AN OPTIMAL CONSUMPTION AND INVESTMENT PROBLEM WITH PARTIAL INFORMATION

HIROAKI HATA,* Shizuoka University
SHUENN-JYI SHEU,** National Central University

Abstract

We consider a finite-time optimal consumption problem where an investor wants to maximize the expected hyperbolic absolute risk aversion utility of consumption and terminal wealth. We treat a stochastic factor model in which the mean returns of risky assets depend linearly on underlying economic factors formulated as the solutions of linear stochastic differential equations. We discuss the partial information case in which the investor cannot observe the factor process and uses only past information of risky assets. Then our problem is formulated as a stochastic control problem with partial information. We derive the Hamilton–Jacobi–Bellman equation. We solve this equation to obtain an explicit form of the value function and the optimal strategy for this problem. Moreover, we also introduce the results obtained by the martingale method.

Keywords: Optimal consumption and investment; HARA utility; stochastic factor model; partial information; Hamilton–Jacobi–Bellman equation

2010 Mathematics Subject Classification: Primary 49L20; 60H30; 91G10; 93E20
Secondary 93E11

1. Introduction

When optimal consumption problems on a finite time horizon are studied, the dynamic programming approach is often used to solve these problems. Adopting this approach, Hamilton–Jacobi–Bellman (HJB) equations are derived. Optimal consumption-investment strategies are obtained from the solutions of the HJB equations. For this approach, we recall the pioneering work of Merton [40] in which the author examined a hyperbolic absolute risk aversion (HARA) utility function for a model with constant interest rate and where the price of the risky asset is an exponential Brownian motion. Karatzas et al. [32] and Liu [39] studied a counterpart of [40] with a general utility and stochastic factor model, respectively. In stochastic factor models, the returns and volatilities of the assets are affected by some economic factors. See [21] for an introduction and [16]–[18], [25], [26], and [43] for stochastic factor models and infinite time horizon optimal consumption problems.

The martingale method is another useful approach to solve optimal investment problems including the consumption problems discussed in this paper; see [12], [31], and [46]. The advantage of the martingale method is that an optimal consumption-investment strategy may be derived without solving a partial differential equation (namely, the HJB equation). In
complete markets, explicit solutions of the consumption problems were obtained by Liu [39] and Wachter [54]. Indeed, in [54] the author considered the models in which the risky stock and the factor process are perfectly negatively correlated. Castaneda-Leyva and Hernandez-Hernandez [8] adopted an incomplete market approach. Using a combination of the martingale method and stochastic control techniques, the authors of [8] solved the consumption problems explicitly for the Kelly utility.

We refer the reader to [14], [23], and [30] for pioneering work on the model in continuous time with partial information. Linear Gaussian models were considered in [4], [14], [15], [23], [35], [41], [44], [47], and [56]. We follow the approach of [41] by considering the filtering equation for the estimated mean based on the observation. We use a different argument starting from the innovation process. We obtain the equation for the estimated mean using an innovation process. Using the estimated mean as a new factor process, we then have the model with complete information. The problem can be solved using either a dynamic programming approach or a duality approach using the martingale method. For the convenience of the reader, we provide the argument necessary to derive the filtering equation in Appendix A. The filtering equation is often considered in the study of models with partial observation; see, e.g., [3], [4], [23], [28], [30], [35], and [37]. See [5], [13], and [50] for the model in which the drift of a risky asset is constant and cannot be observed. In [22], [27], [47], and [52] hidden Markov models were considered. See also [9] and [36] where jump diffusion was used to model the problem.

In this paper we consider the finite time horizon optimal consumption problem with partial information. An investor wants to invest in \( m+1 \) assets including a riskless asset and \( m \) risky assets, also he/she consumes his/her wealth. In particular, we consider the linear Gaussian model in which the factor process is a Gaussian diffusion process and the returns of the assets depend linearly on the factor; see [41], [42], and [44]. In particular, this is the counterpart of [41] for the consumption problem. Note that Nagai [41] studied a risk-sensitive portfolio optimization problem on a finite time horizon.

