Optimal Dividend Policy with Random Interest Rates

Erdinç Akyildirim[*] I. Ethem Güney [†] Jean Charles Rochet[‡§]
H. Mete Soner[¶]

November 30, 2013

Abstract

Several recent papers have studied the impact of macroeconomic shocks on the financial policies of firms. However they only consider the case where these macroeconomic shocks affect the profitability of firms but not the financial markets conditions. We study the polar case where the profitability of firms is stationary, but interest rates and issuance costs are governed by an exogenous Markov chain. We characterize the optimal dividend policy and show that these two macroeconomic factors have opposing effects: all things being equal, firms distribute more dividends when interest rates are high and less when issuing costs are high.

Key words: Dividend Policy, Business Cycles, Financial Frictions.

JEL classifications: $G35$, $E32$, $C61$.

---

[*]University of Zürich, Swiss Finance Institute

[†]University of Zürich, Swiss Finance Institute

[‡]University of Zürich, Swiss Finance Institute and Toulouse School of Economics, Research partly supported by the European Research Council under the grant 249415-RMAC and NCCR Finrisk (project on Banking and Regulation).

[§]Corresponding author. Tel.: +41 44 634 40 55; fax: +41 44 635 57 05. E-mail address: jean-charles.rochet@bf.uzh.ch.

[¶]ETH (Swiss Federal Institute of Technology), Zürich, and Swiss Finance Institute, hmsoner@ethz.ch. Research partly supported by the European Research Council under the grant 228053-FiRM and by the ETH Foundation.

1 Introduction

Since Jeanblanc-Picqué & Shiryaev [8] and Radner & Shepp [10], a sizable literature has investigated the optimal dividend policy problem for a company that is not allowed to issue new securities or obtain a new loan from a bank. The default time is then defined as the first time when the cash reserves of the company fall below zero. In that case, the optimal dividend policy is simple and natural: distribute dividends whenever the level of cash reserves exceeds a certain threshold that depends on the characteristics (drift, volatility) of the cash flow process and the interest rate demanded by shareholders.

An interesting extension of this problem is to investigate how the optimal dividend policy is modified when the profitability of the firm changes over time, due in particular to business cycle fluctuations. As clearly shown by Gertler & Hubbard [5] and more recently by Hackbart, Miao and Morellec [6], macroeconomic conditions have indeed a strong impact on dividend policies through the changes in the profitability of individual firms that they induce. For example, Cadenillas & Sotomayor [2] solve for the optimal dividend policy when the drift and the volatility of the cash flow process are governed by a Markov chain representing macroeconomic fluctuations. Bolton, Chen & Wang [1] study more generally the impact of changing macroeconomic conditions on both the financial and investment policies of the firms. However, Gertler & Hubbard [5] also show that macroeconomic conditions directly influence payments to shareholders, even independently of each firm's specific earnings performance. Two natural channels for this influence are the fluctuations in interest rates demanded by investors, and the conditions of the credit market.

The purpose of this paper is to examine how these macroeconomic fluctuations influence the dividend policies of firms, even in the absence of fluctuations in their earning processes. In other words, we study the polar case to the one considered in the literature: the drift and volatility of the cash flow process are constant, but the interest rate demanded by investors follows a Markov chain. In a recent paper, Jiang and Pistorius [9] consider a similar case where both the profitability of the firm and the discount factor follow a Markov chain. Our paper differs in two respects from Jiang and Pistorius [9]. First we adopt direct approach: we solve the couple of ODEs that characterize the solution by using standard numerical techniques. By contrast, Jiang and Pistorius [9] characterize the solution as the fixed point of a functional operator and find this solution by an iterative algorithm. The second, and more important, difference between our paper and Jiang and Pistorius [9] is that we allow the firm to issue new securities. This possibility is not only realistic, but it also leads to two non-trivial consequences: the ranking of optimal dividend thresholds across the two states is not always the same; issuance may be optimal even when cash reserves are still positive. This shows that introducing possibilities of new issuances is not just a trivial extension, but gives rise to new, economically relevant, results.

Section 2 presents the model and the mathematical characterization of the optimal dividend policy (Theorem 2.1). Section 3 establishes several important properties of the value function. In subsection 3.1, we show that the value function remains concave in the level of cash holdings, even when interest rates are stochastic (Theorem 3.1). The concavity of the value function allows us to prove that it is a smooth solution of the corresponding dynamic programming equation (Proposition 3.1). In particular, it satisfies the *smooth fit* condition which is crucial in the numerical resolution of these types of problems. These mathematical results are necessary to establish an important economic result in subsection 3.3: the firm will distribute dividends more often when interest rates are high than when they are low (Proposition 3.2). This result comes from the fact that the opportunity cost of cash reserves is higher when the interest rates demanded by investors are high. However, it does not fit well with the empirical evidence, given that firms actually tend to distribute less dividends during recessions (when interest rates are high) than during booms (when interest rates are low) even when the changes in firms' individual profitability are corrected for (Gertler & Hubbard [5]). This suggests that other macroeconomic factors, such as the size of frictions on financial markets, must play a role. This is why section 4 introduces the possibility for the firm to make new equity issuances. When the cost of these new issues (a proxy for the size of financial frictions) is substantially higher during recessions than during booms, the ranking of dividend thresholds is reversed, and firms now distribute more dividends during booms than during recessions.

We also provide numerical evidence for the above conclusions. In particular, in subsection 3.4, the sensitivity analysis with respect to mean and volatility of the cash flow rate and jump rates between two different interest rate regimes are presented. The mathematical results proved in Section 3 are also essential in constructing and verifying the numerical algorithm. Section 4 gives several numerical illustrations of the case where new equity issuance is possible.

2 Model and Characterization of the Solution

Uncertainty is described by $(\Omega, \mathbb{F}, \mathbb{P})$, a filtered probability space satisfying the usual assumptions[1]. Let B_t be a one-dimensional standard Brownian motion and $\{i_t\}_{t\geq 0}$ be a simple stationary Markov process taking values in $\{0, 1\}$ with jump rates $\lambda(0), \lambda(1) > 0$. The process $\{i_t\}_{t\geq 0}$ is assumed to be independent from the Brownian motion. The state $i = 0$ is the "good" economic state with a lower interest rate $r_\ell > 0$ and $i = 1$ corresponds to the "bad" state with interest rate $r_h > r_\ell > 0$. We also set $\lambda_\ell := \lambda(0)$ and $\lambda_h := \lambda(1)$.

The cash holdings $\{X_t\}_{t\geq 0}$ of the company follow a diffusion process. Positive dividend payments of any size can be made at any time. However, the cash level is supposed to remain nonnegative at all times. This constraint clearly places a restriction on the possible

[1] See [7] for details.

dividend size. Mathematically,

$$dX_t = \mu dt + \sigma dB_t - dL_t, \tag{2.1}$$

where $\mu, \sigma > 0$ are given constants and the *cumulative dividend payments* L_t is an adaptive, nondecreasing, càdlàg process with $L_{0^-} = 0$. Given a dividend process L and an initial condition $x \in \mathbb{R}$, let $X^{x,L}$ be the unique solution of (2.1), i.e.,

$$X_t^{x,L} = x + \mu t + \sigma B_t - L_t, \quad t \geq 0.$$

Let $\theta = \theta^{x,L}$ be the first exit time of $X^{x,L}$ from the positive real line. This variable θ defines the time of bankruptcy. In what follows we will suppress the dependence on x, L unless this dependence is important. We say that L is *admissible* at the initial level x, if $X_t^{x,L} \geq 0$, for all time $t \in [0, \theta^{x,L}]$ with probability one. We denote the set of all admissible strategies by $\mathcal{A}(x)$. We note that the admissibility condition is relevant only at the exit time. Indeed, we only require that the cash level process does not jump into negative real line. In economic terms, this means that shareholders can never distribute themselves a dividend that exceeds the cash holdings of the firm. Hence, $X_\theta^{x,L} = 0$. Since the dividend policy beyond the exit time is irrelevant, we simply set $L_t = L_\theta$ for all $t \geq \theta$. In particular, $L_\theta - L_{\theta^-} = X_{\theta^-}$.

