but ensures that the value function is polynomially growing—the family of function for which comparison is established.

3.1 | Finiteness

We begin by establishing that the value function is finite. The proof of this result is provided in Section 8.1.

Theorem 3.4. *The value function is finite.*

3.2 | Liquidation threshold

The following theorem establishing the existence of the value μ^* in Figure 1 is proven in Section 8.2.

Theorem 3.5. *If $\mathcal{M}$ has no lower bound, there exists a value μ^* such that it is optimal to liquidate immediately whenever $\mu \leq \mu^*$, that is, $V(x,\mu) \equiv x$.*

3.3 | Continuity

Although the problem formulation bears resemblance to the Merton consumption problem, which has been extensively studied in the mathematical finance literature, the crucial difference here is that the firm is always exposed to the risk of its own operations. In other words, because X is the only asset and also volatile, there is no safe asset, and, 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 2 shows a case where $\theta(\frac{L^1+L^2}{2}) < \max\{\theta(L^1), \theta(L^2)\}$ 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 concavity corresponds to nonconvexity of the set of dividend processes that are constant after ruin.

Despite this lack of concavity, we can prove continuity, given by the following theorem.

Theorem 3.6. *The value function is continuous everywhere.*

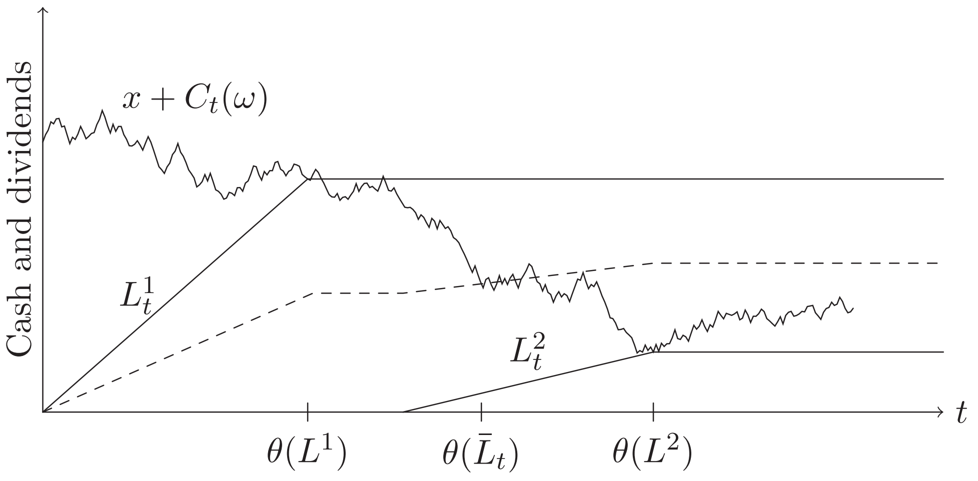

FIGURE 2 Recall that ruin occurs when the dividends reach the total cash accumulated, that is, $x + C_t$. The figure illustrates two dividend policies L^1 and L^2 (solid) as well as their convex combination $\bar{L} = (L^1 + L^2)/2$ (dashed), showing that for some path $\theta(\bar{L}) < \theta(L^1)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

3.4 | Dynamic programming equation

Following the continuity of Theorem 3.6, we refer to Fleming & Soner (2006) for proving the dynamic programming principle. For a general proof of dynamic programming, we refer to Karoui and Tan (2013a, 2013b). Writing

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

where

$$\Sigma = \frac{1}{2}\begin{bmatrix} \sigma^2 & \rho\sigma\tilde{\sigma} \\ \rho\sigma\tilde{\sigma} & \tilde{\sigma}^2 \end{bmatrix}$$

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

$$\min\{rV - \mathcal{L}V, V_x - 1\} = 0, \quad \text{in } \mathbb{R}_{>0} \times \mathcal{M}, \tag{2}$$

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

Theorem 3.7 (**Comparison**). *Let u and v be upper and lower semicontinuous, polynomially growing viscosity sub- and supersolutions of (2). Then, under Assumption 3.3, $u \le v$ for $(x,\mu) \in 0 \times \mathcal{M}$ implies that $u \le v$ everywhere in $\mathcal{O} := \mathbb{R}_{\ge 0} \times \mathcal{M}$.*

Corollary 3.8 (**Uniqueness**). *The value function is the unique subexponentially growing* viscosity solution *to the dynamic programming equation (2).*

Proof. By the dynamic programming principle, the value function V is a viscosity solution to (2). To obtain uniqueness, observe that, by Theorem 3.7, 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.

Note that the comparison principle is essential for the stability property of viscosity solutions, which leads to convergence of monotone numerical schemes to a solution, and, by the comparison theorem, it is unique, cf. Barles and Souganidis (1991). It thus ensures that the stability property gives the correct solution.

4 | 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, calculates the corresponding payoff function, then improves the policy by updating it 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 3.7 and the uniqueness result in Corollary 3.8.

The idea implemented here is to approximate the singularity with increasingly large controls that are absolutely continuous with respect to time. In particular, let $K > 0$ be any large constant and consider control variables $L_t = \int_0^T \ell_t \mathrm{d}t$, where the adaptive control process ℓ_t takes values in $[0, K]$. Then, the problem of optimizing over ℓ amounts to a penalization of the DPE (2) 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

$$\min_{\ell\in[0,K]} \left[rV^K - \mathcal{L}V^K\right] + \ell\left[V_x^K - 1\right] = 0. \tag{3}$$

Letting $\lambda = \ell/K$ 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

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

For $p_x, p_\mu \in \mathbb{R}$ and a matrix Y, let, with some abuse of notation,

$$\mathcal{L}(\mu, p_x, p_\mu, Y) = \mu p_x + \kappa(\mu)p_\mu + \operatorname{Tr}\Sigma(\mu)Y,$$

and

$$F^K(x,\mu,u,p_x,p_\mu,Y) = \min\left\{ru(x,\mu) - \mathcal{L}(\mu,p_x,p_\mu,Y), p_x - 1 + \frac{ru(x,\mu) - \mathcal{L}(\mu,p_x,p_\mu,Y)}{K}\right\}.$$

Letting $K \to \infty$,

$$\begin{aligned}\lim_{K\to\infty} F^K(x,\mu,u,p_x,p_\mu,Y) &= \min\left\{ru(x,\mu) - \mathcal{L}(\mu,p_x,p_\mu,Y), p_x - 1\right\}\\ &= F(x,\mu,u,p_x,p_\mu,Y).\end{aligned}$$

Finally, V^K is bounded from above (by V) and monotonically increasing, and hence converging to some function V^∞ as $K \to \infty$. Moreover, V^K satisfies

$$F^K(x,\mu,V^K,V_x^K,V_\mu^K,D^2V^K) = 0,$$

so, by the viscosity stability property, V^∞ is a viscosity solution to $F = 0$, and, by the comparison theorem, $V^\infty \equiv V$.

