

OPTIMAL ENTRY AND CONSUMPTION UNDER HABIT FORMATION

YUE YANG^{* **} AND
XIANG YU ^{* ***} *The Hong Kong Polytechnic University*

Abstract

This paper studies a composite problem involving decision-making about the optimal entry time and dynamic consumption afterwards. In Stage 1, the investor has access to full market information subject to some information costs and needs to choose an optimal stopping time to initiate Stage 2; in Stage 2, the investor terminates the costly full information acquisition and starts dynamic investment and consumption under partial observation of free public stock prices. Habit formation preferences are employed, in which past consumption affects the investor's current decisions. Using the stochastic Perron method, the value function of the composite problem is proved to be the unique viscosity solution of some variational inequalities.

Keywords: Optimal entry problem; consumption habit formation; stochastic Perron method; viscosity solution

2020 Mathematics Subject Classification: Primary 91G10; 49L25
Secondary 93E11; 93E20; 49L20

1. Introduction

We consider a simple model to incorporate information costs in a continuous-time portfolio-consumption problem. In particular, we study a two-stage composite problem under complete and incomplete filtrations sequentially. The drift process of the stock price is assumed to be of Ornstein–Uhlenbeck type. In the first stage from the initial time, the investor needs to pay information costs to access the full market information generated by both drift and stock price processes, in order to update their dynamic distributions and decide the optimal time to enter the second stage. The information costs may include search costs, storage costs, communication costs, the cost of the investor's attention, or other service costs. In the present paper we consider a simple linear information cost, which is modeled by a constant cost rate and will be subtracted directly from the amount of the investor's wealth. That is, the longer the first stage is, the higher the information costs the investor needs to be able to afford. Some previous works have addressed the impact of information costs on optimal investment from different perspectives; see [18], [29], [1], and [20]. In our first stage, the mathematical problem becomes an optimal stopping problem under the complete market information filtration. The second stage starts from the chosen entry time, when the investor terminates the full observation of the drift

Received 18 February 2020; revision received 30 June 2021.

* Postal address: Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.

** Email address: yue.yy.yang@connect.polyu.hk

*** Email address: xiang.yu@polyu.edu.hk

© The Author(s), 2022. Published by Cambridge University Press on behalf of Applied Probability Trust.

process. From this point on, the investor instead dynamically chooses his investment and consumption levels based on the prior data inputs and free partial observation of the stock price, which can be formulated as an optimal control problem under an incomplete information filtration. As the value function of the interior control problem depends on the stopping time and data inputs of the drift process, the exterior problem can be interpreted as that of waiting in an optimal way so that the input values can achieve the maximum of the interior functional.

Portfolio optimization under partial observation has been extensively studied in past decades; a few examples with different financial motivations can be seen in [22, 34, 8, 25, 6, 7]. As illustrated in these works, the value function under the incomplete information filtration is strictly lower than its counterpart under the full information filtration, and this gap is usually regarded as the value of information. The present paper attempts to study partial observation from a different perspective, where the full market information is available but costly because more data, services, and personal attention are involved. The information costs may change the investor's attitude towards the usage of full observation, because it is no longer true that the more information he observes, the higher profit he can attain. Moreover, from previous work on partial observation, we know that the value function eventually depends on the given initial input of the random factor such as the drift process. As in [22, 8], it is conventionally assumed that the initial data of the unobservable drift is a Gaussian random variable, so that the Kalman–Bucy filtering can be applied. We take this input into account and consider a model in which the investor can wait and dynamically update the distribution of inputs using the full market information, subject to information costs. We show that starting sharp from the initial time to invest and consume under incomplete information may not be optimal.

On the other hand, in recent years, habit formation has provided a new paradigm for modeling consumption rate preferences, which better matches some empirical observations; see [11, 24]. The literature suggests that past consumption patterns may have a continuing impact on an individual's current consumption decisions. In particular, the use of linear habit formation preferences, in which there exists an index term that stands for the accumulated consumption history, has been widely accepted. Habit formation preferences have been well studied by [12, 14, 26] in complete market models and by [35, 36] in incomplete market models. It is noted that the utility function is decreasing in the habit level. In the present paper, we assume that there is no consumption during Stage 1, and the investor starts to gain consumption habits only in Stage 2. Therefore, an early entry time to Stage 2 may not be the optimal decision, because the investor has a longer time to develop a much higher habit level. This is our second motivation for investigating the exterior optimal entry time problem: to see whether waiting longer can benefit the investor more, as the resulting habit level may be much lower and may lead to a higher value function.

We show that the value function of the composite problem is the unique viscosity solution to some variational inequalities. To this end, we can choose to apply either the classical Perron method or the stochastic version of the Perron method introduced in [2]. For the classical Perron method, to establish the equivalence between the value function and the viscosity solution, we have to either prove the dynamic programming principle or upgrade the global regularity of the solution and prove the verification theorem. The convexity (concavity) of the value function with respect to the state variable is usually crucial in standard arguments to conclude global regularity. However, this property is not clear in our composite problem; see Remark 4.1 for details. The global regularity of the value function is not guaranteed, and so the direct verification proof for our exterior problem becomes difficult. Therefore, we instead choose the stochastic Perron method, which allows us to show the equivalence between the

value function and the viscosity solution without global regularity. For some related works on optimal stopping using viscosity solutions, we refer to [31] and [27]. Recent works on stochastic control problems using the stochastic Perron method include [2, 3, 4, 5, 33, 23]. One important step in completing the argument for the stochastic Perron method is the comparison principle for the associated variational inequalities, which is established in the present paper.

The rest of the paper is organized as follows: Section 2 introduces the market model and the habit formation preferences and formulates the two-stage optimization problem. Section 3 gives the main result on the interior utility maximization problem with habit formation and partial observation. Section 4 studies the exterior optimal entry problem with linear information costs. Using the stochastic Perron method, we show that the value function of the composite problem is the unique viscosity solution of some variational inequalities. Some auxiliary results and proofs are reported in Appendices A and B.

2. Mathematical model and preliminaries

2.1. Market model

Given the probability space $(\Omega, \mathbb{F}, \mathbb{P})$ with full information filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ that satisfies the usual conditions, we consider the market with one risk-free bond and one risky asset over a finite time horizon $[0, T]$. It is assumed that the bond process satisfies $S_t^0 \equiv 1$, for $t \in [0, T]$, which amounts to the standard change of numéraire.

The stock price S_t satisfies

$$dS_t = \mu_t S_t dt + \sigma_S S_t dW_t, \quad 0 \leq t \leq T, \quad (2.1)$$

with $S_0 = s > 0$. Some empirical studies, such as [9, 10, 15, 30], have observed that the drift process of many risky assets follows the so-called mean-reverting diffusion. We also consider here that the drift process μ_t in (2.1) satisfies the Ornstein–Uhlenbeck stochastic differential equation (SDE) by

$$d\mu_t = -\lambda(\mu_t - \bar{\mu})dt + \sigma_\mu dB_t, \quad 0 \leq t \leq T. \quad (2.2)$$

Here, $(W_t)_{0 \leq t \leq T}$ and $(B_t)_{0 \leq t \leq T}$ are \mathcal{F}_t -adapted Brownian motions with correlation coefficient $\rho \in [-1, 1]$. For simplicity, the initial value μ_0 of the drift is a given constant. We assume that the market coefficients σ_S , λ , $\bar{\mu}$, and σ_μ are given nonnegative constants based on calibrations from historical data.

It is assumed that the investor starts with initial wealth $x(0) = x_0 > 0$ at time $t = 0$. Also, starting from the initial time $t = 0$, access to the full market information \mathcal{F}_t generated by W and B incurs information costs κt , where $\kappa > 0$ is the constant cost rate per unit time. As stated earlier, the information costs may include storage costs, search costs, communication costs, the cost of the investor's attention, or other service costs involved in fully observing the market information \mathcal{F}_t . Moreover, to simplify the mathematical problem, it is assumed that from $t = 0$ up to a chosen stopping time τ , the investor purely waits and updates the dynamic distributions of the processes μ_t and S_t ; he does not invest or consume at all. This assumption makes sense as long as the value of the optimal entry time τ is short in the model. The dynamic wealth process including the information costs at time t is simply given by a deterministic function $x(t) = x_0 - \kappa t$ for any $t \leq \tau$.

As the full market information filtration is costly, the investor needs to choose optimally an \mathcal{F}_t -adapted stopping time τ to terminate full information acquisition and enter the second stage. From the chosen stopping time τ , he switches to the partial observation filtration $\mathcal{F}_t^S = \mathcal{F}_\tau \vee \sigma(S_u : \tau \leq u \leq t)$ for $\tau \leq t \leq T$, which is the union of the sigma-algebra \mathcal{F}_τ and the natural filtration generated by the stock price S up to time t . Moreover, for any time $\tau \leq t \leq T$, the investor chooses a dynamic consumption rate $c_t \geq 0$ and decides the amounts π_t of his wealth to invest in the risky asset (the rest is invested in the bond). Without paying information costs, he can no longer observe the drift process μ_t and Brownian motions W_t and B_t for $t \geq \tau$. Therefore, the investment–consumption pair (π_t, c_t) is only assumed to be adapted to the partial observation filtration \mathcal{F}_t^S for $\tau \leq t \leq T$. Recall that at the entry time τ , the investor only has wealth $x(\tau) = x_0 - \kappa\tau$ left. Under the incomplete filtration \mathcal{F}_t^S , the investor’s total wealth process \hat{X}_t can be written as

$$d\hat{X}_t = (\pi_t\mu_t - c_t)dt + \sigma_S\pi_t dW_t, \quad \tau \leq t \leq T, \tag{2.3}$$

with the initial value $\hat{X}_\tau = x(\tau) = x_0 - \kappa\tau > 0$. Note that W_t is no longer a Brownian motion under the partial observation filtration \mathcal{F}_t^S ; we have to apply the Kalman–Bucy filtering and consider the *innovation process* defined by

$$d\hat{W}_t := \frac{1}{\sigma_S}[(\mu_t - \hat{\mu}_t)dt + \sigma_S dW_t] = \frac{1}{\sigma_S} \left(\frac{dS_t}{S_t} - \hat{\mu}_t dt \right), \quad \tau \leq t \leq T,$$

which is a Brownian motion under \mathcal{F}_t^S . The best estimate of the unobservable drift process μ_t under \mathcal{F}_t^S is the conditional expectation process $\hat{\mu}_t = \mathbb{E}[\mu_t | \mathcal{F}_t^S]$, for $\tau \leq t \leq T$ with the initial input $\hat{\mu}_\tau = \mu_\tau$, \mathbb{P} -almost surely (a.s.), at the stopping time τ where the distribution of μ_τ is determined via (2.2) by paying information costs up to τ . By standard Kalman–Bucy filtering (see Equation (18) of [8] or Equation (21) of [25]), $\hat{\mu}_t$ satisfies the SDE

$$d\hat{\mu}_t = -\lambda(\hat{\mu}_t - \bar{\mu})dt + \left(\frac{\hat{\Sigma}(t) + \sigma_S\sigma_\mu\rho}{\sigma_S} \right) d\hat{W}_t, \quad \tau \leq t \leq T, \tag{2.4}$$

with $\hat{\mu}_\tau = \mu_\tau$, \mathbb{P} -a.s. Moreover, the conditional variance $\hat{\Sigma}(t) = \mathbb{E}[(\mu_t - \hat{\mu}_t)^2 | \mathcal{F}_t^S]$ satisfies the deterministic Riccati ordinary differential equation (ODE) (see Equation (19) of [8] or Equation (23) of [25])

$$\frac{d\hat{\Sigma}(t)}{dt} = -\frac{1}{\sigma_S^2}\hat{\Sigma}^2(t) + \left(-\frac{2\sigma_\mu\rho}{\sigma_S} - 2\lambda \right)\hat{\Sigma}(t) + (1 - \rho^2)\sigma_\mu^2, \quad \tau \leq t \leq T, \tag{2.5}$$

with the initial value $\hat{\Sigma}(\tau) = \mathbb{E}[(\mu_\tau - \hat{\mu}_\tau)^2 | \mathcal{F}_\tau^S] = 0$, given that $\hat{\mu}_\tau = \mu_\tau$, \mathbb{P} -a.s. This can be solved explicitly by

$$\hat{\Sigma}(t) = \sqrt{k}\sigma_S \frac{k_1 \exp\left(2\left(\frac{\sqrt{k}}{\sigma_S}\right)t\right) + k_2}{k_1 \exp\left(2\left(\frac{\sqrt{k}}{\sigma_S}\right)t\right) - k_2} - \left(\lambda + \frac{\sigma_\mu\rho}{\sigma_S}\right)\sigma_S^2, \quad \tau \leq t \leq T,$$

where $k = \lambda^2\sigma_S^2 + 2\sigma_S\sigma_\mu\lambda\rho + \sigma_\mu^2$, $k_1 = \sqrt{k}\sigma_S + (\lambda\sigma_S^2 + \sigma_S\sigma_\mu\rho)$, and $k_2 = -\sqrt{k}\sigma_S + (\lambda\sigma_S^2 + \sigma_S\sigma_\mu\rho)$.

For the second-stage dynamic control problem, we employ habit formation preferences. In particular, we define $Z_t := Z(c_t)$ as the *habit formation* process or the *standard of living* process, which describes the level of the investor’s consumption habits. It is conventionally assumed that the cumulative preference Z_t satisfies the recursive equation $dZ_t = (\delta(t)c_t - \alpha(t)Z_t)dt$, $\tau \leq t \leq T$ (see [12]), where $Z_\tau = z_0 \geq 0$ is called the *initial consumption habit* of the investor. Equivalently, we have

$$Z_t = z_0 e^{-\int_\tau^t \alpha(u)du} + \int_\tau^t \delta(u) e^{-\int_u^t \alpha(s)ds} c_u du, \quad \tau \leq t \leq T,$$

which is the exponentially weighted average of the initial habit and the past consumption. Here, the deterministic discount factors $\alpha(t) \geq 0$ and $\delta(t) \geq 0$ measure, respectively, the persistence of the past level and the intensity of the consumption history. In the present paper we are interested in *addictive habits*; namely, we require that the investor’s current consumption strategies never fall below the level of standard of living, i.e. $c_t \geq Z_t$ a.s., for $\tau \leq t \leq T$.