The optimal dividend problem is to maximize

$$J(x, i, L) := \mathbb{E}\left[\int_0^\theta \Lambda_t \, dL_t \;\middle|\; i_0 = i, X_{0^-} = x\right], \quad \Lambda_t := \exp\left(-\int_0^t r(i_u) du\right).$$

The *value function* is then defined by

$$v(x, i) := \sup_{\mathcal{A}(x)} J(x, i, L), \quad v_\ell(x) := v(x, 0), \quad v_h(x) := v(x, 1). \tag{2.2}$$

The case of a deterministic (and constant) interest rate (i.e., $r_\ell = r_h$) is exactly the problem studied by Picqué-Jeanblanc & Shiryayev [8] and Radner & Shepp [10]. For future reference, we record that the value function with constant interest rate r is given by

$$V(x, r) := \sup_{L \in \mathcal{A}(x)} \mathbb{E}\left[\int_0^\theta e^{-rt} dL_t | X_{0^-} = x\right]. \tag{2.3}$$

Then, it is clear that

$$0 \leq V(x, r_h) \leq v_h(x) \leq v_\ell(x) \leq V(x, r_l), \qquad \forall\, x \in \mathbb{R}^+. \tag{2.4}$$

2.1 Characterization of the Solution

Our main mathematical result is the following characterization of the value function. The existence part of this theorem will be proved in several steps in the subsequent sections. The uniqueness follows from the classical verification argument (see for instance [4]). This characterization of the value function and the properties of the thresholds are essential in our numerical experiments. Indeed, the numerical algorithm is based on these properties. Moreover, the uniqueness ensures that the computed functions are in fact equal to the value function.

Theorem 2.1 *The value function* $v = (v(\cdot, 0), v(\cdot, 1)) = (v_h, v_\ell)$ *is the unique concave function satisfying the following conditions:*

- $v_\ell, v_h \in C^2([0, \infty)$ *and* $v_l(0) = v_h(0) = 0$*;*
- $v'(x, i) \geq 1$ *for all* x*;*
- *For every* $x > 0$ *and* $i \in \{0, 1\}$, $r(i)v(x, i) - \mathcal{L}v(x, i) \geq 0$, *where*

$$\mathcal{L}v(x, i) := \mu v'(x, i) + \frac{\sigma^2}{2} v''(x, i) + \lambda(i)[v(x, i + 1) - v(x, i)]; \tag{2.5}$$

  with the convention that $i + 1$ *denotes the other state than* i.

- *There are two positive thresholds* $0 < x_h := x(1)$ *and* $x_\ell := x(0) < \infty$ *such that*

$$v'(x, i) = 1, \quad \text{for } x \geq x(i), \quad \text{and} \quad r(i)v(x, i) - \mathcal{L}v(x, i) = 0, \quad \text{for } x \leq x(i).$$

The above characterization of the value function also provides the structure of the optimal dividend policy. The optimal dividend policy is simple: only distribute dividends when cash holdings exceed threshold $x(i)$, which depends on the state i of the economy. This is done exactly as in the deterministic interest rate case. Namely, if the initial cash holdings x exceed $x(i)$, then an initial dividend of $x - x(i)$ is distributed. In later times, dividends are paid only when the cash holdings reach $x(i)$ again. When the state of the economy changes from good to bad (equivalently when i jumps from zero to one), then cash holdings may be larger than $x(1)$ and a dividend payment of the difference is optimal. Then, one proceeds as before.

The above theorem also proves that the value function is a classical solution of the dynamic programming equation,

$$\min \left\{ \ r(i)v(x, i) - \mathcal{L}v(x, i) \ , \ v'(x, i) - 1 \ \right\} = 0, \qquad x > 0, \ i = 1, 2, \tag{2.6}$$

together with boundary condition $v(0, i) = 0$.

2.2 Elementary Properties

In this subsection, we prove several simple properties.

Lemma 2.1 *The value function v is Lipschitz continuous at the origin and*

$$v(0,i)=0, \quad v(x+y,i)\geq v(x,i)+y, \quad \forall\ x,y\geq 0,\ i=0,1.$$

Proof. Since σ is not null, the only admissible process at $x=0$ is $L=0$. This proves that $v(0,i)=0$. We also emphasize that at time zero, L^y has a jump of size at least y. Also, for any given (x,y) and $L\in\mathcal{A}(x)$, we set $L_t^y := L_t + y$ for $t\geq 0$ (with, as it is required $L_{0-}^y=0$).

Then, if one starts with cash holdings $x+y$ at $t=0$ and uses the dividend policy L^y, cash holdings are characterized by $\{\hat{X}_t\}_{t\geq 0}$ defined by

$$\begin{aligned}\hat{X}_t &:= X_t^{x+y,L^y} = x+y+\mu t+\sigma W_t - L_t^y \\ &= x+\mu t+\sigma W_t - L_t = X_t^{x,L} =: X_t,\end{aligned}$$

for all $t\geq 0$. In particular, the exit time $\hat{\theta}$ of $\hat{X}$ from $(0,\infty)$ is the same as that of X. Hence,

$$v(x+y,i)\geq J(x+y,i,L^y)=\mathbb{E}\left[\int_0^\theta \Lambda_t dL_t^y\right] = y+\mathbb{E}\left[\int_0^\theta \Lambda_t dL_t\right].$$

Since $L\in\mathcal{A}(x)$ is arbitrary,

$$v(x+y,i)\geq y+v(x,i), \quad \forall\ (x,y)\in\mathbb{R}^+,\ i=0,1.$$

Recall the deterministic value function defined in (2.3) and the inequality (2.4). Hence for any $x\geq 0$ and i,

$$V(0,r_\ell)=v(0,i)=0\leq v(x,i)\leq V(x,r_\ell).$$

The function V is known explicitly (see [8]) and it is Lipschitz continuous. Hence, v is Lipschitz continuous at the origin, i.e., there is a constant K such that

$$0=v(0,i)\leq v(x,i)\leq Kx$$

for all $x\geq 0$. □

In this context, the standard dynamic programming principle states that for any initial point (x,i) and any stopping time $\tau\leq\theta$,

$$v(x,i)=\sup_{L\in\mathcal{A}(x)}\mathbb{E}\left[\int_0^\tau \Lambda_t dL_t + \Lambda_\tau\ v\left(X_\tau^{x,L}, i_\tau\right)\right]. \tag{2.7}$$

Our next result, is a step towards proving the concavity of the value function. Indeed, the concavity is equivalent to the condition (2.8) below with $c_0=0$.