Thanks 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., Puterman, 1994). Thus, after a suitable space discretization, the problem is solved using standard methods of policy iteration that are known to converge. The convergence to V^K of the solution to the discretized problem can be also be shown with standard viscosity methods.

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 the terms from (3) involving V^K, that is, $rV^K - \mathcal{L}V^K + \ell V^K_x$. Then, starting with any control ℓ_0, the policy iteration scheme with tolerance[7] $\tau \geq 0$ is given by the following steps:[8]

Algorithm 1: Policy iteration algorithm (step *i*)

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

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

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

3. Return to step (i).

Remark 4.1. In the algorithm above, the second step is computationally the most difficult. In general, there is no known structure to be used, so it often has to be solved by brute force. Luckily, however, the problems given for each (x, μ) are fully independent, and the step can thus be entirely parallelized. Thanks to this, the computational burden can be partly mitigated.

[7] Note that the policy iteration scheme halts even for $\tau = 0$.

[8] As we solve in a bounded 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 (approximately) optimal policy is expected to naturally reflect at the upper x-boundary, 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.

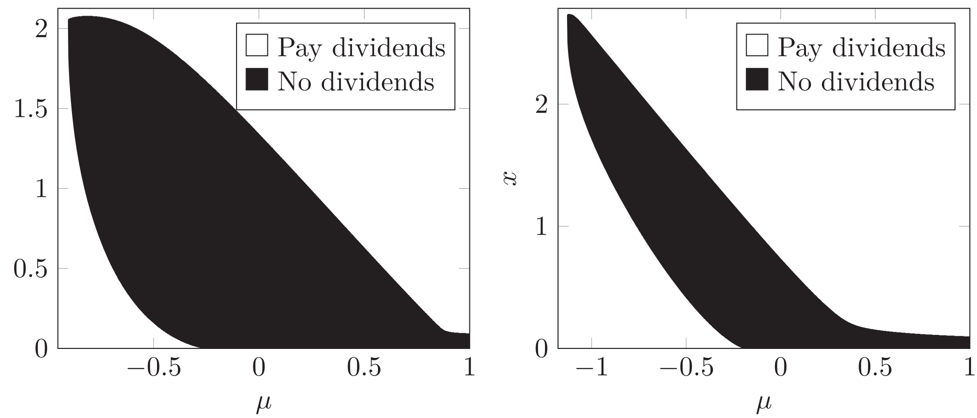

FIGURE 3 The black region corresponds to $\mathrm{d}L = 0$, whereas the white region corresponds to $\mathrm{d}L > 0$. The latter case is interpreted as either dividend payments or liquidation, depending on the position in the state space, see Figure 1

4.1 | Results and comparative statics

What follows is an account of an implementation of the scheme 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 dividend policies are presented in Figure 3, with $k = 0.5$, $\bar{\mu} = 0.15$, and $\tilde{\sigma} = 0.3$ (left) as well as $\tilde{\sigma} = 0.1$ (right). The other parameter choices are $\sigma = 0.1$, $\rho = 0$, and $r = 0.05$. The white regions indicate dividend payments or liquidation, that is, $V_x = 1$ and $\mathrm{d}L > 0$, whereas the black regions indicate that the firm retains all its earnings, that is, $rV - \mathcal{L}V = 0$ and $\mathrm{d}L = 0$. The figures show the (approximately) optimal policy, from which the value function can be obtained by solving a linear system of equations.

Figure 4 shows the effect of changing one parameter at a time. The figures presented here have been generated for $\tau = 0$, without parallelization on an ultrabook laptop.[9] On this system, the computations took around 10 min[10] at this resolution, and were observed to converge. 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.

5 | A SOURCE OF INTUITION: THE DETERMINISTIC PROBLEM

When the two parameters σ and $\tilde{\sigma}$ are zero, the problem can be solved explicitly for a mean-reverting $\kappa(\mu) = k(\bar{\mu} - \mu)$, with $k, \bar{\mu} > 0$. The solution of this deterministic problem provides intuition for the solution of the stochastic problem.

[9] Intel Core i7-7600U, 1866MHz LPDDR3.

[10] For the left solution in Figure 3 (same as "default" values from Figure 4), the runtime was just short of 8 min on a $1,000 \times 1,000$ grid without parallelization. A more modest grid size of 300×300 runs in 7 s on the same hardware. All computations were generated with a starting point with "no information," that is, ℓ identically zero.