Under the partial observation filtration $(\mathcal{F}_t^S)_{\tau \leq t \leq T}$, the stock price dynamics (2.1) can be rewritten as $dS_t = \hat{\mu}_t S_t dt + \sigma_S S_t d\hat{W}_t$, and the wealth dynamics (2.3) can be rewritten as $d\hat{X}_t = (\pi_t \hat{\mu}_t - c_t)dt + \sigma_S \pi_t d\hat{W}_t$, $\tau \leq t \leq T$. To facilitate the formulation of the stochastic control problem and the derivation of the dynamic programming equation, for any $t \in [0, T]$, we denote by $\mathcal{A}_t(x)$ the time-modulated admissible set of the pair of investment and consumption processes $(\pi_s, c_s)_{t \leq s \leq T}$ with the initial wealth $\hat{X}_t = x$, which is \mathcal{F}_t^S -progressively measurable and satisfies the integrability conditions $\int_t^T \pi_s^2 ds < +\infty$, \mathbb{P} -a.s., and $\int_t^T c_s ds < +\infty$, \mathbb{P} -a.s., with the addictive habit formation constraint that $c_s \geq Z_s$, \mathbb{P} -a.s., $t \leq s \leq T$. Moreover, no bankruptcy is allowed, i.e., the investor’s wealth remains nonnegative: $\hat{X}_s \geq 0$, \mathbb{P} -a.s., $t \leq s \leq T$.

2.2. Problem formulation

The two-stage optimal decision-making problem is formulated as a composite problem involving the optimal stopping and the stochastic control afterwards, which is defined by

$$\tilde{V}(0, \mu_0; x_0, z_0) := \sup_{\tau \geq 0} \mathbb{E} \left[\operatorname{esssup}_{(\pi, c) \in \mathcal{A}_\tau(x_0 - \kappa\tau)} \mathbb{E} \left[\int_\tau^T \frac{(c_s - Z_s)^p}{p} ds \middle| \mathcal{F}_\tau^S \right] \right]. \tag{2.6}$$

In particular, starting from the chosen stopping time τ , we are interested in utility maximization on consumption with habit formation, in which the power utility function $U(x) = x^p/p$ is defined on the difference $c_t - Z_t$. To simplify the presentation, in the present paper we only consider the case of a risk aversion coefficient $p < 0$. The indirect utility process of the interior control problem is denoted by

$$\begin{aligned} \hat{V}(t, x_0 - \kappa t, z_0, \mu_t; 0) &:= \operatorname{esssup}_{(\pi, c) \in \mathcal{A}_t(x_0 - \kappa t)} \mathbb{E} \left[\int_t^T \frac{(c_s - Z_s)^p}{p} ds \middle| \mathcal{F}_t^S \right] \\ &= \operatorname{esssup}_{(\pi, c) \in \mathcal{A}_t(x_0 - \kappa t)} \mathbb{E} \left[\int_t^T \frac{(c_s - Z_s)^p}{p} ds \middle| \hat{X}_t = x_0 - \kappa t, \hat{\mu}_t = \mu_t, Z_t = z_0; \hat{\Sigma}(t) = 0 \right]. \end{aligned}$$

To determine the exterior optimal stopping time, we need to maximize over the inputs of the values τ , \hat{X}_τ , Z_τ , and $\hat{\mu}_\tau$. Recall that the investor does not manage his investment and consumption before τ ; it follows that $\hat{X}_\tau = x_0 - \kappa\tau$, $Z_\tau = z_0$, and $\hat{\Sigma}(\tau) = 0$ can all be taken as parameters instead of variables. That is, $\mu_\tau = \hat{\mu}_\tau$ is the only random input, and we can

regard μ_t as the only underlying state process. Therefore, the dynamic counterpart of (2.6) is defined by

$$\tilde{V}(t, \eta; x_0 - \kappa t, z_0) := \operatorname{esssup}_{\tau \geq t} \mathbb{E} \left[\operatorname{esssup}_{(\pi, c) \in \mathcal{A}_\tau(x_0 - \kappa \tau)} \mathbb{E} \left[\int_\tau^T \frac{(c_s - Z_s)^p}{p} ds \middle| \mathcal{F}_\tau^S \right] \middle| \mu_t = \eta \right]. \tag{2.7}$$

Remark 2.1. We focus on the case $p < 0$ in the present paper because then the functions $A(t, s)$, $B(t, s)$, and $C(t, s)$ introduced later, which are solutions to the ODEs (3.4), (3.5), and (3.6), are all bounded, and the utility $U(x)$ is also bounded from above, which significantly simplifies the proof of the verification result in Theorem 3.1 and the comparison results in Proposition 4.1. The other case, $0 < p < 1$, can essentially be handled in a similar way. However, in that case, as the process $\hat{\mu}_t$ in (2.4) is unbounded and the functions $A(t, s)$, $B(t, s)$, and $C(t, s)$ may explode at some $t \in [0, T]$, one needs some additional parameter assumptions to guarantee integrability conditions and martingale properties in the proofs of some of the main results.

Assumption 2.1. *In accordance with Remark 3.1 for the interior control problem, it is assumed from this point onwards that $x_0 - \kappa t > z_0 m(t)$ for any $0 \leq t \leq T$; i.e. the initial wealth is sufficiently large, after paying information costs, so that the interior control problem is well defined for any $0 \leq t \leq T$, where $m(t)$ is defined by*

$$m(t) = \int_t^T \exp \left(\int_t^s (\delta(v) - \alpha(v)) dv \right) ds, \quad 0 \leq t \leq T. \tag{2.8}$$

We note that $m(t)$ in (2.8) represents the cost of subsistence consumption per unit of standard of living at time t , because the interior control problem is solvable if and only if $\hat{X}_t^* \geq m(t)Z_t$, $0 \leq t \leq T$; see Lemma B.1.

The function \hat{V} can be solved in the explicit form given in (3.7) later. The process $\tilde{V}(t, \mu_t; x_0 - \kappa t, z_0)$ with the function \tilde{V} defined in (2.7) is the Snell envelope of the process $\hat{V}(t, x_0 - \kappa t, z_0, \mu_t)$ above. The function \tilde{V} in (2.7) can therefore be written as

$$\tilde{V}(t, \eta; x_0 - \kappa t, z_0) = \operatorname{esssup}_{\tau \geq t} \mathbb{E} \left[\hat{V}(\tau, x_0 - \kappa \tau, z_0, \mu_\tau) \middle| \mu_t = \eta \right].$$

The continuation region, interpreted as the region where the investor continues to use full information observations to update the input value, is denoted by $\mathcal{C} = \{(t, \eta) \in [0, T] \times \mathbb{R} : \tilde{V}(t, \eta; x_0 - \kappa t, z_0) > \hat{V}(t, x_0 - \kappa t, z_0, \eta)\}$, and the free boundary is $\partial \mathcal{C} = \{(t, \eta) \in [0, T] \times \mathbb{R} : \tilde{V}(t, \eta; x_0 - \kappa t, z_0) = \hat{V}(t, x_0 - \kappa t, z_0, \eta)\}$. Let us denote $\tilde{V}(t, \eta; x_0 - \kappa t, z_0)$ by $\tilde{V}(t, \eta)$ for short when there is no possibility of confusion. By heuristic arguments, we can write the Hamilton–Jacobi–Bellman (HJB) variational inequalities with the terminal condition $\tilde{V}(T, \eta) = 0, \eta \in \mathbb{R}$, as

$$\min \left\{ \tilde{V}(t, \eta) - \hat{V}(t, x_0 - \kappa t, z_0, \eta), \quad -\frac{\partial \tilde{V}(t, \eta)}{\partial t} - \mathcal{L}\tilde{V}(t, \eta) \right\} = 0, \tag{2.9}$$

where

$$\mathcal{L}\tilde{V}(t, \eta) = -\lambda(\eta - \bar{\mu}) \frac{\partial \tilde{V}}{\partial \eta}(t, \eta) + \frac{1}{2} \sigma_\mu^2 \frac{\partial^2 \tilde{V}}{\partial \eta^2}(t, \eta).$$

To simplify notation in the following sections, we shall rewrite (2.9) as

$$\begin{cases} F(t, \eta, \tilde{V}, \frac{\partial \tilde{V}}{\partial t}, \frac{\partial \tilde{V}}{\partial \eta}, \frac{\partial^2 \tilde{V}}{\partial \eta^2}) = 0, & \text{on } [0, T] \times \mathbb{R}, \\ v(T, \eta) = 0, & \text{for } \eta \in \mathbb{R}, \end{cases} \tag{2.10}$$

with the operator $F(t, \eta, v, v_t, v_\eta, v_{\eta\eta}) := \min \{ v - \hat{V}, \quad -\frac{\partial v}{\partial t} - \mathcal{L}v \}$.

Remark 2.2. The part $-\frac{\partial \tilde{V}}{\partial t} - \mathcal{L}\tilde{V} = 0$ in (2.9) is a linear parabolic partial differential equation (PDE) and does not depend on the interior control (π, c) . The comparison part $\tilde{V} - \hat{V}$ in (2.9) depends on the optimal control (π, c) because the \hat{V} is the value function of the interior control problem provided the input $\hat{X}_t = x_0 - \kappa t$, $Z_t = z_0$, and $\hat{\mu}_t = \mu_t = \eta$.

The next theorem is the main result of this paper.

Theorem 2.1. *The quantity $\tilde{V}(t, \eta)$ defined in (2.7) is the unique bounded and continuous viscosity solution to the variational inequalities (2.9). In addition, the optimal entry time for the composite problem (2.7) is given by the \mathcal{F}_t -adapted stopping time*

$$\tau^* := T \wedge \inf \{t \geq 0 : \tilde{V}(t, \mu_t; x_0 - \kappa t, z_0) = \hat{V}(t, x_0 - \kappa t, z_0, \mu_t)\}. \tag{2.11}$$

We also have that the process $\tilde{V}(t, \mu_t; x_0 - \kappa t, z_0)$ is a martingale with respect to the full information filtration \mathcal{F}_t for $0 \leq t \leq \tau^*$.

The proof will be provided in Section 4.

3. Interior utility maximization under partial observation

We first solve the interior stochastic control problem under partial observation of stock prices.

3.1. Optimal consumption with Kalman–Bucy filtering

For some fixed time $0 \leq k \leq T$, the dynamic interior stochastic control problem under habit formation is defined by

$$\begin{aligned} \hat{V}(k, x, z, \eta; \theta) &:= \sup_{(\pi, c) \in \mathcal{A}_k(x)} \mathbb{E} \left[\int_k^T \frac{(c_s - Z_s)^p}{p} ds \middle| \mathcal{F}_k^S \right] \\ &= \sup_{(\pi, c) \in \mathcal{A}_k(x)} \mathbb{E} \left[\int_k^T \frac{(c_s - Z_s)^p}{p} ds \middle| \hat{X}_k = x, Z_k = z, \hat{\mu}_k = \eta; \hat{\Sigma}(k) = \theta \right], \end{aligned} \tag{3.1}$$

where $\mathcal{A}_k(x)$ denotes the admissible control space starting from time k . Here, as the conditional variance $\hat{\Sigma}(t)$ is a deterministic function of time, we set θ as a parameter instead of a state variable.

By using the optimality principle and Itô’s formula, we can heuristically obtain the HJB equation as

$$\begin{aligned} V_t - \alpha(t)zV_z - \lambda(\eta - \bar{\mu})V_\eta + \frac{(\hat{\Sigma}(t) + \sigma_S \sigma_\mu \rho)^2}{2\sigma_S^2} V_{\eta\eta} + \max_{(\pi, c) \in \mathcal{A}} \left[-cV_x + c\delta(t)V_z + \frac{(c - z)^p}{p} \right] \\ + \max_{(\pi, c) \in \mathcal{A}} \left[\pi \eta V_x + \frac{1}{2} \sigma_S^2 \pi^2 V_{xx} + V_{x\eta} (\hat{\Sigma}(t) + \sigma_S \sigma_\mu \rho) \pi \right] = 0, \quad k \leq t \leq T, \end{aligned} \tag{3.2}$$

with the terminal condition $V(T, x, z, \eta) = 0$.

3.2. The decoupled solution and main results

If $V(t, x, z, \eta)$ is smooth enough, the first-order condition gives

$$\pi^*(t, x, z, \eta) = \frac{-\eta V_x - (\hat{\Sigma}(t) + \sigma_S \sigma_{\mu\rho}) V_{x\eta}}{\sigma_S^2 V_{xx}},$$

$$c^*(t, x, z, \eta) = z + (V_x - \delta(t) V_z)^{\frac{1}{p-1}}.$$

Thanks to the homogeneity property of the power utility, we conjecture the value function to have the form

$$V(t, x, z, \eta) = \frac{[(x - m(t, \eta)z)]^p}{p} N^{1-p}(t, \eta)$$

for some functions $m(t, \eta)$ and $N(t, \eta)$ to be determined. It also follows that the terminal condition that $N(T, \eta) = 0$ is required. In particular, we find that the simple ansatz of $m(t, \eta) := m(t)$ satisfies the equation (2.8). After substitution, the HJB equation reduces to the linear parabolic PDE for $N(t, \eta)$ as

$$N_t + \frac{p\eta^2}{2(1-p)^2\sigma_S^2} N(t, \eta) + \frac{(\hat{\Sigma}(t) + \sigma_S \sigma_{\mu\rho})^2}{2\sigma_S^2} N_{\eta\eta} + (1 + \delta(t)m(t))^{\frac{p}{p-1}}$$

$$+ \left[-\lambda(\eta - \bar{\mu}) + \frac{\eta(\hat{\Sigma}(t) + \sigma_S \sigma_{\mu\rho})p}{(1-p)\sigma_S^2} \right] N_{\eta}(t, \eta) = 0,$$

with $N(T, \eta) = 0$. We can further solve the linear PDE explicitly by

$$N(t, \eta) = \int_t^T (1 + \delta(s)m(s))^{\frac{p}{p-1}} \exp(A(t, s)\eta^2 + B(t, s)\eta + C(t, s)) ds, \tag{3.3}$$

for $k \leq t \leq s \leq T$. $A(t, s)$, $B(t, s)$ and $C(t, s)$ satisfy the following ODEs:

$$A_t(t, s) + \frac{p}{2(1-p)^2\sigma_S^2} + 2 \left[-\lambda + \frac{p(\hat{\Sigma}(t) + \sigma_S \sigma_{\mu\rho})}{\sigma_S^2(1-p)} \right] A(t, s)$$

$$+ \frac{2(\hat{\Sigma}(t) + \sigma_S \sigma_{\mu\rho})^2}{\sigma_S^2} A^2(t, s) = 0, \tag{3.4}$$

$$B_t(t, s) + \left[-\lambda + \frac{p(\hat{\Sigma}(t) + \sigma_S \sigma_{\mu\rho})}{\sigma_S^2(1-p)} \right] B(t, s) + 2\lambda\bar{\mu}A(t, s)$$

$$+ \frac{2(\hat{\Sigma}(t) + \sigma_S \sigma_{\mu\rho})^2}{\sigma_S^2} A(t, s)B(t, s) = 0, \tag{3.5}$$

$$C_t(t, s) + \lambda\bar{\mu}B(t, s) + \frac{(\hat{\Sigma}(t) + \sigma_S \sigma_{\mu\rho})^2}{2\sigma_S^2} (B^2(t, s) + 2A(t, s)) = 0, \tag{3.6}$$

with terminal conditions $A(s, s) = B(s, s) = C(s, s) = 0$. The explicit solutions of the ODEs (3.4), (3.5), and (3.6) are reported in Appendix A. For fixed $t \in [k, T]$, we can define the

effective domain of the pair (x, z) by $\mathbb{D}_t := \{(x', z') \in (0, +\infty) \times [0, +\infty); x' \geq m(t)z'\}$, where $k \leq t \leq T$. The HJB equation (3.2) admits a classical solution on $[k, T] \times \mathbb{D}_t \times \mathbb{R}$, given by

$$V(t, x, z, \eta) = \left[\int_t^T (1 + \delta(s)m(s))^{\frac{p}{p-1}} \exp(A(t, s)\eta^2 + B(t, s)\eta + C(t, s)) ds \right]^{1-p} \times \frac{[(x - m(t)z)]^p}{p}. \tag{3.7}$$

Remark 3.1. The effective domain of $V(t, x, z, \eta)$ requires some constraints on the optimal wealth process \hat{X}_t^* and habit formation process Z_t^* such that $\hat{X}_t^* \geq m(t)Z_t^*$ for $t \in [k, T]$. In particular, we have to enforce the initial wealth-habit budget constraint that $\hat{X}_k \geq m(k)Z_k$ at the initial time k .