Lemma 2.2 *There exists a constant $c_0 > 0$ such that for all $0 \le x < y$ and $i \in \{0, 1\}$,*

$$v(x, i) + v(y, i) - 2v((x + y)/2, i) \le c_0. \tag{2.8}$$

Proof. Recall the value function defined in (2.3) and the inequality (2.4). Then,

$$v(x, i) + v(y, i) - 2v((x + y)/2, i) \le V(y, r_\ell) + V(x, r_\ell) - 2V((x + y)/2, r_h).$$

The function V is known explicitly and such that there exists a constant $c(r) > 0$ so that

$$x \le V(x, r) \le c(r) + x, \quad \forall\ x, r > 0.$$

We now combine the two inequalities to obtain,

$$v(x, i) + v(y, i) - 2v((x + y)/2, i) \le [c(r_\ell) + x] + [c(r_\ell) + y] - 2((x + y)/2) \le 2c(r_\ell).$$

□

Indeed, the viscosity property is proved exactly as in Theorem 5.1, page 311 in [4]. Moreover, the uniqueness of this solution can be proved by the techniques developed in [4]. But this result is not needed in this paper.

Lemma 2.3 *The value function is a continuous viscosity solution of the dynamic programming equation* (2.6)*.*

3 Value Function

In this section, we establish several important properties of the value function.

3.1 Concavity

In this section, we prove that the value function is concave. We start by showing this is true in an interval near the origin.

Lemma 3.1 *There exists $x_0 > 0$ such that for both $i = 0, 1$,*

$$-v''(\cdot, i) \ge 0, \quad \text{on } (0, x_0),$$

in the viscosity sense.

Proof. We first choose $x_0 > 0$ so that

$$|r(i)v(x, i) - \lambda(i)[v(x, i + 1) - v(x, i)]| \le \mu, \quad \forall\ x \in [0, x_0], i \in \{0, 1\}.$$

This is possible as v is continuous at the origin with value zero.

We need to show that for $\varphi(.,i) \in C^2(\mathbb{R})$ for each i, which depends on the state of the economy i, if

$$(v-\varphi)(x^*,i) = localmin(v-\varphi)(\cdot,i)$$

at some $x^* \in (0,x_0)$, then $\varphi''(x^*) \leq 0$.

Indeed, let φ be as above. Then, by the viscosity supersolution property of v we have

$$r(i)v(x^*,i) - \mu\varphi'(x^*) - \frac{\sigma^2}{2}\varphi''(x^*) - \lambda(i)[v(x^*,i+1) - v(x^*,i)] \geq 0,$$

and $\varphi'(x^*) \geq 1$. Hence,

$$-\varphi''(x^*) \geq \frac{1}{\sigma^2}\left(-r(i)v(x^*,i) + \mu + \lambda(i)[v(x^*,i+1) - v(x^*,i)]\right).$$

By the choice of x_0, the right hand side of the above inequality is non-negative. Therefore, $-\varphi'' \geq 0$. □

The following is an immediate corollary of the above Lemma.

Corollary 3.1 *There exists $x^* > 0$ such that $v(\cdot,i)$ is concave on $[0,x^*]$ and*

$$v'(x,i) \geq v'(x^*,i) > 1, \quad \forall\ i \in \{0,1\},\ x \in [0,x^*].$$

Proof: The concavity of v near the origin follows from the previous results and the theory of viscosity solutions. Also

$$v(h,i) = v(h,i) - v(0,i) \geq V(h,r_h) > (1+\delta)h,$$

for some $\delta > 0$. Hence, $v'(0,i) \geq 1+\delta$. Set

$$x^* = \sup\{x:\ v(\cdot,i) \textit{ is concave on } [0,x] \textit{ and } v'(x,i) \geq 1+\delta/2\}.$$

Then, it is clear that $x^* > 0$. □

The following is proved in the Appendix A.

Theorem 3.1 *$v(\cdot,i)$ is concave for $i \in \{0,1\}$.*

3.2 Smooth Fit

In this section, we use the concavity of the value function to show that it is twice continuously differentiable. This statement is equivalent to the smooth fit property at the thresholds. The smoothness of the value function immediately implies that it is a classical solution of the dynamic programming equation (2.6).

Proposition 3.1 (Smooth Fit) *The value function is twice continuously differentiable in the x variable.*

Proof. Set

$$x(i) = inf\{x:\ 1 \in \partial v(x,i)\}, \quad i = 0,1 \tag{3.1}$$

where $\partial v(x,i)$ denotes the subdifferential of $v(\cdot,i)$ at x (we refer reader to [11] for the definition and the properties of subdifferentials of convex functions). By Lemma 2.1 $x(i) > 0$. Also, since $v' \geq 1$ in the viscosity sense, concavity of v implies,

$$v'(x,i) = 1, \quad \forall\ x \geq x(i), \qquad \text{and} \qquad v'(x,i) > 1, \quad \forall\ x \in [0, x(i)).$$

Then, since v satisfies the dynamic programming equation (2.6),

$$r(i)v(x,i) - \mathcal{L}v(x,i) = 0 \quad \forall x \in (0, x(i)),$$

the elliptic regularity implies that

$$v(\cdot,i) \in C^\infty((0,x(i))).$$

Step 1. First, we show that $\partial v(x(i),i) = \{1\}$.

Suppose to the contrary that

$$\partial v(x(i),i) = [1,p]$$

for some $p > 1$. Then, for any $\varepsilon > 0$, it is straightforward to construct a smooth test function φ_ε so that

$$\sup(v(\cdot,i) - \varphi_\varepsilon(\cdot)) = v(x(i),i) - \varphi_\varepsilon(x(i)) = 0,$$

$\varphi''_\varepsilon(x(i)) = -1/\varepsilon$ and $\varphi'_\varepsilon(x(i)) \in (1,p)$. The viscosity property of $v(\cdot,i)$ implies that

$$r(i)v(x(i),i) - \mu\varphi'_\varepsilon(x(i)) - \frac{\sigma^2}{2}\varphi''_\varepsilon(x(i)) - \lambda(i)[v(x(i),i+1) - v(x(i),i)] \leq 0.$$

For $\varepsilon > 0$ sufficiently small, this is a contradiction. Hence, $\partial v(x(i),i)$ is a singleton $\{1\}$ and $v \in C^1([0,\infty))$.

Step 2. We now show that $v \in C^2$.

The only point at which v may not be twice differentiable is $x(i)$ and

$$v''(x,i) = 0, \ \ \forall x > x(i).$$

Set

$$\gamma = \liminf_{x \uparrow x(i)} v''(x,i).$$

Then there exists $x_n < x(i)$ converging to $x(i)$, so that $v''(x_n,i) \to \gamma$. By the first step, $v'(x_n,i) \to 1$. Moreover, the elliptic equation holds at all x_n's. Hence,

$$\begin{aligned} r(i)v(x(i),i) - \mu - \frac{\sigma^2}{2}\gamma - \lambda(i)[v(x(i),i+1) - v(x(i),i)] \\ = \lim_{n\to\infty} r(i)v(x_n,i) - \mathcal{L}v(x_n,i) = 0. \end{aligned} \tag{3.2}$$

The dynamic programming equation (2.6) implies that at any $x > x(i)$,

$$0 \leq r(i)v(x,i) - \mathcal{L}v(x,i) = r(i)v(x,i) - \mu - \lambda(i)[v(x,i+1) - v(x,i)].$$

Hence as $x \downarrow x(i)$

$$r(i)v(x(i),i) - \mu - \lambda(i)[v(x(i),i+1) - v(x(i),i)] \geq 0.$$

The above inequality, together with (3.2) imply that $\gamma \geq 0$. However, by concavity, $v'' \leq 0$. Hence, $\gamma = 0$ and

$$0 \leq \liminf_{x \uparrow x(i)} v''(x,i) \leq \limsup_{x \uparrow x(i)} v''(x,i) \leq 0.$$

Therefore, v is twice differentiable at $x(i)$. □

3.3 Dividend Thresholds

In the previous sections, we have shown that v is a concave, twice continuously differentiable, classical solution of (2.6). By concavity and Lemma 2.1, there are $x(i) > 0$, $i = 0, 1$ such that

$$v'(x,i) = 1 \quad \text{for} \quad x \geq x(i), \quad \text{and} \quad v'(x,i) > 1, \ r(i)v(x,i) - \mathcal{L}v(x,i) = 0, \quad \text{on} \quad [0, x(i)).$$

Indeed,

$$x(i) := inf\{x : \ v'(x,i) = 1 \ \}, \quad \text{and} \quad x_\ell := x(0), \quad x_h := x(1).$$

The following is proved in Appendix A.