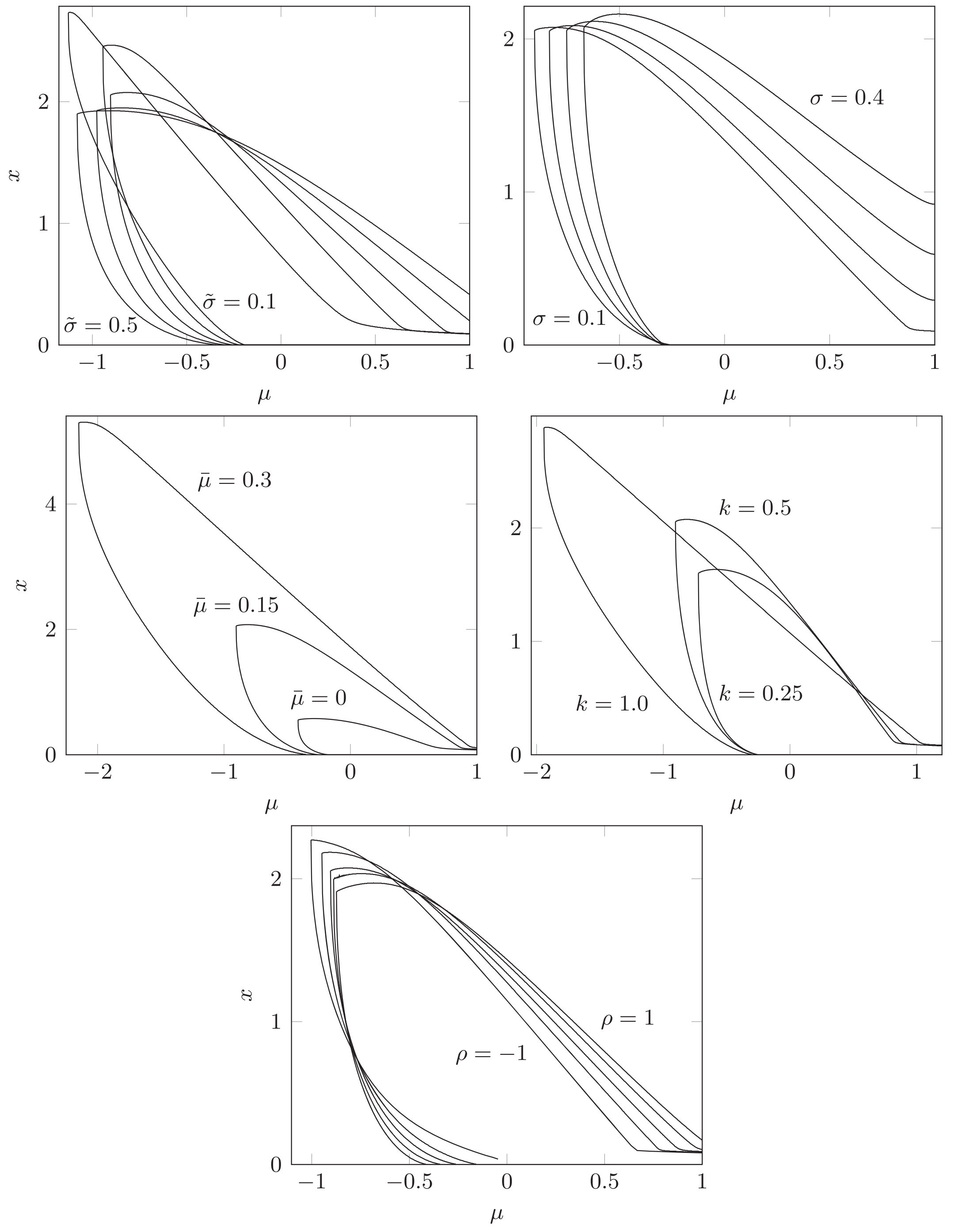

FIGURE 4 Comparative statics. Apart from for the parameter being varied, the chosen values were $r = 0.05$, $k = 0.5$, $\bar{\mu} = 0.15$, $\tilde{\sigma} = 0.3$, $\sigma = 0.1$, and $\rho = 0$. The parameter varied is indicated in the respective figure. The values considered were $k = 0.25, 0.5, 1.0$, $\bar{\mu} = 0.0, 0.15, 0.3$, $\tilde{\sigma} = 0.1, 0.2, 0.3, 0.4, 0.5$, $\sigma = 0.1, 0.2, 0.3, 0.4$, 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

As $(\mu_t)_{t\geq 0}$ 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 does not 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 value of paying the incoming earnings as dividends is

$$\int_0^\infty e^{-rt}\mu_t \mathrm{d}t = \int_0^\infty \bar{\mu}e^{-rt} + (\mu - \bar{\mu})e^{-(k+r)t}\mathrm{d}t = \frac{\bar{\mu}}{r} + \frac{\mu - \bar{\mu}}{r + k}. \tag{4}$$

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 then. 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 a 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 $\mu \geq 0$, that is,

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

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 $x \geq x_b(\mu)$,

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

Because 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 the 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 5.1. *There exist thresholds $x_b(\mu)$ and μ^* such that*

- *it is optimal to liquidate immediately if $\mu \leq \mu^*$;*

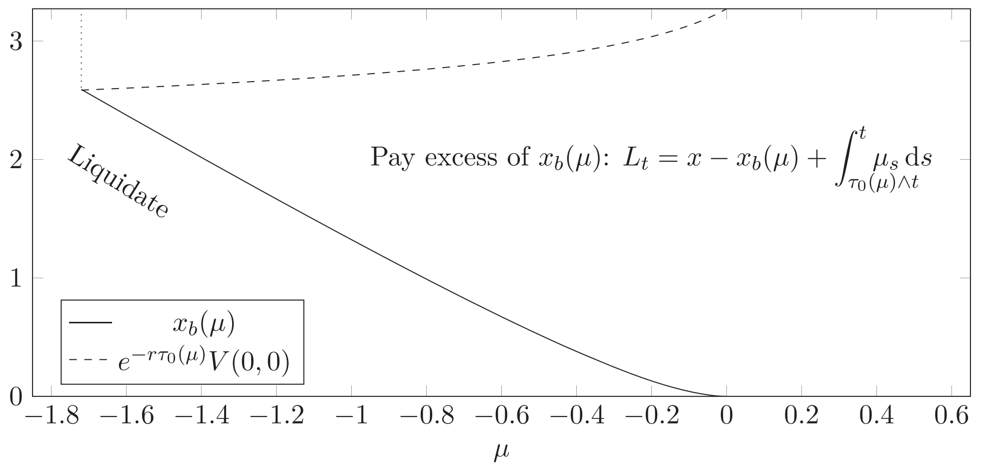

FIGURE 5 The value $x_b(\mu)$ is the cost of waiting for positive cash flow, whereas $e^{-r\tau_0(\mu)}V(0,0)$ is the present value of the future positive cash flow. For μ below the level at which these coincide, liquidation is thus optimal. Liquidation is also optimal when $x < x_b(\mu)$

- *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.*

This is the same type of structure we observe in the numerical solutions. The deterministic solution suggests that liquidation is optimal for two different reasons. When μ is very negative, the firm cannot expect to recover the initial “investment” x even if it avoids ruin, suggesting that the cash flow represents a poor business opportunity. On the other hand, for initial points (x, μ) below the solid line in Figure 5, the net present value of the cash flow net of the investment x is positive, but the firm foresees a liquidity crisis at $\tau_0(\mu)$. This future liquidity crisis could be avoided with more initial capital. However, lacking the extra funds, the optimal decision is immediate liquidation.

6 | MODEL EXTENSIONS

In this section, we present several model extensions that we study numerically. The numerical method used is the same as in Section 4, but in the case of fixed costs, a slight generalization of the approximation of controls is necessary, due to the nonlocal behavior, as mentioned below.

6.1 | 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 that is borne by the firm, and reflects the quality of its access to financial markets. We consider two costs: one cost λ_p *proportional* to the capital received and one *fixed cost* λ_f that is independent of the amount of equity issued. Mathematically, we follow the model in Décamps et al. (2011) and write