Theorem 3.1. (Verification theorem) *If the initial budget constraint $\hat{X}_k \geq m(k)Z_k$ holds at time k , the unique solution (3.7) of the HJB equation equals the value function defined in (3.1), i.e., $V(k, x, z, \eta) = \hat{V}(k, x, z, \eta)$. Moreover, the optimal investment policy π_t^* and optimal consumption policy c_t^* are given in feedback form by $\pi_t^* = \pi^*(t, \hat{X}_t^*, Z_t^*, \hat{\mu}_t)$ and $c_t^* = c^*(t, \hat{X}_t^*, Z_t^*, \hat{\mu}_t)$, $k \leq t \leq T$. The function $\pi^*(t, x, z, \eta) : [k, T] \times \mathbb{D}_t \times \mathbb{R} \rightarrow \mathbb{R}$ is given by*

$$\pi^*(t, x, z, \eta) = \left[\frac{\eta}{(1-p)\sigma_S^2} + \frac{(\hat{\Sigma}(t) + \sigma_S\sigma_\mu\rho) N_\eta(t, \eta)}{\sigma_S^2 N(t, \eta)} \right] (x - m(t)z), \tag{3.8}$$

and the function $c^*(t, x, z, \eta) : [k, T] \times \mathbb{D}_t \times \mathbb{R} \rightarrow \mathbb{R}^+$ is given by

$$c^*(t, x, z, \eta) = z + \frac{(x - m(t)z)}{(1 + \delta(t)m(t))^{1-p} N(t, \eta)}. \tag{3.9}$$

The optimal wealth process \hat{X}_t^* , $k \leq t \leq T$, is given by

$$\hat{X}_t^* = (x - m(k)z) \frac{N(t, \hat{\mu}_t)}{N(k, \eta)} \exp \left(\int_k^t \frac{(\hat{\mu}_u)^2}{2(1-p)\sigma_S^2} du + \int_k^t \frac{\hat{\mu}_u}{(1-p)\sigma_S} d\hat{W}_u \right) + m(t)Z_t^*. \tag{3.10}$$

4. Exterior optimal stopping problem

4.1. Stochastic Perron method

We next study the exterior optimal entry problem. Recall that $\hat{X}_\tau = x_0 - \kappa\tau$, $Z_\tau = z_0$, and $\hat{\Sigma}(\tau) = 0$ are all taken as parameters. Our aim is to solve an optimal stopping problem in which μ_t is the only underlying state process.

Remark 4.1. Recall that the interior value function \hat{V} is of the form in (3.7). Moreover, by Remark A.1, the functions $A(t, s)$ and $B(t, s)$ in (3.7) satisfy $A(t, s) \leq 0$ and $B(t, s) \leq 0$ since $p < 0$. That is, if we take $\hat{V}(\tau, \hat{\mu}_\tau)$ as a functional of the input $\hat{\mu}_\tau$, it is not globally convex or concave in $\hat{\mu}_\tau \in \mathbb{R}$, because the function $\exp(A(t, s)\eta^2 + B(t, s)\eta + C(t, s))$ is not globally convex or concave in the variable $\eta \in \mathbb{R}$, which depends on values of $A(t, s)$ and $B(t, s)$. Therefore, the composite value function $\tilde{V}(t, \eta)$ in (2.7) is not globally convex or concave in $\eta \in \mathbb{R}$, which actually depends on all model parameters.

We choose to apply the stochastic Perron method in the present paper to verify that the value function of the composite problem is the unique viscosity solution of some variational inequalities. We first introduce sets of stochastic semi-solutions \mathcal{V}^+ and \mathcal{V}^- and prove that $v^- \leq \tilde{V} \leq v^+$, where v^- and v^+ are defined later in (4.2) and (4.3). Using the stochastic Perron method, we can show that v^+ is a bounded and upper semi-continuous (u.s.c.) viscosity subsolution and v^- is a bounded and lower semi-continuous (l.s.c.) viscosity supersolution. Finally, we prove the comparison principle: that is, if we have any bounded and u.s.c. viscosity subsolution u and any bounded and l.s.c. viscosity supersolution v of (2.10), we must have the order $u \leq v$. It follows that $v^+ \leq v^-$, which leads to the desired conclusion that $v^- = \tilde{V} = v^+$ and the value function is the unique viscosity solution.

We next present the definitions of stochastic semi-solutions, which are mainly motivated by [4].

Definition 4.1. The set of stochastic supersolutions for the PDE (2.10), denoted by \mathcal{V}^+ , is the set of functions $v : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ which have the following properties:

- (i) The function v is u.s.c. and bounded on $[0, T] \times \mathbb{R}$, and $v(t, \eta) \geq \widehat{V}(t, x_0 - \kappa t, z_0, \eta)$ for any $(t, \eta) \in [0, T] \times \mathbb{R}$.
- (ii) For each $(t, \eta) \in [0, T] \times \mathbb{R}$ and any stopping time $t \leq \tau_1 \in \mathcal{T}$, we have $v(\tau_1, \mu_{\tau_1}) \geq \mathbb{E}[v(\tau_2, \mu_{\tau_2}) | \mathcal{F}_{\tau_1}]$, \mathbb{P} -a.s., for any $\tau_2 \in \mathcal{T}$ and $\tau_2 \geq \tau_1$. That is, the function v along the solution of the SDE (2.2) is a supermartingale under the full information filtration $(\mathcal{F}_t)_{t \in [0, T]}$ between τ_1 and T .

Definition 4.2. The set of stochastic subsolutions for the PDE (2.10), denoted by \mathcal{V}^- , is the set of functions $v : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ which have the following properties:

- (i) The function v is l.s.c. and bounded on $[0, T] \times \mathbb{R}$, and $v(T, \eta) \leq 0$ for any $\eta \in \mathbb{R}$.
- (ii) For each $(t, \eta) \in [0, T] \times \mathbb{R}$ and any stopping time $t \leq \tau_1 \in \mathcal{T}$, we have $v(\tau_1, \mu_{\tau_1}) \leq \mathbb{E}[v(\tau_2 \wedge \zeta, \mu_{\tau_2 \wedge \zeta}) | \mathcal{F}_{\tau_1}]$, \mathbb{P} -a.s., for any $\tau_2 \in \mathcal{T}$ and $\tau_2 \geq \tau_1$. That is, the function v along the solution to (2.2) is a submartingale under the full information filtration $(\mathcal{F}_t)_{t \in [0, T]}$ between τ_1 and ζ , where

$$\zeta := \inf \{t \in [\tau_1, T] : v(t, \mu_t; x_0 - \kappa t, z_0) \geq \widehat{V}(t, x_0 - \kappa t, z_0, \mu_t)\}. \tag{4.1}$$

Remark 4.1. We note that the definitions of stochastic supersolutions and stochastic subsolutions for the optimal stopping problem are not symmetric, which is consistent with the similar definitions in [4]. The main reason for these differences comes from the natural supermartingale property of the Snell envelop process and its martingale property between the initial time and the first hitting time ζ in (4.1). That is, we naturally need $v(t, \eta) \geq \widehat{V}(t, x_0 - \kappa t, z_0, \eta)$ for all $(t, \eta) \in [0, T] \times \mathbb{R}$, including the terminal time T , in item (i) of Definition 4.1 (the definition of stochastic supersolutions), but we only require $v(T, \eta) \leq \widehat{V}(T, x_0 - \kappa T, z_0, \eta) = 0$ at the terminal time T in item (i) of Definition 4.2 (the definition of stochastic subsolutions). These comparison results and the supermartingale and submartingale properties will play important roles in the establishment of the desired sandwich result $v^- \leq \tilde{V} \leq v^+$ in Lemma 4.4.

Lemma 4.1. $\widehat{V}(t, x_0 - \kappa t, z_0, \eta; 0)$ is bounded and continuous for $(t, \eta) \in [0, T] \times \mathbb{R}$.

Proof. For fixed x_0 and z_0 , it is clear that $\widehat{V}(t, x_0 - \kappa t, z_0, \eta)$ in (3.7) is continuous and $\widehat{V}(t, x_0 - \kappa t, z_0, \eta) \leq 0$. Therefore we only need to show that \widehat{V} is bounded below. By Appendix A, we know that $A(u) \leq 0$, $B(u) \leq 0$, and $C(u) \leq K$ for some $K \geq 0$, thanks to

$p < 0$. We hence obtain that $(A(u)\eta^2 + B(u)\eta + C(u)) \leq K_1$ for some $K_1 > 0$, and it follows that $\widehat{V}(t, x_0 - \kappa t, z_0, \eta)$ is bounded below by some constant for $(t, \eta) \in [0, T] \times \mathbb{R}$, again by $p < 0$. \square

As it is trivial to see that $0 \in \mathcal{V}^-$ and $0 \in \mathcal{V}^+$, we have the following result.

Lemma 4.2. \mathcal{V}^+ and \mathcal{V}^- are nonempty.

Definition 4.3. We define

$$v^- := \sup_{p \in \mathcal{V}^-} p, \tag{4.2}$$

$$v^+ := \inf_{q \in \mathcal{V}^+} q. \tag{4.3}$$

The next result is similar to Lemma 2.2 of [2].

Lemma 4.3. We have $v^- \in \mathcal{V}^-$ and $v^+ \in \mathcal{V}^+$.

We now have the first important sandwich result.

Lemma 4.4. We have $v^- \leq \widetilde{V} \leq v^+$.

Proof. For each $v \in \mathcal{V}^+$, let us consider $\tau_1 = t \geq 0$ in Definition 4.1. For any $\tau \geq t$, we have $v(t, \eta) \geq \mathbb{E}[v(\tau, \mu_\tau) | \mathcal{F}_t] \geq \mathbb{E}[\widehat{V}(\tau, x_0 - \kappa\tau, z_0, \mu_\tau) | \mathcal{F}_t]$ thanks to the supermartingale property in Definition 4.1. It follows that $v(t, \eta) \geq \text{esssup}_{t \leq \tau} \mathbb{E}[\widehat{V}(\tau, x_0 - \kappa\tau, z_0, \mu_\tau) | \mathcal{F}_t]$. This implies that $v(t, \eta) \geq \widetilde{V}(t, \eta)$ in view of the definition of $\widetilde{V}(t, \eta)$, and hence $\widetilde{V} \leq v^+$ by the definition in (4.3). On the other hand, for each $v \in \mathcal{V}^-$, by taking $\tau_1 = t \geq 0$ in Definition 4.2, we have $v(t, \eta) \leq \mathbb{E}[v(\tau \wedge \zeta, \mu_{\tau \wedge \zeta}) | \mathcal{F}_t]$ for any $\tau \geq t$ because of the submartingale property in Definition 4.2. In particular, using the definition of ζ , we further have

$$\begin{aligned} v(t, \eta) &\leq \mathbb{E}[v(\tau \wedge \zeta, \mu_{\tau \wedge \zeta}) | \mathcal{F}_t] \leq \mathbb{E}[\widehat{V}(\tau \wedge \zeta, x_0 - f(\tau \wedge \zeta), z_0, \mu_{\tau \wedge \zeta}) | \mathcal{F}_t] \\ &\leq \text{esssup}_{\tau \geq t} \mathbb{E}[\widehat{V}(\tau, x_0 - \kappa\tau, z_0, \mu_\tau) | \mathcal{F}_t] = \widetilde{V}(t, \eta). \end{aligned}$$

It then follows that $\widetilde{V} \geq v^-$ because of (4.2). In conclusion, we have the inequality $v^- \leq \widetilde{V} \leq v^+$. \square

Theorem 4.1. The function v^- in Definition 4.3 is a bounded and l.s.c. viscosity supersolution of

$$\begin{cases} F(t, \eta, v, v_t, v_\eta, v_{\eta\eta}) \geq 0, & \text{on } [0, T] \times \mathbb{R}, \\ v(T, \eta) \geq 0, & \text{for any } \eta \in \mathbb{R}, \end{cases} \tag{4.4}$$

and the function v^+ in Definition 4.3 is a bounded and u.s.c. viscosity subsolution of

$$\begin{cases} F(t, \eta, v, v_t, v_\eta, v_{\eta\eta}) \leq 0, & \text{on } [0, T] \times \mathbb{R}, \\ v(T, \eta) \leq 0, & \text{for any } \eta \in \mathbb{R}. \end{cases} \tag{4.5}$$

Proof. We follow some arguments from [2, 4], modifying them to fit our setting.