Proposition 3.2 *Let $x_\ell, x_h > 0$ be as above. Then, $x_\ell \geq x_h$.*

3.4 Sensitivity Analysis

In this section we give numerical illustrations of the value function and the sensitivities of the dividend thresholds with respect to mean and volatility of the cash flow process and the jump rate between low and high interest rate regimes. The value function is shown in the figure below, for the parameter values

$$\mu = 0.18,\ \sigma = 0.15,\ \lambda = 0.1,\ r_l = 0.02,\ r_h = 0.1,\ x_h = 0.4386,\ x_l = 0.5528. \qquad (3.3)$$

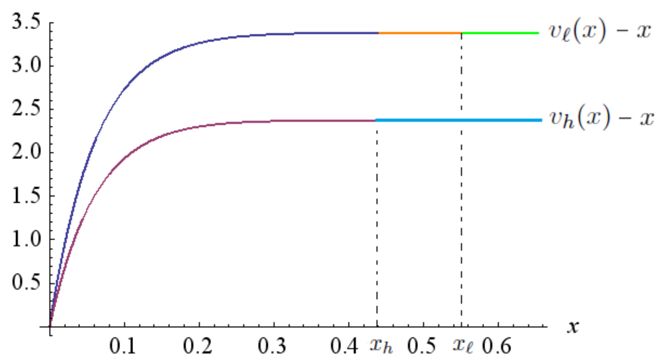

Figure 1: Value function with parameters in (3.3)

$$\mu = 0.18, \quad \lambda = 0.1, \quad r_l = 0.02, \quad r_h = 0.1. \qquad (3.4)$$

$$\sigma = 0.15, \quad \lambda = 0.1, \quad r_l = 0.02, \quad r_h = 0.1. \qquad (3.5)$$

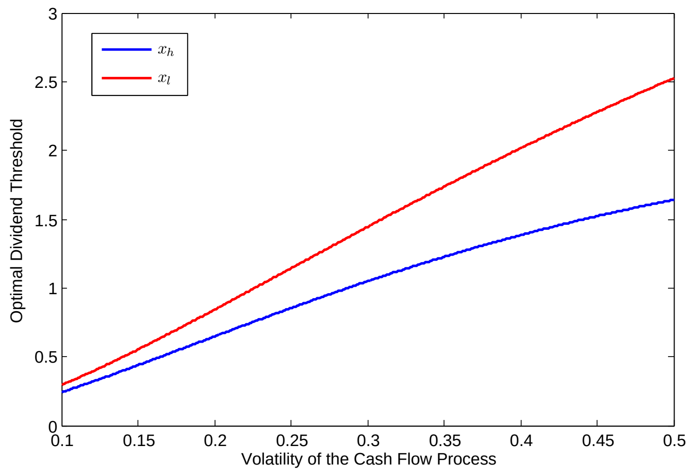

Figure 2: Sensitivities of x_h and x_l wrt σ with parameters in (3.4)

$$\mu = 0.18, \quad \sigma = 0.15, \quad r_l = 0.02, \quad r_h = 0.1. \tag{3.6}$$

4 Issuance

In this section, we enlarge the set of financial policies available to the firm, by allowing it to issue new shares, in addition to distribute dividends. Using the previous notation, the cash level process is now given by

$$X_t = x + \mu t + \sigma B_t - L_t + I_t, \tag{4.1}$$

where I_t is the total amount of cash raised up to time t (cumulated issuance process, net of issuance costs). We assume[2] that I is piecewise constant and has the form

$$I_t = \sum_{k=1}^{\infty} \xi_k \chi_{\{t \geq \tau_k\}}, \tag{4.2}$$

where $0 \leq \tau_1 < \ldots < \tau_k < \tau_{k+1}$ are stopping times at which equity issues are made and $\xi_k \geq 0$ are the issuance sizes. Then, the optimization problem that the firm faces is to maximize[3]

$$J(x, i, L, I) := \mathbb{E}\left[\int_0^\theta \Lambda_t dL_t - \sum_{k=1}^{\infty} \Lambda_{\tau_k} \left(\xi_k + \gamma(i_{\tau_k})\right) \;\Big|\; i_0 = i, X_{0^-} = x\right], \tag{4.3}$$

[2]Given the presence of a fixed issuance cost, such a policy is indeed optimal without loss of generality.
[3]See [3] for a discussion of the objective function.

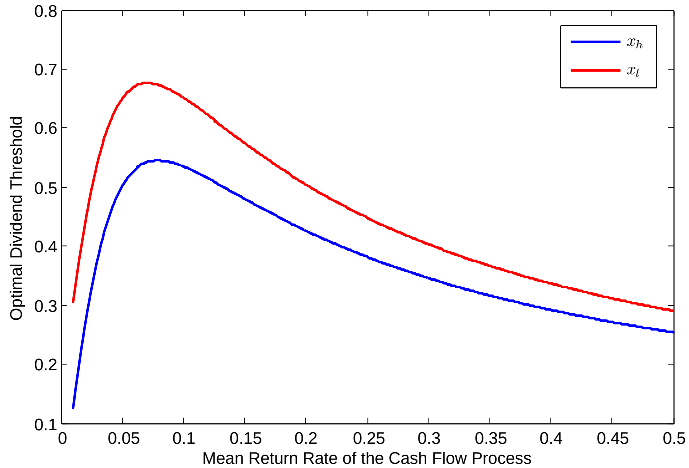

Figure 3: Sensitivities of x_h and x_l wrt μ with parameters in (3.5)

where $\gamma(i) > 0$ is the fixed cost of issuance when the economy is in state i. The interpretation of functional J is straightforward. Since there is a fixed cost $\gamma(i)$ of issuance (which depends on the state i of the economy), new issues will be lumpy and occur at discrete times τ_1, τ_2,... . Since there is no marginal cost of issuance, the total amount of cash raised at date τ_k is just $\xi_k + \gamma(i_{\tau_k})$. Functional J represents expected present value of future dividend payments, net of equity issuances, as in [3].

The value function

$$v(x,i) := \sup_{L,I \in \mathcal{A}(x)} J(x,i,L,I)$$

is the unique viscosity solution of

$$\min \Big\{ r(i)v(x,i) - \mathcal{L}v(x,i) \ ; \ v'(x,i) - 1 \ ; \qquad\qquad (4.4)$$
$$v(x,i) - \sup_{\xi \geq 0} \left(v(x+\xi, i) - \xi - \gamma(i) \right) \Big\} = 0.$$

We distinguish the cases when the cost structure depends on the point process and when not.

4.1 Constant Issuance Cost

The following lemma shows that when $\gamma(i) \equiv \gamma$, it is never optimal to issue new equity before the cash reserves are zero. This is consistent with the results of [3] in the case where interest rates are constant.