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

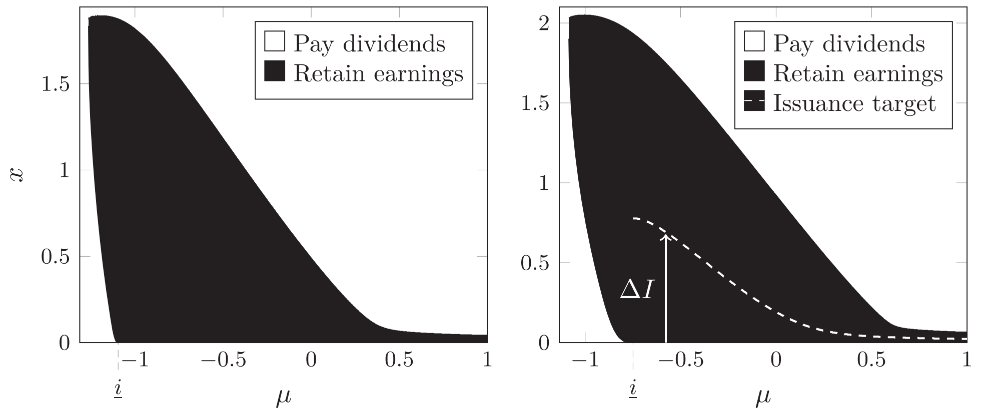

FIGURE 6 In the left figure, proportional issuance costs start at 34% for low μ and decaying to 25%. Equity is issued according to local a time at the boundary $x = 0$ where otherwise ruin would occur, and no equity is issued below $\underline{i} = -1.0953$. In the right figure, fixed costs are present, starting at 0.14 decaying to 0.06 as a function of μ. 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, that is, how much equity was issued

where $I = (I_t)_{t\geq 0}$, just like L, is an adapted, increasing, cádlàg control process. We allow for the costs to be μ-dependent,[11] 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 $r = 0.05$.

6.1.1 | Proportional issuance costs

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

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

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, as 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 do 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.

Because 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, that is, where $x = 0$. However, whether to do so at the boundary depends on the current profitability. Indeed, as seen in Figure 6, equity is only issued when the profitability is above a certain level, below which we still see the band structure of the original problem.

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

6.1.2 | Fixed issuance costs

On the other hand, if the fixed cost is nonzero, we can 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 could not be written in this form, it would imply 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:

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

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

As 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 6. Note that, the numerical method employed in the fixed cost case can be interpreted as the issuance structure in Hugonnier, Malamud, and Morellec (2015) where the arrival of investors might not coincide with the desire for equity issuance. The approximation parameter K from the singular case here describes the arrival rate of investors, and the value V is the limit when the investor arrival rate tends to infinity. This same approximation scheme is presented for another problem in Altarovici, Reppen, and Soner (2017) and turns out to be very accurate.

By letting the fixed cost grow sufficiently fast in $-\mu$, one can obtain substantially different issuance policies. As shown in Figure 7, 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 line, but instead have a jump discontinuity to the right of the gray region.

6.2 | Gambling for resurrection

In Figure 3, we see that in the band region the solution first grows as x (liquidation region) and then reaches the region where $V_x > 1$ (retained earnings region). We thus observe that the value function is not necessarily concave in the x-direction. However, it is sometimes argued that concavity is desirable, because of the possibility to enter (fair) speculative strategies and thus receiving the average

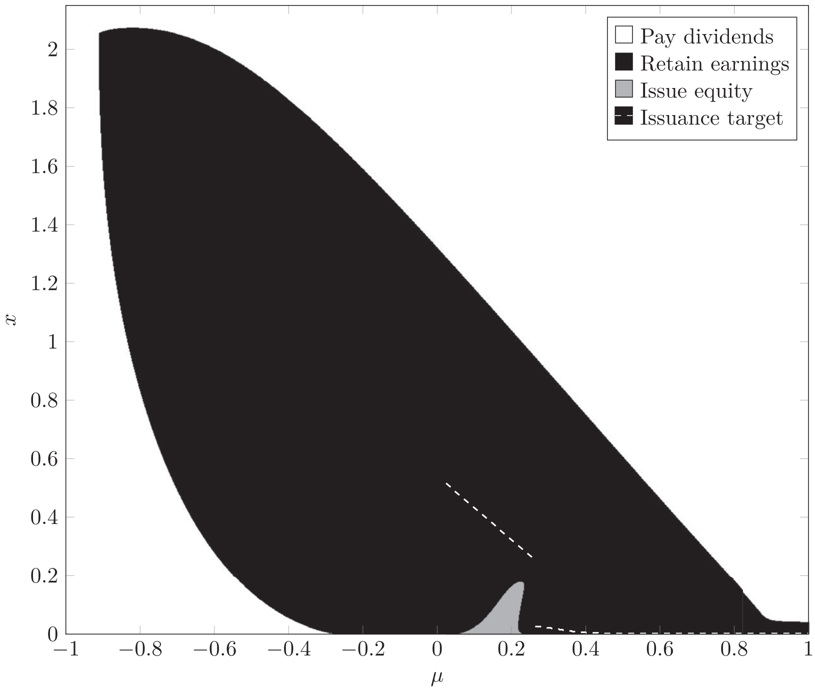

FIGURE 7 When issuance costs are sufficiently high for negative μ, it is optimal to issue equity away from the boundary $x = 0$

of surrounding points.[12] Such 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[13] satisfying $\mathbb{E}[g_k] = 0$ and $X_{\tau_k-} + g_k \geq 0$ P-a.s. This leads to the DPE

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

from which the last conditions makes it directly clear that the value function is now concave. However, as seen in Figure 8, this is not the concave envelope in the x-direction, because the *retained 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 $\mathcal{L}$.

[12] One example of such behavior by FedEx is described by Frock (2006, Chapter 18), one of the firm's co-founders.

[13] 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.

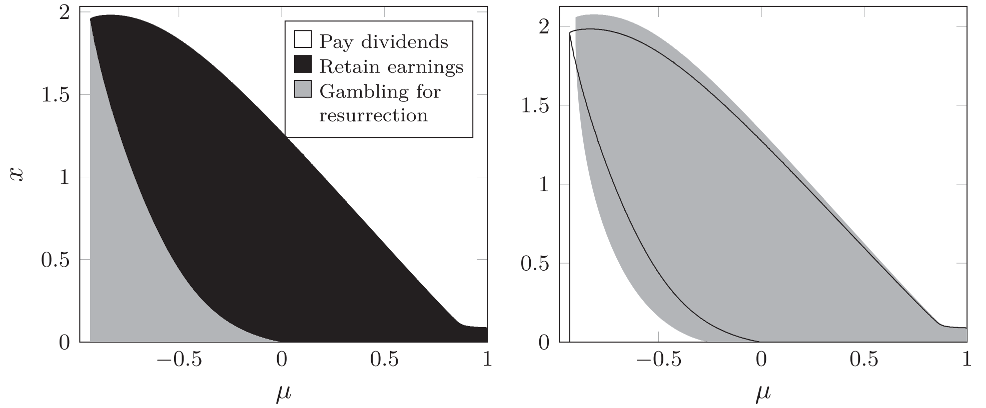

FIGURE 8 The figure on the left shows the model that allows gambling for resurrection. The figure on the right shows the model without gambling in gray, overlaid by the free boundaries of the gambling model in solid lines. When gambling is allowed, voluntary liquidation is no longer optimal in the "band region." This is because entering a "lottery" gives a chance of reaching a higher point in the no activity region, thus concavifying the problem

7 | CONCLUDING COMMENTS

In this paper, we have studied an optimal dividend problem where the profitability of the firm is stochastic. In this context, we have given a comparison result for general dynamics as well as outlined a numerical algorithm to compute the solution. The convergence of the numerical method is a consequence of the comparison result.

