J. Differential Equations 246 (2009) 1445–1469

Contents lists available at ScienceDirect

Journal of Differential Equations

www.elsevier.com/locate/jde

Finite-horizon optimal investment with transaction costs: A parabolic double obstacle problem ☆

Min Dai [a,*], Fahuai Yi [b]

[a] Dept. of Math., National University of Singapore, Singapore
[b] Dept. of Math., South China Normal University, Guangzhou 510631, China

ARTICLE INFO

Article history:
Received 24 May 2007
Revised 1 October 2008
Available online 10 December 2008

Keywords:
Optimal investment
Transaction costs
Finite horizon
Double obstacle problem
Free boundary
Singular stochastic control
Portfolio selection

ABSTRACT

This paper concerns optimal investment problem of a CRRA investor who faces proportional transaction costs and finite time horizon. From the angle of stochastic control, it is a singular control problem, whose value function is governed by a time-dependent HJB equation with gradient constraints. We reveal that the problem is equivalent to a parabolic double obstacle problem involving two free boundaries that correspond to the optimal buying and selling policies. This enables us to make use of the well-developed theory of obstacle problem to attack the problem. The $C^{2,1}$ regularity of the value function is proven and the behaviors of the free boundaries are completely characterized.

1. Introduction

Merton [19] pioneered in applying continuous-time stochastic models to the study of financial markets. In the absence of transaction costs, he showed that an optimal investment problem can be formulated as a Hamilton–Jacobi–Bellman (HJB) equation that allows an explicit solution for a constant relative risk aversion (CRRA) investor. The corresponding optimal investment policy is to keep a constant fraction of total wealth in each asset during the whole investment period. To implement the policy, the investor would have to indulge in incessant trading which is completely unrealistic in

☆ This project is partially supported by Singapore MOE AcRF grant (No. R-146-000-096-112) and NUS RMI grant (No. R-703-000-004-720/646), NNSFs of China (No. 10671075 and No. 10301002), NSF of Guangdong Province (No. 5005930) and University Special Research Fund for PhD Program (20060574002). The authors thank Lishang Jiang, Zuoquan Xu, Xunyu Zhou and seminar participants at Hong Kong University of Science and Technology, Chinese University of Hong Kong, Jilin University and Quantitative Methods in Finance Conference 2005 (Sydney) for helpful discussion and comments. All errors are our own.

* Corresponding author. Fax: +65 67795452.
E-mail addresses: matdm@nus.edu.sg (M. Dai), fhyi@scnu.edu.cn (F. Yi).

0022-0396/$ – see front matter
doi:10.1016/j.jde.2008.11.003

the face of transaction costs and violates the conventional largely buy-and-hold investment strategy as well.

To overcome the shortages, Magill and Constantinides [17] introduced proportional transaction costs to Merton's model. They provided a fundamental insight that there is a no-trading region in the presence of transaction costs and the no-trading region must be a wedge. But, their argument is heuristic at best. In terms of a rigorous mathematical analysis, Davis and Norman [6] showed that for an infinite-horizon investment and consumption with transaction costs, the optimal policies are determined by the solution of a free boundary problem, where the free boundaries correspond to the optimal buying and selling policies. Relying on the concept of viscosity solutions to HJB equations, Shreve and Soner [21] fully characterized the infinite-horizon optimal policies. Using a martingale approach, Cvitanic and Karatzas [4] proved the existence of an optimal solution to the portfolio optimization problem with transaction costs. Other existence results can be found in Bouchard [3], Guasoni [12] and Guasoni and Schachermayer [13]. Akian, Menaldi and Sulem [1] and Kabanov and Kluppelberg [14] considered a multi-asset investment–consumption model with transaction costs.

This paper concerns the finite-horizon optimal investment with transaction costs, and aims to provide a theoretical analysis of the optimal investment policies. From the angle of stochastic control, it is a singular control problem for the displacement of the state variables due to transaction costs might not be continuous. It can be shown that the value function is governed by a time-dependent HJB equation involving gradient constraints which leads to two time-dependent free boundaries. Due to the convexity of utility functions, the value function is expected to be twice continuously differentiable in the spatial direction. Since such regularity of the value function generally leads to free boundary conditions, this regularity property, first observed by Benes et al. [2], is called the *principle of smooth fit* and plays a critical role in the study of singular control problems. For an infinite-horizon problem, Davis and Norman [6] and Shreve and Soner [21] proved the regularity property using an ordinary differential equation approach and the viscosity solution approach, respectively. However, for a finite-horizon problem, the resulting free boundaries change through time such that it is hard to follow their approaches.

One paper related to the present work is Liu and Loewenstein [16] where the authors adopted an indirect method. They first considered an optimal investment problem with a stochastic time horizon following Erlang distribution. For this tractable problem, they derived some analytical properties on the optimal investment policies. They then extended those results to the situation of a deterministic time horizon using the fact that the optimal investment policies of the Erlang distributed case converge to those of the deterministic time case. They obtained some interesting characterization of optimal investment policies (i.e. free boundaries). However, their results are incomplete and some are not sharp. In addition, their approach cannot be extended to including the consumption term.

We will attack the problem directly by virtue of a partial differential equation (PDE) approach. Our key idea is to establish a link between the singular control problem and the obstacle problem. More precisely, we will show that the spatial partial derivative of the value function is the solution to a double obstacle problem. Since the solution to an obstacle problem is once continuously differentiable in the spatial direction, the above link will immediately yield the desired *principle of smooth fit*. The link also enables us to make use of the well-developed theory of obstacle problem to study the present problem and the behaviors of the resulting free boundaries can then be completely characterized.

The rest of this paper is arranged as follows. In the next section, we present the model formulation. In Section 3, we formally derive the parabolic double obstacle problem regarding the spatial partial derivative of the value function, and study the existence and regularity of solution. Section 4 is devoted to the analysis of the behaviors of the free boundaries (i.e. optimal investment policies). The equivalence between the double obstacle problem and the original problem is proven in Section 5. To examine the asymptotic behaviors of the free boundaries as time to maturity goes to infinity, we study the stationary solution to the double obstacle problem in Section 6. The paper ends with conclusive remark in Section 7.

2. Formulation of the model

We take into account the optimal investment problem with transaction costs and a finite horizon T. Except for notational changes, the model formulation is that of Liu and Loewenstein [16]. Most notations are from Davis and Norman [6] and Shreve and Soner [21].

2.1. The asset market

Suppose that there are only two assets available for investment: a riskless asset (bank account) and a risky asset (stock). Their prices, denoted by $P_0(t)$ and $P_1(t)$, respectively, evolve according to the following equations:

$$dP_{0t} = rP_{0t}\,dt,$$

$$dP_{1t} = P_{1t}[\alpha\,dt + \sigma\,d\mathcal{B}_t],$$

where $r > 0$ is the constant riskless rate, $\alpha > r$ and $\sigma > 0$ are constants called the expected rate of return and the volatility, respectively, of the stock. The process $\{\mathcal{B}_t;\ t > 0\}$ is a standard Brownian motion on a filtered probability space $(S, \mathscr{F}, \{\mathscr{F}_t\}_{t\geqslant 0}, P)$ with $\mathcal{B}_0 = 0$ almost surely. We assume $\mathscr{F} = \mathscr{F}_\infty$, the filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ is right-continuous and each $\mathscr{F}_t$ contains all null sets of $\mathscr{F}_\infty$.

Assume that a CRRA investor holds X_t and Y_t in bank and stock respectively, expressed in monetary terms. In the presence of transaction costs, the equations describing their evolution are

$$dX_t = rX_t\,dt - (1+\lambda)\,dL_t + (1-\mu)\,dM_t, \tag{2.1}$$

$$dY_t = \alpha Y_t\,dt + \sigma Y_t\,d\mathcal{B}_t + dL_t - dM_t, \tag{2.2}$$

where L_t and M_t are right-continuous (with left-hand limits), nonnegative, and nondecreasing $\{\mathscr{F}_t\}_{t\geqslant 0}$-adapted processes with $L_0 = M_0 = 0$, representing cumulative dollar values for the purpose of buying and selling stock respectively. The constants $\lambda \in [0, \infty)$ and $\mu \in [0, 1)$ appearing in these equations account for proportional transaction costs incurred on purchase and sale of stock respectively.

2.2. The investor's problem

Due to transaction costs, the investor's net wealth in monetary terms at time t is

$$W_t = \begin{cases} X_t + (1-\mu)Y_t & \text{if } Y_t \geqslant 0, \\ X_t + (1+\lambda)Y_t & \text{if } Y_t < 0. \end{cases}$$

Since it is required that the investor's net wealth be positive, following Davis and Norman [6], we define the solvency region

$$\mathscr{S} = \big\{(x, y) \in R^2:\ x + (1+\lambda)y > 0,\ x + (1-\mu)y > 0\big\}.$$

Assume that the investor is given an initial position in $\mathscr{S}$. An investment strategy (L, M) is admissible for (x, y) starting from $s \in [0, T)$ if (X_t, Y_t) given by (2.1)–(2.2) with $X_s = x$ and $Y_s = y$ is in $\mathscr{S}$ for all $t \in [s, T]$. We let $A_s(x, y)$ be the set of admissible investment strategies.

The investor's problem is to choose an admissible strategy so as to maximize the expected utility of terminal wealth, that is,

$$\sup_{(L,M)\in A_0(x,y)} E_0^{x,y}\big[U(W_T)\big]$$

subject to (2.1)–(2.2). Here $E_t^{x,y}$ denotes the conditional expectation at time t given that initial endowment $X_t = x$, $Y_t = y$, and the utility function

$$U(W) = \begin{cases} \frac{W^\gamma}{\gamma} & \text{if } \gamma < 1,\ \gamma \neq 0, \\ \log W & \text{if } \gamma = 0. \end{cases}$$

We define the value function by

$$\varphi(x, y, t) = \sup_{(L,M)\in A_t(x,y)} E_t^{x,y}\big[U(W_T)\big], \quad (x, y) \in \mathscr{S},\ t \in [0, T).$$

2.3. Merton's result with no transaction costs

If $\lambda = \mu = 0$, the problem reduces to the case of no transaction costs in which the wealth process W_t can be chosen as state variable, and Y_t be control variable (cf. [19]). Consequently, the control problem becomes a classical one and allows an explicit expression of the value function:

$$\varphi(x, y, t) = \begin{cases} e^{\gamma(r+\frac{(\alpha-r)^2}{2\sigma^2(1-\gamma)})(T-t)} \frac{(x+y)^\gamma}{\gamma} & \text{if } \gamma < 1,\ \gamma \neq 0, \\ (r + \frac{(\alpha-r)^2}{2\sigma^2})(T - t) + \log(x + y) & \text{if } \gamma = 0. \end{cases}$$

Correspondingly, the optimal policy is to keep a constant proportion of wealth in the bank account and the stock, namely,

$$\frac{x}{y} = -\frac{\alpha - r - (1 - \gamma)\sigma^2}{\alpha - r} \triangleq x_M. \tag{2.3}$$

Here x_M is called "Merton line."