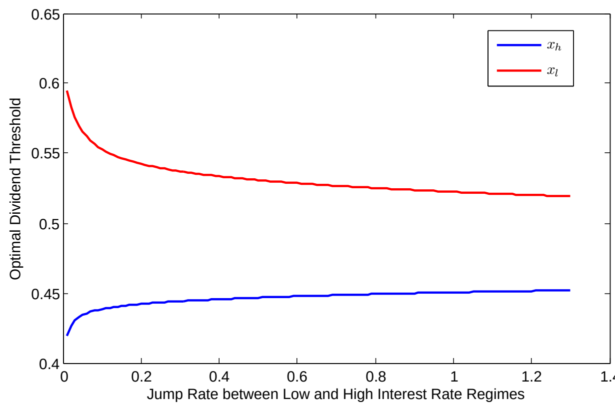

Figure 4: Sensitivities of x_h and x_l wrt μ with parameters in (3.6)

Lemma 4.1 *Suppose γ is independent of i. Then, it is never optimal to issue new equity when the cash level is non zero. Hence, v is the unique solution of*

$$\min\left\{r(i)v(x,i) - \mathcal{L}v(x,i)\ ;\ v'(x,i) - 1\ \right\} = 0,$$

with boundary condition

$$v(0,i) = \max\{0\ ;\ \sup_{\xi \geq 0}\left(v(\xi,i) - \xi - \gamma\right)\}.$$

Moreover for any $x > 0$,

$$v(x,i) > \sup_{\xi \geq 0}\left(v(x+\xi,i) - \xi - \gamma\right).$$

Proof.

Fix $x \geq 0$ and let $(L, I) \in \mathcal{A}(x)$ be any admissible dividend-issuance policy. Then, I is as in (4.2). Suppose that $X_{\tau_1} > 0$. Define $\tilde{I}$ simply by removing the first issuance, i.e.,

$$\tilde{I}_t = \sum_{k=2}^{\infty} \xi_k \chi_{\{t \geq \tau_k\}} = I_t - \xi_1 \chi_{\{t \geq \tau_1\}}.$$

The new strategy $(L, \tilde{I})$ may not be admissible, but the corresponding cash flow process $\tilde{X}$ exists and is given by

$$\tilde{X}_t = x + \mu t + \sigma B_t - L_t + \tilde{I}_t.$$

Set

$$\tau := \inf\{t \geq \tau_1\ :\ \tilde{X}_t \leq 0\},$$

or infinity, if the above set is empty. Since we have assumed that $X_{\tau_1} > 0$, $\tau > \tau_1$.

We now define another issuance strategy $\hat{I}$ by

$$\hat{I}_t = \tilde{I}_t + \xi_1 \chi_{\{t \geq \tau\}}.$$

Then, it is clear that $\hat{I}_t = I_t$ for all $t \geq \tau$. Let $\hat{X}$ be the corresponding cash level process, i.e.,

$$\tilde{X}_t = x + \mu t + \sigma B_t - L_t + \hat{I}_t.$$

Then,

$$\hat{X}_t = \begin{cases} \tilde{X}_t, & \text{for } t \in [0, \tau), \\ X_t, & \text{for } t \geq \tau. \end{cases}$$

The above characterization of $\hat{X}$ shows that $\hat{X}_t \geq 0$ for all $t \geq 0$. Hence, $(L, \hat{I})$ is indeed admissible. Moreover,

$$J(x, i, L, \hat{I}) = J(x, i, L, I) + \mathbb{E}\left[(\Lambda_{\tau_1} - \Lambda_\tau)\, \xi_1\right] > J(x, i, L, I),$$

where the final inequality follows from the fact that $\tau > \tau_1$.

The above argument shows that it is enough to consider issuance strategies for which $X_{\tau_1} = 0$. By induction we can show that this result extends to all issuance times and we need only to consider strategies with $X_{\tau_k} = 0$ for every k. This is exactly the statement of the Lemma. □

4.2 Issuance with Random Costs

If the cost structure γ depends on i, then the above result no longer holds. This is illustrated in the following numerical example where $\gamma(1)$ is much larger than $\gamma(0)$. We use the following parameter values:

$$\mu = 0.18,\ \sigma = 0.5,\ \lambda = 0.1,\ r(0) = 0.02,\ r(1) = 0.1.$$

For this set of parameter values the value function is twice continuously differentiable except one point, x_I, and has the following form. There are thresholds $0 < x_I < x_\ell < x_h$. Set

$$\text{Region 1} := (0, x_I), \quad \text{Region 2} := (x_I, x_\ell), \quad \text{Region 3} := (x_\ell, x_h).$$

In region 1, the firm issues new equity when the interest rate is low (but not when it is high). The two other regions are associated with dividend thresholds x_l and x_h like before. Thus, the value function satisfies

$$\begin{aligned}
&v(x,0) = v(x_\ell, 0) - (x_\ell - x) - \gamma(0), \quad x \in \text{Region 1}, \\
&r(0)v(x,0) = \mathcal{L}v(x,0), \quad x \in \text{Region 2}, \\
&v'(x,0) = 1, \quad x \geq x_\ell, \\
&r(1)v(x,1) = \mathcal{L}v(x,1), \quad x \leq x_h, \\
&v'(x,1) = 1, \quad x \geq x_h.
\end{aligned}$$

Therefore the optimal strategy is given as follows. The fixed cost $\gamma(1)$ is so high that it is never optimal to issue new equity if the state i is equal to one (equivalently, if the interest rate is high). The dividend threshold for $r = r_h$ is x_h and when $r = r_l$ it is x_ℓ. Interestingly, $x_\ell < x_h$ while without issuance the opposite inequality always holds, c.f, Proposition 3.2. For $i = 0$, if the cash level is sufficiently small, i.e., if in Region 1, then the firm issues new equity. In Region 2, the firm does not take any action and pays dividends when $x > x_\ell$. The value function is shown in the figure below, for the parameter values

$$\gamma(0) = 0.48, \quad r(0) = 0.02, \quad r(1) = 0.1, \quad \lambda = 0.1, \quad \sigma = 0.5, \quad \mu = 0.18. \tag{4.5}$$

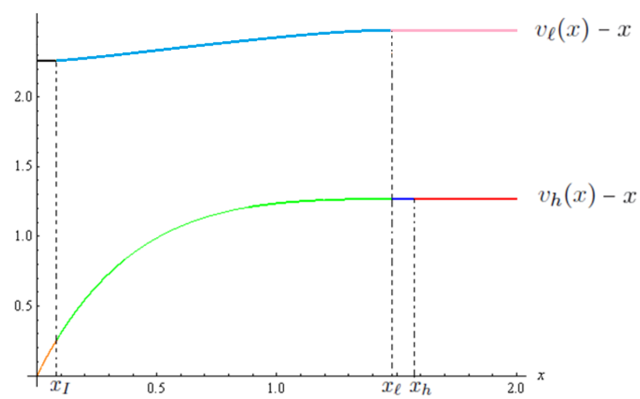

Figure 5: Value function with parameters in (4.5)