The numerical solution indicates that the problem exhibits both barrier and band structure, depending on the current profitability. The band structure is especially interesting, as it is unusual for Brownian dynamics. We also prove that when the profitability falls below a certain threshold, liquidation becomes optimal. Although mean reversion indicates that the profitability of the company will improve over time, that is not enough; when the initial profitability is too low, the expected cumulative losses are too large to be covered by the future income, so immediate closure is optimal.

The flexibility of the numerical method from Section 4 allows us to numerically study a number of extensions to our original model. The extended models allow for issuance with proportional and/or fixed costs as well as the possibility of gambling for resurrection. This last extension provides a way to concavify the problem (in the x-dimension) by giving the firm the opportunity of entering fair bets.

8 | PROOFS

This section is dedicated to the proofs 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 be satisfied without loss of generality where the lemma is needed.

Lemma 8.1. *If, in addition to Assumption 3.2,* $\inf \mathcal{M} > -\infty$ *or* $\mu \in O(\kappa(\mu))$ *as* $\mu \to -\infty$, *there exists a sublinearly growing function* H *and a constant* C *such that, for any stopping time* τ,[14]

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

[14] Recall that $(\mu_t^0)_{t\ge 0}$ is the process started in 0.

Proof. We will use the result by Peskir (2001) to obtain a function H that is sublinearly growing. For some $c \in \mathcal{M}^o$, 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'Hpital'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, Grnwall'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'Hpital's rule again

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

because $S'(\mu)/\mu S''(\mu) = \tilde{\sigma}(\mu)^2/(-2\kappa(\mu)\mu) \to 0$. Thus, with $H = F^{-1}$, Peskir (2001) 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. □

8.1 | Finiteness

In light of Lemma 8.1, the proof of the following theorem is natural.

Theorem 8.2. *The value function is finite.*

Proof. As $X_t^{(x,\mu),L} \geq 0$ for $t \leq \theta^{(x,\mu)}(L)$,

$$L_t \leq x + \int_0^t |\mu_s|\mathrm{d}s + \sigma W_t \leq x + |\mu_0|t + \int_0^t |\mu_s^0|\mathrm{d}s + \sigma W_t =: Y_t.$$

Note that this bound is uniform in L. Hence, with $\theta := \theta^{(x,\mu)}(L)$ it follows from integration by parts and Lemma 8.1 that

$$\begin{aligned} J(x,\mu;L) &= \mathbb{E}\left[L_\theta e^{-r\theta} - L_0 + r\int_0^\theta L_t e^{-rt}\mathrm{d}t\right] \\ &\leq \mathbb{E}\left[Y_\theta e^{-r\theta} - Y_0 + x + r\int_0^\theta Y_t e^{-rt}\mathrm{d}t\right] \\ &= \mathbb{E}\left[x + \int_0^\theta e^{-rt}\mathrm{d}Y_t\right] \\ &\leq x + \frac{|\mu_0|}{r} + \int_0^\infty e^{-rt}\mathbb{E}[|\mu_t^0|]\mathrm{d}t < \infty. \end{aligned}$$

8.2 | Liquidation threshold

Before we present the proof of Theorem 3.5, we present an auxiliary problem whose properties are especially useful to prove the existence of a liquidation threshold.

Consider the case where L is not restricted to be nondecreasing. This 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

$$V_a(x,\mu) = x + V_a(\mu),$$

where the auxiliary function V_a is given by

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

The liquidation time τ is chosen freely by the shareholders of the firm. This is a real option problem (see, for example, Dixit & Pindyck, 1994). 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, V_a and μ^* satisfies the following boundary value problem:

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

for all $\mu > \mu^*$, subject to

$$V_a(\mu) = 0,$$

for all $\mu \leq \mu^*$.

The second part of this theorem characterizes an optimal liquidation time. This proves useful in the original problem as $V(x,\mu) \leq V_a(x,\mu)$, thus showing that $V(x,\mu) \equiv x$ for levels of μ below some threshold.

Lemma 8.3. *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 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. Let L be any strategy for which $X_t^L \geq 0$ until some (liquidation) time τ. Then, define

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

Because 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.

It remains to prove the existence of μ^*. 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

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

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

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

We now try to find a point in which V_a is 0. Consider the equation

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

and suppose it has a solution that never touches 0, that is, $\phi > 0$. Then,

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

Using Itô's formula, the bounds on κ and $\tilde{\sigma}$, as well as (5), for $\mu < 0$, we obtain

$$\mathbb{E}[\mu_t] \leq \mu + \int_0^t C_1\left(1 + \mathbb{E}\left[\sup_{0\leq s\leq t}|\mu_s^0|\right]\right)\mathrm{d}s \leq \mu + tC_2(1 + H(t)),$$

for some constants C_1 and C_2. Hence,

$$\phi(\mu) \leq \int_0^\infty e^{-rt}\big(\mu + C_2 H(t)\big)\mathrm{d}t \leq \frac{\mu}{r} + C_3,$$

for another constant C_3. 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 V_a satisfies the dynamic programming equation (6). By Lemma 8.1 and Kobylanski and Quenez (2012), the optimal stopping time is the hitting time of $A_0 = \{\mu : V_a(\mu) = 0\} \neq \emptyset$. Hence, the function is smooth everywhere, except possibly at $\mu^* := \sup A_0$. However, as $\tilde{\sigma}$ never vanishes, an argument analogous to in the proof of Theorem 8.5 yields continuity also at μ^*, from which (6) can be derived.

Theorem 3.5. *If $\mathcal{M}$ has no lower bound, there exists a value μ^* such that it is optimal to liquidate immediately whenever $\mu \leq \mu^*$, that is, $V(x, \mu) \equiv x$.*

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 like in the proof of Theorem 3.4,

$$V(x,\mu) = \sup_{L \in \mathcal{A}(x,\mu)} \mathbb{E}\left[\int_0^{\theta(L)} e^{-rt} \mathrm{d}L_t\right] \leq x + \sup_{L \in \mathcal{A}(x,\mu)} \mathbb{E}\left[\int_0^{\theta(L)} e^{-rt} \mu_t \mathrm{d}t\right] + \mathbb{E}\left[\int_0^{\theta(L)} e^{-rt} \sigma \mathrm{d}W_t\right].$$

First observe that the last term is equal to 0. Then, because the second term is smaller than or equal to $V_a(\mu)$, the result is a direct consequence of Theorem 8.3. □

8.3 | Continuity

When proving the continuity of the value function, we need the following weak form of a dynamic programming *inequality*.

