Optimal Investment and Consumption with Proportional Transaction Costs in Regime-Switching Model

Ruihua Liu

Journal of Optimization Theory and Applications

ISSN 0022-3239
Volume 163
Number 2

J Optim Theory Appl (2014) 163:614-641
DOI 10.1007/s10957-013-0445-y

Springer

J Optim Theory Appl (2014) 163:614–641
DOI 10.1007/s10957-013-0445-y

Optimal Investment and Consumption with Proportional Transaction Costs in Regime-Switching Model

Ruihua Liu

Received: 21 January 2013 / Accepted: 1 October 2013 / Published online: 16 October 2013

Abstract This paper is concerned with an infinite-horizon problem of optimal investment and consumption with proportional transaction costs in continuous-time regime-switching models. An investor distributes his/her wealth between a stock and a bond and consumes at a non-negative rate from the bond account. The market parameters (the interest rate, the appreciation rate, and the volatility rate of the stock) are assumed to depend on a continuous-time Markov chain with a finite number of states (also known as regimes). The objective of the optimization problem is to maximize the expected discounted total utility of consumption. We first show that for a class of hyperbolic absolute risk aversion utility functions, the value function is a viscosity solution of the Hamilton–Jacobi–Bellman equation associated with the optimization problem. We then treat a power utility function and generalize the existing results to the regime-switching case.

Keywords Optimal investment and consumption problem · Transaction cost · Regime-switching model · Hamilton-Jacobi-Bellman equation · Power utility

1 Introduction

The continuous-time optimal investment and consumption model with proportional transaction costs was first introduced in [1]. A rigorous study of an infinite-horizon problem for a power utility function was provided in [2] which partitioned the solvency region into three non-overlapped cones, corresponding to the buy, sell, and no-transaction regions, respectively; the optimal policy is to keep the system state

Communicated by Vladimir Veliov.

R. Liu (✉)
Department of Mathematics, University of Dayton, 300 College Park, Dayton, OH 45469, USA
e-mail: rliu01@udayton.edu

within the no-transaction region for all the time. In [3] Shreve and Soner presented a thorough analysis of the same model of [2]. They successfully removed some assumptions in [2] that are too restrictive and not fully verifiable. By using the method of viscosity solution to the Hamilton–Jacobi–Bellman (HJB) equation, [3] established the regularity of the value function of the optimization problem and characterized its behavior within and across the boundaries of the three regions for buy, sell, and no-transaction.

In this paper we study an infinite-horizon problem of optimal investment and consumption with proportional transaction costs in continuous-time regime-switching models. The problem formulation is almost the same as that in [3] except that a Markov regime-switching model is assumed for the underlying asset prices. We adopt the methodology in [3] and extend many results of [3] to the regime-switching case. The introduction of regime switching brings substantial differences into the model. First, aiming to include the impact of macroeconomic conditions on the individual asset price, the regime-switching models are capable of capturing the dynamical change of the asset prices across different stages of business cycles. This is particularly necessary when problems involving a long time horizon are considered, as in the case of investment and consumption problems. Second, the associated HJB equation to the optimal control problem is a system of coupled variational inequalities (in contrast with a single one without regime switching as in [3]) for which the analysis is much more challenging due to regime coupling. Third, we show that the three regions (buy, sell and no-transaction) become regime dependent, that is, they can be different for different states of the Markov chain and, when a regime switching occurs, those three regions may also change. This is a fundamental difference of the characterization of the optimal policies resulted from regime switching.

The paper is organized as follows. Section 2 presents the problem formulation and the associated HJB system. Section 3 presents a number of fundamental properties of the value function. Section 4 establishes the viscosity solution property of the value function to the HJB system. Section 5 treats a power utility function and derives both upper and lower bounds for the value function that are necessary for the characterization of the optimal trading regions. Section 6 combines the viscosity solution property and the convexity property of the value function to characterize the (regime-dependent) buy, sell, and no-transaction regions associated with the optimal trading policy. Section 7 provides further remarks and concludes the paper. Additionally, an Appendix is provided that contains two technical proofs and the results from a convex analysis of the value function.

2 Problem Formulation

Given a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ equipped with a filtration $\mathbb{F} = \{\mathcal{F}_t, t \geq 0\}$ satisfying the usual conditions of right continuity and completeness. Let $W(t)$ be a standard Brownian motion. Let α_t be a continuous-time Markov chain valued in a finite set $\mathcal{M} := \{1, \ldots, m_0\}$ with $m_0 > 0$ fixed. Let $Q = (q_{ij})_{m_0 \times m_0}$ be the intensity matrix (or the generator) of α_t. Q has the properties: $q_{ij} \geq 0$ if $i \neq j$, $q_{ii} \leq 0$ and $q_{ii} = -\sum_{j \neq i} q_{ij}$ for each $i = 1, \ldots, m_0$. We assume that α_t is observable and is independent of the Brownian motion $W(t)$.

We consider a market with one bond and one stock. The bond price follows

$$dB(t) = r(\alpha_t)B(t)\,dt, \tag{1}$$

and the stock price follows

$$dS(t) = S(t)\big[\eta(\alpha_t)\,dt + \sigma(\alpha_t)\,dW(t)\big], \tag{2}$$

where $r : \mathcal{M} \to \mathbb{R}$ denotes the interest rate, $\eta : \mathcal{M} \to \mathbb{R}$ and $\sigma : \mathcal{M} \to \mathbb{R}$ denote the appreciation rate and the volatility rate of the stock, respectively.

The following notations will be used throughout the paper. For any function $a : \mathcal{M} \to \mathbb{R}$, we let $a^* = \max_{i\in\mathcal{M}} a(i)$ and $a_* = \min_{i\in\mathcal{M}} a(i)$.

We assume that $r(i) > 0$ and $\sigma(i) > 0$ for each $i \in \mathcal{M}$. Moreover, we assume that $\eta^* > r_*$, implying that the best return from the risky asset beats the worst interest rate over the entire time horizon. Note that we do not require $\eta(i) > r(i)$ for every $i \in \mathcal{M}$. Recall that the literature dealing with models with constant interest rate r and appreciation rate η usually assumes $\eta > r$ (see for example [3]). The main motivation for introducing regime switching is to have a model capable of describing the asset price under different market conditions. It has been witnessed that during periods of recessions, the returns from stocks often are lower than the interest rate. The assumption $\eta^* > r_*$ certainly takes this consideration into account.

An investor distributes his/her wealth between the bond and the stock and consumes from the bond account. The optimization problem we will study is to maximize the total utility from consumption over an infinite time horizon.

Let $x(t)$ and $y(t)$ denote the amounts invested at time t in the bond and stock, respectively. Then the system state $(x(t), y(t))$ evolve according to the equations

$$\begin{cases} dx(t) = \big[r(\alpha_t)x(t) - C(t)\big]\,dt - (1+\lambda)\,dM(t) + (1-\mu)\,dN(t), \\ dy(t) = y(t)\big[\eta(\alpha_t)\,dt + \sigma(\alpha_t)\,dW(t)\big] + dM(t) - dN(t), \\ x(0-) = x, \qquad y(0-) = y, \qquad \alpha_0 = i, \end{cases} \tag{3}$$

where $C(t)$ denotes the consumption rate at time t, $M(t)$ and $N(t)$ denote the cumulative purchases and sales of the stock, respectively, by time t. The triple $(C, M, N) = \{(C(t), M(t), N(t)), t \geq 0\}$ is known as a control policy. The constants λ and μ stand for the proportional transaction costs associated with, respectively the purchases and sales of the stock. We assume that $0 \leq \mu < 1$ and $0 \leq \lambda < 1$.

The solvency region $\mathit{\Pi}$ is the open subset of $\mathbb{R}^2$ defined by

$$\mathit{\Pi} = \big\{(x, y) \in \mathbb{R}^2 : x + (1+\lambda)y > 0, x + (1-\mu)y > 0\big\}. \tag{4}$$

Moreover, let

$$\begin{aligned} \partial_1\mathit{\Pi} &= \big\{(x, y) \in \mathbb{R}^2 : y \geq 0, x + (1-\mu)y = 0\big\}, \\ \partial_2\mathit{\Pi} &= \big\{(x, y) \in \mathbb{R}^2 : y \leq 0, x + (1+\lambda)y = 0\big\}. \end{aligned} \tag{5}$$

Then we have $\partial\mathit{\Pi} = \partial_1\mathit{\Pi} \cup \partial_2\mathit{\Pi}$.

Let $U:[0,\infty[\to\mathbb{R}$ be the utility function. We assume that U has the following properties: $U(0)=0$, $U'(c)>0$, $U''(c)<0$ for $c>0$, $\lim_{c\to 0^+}U'(c)=\infty$ and $\lim_{c\to\infty}U'(c)=0$. Moreover, there exist constants $K>0$ and $0<\gamma<1$ such that

$$U(c)\le Kc^{\gamma},\quad \forall c\ge 0. \tag{6}$$

We now introduce the notion of admissible control.

Definition 2.1 Given $(x,y)\in\overline{\Pi}$ and $i\in\mathcal{M}$, a control policy (C,M,N) is called admissible with respect to the initial state (x,y,i) iff the following conditions hold:

1. C,M,N are $\mathbb{F}$-adapted processes.
2. The process C satisfies

$$C(t)\ge 0,\quad \text{a.s.},\quad \int_0^t C(s)\,ds<\infty,\quad \text{a.s.}, \forall t\ge 0. \tag{7}$$

3. The processes M and N are right continuous with left limits, non-negative and nondecreasing.
4. $(x(t),y(t))\in\overline{\Pi}$, $\forall t\ge 0$, where $(x(t),y(t))$ is the solution of (3) when the control (C,M,N) is being used.

Let $\mathcal{A}(x,y,i)$ denote the collection of admissible controls with the initial state (x,y,i).

The following assumption is made in this paper.

Assumption 2.1 For all $(x,y)\in\overline{\Pi}$ and $i\in\mathcal{M}$,

$$\sup_{(C,M,N)\in\mathcal{A}(x,y,i)} E\left[\int_0^\infty e^{-\beta t}U\big(C(t)\big)\,dt\right]<\infty, \tag{8}$$

where $\beta>0$ is a discount factor.

Remark 2.1 It will be interesting to find conditions on the market parameters and the discount factor β so that (8) holds, particularly for the power utility function. We leave it for future research and use Assumption 2.1 as a standing assumption in the current setting. See [3] for comments on the single-regime model.

Given $(x,y,i)\in\overline{\Pi}\times\mathcal{M}$, and an admissible control $(C,M,N)\in\mathcal{A}(x,y,i)$, we define the total expected discounted utility J from consumption by

$$J(x,y,i,C,M,N)=E\left[\int_0^\infty e^{-\beta t}U\big(C(t)\big)\,dt\right]. \tag{9}$$

The value function V is defined by

$$V(x,y,i)=\sup_{(C,M,N)\in\mathcal{A}(x,y,i)} J(x,y,i,C,M,N). \tag{10}$$

Remark 2.2 We note that if $(x, y) \in \partial\Pi$, then for any $i \in \mathcal{M}$, the only admissible control policy is the one that jumps immediately to $(0, 0)$ and remains there and has $C \equiv 0$. Consequently $V(x, y, i) = 0$. See [3, Remark 2.1] for a mathematical verification that can be easily generalized to the regime-switching case.

For the sake of notational simplicity, in what follows we will use $a_i = a(i)$, $i \in \mathcal{M}$ for any function $a : \mathcal{M} \to \mathbb{R}$, and $w_i(x, y) = w(x, y, i)$ for any function $w : \overline{\Pi} \times \mathcal{M} \to \mathbb{R}$.