4.3 Different Cost but Same Interest Rate

In the example above, the possibility to issue new equity in the good state allows to reverse the ranking of the thresholds. So, even if the opportunity cost of cash is lower ($r_l < r_h$) the firm will issue dividends more often in the good state. In order to understand the impact of issuing costs, we now study this particular case to understand the affect of the cost alone. Indeed, let

$$r(i) = r > 0, \qquad i = 0, 1, \qquad \gamma(0) \leq \gamma(1). \tag{4.6}$$

It is clear that when both $\gamma(0)$ and $\gamma(1)$ are very large, then there will not be any issuance and the problem is the same as the one studied in [3]. In fact, we have an easy quantification of this statement. Let $V(x, r)$ be the Jeanblanc-Picqué & Shiryaev value function defined in (2.3). Let $x^*(r)$ be the dividend payment threshold for this problem and set

$$\gamma^*(r) := V(x^*(r), r) - x^*(r).$$

Lemma 4.2 *Assume* (4.6). *Then, new equity issues are never optimal and* $v(x, i) = V(x, r)$*, if and only if*

$$\gamma(i) \geq \gamma^*(r), \qquad i = 0, 1.$$

Proof. Since V is concave, we directly verify that for every $x, \xi \geq 0$ and $i = 0, 1$,

$$\begin{aligned} V(x+\xi, r) - V(x, r) &\leq V(\xi, r) - V(0, r) = V(\xi, r) \\ &< \xi + \gamma^* \leq \xi + \gamma(i). \end{aligned}$$

Using this it is straightforward to show that the value function $V(x, r)$ solves the dynamic programming equation (4.4). Hence by uniqueness $v = V$. In particular there are never new equity issues.

To prove the converse, assume that there are never new equity issues. Then, $v = V$ where V solves the dynamic programming equation (4.4). In particular,

$$V(x, r) \geq V(x + \xi, r) - \xi - \gamma(i),$$

for all $x, \xi \geq 0$ and $i = 0, 1$. We take $\xi = x^*(r)$ and $x = 0$ to conclude. □

Based on the above result, we computed the value functions for the following parameter values

$$r(0) = r(1) = 0.05, \quad \lambda = 0.3, \quad \sigma = 0.25, \quad \mu = 0.18, \tag{4.7}$$

with two different issuance costs:

$$\gamma(0) = 0.1489 < \gamma^*(r) = 2.60748 << \gamma(1),$$

$$\gamma(0) = 0.7756 < \gamma^*(r) = 2.60748 << \gamma(1).$$

In both cases, we decreased $\gamma(0)$ from γ^*. In all examples, there is issuance as proved in Lemma 4.2. There are three critical thresholds:

$$0 \leq z_0 := \text{issuance threshold},$$

i.e., it is optimal to make an issuance whenever the cash reserves are less than or equal to z_0 and when we are in state $i = 0$. Numerically we observed that of relatively high values of $\gamma(0)$ (i.e, values less than but close to γ^*), $z_0 = 0$. However, $z_0 > 0$ for sufficiently small values of $\gamma(0)$. Hence, there is a balance between the probability of going to a bad state in which issuance is too costly and the probability of recovery.

The other common features of the numerical results is that the dividend payment threshold $x(i)$ is smaller in the "good" state of the economy, i.e., we always find:

$$x(0) < x(1).$$

In other words, dividend payment starts at lower cash reserves when the economy is in a good state.

Below are the tables of these results and two representative graphs. In the first graph $z_0 > 0$ and the black curve is the issuance part. In the second $z_0 = 0$. In both graphs red parts correspond to the dividend payment region.

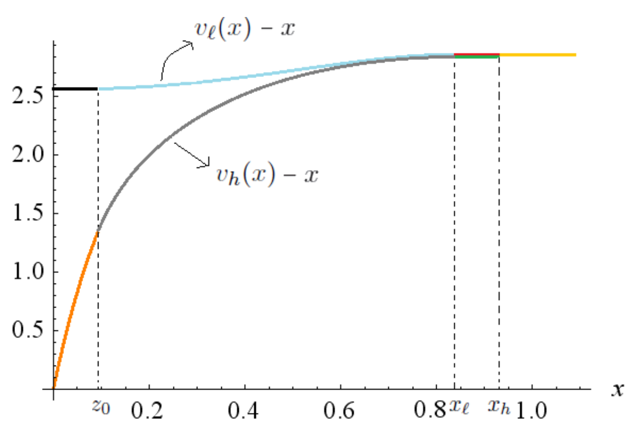

Figure 6: Value function with parameters in (4.7) and $\gamma(0) = 0.1489$

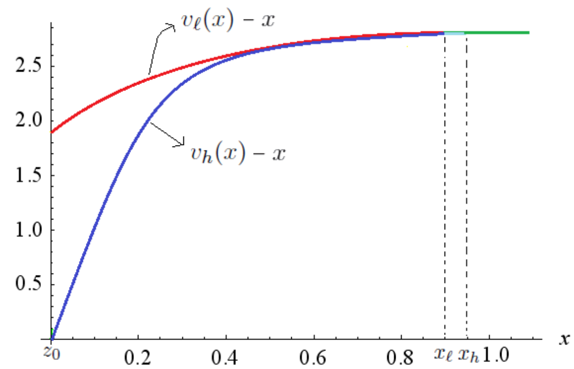

Figure 7: Value function with parameters in (4.7) and $\gamma(0) = 0.7756$

Table 1: Optimal values for the set of parameters $\sigma = 0.25$, $\mu = 0.18$, $r = 0.05$, $\lambda = 0.3$.

$\gamma(0)$	z_0	x_l	x_h
0.0002	0.4990	0.6726	0.9226
0.0033	0.3958	0.7229	0.9229
0.1236	0.1153	0.8327	0.9327
0.1490	0.0954	0.8390	0.9340
0.2691	0.0286	0.8582	0.9382
0.7756	0	0.9003	0.9503
1.0087	0	0.9159	0.9559
1.6265	0	0.9504	0.9704
2.0527	0	0.9702	0.9802

5 Conclusion

This paper has studied the specific impact of macroeconomic variables on the dividend policies of firms by considering the extreme case of a firm whose profitability is constant, but evolves in a stochastic macroeconomic environment, where interest rates and/or issuance costs are governed by an exogenous Markov chain.

Interestingly, we show that these two variables have opposed effects on the dividend policies of firms. Specifically, firms tend to distribute more dividends when interest rates are high and less dividends when issuing costs are high. We also find that stochastic issuing costs allow to get rid of the unfortunate prediction of previous models to which firms wait until the last moment (i.e. until they run out of cash) to issue new equity. Like Bolton, Chen & Wang [1], we obtain a market timing effect: when issuing costs are very high during recessions (so that shareholders refuse to recapitalize firms when they run out of cash) it becomes optimal to issue new equity in the good state even if the firm still has cash reserves, due to the fear that a recession might occur, leading to the forced closure of a profitable company.

References

[1] Bolton, P., Chen, H. and Wang, N. (2011). Market Timing, Investment and Risk Management. Working paper, Columbia University.

[2] Cadenillas, A. and Sotomayor, L.R. (2008). Classical Singular and Impulse Stochastic Control for the Optimal Dividend Policy when there is Regime Switching. Working paper, GSU and University of Alberta.

[3] Décamps, J-P, Mariotti, T., Rochet, J-C., and Villeneuve, S. (2011). Free Cash-Flow, Issuance Costs and Stock Price Volatility, Journal of Finance, 66(5), $1501 - 1544$.

[4] Fleming, W.H., Soner, H.M. (1993). *Controlled Markov Processes and Viscosity Solutions*. Applications of Mathematics 25. Springer-Verlag, New York.

[5] Gertler, M., and Hubbard, R.G. (1993). Corporate Financial Policy, Taxation and Macroeconomic Risk. The RAND Journal of Economics, 24(2), $286 - 303$.