Lemma 8.4. *For any control* L *and stopping time τ with values between 0 and $\theta(L)$,*

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

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

Proof. By the definition of J,

$$\mathbb{E}\left[\int_\tau^{\theta(L)} e^{-r(t-\tau)} \mathrm{d}L_t \middle| \mathcal{F}_\tau\right] = J(X_\tau, \mu_\tau; L_{\cdot+\tau}), \quad P\text{-a.s.}$$

As for every x, μ, and $L \in \mathcal{A}(x, \mu)$

$$J(x, \mu; L) \leq V(x, \mu) \leq \bar{V}(x, \mu),$$

the lemma follows directly.

The necessity of this lemma stems from the a priori lack of the dynamic programming principle and the measurability of V. This inequality is much related to the weak dynamic programming principle (Bouchard & Touzi, 2011) that also establishes a similar inequality in the other direction. However, as seen in the proof above, this equality is more primitive than that in the other direction.

The 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 (Fleming & Soner, 2006). For the general case, we once again refer to Karoui and Tan (2013a, 2013b).

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

Theorem 8.5. *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. As $V \equiv 0$ 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

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

Now, make the observation that

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

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

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

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

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

As V is nonnegative and zero where $x = 0$,

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

Lemma 8.6. *For each starting point* (x, μ)*,*

$$\sup_{L\in\mathcal{A}(x,\mu)} \left\| \int_0^{\theta(L)} e^{-\gamma t}\mathrm{d}L_t \right\| < \infty,$$

for every $\gamma > 0$.

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

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

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, because, by Lemma 8.1,

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

Similarly, by Doob's inequality, Itô isometry, and Lemma 8.1, we obtain, for some C,

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

□

Theorem 3.6. *The value function is continuous everywhere.*

Proof. Consider a sequence (x^n, μ^n) that is nondecreasing in both coordinates and converges to some point (x^∞, μ^∞). The goal is to show that $V(x^\infty, \mu^\infty) - \varepsilon \le \lim_{n\to\infty} V(x^n, \mu^n)$ for any $\varepsilon > 0$. For the sake of readability, we introduce some notation. Let $\theta^n(L)$ be the ruin time of starting in (x^n, μ^n) and using the dividend policy L, and let $X^{n,L} = X^{(x^n,\mu^n),L}$ be the process associated with the starting point (x^n, μ^n) and dividend process L.

Denote by L^∞ an ε-optimal strategy starting in (x^∞, μ^∞). As L^∞ is the only strategy appearing in the first part of the proof, let θ^n and X^n are shorthand $\theta^n(L^\infty)$ and X^{n,L^∞}, respectively. Then, as $\Delta L^\infty_{\theta^n} = X^n_{\theta^n-} + \Delta L^\infty_{\theta^n} - X^n_{\theta^n-}$,

$$\begin{aligned} V(x^\infty, \mu^\infty) - \varepsilon \le \mathbb{E}\Bigg[&\int_{[0,\theta^n)} e^{-rt}\mathrm{d}L^\infty_t \\ &+ e^{-r\theta^n}\left(X^n_{\theta^n} + \Delta L^\infty_{\theta^n} - X^n_{\theta^n-}\right) + \int_{(\theta^n,\theta^\infty]} e^{-rt}\mathrm{d}L^\infty_t\Bigg] \\ \le\ & V(x^n, \mu^n) \\ &+ \mathbb{E}\left[e^{-r\theta^n}\left(\Delta L^\infty_{\theta^n} - X^n_{\theta^n-}\right)\right] + \mathbb{E}\left[\int_{(\theta^n,\theta^\infty]} e^{-rt}\mathrm{d}L^\infty_t\right]. \end{aligned} \tag{8}$$

The middle term is needed to cover the case that θ^n occurs as a result of a jump in L^∞. Note that it is nonnegative, because either θ^n is caused by a jump larger than $X^n_{\theta^n-}$, or the latter is zero. We want to show that the last two terms tend to zero as n tends to infinity.

By Lebesgue's dominated convergence theorem (see Lemma 8.6),

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

Note that the middle term is also dominated by an integrable expression, as it is itself dominated by the total discounted payoff of L^∞. Hence, we may take the limit inside both of the expectations.

We will prove that the limits inside the expectations are 0 on the following set

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

where $\bar{V}$ denotes the upper semicontinuous envelope of V. Note that by Lemma 8.1 and (7), $P(\Omega')=1$. For any $\omega\in\Omega'$, consider the following two cases.

First, consider any strictly increasing, divergent subsequence $\theta^{n(k)}=\theta^{n(k)}(L^\infty)(\omega)$. Then,

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

because $\omega\in\Omega'$ ensures that the integral factor is bounded, and the exponential factor converges to 0. Moreover, because $\theta^{n(k)}$ is strictly increasing,

$$\sum_{k\in\mathbb{N}} e^{-r\theta^{n(k)}}\Delta L^\infty_{\theta^{n(k)}}(\omega)\le\int_0^{\theta^\infty} e^{-rt}\mathrm{d}L_t^\infty(\omega)<\infty,$$

which shows that the sum is convergent, so its terms must converge to zero. In other words,

$$\lim_{k\to\infty} e^{-r\theta^{n(k)}}\left(\Delta L^\infty_{\theta^{n(k)}}-X^n_{\theta^{n(k)}-}\right)\le\lim_{k\to\infty} e^{-r\theta^{n(k)}}\Delta L^\infty_{\theta^{n(k)}}=0.$$

On the other hand, let $\theta^{n(k)}=\theta^{n(k)}(L^\infty)(\omega)$ be a bounded subsequence. Then, because $\sup_k\theta^{n(k)}(\omega)<\infty$,

$$X^\infty_{\theta^{n(k)}-}-X^{n(k)}_{\theta^{n(k)}-}=x^\infty-x^{n(k)}+\int_0^{\theta^{n(k)}}\mu_t^\infty-\mu_t^{n(k)}\mathrm{d}t\xrightarrow{k\to\infty}0,$$

because of continuity with respect to initial points. Thus,

$$0\le\Delta L^\infty_{\theta^{n(k)}}-X^{n(k)}_{\theta^{n(k)}-}\le X^\infty_{\theta^{n(k)}-}-X^{n(k)}_{\theta^{n(k)}-}\xrightarrow{k\to\infty}0.$$

For the other term, first observe that

$$0\le X^\infty_{\theta_k}=X^\infty_{\theta_k-}-\Delta L^\infty_{\theta_k}+X^{n(k)}_{\theta^{n(k)}-}-X^{n(k)}_{\theta^{n(k)}-}\le X^\infty_{\theta_k-}-X^{n(k)}_{\theta^{n(k)}-}\xrightarrow{k\to\infty}0.$$