(i) *The subsolution property of v^+ .* First, the definition in (4.3) and Lemma 4.3 imply that v^+ is bounded and u.s.c. Suppose v^+ is not a viscosity subsolution; then there exists some interior point $(\bar{t}, \bar{\eta}) \in (0, T) \times \mathbb{R}$ and a $C^{1,2}$ test function $\varphi : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ such that $v^+ - \varphi$ attains a strict local maximum that is equal to zero and $F(\bar{t}, \bar{\eta}, v, v_{\bar{t}}, v_{\bar{\eta}}, v_{\bar{\eta}\bar{\eta}}) > 0$. It follows that

$$\begin{cases} v^+(\bar{t}, \bar{\eta}) - \widehat{V}(\bar{t}, x_0 - f(\bar{t}), z_0, \bar{\eta}) > 0, \\ -\frac{\partial \varphi}{\partial t}(\bar{t}, \bar{\eta}) - \mathcal{L}\varphi(\bar{t}, \bar{\eta}) > 0. \end{cases}$$

There exists a ball $B(\bar{t}, \bar{\eta}, \varepsilon)$ small enough that

$$\begin{cases} -\frac{\partial \varphi}{\partial t} - \mathcal{L}\varphi > 0 \text{ on } \overline{B(\bar{t}, \bar{\eta}, \varepsilon)}, \\ \varphi > v^+ \text{ on } \overline{B(\bar{t}, \bar{\eta}, \varepsilon)} \setminus (\bar{t}, \bar{\eta}). \end{cases}$$

In addition, as $\varphi(\bar{t}, \bar{\eta}) = v^+(\bar{t}, \bar{\eta}) > \widehat{V}(\bar{t}, x_0 - f(\bar{t}), z_0, \bar{\eta})$, φ is continuous, and \widehat{V} is continuous, we can derive that for some ε small enough, we have $\varphi - \varepsilon \geq \widehat{V}$ on $\overline{B(\bar{t}, \bar{\eta}, \varepsilon)}$. Because $v^+ - \varphi$ is u.s.c. and $\overline{B(\bar{t}, \bar{\eta}, \varepsilon)} \setminus B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})$ is compact, it then follows that there exists a $\delta > 0$ such that $\varphi - \delta \geq v^+$ on $\overline{B(\bar{t}, \bar{\eta}, \varepsilon)} \setminus B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})$.

If we choose $0 < \xi < \delta \wedge \varepsilon$, the function $\varphi^\xi = \varphi - \xi$ satisfies

$$\begin{cases} -\frac{\partial \varphi^\xi}{\partial t} - \mathcal{L}\varphi^\xi > 0 \text{ on } \overline{B(\bar{t}, \bar{\eta}, \varepsilon)}, \\ \varphi^\xi > v^+ \text{ on } \overline{B(\bar{t}, \bar{\eta}, \varepsilon)} \setminus B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2}), \\ \varphi^\xi \geq \widehat{V} \text{ on } \overline{B(\bar{t}, \bar{\eta}, \varepsilon)}, \end{cases}$$

and $\varphi^\xi(\bar{t}, \bar{\eta}) = v^+(\bar{t}, \bar{\eta}) - \xi$.

Let us define an auxiliary function by

$$v^\xi := \begin{cases} v^+ \wedge \varphi^\xi & \text{on } \overline{B(\bar{t}, \bar{\eta}, \varepsilon)}, \\ v^+ & \text{outside } \overline{B(\bar{t}, \bar{\eta}, \varepsilon)}. \end{cases}$$

It is easy to check that v^ξ is u.s.c. and $v^\xi(\bar{t}, \bar{\eta}) = \varphi^\xi(\bar{t}, \bar{\eta}) < v^+(\bar{t}, \bar{\eta})$. We claim that v^ξ satisfies the terminal condition. To this end, we pick some $\varepsilon > 0$ that satisfies $T > \bar{t} + \varepsilon$ and recall that v^+ satisfies the terminal condition. We then continue to show that $v^\xi \in \mathcal{V}^+$ to obtain a contradiction.

Let us fix (t, η) and recall that $((\mu_s)_{t \leq s \leq T}, (W_s, B_s)_{t \leq s \leq T}, \Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_s)_{t \leq s \leq T}) \in \chi$, where χ is the nonempty set of all weak solutions. We need to show that the process $(v^\xi(s, \mu_s))_{t \leq s \leq T}$ is a supermartingale on (Ω, \mathbb{P}) with respect to $(\mathcal{F}_s)_{t \leq s \leq T}$. We first assume that $(v^+(s, \mu_s))_{t \leq s \leq T}$ has right-continuous paths. In this case, v^ξ is a supermartingale locally in the region $[t, T] \times \mathbb{R} \setminus B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})$ because it equals the right-continuous supermartingale $(v^+(s, \mu_s))_{t \leq s \leq T}$. As the process $(v^\xi(s, \mu_s))_{t \leq s \leq T}$ is the minimum between two local supermartingales in the region $B(\bar{t}, \bar{\eta}, \varepsilon)$, it is a local supermartingale. As the two regions $[t, T] \times \mathbb{R} \setminus B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})$ and $B(\bar{t}, \bar{\eta}, \varepsilon)$ overlap over an open region, $(v^\xi(s, \mu_s))_{t \leq s \leq T}$ is actually a supermartingale.

If the process $(v^+(s, \mu_s))_{t \leq s \leq T}$ is not right-continuous, we can consider its right-continuous limit over rational times to transform it to the special case discussed above. In particular, for a given rational number r and fixed $0 \leq t \leq r \leq s \leq T$ and $\eta \in \mathbb{R}$, it remains to show the process $(Y_u)_{t \leq u \leq T} := (v^\xi(u, \mu_u))_{t \leq u \leq T}$ between r and s is a supermartingale, which is equivalent to showing that $Y_r \geq \mathbb{E}[Y_s | \mathcal{F}_r]$.

Let us define $G_u := v^+(u, \mu_u)$, $r \leq u \leq s$, and freeze the process G after time s , i.e. $G_u := v^+(s, \mu_s)$, $s \leq u \leq T$. As $(G_u)_{r \leq u \leq T}$ may not be right-continuous, by Proposition 1.3.14 in [19], we can consider its right-continuous modification

$$G_u^+(\omega) := \lim_{u' \rightarrow u, u' > u, u' \in \mathbb{Q}} G_{u'}(\omega), \quad r \leq u \leq T.$$

Note that G^+ is a right-continuous supermartingale with respect to \mathcal{F} that satisfies the usual conditions. Because v^+ is u.s.c. and the process remains the same after s , we conclude that $G_r \geq G_r^+$, $G_s = G_s^+$. Recall that $v^+ < \varphi - \delta$ in the open region $B(\bar{t}, \bar{\eta}, \varepsilon) \setminus \overline{B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})}$; if we take right limits inside this region and use the continuous function φ , we have

$$G_u^+ < \varphi^\xi(u, \mu_u), \quad \text{if } (u, \mu_u) \in B(\bar{t}, \bar{\eta}, \varepsilon) \setminus \overline{B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})}.$$

Thus, if we consider the process

$$Y_u^+ := \begin{cases} G_u^+, & (u, \mu_u) \notin \overline{B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})}, \\ G_u^+ \wedge \varphi^\xi(u, \mu_u), & (u, \mu_u) \in B(\bar{t}, \bar{\eta}, \varepsilon), \end{cases}$$

we also have $Y_r \geq Y_r^+$, $Y_s = Y_s^+$.

Because G^+ has right-continuous paths, we can conclude that Y is a supermartingale such that

$$Y_r \geq Y_r^+ \geq \mathbb{E}[Y_s^+ | \mathcal{F}_r] = \mathbb{E}[Y_s | \mathcal{F}_r].$$

(ii) *The terminal condition for v^+ .*

For some $\eta_0 \in \mathbb{R}$, we assume that $v^+(T, \eta_0) > 0$ and will show a contradiction. As \widehat{V} is continuous on \mathbb{R} , we can choose an $\varepsilon > 0$ such that $0 \leq v^+(T, \eta_0) - \varepsilon$ and $|\eta - \eta_0| \leq \varepsilon$. On the compact set $(\overline{B(T, \eta_0, \varepsilon)} \setminus B(T, \eta_0, \frac{\varepsilon}{2})) \cap ([0, T] \times \mathbb{R})$, v^+ is bounded above by the definition of \mathcal{V}^+ and the fact that $v^+ \in \mathcal{V}^+$. Moreover, as v^+ is u.s.c. on this compact set, we can find $\delta > 0$ small enough so that

$$v^+(T, \eta_0) + \frac{\varepsilon^2}{4\delta} \geq \varepsilon + \sup_{(t, \eta) \in (\overline{B(T, \eta_0, \varepsilon)} \setminus B(T, \eta_0, \frac{\varepsilon}{2})) \cap ([0, T] \times \mathbb{R})} v^+(t, \eta). \tag{4.6}$$

Next, for $k > 0$, we define the function

$$\varphi^{\delta, \varepsilon, k}(t, \eta) := v^+(T, \eta_0) + \frac{|\eta - \eta_0|^2}{\delta} + k(T - t).$$

For k large enough, we derive that $-\varphi_t^{\delta, \varepsilon, k} - \mathcal{L}\varphi^{\delta, \varepsilon, k} > 0$ on $\overline{B(T, \eta_0, \varepsilon)}$. Moreover, in view of (4.6), we have

$$\varphi^{\delta, \varepsilon, k} \geq \varepsilon + v^+ \text{ on } \left(\overline{B(T, \eta_0, \varepsilon)} \setminus B\left(T, \eta_0, \frac{\varepsilon}{2}\right) \right) \cap ([0, T] \times \mathbb{R}),$$

and $\varphi^{\delta, \varepsilon, k}(T, \eta) \geq v^+(T, \eta_0) \geq 0 + \varepsilon$ for $|\eta - \eta_0| \leq \varepsilon$.

Now, we can find $\xi < \varepsilon$ and define the following function:

$$v^{\delta, \varepsilon, k, \xi} := \begin{cases} v^+ \wedge (\varphi^{\delta, \varepsilon, k} - \xi) & \text{on } \overline{B(T, \eta_0, \varepsilon)}, \\ v^+ & \text{outside } \overline{B(T, \eta_0, \varepsilon)}. \end{cases}$$

By following a similar argument to that used in Step (i), one can obtain that $v^{\delta, \varepsilon, k, \xi} \in \mathcal{V}^+$, but $v^{\delta, \varepsilon, k, \xi}(T, \eta_0) = v^+(T, \eta_0) - \xi$, which leads to a contradiction.

(iii) *The supersolution property of v^- .*

We provide only a sketch of the proof, as it is essentially similar to that of Step (i). Suppose that v^- is not a viscosity supersolution; then there exist some interior point $(\bar{t}, \bar{\eta}) \in (0, T) \times \mathbb{R}$ and a $C^{1,2}$ test function $\psi : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ such that $v^- - \psi$ attains a strict local minimum that is equal to zero. As $F(\bar{t}, \bar{\eta}, v, v_{\bar{t}}, v_{\bar{\eta}}, v_{\bar{\eta}\bar{\eta}}) < 0$, there are two separate cases to check.

Case (i): $v^-(\bar{t}, \bar{\eta}) - \widehat{V}(\bar{t}, x_0 - f(\bar{t}), z_0, \bar{\eta}) < 0$. This already leads to a contradiction with $v^-(\bar{t}, \bar{\eta}) \geq \widehat{V}(\bar{t}, x_0 - f(\bar{t}), z_0, \bar{\eta})$ by the definition of v^- .

Case (ii): $-\frac{\partial \psi}{\partial t}(\bar{t}, \bar{\eta}) - \mathcal{L}\psi(\bar{t}, \bar{\eta}) < 0$. In this case we can find a ball $B(\bar{t}, \bar{\eta}, \varepsilon)$ small enough so that $-\frac{\partial \psi}{\partial t} - \mathcal{L}\psi < 0$ on $\overline{B(\bar{t}, \bar{\eta}, \varepsilon)}$. Moreover, as $v^- - \psi$ is l.s.c. and $\overline{B(\bar{t}, \bar{\eta}, \varepsilon)} \setminus B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})$ is compact, there exists a $\delta > 0$ such that $\psi + \delta \leq v^-$ on $\overline{B(\bar{t}, \bar{\eta}, \varepsilon)} \setminus B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})$. We can then choose $\xi \in (0, \frac{\delta}{2})$ small such that $\psi^\xi = \psi + \xi$ satisfies three properties:

- (i) $-\frac{\partial \psi^\xi}{\partial t} - \mathcal{L}\psi^\xi < 0$ on $\overline{B(\bar{t}, \bar{\eta}, \varepsilon)}$;
- (ii) we have $v^- \geq \psi + \delta > \psi + \xi = \psi^\xi$ on $\overline{B(\bar{t}, \bar{\eta}, \varepsilon)} \setminus B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})$;
- (iii) $\psi^\xi(\bar{t}, \bar{\eta}) = \psi(\bar{t}, \bar{\eta}) + \xi = v^-(\bar{t}, \bar{\eta}) + \xi > v^-(\bar{t}, \bar{\eta})$.

Thus, we can define an auxiliary function by

$$v^\xi := \begin{cases} v^- \vee \psi^\xi & \text{on } \overline{B(\bar{t}, \bar{\eta}, \varepsilon)}, \\ v^- & \text{outside } \overline{B(\bar{t}, \bar{\eta}, \varepsilon)}. \end{cases}$$

By repeating an argument similar to that of Step (i), we obtain that $v^\xi \in \mathcal{V}^-$ by showing that $(v^\xi(s, \mu_s))_{t \leq s \leq T}$ is a submartingale. If v^- has right-continuous paths, the proof is trivial. In general, by Proposition 1.3.14 in [19], we can consider the right-continuous submartingale $G_u^+(\omega) := \lim_{u' \rightarrow u, u' > u, u' \in \mathbb{Q}} G_{u'}(\omega)$, $\omega \in \Omega^*$, $r \leq u \leq T$, where $G_u := v^-(u, \mu_u)$, $r \leq u \leq s$ and we stop it at time t . Similarly to Step (i), we note that G^+ is a right-continuous submartingale and therefore $G_r \leq G_r^+$, $G_s = G_s^+$. As $G_u^+ > \psi^\xi(u, \mu_u)$, if $(u, \mu_u) \in B(\bar{t}, \bar{\eta}, \varepsilon) \setminus \overline{B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})}$, we can define the process

$$Y_u^+ := \begin{cases} G_u^+, & (u, \mu_u) \notin \overline{B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})}, \\ G_u^+ \vee \psi^\xi(u, \mu_u), & (u, \mu_u) \in \overline{B(\bar{t}, \bar{\eta}, \frac{\varepsilon}{2})}. \end{cases}$$

We can conclude that $Y_r \leq Y_r^+$, $Y_s = Y_s^+$, and Y is a submartingale such that $Y_r \leq Y_r^+ \leq \mathbb{E}[Y_s^+ | \mathcal{F}_r] = \mathbb{E}[Y_s | \mathcal{F}_r]$, which completes the proof.