[6] Hackbarth, D., Miao, J. and Morellec, E. (2006). Capital Structure, Credit Risk and Macroeconomic Conditions. Journal of Financial Economics, 82, $519 - 550$.

[7] Karatzas, I., and S.E. Shreve (1991). *Brownian Motion and Stochastic Calculus, New York*, Springer.

[8] Jeanblanc-Picqué, M., and A.N. Shiryaev (1995). Optimization of the Flow of Dividends, Russian Mathematical Surveys, 50(2), $257 - 277$.

[9] Jiang, Z., Pistorius, M. (2012). Optimal Dividend Distribution under Markov Regime Switching. Finance and Stochastics, 16(3), $449 - 476$.

[10] Radner, R. and Shepp, L. (1996). Risk vs Profit Potential: A Model for Corporate Strategy. Journal of Economic Dynamics and Control, 20, $1373 - 1393$.

[11] Rockefellar, R.T. 1970. *Convex Analysis*, Princeton University Press.

Appendix A

In this Appendix, we prove the concavity of the value function. Firstly, in view of Lemma 3.1 and Corollary 3.1, there are constants $c_1, c_2 > 0$ such that

$$v(x,i) \geq x + c_1 \quad \forall\, x \geq x^*/2,\ i \in \{0,1\} \tag{5.1}$$

$$v(x,i) \leq V(x, r_\ell) \leq x + c_2 \quad \forall\, x \geq 0,\ i \in \{0,1\}. \tag{5.2}$$

The following technical result is needed in the proof of concavity. Let x^* be as in the previous result. Also recall that $\theta^{x,L}$ is the exit time of $X^{x,L}$ from the interval $(0,\infty)$.

Lemma. There are $\hat{T} \geq 1$ and $\hat{\Lambda} < 1$ such that

$$\mathbb{E}[\Lambda_{\hat{T}\wedge\theta^{x,L}}] \leq \hat{\Lambda},$$

for all $x \geq x^*/2$, $L \in \mathcal{A}(x)$ satisfying

$$J(x,i;L) \geq x + \frac{c_1}{2},$$

where c_1 is as in (5.1).

Proof. Fix x and L as in the statement and set $X = X^{x,L}$. For $T > 0$ to be determined, set $\theta = \theta^{x,L}$ and $\tau := \theta \wedge T$. By dynamic programming,

$$J(x,i,L) \leq \mathbb{E}\left[\int_0^\tau \Lambda_t dL_t + \Lambda_\tau v(X_\tau, i_\tau)\right].$$

Set $\tilde{X}_t = x + \mu t + \sigma W_t$, so that $X_t = \tilde{X}_t - L_t$. Since $\Lambda_t \leq 1$, (5.2) implies

$$\begin{aligned} J(x,i,L) &\leq \mathbb{E}\left[\int_0^\tau dL_t + \chi_{\{\theta\geq T\}}(\tilde{X}_T - L_T + c_2)e^{-r_\ell T}\right] \\ &= \mathbb{E}\left[L_\tau\left(1 - \chi_{\{\theta\geq T\}}e^{-r_\ell T}\right) + \chi_{\{\theta\geq T\}}(\tilde{X}_T + c_2)e^{-r_\ell T}\right]. \end{aligned}$$

On $\{\theta < T\}$, $L_\theta = \tilde{X}_\theta$ and on $\{\theta \geq T\}$, we have $\tau = T$ and $L_T = \tilde{X}_T - X_T$. Then, since $J(x,i;L) \geq x + c_1/2$,

$$\begin{aligned} x + \frac{1}{2}c_1 &\leq J(x,i;L) \\ &\leq \mathbb{E}\left[\tilde{X}_\theta\ \chi_{\{\theta<T\}} + \Big(\tilde{X}_T - X_T + e^{-r_\ell T}(X_T + c_2)\Big)\chi_{\{\theta\geq T\}}\right] \\ &= \mathbb{E}\left[\tilde{X}_\tau + \left(-X_T + e^{-r_\ell T}(X_T + c_2)\right)\chi_{\{\theta\geq T\}}\right] \\ &= \mathbb{E}\left[\tilde{X}_\tau + \left(e^{-r_\ell T}c_2 - X_T(1 - e^{-r_\ell T})\right)\chi_{\{\theta\geq T\}}\right] \\ &\leq \mathbb{E}\left[\tilde{X}_\tau + e^{-r_\ell T}c_2\ \chi_{\{\theta\geq T\}}\right] \leq (x + \mu\mathbb{E}[\tau]) + e^{-r_\ell T}c_2. \end{aligned}$$

We now set $T = \hat{T}$ where $\hat{T}$ is so that $e^{-r_\ell \hat{T}} c_2 = \frac{c_1}{4}$. Then,

$$x + \frac{c_1}{2} \leq x + \mu \mathbb{E}(\tau) + \frac{c_1}{4}.$$

Hence,

$$\mathbb{E}[\theta^{x,L} \wedge \hat{T}] = \mathbb{E}[\tau] \geq \frac{c_1}{4\mu}.$$

Set $f(t) = e^{-r_\ell t}$ so that $\Lambda_t \leq f(t)$. Since f is convex and $f(0) = 1$,

$$\begin{aligned}
\mathbb{E}[\Lambda_\tau] &\leq \mathbb{E}[f(\tau)] \leq \mathbb{E}\left[\frac{\tau}{\hat{T}} f(\hat{T}) + (1 - \frac{\tau}{\hat{T}}) f(0)\right] \\
&= \frac{f(\hat{T})}{\hat{T}} \mathbb{E}[\tau] + \left(1 - \frac{1}{\hat{T}} \mathbb{E}[\tau]\right) \\
&= 1 - \frac{1}{\hat{T}} (1 - f(\hat{T})) \mathbb{E}[\tau] \\
&\leq 1 - \frac{1}{\hat{T}} (1 - f(\hat{T})) \frac{c_1}{4\mu} =: \hat{\Lambda}.
\end{aligned}$$

□

We are now ready to prove the concavity of the value function.

Proof of Theorem 3.1. For $x, y \geq 0$, $i \in \{0, 1\}$, set

$$I(x, y, i) := v(x, i) + v(y, i) - 2v\left(\frac{x + y}{2}, i\right).$$

In view of Corollary 3.1, $I(x, y, i) \leq 0$, for all $x, y \in [0, x^*]$. Set

$$\hat{\alpha} := \sup\left\{I(x, y, i) : \ \frac{x^*}{2} \leq x \leq y,\ i = 0, 1\right\}.$$

By Lemma 2.2, $\hat{\alpha} < \infty$. Hence, for every $\varepsilon > 0$ there are $x_\varepsilon, y_\varepsilon$, $i_\varepsilon \in \{0, 1\}$ such that

$$\hat{\alpha} \leq I(x_\varepsilon, y_\varepsilon, i_\varepsilon) + \varepsilon, \quad \text{and} \quad \frac{x^*}{2} \leq x_\varepsilon \leq y_\varepsilon.$$

In view of Lemma 3.1, to prove the concavity of v, it suffices to show that $\hat{\alpha} \leq 0$.