For each $i \in \mathcal{M}$ and function $\varphi \in C^2(\Pi)$, define an operator $\mathcal{L}_i$ by

$$\mathcal{L}_i\varphi(x, y) = \beta\varphi(x, y) - \frac{1}{2}\sigma_i^2 y^2\varphi_{yy}(x, y) - \eta_i y\varphi_y(x, y) - r_i x\varphi_x(x, y). \tag{11}$$

The HJB equation associated with the optimal control problem is a system of m_0 coupled variational inequalities given by

$$\begin{aligned}\min\Big\{&\mathcal{L}_i w(x, y, i) - Qw(x, y, \cdot)(i) - \widetilde{U}\big(w_x(x, y, i)\big),\\ &(1+\lambda)w_x(x, y, i) - w_y(x, y, i),\\ &-(1-\mu)w_x(x, y, i) + w_y(x, y, i)\Big\} = 0, \quad (x, y) \in \Pi, i \in \mathcal{M},\end{aligned} \tag{12}$$

where

$$Qw(x, y, \cdot)(i) = \sum_{j\neq i} q_{ij}\big[w(x, y, j) - w(x, y, i)\big] = \sum_{j=1}^{m_0} q_{ij} w(x, y, j) \tag{13}$$

and

$$\widetilde{U}(x) = \sup_{c\geq 0}\big\{U(c) - cx\big\}. \tag{14}$$

3 Properties of the Value Function

In this section we present fundamental properties for the value function V.

Proposition 3.1 *The value function V is jointly concave in x and y for each $i \in \mathcal{M}$.*

This can be easily shown by using the concavity of the utility function U and the linearity of the state equation (3).

An immediate consequence of the concavity is the following continuity property.

Corollary 3.1 *The value function V is continuous in Π for each $i \in \mathcal{M}$.*

We then establish the continuity of the value function V on $\overline{\Pi}$.

Theorem 3.1 *The value function V is continuous on $\overline{\Pi}$ for each $i \in \mathcal{M}$.*

The proof uses the upper bound property established in Proposition 5.3 for the power utility function and is postponed to the Appendix.

Proposition 3.2 *The value function V is strictly increasing in x and increasing in y for each $i \in \mathcal{M}$.*

Proof First, note that $\mathcal{A}(x_1, y, i) \subset \mathcal{A}(x_2, y, i)$ if $x_1 \le x_2$. It then follows that $V(x_1, y, i) \le V(x_2, y, i)$, i.e., V is increasing in x for each $i \in \mathcal{M}$. Similarly we can show that V is increasing in y for each $i \in \mathcal{M}$.

Now suppose there exist an i_0 and a y_0 such that $V(x, y_0, i_0)$ is not strictly increasing in x. Consequently, there exist $x_1 < x_2$ such that $V(x_1, y_0, i_0) = V(x_2, y_0, i_0)$. It follows that $V(x, y_0, i_0) = V(x_1, y_0, i_0)$ for all $x \in [x_1, x_2]$ since $V(x, y_0, i_0)$ is increasing in x. Note that $V(x, y_0, i_0)$ is concave in x, hence the interval $[x_1, x_2]$ cannot be finite. Therefore there exists $x_0 > 0$ such that

$$V(x, y_0, i_0) = V(x_0, y_0, i_0), \quad \forall x \in [x_0, \infty[. \tag{15}$$

Consider the triple (x_0, y_0, i_0) and let $(C^\varepsilon, M^\varepsilon, N^\varepsilon) \in \mathcal{A}(x_0, y_0, i_0)$ be an ε-optimal control for given $\varepsilon > 0$. Then we have

$$V(x_0, y_0, i_0) \le E\left[\int_0^\infty e^{-\beta t} U\big(C^\varepsilon(t)\big)\, dt\right] + \varepsilon. \tag{16}$$

Depending on the sign of y_0, we consider two cases.

Case 1: $y_0 \ge 0$. Pick x_1 such that

$$x_1 > \max\left(x_0, U^{-1}\left[\beta\left(E\left[\int_0^\infty e^{-\beta t} U\big(C^\varepsilon(t)\big)\, dt\right] + \varepsilon\right)\right] \Big/ r_*\right). \tag{17}$$

We claim that the policy $(r_* x_1, 0, 0) \in \mathcal{A}(x_1, y_0, i_0)$. To prove this, applying the policy $(r_* x_1, 0, 0)$ to the system equation (3), we have

$$\begin{cases} dx(t) = \big[r(\alpha_t)x(t) - r_* x_1\big]\, dt, \\ dy(t) = y(t)\big[\eta(\alpha_t)\, dt + \sigma(\alpha_t)\, dW(t)\big], \quad x(0) = x_1, y(0) = y_0, \alpha_0 = i_0. \end{cases}$$

It is readily seen that $x(t) \ge x_1$ and $y(t) \ge 0$ for all $t \ge 0$. Hence $(x(t), y(t)) \in \overline{\Pi}$ for all $t \ge 0$. We obtain from (16) and (17) that

$$V(x_0, y_0, i_0) < \frac{1}{\beta} U(r_* x_1) = \int_0^\infty e^{-\beta t} U(r_* x_1)\, dt \le V(x_1, y_0, i_0),$$

which is a contradiction to (15).

Case 2: $y_0 < 0$. Pick x_1 such that

$$x_1 + (1+\lambda) y_0 > \max\left(x_0, U^{-1}\left[\beta\left(E\left[\int_0^\infty e^{-\beta t} U\big(C^\varepsilon(t)\big)\, dt\right] + \varepsilon\right)\right] \Big/ r_*\right). \tag{18}$$

In this case, the policy $(r_*(x_1+(1+\lambda)y_0), -y_0, 0) \in \mathcal{A}(x_1, y_0, i_0)$. Indeed, applying this policy, we have

$$\begin{cases} dx(t) = \big[r(\alpha_t)x(t) - r_*\big(x_1+(1+\lambda)y_0\big)\big]\,dt, \\ dy(t) = y(t)\big[\eta(\alpha_t)\,dt + \sigma(\alpha_t)\,dW(t)\big], \quad x(0) = x_1+(1+\lambda)y_0, y(0)=0, \alpha_0 = i_0. \end{cases}$$

Then $x(t) \geq x_1+(1+\lambda)y_0 > 0$ and $y(t)=0$ for all $t \geq 0$. Hence $(x(t), y(t)) \in \overline{\Pi}$ for all $t \geq 0$. It follows from (16) and (18) that

$$\begin{aligned} V(x_0, y_0, i_0) &< \frac{1}{\beta} U\big(r_*\big(x_1+(1+\lambda)y_0\big)\big) \\ &= \int_0^\infty e^{-\beta t} U\big(r_*\big(x_1+(1+\lambda)y_0\big)\big)\,dt \leq V(x_1, y_0, i_0), \end{aligned}$$

which again is a contradiction to (15). □

Remark 3.1 The proof of Proposition 3.2 for V being strictly increasing in x generalizes [4, Proposition 2.1] by considering both the cases of $y_0 \geq 0$ and $y_0 < 0$. Since [4] restricts the solvency region to the first quadrat of the xy-plane, the state x and y are not allowed to be negative. Thus [4, Proposition 2.1] does not need to consider the $y_0 < 0$ case while we do so in the proof of Proposition 3.2.

Given $i \in \mathcal{M}$ and any two points $(x_1, y_1), (x_2, y_2) \in \overline{\Pi}$. If there exist $m, n \geq 0$ such that $x_2 = x_1 - (1+\lambda)m + (1-\mu)n$, $y_2 = y_1 + m - n$, then we say that (x_2, y_2) can be reached from (x_1, y_1) by a transaction. In this case, for each policy $(C, M, N) \in \mathcal{A}(x_2, y_2, i)$, we can construct a policy in $\mathcal{A}(x_1, y_1, i)$ by simply adding this transaction to (C, M, N) at time $t=0$. This results in the following proposition.

Proposition 3.3 *If $(x_2, y_2) \in \overline{\Pi}$ can be reached from $(x_1, y_1) \in \overline{\Pi}$ by a transaction, then we have $V(x_2, y_2, i) \leq V(x_1, y_1, i)$ for any $i \in \mathcal{M}$.*

We conclude this section by presenting the dynamic programming (DP) equation for the optimal control problem, which can be established by generalizing the arguments in the literature (e.g., [3, Sect. 4]) to the Markovian-modulated systems.

Let τ be an almost surely finite stopping time. Then

$$V(x, y, i) = \sup_{(C,M,N)\in A(x,y,i)} E\bigg[\int_0^\tau e^{-\beta t} U\big(C(t)\big)\,dt + e^{-\beta\tau} V\big(x(\tau), y(\tau), \alpha_\tau\big)\bigg]. \tag{19}$$

Moreover, if (C^*, M^*, N^*) is an optimal policy, then

$$V(x, y, i) = E\bigg[\int_0^\tau e^{-\beta t} U\big(C^*(t)\big)\,dt + e^{-\beta\tau} V\big(x^*(\tau), y^*(\tau), \alpha_\tau\big)\bigg], \tag{20}$$

where $(x^*(t), y^*(t)), t \geq 0$ is the corresponding solution to (3).

4 Viscosity Solution

In this section we show that the value function (10) is a viscosity solution to the HJB system (12). We first give the definition of the viscosity solution to (12), which is a natural modification of that for the second-order nonlinear partial differential equations introduced in [5].

Consider function $v:\overline{\Pi}\times\mathcal{M}\to\mathbb{R}$ with $v_i(x,y)=v(x,y,i)\in\mathbb{R}$, $(x,y)\in\overline{\Pi}$, $i\in\mathcal{M}$. We say that v is continuous if v_i is continuous in (x,y) on $\overline{\Pi}$ for each $i\in\mathcal{M}$.

Definition 4.1 A continuous function $v:\overline{\Pi}\times\mathcal{M}\to\mathbb{R}$ is a viscosity solution of the HJB system (12) iff both 1 and 2 hold.

1. v is a viscosity subsolution of (12) in Π, that is, for each $i\in\mathcal{M}$,

$$\begin{aligned}\min\bigl\{&\mathcal{L}_i\varphi(x_0,y_0)-Qv(x_0,y_0,\cdot)(i)-\widetilde{U}\bigl(\varphi_x(x_0,y_0)\bigr),\\&(1+\lambda)\varphi_x(x_0,y_0)-\varphi_y(x_0,y_0),-(1-\mu)\varphi_x(x_0,y_0)+\varphi_y(x_0,y_0)\bigr\}\le 0,\end{aligned}\tag{21}$$

whenever $\varphi\in C^2(\Pi)$ such that $v_i-\varphi$ has a maxima at $(x_0,y_0)\in\Pi$ and that $v_i(x_0,y_0)=\varphi(x_0,y_0)$.

2. v is a viscosity supersolution of (12) in Π, that is, for each $i\in\mathcal{M}$,

$$\begin{aligned}\min\bigl\{&\mathcal{L}_i\psi(x_0,y_0)-Qv(x_0,y_0,\cdot)(i)-\widetilde{U}\bigl(\psi_x(x_0,y_0)\bigr),\\&(1+\lambda)\psi_x(x_0,y_0)-\psi_y(x_0,y_0),-(1-\mu)\psi_x(x_0,y_0)+\psi_y(x_0,y_0)\bigr\}\ge 0,\end{aligned}\tag{22}$$

whenever $\psi\in C^2(\Pi)$ such that $v_i-\psi$ has a minima at $(x_0,y_0)\in\Pi$ and that $v_i(x_0,y_0)=\psi(x_0,y_0)$.