2.4. HJB equation

Throughout the rest of this paper, we take into account the case of transaction costs, namely, $\lambda + \mu > 0$. The problem is indeed a singular control problem for the displacement of the state variables (X_t, Y_t) due to control effort might be discontinuous. The value function is shown to be the viscosity solution to the following HJB equation (cf. [9] or [21])

$$\min\big\{-\varphi_t - \mathscr{L}\varphi, -(1 - \mu)\varphi_x + \varphi_y, (1 + \lambda)\varphi_x - \varphi_y\big\} = 0, \quad (x, y) \in \mathscr{S},\ t \in [0, T), \tag{2.4}$$

with the terminal condition

$$\varphi(x, y, T) = \begin{cases} U(x + (1 - \mu)y) & \text{if } y > 0, \\ U(x + (1 + \lambda)y) & \text{if } y \leqslant 0, \end{cases} \tag{2.5}$$

where

$$\mathscr{L}\varphi = \frac{1}{2}\sigma^2 y^2\varphi_{yy} + \alpha y\varphi_y + rx\varphi_x.$$

The uniqueness of solution of problem (2.4)–(2.5) is proven by Davis et al. [7] in the sense of viscosity solution.

Due to the homotheticity of the utility function, we have for any positive constant ρ,

$$\varphi(\rho x, \rho y, t) = \begin{cases} \rho^\gamma\varphi(x, y, t) & \text{if } \gamma < 1,\ \gamma \neq 0, \\ \varphi(x, y, t) + \log\rho & \text{if } \gamma = 0. \end{cases} \tag{2.6}$$

It is well known that under the assumption $\alpha > r$, short selling is always suboptimal. Hence, we only need to consider $y > 0$. Denote

$$V(x,t) = \varphi(x,1,t).$$

Then, it follows from (2.6) that for $y > 0$,

$$\varphi(x,y,t) = \begin{cases} y^\gamma V(\frac{x}{y},t) & \text{if } \gamma < 1,\ \gamma \neq 0, \\ V(\frac{x}{y},t) + \log y & \text{if } \gamma = 0. \end{cases}$$

Accordingly, (2.4)–(2.5) are reduced to

$$\begin{cases} \min\{-V_t - \mathcal{L}_1 V, -(x+1-\mu)V_x + \gamma V, (x+1+\lambda)V_x - \gamma V\} = 0 \quad \text{in } \Omega, \\ V(x,T) = \dfrac{1}{\gamma}(x+1-\mu)^\gamma, \end{cases} \quad \text{if } \gamma < 1,\ \gamma \neq 0, \tag{2.7}$$

or

$$\begin{cases} \min\{-V_t - \mathcal{L}_2 V, -(x+1-\mu)V_x + 1, (x+1+\lambda)V_x - 1\} = 0 \quad \text{in } \Omega, \\ V(x,T) = \log(x+1-\mu), \end{cases} \quad \text{if } \gamma = 0, \tag{2.8}$$

where $\Omega = (-(1-\mu), +\infty) \times [0,T)$,

$$\mathcal{L}_1 V = \frac{1}{2}\sigma^2 x^2 V_{xx} + \beta_2 x V_x + \beta_1 V$$

with $\beta_1 = \gamma(\alpha - \frac{1}{2}\sigma^2(1-\gamma))$ and $\beta_2 = -(\alpha - r - \sigma^2(1-\gamma))$, and

$$\mathcal{L}_2 V = \frac{1}{2}\sigma^2 x^2 V_{xx} - (\alpha - r - \sigma^2) x V_x + \alpha - \frac{1}{2}\sigma^2.$$

In the following, we will concentrate on the problem (2.7) and the problem (2.8).

3. A parabolic double obstacle problem

In this section, we will formally derive and study the parabolic double obstacle problem regarding the spatial partial derivative of the value function.

3.1. Derivation

Let us first take into account the case of $\gamma \neq 0$, $\gamma < 1$. Eq. (2.7) can be rewritten as

$$\begin{cases} -V_t - \mathcal{L}_1 V = 0 \quad \text{if } \dfrac{1}{x+1+\lambda} < \dfrac{V_x}{\gamma V} < \dfrac{1}{x+1-\mu}, \\ -V_t - \mathcal{L}_1 V \geqslant 0 \quad \text{if } \dfrac{V_x}{\gamma V} = \dfrac{1}{x+1+\lambda} \text{ or } \dfrac{V_x}{\gamma V} = \dfrac{1}{x+1-\mu}, \\ V(x,T) = \dfrac{1}{\gamma}(x+1-\mu)^\gamma. \end{cases} \tag{3.1}$$

Consider the transformation

$$w(x,t) = \frac{1}{\gamma}\log(\gamma V).$$

Apparently $w_x = \frac{V_x}{\gamma V}$. It is easy to see that $w(x,t)$ satisfies

$$\begin{cases} -w_t - \mathcal{L}_3 w = 0 & \text{if } \dfrac{1}{x+1+\lambda} < w_x < \dfrac{1}{x+1-\mu}, \\ -w_t - \mathcal{L}_3 w \geqslant 0 & \text{if } w_x = \dfrac{1}{x+1+\lambda} \text{ or } w_x = \dfrac{1}{x+1-\mu}, \\ w(x,T) = \log(x+1-\mu), \end{cases} \tag{3.2}$$

where

$$\begin{aligned} \mathcal{L}_3 w &= \frac{1}{2}\sigma^2 x^2\big(w_{xx} + \gamma w_x^2\big) + \beta_2 x w_x + \frac{1}{\gamma}\beta_1 \\ &= \frac{1}{2}\sigma^2 x^2\big(w_{xx} + \gamma w_x^2\big) - \big(\alpha - r - \sigma^2(1-\gamma)\big)x w_x + \alpha - \frac{1}{2}\sigma^2(1-\gamma). \end{aligned}$$

It is worth pointing out that (3.2) reduces to (2.8) when $\gamma = 0$. Hence the following arguments cover the case of $\gamma = 0$.

Set

$$v(x,t) = w_x(x,t).$$

Formally we have

$$\begin{aligned} \frac{\partial}{\partial x}(\mathcal{L}_3 w) &= \frac{1}{2}\sigma^2 x^2 (w_{xxx} + 2\gamma w_x w_{xx}) + \sigma^2 x\big(w_{xx} + \gamma w_x^2\big) + \beta_2 x w_{xx} + \beta_2 w_x \\ &= \frac{1}{2}\sigma^2 x^2 v_{xx} - \big(\alpha - r - (2-\gamma)\sigma^2\big)x v_x - \big(\alpha - r - (1-\gamma)\sigma^2\big)v + \gamma\sigma^2\big(x^2 v v_x + x v^2\big) \\ &\triangleq \mathcal{L}v. \end{aligned} \tag{3.3}$$

We then postulate that v satisfies the following parabolic double obstacle problem:

$$\begin{cases} -v_t - \mathcal{L}v = 0 & \text{if } \dfrac{1}{x+1+\lambda} < v < \dfrac{1}{x+1-\mu}, \\ -v_t - \mathcal{L}v \geqslant 0 & \text{if } v = \dfrac{1}{x+1+\lambda}, \\ -v_t - \mathcal{L}v \leqslant 0 & \text{if } v = \dfrac{1}{x+1-\mu}, \\ v(x,T) = \dfrac{1}{x+1-\mu} \end{cases} \tag{3.4}$$

in Ω. Here $\frac{1}{x+1+\lambda}$ and $\frac{1}{x+1-\mu}$ correspond to lower and upper obstacles, respectively. We stress that $-v_t - \mathcal{L}v \geqslant 0$ on the lower obstacle and $-v_t - \mathcal{L}v \leqslant 0$ on the upper obstacle, which has a clear physical interpretation (cf. [11]).

Remark 3.1. It is well known that an optimal stopping time problem can be described as an obstacle problem (variational inequality problem) (cf. [20]). To some extent, our approach is identical to finding a connection between the singular control problem and the optimal stopping time problem pursued by Karatzas and Shreve [15].

The proof of the equivalence between the double obstacle problem (3.4) and the original problem (3.2) is deferred to Section 5. Let us first study the double obstacle problem (3.4). For later use, we define

$$\mathbf{SR} = \left\{(x,t) \in \Omega\colon\ v(x,t) = \frac{1}{x+1-\mu}\right\},$$

$$\mathbf{BR} = \left\{(x,t) \in \Omega\colon\ v(x,t) = \frac{1}{x+1+\lambda}\right\},$$

$$\mathbf{NT} = \left\{(x,t) \in \Omega\colon\ \frac{1}{x+1+\lambda} < v(x,t) < \frac{1}{x+1-\mu}\right\}.$$

In finance, the three regions defined above stand for the selling region, buying region and no transaction region, respectively.

3.2. Existence and regularity of solution to problem (3.4)

We aim to prove the existence and regularity of solution to the double obstacle problem. One technical difficulty is that the upper obstacle is infinite on the boundary $x = -(1-\mu)$. To avoid the singularity, we confine ourselves within $\widetilde{\Omega} = \{x > x^*,\ 0 < t < T\}$, where $x^* > -(1-\mu)$ is sufficiently close to $-(1-\mu)$, and the following boundary condition will be imposed on $x = x^*$:

$$v(x^*,t) = \frac{1}{x^*+1-\mu}, \quad t \in [0,T). \tag{3.5}$$

Later we will see that (3.5) is indeed true because $\{x \leqslant x^*\}$ is contained in **SR** when $x^* \leqslant x_{s,\infty}$ defined in Section 4.