(iv) *The terminal condition for v^- .*

For some $\eta_0 \in \mathbb{R}$, suppose that $v^-(T, \eta_0) < 0$; we will derive a contradiction. As \widehat{V} is continuous on \mathbb{R} , we can choose an $\varepsilon > 0$ such that $0 \geq v^-(T, \eta_0) + \varepsilon$ and $|\eta - \eta_0| \leq \varepsilon$. Similarly

to Step (ii), we can find $\delta > 0$ small enough so that

$$v^-(T, \eta_0) - \frac{\varepsilon^2}{4\delta} \leq \inf_{(t, \eta) \in (\overline{B(T, \eta_0, \varepsilon)} \setminus B(T, \eta_0, \frac{\varepsilon}{2})) \cap ([0, T] \times \mathbb{R})} v^-(t, \eta) - \varepsilon. \tag{4.7}$$

Then, for $k > 0$, we consider

$$\psi^{\delta, \varepsilon, k}(t, \eta) := v^-(T, \eta_0) - \frac{|\eta - \eta_0|^2}{\delta} - k(T - t).$$

For k large enough, we have that $-\psi_t^{\delta, \varepsilon, k} - \mathcal{L}\psi^{\delta, \varepsilon, k} < 0$ on $\overline{B(T, \eta_0, \varepsilon)}$. Furthermore, in view of (4.7), we have

$$\psi^{\delta, \varepsilon, k} \leq v^- - \varepsilon \text{ on } \left(\overline{B(T, \eta_0, \varepsilon)} \setminus B\left(T, \eta_0, \frac{\varepsilon}{2}\right) \right) \cap ([0, T] \times \mathbb{R}),$$

and $\psi^{\delta, \varepsilon, k}(T, \eta) \leq v^-(T, \eta_0) \leq -\varepsilon$ for $|\eta - \eta_0| \leq \varepsilon$.

Next, we can find $\xi < \varepsilon$ and define the function

$$v^{\delta, \varepsilon, k, \xi} := \begin{cases} v^- \vee (\psi^{\delta, \varepsilon, k} + \xi) & \text{on } \overline{B(T, \eta_0, \varepsilon)}, \\ v^- & \text{outside } \overline{B(T, \eta_0, \varepsilon)}. \end{cases}$$

Similarly to Step (iii), we obtain that $v^{\delta, \varepsilon, k, \xi} \in \mathcal{V}^-$, but $v^{\delta, \varepsilon, k, \xi}(T, \eta_0) = v^-(T, \eta_0) + \xi$, which gives a contradiction. \square

We now reverse the time and consider $s := T - t$. However, for simplicity of presentation, we continue to use t in place of s whenever there is no possibility of confusion. The variational inequalities can be written as

$$\min \left\{ \tilde{V}(t, \eta; x_0 - f(T - t), z_0) - \widehat{V}(t, x_0 - f(T - t), z_0, \eta), \frac{\partial \tilde{V}(t, \eta)}{\partial t} - \mathcal{L}\tilde{V}(t, \eta) \right\} = 0, \tag{4.8}$$

where

$$\mathcal{L}\tilde{V}(t, \eta) = -\lambda(\eta - \bar{\mu}) \frac{\partial \tilde{V}}{\partial \eta}(t, \eta) + \frac{1}{2} \sigma_\mu^2 \frac{\partial^2 \tilde{V}}{\partial \eta^2}(t, \eta)$$

with the condition $\tilde{V}(0, \eta) = 0$.

Let us write this equivalently as

$$\begin{cases} F(t, \eta, v, v_t, v_\eta, v_{\eta\eta}) = 0, & \text{on } (0, T] \times \mathbb{R}, \\ v(0, \eta) = \widehat{V}(0, x_0 - f(0), z_0, \eta), & \text{for any } \eta \in \mathbb{R}, \end{cases} \tag{4.9}$$

where

$$F(t, \eta, v, v_t, v_\eta, v_{\eta\eta}) := \min \left\{ v - \widehat{V}, \frac{\partial v}{\partial t} - \mathcal{L}v \right\}.$$

We also have the continuation region as $\mathcal{C} = \{(t, \eta) \in (0, T] \times \mathbb{R} : \tilde{V}(t, \eta; x_0 - f(T - t), z_0) > \widehat{V}(t, x_0 - f(T - t), z_0, \eta)\}$.

Proposition 4.1. (Comparison principle) *Let u, v respectively be a u.s.c. viscosity subsolution and an l.s.c. viscosity supersolution of (4.9). If $u(0, \eta) \leq v(0, \eta)$ on \mathbb{R} , then we have $u \leq v$ on $(0, T] \times \mathbb{R}$.*

Proof. We will follow some arguments from [5, 28], modifying them to fit our setting. Suppose that $u(0, \eta) \leq v(0, \eta)$ on \mathbb{R} ; we will prove that $u \leq v$ on $[0, T] \times \mathbb{R}$. We first construct the strict supersolution to the system (4.9) with suitable perturbations of v . Let us recall that $A \leq 0, B \leq 0$, and C is bounded above by some constant as in Remark A.1. Moreover, we know that $\widehat{V}(t, x_0 - \kappa t, z_0, \eta) \leq 0$. Let us fix a constant $C_2 > 0$ small enough so that $\lambda > C_2\sigma_\mu^2$ and set $\psi(t, \eta) = C_0e^t + e^{C_2\eta^2}$ with some $C_0 > 1$. We have that

$$\begin{aligned} \frac{\partial \psi}{\partial t} - \mathcal{L}\psi &= C_0e^t + C_2 \left[2(\lambda - C_2\sigma_\mu^2)\eta^2 - 2\lambda\bar{\mu}\eta - \sigma_\mu^2 \right] e^{C_2\eta^2} \\ &\geq C_0e^t + C_2 \frac{-2(\lambda - C_2\sigma_\mu^2)\sigma_\mu^2 - \lambda^2\bar{\mu}^2}{2(\lambda - C_2\sigma_\mu^2)} \\ &> C_0 + C_2 \frac{-2(\lambda - C_2\sigma_\mu^2)\sigma_\mu^2 - \lambda^2\bar{\mu}^2}{2(\lambda - C_2\sigma_\mu^2)}. \end{aligned}$$

We can then choose $C_0 > 1$ large enough so that

$$C_0 + C_2 \frac{-2(\lambda - C_2\sigma_\mu^2)\sigma_\mu^2 - \lambda^2\bar{\mu}^2}{2(\lambda - C_2\sigma_\mu^2)} > 1,$$

which guarantees that

$$\frac{\partial \psi}{\partial t} - \mathcal{L}\psi > 1. \tag{4.10}$$

Let us define $v^\Lambda := (1 - \Lambda)v + \Lambda\psi$ on $[0, T] \times \mathbb{R}$ for any $\Lambda \in (0, 1)$. It follows that

$$\begin{aligned} v^\Lambda - \widehat{V} &= (1 - \Lambda)v + \Lambda\psi - \widehat{V} = (1 - \Lambda)v + \Lambda \left(C_0e^t + e^{C_2\eta^2} \right) - \widehat{V} \\ &\geq (1 - \Lambda)v + \Lambda \left(C_0e^t + e^{C_2\eta^2} \right) + \Lambda\widehat{V} - \widehat{V} \\ &> (1 - \Lambda)(v - \widehat{V}) + \Lambda C_0 > \Lambda, \end{aligned} \tag{4.11}$$

where we used $v - \widehat{V} \geq 0$ in the last inequality. From (4.10) and (4.11), we can deduce that for $\Lambda \in (0, 1)$, v^Λ is a supersolution to

$$\min \left\{ v^\Lambda - \widehat{V}, \frac{\partial v^\Lambda}{\partial t} - \mathcal{L}v^\Lambda \right\} \geq \Lambda. \tag{4.12}$$

To prove the comparison principle, it suffices to prove the claim that $\sup(u - v^\Lambda) \leq 0$ for all $\Lambda \in (0, 1)$, as the required result is then obtained by letting Λ go to 0. To this end, we will suppose that there exists some $\Lambda \in (0, 1)$ such that $M := \sup(u - v^\Lambda) > 0$, and derive a contradiction.

It is clear that u, v , and \widehat{V} have the same growth conditions: in view of the explicit forms of A, B, C , and \widehat{V} , it follows that \widehat{V} has growth condition in t as $e^{e^{K_1 t}}$ for some $K_1 < 0$ and has growth condition in η as $e^{K_2\eta^2}$ for some $K_2 < 0$; on the other hand, ψ has growth condition in t as e^t and has growth condition in η as $e^{C_2\eta^2}$. Thus, we have that $u(t, \eta) - v^\Lambda(t, \eta) = (u - (1 - \Lambda)v - \Lambda\psi)(t, \eta)$ goes to $-\infty$ as $t \rightarrow T, \eta \rightarrow \infty$. Consequently, the u.s.c. function $(u - v^\Lambda)$ attains its maximum M .

Let us consider the u.s.c. function $\Phi_\varepsilon(t, t', \eta, \eta') = u(t, \eta) - v^\Lambda(t', \eta') - \phi_\varepsilon(t, t', \eta, \eta')$, where $\phi_\varepsilon(t, t', \eta, \eta') = \frac{1}{2\varepsilon}((t - t')^2 + (\eta - \eta')^2)$, $\varepsilon > 0$, and $(t_\varepsilon, t'_\varepsilon, \eta_\varepsilon, \eta'_\varepsilon)$ attains the maximum of Φ_ε . We have

$$M_\varepsilon = \max \Phi_\varepsilon = \Phi_\varepsilon(t_\varepsilon, t'_\varepsilon, \eta_\varepsilon, \eta'_\varepsilon) \rightarrow M \text{ and } \phi_\varepsilon(t_\varepsilon, t'_\varepsilon, \eta_\varepsilon, \eta'_\varepsilon) \rightarrow 0 \text{ when } \varepsilon \rightarrow 0. \tag{4.13}$$

Let us recall the equivalent definition of viscosity solutions in terms of superjets and subjets. In particular, we define $\bar{\mathcal{P}}^{2,+}u(\bar{t}, \bar{\eta})$ as the set of elements $(\bar{q}, \bar{k}, \bar{M}) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ satisfying $u(t, \eta) \leq u(\bar{t}, \bar{\eta}) + \bar{q}(t - \bar{t}) + \bar{k}(\eta - \bar{\eta}) + \frac{1}{2}\bar{M}(\eta - \bar{\eta})^2 + o((t - \bar{t}) + (\eta - \bar{\eta})^2)$. We define $\bar{\mathcal{P}}^{2,-}v^\Lambda(\bar{t}, \bar{\eta})$ similarly. Thanks to the Crandall–Ishii lemma, we can find $A_\varepsilon, B_\varepsilon \in \mathbb{R}$ such that

$$\begin{aligned} \left(\frac{t_\varepsilon - t'_\varepsilon}{\varepsilon}, \frac{\eta_\varepsilon - \eta'_\varepsilon}{\varepsilon}, A_\varepsilon \right) &\in \bar{\mathcal{P}}^{2,+}u(t_\varepsilon, \eta_\varepsilon), \\ \left(\frac{t_\varepsilon - t'_\varepsilon}{\varepsilon}, \frac{\eta_\varepsilon - \eta'_\varepsilon}{\varepsilon}, B_\varepsilon \right) &\in \bar{\mathcal{P}}^{2,-}v^\Lambda(t'_\varepsilon, \eta'_\varepsilon), \\ \sigma^2(\eta_\varepsilon)A_\varepsilon - \sigma^2(\eta'_\varepsilon)B_\varepsilon &\leq \frac{3}{\varepsilon}(\sigma(\eta_\varepsilon) - \sigma(\eta'_\varepsilon))^2. \end{aligned}$$

By combining the viscosity subsolution property (4.5) of u and the viscosity strict supersolution property (4.12) of v^Λ , we have that

$$\min \left\{ u(t_\varepsilon, \eta_\varepsilon) - \widehat{V}(t_\varepsilon, x_0 - f(t_\varepsilon), z_0, \eta_\varepsilon), \frac{t_\varepsilon - t'_\varepsilon}{\varepsilon} - \frac{\eta_\varepsilon - \eta'_\varepsilon}{\varepsilon} b(t_\varepsilon, \eta_\varepsilon) - \frac{1}{2}\sigma^2(\eta_\varepsilon)A_\varepsilon \right\} \leq 0, \tag{4.14}$$

$$\min \left\{ v^\Lambda(t'_\varepsilon, \eta'_\varepsilon) - \widehat{V}(t'_\varepsilon, x_0 - f(t'_\varepsilon), z_0, \eta'_\varepsilon), \frac{t_\varepsilon - t'_\varepsilon}{\varepsilon} - \frac{\eta_\varepsilon - \eta'_\varepsilon}{\varepsilon} b(t'_\varepsilon, \eta'_\varepsilon) - \frac{1}{2}\sigma^2(\eta'_\varepsilon)B_\varepsilon \right\} \geq \Lambda, \tag{4.15}$$

where $b(t_\varepsilon, \eta_\varepsilon) = -\lambda(\eta_\varepsilon - \bar{\mu})$, $\sigma^2(\eta_\varepsilon) = \sigma_\mu^2$, $b(t'_\varepsilon, \eta'_\varepsilon) = -\lambda(\eta'_\varepsilon - \bar{\mu})$, and $\sigma^2(\eta'_\varepsilon) = \sigma_\mu^2$.