Using Definition 4.1, we can easily establish the following result. The proof is omitted.

Corollary 4.1 *Let the continuous function $v:\overline{\Pi}\times\mathcal{M}\to\mathbb{R}$ be a viscosity solution of the HJB system (12). If for some $i\in\mathcal{M}$, v_i is twice differentiable at $(x_0,y_0)\in\Pi$, then*

$$\begin{aligned}\min\bigl\{&\mathcal{L}_iv(x_0,y_0,i)-Qv(x_0,y_0,\cdot)(i)-\widetilde{U}\bigl(v_x(x_0,y_0,i)\bigr),\\&(1+\lambda)v_x(x_0,y_0,i)-v_y(x_0,y_0,i),-(1-\mu)v_x(x_0,y_0,i)+v_y(x_0,y_0,i)\bigr\}=0,\end{aligned}\tag{23}$$

that is, the function v_i satisfies equation (23) in the classical sense at (x_0,y_0).

Theorem 4.1 *The value function V is a viscosity solution of the HJB system (12) in Π.*

The proof of Theorem 4.1 is done by using the generalized Itô formula for Markovian-modulated processes and the dynamic programming (DP) equations (19), (20), with distinct treatment of the coupling terms in (12). It is not included due to length limitation (available upon request).

5 Power Utility Function

In this section we present further properties of the value function V for a power utility function $U(c) = c^\gamma/\gamma$, $0 < \gamma < 1$. Note that in this case, $\widetilde{U}(x)$ defined in (14) is given by

$$\widetilde{U}(x) = \frac{1-\gamma}{\gamma} x^{\frac{\gamma}{\gamma-1}}. \tag{24}$$

First, similar to [3, Proposition 3.3], we can show:

Proposition 5.1 *The value function V has the homotheticity property,*

$$V(\theta x, \theta y, i) = \theta^\gamma V(x, y, i), \tag{25}$$

for all $\theta > 0$, $(x, y) \in \overline{\Pi}$ and for each $i \in \mathcal{M}$.

Next we present lower bounds for the value function V. This result generalizes [3, Proposition 3.4] to the regime-switching system for the case of power utility function.

Proposition 5.2 *Define*

$$L_0 = \frac{\beta - \gamma r_*}{1-\gamma}. \tag{26}$$

Then $L_0 > 0$ and

$$V(x, y, i) \geq \begin{cases} \frac{1}{\gamma} L_0^{\gamma-1}[x + (1-\mu)y]^\gamma, & \forall i \in \mathcal{M}, \forall (x, y) \in \overline{\Pi}, y \geq 0, \\ \frac{1}{\gamma} L_0^{\gamma-1}[x + (1+\lambda)y]^\gamma, & \forall i \in \mathcal{M}, \forall (x, y) \in \overline{\Pi}, y < 0. \end{cases} \tag{27}$$

Proof First we note that in view of Remark 2.2, $V(x, y, i) = 0$ on the boundaries $\partial\Pi = \partial_1\Pi \cup \partial_2\Pi$, which agrees with the lower bounds given in (27).

We now show that the lower bounds (27) hold in Π. To this end, given $i \in \mathcal{M}$ and $(x, y) \in \Pi$. We construct an admissible control policy (C, M, N) in $\mathcal{A}(x, y, i)$ that takes one initial jump to a point on the x-axis (if $y \neq 0$), and after that makes no further transfers between the bond and the stock, and consumes at a rate proportional to the wealth. Specifically, if $y > 0$, we set $M(0) = 0$, $N(0) = y$, resulting in $x(0) = x + (1-\mu)y$, $y(0) = 0$; if $y < 0$, we set $M(0) = -y$, $N(0) = 0$, resulting in $x(0) = x + (1+\lambda)y$, $y(0) = 0$. For both cases, we set $C(t) = \theta x(t)$, $t \geq 0$, where $\theta > 0$ is a positive constant. Applying the policy (C, M, N) to (3), we have $y(t) = 0$, $t \geq 0$ and $dx(t) = [r(\alpha_t)x(t) - \theta x(t)]\, dt$, which leads to

$$x(t) = x(0)e^{\int_0^t [r(\alpha_s)-\theta]\, ds} \quad \text{with } x(0) = \begin{cases} x + (1-\mu)y, & \text{if } y \geq 0, \\ x + (1+\lambda)y, & \text{if } y < 0. \end{cases}$$

It follows that

$$\begin{aligned} V(x,y,i) &\geq E\biggl[\int_0^\infty e^{-\beta t} U\bigl(C(t)\bigr)\,dt\biggr] = E\biggl[\int_0^\infty e^{-\beta t} \frac{\theta^\gamma x^\gamma(t)}{\gamma}\,dt\biggr] \\ &= \frac{1}{\gamma}\theta^\gamma x^\gamma(0) E\biggl[\int_0^\infty e^{-\beta t} e^{\int_0^t \gamma[r(\alpha_s)-\theta]\,ds}\,dt\biggr] \\ &\geq \frac{1}{\gamma}\theta^\gamma x^\gamma(0)\int_0^\infty e^{-\beta t} e^{\gamma[r_*-\theta]t}\,dt = \frac{\theta^\gamma x^\gamma(0)}{\gamma(\beta-\gamma r_*+\gamma\theta)}, \end{aligned} \tag{28}$$

where for the last equality, we assumed $\beta - \gamma r_* + \gamma\theta > 0$. If $\frac{\beta-\gamma r_*}{\gamma} \leq 0$, we then let $\theta \downarrow -\frac{\beta-\gamma r_*}{\gamma}$ in (28) and obtain $V(x,y,i) = \infty$ for all $(x,y) \in \Pi$ and for each $i \in \mathcal{M}$, which is a violation of Assumption 2.1. Hence $\frac{\beta-\gamma r_*}{\gamma} > 0$ which is equivalent to $L_0 > 0$.

Finally, the maximum value of $\frac{\theta^\gamma}{(\beta-\gamma r_*+\gamma\theta)}$ is attained at $\theta = \frac{\beta-\gamma r_*}{1-\gamma} = L_0$ and the maximum value is equal to $L_0^{\gamma-1}$. Consequently, we obtain (27). □

Next we study the upper bounds for V near the boundary $\partial\Pi$. The introduction of regime switching in the model significantly complicates the analysis comparing to the case without regime switching as in [3, Sect. 5].

To proceed, let $0 < \delta_0 < 1 - \mu$ be such that

$$0 < 1 - \mu - \delta_0 < \frac{(1-\gamma)(1-\mu)\sigma_*^2}{2(\eta^* - r_*)}. \tag{29}$$

Define

$$A_0 = \left(\frac{\gamma V^*(-\delta_0, 1)}{(1-\mu-\delta_0)^\gamma}\right)^{\frac{1}{\gamma-1}}, \tag{30}$$

where $V^*(x,y) = \max_{i\in\mathcal{M}} V(x,y,i)$. In view of Proposition 5.2 and that $\gamma < 1$, we have

$$\frac{1}{\gamma}L_0^{\gamma-1}(1-\mu-\delta_0)^\gamma \leq V(-\delta_0,1,i) \leq V^*(-\delta_0,1) = \frac{1}{\gamma}A_0^{\gamma-1}(1-\mu-\delta_0)^\gamma$$

which implies

$$A_0 \leq L_0. \tag{31}$$

Define an open subset D_0 in Π by

$$D_0 = \left\{(x,y) : x < 0, -\frac{x}{1-\mu} < y < -\frac{x}{\delta_0}\right\} \subset \Pi, \tag{32}$$

and define

$$\varphi(x,y) = \frac{1}{\gamma}A_0^{\gamma-1}\bigl[x + (1-\mu)y\bigr]^\gamma, \quad (x,y) \in \overline{D}_0. \tag{33}$$

It is readily seen that

$$-(1-\mu)\varphi_x(x,y)+\varphi_y(x,y)=0, \qquad (1+\lambda)\varphi_x(x,y)-\varphi_y(x,y)>0, \quad \forall (x,y)\in D_0. \tag{34}$$

Moreover, we have

Lemma 5.1 *For each $i \in \mathcal{M}$ and all $(x,y)\in D_0$,*

$$\mathcal{L}_i\varphi(x,y)-\widetilde{U}\big(\varphi_x(x,y)\big)\geq 0. \tag{35}$$

Proof In view of (11), (24) and (33), we have, for $i\in\mathcal{M}$ and $(x,y)\in D_0$,

$$\begin{aligned}
&\mathcal{L}_i\varphi(x,y)-\widetilde{U}\big(\varphi_x(x,y)\big)\\
&\quad = A_0^{\gamma-1}\big[x+(1-\mu)y\big]^{\gamma}\bigg[\frac{\beta-\gamma r_i}{\gamma}-\frac{1-\gamma}{\gamma}A_0\\
&\qquad + \frac{(1-\mu)y}{x+(1-\mu)y}\bigg(\frac{1}{2}(1-\gamma)\sigma_i^2\frac{(1-\mu)y}{x+(1-\mu)y}-(\eta_i-r_i)\bigg)\bigg].
\end{aligned} \tag{36}$$

Since $A_0\leq L_0$, we have, in view of (26),

$$\frac{\beta-\gamma r_i}{\gamma}-\frac{1-\gamma}{\gamma}A_0\geq\frac{\beta-\gamma r_i}{\gamma}-\frac{1-\gamma}{\gamma}L_0=r_*-r_i. \tag{37}$$

For $(x,y)\in D_0$, $x<-\delta_0 y$, then $0<x+(1-\mu)y<(1-\mu-\delta_0)y$. It follows that

$$\frac{(1-\mu)y}{x+(1-\mu)y}>\frac{1-\mu}{1-\mu-\delta_0}>1. \tag{38}$$

Also (29) implies

$$\frac{1-\mu}{1-\mu-\delta_0}>\frac{2(\eta^*-r_*)}{(1-\gamma)\sigma_*^2}\geq\frac{2(\eta_i-r_*)}{(1-\gamma)\sigma_i^2}, \quad \forall i\in\mathcal{M}. \tag{39}$$

Using (37), (38), (39) in (36), we have

$$\begin{aligned}
&\mathcal{L}_i\varphi(x,y)-\widetilde{U}\big(\varphi_x(x,y)\big)\\
&\quad\geq A_0^{\gamma-1}\big[x+(1-\mu)y\big]^{\gamma}\\
&\qquad\times\bigg[(r_*-r_i)+\frac{(1-\mu)y}{x+(1-\mu)y}\bigg(\frac{1}{2}(1-\gamma)\sigma_i^2\frac{2(\eta_i-r_*)}{(1-\gamma)\sigma_i^2}-(\eta_i-r_i)\bigg)\bigg]\\
&\quad= A_0^{\gamma-1}\big[x+(1-\mu)y\big]^{\gamma}\bigg[(r_*-r_i)+\frac{(1-\mu)y}{x+(1-\mu)y}(r_i-r_*)\bigg]\\
&\quad\geq A_0^{\gamma-1}\big[x+(1-\mu)y\big]^{\gamma}\big[(r_*-r_i)+(r_i-r_*)\big]=0.
\end{aligned}$$

This completes the proof. □

Next, we introduce a function $h_0 :]1, \infty[\to \mathbb{R}$ defined by

$$h_0(z) = \frac{(1-\gamma)\sigma_*^2}{2(z-1)} + \big(r_* - r^*\big)z + (\eta_* - r_*), \quad z \in]1, \infty[. \tag{40}$$

Choose $\xi_0 > 1 + \lambda$ such that $h_0(\xi_0/(1+\lambda)) \ge 0$. Define

$$B_0 = \left(\frac{\gamma V^*(\xi_0, -1)}{(\xi_0 - 1 - \lambda)^\gamma}\right)^{\frac{1}{\gamma-1}}. \tag{41}$$