Proposition 3.2. *The double obstacle problem* (3.4) *has a unique solution* $v(x,t) \in W_p^{2,1}(\widetilde{\Omega}_N \setminus \{|x| < \delta\})$*, for any* $\delta > 0$, $1 < p < +\infty$*, where* $\widetilde{\Omega}_N$ *is any bounded set in* $\widetilde{\Omega}$*. Moreover,*

$$v(x,t) \in C^\infty(\mathbf{NT}), \tag{3.6}$$

$$v_t(x,t) \geqslant 0, \tag{3.7}$$

and

$$v(0,t) = \frac{1}{1-\mu} \quad \text{when } \alpha - r - (1-\gamma)\sigma^2 \leqslant 0, \tag{3.8}$$

$$v(0,t) = \begin{cases} e^{-(\alpha-r-(1-\gamma)\sigma^2)(T-t)}\frac{1}{1-\mu} & \text{for } t_1 < t \leqslant T, \\ \frac{1}{1+\lambda} & \text{for } 0 \leqslant t \leqslant t_1, \end{cases} \quad \text{when } \alpha - r - (1-\gamma)\sigma^2 > 0, \tag{3.9}$$

where

$$t_1 = T - \frac{1}{\alpha - r - (1-\gamma)\sigma^2}\log\frac{1+\lambda}{1-\mu}. \tag{3.10}$$

The main difficulty of the proof lies in the degeneracy of the operator $\mathcal{L}$ at $x=0$. Before providing a proof, we would like to give its sketch. Thanks to the Fichera criterion (cf. [8]), we are able to deal with the problem in $\{x^*<x<0\}$ and in $\{x>0\}$ independently in order to avoid the degeneracy of the operator $\mathcal{L}$ at $x=0$, and no boundary value is required on $x=0$. The standard penalty method can be adopted to show the $W_p^{2,1}$ regularity of solution (cf. [11]), and (3.7) can be deduced from maximum principle. Regarding (3.6), we only need to show the smoothness on $\{x=0\}\cap\mathbf{NT}$. For $\gamma=0$ in which case the operator $\mathcal{L}$ is linear, it can be directly obtained by the hypoellipticity of the operator $\mathcal{L}$ (cf. [18]). In the case of $\gamma<1$ and $\gamma\neq 0$, a careful analysis will be made due to the nonlinearity of the operator $\mathcal{L}$.

3.3. *The proof of Proposition 3.2*

We will only confine our attention to $\{x^*<x<0\}$, and the case of $\{x^*<x<0\}$ is similar. By transformation $x=-e^y$ and $u(y,t)=v(x,t)$, (3.4) and (3.5) become

$$\begin{cases} -u_t-\mathcal{L}_y u=0 \quad \text{if } \dfrac{1}{-e^y+1+\lambda}<u<\dfrac{1}{-e^y+1-\mu}, \\ -u_t-\mathcal{L}_y u\geqslant 0 \quad \text{if } u=\dfrac{1}{-e^y+1+\lambda}, \\ -u_t-\mathcal{L}_y u\leqslant 0 \quad \text{if } u=\dfrac{1}{-e^y+1-\mu}, & \begin{array}{l} -\infty<y<y^*, \\ 0\leqslant t<T, \end{array} \\ u(y,T)=\dfrac{1}{-e^y+1-\mu}, \\ u(y^*,t)=\dfrac{1}{-e^{y^*}+1-\mu}, \end{cases} \tag{3.11}$$

where $y^*=\log(-x^*)$, and

$$\mathcal{L}_y u=\frac{1}{2}\sigma^2 u_{yy}-\left(\alpha-r-\left(\frac{3}{2}-\gamma\right)\sigma^2\right)u_y-\left(\alpha-r-(1-\gamma)\sigma^2\right)u-\gamma\sigma^2 e^y u(u_y+u).$$

First we prove the comparison principle for the problem (3.11).

Lemma 3.3. *Let u_i, $i=1,2$, satisfy*

$$\begin{cases} -u_{it}-\mathcal{L}_y u_i=0 \quad \textit{if } \dfrac{1}{-e^y+1+\lambda}<u_i<\dfrac{1}{-e^y+1-\mu}, \\ -u_{it}-\mathcal{L}_y u_i\geqslant 0 \quad \textit{if } u_i=\dfrac{1}{-e^y+1+\lambda}, \\ -u_{it}-\mathcal{L}_y u_i\leqslant 0 \quad \textit{if } u_i=\dfrac{1}{-e^y+1-\mu}, & \begin{array}{l} -\infty<y<y^*, \\ 0\leqslant t<T. \end{array} \\ u_i(y,T)=\psi_i(y), \\ u_i(y^*,t)=\dfrac{1}{-e^{y^*}+1-\mu}, \end{cases} \tag{3.12}$$

Assume that (a) $\psi_i(y^*)=\frac{1}{-e^{y^*}+1-\mu}$; (b) *$\psi_i(y)$ is bounded*; (c) *for any $N>0$, $u_i(y,t)\in W_p^{2,1}((-N,y^*)\times[0,T))$*; (d) *$\partial_y u_1$ is bounded. If $\psi_1(y)\geqslant\psi_2(y)$, then $u_1(y,t)\geqslant u_2(y,t)$.*

Proof. Denote

$$\mathcal{N}=\big\{(y,t):\ u_1(y,t)<u_2(y,t),\ -\infty<y<y^*,\ 0\leqslant t<T\big\}.$$

We use the method of contradiction. Suppose not. Then $\mathcal{N}$ must be a nonempty open set, and

$$u_1 < \frac{1}{-e^y + 1 - \mu}, \qquad u_2 > \frac{1}{-e^y + 1 + \lambda} \quad \text{in } \mathcal{N}.$$

It follows that

$$-u_{1t} - \mathcal{L}_y u_1 \geqslant 0, \qquad -u_{2t} - \mathcal{L}_y u_2 \leqslant 0 \quad \text{in } \mathcal{N}.$$

Let $\bar{w} = u_1 - u_2$, which satisfies

$$\begin{cases} -\bar{w}_t - \overline{\mathcal{L}}_y \bar{w} \geqslant 0 & \text{in } \mathcal{N}, \\ \bar{w} = 0 & \text{on } \partial_p \mathcal{N}, \end{cases}$$

where $\partial_p \mathcal{N}$ is the parabolic boundary of $\mathcal{N}$, and

$$\overline{\mathcal{L}}_y \bar{w} = \frac{1}{2}\sigma^2 \bar{w}_{yy} - \left(\alpha - r - \left(\frac{3}{2} - \gamma\right)\sigma^2\right)\bar{w}_y - \left(\alpha - r - (1-\gamma)\sigma^2\right)\bar{w}$$
$$- \gamma\sigma^2 e^y \left[u_2 \bar{w}_y + (u_{1y} + u_1 + u_2)\bar{w}\right].$$

Since $\bar{w}$ is bounded and all coefficients in the above equation are bounded as well, applying the maximum principle, we have $\bar{w} \geqslant 0$ in $\mathcal{N}$, namely $u_1 - u_2 \geqslant 0$ in $\mathcal{N}$, which contradicts the definition of $\mathcal{N}$. $\square$

Since the interval $(-\infty, y^*)$ is unbounded, we confine our attention to (3.11) in a finite domain $(-N, y^*) \times [0, T)$ with $N > 0$, namely,

$$\begin{cases} -u_t^N - \mathcal{L}_y u^N = 0 & \text{if } \dfrac{1}{-e^y + 1 + \lambda} < u^N < \dfrac{1}{-e^y + 1 - \mu}, \\ -u_t^N - \mathcal{L}_y u^N \geqslant 0 & \text{if } u^N = \dfrac{1}{-e^y + 1 + \lambda}, \\ -u_t^N - \mathcal{L}_y u^N \leqslant 0 & \text{if } u^N = \dfrac{1}{-e^y + 1 - \mu}, \\ u^N(y, T) = \dfrac{1}{-e^y + 1 - \mu}, \\ u^N(y^*, t) = \dfrac{1}{-e^{y^*} + 1 - \mu}, \quad u^N(-N, t) = \dfrac{1}{-e^{-N} + 1 - \mu}, \end{cases} \qquad \begin{array}{l} -N < y < y^*, \\ 0 \leqslant t < T, \end{array} \tag{3.13}$$

where a boundary condition on $y = -N$ is imposed.

Lemma 3.4. *For any $N > 0$ given, the problem* (3.13) *has a solution $u^N(y,t) \in W_p^{2,1}((-N, y^*) \times [0, T))$, $1 < p < +\infty$, and*

$$u_t^N \geqslant 0, \tag{3.14}$$

$$\left|u^N\right|_{W_p^{2,1}((-\widetilde{N}, y^*) \times [0,T))} \leqslant c, \tag{3.15}$$

where $\widetilde{N} < N$ and c depends only on $\widetilde{N}$ but is independent of N. Moreover

$$\left|u_y^N\right|_{L^\infty((-N, y^*) \times [0,T))} \leqslant M, \tag{3.16}$$

where M is independent of N.

Proof. Following Friedman [11], we consider a penalty approximation of the problem (3.13):

$$\begin{cases} -u_t^{N,\varepsilon} - \mathcal{L}_y u^{N,\varepsilon} + \beta_\varepsilon\left(u^{N,\varepsilon} - \dfrac{1}{-e^y+1+\lambda}\right) + \gamma_\varepsilon\left(u^{N,\varepsilon} - \dfrac{1}{-e^y+1-\mu}\right) = 0, & \\ u^{N,\varepsilon}(y,T) = \dfrac{1}{-e^y+1-\mu}, & \begin{array}{l} -N < y < y^*, \\ 0 \leqslant t < T, \end{array} \\ u^{N,\varepsilon}(y^*,t) = \dfrac{1}{-e^{y^*}+1-\mu}, \quad u^{N,\varepsilon}(-N,t) = \dfrac{1}{-e^{-N}+1-\mu}, & \end{cases} \tag{3.17}$$

where

$$\begin{array}{ll} \beta_\varepsilon(\xi) \leqslant 0, & \gamma_\varepsilon(\xi) \geqslant 0, \\ \beta_\varepsilon(\xi) = 0 \quad \text{if } \xi \geqslant \varepsilon, & \gamma_\varepsilon(\xi) = 0 \quad \text{if } \xi \leqslant -\varepsilon, \\ \beta_\varepsilon(0) = -c_1 \quad (c_1 > 0), & \gamma_\varepsilon(0) = c_2 \quad (c_2 > 0), \\ \beta'_\varepsilon(\xi) \geqslant 0, & \gamma''_\varepsilon(\xi) \geqslant 0, \\ \beta''_\varepsilon(\xi) \leqslant 0, & \gamma''_\varepsilon(\xi) \geqslant 0 \end{array}$$

with constants c_1 and c_2 to be chosen later. For any $\varepsilon > 0$ given, it is not hard to show by the fixed point theorem that the above semi-linear problem has a unique solution $u^{N,\varepsilon} \in W_p^{2,1}((-N, y^*) \times [0,T))$.