We wish to use the continuity of V, and therefore also of $\bar{V}$, at 0 to argue that

$$\lim_{k\to\infty}\bar{V}\left(X^\infty_{\theta^{n(k)}}(\omega),\mu^\infty_{\theta^{n(k)}}(\omega)\right)=0.$$

As V is increasing, it is sufficient to find a bound to $\mu^\infty_{\theta^{n(k)}}(\omega)$. Begin by considering the process $M_t^\infty=\sup_{0\le s\le t}\mu_s^\infty$. Then, because $\theta^{n(k)}$ is a bounded sequence and M_t^∞ is continuous, $M^\infty_{\theta^{n(k)}}(\omega)$ is bounded by some constant C. Therefore, by Theorem 8.5,

$$\lim_{k\to\infty}\bar{V}\left(X^\infty_{\theta^{n(k)}}(\omega),\mu^\infty_{\theta^{n(k)}}(\omega)\right)\le\lim_{k\to\infty}\bar{V}\left(X^\infty_{\theta^{n(k)}}(\omega),C\right)=\bar{V}(0,C)=0.$$

By the preceding arguments, it holds that for any bounded subsequence and any divergent subsequence,

$$\left(e^{-r\theta^n}\left(\Delta L^\infty_{\theta^n}-X^n_{\theta^n-}\right)+\int_{(\theta^n,\theta^\infty]} e^{-rt}\mathrm{d}L_t^\infty\right)(\omega)$$

converges to 0 for every $\omega \in \Omega'$. As a consequence of this, the whole sequence converges to 0, for every $\omega \in \Omega'$, that is, P-a.s. Returning to (8), this yields

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

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

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

$$\begin{aligned}\lim_{n\to\infty} V(x^n, \mu^n) - \varepsilon &\leq V(x^\infty, \mu^\infty)\\ &+\mathbb{E}\left[\lim_{n\to\infty} e^{-r\theta^\infty(L^n)}\left(\Delta L^n_{\theta^\infty(L^n)} - X^{\infty,L^n}_{\theta^\infty(L^n)}\right)\right] + \mathbb{E}\left[\lim_{n\to\infty}\int_{(\theta^\infty(L^n),\theta^n(L^n)]} e^{-rt}\mathrm{d}L^n_t\right],\end{aligned}$$

for ε-optimal strategies L^n starting in (x^n, μ^n). 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,

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

As 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 nondecreasing and nonincreasing, respectively, it follows that

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

and V is continuous everywhere. □

8.4 | Comparison principle

Lemma 8.7. *If a function* u *is a viscosity subsolution to (2), then*

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

is a viscosity subsolution to

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

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 (2). 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 8.8. *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 differentiable function with

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

for strictly positive constants η_- and η_+ as well as $\mu \in [-y, y]^c$.[15] 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, as 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 8.9. Assumption 3.2 could possibly be generalized by finding strict supersolutions to the equation

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

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

Theorem 3.7 (Comparison). *Let u and v be upper and lower semicontinuous, polynomially growing viscosity sub- and supersolutions of (2). Then, under Assumption 3.3, $u \le v$ for $(x, \mu) \in 0 \times \mathcal{M}$ implies that $u \le v$ everywhere in $\mathcal{O} := \mathbb{R}_{\ge 0} \times \mathcal{M}$.*

Proof of Theorem 3.7. Comparison is shown for the transformed DPE (11) with η chosen as in Lemma 8.8. Because the transformation (10) is sign-preserving, this is sufficient to establish comparison for (2) thanks to Lemma 8.7. For the sake of simpler presentation later on, (11) is shortened to

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

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

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 to argue that, without loss of generality, we may assume that the maximum of $\tilde{u} - \tilde{v}$ is attained in the interior $\mathcal{O}^{\mathrm{o}}$:

1. The function $\tilde{u} - \tilde{v}$ attains a maximum in $[0, \infty) \times \mathcal{M}^{\mathrm{o}}$. If the maximum is at $x = 0$, we are done. Otherwise, the maximum is attained in the interior $\mathcal{O}^{\mathrm{o}}$.
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 Amadori (2007). 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

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

with

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

Note that by Assumption 3.3, $C < \infty$. 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 $\mathcal{O}^{\circ}$.

By the preceding argument, we may assume that $\tilde{u} - \tilde{v}$ attains local maximum $(x_0, \mu_0) \in \mathcal{O}^{\circ}$. Let $N \subseteq N' \subseteq \mathcal{O}$ be two neighborhoods of (x_0, μ_0) in which (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 $\overline{N} \times \overline{N}$ that 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

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

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

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

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

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

and

$$\min\left\{ f(\nu_\epsilon)\tilde{v}(y_\epsilon, \nu_\epsilon) - f^x(\nu_\epsilon)\frac{x_\epsilon - y_\epsilon}{\epsilon} - f^\mu(\nu_\epsilon)\frac{\mu_\epsilon - \nu_\epsilon}{\epsilon} - \operatorname{Tr}\Sigma(\nu_\epsilon)\tilde{Y}_\epsilon, \right.$$
$$\left. \eta\tilde{v}(y_\epsilon, \nu_\epsilon) + \frac{x_\epsilon - y_\epsilon}{\epsilon} - e^{-\eta y_\epsilon - \eta g(\nu_\epsilon)} \right\} \geq 0. \tag{12}$$

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

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

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

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

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 $\tilde{u}-\tilde{v}$ attains its maximum on the interior. The theorem statement follows from the fact that

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

□

REFERENCES

Akyildirim, E., Güney, I. E., Rochet, J.-C., & Soner, H. M. (2014). Optimal dividend policy with random interest rates. *Journal of Mathematical Economics*, *51*, 93–101.

Albrecher, H., & Cani, A. (2017). Risk theory with affine dividend payment strategies. In C. Elsholtz & P. Grabner (Eds.). *Number theory–diophantine problems, uniform distribution and applications* (pp. 25–60). Cham, Switzerland: Springer.

Albrecher, H., & Thonhauser, S. (2009). Optimality results for dividend problems in insurance. *RACSAM-Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas*, *103*(2), 295–320.

Altarovici, A., Reppen, M., & Soner, H. M. (2017). Optimal consumption and investment with fixed and proportional transaction costs. *SIAM Journal on Control and Optimization*, *55*(3), 1673–1710.

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

Anderson, R. W., & Carverhill, A. (2012). Corporate liquidity and capital structure. *Review of Financial Studies*, *25*(3), 797–837.

Asmussen, S., & Albrecher, H. (2010). *Ruin probabilities (advanced series on statistical science & applied probability)* (Vol. 14). Singapore, Singapore: World Scientific.