In view of Proposition 5.2, we have

$$\frac{1}{\gamma} L_0^{\gamma-1}(\xi_0 - 1 - \lambda)^\gamma \le V(\xi_0, -1, i) \le V^*(\xi_0, -1) = \frac{1}{\gamma} B_0^{\gamma-1}(\xi_0 - 1 - \lambda)^\gamma$$

which implies

$$B_0 \le L_0. \tag{42}$$

Define an open subset E_0 in Π by

$$E_0 = \left\{(x, y) : x > 0, -\frac{x}{1+\lambda} < y < -\frac{x}{\xi_0}\right\} \subset \Pi. \tag{43}$$

Define

$$\psi(x, y) = \frac{1}{\gamma} B_0^{\gamma-1}\big[x + (1+\lambda)y\big]^\gamma, \quad (x, y) \in \overline{E}_0. \tag{44}$$

It is readily seen that

$$-(1-\mu)\psi_x(x, y) + \psi_y(x, y) > 0, \qquad (1+\lambda)\psi_x(x, y) - \psi_y(x, y) = 0, \quad \forall (x, y) \in E_0. \tag{45}$$

Moreover, we have

Lemma 5.2 *For each $i \in \mathcal{M}$ and all $(x, y) \in E_0$,*

$$\mathcal{L}_i \psi(x, y) - \widetilde{U}\big(\psi_x(x, y)\big) \ge 0. \tag{46}$$

Proof For $i \in \mathcal{M}$ and $(x, y) \in E_0$, we have

$$\begin{aligned}\mathcal{L}_i \psi(x, y) - \widetilde{U}\big(\psi_x(x, y)\big) &= B_0^{\gamma-1}\big[x + (1+\lambda)y\big]^{\gamma-1}\big(-(1+\lambda)y\big) \\ &\quad \times \left[\left(\frac{\beta - \gamma r_i}{\gamma} - \frac{1-\gamma}{\gamma} B_0\right)\frac{x + (1+\lambda)y}{-(1+\lambda)y}\right. \\ &\quad \left. + \left(\frac{1}{2}(1-\gamma)\sigma_i^2 \frac{-(1+\lambda)y}{x + (1+\lambda)y} + (\eta_i - r_i)\right)\right]. \end{aligned} \tag{47}$$

Since $B_0 \le L_0$, we have, in view of (26),

$$\frac{\beta-\gamma r_i}{\gamma}-\frac{1-\gamma}{\gamma}B_0 \ge \frac{\beta-\gamma r_i}{\gamma}-\frac{1-\gamma}{\gamma}L_0 = r_* - r_i. \tag{48}$$

For $(x,y)\in E_0$, $x<-\xi_0 y$, then $0<x+(1+\lambda)y<(\xi_0-1-\lambda)(-y)$. It follows that

$$0<\frac{x+(1+\lambda)y}{-(1+\lambda)y}<\frac{\xi_0-1-\lambda}{1+\lambda}=\frac{\xi_0}{1+\lambda}-1, \tag{49}$$

and then

$$\frac{-(1+\lambda)y}{x+(1+\lambda)y}>\frac{1}{\frac{\xi_0}{1+\lambda}-1}. \tag{50}$$

In view of (48), (49) and noting that $r_*-r_i\le 0$, we have

$$\left(\frac{\beta-\gamma r_i}{\gamma}-\frac{1-\gamma}{\gamma}B_0\right)\frac{x+(1+\lambda)y}{-(1+\lambda)y}\ge (r_*-r_i)\left(\frac{\xi_0}{1+\lambda}-1\right). \tag{51}$$

Using (50) and (51) in (47), we obtain

$$\begin{aligned}
&\mathcal{L}_i\psi(x,y)-\widetilde{U}\big(\psi_x(x,y)\big)\\
&\quad\ge B_0^{\gamma-1}\big[x+(1+\lambda)y\big]^{\gamma-1}\big(-(1+\lambda)y\big)\\
&\qquad\times\left[(r_*-r_i)\left(\frac{\xi_0}{1+\lambda}-1\right)+\left(\frac{1}{2}(1-\gamma)\sigma_i^2\frac{1}{\frac{\xi_0}{1+\lambda}-1}+(\eta_i-r_i)\right)\right]\\
&\quad\ge B_0^{\gamma-1}\big[x+(1+\lambda)y\big]^{\gamma-1}\big(-(1+\lambda)y\big)\\
&\qquad\times\left[\left(r_*-r^*\right)\frac{\xi_0}{1+\lambda}+(\eta_*-r_*)+\frac{1}{2}(1-\gamma)\sigma_*^2\frac{1}{\frac{\xi_0}{1+\lambda}-1}\right]\\
&\quad= B_0^{\gamma-1}\big[x+(1+\lambda)y\big]^{\gamma-1}\big(-(1+\lambda)y\big)h_0\left(\frac{\xi_0}{1+\lambda}\right)\ge 0.
\end{aligned}$$

This completes the proof. □

Using Lemmas 5.1 and 5.2, we can establish the following upper bounds for the value function V.

Proposition 5.3 *For each $i\in\mathcal{M}$, we have*

$$V(x,y,i)\le\begin{cases}\frac{1}{\gamma}A_0^{\gamma-1}[x+(1-\mu)y]^{\gamma}, & \textit{if } (x,y)\in\overline{D}_0,\\ \frac{1}{\gamma}B_0^{\gamma-1}[x+(1+\lambda)y]^{\gamma}, & \textit{if } (x,y)\in\overline{E}_0.\end{cases} \tag{52}$$

The proof is almost the same as that of [3, Proposition 5.7] since the functions φ in Lemma 5.1 and ψ in Lemma 5.2 are independent of i. In Proposition 5.4 presented next we show that in fact we can find regime-dependent coefficients A_i and B_i so that

(52) is replaced by an equality in appropriately chosen subsets $D \subset D_0$ and $E \subset E_0$, respectively.

To this end, we first establish results similar to Lemmas 5.1 and 5.2 for appropriately defined regime-dependent functions. The proofs are similar and omitted.

Lemma 5.3 *Let* $q^* = \max_{i \in \mathcal{M}}(-q_{ii})$. *Choose* $\delta \in [\delta_0, 1-\mu[$ *such that*

$$0 < 1-\mu-\delta < \frac{(1-\gamma)(1-\mu)\sigma_*^2}{2[(\eta^* - r_*) + \frac{q^*}{\gamma}(\frac{A_0^{\gamma-1}}{L_0^{\gamma-1}} - 1)]}. \tag{53}$$

Define an open subset D *in* Π *by*

$$D = \left\{(x,y) : x < 0, -\frac{x}{1-\mu} < y < -\frac{x}{\delta}\right\} \subset D_0. \tag{54}$$

For each $i \in \mathcal{M}$, *define*

$$\varphi(x,y,i) = \frac{1}{\gamma} A_i^{\gamma-1}\big[x + (1-\mu)y\big]^\gamma, \quad (x,y) \in \overline{D}, \tag{55}$$

where

$$A_i = \left(\frac{\gamma V(-\delta,1,i)}{(1-\mu-\delta)^\gamma}\right)^{\frac{1}{\gamma-1}}, \quad i \in \mathcal{M}. \tag{56}$$

Then we have

$$\begin{cases} -(1-\mu)\varphi_x(x,y,i) + \varphi_y(x,y,i) = 0, \\ (1+\lambda)\varphi_x(x,y,i) - \varphi_y(x,y,i) > 0, \end{cases} \quad \forall (x,y) \in D, i \in \mathcal{M}, \tag{57}$$

and

$$\mathcal{L}_i\varphi(x,y,i) - Q\varphi(x,y,\cdot)(i) - \widetilde{U}\big(\varphi_x(x,y,i)\big) \ge 0, \quad \forall (x,y) \in D, i \in \mathcal{M}. \tag{58}$$

Lemma 5.4 *Choose* $\xi \in]1+\lambda, \xi_0]$ *such that* $h(\xi/(1+\lambda)) \ge 0$ *where the function* $h :]1,\infty[\to \mathbb{R}$ *is defined by*

$$h(z) = \frac{(1-\gamma)\sigma_*^2}{2(z-1)} + \big(r_* - r^*\big)z + (\eta_* - r_*) - \frac{q^*}{\gamma}\left(\frac{B_0^{\gamma-1}}{L_0^{\gamma-1}} - 1\right)(z-1), \quad z \in]1,\infty[. \tag{59}$$

Define an open subset E *in* Π *by*

$$E = \left\{(x,y) : x > 0, -\frac{x}{1+\lambda} < y < -\frac{x}{\xi}\right\} \subset E_0. \tag{60}$$

For each $i \in \mathcal{M}$, *define*

$$\psi(x,y,i) = \frac{1}{\gamma} B_i^{\gamma-1}\big[x + (1+\lambda)y\big]^\gamma, \quad (x,y) \in \overline{E}, \tag{61}$$

where

$$B_i = \left(\frac{\gamma V(\xi, -1, i)}{(\delta - 1 - \xi)^\gamma} \right)^{\frac{1}{\gamma-1}}, \quad i \in \mathcal{M}. \tag{62}$$

Then we have

$$\begin{cases} -(1-\mu)\psi_x(x,y,i) + \psi_y(x,y,i) > 0, \\ (1+\lambda)\psi_x(x,y,i) - \psi_y(x,y,i) = 0, \end{cases} \quad \forall (x,y) \in E, i \in \mathcal{M}, \tag{63}$$

and

$$\mathcal{L}_i \psi(x,y,i) - Q\psi(x,y,\cdot)(i) - \widetilde{U}\big(\psi_x(x,y,i)\big) \geq 0, \quad \forall (x,y) \in E, i \in \mathcal{M}. \tag{64}$$

Based on Lemmas 5.3 and 5.4, we can show the following assertion.

Proposition 5.4 *For each $i \in \mathcal{M}$, we have*

$$V(x,y,i) = \begin{cases} \frac{1}{\gamma} A_i^{\gamma-1}[x + (1-\mu)y]^\gamma, & \text{if } (x,y) \in \overline{D}, \\ \frac{1}{\gamma} B_i^{\gamma-1}[x + (1+\lambda)y]^\gamma, & \text{if } (x,y) \in \overline{E}, \end{cases} \tag{65}$$

where A_i, B_i, $i \in \mathcal{M}$, D and E are introduced in Lemmas 5.3 and 5.4.

Proof We show the $\geq$ part here. The proof of the $\leq$ part generalizes the arguments of [3, Proposition 5.7] by taking into account the regime-switching component and is deferred to the Appendix.