Next we want to prove

$$\frac{1}{-e^y+1+\lambda} \leqslant u^{N,\varepsilon} \leqslant \frac{1}{-e^y+1-\mu} \quad \text{in } (-N, y^*) \times [0,T). \tag{3.18}$$

Let

$$g(y) = \frac{1}{-e^y+1-\mu}.$$

Then

$$g'(y) = \frac{e^y}{(-e^y+1-\mu)^2}, \qquad g''(y) = \frac{e^{2y}+(1-\mu)e^y}{(-e^y+1-\mu)^3}.$$

Notice

$$\begin{aligned} &\left(-\frac{\partial}{\partial t} - \mathcal{L}_y\right) g(y) + \beta_\varepsilon\left(g(y) - \frac{1}{-e^y+1+\lambda}\right) + \gamma_\varepsilon\left(g(y) - \frac{1}{-e^y+1-\mu}\right) \\ &\quad = \frac{1-\mu}{(-e^y+1-\mu)^3}\left[-(\alpha-r)e^y + \left(\alpha - r - (1-\gamma)\sigma^2\right)(1-\mu)\right] \\ &\qquad + \beta_\varepsilon\left(\frac{\lambda+\mu}{(-e^y+1-\mu)(-e^y+1+\lambda)}\right) + \gamma_\varepsilon(0) \\ &\quad \geqslant -\frac{(1-\mu)^2(1-\gamma)\sigma^2}{(x^*+1-\mu)^3} + \beta_\varepsilon\left(\frac{\lambda+\mu}{(-e^y+1-\mu)(-e^y+1+\lambda)}\right) + \gamma_\varepsilon(0) \quad \text{in } (-N, y^*) \times [0,T). \end{aligned}$$

When ε is sufficiently small, $\beta_\varepsilon(\frac{\lambda+\mu}{(-e^y+1-\mu)(-e^y+1+\lambda)}) \equiv 0$. Take

$$c_2 = \gamma_\varepsilon(0) = \frac{(1-\mu)^2(1-\gamma)\sigma^2}{(x^*+1-\mu)^3}.$$

Then $\frac{1}{-e^y+1-\mu}$ is a supersolution, and thus

$$u^{N,\varepsilon} \leqslant \frac{1}{-e^y+1-\mu} \quad \text{in } (-N, y^*) \times [0, T).$$

In the same way, we can choose

$$c_1 = -\beta_\varepsilon(0) = \frac{|\alpha - r - (1-\gamma)\sigma^2|(1+\lambda)^2}{(x^*+1+\lambda)^3}$$

such that $\frac{1}{-e^y+1+\lambda}$ is a subsolution. So, we obtain (3.18).

Due to (3.18), we get

$$-c_1 \leqslant \beta_\varepsilon\left(u^{N,\varepsilon} - \frac{1}{-e^y+1+\lambda}\right) \leqslant 0 \quad \text{and} \quad 0 \leqslant \gamma_\varepsilon\left(u^{N,\varepsilon} - \frac{1}{-e^y+1-\mu}\right) \leqslant c_2.$$

We then deduce from (3.17) that $|u^{N,\varepsilon}|_{W_p^{2,1}((-N,y^*)\times[0,T))} \leqslant c$, where c is independent of ε because both c_1 and c_2 are independent of ε. Using $W_p^{2,1}$ interior estimate, we further have for any $\widetilde{N} < N$,

$$\left|u^{N,\varepsilon}\right|_{W_p^{2,1}((-\widetilde{N},y^*)\times[0,T))} \leqslant c,$$

where c depends on $\widetilde{N}$ but is independent of ε and N.

Due to (3.18), we infer $u_t^{N,\varepsilon}|_{t=T} \geqslant 0$. Differentiating the equation in (3.17) w.r.t. t, we get an equation that $u_t^{N,\varepsilon}$ satisfies. Applying the maximum principle, we then deduce

$$u_t^{N,\varepsilon} \geqslant 0,$$

which gives (3.14) by letting $\varepsilon \to 0$.

Now we prove (3.16). Since the bound of $u^{N,\varepsilon}$ and the C^2 norm of terminal value are independent of N and ε, we obtain by the $W_p^{2,1}$ interior estimate

$$\left|u^{N,\varepsilon}\right|_{W_p^{2,1}((-y-1,y)\times[0,T))} \leqslant M,$$

for any $y \leqslant 0$, where M is independent of N and ε. Applying the embedding theorem, we have

$$\left|u_y^{N,\varepsilon}\right|_{L^\infty((-y-1,y)\times[0,T))} \leqslant M.$$

Since y is arbitrary, it follows

$$\left|u_y^{N,\varepsilon}\right|_{L^\infty((-N,y^*)\times[0,T))} \leqslant M,$$

which yields (3.16) by letting $\varepsilon \to 0$. The proof is complete. □

In Lemma 3.4, we let N go to $+\infty$, which immediately leads to the existence of solution to the problem (3.11). The uniqueness of solution can be deduced from the comparison principle (Lemma 3.3). Therefore, we arrive at the following lemma.

Lemma 3.5. *The problem* (3.11) *has a unique solution* $u(y,t)$,

$$u(y,t) \in W_p^{2,1}\big((-N, y^*) \times [0,T)\big) \quad \textit{for any } N > 0,$$

$$u_t \geqslant 0,$$

$$|u_y|_{L^\infty((-\infty,y^*)\times[0,T))} \leqslant M.$$

Now we come back to the problem (3.4) with (3.5). By Lemma 3.5, (3.4) with (3.5) has a unique solution $v(x,t) \in W_p^{2,1}(\widetilde{\Omega} \cap \{x < -\delta\})$, for any $\delta > 0$. Similarly we can show that $v(x,t) \in W_p^{2,1}(\widetilde{\Omega} \cap \{x > \delta\})$, for any $\delta > 0$. So, we have the desired $W_p^{2,1}$ regularity. (3.7) also follows from Lemma 3.5.

In the following, we will prove (3.6). As mentioned earlier, we only need to show the C^∞ smoothness on $\{x = 0\} \cap \mathbf{NT}$ if it is nonempty. Let us first prove

$$v(x,t) \in C^\infty\big(\mathbf{NT} \cap \{x^* \leqslant x \leqslant 0\}\big). \tag{3.19}$$

Assume that $\mathbf{NT} \cap \{x^* \leqslant x \leqslant 0\}$ is nonempty. It is sufficient to prove that for any compact set $E \subset \mathbf{NT} \cap \{x^* \leqslant x \leqslant 0\}$, all derivatives of v are bounded on E. Notice that E may contain a part of line $x = 0$.

Let us start from the first equation in (3.11). Observe that the bound of the coefficient $-\gamma\sigma^2 e^y u$ is independent of y, and so are the bound of $u(y,t)$ and the C^1 norm of the initial value function. As a result, we infer by the $C^{\theta,\theta/2}$ estimate

$$|u|_{C^{\theta,\theta/2}(E_y)} \leqslant c,$$

where E_y denotes the counterpart of E w.r.t. y, and c depends only on the C^1 norm of the initial value function and the bounds of $u(y,t)$ and of those coefficients appeared in $\mathcal{L}_y$. Using the $C^{2+\theta,1+\theta/2}$ estimate, we have

$$|u|_{C^{2+\theta,1+\theta/2}(E_y)} \leqslant c.$$

We can further use the bootstrap argument to get the boundedness of any order partial derivative of $u(y,t)$ w.r.t. y.

We assert that $v_x = -e^{-y}u_y$ is bounded. Indeed, denote $\tilde{u}(y,t) = -e^{-y}u_y$, which satisfies

$$\begin{cases} -\tilde{u}_t - \dfrac{1}{2}\sigma^2\tilde{u}_{yy} + \left(\alpha - r - \left(\dfrac{5}{2} - \gamma\right)\sigma^2\right)\tilde{u}_y + 2\left(\alpha - r - \left(\dfrac{3}{2} - \gamma\right)\sigma^2\right)\tilde{u} \\ \quad = -\gamma\sigma^2\big(uu_{yy} + u_y^2 + 3uu_y + u^2\big) \quad \text{in } E_y, \\ \tilde{u}(y,T) = -\dfrac{1}{(-e^y + 1 - \mu)^2}. \end{cases}$$

Because both the right-hand side terms of the above equation and the terminal value are bounded, we deduce that $\tilde{u}$ is bounded, so is v_x. Using the same argument, we can show the boundedness of $v_{xx} = e^{-2y}(u_{yy} - u_y)$ as well as of any order partial derivatives in x. (3.19) then follows.

Due to (3.19), we obtain an *ordinary* differential inequality by letting $x \to 0^-$ in (3.4)

$$\begin{cases} -v_t(0,t) + \big(\alpha - r - (1-\gamma)\sigma^2\big)v(0,t) = 0 & \text{if } \dfrac{1}{1+\lambda} < v(0,t) < \dfrac{1}{1-\mu}, \\ -v_t(0,t) + \big(\alpha - r - (1-\gamma)\sigma^2\big)v(0,t) \geqslant 0 & \text{if } v(0,t) = \dfrac{1}{1+\lambda}, \\ -v_t(0,t) + \big(\alpha - r - (1-\gamma)\sigma^2\big)v(0,t) \leqslant 0 & \text{if } v(0,t) = \dfrac{1}{1-\mu}, \\ v(0,T) = \dfrac{1}{1-\mu}. \end{cases} \quad 0 \leqslant t < T, \tag{3.20}$$

In the same way, we can show $v(x,t) \in C^\infty(\mathbf{NT} \cap \{x \geqslant 0\})$ and obtain the same differential inequality (3.20) by letting $x \to 0^+$. This implies the continuity of $v(x,t)$ across $\mathbf{NT} \cap \{x = 0\}$. Solving (3.20) gives (3.8)–(3.9), from which we infer that $\mathbf{NT} \cap \{x = 0\} = \emptyset$ when $\alpha - r - (1-\gamma)\sigma^2 \leqslant 0$, and $\mathbf{NT} \cap \{x = 0\} = \{x = 0,\ t_1 < t < T\}$ when $\alpha - r - (1-\gamma)\sigma^2 > 0$. Assume that $\mathbf{NT} \cap \{x = 0\} \neq \emptyset$ (i.e., $\alpha - r - (1-\gamma)\sigma^2 > 0$). We differentiate the first equation of (3.4) w.r.t. x and let $x \to 0^+$ and $x \to 0^-$ respectively. Then we obtain the same ordinary differential equation and terminal condition for $v_x(0^+,t)$ and $v_x(0^-,t)$ in $\mathbf{NT} \cap \{x = 0\}$, which implies the continuity of $v_x(x,t)$ across $\mathbf{NT} \cap \{x = 0\}$. Using the same argument we can show that any order derivative of $v(x,t)$ is continuous across $\mathbf{NT} \cap \{x = 0\}$. So, $v(x,t) \in C^\infty(\mathbf{NT})$ and the proof is complete.

4. Characterization of free boundaries

This section is devoted to the theoretical analysis of free boundaries. A double obstacle problem usually gives rise to two free boundaries. We will first show that each free boundary can be expressed as a single-value function of time t. Then we will examine the properties of the free boundaries.

To begin with, we introduce a lemma which will play a critical role in the existence proof of free boundaries.