Avanzi, B., & Wong, B. (2012). On a mean reverting dividend strategy with brownian motion. *Insurance: Mathematics and Economics*, *51*(2), 229–238.

Barles, G., & Souganidis, P. E. (1991). Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Analysis*, *4*(3), 271–283.

Barth, A., Moreno-Bromberg, S., & Reichmann, O. (2016). A non-stationary model of dividend distribution in a stochastic interest-rate setting. *Computational Economics*, *47*(3), 447–472.

Bolton, P., Chen, H., & Wang, N. (2013). Market timing, investment, and risk management. *Journal of Financial Economics*, *109*(1), 40–62.

Bouchard, B., & Touzi, N. (2011). Weak dynamic programming principle for viscosity solutions. *SIAM Journal on Control and Optimization*, *49*(3), 948–962.

Cadenillas, A., Sarkar, S., & Zapatero, F. (2007). Optimal dividend policy with mean-reverting cash reservoir. *Mathematical Finance*, *17*(1), 81–109.

Cai, J., Gerber, H. U., & Yang, H. (2006). Optimal dividends in an Ornstein–Uhlenbeck type model with credit and debit interest. *North American Actuarial Journal*, *10*(2), 94–108.

Chiarolla, M. B., & Haussmann, U. G. (1994). The free boundary of the monotone follower. *SIAM Journal on Control and Optimization*, *32*(3), 690–727.

De Angelis, T., Federico, S., & Ferrari, G. (2017). Optimal boundary surface for irreversible investment with stochastic costs. *Mathematics of Operations Research*, *42*(4), 1135–1161.

de Finetti, B. (1957). Su un'impostazione alternativa della teoria collettiva del rischio. In *Transactions of the XVth international congress of Actuaries* (Vol. 2, pp. 433–443). New York, NY: Actuarial Society of America.

Décamps, J.-P., Mariotti, T., Rochet, J.-C., & Villeneuve, S. (2011). Free cash flow, issuance costs, and stock prices. *Journal of Finance*, *66*(5), 1501–1544.

Dixit, A. K., & Pindyck, R. S. (1994). *Investment under uncertainty*. Princeton, NJ: Princeton University Press.

Fama, E. F., & French, K. R. (2000). Forecasting Profitability and earnings. *Journal of Business*, *73*(2), 161–175.

Ferrari, G. (2015). On an integral equation for the free-boundary of stochastic, irreversible investment problems. *Annals of Applied Probability*, *25*(1), 150–176.

Ferrari, G. (2018). On the optimal management of public debt: A singular stochastic control problem. *SIAM Journal on Control and Optimization*, *56*(3), 2036–2073.

Fleming, W. H., & Soner, H. M. (2006). *Controlled Markov processes and viscosity solutions* (Vol. 25). Berlin, Germany: Springer Science & Business Media.

Frock, R. (2006). *Changing how the world does business*. Oakland, CA: Berrett-Koehler.

Gerber, H.-U. (1969). *Entscheidungskriterien für den zusammengesetzten Poisson-Prozess* (Doctoral thesis, ETH Zurich).

Gertler, M., & Hubbard, R. G. (1991). *Corporate financial policy, taxation, and macroeconomic risk*. Technical Report. Cambridge, MA: National Bureau of Economic Research.

Hackbarth, D., Miao, J., & Morellec, E. (2006). Capital structure, credit risk, and macroeconomic conditions. *Journal of Financial Economics*, *82*(3), 519–550.

Hugonnier, J., Malamud, S., & Morellec, E. (2015). Capital supply uncertainty, cash holdings, and investment. *Review of Financial Studies*, *28*(2), 391–445.

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

Jeanblanc-Picqué, M., & Shiryaev, A. N. (1995). Optimization of the flow of dividends. *Russian Mathematical Surveys*, *50*(2), 257.

Jiang, Z., & Pistorius, M. (2012). Optimal dividend distribution under Markov regime switching. *Finance and Stochastics*, *16*(3), 449–476.

Karoui, N. E., & Tan, X. (2013a). Capacities, measurable selection and dynamic programming part I: abstract framework. *arXiv:1310.3363*.

Karoui, N. E., & Tan, X. (2013b). Capacities, measurable selection and dynamic programming part II: application in stochastic control problems. *arXiv:1310.3364*.

Kobylanski, M., & Quenez, M.-C. (2012). Optimal stopping in a general framework. *Electronic Journal of Probability*, *17*(72), 1–28.

Kolodner, I. I. (1956). Free boundary problem for the heat equation with applications to problems of change of phase. I. general method of solution. *Communications on Pure and Applied Mathematics*, *9*(1), 1–31.

Liang, Z., & Young, V. R. (2012). Dividends and reinsurance under a penalty for ruin. *Insurance: Mathematics and Economics*, *50*(3), 437–445.

Lundberg, F. (1903). *1. Approximerad framställning af sannolikhetsfunktionen: 2. Återförsäkring af kollektivrisker. Akademisk afhandling*. Almqvist & Wiksells.

Murto, P., & Terviö, M. (2014). Exit options and dividend policy under liquidity constraints. *International Economic Review*, *55*(1), 197–221.

Myers, S. C. (1977). Determinants of corporate borrowing. *Journal of Financial Economics*, *5*(2), 147–175.

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

Puterman, M. L. (1994). *Markov decision processes: discrete stochastic dynamic programming*. Hoboken, NJ: John Wiley & Sons.

Radner, R., & Shepp, L. (1996). Risk vs. profit potential: A model for corporate strategy. *Journal of economic dynamics and Control*, *20*(8), 1373–1393.

Revuz, D., & Yor, M. (1999). *Continuous martingales and Brownian motion*. Grundlehren der mathematischen Wissenschaften (Vol. 293). Berlin/Heidelberg, Germany: Springer.

Shreve, S. E., & Soner, H. M. (1991a). A free boundary problem related to singular stochastic control: The parabolic case. *Communications in partial differential equations*, *16*(2-3), 373–424.

Shreve, S. E., & Soner, H. M. (1991b). A free boundary problem related to singular stochastic control. *Applied Stochastic Analysis*, *5*, 265–301.

Shreve, S. E., Lehoczky, J. P., & Gaver, D. P. (1984). Optimal consumption for general diffusions with absorbing and reflecting barriers. *SIAM Journal on Control and Optimization*, *22*(1), 55–75.

Soner, H. M., & Shreve, S. E. (1989). Regularity of the value function for a two-dimensional singular stochastic control problem. *SIAM Journal on Control and Optimization*, *27*(4), 876–907.

How to cite this article: Reppen AM, Rochet J-C, Soner HM. Optimal dividend policies with random profitability. *Mathematical Finance*. 2019;1–32. https://doi.org/10.1111/mafi.12223