If $u - \widehat{V} \leq 0$ in (4.14), then because $v^\Lambda - \widehat{V} \geq \Lambda$ in (4.15), we obtain that $u - v^\Lambda \leq -\Lambda < 0$ by contradiction with $\sup(u - v^\Lambda) = M > 0$. On the other hand, if $u - \widehat{V} > 0$ in (4.14), then we have

$$\begin{cases} \frac{t_\varepsilon - t'_\varepsilon}{\varepsilon} - \frac{\eta_\varepsilon - \eta'_\varepsilon}{\varepsilon} b(t_\varepsilon, \eta_\varepsilon) - \frac{1}{2}\sigma^2(\eta_\varepsilon)A_\varepsilon \leq 0, \\ \frac{t_\varepsilon - t'_\varepsilon}{\varepsilon} - \frac{\eta_\varepsilon - \eta'_\varepsilon}{\varepsilon} b(t'_\varepsilon, \eta'_\varepsilon) - \frac{1}{2}\sigma^2(\eta'_\varepsilon)B_\varepsilon \geq \Lambda. \end{cases}$$

Furthermore, combining two inequalities above, we derive that

$$\begin{aligned} &\frac{\eta_\varepsilon - \eta'_\varepsilon}{\varepsilon} (b(t_\varepsilon, \eta_\varepsilon) - b(t'_\varepsilon, \eta'_\varepsilon)) + \frac{3}{2\varepsilon} (\sigma(\eta_\varepsilon) - \sigma(\eta'_\varepsilon))^2 \\ &\geq \frac{\eta_\varepsilon - \eta'_\varepsilon}{\varepsilon} (b(t_\varepsilon, \eta_\varepsilon) - b(t'_\varepsilon, \eta'_\varepsilon)) + \frac{1}{2} (\sigma^2(\eta_\varepsilon)A_\varepsilon - \sigma^2(\eta'_\varepsilon)B_\varepsilon) \geq \Lambda. \end{aligned}$$

The first inequality holds by the Crandall–Ishii lemma. In addition, by letting $\varepsilon \rightarrow 0$, we get

$$\frac{\eta_\varepsilon - \eta'_\varepsilon}{\varepsilon} (b(t_\varepsilon, \eta_\varepsilon) - b(t'_\varepsilon, \eta'_\varepsilon)) + \frac{3}{2\varepsilon} (\sigma(\eta_\varepsilon) - \sigma(\eta'_\varepsilon))^2 = 0$$

thanks to (4.13). It follows that we have $0 \geq \Lambda > 0$, which leads to a contradiction; therefore our claim holds. \square

Lemma 4.5. *For all $(t, \eta) \in \mathcal{C}$ in the continuation region, \tilde{V} in (2.7) has Hölder continuous derivatives.*

Proof. The proof closely follows the argument in Section 6.3 of [16]. First, let us recall that

$$\frac{\partial \tilde{V}}{\partial t}(t, \eta) + \lambda(\eta - \bar{\mu}) \frac{\partial \tilde{V}}{\partial \eta}(t, \eta) - \frac{1}{2} \sigma_\mu^2 \frac{\partial^2 \tilde{V}}{\partial \eta^2}(t, \eta) = 0 \text{ on } \mathcal{C}. \tag{4.16}$$

The fact that \tilde{V} is a viscosity solution to (4.8) gives that \tilde{V} is a supersolution to (4.16). On the other hand, for any $(\bar{t}, \bar{\eta}) \in \mathcal{C}$, let φ be a C^2 test function such that $(\bar{t}, \bar{\eta})$ is a maximum of $\tilde{V} - \varphi$ with $\tilde{V}(\bar{t}, \bar{\eta}) = \varphi(\bar{t}, \bar{\eta})$. By definition of \mathcal{C} , we have $\tilde{V}(\bar{t}, \bar{\eta}) > \widehat{V}(\bar{t}, x_0 - f(\bar{t}), z_0, \bar{\eta})$, so that

$$\frac{\partial \varphi}{\partial t}(\bar{t}, \bar{\eta}) + \lambda(\eta - \bar{\mu}) \frac{\partial \varphi}{\partial \eta}(\bar{t}, \bar{\eta}) - \frac{1}{2} \sigma_\mu^2 \frac{\partial^2 \varphi}{\partial \eta^2}(\bar{t}, \bar{\eta}) \leq 0,$$

owing to the property that \tilde{V} is a viscosity subsolution to (4.8). It follows that \tilde{V} is a viscosity subsolution and therefore a viscosity solution to (4.16).

Let us consider an initial boundary value problem:

$$\begin{aligned} -\frac{\partial w}{\partial t}(t, \eta) - \lambda(\eta - \bar{\mu}) \frac{\partial w}{\partial \eta}(t, \eta) + \frac{1}{2} \sigma_\mu^2 \frac{\partial^2 w}{\partial \eta^2}(t, \eta) &= 0 \text{ on } Q \cup B_T, \\ w(0, \eta) &= 0 \text{ on } B, \\ w(t, \eta) &= \widehat{V}(t, x_0 - \kappa t, z_0, \eta) \text{ on } S. \end{aligned} \tag{4.17}$$

Here, Q is an arbitrary bounded open region in \mathcal{C} , Q lies in the strip $0 < t < T$, $\tilde{B} = \bar{Q} \cap \{t = 0\}$, $\tilde{B}_T = \bar{Q} \cap \{t = T\}$, B_T denotes the interior of \tilde{B}_T , B denotes the interior of \tilde{B} , S_0 denotes the boundary of Q lying in the strip $0 \leq t \leq T$, and $S = S_0 \setminus B_T$. Theorem 3.6 in [16] gives the existence and uniqueness of a solution w on $Q \cup B_T$ to (4.17), and the solution w has Hölder continuous derivatives w_t , w_η , and $w_{\eta\eta}$. Because the solution w is a viscosity solution to (4.16) on $Q \cup B_T$, from standard uniqueness results on viscosity solutions, we know that $\tilde{V} = w$ on $Q \cup B_T$. As $Q \subset \mathcal{C}$ is arbitrary, it follows that \tilde{V} has the same property in the continuation region \mathcal{C} . Therefore, \tilde{V} has Hölder continuous derivatives \tilde{V}_t , \tilde{V}_η , and $\tilde{V}_{\eta\eta}$. \square

Finally, we can prove Theorem 2.1.

Proof. We have proved the inequality $v^- = \sup_{p \in \mathcal{V}^-} p \leq \tilde{V} \leq v^+ = \inf_{q \in \mathcal{V}^+} q$ in Lemma 4.4. Using the comparison result in Proposition 4.1, we also have $v^+ \leq v^-$. Putting all the pieces together, we conclude that $v^+ = \tilde{V}(t, \eta) = v^-$, and therefore the value function $\tilde{V}(t, \eta)$ is the unique viscosity solution of the HJB variational inequality (2.9). Following an argument similar to that given for Theorem 1 in [13], we fix the \mathcal{F}_t -adapted stopping time τ^* defined in (2.11); the Itô–Tanaka formula (see Theorem IV.1.5 and Corollary IV.1.6 of [32]) can be applied to $\tilde{V}(t, \mu_t)$ in view of the Hölder continuous derivatives of $\tilde{V}(t, \eta)$, and we get that

$$\begin{aligned} &\widehat{V}(\tau^* \wedge \tau_n, x_0 - \kappa \tau^* \wedge \tau_n, z_0, \mu_{\tau^* \wedge \tau_n}) \\ &= \tilde{V}(t, \mu_t) + \left[\widehat{V}(\tau^* \wedge \tau_n, x_0 - \kappa \tau^* \wedge \tau_n, z_0, \mu_{\tau^* \wedge \tau_n}) - \tilde{V}(\tau^* \wedge \tau_n, \mu_{\tau^* \wedge \tau_n}) \right] \\ &+ \int_t^{\tau^* \wedge \tau_n} \sigma_\mu \frac{\partial \tilde{V}}{\partial \eta}(s, \mu_s) dB_s + \int_t^{\tau^* \wedge \tau_n} \left[\frac{\partial \tilde{V}(s, \mu_s)}{\partial t} + \mathcal{L} \tilde{V}(s, \mu_s) \right] ds, \end{aligned}$$

where $\tau_n \uparrow T$ is the localizing sequence. As $\tilde{V}(t, \eta)$ satisfies the HJB variational inequality (2.9), by taking conditional expectations and using the definition of τ^* in (2.11), we obtain that

$$\mathbb{E}_t [\widehat{V}(\tau^* \wedge \tau_n, x_0 - \kappa \tau^* \wedge \tau_n, z_0, \mu_{\tau^* \wedge \tau_n}) \mathbf{1}_{\{\tau^* \leq \tau_n\}}] + \mathbb{E}_t [\tilde{V}(\tau_n, \mu_{\tau_n}) \mathbf{1}_{\{\tau^* > \tau_n\}}] = \tilde{V}(t, \mu_t).$$

By taking the limit of τ_n and using the dominated convergence theorem, we verify that

$$\mathbb{E}_t [\widehat{V}(\tau^*, x_0 - \kappa \tau^*, z_0, \mu_{\tau^*})] = \tilde{V}(t, \mu_t),$$

and therefore τ^* is the optimal entry time.

Finally, the martingale property between $t=0$ and τ^* follows from the definition of stochastic subsolutions and stochastic supersolutions. □

Moreover, we can also easily verify the following sensitivity results for the composite value function.

Lemma 4.6. *The value function $\tilde{V}(t, \eta)$ has the following sensitivity properties:*

- (i) *Suppose that $\alpha > 0$ and $\delta > 0$ are both constants in the definition of a habit formation process such that $\delta > \alpha$. We have that $\tilde{V}(t, \eta; \alpha, \delta)$ is decreasing in δ and increasing in α .*
- (ii) *If the initial habit z_0 increases, the value function $\tilde{V}(t, \eta)$ decreases.*
- (iii) *If the information cost rate κ increases, the value function $\tilde{V}(t, \eta)$ decreases for any $t < T$.*

Proof. By the definition of $\tilde{V}(t, \eta)$ and the explicit form of $\widehat{V}(t, x_0 - \kappa t, z_0, \eta)$ in (3.7) and $m(t)$ in (2.8), for given $\delta > \alpha$, it is clear that $\widehat{V}(t, x_0 - \kappa t, z_0, \eta)$ is decreasing in δ and increasing in α , which implies that $\tilde{V}(t, \eta)$ has the same sensitivity property. Similarly, it is clear that $\widehat{V}(t, x_0 - \kappa t, z_0, \eta)$ decreases as z_0 increases, and hence $\tilde{V}(t, \eta)$ is decreasing in z_0 . Finally, $\widehat{V}(t, x_0 - \kappa t, z_0, \eta)$ decreases if $x_0 - \kappa t$ decreases; it readily follows that $\tilde{V}(t, \eta)$ is decreasing in κ . □

Appendix A. Explicit solution to the auxiliary ODEs

Our ODE problems (3.4), (3.5), (3.6) are similar to ODEs for the terminal wealth optimization problem in [8], in which the insightful observation is made that we can solve these ODEs with coefficients depending on time t by solving five auxiliary ODEs with constant coefficients; see Section 4 of [8] for detailed discussions.

Lemma A.1. *For $k \leq t \leq s \leq T$, let us consider the following auxiliary ODEs for $a(t, s)$, $b(t, s)$, $l(t, s)$, $w(t, s)$, and $g(t, s)$:*

$$a_t = - \frac{2(1-p+p\rho^2)}{1-p} \sigma_\mu^2 a^2 + \left(2\lambda - \frac{2p\rho\sigma_\mu}{(1-p)\sigma_S} \right) a - \frac{p}{2(1-p)\sigma_S^2}, \tag{A.1}$$

$$b_t = - \frac{2(1-p+p\rho^2)}{1-p} \sigma_\mu^2 ab - 2\lambda\bar{\mu}a + \left(\lambda - \frac{p\rho\sigma_\mu}{(1-p)\sigma_S} \right) b, \tag{A.2}$$

$$l_t = - \sigma_\mu^2 a - \frac{(1-p+p\rho^2)\sigma_\mu^2}{2(1-p)} b^2 - \lambda\bar{\mu}b, \tag{A.3}$$

$$w_t = -2(1 - \rho^2)\sigma_\mu^2 w^2 + 2 \frac{\lambda\sigma_S + \rho\sigma_\mu}{\sigma_S} w + \frac{1}{2\sigma_S^2}, \tag{A.4}$$

$$g_t = \sigma_\mu^2(1 - \rho^2)(w - a), \tag{A.5}$$

with the terminal conditions $a(s, s) = b(s, s) = l(s, s) = w(s, s) = g(s, s) = 0$. Direct substitutions and computations show that the solutions of the ODEs (3.4), (3.5), (3.6) are given respectively by

$$\begin{aligned} A(t, s) &:= \frac{a(t, s)}{(1 - p)(1 - 2a(t, s)\hat{\Sigma}(t))}, & B(t, s) &:= \frac{b(t, s)}{(1 - p)(1 - 2a(t, s)\hat{\Sigma}(t))}, \\ C(t, s) &:= \frac{1}{1 - p} \left[l(t, s) + \frac{\hat{\Sigma}(t)}{(1 - 2a(t, s)\hat{\Sigma}(t))} b^2(t, s) - \frac{1 - p}{2} \log(1 - 2a(t, s)\hat{\Sigma}(t)) \right. \\ &\quad \left. - \frac{p}{2} \log(1 - 2w(t, s)\hat{\Sigma}(t)) - pg(t, s) \right]. \end{aligned} \tag{A.6}$$