Lemma 4.1. *Let $v(x,t)$ be the solution to the double obstacle problem* (3.4)*. Then*

$$v_x + v^2 \leqslant 0 \quad \text{in } \Omega.$$

Proof. It is clear that $v_x + v^2 = 0$ in **BR** and **SR**. So, the rest is to show $v_x + v^2 \leqslant 0$ in **NT**. Denote

$$p(x,t) = v_x(x,t) \quad \text{and} \quad q(x,t) = v^2(x,t).$$

It is not hard to check that

$$\begin{aligned} &-p_t - \frac{1}{2}\sigma^2 x^2 p_{xx} + \big(\alpha - r - (3-\gamma)\sigma^2\big)xp_x + \big(2\alpha - 2r - (3-2\gamma)\sigma^2\big)p \\ &\quad = \gamma\sigma^2\big(4xvv_x + x^2v_x^2 + x^2vv_{xx} + v^2\big) \quad \text{in } \mathbf{NT} \end{aligned}$$

and

$$\begin{aligned} &-q_t - \frac{1}{2}\sigma^2 x^2 q_{xx} + \big(\alpha - r - (2-\gamma)\sigma^2\big)xq_x + \big(2\alpha - 2r - (2-2\gamma)\sigma^2\big)q \\ &\quad = -\sigma^2x^2v_x^2 + \gamma\sigma^2\big(2x^2v^2v_x + 2xv^3\big) \quad \text{in } \mathbf{NT}. \end{aligned}$$

Let $H(x,t) = v_x(x,t) + v^2(x,t) = p(x,t) + q(x,t)$. Then $H(x,t)$ satisfies

$$
\begin{aligned}
&-H_t - \frac{1}{2}\sigma^2 x^2 H_{xx} + \big(\alpha - r - (3-\gamma)\sigma^2\big)xH_x + \big(2\alpha - 2r - (3-2\gamma)\sigma^2\big)H \\
&\quad = -\sigma^2 x^2 v_x^2 - \sigma^2 x q_x - \sigma^2 q + \gamma\sigma^2\big(2x^2 v^2 v_x + 2xv^3 + 4xvv_x + x^2 v_x^2 + x^2 v v_{xx} + v^2\big) \\
&\quad = -\sigma^2\big(x^2 v_x^2 + 2xvv_x + \sigma^2 v^2\big) + \gamma\sigma^2\big[x^2 v\big(v_x + v^2\big)_x + 2xv\big(v^2 + v_x\big) + (xv_x + v)^2\big] \\
&\quad = -(1-\gamma)\sigma^2 (xv_x + v)^2 + \gamma\sigma^2\big[x^2 v H_x + 2xvH\big] \quad \text{in } \mathbf{NT},
\end{aligned}
$$

namely,

$$
\begin{aligned}
&-H_t - \frac{1}{2}\sigma^2 x^2 H_{xx} + \big(\alpha - r - (3-\gamma)\sigma^2 - \gamma\sigma^2 xv\big)xH_x + \big(2\alpha - 2r - (3-2\gamma)\sigma^2 - 2\gamma\sigma^2 xv\big)H \\
&\quad = -(1-\gamma)\sigma^2 (xv_x + v)^2 \leqslant 0 \quad \text{in } \mathbf{NT}.
\end{aligned}
$$

Obviously $H(x,t) = 0$ on the parabolic boundary of **NT**. Applying the maximum principle yields the desired result. □

Remark 4.2. Assume that we already have $w_x = v$, then it is easy to verify that Lemma 4.1 is equivalent to

$$\varphi_{xx}\varphi_{yy} - \varphi_{xy}^2 \leqslant 0,$$

which implies the concavity of the original value function $\varphi(x,y,t)$ in x and y. Note that the concavity is obvious from the definition of $\varphi(x,y,t)$ for the dynamics are linear and the utility function is concave. However, we emphasize that at the moment the relation $w_x = v$ has not yet been established, so we cannot utilize the concavity to prove Lemma 4.1.

Now we can prove the existence of two free boundaries.

Theorem 4.3. *There are two monotonically increasing functions $x_s(t) : [0,T] \to [-(1-\mu), +\infty]$ and $x_b(t) : [0,T] \to [-(1-\mu), +\infty]$, such that*

$$\mathbf{SR} = \big\{(x,t) \in \Omega:\ x \leqslant x_s(t),\ t \in [0,T)\big\}$$

and

$$\mathbf{BR} = \big\{(x,t) \in \Omega:\ x \geqslant x_b(t),\ t \in [0,T)\big\}.$$

Moreover,

$$x_s(t) < x_b(t) \quad \text{for all } t \in [0,T). \tag{4.1}$$

Proof. Notice

$$\frac{\partial}{\partial x}\left(v - \frac{1}{x+1+\lambda}\right) = v_x + \frac{1}{(x+1+\lambda)^2} \leqslant v_x + v^2.$$

It follows from the above lemma

$$\frac{\partial}{\partial x}\left(v - \frac{1}{x+1+\lambda}\right) \leqslant 0.$$

As a consequence, if $(x_1, t) \in \mathbf{BR}$, i.e., $v(x_1, t) = \frac{1}{x_1+1+\lambda}$, then for any $x_2 > x_1$,

$$0 \leqslant v(x_2, t) - \frac{1}{x_2 + 1 + \lambda} \leqslant v(x_1, t) - \frac{1}{x_1 + 1 + \lambda} = 0,$$

from which we infer $v(x_2, t) = \frac{1}{x_2+1+\lambda}$, i.e., $(x_2, t) \in \mathbf{BR}$. This indicates the connection of each t-section of $\mathbf{BR}$. The existence of $x_b(t)$ (as a single-value function) follows.

To prove the existence of $x_s(t)$, we instead consider $\bar{v}(x, t) = (x + 1 - \mu)^2 v(x, t)$. Notice $\mathbf{BR} = \{(x, t) \in \Omega\colon\ \bar{v}(x, t) = x + 1 - \mu\}$ and

$$\begin{aligned}
\frac{\partial}{\partial x}\big[\bar{v} - (x + 1 - \mu)\big] &= \frac{\partial}{\partial x}\left[(x + 1 - \mu)^2\left(v - \frac{1}{x + 1 - \mu}\right)\right] \\
&= 2(x + 1 - \mu)\left(v - \frac{1}{x + 1 - \mu}\right) + (x + 1 - \mu)^2\left(v_x + \frac{1}{(x + 1 - \mu)^2}\right) \\
&= -\big[(x + 1 - \mu)v - 1\big]^2 + (x + 1 - \mu)^2\big(v_x + v^2\big) \\
&\leqslant 0.
\end{aligned}$$

The desired result then follows.

The monotonicity of $x_s(t)$ and $x_b(t)$ can be similarly deduced by virtue of

$$\frac{\partial}{\partial t}\left(v - \frac{1}{x + 1 + \lambda}\right) = \frac{\partial}{\partial t}\left(v - \frac{1}{x + 1 - \mu}\right) = v_t \geqslant 0.$$

(4.1) is clear because $\mathbf{SR} \cap \mathbf{BR} = \emptyset$. $\square$

Remark 4.4. The monotonicity of $x_s(t)$ and $x_b(t)$ indicates that the shorter the maturity, the less the chance of buying risky asset and the more the chance of selling risky asset. This is consistent with the investment criterion that younger investors should allocate a greater share of wealth to stocks than older investors.

In finance, $x_s(t)$ and $x_b(t)$ stand for the optimal selling and buying boundaries, respectively. In what follows we study their behaviors. To begin with, we study the selling boundary $x_s(t)$.

Theorem 4.5. *Let $x_s(t)$ be the optimal selling boundary in Theorem 4.3. Then*

(i)
$$x_s(t) \leqslant (1 - \mu)x_M, \tag{4.2}$$

where x_M is defined in (2.3), and

$$x_s(T^-) \triangleq \lim_{t \to T^-} x_s(t) = (1 - \mu)x_M; \tag{4.3}$$

(ii)
$$x_s(t) \equiv 0 \quad \text{when } \alpha - r - (1 - \gamma)\sigma^2 = 0, \tag{4.4}$$

$$x_s(t) > 0 \quad \text{when } \alpha - r - (1 - \gamma)\sigma^2 < 0, \tag{4.5}$$

$$x_s(t) < 0 \quad \text{when } \alpha - r - (1 - \gamma)\sigma^2 > 0; \tag{4.6}$$

(iii) *$x_s(t)$ is continuous. Moreover, $x_s(t) \in C^\infty[0, T)$.*

Proof. For any $(x,t)\in \mathbf{SR}$,

$$\begin{aligned}0 &\geqslant \left(-\frac{\partial}{\partial t}-\mathcal{L}\right)\left(\frac{1}{x+1-\mu}\right)=-\mathcal{L}\left(\frac{1}{x+1-\mu}\right)\\ &=\frac{1-\mu}{(x+1-\mu)^3}\big[(\alpha-r)x+(1-\mu)\big(\alpha-r-(1-\gamma)\sigma^2\big)\big].\end{aligned}\tag{4.7}$$

Then

$$x\leqslant -\frac{\alpha-r-(1-\gamma)\sigma^2}{\alpha-r}(1-\mu)=(1-\mu)x_M,$$

which implies (4.2).

To prove the rest part, we will utilize (3.7) many times. Let us first consider (4.3). Suppose not, we then have $x_s(T^-)<(1-\mu)x_M$. For any $x_0\in (x_s(T^-),(1-\mu)x_M)$, applying the equation $-v_t-\mathcal{L}v=0$ at $t=T$ gives $v_t|_{t=T,\ x=x_0}=-\mathcal{L}v|_{t=T,\ x=x_0}=-\mathcal{L}(\frac{1}{x+1-\mu})|_{x=x_0}<0$, where the last (strict) inequality is due to (4.7). This, however, contradicts (3.7).

(4.6) is a corollary of (4.2). Now we prove (4.5). Again we use the method of contradiction. Suppose not. Since $x_s(T^-)>0$ when $\alpha-r-(1-\gamma)\sigma^2<0$, there must be a $\tilde{t}<T$ such that $x_s(t)\leqslant 0$ for $t\leqslant\tilde{t}$. Then Eq. (3.4) on $x=0$ reduces to

$$\begin{cases}-v_t+\big(\alpha-r-(1-\gamma)\sigma^2\big)v\big|_{x=0}\geqslant 0 & \text{for } t<\tilde{t},\\ v(0,t)=\dfrac{1}{x+1-\mu} & \text{for } \tilde{t}\leqslant t\leqslant T.\end{cases}$$

So, $v_t(0,t)\leqslant(\alpha-r-(1-\gamma)\sigma^2)v(0,t)<0$ for $t<\tilde{t}$, which also contradicts (3.7). So (4.5) follows.