Given $(x,y) \in D$. Let

$$x_1 = x + (1-\mu)\left(\frac{-x - \delta y}{1 - \mu - \delta}\right) = -\delta\left(\frac{x + (1-\mu)y}{1-\mu-\delta}\right),$$
$$y_1 = y - \left(\frac{-x - \delta y}{1-\mu-\delta}\right) = \frac{x + (1-\mu)y}{1-\mu-\delta}.$$

Then $(x_1, y_1) \in \partial D$ and (x_1, y_1) can be reached from (x,y) by a transaction. In view of Propositions 3.3, 5.1 and (56), we have, for each $i \in \mathcal{M}$,

$$V(x,y,i) \geq V(x_1,y_1,i) = V(-\delta y_1, y_1, i) = V(-\delta, 1, i) y_1^\gamma$$
$$= \frac{1}{\gamma} A_i^{\gamma-1}\big(y_1(1-\mu-\delta)\big)^\gamma = \frac{1}{\gamma} A_i^{\gamma-1}\big(x + (1-\mu)y\big)^\gamma.$$

Similarly, given $(x,y) \in E$. Let

$$x_2 = x - (1+\lambda)\left(\frac{-x - \xi y}{\xi - 1 - \lambda}\right) = \xi\left(\frac{x + (1+\lambda)y}{\xi - 1 - \lambda}\right),$$
$$y_2 = y + \left(\frac{-x - \xi y}{\xi - 1 - \lambda}\right) = -\frac{x + (1+\lambda)y}{\xi - 1 - \lambda}.$$

Then $(x_2, y_2) \in \partial E$ and (x_2, y_2) can be reached from (x, y) by a transaction. It follows that, for each $i \in \mathcal{M}$,

$$\begin{aligned} V(x, y, i) &\geq V(x_2, y_2, i) = V(-\xi y_2, y_2, i) = V(\xi, -1, i)(-y_2)^\gamma \\ &= \frac{1}{\gamma} B_i^{\gamma-1}\big(-y_2(\xi - 1 - \lambda)\big)^\gamma = \frac{1}{\gamma} B_i^{\gamma-1}\big(x + (1+\lambda)y\big)^\gamma. \end{aligned}$$

This completes the proof. □

The following properties will be used in the next section. Their proofs are similar to [3, Proposition 3.6, Corollary 3.7] and hence omitted.

Proposition 5.5 *Give $i \in \mathcal{M}$ and $(x, y) \in \Pi$. Then for all $h \geq 0$, we have*

$$V(x+h, y, i) \geq \begin{cases} [1 + \frac{h}{x+(1+\lambda)y}]^\gamma V(x, y, i), & \text{if } y \geq 0, \\ [1 + \frac{h}{x+(1-\mu)y}]^\gamma V(x, y, i), & \text{if } y < 0. \end{cases} \tag{66}$$

Corollary 5.1 *Give $i \in \mathcal{M}$ and $(x_0, y_0) \in \Pi$. Let $\phi : \Pi \to \mathbb{R}$ be a differentiable function satisfying $\phi(x_0, y_0) = V(x_0, y_0, i)$. If $\phi \geq V_i$ or $\phi \leq V_i$ on Π, then,*

$$\phi_x(x_0, y_0) \geq \begin{cases} \frac{\gamma V(x_0, y_0, i)}{x_0+(1+\lambda)y_0}, & \text{if } y_0 \geq 0, \\ \frac{\gamma V(x_0, y_0, i)}{x_0+(1-\mu)y_0}, & \text{if } y_0 < 0. \end{cases} \tag{67}$$

6 Further Analysis for the Power Utility Function

In this section we continue the analysis of the value function V for the power utility. By the methods of convex analysis, we can show that for each regime $i \in \mathcal{M}$, the solvency region Π can be partitioned into three cones corresponding to the three arguments of the minimum operator in (12). This part of analysis is almost the same as that in [3, Sect. 6] since each of the m_0 functions V_i, $1 \leq i \leq m_0$ can be dealt with almost separately and coupling due to regime switching does not complicate the analysis. We include in the Appendix the main results that are needed in this section, using our notations.

Due to the homotheticity property (25), for each $i \in \mathcal{M}$, the function V_i can be transformed into a single-variable function by making a proper change of variable. For this purpose, define $\mathcal{I} =]-\frac{1}{\lambda}, \frac{1}{\mu}[$ and

$$u_i(z) = u(z, i) = V(1-z, z, i), \quad \forall z \in \overline{\mathcal{I}}, \forall i \in \mathcal{M}. \tag{68}$$

In view of (25), we have

$$V(x, y, i) = (x+y)^\gamma u\left(\frac{y}{x+y}, i\right), \quad \forall (x, y) \in \overline{\Pi} \setminus \{(0,0)\}, \forall i \in \mathcal{M}. \tag{69}$$

Let $z = \frac{y}{x+y}$. With the connection $w_i(x, y) = (x+y)^\gamma \varpi_i(z)$, direct computations show that the system (12) for w_i is transformed into the following system for ϖ_i, $1 \le i \le m_0$:

$$\min\Bigg\{(\beta - q_{ii})\varpi_i(z) - \gamma d_{i1}(z)\varpi_i(z) - d_{i2}(z)\varpi_i'(z) - d_{i3}(z)\varpi_i''(z) \\ - \sum_{j\neq i} q_{ij}\varpi_j(z) - \widetilde{U}\big(\gamma\varpi_i(z) - z\varpi_i'(z)\big), \\ \gamma\varpi_i(z) + d_4(z)\varpi_i'(z), \gamma\varpi_i(z) - d_5(z)\varpi_i'(z)\Bigg\} = 0 \tag{70}$$

for $z \in \mathcal{I}$, where

$$d_{i1}(z) = r_i + (\eta_i - r_i)z - \frac{1}{2}(1-\gamma)\sigma_i^2 z^2,$$
$$d_{i2}(z) = (\eta_i - r_i)z(1-z) - (1-\gamma)\sigma_i^2 z^2(1-z),$$
$$d_{i3}(z) = \frac{1}{2}\sigma_i^2 z^2(1-z)^2,$$

for $i \in \mathcal{M}$ and

$$d_4(z) = \frac{1-\mu z}{\mu}, \qquad d_5(z) = \frac{1+\lambda z}{\lambda},$$

which are independent of i.

Proposition 6.1 *The function u defined in (68) is a viscosity solution of the system (70) in $\mathcal{I}$.*

The proof of Proposition 6.1 is straightforward and hence omitted.

Note that u_i is the function V_i evaluated along a line segment of $x + y = 1$ in $\overline{\Pi}$. Therefore u_i is a concave function in $z \in \mathcal{I}$ for each $i \in \mathcal{M}$. Applying the arguments in [3, pp. 648–649] to each function u_i, $i \in \mathcal{M}$ produces the following results. For any given $z \in \mathcal{I}$, the difference quotient $\frac{u_i(z+h)-u_i(z)}{h}$ is a non-increasing function of h, so the right derivative $D^+u_i(z)$ and left derivative $D^-u_i(z)$ exist and are finite, where

$$D^{\pm}u_i(z) = \lim_{h\to 0\pm}\frac{u_i(z+h) - u_i(z)}{h}.$$

Moreover, D^+ is right continuous and D^- is left continuous; they are both non-increasing and agree except on a countable subset $\mathcal{N}_i$ of $\mathcal{I}$. As a result, u_i is differentiable on $\mathcal{I} \setminus \mathcal{N}_i$ and $\partial u_i(z) = [D^+u_i(z), D^-u_i(z)]$ for all $z \in \mathcal{I}$, where ∂u_i denotes the sub-differential of u_i which can be defined similar to (101). Since u_i' is defined almost everywhere in $\mathcal{I}$ and is non-increasing, its pointwise derivative u_i'' is also defined almost everywhere in $\mathcal{I}$ and is locally integrable. Similar to Corollary 4.1, one can show that u_i satisfies (70) almost everywhere in the classical sense.

The next proposition is concerned with the C^1 property of the function u_i. It generalizes [3, Proposition 8.2] to the regime-switching case for the power utility function. The proof is not included due to length limitation (available upon request).

Proposition 6.2 *For each $i \in \mathcal{M}$, the function u_i is C^1 on $\mathcal{I} \setminus \{0\}$. If for some $i \in \mathcal{M}$, u_i is not continuously differentiable at the point* 0, *then we have*

$$V(x,0,i) = x^\gamma u_i(0), \quad \forall x > 0, \tag{71}$$

where $u_i(0)$ satisfies the equation

$$(\beta - q_{ii} - \gamma r_i)u_i(0) - \sum_{j \neq i} q_{ij} u_j(0) - \frac{1-\gamma}{\gamma}\big(\gamma u_i(0)\big)^{\frac{\gamma}{\gamma-1}} = 0. \tag{72}$$

Moreover, its left and right derivatives exist and are limits of its derivatives from the left and right sides of 0, *respectively.*

Remark 6.1 What we know from Proposition 6.2 is that if u_i is not continuously differentiable at 0 for some $i \in \mathcal{M}$, then this u_i satisfies (72) for this i. However, it is possible that u_k is continuously differentiable at 0 for some $k \neq i$ and we do not know whether u_k satisfies (72) (with i replaced by k) or not. In fact, we can only show that in this case u_k satisfies the inequality

$$(\beta - q_{kk} - \gamma r_k)u_k(0) - \sum_{j \neq k} q_{kj} u_j(0) - \widetilde{U}\big(\gamma u_k(0)\big) \geq 0. \tag{73}$$

Corollary 6.1 *The value function $V_i(x,y)$ is C^1 on $\Pi \setminus \{(x,0) : x > 0\}$ for each $i \in \mathcal{M}$. If V_i is not continuously differentiable on $\{(x,0) : x > 0\}$ for some $i \in \mathcal{M}$, then $V_i(x,0)$ is given by (71) for this i. Moreover, even if V_i is not continuously differentiable on $\{(x,0) : x > 0\}$ for some $i \in \mathcal{M}$, the partial derivative $V_x(x,0,i)$ is defined and continuous there, and the two one-sided partial derivatives*

$$V_y(x,0\pm,i) = \lim_{h \to 0\pm} \frac{V(x,h,i) - V(x,0,i)}{h}, \quad x > 0$$

are defined and satisfy the one-sided continuity conditions

$$V_y(x,0\pm,i) = \lim_{(\tilde{x},\tilde{y}) \to (x,0\pm)} V_y(\tilde{x},\tilde{y},i), \quad x > 0.$$

From the convex analysis results of the value function (see the Appendix) we have, for each $i \in \mathcal{M}$, we can partition Π into three open, nonintersecting (possibly empty) cones SA^i, BU^i and NT^i. Let $\Pi^* = \Pi \setminus \{(x,0) : x > 0\}$. Then from Corollary 6.1, V_i is C^1 on Π^* for each $i \in \mathcal{M}$. In view of Corollary A.2 and Proposition A.4, We can characterize those cones by

$$\overline{SA^i} = \overline{\big\{(x,y) \in \Pi^* : -(1-\mu)V_x(x,y,i) + V_y(x,y,i) = 0\big\}}, \tag{74}$$

$$\overline{BU^i} = \overline{\big\{(x,y) \in \Pi^* : (1+\lambda)V_x(x,y,i) - V_y(x,y,i) = 0\big\}}, \tag{75}$$

and

$$NT^i \setminus \{(x,0): x>0\} = \{(x,y)\in \Pi^*: -(1-\mu)V_x(x,y,i)+V_y(x,y,i)>0,\ (1+\lambda)V_x(x,y,i)-V_y(x,y,i)>0\}. \tag{76}$$

Proposition 6.3 *If for some $i \in \mathcal{M}$, $\eta_i > r_i$, then $NT^i \neq \emptyset$.*

Proof Recall that both SA^i and BU^i are nonempty and $SA^i \cap BU^i = \emptyset$ for each $i \in \mathcal{M}$. Suppose $\eta_i > r_i$ for some $i \in \mathcal{M}$ and $NT^i = \emptyset$. Then the two cones SA^i and BU^i would share a half-line boundary $H^i = \overline{SA^i} \cap \overline{BU^i} \setminus \{(0,0)\}$ in Π. We divide the discussion of H^i into two cases.

Case 1: If V_i is C^1 on H^i, then (74) and (75) imply

$$\begin{cases} -(1-\mu)V_x(x,y,i)+V_y(x,y,i)=0, \\ (1+\lambda)V_x(x,y,i)-V_y(x,y,i)=0, \end{cases} \quad (x,y)\in H^i$$

due to the continuity of V_x and V_y. This leads to $V_x(x,y,i) = V_y(x,y,i) = 0$, $(x,y) \in H^i$, a contradiction with (104).