Following the same arguments as in [21], we can actually solve the auxiliary ODEs (A.1), (A.2), (A.3), (A.4) and (A.5) explicitly, in the following order: we first solve the simple ODEs (A.1) and (A.4) to get $a(t, s)$ and $w(t, s)$, and then obtain $b(t, s)$ and $g(t, s)$ by solving the ODEs (A.2) and (A.5). Finally, we solve the ODE (A.3) to get $l(t, s)$. We thus obtain

$$\begin{aligned} a(t, s) &= \frac{p(1 - e^{2\xi(t-s)})}{2(1 - p)\sigma_S^2 \left[2\xi - (\xi + \gamma_2)(1 - e^{2\xi(t-s)}) \right]}, \\ b(t, s) &= \frac{p\lambda\bar{\mu}(1 - e^{\xi(t-s)})^2}{(1 - p)\sigma_S^2 \xi \left[2\xi - (\xi + \gamma_2)(1 - e^{2\xi(t-s)}) \right]}, \\ l(t, s) &= \frac{p}{2(1 - p)\sigma_S^2} \left(\frac{\lambda^2\bar{\mu}^2}{\xi^2} - \frac{\sigma_\mu^2\gamma_2}{\gamma_2^2 - \xi^2} \right) (s - t) \\ &\quad + \frac{p\lambda^2\bar{\mu}^2 \left[(\xi + 2\gamma_2)e^{2\xi(t-s)} - 4\gamma_2e^{\xi(t-s)} + 2\gamma_2 - \xi \right]}{2(1 - p)\sigma_S^2 \xi^3 \left[2\xi - (\xi + \gamma_2)(1 - e^{2\xi(t-s)}) \right]} \\ &\quad + \frac{p\sigma_\mu^2}{2(1 - p)\sigma_S^2 (\xi^2 - \gamma_2^2)} \log \left| \frac{2\xi - (\xi + \gamma_2)(1 - e^{2\xi(t-s)})}{2\xi e^{\xi(t-s)}} \right|, \\ w(t, s) &= -\frac{1}{2\sigma_S} \frac{1 - e^{2\xi_1(t-s)}}{(\sigma_S\xi_1 + \lambda\sigma_S + \rho\sigma_\mu) + (\sigma_S\xi_1 - \lambda\sigma_S - \rho\sigma_\mu)e^{2\xi_1(t-s)}}, \\ g(t, s) &= \frac{1}{2} \log \left(\frac{(\sigma_S\xi_1 + \lambda\sigma_S + \rho\sigma_\mu) + (\sigma_S\xi_1 - \lambda\sigma_S - \rho\sigma_\mu)e^{2\xi_1(t-s)}}{2\sigma_S\xi_1 e^{\xi_1(t-s)}} \right) \\ &\quad - \frac{(1 - p)(1 - \rho^2)}{2(1 - p + p\rho^2)} \log \left(\frac{(\sigma_S\xi + \lambda\sigma_S - \frac{\rho\sigma_\mu p}{1 - p}) + (\sigma_S\xi - \lambda\sigma_S + \frac{\rho\sigma_\mu p}{1 - p})e^{2\xi(t-s)}}{2\sigma_S\xi e^{\xi(t-s)}} \right) \\ &\quad - \frac{\rho^2\lambda(s - t)}{2(1 - p + p\rho^2)} - \frac{\rho\sigma_\mu(s - t)}{2(1 - p + p\rho^2)\sigma_S}, \end{aligned}$$

where

$$\Delta := \lambda^2 - \frac{2\lambda p\rho\sigma_\mu}{(1-p)\sigma_S} - \frac{p\sigma_\mu^2}{(1-p)\sigma_S^2} > 0, \tag{A.7}$$

and

$$\begin{aligned} \xi &:= \sqrt{\Delta} = \sqrt{\gamma_2^2 - \gamma_1\gamma_3}, \quad \xi_1 := \frac{\sqrt{(1-\rho^2)\sigma_\mu^2 + (\lambda\sigma_S + \rho\sigma_\mu)^2}}{\sigma_S}, \\ \gamma_1 &:= \frac{(1-p+p\rho^2)}{1-p}\sigma_\mu^2, \quad \gamma_2 := -\lambda + \frac{p\rho\sigma_\mu}{(1-p)\sigma_S}, \quad \gamma_3 := \frac{p}{(1-p)\sigma_S^2}. \end{aligned}$$

Moreover, it is straightforward to see that $a, b, l, w,$ and g are globally bounded if we have that $\gamma_3 > 0,$ or $\gamma_1 > 0,$ or $\gamma_2 < 0.$

Remark A.1 Under the assumption that $p < 0,$ (A.7) clearly holds and we have $\gamma_2 < 0.$ We can see that $a(t, s) \leq 0$ and $b(t, s) \leq 0$ are bounded and that $1 - 2a(t, s)\hat{\Sigma}(t) > 1$ and $1 - w(t, s)\hat{\Sigma}(t) > 1.$ From the expressions in (A.6), we can conclude that $A(t, s), B(t, s),$ and $C(t, s)$ are all bounded on $k \leq t \leq s \leq T,$ and that

$$A(t, s) = \frac{a(t, s)}{(1-p)(1 - 2a(t, s)\hat{\Sigma}(t))} \leq 0$$

and

$$B(t, s) = \frac{b(t, s)}{(1-p)(1 - 2a(t, s)\hat{\Sigma}(t))} \leq 0$$

for $k \leq t \leq s \leq T.$

Appendix B. Proof of the verification theorem

We first show that the consumption constraint $c_t \geq Z_t$ implies the constraint on the controlled wealth process in the next lemma.

Lemma B.1. *The admissible space \mathcal{A} is non-empty if and only if the initial budget constraint $x \geq m(k)z$ is fulfilled. Moreover, for each pair $(\pi, c) \in \mathcal{A},$ the controlled wealth process $\hat{X}_t^{\pi,c}$ satisfies the constraint*

$$\hat{X}_t^{\pi,c} \geq m(t)Z_t, \quad k \leq t \leq T, \tag{B.1}$$

where the deterministic function $m(t)$ is defined in (2.8) and refers to the cost of subsistence consumption per unit of standard of living at time $t.$

Proof. Let us first assume that $x \geq m(k)z;$ we can always take $\pi_t \equiv 0,$ and

$$c_t = ze^{\int_k^t (\delta(v) - \alpha(v))dv}$$

for $t \in [k, T].$ It is easy to verify that $\hat{X}_t^{\pi,c} \geq 0$ and $c_t \equiv Z_t,$ so that $(\pi, c) \in \mathcal{A},$ and hence \mathcal{A} is non-empty.

On the other hand, starting from $t = k$ with wealth x and standard of living $z,$ the additive habits constraint $c_t \geq Z_t, k \leq t \leq T,$ implies that the consumption must always exceed the subsistence consumption $\bar{c}_t = Z(t; \bar{c}_t)$ which satisfies

$$d\bar{c}_t = (\delta(t) - \alpha(t))\bar{c}_t dt, \quad \bar{c}_k = z, \quad k \leq t \leq T. \tag{B.2}$$

Indeed, since Z_t satisfies $dZ_t = (\delta_t c_t - \alpha_t Z_t)dt$ with $Z_k = z \geq 0$, the constraint $c_t \geq Z_t$ implies that

$$dZ_t \geq (\delta_t Z_t - \alpha_t Z_t)dt, \quad Z_k = z. \tag{B.3}$$

By (B.2) and (B.3), one can get $d(Z_t - \bar{c}_t) \geq (\delta_t - \alpha_t)(Z_t - \bar{c}_t)dt$ and $Z_k - \bar{c}_k = 0$, from which we can derive that

$$e^{\int_k^t (\delta_s - \alpha_s)ds} (Z_t - \bar{c}_t) \geq 0, \quad k \leq t \leq T.$$

It follows that $c_t \geq \bar{c}_t$, which is equivalent to

$$c_t \geq z e^{\int_k^t (\delta(v) - \alpha(v))dv}, \quad k \leq t \leq T. \tag{B.4}$$

Define the exponential local martingale

$$\tilde{H}_t = \exp \left(- \int_k^t \frac{\hat{\mu}_v}{\sigma_S} d\hat{W}_v - \frac{1}{2} \int_k^t \frac{\hat{\mu}_v^2}{\sigma_S^2} dv \right), \quad k \leq t \leq T.$$

As $\hat{\mu}_t$ follows the dynamics (2.4), we derive that

$$\hat{\mu}_t = e^{-t\lambda} \eta + \bar{\mu} (1 - e^{-t\lambda}) + \int_k^t e^{\lambda(u-t)} \frac{(\hat{\Sigma}(u) + \sigma_S \sigma_{\mu\rho})}{\sigma_S} d\hat{W}_u.$$

Similarly to the proof of Corollary 3.5.14 and Corollary 3.5.16 in [19], the Beneš condition implies that \tilde{H} is a true martingale with respect to $(\Omega, \mathcal{F}^S, \mathbb{P})$.

Now, define the probability measure $\tilde{\mathbb{P}}$ by $\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \tilde{H}_T$. Girsanov's theorem states that

$$\tilde{W}_t := \hat{W}_t + \int_k^t \frac{\hat{\mu}_v}{\sigma_S} dv, \quad k \leq t \leq T,$$

is a Brownian motion under $(\tilde{\mathbb{P}}, (\mathcal{F}_t^S)_{k \leq t \leq T})$. We can rewrite the wealth process as

$$\hat{X}_T + \int_k^T c_v dv = x + \int_k^T \pi_v \sigma_S d\tilde{W}_v.$$

As we have $\hat{X}_T \geq 0$, it is easy to see that $\int_k^t \pi_v \sigma_S d\tilde{W}_v$ is a supermartingale under $(\Omega, \mathbb{F}^S, \tilde{\mathbb{P}})$. By taking the expectation under $\tilde{\mathbb{P}}$, we have $x \geq \tilde{\mathbb{E}} \left[\int_k^T c_v dv \right]$. Thanks to the inequality (B.4), we further have $x \geq z \tilde{\mathbb{E}} \left[\int_k^T \exp \left(\int_k^v (\delta(u) - \alpha(u)) du \right) dv \right]$. Because $\delta(t)$ and $\alpha(t)$ are deterministic

functions, we obtain that $x \geq m(k)z$. In general, for any $t \in [k, T]$, following the same procedure, we can take the conditional expectation under the filtration \mathcal{F}_t^S and get

$$\hat{X}_t \geq Z_t \tilde{\mathbb{E}} \left[\int_t^T \exp \left(\int_t^v (\delta(u) - \alpha(u)) du \right) dv \middle| \mathcal{F}_t^S \right].$$

Again, as $\delta(t), \alpha(t)$ are deterministic, we get $\hat{X}_t \geq m(t)Z_t, k \leq t \leq T$. □

We can finally prove Theorem 3.1 for the interior control problem.

Proof. For any pair of admissible controls $(\pi_t, c_t) \in \mathcal{A}$, Itô’s lemma gives

$$d[V(t, \hat{X}_t, Z_t, \hat{\mu}_t)] = [\mathcal{G}^{\pi_t, c_t} V(t, \hat{X}_t, Z_t, \hat{\mu}_t)] dt + \left[V_x \sigma_S \pi_t + V_\eta \frac{(\hat{\Sigma}(t) + \sigma_S \sigma_\mu \rho)}{\sigma_S} \right] d\hat{W}_t, \tag{B.5}$$

where we define the process $\mathcal{G}^{\pi_t, c_t} V(t, \hat{X}_t, Z_t, \hat{\mu}_t)$ by

$$\begin{aligned} \mathcal{G}^{\pi_t, c_t} V(t, \hat{X}_t, Z_t, \hat{\mu}_t) &= V_t - \alpha(t)Z_t V_t - \lambda(\hat{\mu}_t - \bar{\mu})V_\eta + \frac{(\hat{\Sigma}(t) + \sigma_S \sigma_\mu \rho)^2}{2\sigma_S^2} V_{\eta\eta} - c_t V_x \\ &+ c_t \delta(t) V_z + \frac{(c_t - Z_t)^p}{p} + \pi_t \hat{\mu}_t V_x + \frac{1}{2} \sigma_S^2 \pi_t^2 V_{xx} + V_{x\eta} (\hat{\Sigma}(t) + \sigma_S \sigma_\mu \rho) \pi_t. \end{aligned}$$

For any localizing sequence τ_n , by integrating (B.5) on $[k, \tau_n \wedge T]$ and taking the expectation, we have

$$V(k, x, z, \eta) \geq \mathbb{E} \left[\int_k^{\tau_n \wedge T} \frac{(c_s - Z_s)^p}{p} ds \right] + \mathbb{E} [V(\tau_n \wedge T, \hat{X}_{\tau_n \wedge T}, Z_{\tau_n \wedge T}, \hat{\mu}_{\tau_n \wedge T})]. \tag{B.6}$$

Similarly to the argument in [17], let us consider a fixed pair of controls $(\pi_t, c_t) \in \mathcal{A} = \mathcal{A}_x$, where we denote by \mathcal{A}_x the admissible space with initial endowment x . For any $\epsilon > 0$, it is clear that $\mathcal{A}_x \subseteq \mathcal{A}_{x+\epsilon}$ and $(\pi_t, c_t) \in \mathcal{A}_{x+\epsilon}$. Also, it is easy to see that $\hat{X}_t^{x+\epsilon} = \hat{X}_t^x + \epsilon = \hat{X}_t + \epsilon, k \leq t \leq T$. As the process Z_t is defined using this consumption policy c_t , under the probability measure $\mathbb{P}_{x,z,\eta}$, we obtain

$$V(k, x + \epsilon, z, \eta) \geq \mathbb{E} \left[\int_k^{\tau_n \wedge T} \frac{(c_s - Z_s)^p}{p} ds \right] + \mathbb{E} [V(\tau_n \wedge T, \hat{X}_{\tau_n \wedge T} + \epsilon, Z_{\tau_n \wedge T}, \hat{\mu}_{\tau_n \wedge T})]. \tag{B.7}$$

The monotone convergence theorem first leads to

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_k^{\tau_n \wedge T} \frac{(c_s - Z_s)^p}{p} ds \right] = \mathbb{E} \left[\int_k^T \frac{(c_s - Z_s)^p}{p} ds \right].$$

For simplicity, let us write $Y_t = (\hat{X}_t - m(t)Z_t)$. The definition (3.7) implies that

$$V(\tau_n \wedge T, \hat{X}_{\tau_n \wedge T} + \epsilon, Z_{\tau_n \wedge T}, \hat{\mu}_{\tau_n \wedge T}) = \frac{1}{p} (Y_{\tau_n \wedge T} + \epsilon)^p N_{\tau_n \wedge T}^{1-p}.$$

Lemma B.1 gives $\hat{X}_t \geq m(t)Z_t$ for $k \leq t \leq T$ under any admissible control (π_t, c_t) , so we get that $Y_{\tau_n \wedge T} + \epsilon \geq \epsilon > 0$ for all $k \leq t \leq T$. As $p < 0$, it follows that

$$\sup_n (Y_{\tau_n \wedge T} + \epsilon)^p < \epsilon^p < +\infty. \tag{B.8}$$

Remark A.1 gives that $A(t, s) \leq 0, B(t, s)$, and $C(t, s)$ are all bounded on $k \leq t \leq s \leq T$. Also, $m(s)$ and $\delta(s)$ are continuous functions and hence bounded on $[k, T]$. Hence $N(k, \eta) \leq k_1 \exp(k_2 \eta)$, for some constants $k_2, k_1 > 1$. It follows that there exist some constants $\bar{k}_2, \bar{k}_1 > 1$ such that

$$\sup_n N_{\tau_n \wedge T}^{1-p} \leq \sup_{t \in [k, T]} (k_1 \exp(k_2 \hat{\mu}_t))^{1-p} \leq \bar{k}_1 \exp\left(\bar{k}_2 \sup_{t \in [k, T]} \hat{\mu}_t\right).$$