Combining (4.2) and (4.5), we get

$$0<x_s(t)\leqslant(1-\mu)x_M=-(1-\mu)\frac{\alpha-r-(1-\gamma)\sigma^2}{\alpha-r}\quad\text{when } \alpha-r-(1-\gamma)\sigma^2<0.$$

We then obtain (4.4) by letting $\alpha-r-(1-\gamma)\sigma^2\to 0$.

At last, let us prove the continuity of $x_s(t)$. Suppose not, i.e., there is a time $t^*<T$ at which $x_s(t)$ is discontinuous. Then

$$v(x,t^*)=\frac{1}{x+1-\mu},\quad x\in\big[x_s(t^{*-}),x_s(t^{*+})\big].$$

Applying the equation $-v_t-\mathcal{L}v=0$ on $t=t^*$, $x\in(x_s(t^{*-}),x_s(t^{*+}))$, we have

$$\begin{aligned}v_t(x,t^*)&=-\mathcal{L}\left(\frac{1}{x+1-\mu}\right)\\ &=\frac{1-\mu}{(x+1-\mu)^3}\big[(\alpha-r)x+(1-\mu)\big(\alpha-r-(1-\gamma)\sigma^2\big)\big]<0,\end{aligned}$$

which is again in contradiction with (3.7). So, we obtain the continuity of $x_s(t)$. Thanks to (3.7), we can take advantage of the same arguments as in Friedman [10] to obtain $x_s(t)\in C^\infty[0,T)$. The proof is complete. □

Remark 4.6. From (4.4)–(4.5), we deduce $x_s(t)\geqslant 0$ (i.e. $\mathbf{NT}\subset\{x>0\}$) if and only if $\alpha-r-(1-\gamma)\sigma^2\leqslant 0$, which indicates that the leverage is always suboptimal when $\alpha-r-(1-\gamma)\sigma^2\leqslant 0$. This conclusion is the same as in the absence of transaction costs.

Now we move on to the buying boundary $x_b(t)$.

Theorem 4.7. *Let $x_b(t)$ be the optimal buying boundary in Theorem 4.3. Denote*

$$t_0 = T - \frac{1}{\alpha - r}\log\left(\frac{1+\lambda}{1-\mu}\right). \tag{4.8}$$

Then,

(i)
$$x_b(t) \geqslant (1+\lambda)x_M, \tag{4.9}$$

where x_M is defined in (2.3), *and*

$$x_b(t) = \infty \quad \text{if and only if} \quad t_0 \leqslant t < T; \tag{4.10}$$

(ii)
$$x_b(t) > 0 \quad \text{when } \alpha - r - (1-\gamma)\sigma^2 \leqslant 0 \tag{4.11}$$

and

$$\begin{gathered} x_b(t) > 0 \quad \text{for } t \in (t_1, T), \qquad x_b(t_1) = 0, \\ x_b(t) < 0 \quad \text{for } t \in (0, t_1) \text{ when } \alpha - r - (1-\gamma)\sigma^2 > 0. \end{gathered} \tag{4.12}$$

Here t_1 is defined in (3.10).

(iii) *$x_b(t)$ is continuous.*

Proof. (4.9) is a counterpart of (4.2) and can be similarly deduced by

$$\left(-\frac{\partial}{\partial t} - \mathcal{L}\right)\left(\frac{1}{x+1+\lambda}\right) = \frac{1+\lambda}{(x+1+\lambda)^3}\left[(\alpha - r)x + (1+\lambda)\big(\alpha - r - (1-\gamma)\sigma^2\big)\right] \geqslant 0.$$

Combination of (4.1), (4.4) and (4.5) yields (4.11), and (4.12) can be directly inferred from (3.9). The proof of the continuity of $x_b(t)$ is the same as that of $x_s(t)$. Hence, what remains is to prove (4.10). By the transformation

$$z = \frac{x}{x+1+\lambda}, \qquad \tilde{v}(z,t) = \left(v(x,t) - \frac{1}{x+1+\lambda}\right)\frac{(x+1+\lambda)^2}{1+\lambda},$$

the problem (3.4) becomes

$$\begin{cases} -\tilde{v}_t - \widetilde{\mathcal{L}}\tilde{v} = 0 & \text{if } 0 < \tilde{v} < \dfrac{\lambda+\mu}{(1-\mu)+(\lambda+\mu)z}, \\ -\tilde{v}_t - \widetilde{\mathcal{L}}\tilde{v} \geqslant 0 & \text{if } \tilde{v} = 0, \qquad\qquad \dfrac{1-\mu}{\lambda+\mu} < z < 1, \\ -\tilde{v}_t - \widetilde{\mathcal{L}}\tilde{v} \leqslant 0 & \text{if } \tilde{v} = \dfrac{\lambda+\mu}{(1-\mu)+(\lambda+\mu)z}, \qquad 0 \leqslant t < T, \\ \tilde{v}(z,T) = \dfrac{\lambda+\mu}{(1-\mu)+(\lambda+\mu)z}. \end{cases} \tag{4.13}$$

Here

$$\begin{aligned} \widetilde{\mathcal{L}}\tilde{v} = {} & \frac{1}{2}\sigma^2 z^2(1-z)^2\tilde{v}_{zz} - \big(\big(\alpha - r - (2-\gamma)\sigma^2\big) + 3\sigma^2 z\big)z(1-z)\tilde{v}_z \\ & - \big(\alpha - r - (1-\gamma)\sigma^2 - 2\big(\alpha - r - (2-\gamma)\sigma^2\big)z - 3\sigma^2 z^2\big)\tilde{v} \\ & - \big(\sigma^2 z + \alpha - r - (1-\gamma)\sigma^2\big) + \gamma\sigma^2 z\big[1 + (1-z)\tilde{v}\big]\big[z(1-z)\tilde{v}_z + (1-2z)\tilde{v} + 1\big]. \end{aligned}$$

It is easy to verify

$$\tilde{v}_z = v_x + v^2 - \left(v - \frac{1}{x+1+\lambda}\right)^2 \leqslant 0. \tag{4.14}$$

Define

$$z_b(t) = \sup\left\{z \in \left[\frac{\lambda+\mu}{1-\mu}, 1\right]:\ \tilde{v}(z,t) > 0\right\}, \quad t \in [0,T).$$

Clearly

$$z_b(t) = \frac{x_b(t)}{x_b(t)+1+\lambda}.$$

In order to prove (4.10), it suffices to show

$$z_b(t) = 1 \quad \text{if and only if} \quad t \in [t_0, T).$$

Note that at $z = 1$, the problem (4.13) is reduced to

$$\begin{cases} -\tilde{v}_t(1,t) - (\alpha - r)\tilde{v}(1,t) + \alpha - r = 0 & \text{if } 0 < \tilde{v}(1,t) < \dfrac{\lambda+\mu}{1+\lambda}, \\ -\tilde{v}_t(1,t) - (\alpha - r)\tilde{v}(1,t) + \alpha - r \geqslant 0 & \text{if } \tilde{v}(1,t) = 0, \\ -\tilde{v}_t(1,t) - (\alpha - r)\tilde{v}(1,t) + \alpha - r \leqslant 0 & \text{if } \tilde{v}(1,t) = \dfrac{\lambda+\mu}{1+\lambda}, \\ \tilde{v}(1,T) = \dfrac{\lambda+\mu}{1+\lambda}, \end{cases} \qquad 0 \leqslant t < T,$$

whose solution is

$$\tilde{v}(1,t) = \max\left(1 - e^{(\alpha-r)(T-t)}\frac{1-\mu}{1+\lambda}, 0\right) = \begin{cases} 1 - e^{(\alpha-r)(T-t)}\frac{1-\mu}{1+\lambda} & \text{when } t_0 < t \leqslant T, \\ 0 & \text{when } 0 \leqslant t \leqslant t_0. \end{cases}$$

Combining with (4.14), we immediately obtain $z_b(t) = 1$ for $t \in (t_0, T)$ and $z_b(t) < 1$ for $t \in [0, t_0)$. So, we only need to further prove $z_b(t_0) = 1$. Using the strong maximum principle, we can deduce $v_x(x,t_0) + v^2(x,t_0) < 0$ for $-(1-\mu) < x < +\infty$ and thus $\tilde{v}_z(z,t_0) < 0$ for $z < 1$, which implies $\tilde{v}(z,t_0) > 0$ for $z < 1$. Hence, we must have $z_b(t_0) = 1$. This completes the proof. $\square$

Remark 4.8. (4.2) and (4.4) in Theorem 4.5 and part (i) in Theorem 4.7 are also obtained by Liu and Loewenstein [16] in terms of another approach.

5. Equivalence

This section is devoted to the equivalence between the double obstacle problem (3.4) and the original problem (3.2).

Theorem 5.1. *Let $v(x,t)$ be the solution to the double obstacle problem* (3.4)*. Define*

$$w(x,t) = A(t) + \log\bigl(x_s(t) + 1 - \mu\bigr) + \int_{x_s(t)}^{x} v(\xi,t)\,d\xi, \tag{5.1}$$

where

$$A(t)=\int_t^T \frac{rx^2+(\alpha+r)(1-\mu)x+(\alpha-\frac{1}{2}\sigma^2(1-\gamma))(1-\mu)^2}{(x+1-\mu)^2}\bigg|_{x=x_s(\tau)}d\tau. \tag{5.2}$$

Then $w(x,t)$ is the solution to the problem (3.2)*. Moreover,*

$$w(x,t)\in C^{2,1}(\Omega\setminus F), \tag{5.3}$$

where F is the intersection of the free boundaries and the line $x=0$, i.e., $F=\{(0,t)\mid x_s(t)=0$ or $x_b(t)=0$, $t\in[0,T)\}$.

Remark 5.2. We exclude the set F on which some partial derivatives of $v(x,t)$ or $w(x,t)$ are discontinuous because of the degeneracy of the differential operator $\mathcal{L}$ or $\mathcal{L}_3$. Thanks to Theorems 4.5 and 4.7,

$$F=\begin{cases}\emptyset & \text{if } \alpha-r-(1-\gamma)\sigma^2<0,\\ \{x=0\} & \text{if } \alpha-r-(1-\gamma)\sigma^2=0,\\ (0,t_1) & \text{if } \alpha-r-(1-\gamma)\sigma^2>0,\end{cases}$$

where t_1 is defined in (3.10).