Case 2: If V_i is not C^1 on H^i, then we must have $H^i = \{(x,0): x>0\}$. In view of (71), (118) and the continuity of V_i, we must have

$$\frac{1}{\gamma}A_i^{\gamma-1} = \frac{1}{\gamma}B_i^{\gamma-1} = u_i(0). \tag{77}$$

If $x>0$ and $y>0$ (the proof for $x>0$ and $y<0$ is completely analogous), then $(x,y)\in SA^i$ and $V_i = u_i(0)[x+(1-\mu)y]^\gamma$ in view of (118) and (77). Note that V_i is twice differentiable at (x,y). In light of Corollary 4.1, V_i satisfies (23) in the classical sense at the point (x,y). Then we have

$$\mathcal{L}_i V(x,y,i) - QV(x,y,\cdot)(i) - \widetilde{U}\big(V_x(x,y,i)\big) \geq 0. \tag{78}$$

On the other hand, direct calculation produces

$$\begin{aligned}
&\mathcal{L}_i V(x,y,i) - QV(x,y,\cdot)(i) - \widetilde{U}\big(V_x(x,y,i)\big) \\
&\quad = \gamma u_i(0)\big[x+(1-\mu)y\big]^\gamma \\
&\qquad \times \bigg[\frac{\beta-\gamma r_i}{\gamma} - \frac{1-\gamma}{\gamma}\big(\gamma u_i(0)\big)^{\frac{1}{\gamma-1}} + \frac{(1-\mu)y}{[x+(1-\mu)y]^2} \\
&\qquad \times \bigg(\frac{1}{2}(1-\gamma)\sigma_i^2(1-\mu)y - (\eta_i - r_i)\big[x+(1-\mu)y\big]\bigg)\bigg] \\
&\qquad - \sum_{j\neq i} q_{ij}\big(V(x,y,j) - u_i(0)\big[x+(1-\mu)y\big]^\gamma\big)
\end{aligned}$$

$$= \big[x+(1-\mu)y\big]^{\gamma}\bigg[(\beta-\gamma r_i-q_{ii})u_i(0)-\frac{1-\gamma}{\gamma}\big(\gamma u_i(0)\big)^{\frac{\gamma}{\gamma-1}}+\frac{\gamma u_i(0)(1-\mu)y}{[x+(1-\mu)y]^2}$$
$$\times\left(\frac{1}{2}(1-\gamma)\sigma_i^2(1-\mu)y-(\eta_i-r_i)\big[x+(1-\mu)y\big]\right)\bigg]-\sum_{j\neq i}q_{ij}V(x,y,j)$$
$$=\sum_{j\neq i}q_{ij}\big(\big[x+(1-\mu)y\big]^{\gamma}u_j(0)-V(x,y,j)\big)$$
$$+\big[x+(1-\mu)y\big]^{\gamma-2}\gamma u_i(0)(1-\mu)y$$
$$\times\left(\frac{1}{2}(1-\gamma)\sigma_i^2(1-\mu)y-(\eta_i-r_i)\big[x+(1-\mu)y\big]\right), \tag{79}$$

where (72) is used for the last equality. Using (69) and the continuity of function u_j, we have, as $y\to 0+$,

$$\big[x+(1-\mu)y\big]^{\gamma}u_j(0)-V(x,y,j)$$
$$=\big[x+(1-\mu)y\big]^{\gamma}u_j(0)-(x+y)^{\gamma}u_j\left(\frac{y}{x+y}\right)\to 0.$$

It then follows that for y sufficiently close to $0+$,

$$\mathcal{L}_iV(x,y,i)-QV(x,y,\cdot)(i)-\widetilde{U}\big(V_x(x,y,i)\big)<0$$

because $\eta_i-r_i>0$. This contradicts (78). □

Remark 6.2 Recall that for the regime-switching model, we do not require $\eta_i>r_i$ for all $i\in\mathcal{M}$. Consequently, if for some $i\in\mathcal{M}$, $\eta_i\le r_i$, then it may be possible to have $NT^i=\emptyset$ and $H^i=\{(x,0):x>0\}$. In this case, the optimal trading policy is to clean up positions in the stock immediately at $t=0$ (sell if $y(0-)>0$ and buy if $y(0-)<0$) and never hold the stock, as long as the regime stays at i.

If $NT^i\neq\emptyset$ for $i\in\mathcal{M}$, then there exist numbers θ_1^i,θ_2^i in $\mathcal{I}$ with $\theta_1^i<\theta_2^i$ such that

$$NT^i=\left\{(x,y)\in\Pi:\theta_1^i<\frac{y}{x+y}<\theta_2^i\right\}. \tag{80}$$

We have the following property.

Proposition 6.4 *If $NT^i\neq\emptyset$ for some $i\in\mathcal{M}$, then the function u_i defined by (68) is C^2 on $]\theta_1^i,\theta_2^i[\setminus\{0,1\}$. Moreover, u_i satisfies the following equation in the classical sense on this set*:

$$(\beta-q_{ii})u_i(z)-\gamma d_{i1}(z)u_i(z)-d_{i2}(z)u_i'(z)-d_{i3}(z)u_i''(z)$$
$$-\sum_{j\neq i}q_{ij}u_j(z)-\widetilde{U}\big(\gamma u_i(z)-zu_i'(z)\big)=0,\quad z\in]\theta_1^i,\theta_2^i[\setminus\{0,1\}. \tag{81}$$

The proof of Proposition 6.4 generalizes that of [3, Proposition 8.5] to the regime-switching case and is not included due to length limitation (available upon request).

Corollary 6.2 *If $NT^i \neq \emptyset$ for some $i \in \mathcal{M}$, then the function V_i is C^2 in the set $NT^i \setminus \{(x, y) : x = 0$ or $y = 0\}$ and satisfies the equation*

$$\mathcal{L}_i V(x, y, i) - QV(x, y, \cdot)(i) - \widetilde{U}\big(V_x(x, y, i)\big) = 0$$

in the classical sense.

Remark 6.3 Shreve and Soner show in [3, Corollary 8.8] that the buy region always contains the cone $\{(x, y) \in \Pi, y < 0\}$ for their model. For the regime-switching model considered in this paper, we are unable to reach the same assertion. What we have obtained is this: both the sell region SA^i and the buy region BU^i are nonempty for any regime $i \in \mathcal{M}$.

7 Conclusions

In this paper we studied an infinite-horizon problem of optimal investment and consumption with proportional transaction costs in continuous-time regime-switching models. In particular, we extended the analysis and results of [3] for a power utility function to the regime-switching case. It has been well documented in the literature that including a regime-switching component in the models can better describe the market changes under different macroeconomic conditions. Therefore the results obtained in this work can help us get a better insight into the complicated behavior of the optimal dynamics of the investment and consumption problems. We note that the market regime α_t is assumed to be observable in the current setting. However, in practice, information of α_t is usually hidden in noisy market data and not directly observable. Therefore it will be a very interesting project for future research to develop nonlinear filtering methods (for example, Wonham filter) for estimating the conditional probabilities of α_t given the stock process. See [6] and [7] for related works that will be helpful in this direction. Other interesting projects for future research include finite-horizon problems, numerical schemes and asymptotic analysis for the investment and consumption problems in regime-switching models with transaction costs.

Acknowledgements The author would like to thank the two anonymous referees and the editors for their valuable comments, which helped to improve the exposition of this paper.

Appendix

Proof of Theorem 3.1 In view of Remark 2.1 and Corollary 3.1, it suffices to show that for each $i \in \mathcal{M}$, $V(x, y, i)$ has limit 0 at any point $(x_0, y_0) \in \partial\Pi$. Using (6), we obtain $V \leq K_1 V_p$ where V_p denotes the value function for the power utility function

$U(c) = c^\gamma/\gamma$ where γ is specified in (6) and K_1 is some positive constant. In view of the upper bound property established in Proposition 5.3, we have

$$V(x,y,i) \le \begin{cases} K_2[x+(1-\mu)y]^\gamma, & \text{if } (x,y) \in \overline{D}_0, \\ K_3[x+(1+\lambda)y]^\gamma, & \text{if } (x,y) \in \overline{E}_0 \end{cases} \tag{82}$$

for some constants $K_2, K_3 > 0$, where D_0 and E_0 are defined in (32) and (43), respectively.

Given $(x_0, y_0) \in \partial_1 \Pi$ and $(x_0, y_0) \neq (0,0)$. Consider an δ-neighborhood of (x_0, y_0): $B_\delta(x_0, y_0) = \{(x,y) : |(x,y) - (x_0, y_0)| < \delta\}$ for some $\delta > 0$. Then for any point $(x,y) \in B_\delta(x_0,y_0) \cap \overline{D}_0$, we have $0 \le V(x,y,i) \le K_2[x+(1-\mu)y]^\gamma$. Sending $(x,y) \to (x_0, y_0)$, we obtain

$$\lim_{(x,y)\to(x_0,y_0)} V(x,y,i) = 0. \tag{83}$$

Similarly, we can show that (83) holds at any point $(x_0, y_0) \in \partial_2 \Pi \setminus \{(0,0)\}$.

Finally, we consider the origin $(0,0)$. For any point $(x,y) \in B_\delta(0,0) \cap \overline{\Pi}$ where $B_\delta(0,0) = \{(x,y) : |(x,y)| < \delta\}$ for some $\delta > 0$, we consider three cases:

(I) If $(x,y) \in \overline{D}_0$, then

$$0 \le V(x,y,i) \le K_2\big[x+(1-\mu)y\big]^\gamma; \tag{84}$$

(II) If $(x,y) \in \overline{E}_0$, then

$$0 \le V(x,y,i) \le K_3\big[x+(1+\lambda)y\big]^\gamma; \tag{85}$$

(III) If $(x,y) \in \Pi \setminus (\overline{D}_0 \cup \overline{E}_0)$, then in view of the homotheticity property (25) of the function V_p, we have

$$V(x,y,i) \le K_1 V_p(x,y,i) = K_1\big(\sqrt{x^2+y^2}\big)^\gamma V_p\left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}, i\right). \tag{86}$$

Note that $(x/\sqrt{x^2+y^2}, y/\sqrt{x^2+y^2})$ is on the arc segment of the unit circle $r = 1$, $-\arctan(1/\xi_0) \le \theta \le \pi - \arctan(1/\delta_0)$, which is a closed subset in Π. In view of Corollary 3.1, V_p is continuous on this arc segment and therefore bounded. It follows that

$$V(x,y,i) \le K_4\big(\sqrt{x^2+y^2}\big)^\gamma, \quad \forall (x,y) \in \Pi \setminus (\overline{D}_0 \cup \overline{E}_0), \tag{87}$$

for some constant $K_4 > 0$. Sending $(x,y) \to (0,0)$ in (84), (85), and (87), we obtain

$$\lim_{(x,y)\to(0,0)} V(x,y,i) = 0. \tag{88}$$

This completes the proof. □

Proof of the $\le$ part of Proposition 5.4 We only show the inequality for $x \le 0$ since that for $x \ge 0$ can be proved analogously and therefore omitted. For this purpose, we introduce a sequence of closed subsets in Π. For $n = 1, 2, \ldots,$ define

$$\Pi_n = \left\{ (x, y) : x + (1 - \mu)y \geq \frac{1}{n}, x + (1 + \lambda)y \geq \frac{1}{n} \right\} \subset \Pi. \tag{89}$$

Given $(x_0, y_0) \in \Pi$, $i \in \mathcal{M}$, and $(C, M, N) \in \mathcal{A}(x_0, y_0, i)$, define

$$\nu_n = \inf\{t \geq 0 : \big(x(t), y(t)\big) \notin \Pi_n\} \tag{90}$$

and

$$\nu = \inf\{t \geq 0 : \big(x(t), y(t)\big) = (0, 0)\} \tag{91}$$

where $(x(t), y(t))$ are given by (3). It can be shown as in [3, Lemma 5.6] that $\nu_n \uparrow \nu$ a.s. as $n \to \infty$.