The process $\hat{\mu}_t$ satisfies (2.4), which leads to

$$\hat{\mu}_t = e^{-t\lambda} \eta + \bar{\mu}(1 - e^{-t\lambda}) + \int_k^t e^{\lambda(u-t)} \frac{(\hat{\Sigma}(u) + \sigma_S \sigma_{\mu} \rho)}{\sigma_S} d\hat{W}_u.$$

Hence, there exist positive constants l and $l_1 > 1$ large enough so that

$$\sup_{t \in [k, T]} \hat{\mu}_t \leq l + \sup_{t \in [k, T]} l_1 \hat{W}_t, \quad t \in [k, T].$$

Using the distribution of the running maximum of the Brownian motion, there exist some positive constants $\bar{l} > 1$ and \bar{l}_1 such that

$$\mathbb{E} \left[\sup_n N_{\tau_n \wedge T}^{1-p} \right] \leq \bar{l}_1 \mathbb{E} \left[\exp \left(\sup_{t \in [k, T]} \bar{l} \hat{B}_t \right) \right] < +\infty. \tag{B.9}$$

Finally, by (B.8) and (B.9), we can conclude that

$$\mathbb{E} \left[\sup_n V(\tau_n \wedge T, \hat{X}_{\tau_n \wedge T} + \epsilon, Z_{\tau_n \wedge T}, \hat{\mu}_{\tau_n \wedge T}) \right] < +\infty.$$

The dominated convergence theorem and $N(T, \hat{\mu}_T) = 0$ imply that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[V(\tau_n \wedge T, \hat{X}_{\tau_n \wedge T} + \epsilon, Z_{\tau_n \wedge T}, \hat{\mu}_{\tau_n \wedge T}) \right] = \mathbb{E} \left[\frac{1}{p} (Y_T + \epsilon)^p N^{1-p}(T, \hat{\mu}_T) \right] = 0.$$

Combining this with (B.7) and $(\pi_t, c_t) \in \mathcal{A}$, we have that

$$V(k, x + \epsilon, z, \eta; \theta) \geq \sup_{\pi, c \in \mathcal{A}} \mathbb{E} \left[\int_k^T \frac{(c_s - Z_s)^p}{p} ds \right] = \hat{V}(k, x, z, \eta, \theta).$$

Note that $V(t, x, z, \eta; \theta)$ is continuous in the variable x . By letting $\epsilon \rightarrow 0$, we deduce that

$$V(k, x, z, \eta; \theta) = \lim_{\epsilon \rightarrow 0} V(k, x + \epsilon, z, \eta) \geq \hat{V}(k, x, z, \eta, \theta).$$

On the other hand, for π_t^* and c_t^* given in (3.8) and (3.9), we first need to show that the SDE

$$d\hat{X}_t^* = (\pi_t^* \mu_t - c_t^*) dt + \sigma_S \pi_t^* d\hat{W}_t, \quad k \leq t \leq T, \tag{B.10}$$

with initial condition $x > m(k)z$ admits a unique strong solution that satisfies the constraint $\hat{X}_t^* > m(t)Z_t^*$ for all $k \leq t \leq T$. Let $Y_t^* = \hat{X}_t^* - m(t)Z_t^*$. By Itô's lemma and substitution of c_t^* using (3.9), we obtain that

$$dY_t^* = \left[-\frac{(1 + \delta(t)m(t))^{\frac{-p}{1-p}}}{N} + \frac{\hat{\mu}_t^2}{(1-p)\sigma_S^2} + \frac{(\hat{\Sigma}(t) + \sigma_S\sigma_\mu\rho) \frac{N_\eta}{N} \hat{\mu}_t}{\sigma_S^2} \right] Y_t^* dt + \left[\frac{\hat{\mu}_t}{(1-p)\sigma_S} + \frac{(\hat{\Sigma}(t) + \sigma_S\sigma_\mu\rho) \frac{N_\eta}{N}}{\sigma_S} \right] Y_t^* d\hat{W}_t.$$

In order to solve for X_t^* explicitly, we define the auxiliary process $\Gamma_t := \frac{N(t, \hat{\mu}_t)}{Y_t^*}$, for $k \leq t \leq T$. Itô's lemma gives that

$$d\Gamma_t = \frac{\Gamma_t}{N_t} \left[N_t - \lambda(\hat{\mu}_t - \bar{\mu})N_\eta + \frac{(\hat{\Sigma}(t) + \sigma_S\sigma_\mu\rho)^2}{2\sigma_S^2} N_{\eta\eta} + \frac{\hat{\mu}_t(\hat{\Sigma}(t) + \sigma_S\sigma_\mu\rho)p}{(1-p)\sigma_S^2} N_\eta + (1 + \delta(t)m(t))^{\frac{-p}{1-p}} + \frac{p\hat{\mu}_t^2}{(1-p)^2\sigma_S^2} N \right] dt + \Gamma_t \left[\frac{-\hat{\mu}_t}{(1-p)\sigma_S} \right] d\hat{W}_t. \tag{B.11}$$

As $N(t, \eta)$ satisfies the linear PDE (3.3), (B.11) is reduced to

$$d\Gamma_t = \Gamma_t \left[\frac{p\hat{\mu}_t^2}{2(1-p)^2\sigma_S^2} \right] dt + \Gamma_t \left[\frac{-\hat{\mu}_t}{(1-p)\sigma_S} \right] d\hat{W}_t;$$

the existence of a unique strong solution is thus verified, and $\Gamma_k = \frac{N(k, \eta)}{x - m(k)z} > 0$ implies that $\Gamma_t > 0, \forall k \leq t \leq T$. Therefore, it holds that the SDE (B.10) admits a unique strong solution as defined in (3.10), and the solution \hat{X}_t^* satisfies the constraint (B.1).

Next, we verify that the pair (π_t^*, c_t^*) is indeed in the admissible space \mathcal{A} . First, by the definitions in (3.8) and (3.9), it is clear that π_t^* and c_t^* are \mathcal{F}_t^S -progressively measurable, and by the path continuity of $Y_t^* = \hat{X}_t^* - m(t)Z_t^*$ and of π_t^* and c_t^* , it is easy to show that $\int_k^T (\pi_t^*)^2 dt < +\infty$ and $\int_k^T c_t^* dt < +\infty$, a.s. Also, because $\hat{X}_t^* > m(t)Z_t^*$ for all $t \in [k, T]$, by the definition of c_t^* , the consumption constraint $c_t^* > Z_t^*$ for all $t \in [k, T]$ is satisfied. It follows that $(\pi_t^*, c_t^*) \in \mathcal{A}$.

Given (π_t^*, c_t^*) as above, instead of (B.6), the following equality is proved:

$$V(k, x, z, \eta; \theta) = \mathbb{E} \left[\int_k^{\tau_n \wedge T} \frac{(c_t^* - Z_t^*)^p}{p} dt \right] + \mathbb{E} \left[V(\tau_n \wedge T, \hat{X}_{\tau_n \wedge T}^*, Z_{\tau_n \wedge T}^*, \hat{\mu}_{\tau_n \wedge T}) \right].$$

The monotone convergence theorem gives

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_k^{\tau_n \wedge T} \frac{(c_t^* - Z_t^*)^p}{p} dt \right] = \mathbb{E} \left[\int_k^T \frac{(c_t^* - Z_t^*)^p}{p} dt \right].$$

Moreover, as we have $V(t, x, z, \eta) < 0$ by $p < 0$, Fatou's lemma implies that

$$\limsup_{n \rightarrow +\infty} \mathbb{E} \left[V(\tau_n \wedge T, \hat{X}_{\tau_n \wedge T}^*, Z_{\tau_n \wedge T}^*, \hat{\mu}_{\tau_n \wedge T}) \right] \leq \mathbb{E} \left[V(T, \hat{X}_T^*, Z_T^*, \hat{\mu}_T) \right] = 0.$$

It follows that

$$V(k, x, z, \eta; \theta) \leq \mathbb{E} \left[\int_k^T \frac{(c_t^* - Z_t^*)^p}{p} dt \right] \leq \widehat{V}(k, x, z, \eta, \theta),$$

which completes the proof. \square

Funding information

Y. Yang and X. Yu are supported by the Hong Kong Polytechnic University under research grant no. P0031417.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process for this article.

References

- [1] AHEARNE, A. G., GRIEVER, W. L. AND WARNOCK, F. E. (2004). Information costs and home bias: an analysis of US holdings of foreign equities. *J. Internat. Econom.* **62**, 313–336.
- [2] BAYRAKTAR, E. AND SIRBU, M. (2012). Stochastic Perron's method and verification without smoothness using viscosity comparison: the linear case. *Proc. Amer. Math. Soc.* **140**, 3645–3654.
- [3] BAYRAKTAR, E. AND SIRBU, M. (2013). Stochastic Perron's method for Hamilton–Jacobi–Bellman equations. *SIAM J. Control Optimization* **51**, 4274–4294.
- [4] BAYRAKTAR, E. AND SIRBU, M. (2014). Stochastic Perron's method and verification without smoothness using viscosity comparison: obstacle problems and Dynkin games. *Proc. Amer. Math. Soc.* **142**, 1399–1412.
- [5] BAYRAKTAR, E. AND ZHANG, Y. (2015). Stochastic Perron's method for the probability of lifetime ruin problem under transaction costs. *SIAM J. Control Optimization* **53**, 91–113.
- [6] BJÖRK, T., DAVIS, M. AND LANDÉN, C. (2010). Optimal investment under partial information. *Math. Meth. Operat. Res.* **71**, 371–399.
- [7] BO, L., LIAO, H. AND YU, X. (2019). *Risk-sensitive credit portfolio optimization under partial information and contagion risk*. Preprint. Available at <https://arxiv.org/abs/1905.08004>.
- [8] BRENDLE, S. (2006). Portfolio selection under incomplete information. *Stoch. Process. Appl.* **116**, 701–723.
- [9] BRENNAN, M. J. AND XIA, Y. (2010). Persistence, predictability, and portfolio planning. In *Handbook of Quantitative Finance and Risk Management*, Springer, Boston, pp. 289–318.
- [10] CAMPBELL, J. Y. *et al.* (1997). *The Econometrics of Financial Markets*. Princeton University Press.
- [11] CONSTANTINIDES, G. M. (1990). Habit formation: a resolution of the equity premium puzzle. *J. Political Econom.* **98**, 519–543.
- [12] DETEMPLE, J. AND ZAPATERO, F. (1992). Optimal consumption–portfolio policies with habit formation. *Math. Finance* **2**, 251–274.
- [13] DUCKWORTH, J. K. AND ZERVOS, M. (2000). An investment model with entry and exit decisions. *J. Appl. Prob.* **37**, 547–559.
- [14] ENGLEZOS, N. AND KARATZAS, I. (2009). Utility maximization with habit formation: dynamic programming and stochastic PDEs. *SIAM J. Control Optimization* **48**, 481–520.
- [15] FAMA, E. F. AND FRENCH, K. R. (1989). Business conditions and expected returns on stocks and bonds. *J. Financial Econom.* **25**, 23–49.
- [16] FRIEDMAN, A. (2012). *Stochastic Differential Equations and Applications*. Dover, Mineola, NY.
- [17] JANEČEK, K. AND SIRBU, M. (2012). Optimal investment with high-watermark performance fee. *SIAM J. Control Optimization* **50**, 790–819.
- [18] KANG, J. AND STULZ, R. M. (1997). Why is there a home bias? An analysis of foreign portfolio equity ownership in Japan. *J. Financial Econom.* **46**, 3–28.
- [19] KARATZAS, I. AND SHREVE, S. (1991). *Brownian Motion and Stochastic Calculus*, 2nd edn. Springer, New York.
- [20] KEPPO, J., TAN, H. M. AND ZHOU, C. (2019). Smart city investments. Preprint. Available at <https://doi.org/10.2139/ssrn.3141043>.
- [21] KIM, T. S. AND OMBERG, E. (1996). Dynamic nonmyopic portfolio behavior. *Rev. Financial Studies* **9**, 141–161.

- [22] LAKNER, P. (1998). Optimal trading strategy for an investor: the case of partial information. *Stoch. Process. Appl.* **76**, 77–97.
- [23] LEE, J., YU, X. AND ZHOU, C. (2021). Lifetime ruin under high-water mark fees and drift uncertainty. *Appl. Math. Optimization* **84**, 2743–2773.
- [24] MEHRA, R. AND PRESCOTT, E. C. (1985). The equity premium: a puzzle. *J. Monetary Econom.* **15**, 145–161.
- [25] MONOYIOS, M. (2009). Optimal investment and hedging under partial and inside information. *Adv. Financial Model.* **8**, 371–410.
- [26] MUNK, C. (2008). Portfolio and consumption choice with stochastic investment opportunities and habit formation in preferences. *J. Econom. Dynam. Control* **32**, 3560–3589.
- [27] PHAM, H. (1997). Optimal stopping, free boundary, and American option in a jump-diffusion model. *Appl. Math. Optimization* **35**, 145–164.
- [28] PHAM, H. (2009). *Continuous-Time Stochastic Control and Optimization with Financial Applications*. Springer, Berlin, Heidelberg.
- [29] PORTES, R. AND REY, H. (2005). The determinants of cross-border equity flows. *J. Internat. Econom.* **65**, 269–296.
- [30] POTERBA, J. M. AND SUMMERS, L. H. (1988). Mean reversion in stock prices: evidence and implications. *J. Financial Econom.* **22**, 27–59.
- [31] REIKVAM, K. (1998). Viscosity solutions of optimal stopping problems. *Stoch. Stoch. Reports* **62**, 285–301.
- [32] D. REVUZ AND M. YOR (1991). *Continuous Martingales and Brownian Motion*. Springer, Berlin, Heidelberg.
- [33] SIRBU, M. (2014). Stochastic Perron’s method and elementary strategies for zero-sum differential games. *SIAM J. Control Optimization* **52**, 1693–1711.
- [34] XIA, Y. (2001). Learning about predictability: the effects of parameter uncertainty on dynamic asset allocation. *J. Finance* **56**, 205–246.
- [35] YU, X. (2015). Utility maximization with addictive consumption habit formation in incomplete semimartingale markets. *Ann. Appl. Prob.* **25**, 1383–1419.
- [36] YU, X. (2017). Optimal consumption under habit formation in markets with transaction costs and random endowments. *Ann. Appl. Prob.* **27**, 960–1002.