Proof of Theorem 5.1. Since $v(x,t)=\frac{1}{x+1-\mu}$ for $x\leqslant x_s(t)$, it is not hard to get by (5.1)

$$w(x,t)=A(t)+\log(x+1-\mu),\quad x\leqslant x_s(t). \tag{5.4}$$

Clearly $w(x,t)$ satisfies the terminal condition. Therefore, to prove that $w(x,t)$ is the solution to the problem (3.2), it suffices to show

$$\begin{cases}-w_t-\mathcal{L}_3 w\geqslant 0 & \text{in } \mathbf{SR} \text{ and } \mathbf{BR},\\ -w_t-\mathcal{L}_3 w=0 & \text{in } \mathbf{NT}.\end{cases} \tag{5.5}$$

Observe

$$w_x(x,t)=v(x,t). \tag{5.6}$$

According to the definition of $A(t)$, we claim

$$-w_t-\mathcal{L}_3 w=0\quad \text{on } x=x_s(t). \tag{5.7}$$

Indeed, because of (5.6),

$$\begin{aligned}\mathcal{L}_3 w|_{x=x_s(t)}&=\frac{1}{2}\sigma^2x^2\big(v_x+\gamma v^2\big)-\big(\alpha-r-\sigma^2(1-\gamma)\big)xv+\alpha-\frac{1}{2}\sigma^2(1-\gamma)\Big|_{x=x_s(t)}\\ &=\frac{1}{2}\sigma^2x_s(t)^2\left(-\frac{1}{(x_s(t)+1-\mu)^2}+\frac{\gamma}{(x_s(t)+1-\mu)^2}\right)\\ &\quad-\big(\alpha-r-\sigma^2(1-\gamma)\big)\frac{x_s(t)}{x_s(t)+1-\mu}+\alpha-\frac{1}{2}\sigma^2(1-\gamma)\\ &=\frac{1}{(x_s(t)+1-\mu)^2}\left[rx_s(t)^2+(\alpha+r)(1-\mu)x_s(t)+\left(\alpha-\frac{1}{2}\sigma^2(1-\gamma)\right)(1-\mu)^2\right]\\ &=-A'(t)=-w_t\big(x_s(t),t\big),\end{aligned} \tag{5.8}$$

where the last equality is due to (5.4).

Furthermore, due to (3.3), (5.6) and the fact that $v(x,t)$ is the solution to the problem (3.4), we have

$$\frac{\partial}{\partial x}(-w_t - \mathcal{L}_3 w) \leqslant 0, \qquad w_x = \frac{1}{x+1-\mu} \quad \text{if } x \leqslant x_s(t) \text{ (i.e. in } \mathbf{SR}\text{)},$$
$$\frac{\partial}{\partial x}(-w_t - \mathcal{L}_3 w) = 0 \quad \text{if } x_s(t) < x < x_b(t) \text{ (i.e. in } \mathbf{NT}\text{)},$$
$$\frac{\partial}{\partial x}(-w_t - \mathcal{L}_3 w) \geqslant 0, \qquad w_x = \frac{1}{x+1+\lambda} \quad \text{if } x \geqslant x_b(t) \text{ (i.e. in } \mathbf{BR}\text{)}.$$

Combining with (5.7), we then deduce (5.5).

Now we prove (5.3). By Proposition 3.2, $v \in C^{1,0}(\Omega \setminus F)$ and then $w \in C^{2,0}(\Omega \setminus F)$. What remains is to show $w_t \in C(\Omega \setminus F)$. Owing to (5.1),

$$\begin{aligned}
w_t(x,t) &= A'(t) + \frac{x_s'(t)}{x_s(t)+1-\mu} - v\big(x_s(t),t\big)x_s'(t) + \int_{x_s(t)}^{x} v_t(\xi,t)\,d\xi \\
&= A'(t) + \int_{x_s(t)}^{\max(\min(x,x_b(t)),x_s(t))} v_t(\xi,t)\,d\xi = A'(t) - \int_{x_s(t)}^{\max(\min(x,x_b(t)),x_s(t))} \mathcal{L}v(\xi,t)\,d\xi \\
&= A'(t) - \int_{x_s(t)}^{\max(\min(x,x_b(t)),x_s(t))} d\mathcal{L}_3 w = A'(t) + \mathcal{L}_3 w|_{x_s(t)} - \mathcal{L}_3 w|_{\max(\min(x,x_b(t)),x_s(t))} \\
&= -\mathcal{L}_3 w|_{\max(\min(x,x_b(t)),x_s(t))},
\end{aligned} \tag{5.9}$$

which implies the continuity of w_t. The proof is complete. □

6. A semi-explicit stationary solution to the obstacle problem

In this section, we aim to investigate the asymptotic behavior of $x_s(t)$ and $x_b(t)$ as $T-t \to +\infty$ through the stationary double obstacle problem, which can be written as a stationary free boundary problem:

$$\mathcal{L}v_\infty = 0, \quad x \in (x_{s,\infty}, x_{b,\infty}), \tag{6.1}$$

$$v_\infty(x_{s,\infty}) = \frac{1}{x_{s,\infty}+1-\mu}, \qquad v_\infty'(x_{s,\infty}) = -\frac{1}{(x_{s,\infty}+1-\mu)^2}, \tag{6.2}$$

$$v_\infty(x_{b,\infty}) = \frac{1}{x_{b,\infty}+1+\lambda}, \qquad v_\infty'(x_{b,\infty}) = -\frac{1}{(x_{b,\infty}+1+\lambda)^2}. \tag{6.3}$$

Here

$$x_{s,\infty} = \lim_{T-t\to+\infty} x_s(t) \quad \text{and} \quad x_{b,\infty} = \lim_{T-t\to+\infty} x_b(t)$$

and clearly

$$x_{s,\infty} < x_{b,\infty} < 0 \quad \text{when } \alpha - r - (1-\gamma)\sigma^2 > 0;$$
$$0 < x_{s,\infty} < x_{b,\infty} \quad \text{when } \alpha - r - (1-\gamma)\sigma^2 < 0.$$

We will reduce Eq. (6.1) to a Riccati equation, through which we can find a semi-explicit solution to problem (6.1)–(6.3) if $\alpha - r - (1-\gamma)\sigma^2 \neq 0$. The main result is summarized as follows.

Theorem 6.1. *Assume $\alpha - r - (1-\gamma)\sigma^2 \neq 0$. The solution of the stationary problem* (6.1)–(6.3) *takes the form*

$$v_\infty(x) = \begin{cases} \frac{C}{x} + \frac{1}{g(x)} & \text{if } x_{s,\infty} < x < x_{b,\infty}, \\ \frac{1}{x+1-\mu} & \text{if } x \leqslant x_{s,\infty}, \\ \frac{1}{x+1+\lambda} & \text{if } x \geqslant x_{b,\infty}, \end{cases}$$

where

$$g(x) = \left(\frac{x_{s,\infty}}{x}\right)^\beta \left(\frac{x_{s,\infty}(x_{s,\infty}+1-\mu)}{(1-C)x_{s,\infty} - (1-\mu)C} - \frac{\gamma x_{s,\infty}}{\beta+1}\right) + \frac{\gamma x}{\beta+1} \quad \text{if } \beta \neq -1, \tag{6.4}$$

$$g(x) = \frac{x(x_{s,\infty}+1-\mu)}{(1-C)x_{s,\infty} - (1-\mu)C} + \gamma x \log \frac{x}{x_{s,\infty}} \quad \text{if } \beta = -1, \tag{6.5}$$

$$x_{s,\infty} = -\frac{a}{a+k}(1-\mu), \qquad x_{b,\infty} = -\frac{a}{a+\frac{k}{k-1}}(1+\lambda), \tag{6.6}$$

$$a = \frac{\alpha - r - (1-\gamma)\sigma^2}{\frac{1}{2}(1-\gamma)\sigma^2}, \tag{6.7}$$

$$\beta = (1-\gamma)a - 2\gamma C, \tag{6.8}$$

$$C = -\frac{2(k-1)a^2}{k^2\left(a + \frac{1}{1-\gamma} + \sqrt{(a+\frac{1}{1-\gamma})^2 + 4\frac{\gamma}{1-\gamma}\frac{k-1}{k^2}a^2}\right)}, \tag{6.9}$$

and k is the root to the following equation in $(1,2)$ *when $a > 0$, or in* $(2,\infty)$ *when $a < 0$,*

$$\frac{a+\frac{k}{k-1}}{a+k}\left(\frac{\frac{\gamma}{\beta+1} + \frac{1}{\frac{1}{k}a+C}}{\frac{\gamma}{\beta+1} + \frac{1}{\frac{k-1}{k}a+C}}\right)^{\frac{1}{\beta+1}} = \frac{1+\lambda}{1-\mu} \quad \text{if } \beta \neq -1, \tag{6.10}$$

$$\frac{a+\frac{k}{k-1}}{a+k}\exp\left(\frac{1}{\gamma}\left(\frac{1}{\frac{1}{k}a+C} - \frac{1}{\frac{k-1}{k}a+C}\right)\right) = \frac{1+\lambda}{1-\mu} \quad \text{if } \beta = -1. \tag{6.11}$$

Let us give a remark before the proof.

Remark 6.2. For $\alpha - r - (1-\gamma)\sigma^2 > 0$ (i.e. $a > 0$), using (6.6) and noticing $k \in (1,2)$, we immediately obtain

$$x_s(t) \geqslant \left[\frac{(1-\gamma)\sigma^2}{2(\alpha - r) - (1-\gamma)\sigma^2} - 1\right](1-\mu). \tag{6.12}$$

Liu and Loewenstein [16] obtained another estimate:

$$x_s(t) \geqslant \left[\frac{(1-\gamma)\sigma^2}{2(\alpha - r)} - 1\right](1-\mu), \tag{6.13}$$

which is obviously not sharp because we have shown $x_s(t) \geqslant 0$ when $\alpha - r - (1-\gamma)\sigma^2 \leqslant 0$. For $\alpha - r - (1-\gamma)\sigma^2 > 0$, our estimate (6.12) is also better than (6.13).

Proof of Theorem 6.1. Define

$$f(x)=-\frac{x^2}{1-\gamma}\big(v_\infty'(x)+\gamma v_\infty^2(x)\big)+axv_\infty(x).$$

Note that

$$\mathcal{L}v_\infty=-\frac{1}{2}(1-\gamma)\sigma^2\frac{d}{dx}f(x).$$

Owing to (6.1),

$$-\frac{x^2}{1-\gamma}\big(v_\infty'(x)+\gamma v_\infty^2(x)\big)+axv_\infty(x)=f(x_{s,\infty}),\quad x\in(x_{s,\infty},x_{b,\infty}). \tag{6.14}$$

This is a Riccati equation that we need to solve subject to the free boundary conditions (6.2)–(6.3).