Let $L^x \in \mathcal{A}(x_\varepsilon)$, $L^y \in \mathcal{A}(y_\varepsilon)$ be arbitrary dividend strategies satisfying

$$J(x_\varepsilon, i; L^x) \geq x_\varepsilon + \frac{c_1}{2}, \qquad J(y_\varepsilon, i; L^y) \geq y_\varepsilon + \frac{c_1}{2}. \tag{5.3}$$

In view of (5.1), such processes exist, and

$$v(x_\varepsilon, i) = \sup\{J(x_\varepsilon, i; L^x) \mid L^x \in \mathcal{A}(x_\varepsilon) \text{ and } L^x \text{ satisfies (5.3) }\}.$$

The same also holds at y_ε. Set

$$\bar{L} := \frac{L^x + L^y}{2}, \qquad \bar{x} := \frac{x_\varepsilon + y_\varepsilon}{2}.$$

Finally, let $\hat{T}$ be as in the Lemma above. Set $\theta^x := \theta^{x_\epsilon, L^x}$. Without loss of generality assume that

$$X_t^\varepsilon := x_\varepsilon + \mu t + \sigma W_t - L_t^x \le Y_t^\varepsilon := y_\varepsilon + \mu t + \sigma W_t - L_t^y, \;\; \forall\, t \le \theta^x.$$

Otherwise, one may simply redefine L^x and L^y so that $X_t^\varepsilon = Y_t^\varepsilon$ after the first time they are equal.

Set $\tau := \theta^x \wedge \hat{T}$. By the dynamic programming principle (2.7),

$$\begin{aligned} J(x_\varepsilon, i; L^x) + J(y_\varepsilon, i; L^y) &\le 2\mathbb{E}\left[\int_0^\tau \Lambda_t d\bar{L}_t\right] + \mathbb{E}\left[\Lambda_\tau (v(X_\tau^\varepsilon, i_\tau) + v(Y_\tau^\varepsilon, i_\tau))\right] \\ &= 2\mathbb{E}\left[\int_0^\tau \Lambda_t d\bar{L}_t + \Lambda_\tau v(X_\tau^{\bar{x},\bar{L}}, i_\tau)\right] \\ &\quad + \mathbb{E}\left(\Lambda_\tau \left[v(X_\tau^\varepsilon, i_\tau) + v(Y_\tau^\varepsilon, i_\tau) - 2v(X_\tau^{\bar{x},\bar{L}}, i_\tau)\right]\right) \\ &\le 2v(\bar{x}, i) + \mathbb{E}[\Lambda_\tau]\hat{\alpha}. \end{aligned}$$

By the Lemma above, $\mathbb{E}[\Lambda_\tau] \le \hat{\Lambda} < 1$. Also,

$$v(x_\varepsilon, i_\varepsilon) + v(y_\varepsilon, i_\varepsilon) = \sup\{J(x_\varepsilon, i_\varepsilon; L^x) + J(y_\varepsilon, i_\varepsilon; L^y) \mid (L^x, L^y) \text{ satisfying (5.3)}\}.$$

Hence,

$$v(x_\varepsilon, i_\varepsilon) + v(y_\varepsilon, i_\varepsilon) \le 2v(\bar{x}, i_\varepsilon) + \hat{\Lambda}\hat{\alpha}.$$

By the choice of $(x_\varepsilon, y_\varepsilon)$,

$$\hat{\alpha} \le v(x_\varepsilon, i_\varepsilon) + v(y_\varepsilon, i_\varepsilon) - 2v(\bar{x}, i_\varepsilon) + \varepsilon \le \hat{\Lambda}\hat{\alpha} + \varepsilon.$$

Hence $\hat{\alpha} \le \varepsilon/(1 - \hat{\Lambda})$, for all $\varepsilon > 0$. Therefore, $\hat{\alpha} \le 0$ and consequently v is concave.

□

Appendix B

Proof of Proposition 3.2. Towards a contradiction, suppose that $x_\ell < x_h$. Set

$$u(x) := v_\ell'(x), \quad w(x) := v_h'(x), \qquad \lambda_\ell := \lambda(0), \quad \lambda_h := \lambda(1).$$

Differentiating the original system once and using the above definitions yield the following coupled ordinary differential equations for u and w, on the interval $(0, x_\ell)$,

$$r_h w(x) \;=\; \mu w'(x) + (1/2)\sigma^2 w''(x) - \lambda_h[w(x) - u(x)], \tag{5.4}$$

$$r_\ell u(x) \;=\; \mu u'(x) + (1/2)\sigma^2 u''(x) + \lambda_\ell[w(x) - u(x)]. \tag{5.5}$$

Since $v_\ell(0) = v_h(0) = 0$ and $v_\ell(x) \geq v_h(x)$ for all $x \in [0, \infty)$, we conclude that $u(0) \geq w(0)$.

Our goal is to show that $u(x) \geq w(x)$ for all $x \in [0, x_\ell]$. Indeed, by our hypothesis $x_\ell < x_h$, $w(x_\ell) > w(x_h) = 1$. So if we can prove that $u \geq w$ on $[0, x_\ell]$, then

$$1 = u(x_\ell) \geq w(x_\ell) > 1$$

will provide the desired contradiction. Hence it suffices to prove that $u \geq w$ on $[0, x_\ell]$.

Set $\Phi(x) = (u - w)(x)$ and choose $y \in [0, x_\ell]$ so that

$$(u - w)(y) = \min_{x \in [0, x_\ell]} (u - w)(x) =: \alpha. \tag{5.6}$$

Our goal is to show that $\alpha \geq 0$. We analyze three cases separately.

Case 1: $y = 0$. In this case, $\alpha = u(0) - w(0) = 0$.

Case 2: $y \in (0, A)$. Since y is a local minimum of Φ,

$$\Phi'(y) = u'(y) - w'(y) = 0, \qquad \Phi''(y) = u''(y) - w''(y) \geq 0.$$

We use these first in (5.4) and then in (5.5) at the point y. The result is the following,

$$\begin{aligned} r_\ell u(y) \;&=\; \mu u'(y) + \frac{1}{2}\sigma^2 u''(y) - \lambda_\ell \alpha \geq \mu w'(y) + \frac{1}{2}\sigma^2 w''(y) - \lambda_\ell \alpha \\ &=\; r_h w(y) - [\lambda_h + \lambda_\ell]\alpha \geq r_\ell w(y) - [\lambda_h + \lambda_\ell]\alpha. \end{aligned}$$

In the a last step we used the fact that $w \geq 0$. Since $\alpha = u(y) - w(y)$, the above implies that $\alpha \geq 0$.

Case 3: $y = A$. By the smooth fit, we know that $v''(x_\ell) = u'(x_\ell) = 0$. We directly conclude that

$$\Phi'(x_\ell) = u'(x_\ell) - w'(x_\ell) = v_\ell''(x_\ell) - v_h''(x_\ell) = -v_h''(x_\ell) \geq 0.$$

Since $y = x_\ell$ is the minimum of Φ on the interval $[0, x_\ell]$, $\Phi'(x_\ell) \leq 0$. Hence, $\Phi''(x_\ell) = -v_h'(x_\ell) = 0$.

Recall that we have assumed that $x_h > x_\ell$. Set $f(x) := v_h''(x)$ and differentiate the dynamic programming equation (2.6) for v_h twice. The result is,

$$r_h f(x) = \mu f'(x) + \frac{1}{2}\sigma^2 f''(x) - \lambda_h f(x), \quad x \in (x_\ell, x_h),$$

together with boundary conditions $f(x_\ell) = f(x_h) = 0$. However, the zero function is the unique solution of this equation. Hence, $f(x) = v_h''(x) = 0$ for $x \in [x_\ell, x_h]$. So, v_h' is constant on $[x_\ell, x_h]$ as well. Since $v_h'(x_\ell) > 1$, we conclude that $x_h = \infty$. But this implies that $v_h(x) > v_\ell(x)$ for all sufficiently large x.

Hence, $x_\ell \geq x_h$.

□