We consider φ defined in (55). It is readily seen that $\varphi(x, y, i) = V(x, y, i)$, $\forall (x, y) \in \partial D$, $i \in \mathcal{M}$. We extend the definition of φ to $\overline{\Pi}$ by letting $\varphi = V$ on $\overline{\Pi} \setminus \overline{D}$.

Next, we introduce a sequence of closed subsets in $\overline{D}$. For $n = 1, 2, \ldots,$ define

$$D_n = \{(x, y) \in \overline{D} : y \leq n\} \subset \overline{D}. \tag{92}$$

Let $(x_0, y_0) \in D$. Define

$$\tau_n = \inf\{t \geq 0 : \big(x(t), y(t)\big) \notin D_n\} \tag{93}$$

and

$$\tau = \inf\{t \geq 0 : \big(x(t), y(t)\big) \notin \overline{D}\}. \tag{94}$$

Then $\tau_n \uparrow \tau$ a.s. as $n \to \infty$. Moreover, similar to [3, Proposition 5.7], we have

$$\{\tau < \infty\} = \bigcup_{n=1}^{\infty} \{\nu_n \wedge \tau_n = \tau \leq n\}. \tag{95}$$

Now consider $0 \leq t < n \wedge \nu_n \wedge \tau_n$. Note that (57) implies that for each $i \in \mathcal{M}$, the function φ_i is non-increasing in the jump directions $(1 - \mu, -1)$ and $(-(1 + \lambda), 1)$ at any point in D. Thus,

$$\varphi\big(x(t), y(t), i\big) \leq \varphi\big(x(t-), y(t-), i\big), \quad 0 \leq t < n \wedge \nu_n \wedge \tau_n, i \in \mathcal{M}. \tag{96}$$

Applying the Itô formula to $e^{-\beta t}\varphi(x(t), y(t), \alpha_t)$ for $0 \leq t \leq n \wedge \nu_n \wedge \tau_n$, noting that $\varphi(x(t), y(t), \alpha_t)$ and $y(t)\varphi_y(x(t), y(t), \alpha_t)$ are bounded on $[0, n \wedge \nu_n \wedge \tau_n[$, we obtain

$$\begin{aligned} &E\big[e^{-\beta(n\wedge\nu_n\wedge\tau_n)}\varphi\big(x(n \wedge \nu_n \wedge \tau_n), y(n \wedge \nu_n \wedge \tau_n), \alpha_{n\wedge\nu_n\wedge\tau_n}\big)\big] \\ &\quad \leq \varphi(x_0, y_0, i) - E\bigg[\int_0^{n\wedge\nu_n\wedge\tau_n} e^{-\beta t}\big[\mathcal{L}_{\alpha_t}\varphi\big(x(t), y(t), \alpha_t\big) \\ &\qquad + C(t)\varphi_x\big(x(t), y(t), \alpha_t\big) - Q\varphi\big(x(t), y(t), \cdot\big)(\alpha_t)\big]\,dt\bigg]. \end{aligned} \tag{97}$$

In view of (14) and (58), we have, for $t \in]0, n \wedge \nu_n \wedge \tau_n[$,

$$\mathcal{L}_{\alpha_t}\varphi\big(x(t), y(t), \alpha_t\big) + C(t)\varphi_x\big(x(t), y(t), \alpha_t\big) - Q\varphi\big(x(t), y(t), \cdot\big)(\alpha_t) \geq U\big(C(t)\big). \tag{98}$$

Using (98) and the fact that $\varphi = V$ on $\overline{\Pi} \setminus \overline{D}$, we get from (97)

$$\begin{aligned} \varphi(x_0, y_0, i) &\geq E\big[e^{-\beta(n\wedge\nu_n\wedge\tau_n)}\varphi\big(x(n\wedge\nu_n\wedge\tau_n), y(n\wedge\nu_n\wedge\tau_n), \alpha_{n\wedge\nu_n\wedge\tau_n}\big)\big] \\ &\quad + E\bigg[\int_0^{n\wedge\nu_n\wedge\tau_n} e^{-\beta t}U\big(C(t)\big)\,dt\bigg] \\ &\geq E\big[I_{\{\nu_n\wedge\tau_n=\tau\leq n\}}e^{-\beta\tau}\varphi\big(x(\tau), y(\tau), \alpha_\tau\big)\big] \\ &\quad + E\bigg[\int_0^{n\wedge\nu_n\wedge\tau_n} e^{-\beta t}U\big(C(t)\big)\,dt\bigg] \\ &= E\big[I_{\{\nu_n\wedge\tau_n=\tau\leq n\}}e^{-\beta\tau}V\big(x(\tau), y(\tau), \alpha_\tau\big)\big] \\ &\quad + E\bigg[\int_0^{n\wedge\nu_n\wedge\tau_n} e^{-\beta t}U\big(C(t)\big)\,dt\bigg]. \end{aligned} \tag{99}$$

Note that $n \wedge \nu_n \wedge \tau_n \to \nu \wedge \tau$ as $n \to \infty$. We have two cases. (I) If $\tau(\omega) < \infty$, then we must have $\tau(\omega) < \nu(\omega)$. Otherwise the state $(x(t,\omega), y(t,\omega))$ would reach the origin $(0,0)$ and stay there forever and hence never exit $\overline{D}$, resulting in $\tau(\omega) = \infty$, a contradiction! (II) If $\tau(\omega) = \infty$, then $\nu(\omega) \wedge \tau(\omega) = \nu(\omega)$. For this case, we have

$$\int_0^{\nu(\omega)} e^{-\beta t}U\big(C(t,\omega)\big)\,dt = \int_0^{\infty} e^{-\beta t}U\big(C(t,\omega)\big)\,dt = \int_0^{\tau(\omega)} e^{-\beta t}U\big(C(t,\omega)\big)\,dt,$$

where the first equality is due to the fact that $C(t,\omega) = 0$ for $t > \nu(\omega)$. It follows by using the monotone convergence theorem that

$$\lim_{n\to\infty} E\bigg[\int_0^{n\wedge\nu_n\wedge\tau_n} e^{-\beta t}U\big(C(t)\big)\,dt\bigg] = E\bigg[\int_0^{\tau} e^{-\beta t}U\big(C(t)\big)\,dt\bigg].$$

On the other hand, we can show that $I_{\{\nu_n\wedge\tau_n=\tau\leq n\}} \uparrow I_{\{\tau<\infty\}}$ as $n \to \infty$. Again by the monotone convergence theorem we have

$$\lim_{n\to\infty} E\big[I_{\{\nu_n\wedge\tau_n=\tau\leq n\}}e^{-\beta\tau}V\big(x(\tau), y(\tau), \alpha_\tau\big)\big] = E\big[I_{\{\tau<\infty\}}e^{-\beta\tau}V\big(x(\tau), y(\tau), \alpha_\tau\big)\big].$$

Therefore we obtain

$$\varphi(x_0, y_0, i) \geq E\big[I_{\{\tau<\infty\}}e^{-\beta\tau}V\big(x(\tau), y(\tau), \alpha_\tau\big)\big] + E\bigg[\int_0^{\tau} e^{-\beta t}U\big(C(t)\big)\,dt\bigg]. \tag{100}$$

Taking the supremum of the RHS of (100) over $(C, M, N) \in \mathcal{A}(x_0, y_0, i)$ and using (19), we obtain $\varphi(x_0, y_0, i) \geq V(x_0, y_0, i)$. □

Partition of the Region $\varPi$ by Convex Analysis

For each $(x,y)\in\varPi$ and each $i\in\mathcal{M}$, define the sub-differential of V by

$$\partial V(x,y,i)=\bigl\{\bigl(\delta_x^i,\delta_y^i\bigr)\in\mathbb{R}^2 : V(\tilde{x},\tilde{y},i)\le V(x,y,i)+\delta_x^i(\tilde{x}-x)+\delta_y^i(\tilde{y}-y),\\ \forall(\tilde{x},\tilde{y})\in\varPi\bigr\}. \tag{101}$$

Since $V_i(x,y)$ is concave and finite on $\varPi$, $\partial V(x,y,i)$ is a nonempty, compact and convex subset of $\mathbb{R}^2$. The function $V_i(x,y)$ is differentiable at the point $(x,y)\in\varPi$ iff $\partial V(x,y,i)$ is a singleton. In this case we have $\partial V(x,y,i)=\{V_x(x,y,i),V_y(x,y,i)\}$.

The following two results can be established by applying [3, Lemma 6.1 and Proposition 6.2] to each function V_i, $i\in\mathcal{M}$.

Lemma A.1 *Given $i\in\mathcal{M}$. Let $\{(x_n,y_n),n\ge 1\}$ be a sequence of points in $\varPi$ with limit $(x_0,y_0)\in\varPi$. If $(\delta_x^{i,n},\delta_y^{i,n})\in\partial V(x_n,y_n,i)$ for every $n\ge 1$, then the sequence $\{(\delta_x^{i,n},\delta_y^{i,n}),n\ge 1\}$ is bounded and its every limit point is in $\partial V(x_0,y_0,i)$.*

Proposition A.1 *Let O be an open subset of $\varPi$. Then V_i is of C^1 in O if and only if $\partial V(x,y,i)$ is a singleton for every $(x,y)\in O$.*

Given $i\in\mathcal{M}$, $(x,y)\in\varPi$, and $(\delta_x^i,\delta_y^i)\in\partial V(x,y,i)$. We define function $\phi(\tilde{x},\tilde{y},i)=V(x,y,i)+\delta_x^i(\tilde{x}-x)+\delta_y^i(\tilde{y}-y)$, $(\tilde{x},\tilde{y})\in\varPi$. Then $\phi(\tilde{x},\tilde{y},i)\ge V(\tilde{x},\tilde{y},i),\forall(\tilde{x},\tilde{y})\in\varPi$. For each $h>0$ satisfying $(x-(1-\mu)h,y+h)\in\varPi$ from which (x,y) can be reached by a transaction, by Proposition 3.3, we have

$$\phi(x,y,i)=V(x,y,i)\le V\bigl(x-(1-\mu)h,y+h,i\bigr)\\ \le\phi\bigl(x-(1-\mu)h,y+h,i\bigr)=\phi(x,y,i)+h\bigl[-(1-\mu)\delta_x^i+\delta_y^i\bigr].$$

It follows that

$$-(1-\mu)\delta_x^i+\delta_y^i\ge 0,\quad \forall\bigl(\delta_x^i,\delta_y^i\bigr)\in\partial V(x,y,i),\forall(x,y)\in\varPi. \tag{102}$$

Similarly, we can show

$$(1+\lambda)\delta_x^i-\delta_y^i\ge 0,\quad \forall\bigl(\delta_x^i,\delta_y^i\bigr)\in\partial V(x,y,i),\forall(x,y)\in\varPi. \tag{103}$$