Applying the boundary condition (6.2), we have

$$\begin{aligned} f(x_{s,\infty}) &= -\frac{x_{s,\infty}^2}{1-\gamma}\left(-\frac{1}{(x_{s,\infty}+1-\mu)^2}+\frac{\gamma}{(x_{s,\infty}+1-\mu)^2}\right)+\frac{ax_{s,\infty}}{x_{s,\infty}+1-\mu} \\ &= \frac{(a+1)x_{s,\infty}^2+a(1-\mu)x_{s,\infty}}{(x_{s,\infty}+1-\mu)^2}. \end{aligned} \tag{6.15}$$

Similarly, the boundary condition (6.3) gives

$$f(x_{b,\infty})=\frac{(a+1)x_{b,\infty}^2+a(1+\lambda)x_{b,\infty}}{(x_{b,\infty}+1+\lambda)^2}.$$

Due to (6.14), $f(x_{s,\infty})=f(x_{b,\infty})$ or

$$\frac{(a+1)x_{s,\infty}^2+a(1-\mu)x_{s,\infty}}{(x_{s,\infty}+1-\mu)^2}=\frac{(a+1)x_{b,\infty}^2+a(1+\lambda)x_{b,\infty}}{(x_{b,\infty}+1+\lambda)^2}. \tag{6.16}$$

To illustrate method, we only consider $a>0$ in which case $x_{s,\infty}<x_{b,\infty}<0$ and $x_{b,\infty}\geqslant(1+\lambda)x_M=-\frac{a}{a+2}(1+\lambda)$ by Theorem 4.7. So

$$(a+1)x_{b,\infty}+a(1+\lambda)\geqslant(a+2)x_{b,\infty}+a(1+\lambda)\geqslant 0.$$

Then, it follows from (6.16) that $(a+1)x_{s,\infty}+a(1-\mu)\geqslant 0$, namely

$$x_{s,\infty}\geqslant-\frac{a}{a+1}(1-\mu).$$

By Theorem 4.5, we know

$$x_{s,\infty}\leqslant x_M(1-\mu)=-\frac{a}{a+2}(1-\mu).$$

Therefore $x_{s,\infty}$ must take the form of

$$x_{s,\infty}=-\frac{a}{a+k}(1-\mu), \tag{6.17}$$

where $k \in [1,2]$ is to be determined. Substituting this into (6.15), we have

$$f(x_{s,\infty}) = \frac{(a+1)(\frac{a}{a+k})^2(1-\mu)^2 - a\frac{a}{a+k}(1-\mu)^2}{(-\frac{a}{a+k}(1-\mu)+1-\mu)^2} = -\frac{k-1}{k^2}a^2. \tag{6.18}$$

In virtue of (6.16) and (6.18), $x_{b,\infty}$ satisfies the quadratic equation

$$(a+1)x_{b,\infty}^2 + a(1+\lambda)x_{b,\infty} = -\frac{k-1}{k^2}a^2(x_{b,\infty}+1+\lambda)^2,$$

which has two solutions

$$-\frac{a}{a+\frac{k}{k-1}}(1+\lambda) \quad \text{or} \quad -\frac{a}{a+k}(1+\lambda).$$

Because the latter is less than $x_{s,\infty} = -\frac{a}{a+k}(1-\mu)$ (remember $x_{s,\infty} < x_{b,\infty}$), we choose the former, namely,

$$x_{b,\infty} = -\frac{a}{a+\frac{k}{k-1}}(1+\lambda). \tag{6.19}$$

Now we come back to (6.14). Due to (6.18), we are able to rewrite (6.14) as

$$-\frac{x^2}{1-\gamma}\left(v'_\infty + \gamma v_\infty^2\right) + axv_\infty = -\frac{k-1}{k^2}a^2, \quad x \in (x_{s,\infty}, x_{b,\infty}). \tag{6.20}$$

This is a Riccati equation which has a special solution $\frac{C}{x}$. Here C satisfies the quadratic algebraic equation

$$\frac{\gamma}{1-\gamma}C^2 - \left(a + \frac{1}{1-\gamma}\right)C - \frac{k-1}{k^2}a^2 = 0,$$

solving which we get (6.9) (the smaller root is chosen).

Then, the general solution of the Riccati equation (6.20) in $(x_{s,\infty}, x_{b,\infty})$ is

$$v_\infty(x) = \frac{C}{x} + \frac{1}{g(x)}, \tag{6.21}$$

where $g(x)$ satisfies

$$g'(x) + \beta\frac{g(x)}{x} - \gamma = 0$$

or

$$\left(x^\beta g(x)\right)' = \gamma x^\beta. \tag{6.22}$$

Integrating the equation, we obtain

$$g(x) = \begin{cases} (\frac{x_{s,\infty}}{x})^\beta(g(x_{s,\infty}) - \frac{\gamma}{\beta+1}x_{s,\infty}) + \frac{\gamma}{\beta+1}x & \text{if } \beta \neq -1, \\ x(\frac{g(x_{s,\infty})}{x_{s,\infty}} + \gamma\log\frac{x}{x_{s,\infty}}) & \text{if } \beta = -1. \end{cases} \tag{6.23}$$

Let us focus on $\beta \neq -1$. At $x = x_{b,\infty}$,

$$g(x_{b,\infty}) = \left(\frac{x_{s,\infty}}{x_{b,\infty}}\right)^{\beta}\left(g(x_{s,\infty}) - \frac{\gamma}{\beta+1}x_{s,\infty}\right) + \frac{\gamma}{\beta+1}x_{b,\infty},$$

or equivalently,

$$\left(\frac{\frac{\gamma}{\beta+1} - \frac{g(x_{s,\infty})}{x_{s,\infty}}}{\frac{\gamma}{\beta+1} - \frac{g(x_{b,\infty})}{x_{b,\infty}}}\right)^{\frac{1}{\beta+1}} = \frac{x_{b,\infty}}{x_{s,\infty}}. \tag{6.24}$$

Thanks to (6.21) and boundary conditions (6.2) and (6.3), we have

$$\frac{g(x_{s,\infty})}{x_{s,\infty}} = \frac{1}{\frac{x_{s,\infty}}{x_{s,\infty}+1-\mu} - C}, \tag{6.25}$$

$$\frac{g(x_{b,\infty})}{x_{b,\infty}} = \frac{1}{\frac{x_{b,\infty}}{x_{b,\infty}+1+\lambda} - C}. \tag{6.26}$$

Substitution of (6.25) into (6.23) gives (6.4). Combining (6.24)–(6.26) with (6.17) and (6.19), we obtain (6.10).

When $\beta = -1$, we can obtain (6.5) and (6.11) using a similar argument. The proof is complete. $\square$

7. Conclusion

In this paper, we study the optimal investment problem for a CRRA investor who faces a finite horizon and transaction costs. From the angle of stochastic control, it is a singular control problem, whose value function is governed by a time-dependent HJB equation with gradient constraints. Using an elegant transformation, we reveal that the problem is equivalent to a parabolic double obstacle problem involving two free boundaries which correspond to the optimal buying and selling boundaries, respectively. This enables us to make use of the well-developed theory of obstacle problem to attack the problem. The $C^{2,1}$ regularity of the value function is proven.

Another purpose of the paper is to characterize the free boundaries (i.e. optimal investment policies). Relying on the double obstacle problem, the behaviors of the free boundaries can be completely characterized. In addition to the results obtained by Liu and Loewenstein [16], we show that the free boundaries increase with time and their behaviors depend sensitively on the relative magnitude of $\alpha - r$ and $(1-\gamma)\sigma^2$. When the maturity goes to infinity, the asymptotic behaviors of the free boundaries are determined by the solution of a Riccati equation with free boundary conditions for which a semi-explicit solution is gained.

Our approach can be generalized to a larger class of problems. For example, it can be extended to including the consumption term (see Dai et al. [5]). Also, the approach can be used to deal with the infinite-horizon problem discussed by Davis and Norman [6] and Shreve and Soner [21]. However, it seems to us that it is not straightforward to extend our approach to the case of multiple risky assets or general utility functions, and more efforts should be made to tackle a general setting.

References

[1] M. Akian, J.L. Menaldi, A. Sulem, On an investment–consumption model with transaction costs, SIAM J. Control Optim. 34 (1996) 329–364.

[2] V.E. Benes, L.A. Shepp, H.S. Witsenhausen, Some solvable stochastic control problems, Stochastics 4 (1980) 39–83.

[3] B. Bouchard, Utility maximization on the real line under proportional transaction costs, Finance Stoch. 6 (4) (2002) 495–516.

[4] J. Cvitanic, I. Karatzas, Hedging and portfolio optimization under transaction costs: A martingale approach, Math. Finance 6 (2) (1996) 133–165.

[5] M. Dai, L.S. Jiang, P.F. Li, F.H. Yi, Finite horizon optimal investment and consumption with transaction costs, Working paper, National University of Singapore, 2007.
[6] M.H.A. Davis, A.R. Norman, Portfolio selection with transaction costs, Math. Methods Oper. Res. 15 (1990) 676–713.
[7] M.H.A. Davis, V.G. Panas, T. Zariphopoulou, European option pricing with transaction costs, SIAM J. Control Optim. 31 (2) (1993) 470–493.
[8] G. Fichera, Sulle equazioni differenziali lineari ellittico-paraboliche del secondo ordine, Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez. I 5 (1956) 1–30.
[9] W.H. Fleming, H.M. Soner, Controlled Markov Processes and Viscosity Solutions, second ed., Springer-Verlag, New York, 2006.
[10] A. Friedman, Parabolic variational inequality in one space dimension and smoothness of the free boundary, J. Funct. Anal. 18 (3) (1975) 151–176.
[11] A. Friedman, Variational Principles and Free-Boundary Problems, Wiley, New York, 1982.
[12] P. Guasoni, Optimal investment with transaction costs and without semimartingales, Ann. Appl. Probab. 12 (4) (2002) 1227–1246.
[13] P. Guasoni, W. Schachermayer, Necessary conditions for the existence of utility maximizing strategies under transaction costs, Statist. Decisions 22 (2) (2004) 153–170.
[14] Y. Kabanov, C. Kluppelberg, A geometric approach to portfolio optimization in models with transaction costs, Finance Stoch. 8 (2004) 207–227.
[15] I. Karatzas, S.E. Shreve, Connections between optimal stopping and singular stochastic control I: Monotone follower problems, SIAM J. Control Optim. 22 (6) (1984) 856–877.
[16] H. Liu, M. Loewenstein, Optimal portfolio selection with transaction costs and finite horizons, Rev. Finan. Stud. 15 (3) (2002) 805–835.
[17] M.J.P. Magill, G.M. Constantinides, Portfolio selection with transaction costs, J. Econom. Theory 13 (1976) 264–271.
[18] T. Matsuzawa, On some degenerate parabolic equations II, Nagoya Math. J. 52 (1973) 61–84.
[19] R.C. Merton, Optimal consumption and portfolio rules in a continuous time model, J. Econom. Theory 3 (1971) 373–413.
[20] A.N. Shiryaev, Optimal Stopping Rules, Springer-Verlag, New York, 1978.
[21] S.E. Shreve, H.M. Soner, Optimal investment and consumption with transaction costs, Ann. Appl. Probab. 4 (1994) 609–692.