Since for $(x,y)\in\varPi$, $x+(1-\mu)y>0$ and $x+(1+\lambda)y>0$, we have $V(x,y,i)>0$ in view of Proposition 5.2. It follows from Corollary 5.1 that $\delta_x^i=\phi_x(x,y)>0$, and then from (102) that $\delta_y^i>0$. To summarize, we have

$$\delta_x^i>0,\qquad \delta_y^i>0,\quad \forall\bigl(\delta_x^i,\delta_y^i\bigr)\in\partial V(x,y,i),\forall(x,y)\in\varPi. \tag{104}$$

For θ sufficiently close to 1, using the homotheticity property (25), we have

$$\bigl(\theta^\gamma-1\bigr)V(x,y,i)=V(\theta x,\theta y,i)-V(x,y,i)\\ \le\phi(\theta x,\theta y,i)-V(x,y,i)=(\theta-1)\bigl(x\delta_x^i+y\delta_y^i\bigr). \tag{105}$$

Dividing both sides of (105) by $\theta - 1$, sending $\theta \to 1+$ and $\theta \to 1-$, respectively, we obtain

$$x\delta_x^i + y\delta_y^i = \gamma V(x, y, i), \quad \forall(\delta_x^i, \delta_y^i) \in \partial V(x, y, i), \forall(x, y) \in \Pi. \tag{106}$$

Now, for $(\tilde{x}, \tilde{y}) \in \Pi$, $\theta > 0$, we have

$$\begin{aligned} V(\tilde{x}, \tilde{y}, i) &= \theta^\gamma V\left(\frac{\tilde{x}}{\theta}, \frac{\tilde{y}}{\theta}, i\right) \le \theta^\gamma \phi\left(\frac{\tilde{x}}{\theta}, \frac{\tilde{y}}{\theta}, i\right) \\ &= \theta^\gamma \left[V(x, y, i) + \delta_x^i\left(\frac{\tilde{x}}{\theta} - x\right) + \delta_y^i\left(\frac{\tilde{y}}{\theta} - y\right)\right] \\ &= V(\theta x, \theta y, i) + \left(\theta^{\gamma-1}\delta_x^i\right)(\tilde{x} - \theta x) + \left(\theta^{\gamma-1}\delta_y^i\right)(\tilde{y} - \theta y). \end{aligned}$$

It follows that $(\theta^{\gamma-1}\delta_x^i, \theta^{\gamma-1}\delta_y^i) \in \partial V(\theta x, \theta y, i)$. Thus

$$\theta^{\gamma-1}\partial V(x, y, i) \subset \partial V(\theta x, \theta y, i).$$

The reverse set containment can be established by simply replacing θ by $\frac{1}{\theta}$, x by θx and y by θy. Therefore we have

$$\theta^{\gamma-1}\partial V(x, y, i) = \partial V(\theta x, \theta y, i), \quad \forall(x, y) \in \Pi, \forall\theta > 0, \forall i \in \mathcal{M}. \tag{107}$$

Using the sub-differentials, we can partition the solvency region Π into three convex cones. First, for $i \in \mathcal{M}$, $(x, y) \in \Pi$, $(\bar{x}, \bar{y}) \in \Pi$, $(\delta_x^i, \delta_y^i) \in \partial V(x, y, i)$ and $(\delta_{\bar{x}}^i, \delta_{\bar{y}}^i) \in \partial V(\bar{x}, \bar{y}, i)$, in view of (101), we have

$$\begin{aligned} V(\bar{x}, \bar{y}, i) &\le V(x, y, i) + \delta_x^i(\bar{x} - x) + \delta_y^i(\bar{y} - y) \\ &\le V(\bar{x}, \bar{y}, i) + \delta_{\bar{x}}^i(x - \bar{x}) + \delta_{\bar{y}}^i(y - \bar{y}) + \delta_x^i(\bar{x} - x) + \delta_y^i(\bar{y} - y). \end{aligned}$$

It follows that

$$\left(\delta_x^i - \delta_{\bar{x}}^i\right)(x - \bar{x}) + \left(\delta_y^i - \delta_{\bar{y}}^i\right)(y - \bar{y}) \le 0. \tag{108}$$

For $i \in \mathcal{M}$, $(x, y) \in \Pi$, define

$$\begin{aligned} \vartheta^+(x, y, i) &= \max\left\{-(1-\mu)\delta_x^i + \delta_y^i : \left(\delta_x^i, \delta_y^i\right) \in \partial V(x, y, i)\right\}, \\ \vartheta^-(x, y, i) &= \min\left\{-(1-\mu)\delta_x^i + \delta_y^i : \left(\delta_x^i, \delta_y^i\right) \in \partial V(x, y, i)\right\}. \end{aligned} \tag{109}$$

Note that the maxima and minima in (109) are attained since $\partial V(x, y, i)$ is compact. In addition, $\vartheta^+(x, y, i) \ge \vartheta^-(x, y, i) \ge 0$ due to (102).

Consider the parametric equations of the half line L_1 originating at the point $(1 + \lambda, -1)$ on $\partial_2\Pi$ and parallel to $\partial_1\Pi$, defined by

$$L_1 : \begin{cases} x(\rho) = 1 + \lambda - (1-\mu)\rho, \\ y(\rho) = -1 + \rho, \end{cases} \quad \forall\rho \ge 0. \tag{110}$$

Define

$$\rho_0^i = \inf\{\rho > 0 : \vartheta^-\big(x(\rho), y(\rho), i\big) = 0\}, \tag{111}$$

where $\rho_0^i = \infty$ if the above set is empty.

Lemma A.2 *Given $i \in \mathcal{M}$. For $0 < \rho < \bar{\rho} < \infty$,*

$$\vartheta^+\big(x(\bar{\rho}), y(\bar{\rho}), i\big) \le \vartheta^-\big(x(\rho), y(\rho), i\big). \tag{112}$$

If $0 < \rho_0^i < \infty$, then $\vartheta^-(x(\rho_0^i), y(\rho_0^i), i) = 0$ and

$$\vartheta^+\big(x(\rho), y(\rho), i\big) = 0, \quad \forall \rho > \rho_0^i. \tag{113}$$

Let $i \in \mathcal{M}$ be fixed. We partition Π into two open, convex (possibly empty) cones as follows:

$$SA^i = \big\{(x, y) \in \Pi : (\theta x, \theta y) = \big(x(\rho), y(\rho)\big) \text{ for some } \theta > 0 \text{ and some } \rho \text{ with } \rho_0^i < \rho < \infty\big\},$$

$$\Pi \setminus \overline{SA^i} = \big\{(x, y) \in \Pi : (\theta x, \theta y) = \big(x(\rho), y(\rho)\big) \text{ for some } \theta > 0 \text{ and some } \rho \text{ with } 0 < \rho < \rho_0^i\big\}.$$

Proposition A.2 *Given $i \in \mathcal{M}$. We have*

$$-(1-\mu)\delta_x^i + \delta_y^i = 0, \quad \forall\big(\delta_x^i, \delta_y^i\big) \in \partial V(x, y, i), \forall (x, y) \in SA^i. \tag{114}$$

In a similar way, we introduce the parametric equations of the half line L_2 originating at the point $(-(1-\mu), 1)$ on $\partial_1 \Pi$ and parallel to $\partial_2 \Pi$:

$$L_2 : \begin{cases} \tilde{x}(\rho) = -(1-\mu) + (1+\lambda)\rho, \\ \tilde{y}(\rho) = 1 - \rho, \end{cases} \quad \forall \rho \ge 0. \tag{115}$$

Define

$$\tilde{\rho}_0^i = \inf\big\{\rho > 0 : (1+\lambda)\delta_x^i - \delta_y^i = 0 \text{ for some } \big(\delta_x^i, \delta_y^i\big) \in \partial V\big(\tilde{x}(\rho), \tilde{y}(\rho), i\big)\big\}. \tag{116}$$

Let $i \in \mathcal{M}$ be fixed. We partition Π into two open, convex (possibly empty) cones as follows:

$$BU^i = \big\{(x, y) \in \Pi : (\theta x, \theta y) = \big(\tilde{x}(\rho), \tilde{y}(\rho)\big) \text{ for some } \theta > 0 \text{ and some } \rho \text{ with } \tilde{\rho}_0^i < \rho < \infty\big\},$$

$$\Pi \setminus \overline{BU^i} = \big\{(x, y) \in \Pi : (\theta x, \theta y) = \big(\tilde{x}(\rho), \tilde{y}(\rho)\big) \text{ for some } \theta > 0 \text{ and some } \rho \text{ with } 0 < \rho < \tilde{\rho}_0^i\big\}.$$

Proposition A.3 *Given $i \in \mathcal{M}$. We have*

$$(1+\lambda)\delta_x^i - \delta_y^i = 0, \quad \forall\big(\delta_x^i, \delta_y^i\big) \in \partial V(x, y, i), \forall(x, y) \in BU^i. \tag{117}$$

Corollary A.1 $SA^i \cap BU^i = \emptyset, \forall i \in \mathcal{M}$.

Corollary A.2 *For each $i \in \mathcal{M}$, the value function V_i is C^1 in $SA^i \cup BU^i$.*

Note that Corollary A.2 implies that $-(1-\mu)V_x(x, y, i) + V_y(x, y, i) = 0$ for $(x, y) \in SA^i$ and $(1+\lambda)V_x(x, y, i) - V_y(x, y, i) = 0$ for $(x, y) \in BU^i$. In view of Proposition 5.4, we conclude that both SA^i and BU^i are nonempty, since the two equations are satisfied by V in D and E, respectively, where D and E are defined by (54) and (60), respectively. Clearly, $SA^i \supset D$ and $BU^i \supset E$. Moreover, it follows from these equations, the homotheticity of the value function, and the functions presented in Proposition 5.4 that

$$V(x, y, i) = \begin{cases} \frac{1}{\gamma} A_i^{\gamma-1}[x + (1-\mu)y]^\gamma, & \forall(x, y) \in SA^i, \\ \frac{1}{\gamma} B_i^{\gamma-1}[x + (1+\lambda)y]^\gamma, & \forall(x, y) \in BU^i, \end{cases} \tag{118}$$

where A_i, B_i are given by (56) and (62), respectively.

For each $i \in \mathcal{M}$, let $NT^i = \Pi \setminus \overline{SA^i \cup BU^i}$. Then we have

Proposition A.4 *Given $i \in \mathcal{M}$. we have*

$$\begin{cases} -(1-\mu)\delta_x^i + \delta_y^i > 0, \\ (1+\lambda)\delta_x^i - \delta_y^i > 0, \end{cases} \quad \forall\big(\delta_x^i, \delta_y^i\big) \in \partial V(x, y, i), \forall(x, y) \in NT^i. \tag{119}$$

References

1. Magill, M.J.P., Constantinides, G.M.: Portfolio selection with transaction costs. J. Econ. Theor. **13**, 245–263 (1976)
2. Davis, M.H.A., Norman, A.R.: Portfolio selection with transaction costs. Math. Oper. Res. **15**, 676–713 (1990)
3. Shreve, S.E., Soner, H.M.: Optimal investment and consumption with transaction costs. Ann. Appl. Probab. **4**, 609–692 (1994)
4. Zariphopoulou, T.: Investment-consumption models with transaction fees and Markov-chain parameters. SIAM J. Control Optim. **30**, 613–636 (1992)
5. Crandall, M.G., Ishii, H., Lions, P.L.: User's guide to viscosity solutions of second order partial differential equations. Bull. Am. Math. Soc. **27**, 1–67 (1992)
6. Rishel, R.: Whether to sell or hold a stock. Commun. Inf. Syst. **6**, 193–202 (2006)
7. Dai, M., Zhang, Q., Zhu, Q.J.: Trend following trading under a regime switching model. SIAM J. Financ. Math. **1**, 780–810 (2010)