Now we consider the following setting. The market consists of one bank account and \( m \) risky stocks. We assume that the bank account process \( S^0 \) and the price process of the risky stocks \( S := (S^1, \ldots, S^m)^\top \) are governed by:

\[
dS^0_t = r S^0_t \, dt, \quad S^0_0 = s^0, \\
dS^i_t = S^i_t \left\{ (a + AY_t)^\top \, dt + \sum_{k=1}^{n+m} \Sigma^i_k \, dW^k_t \right\}, \quad S^i_0 = s^i, \quad i = 1, \ldots, m, \\
dY_t = (b + BY_t) \, dt + \Lambda \, dW_t, \quad Y_0 = y \in \mathbb{R}^n, \tag{1.1}
\]

where \( W_t = (W^k_t)_{k=1,\ldots,(n+m)} \) is an \((m+n)\)-dimensional standard Brownian motion process. Here \(^\top\) denotes the transpose of a vector or matrix. For the coefficients, we have \( r \geq 0, \ a \in \mathbb{R}^m, b \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times n}, \Sigma \in \mathbb{R}^{m 	imes (n+m)}, \) and \( \Lambda \in \mathbb{R}^{n \times (n+m)} \). In this paper we always assume that

\[
(\text{H}) \quad \Sigma \Sigma^\top > 0.
\]

We consider an investor who invests at time \( t \) a proportion \( h^i_t \) of his/her wealth in the \( i \)th risky stock \( S^i \), \( i = 1, \ldots, m \). With \( h^i_t = (h^i_1, \ldots, h^i_m)^\top \) chosen, the proportion of the wealth invested in the bank account is \( 1 - h^\top_t \mathbf{1} \). Here \( \mathbf{1} = (1, \ldots, 1)^\top \). Let \( X^{c,h}_t \) and \( c_t X^{c,h} \) be the investor’s wealth and the rate at which wealth is consumed, respectively.

We set

\[
\tilde{g}_t := \sigma (S_u; \ u \leq t),
\]
An optimal consumption and investment problem

and we consider a strategy \((c_t, h_t)\) in the space \(\mathcal{H}_T\) of investment strategies:

\[
\mathcal{H}_T := \left\{ (c_t, h_t) \in [0, T] \times \mathbb{R}^m \text{-valued } \mathcal{g}_t \text{-progressively measurable stochastic process such that } \int_0^T c_t \, dt < \infty, \int_0^T |h_t|^2 \, dt < \infty, \mathbb{P}\text{-a.s.} \right\},
\]

where we abbreviate \(\mathbb{P}\)-almost surely to \(\mathbb{P}\)-a.s. By the self-financing condition, the investor’s wealth \(X_{c, h}^t\), starting with the initial capital \(x\), satisfies

\[
\frac{dX_{c, h}^t}{X_{c, h}^t} = \left(1 - h_t^\top \mathbf{1} \right) \frac{dS_t^0}{S_t^0} + \sum_{i=1}^m h_t^i \frac{dS_t^i}{S_t^i} - c_t \, dt, \quad X_{c, h}^0 = x,
\]

or, equivalently

\[
\frac{dX_{c, h}^t}{X_{c, h}^t} = \left[ r + h_t^\top (AY_t + a - r \mathbf{1}) - c_t \right] \, dt + h_t^\top \Sigma \, dW_t. \tag{1.2}
\]

Our goal is to select consumption and investment controls which maximize the finite horizon discounted expected HARA utility of consumption and terminal wealth:

\[
(\text{IC}) \quad V(0, x, y) := \sup_{(c, \pi) \in A_T} J(x, y; c, h; T), \rho > 0,
\]

where \(J(x, y; c, h; T)\) is defined by

\[
J(x, y; c, h; T) := \mathbb{E} \left[ \int_0^T e^{-\rho t} \frac{1}{\gamma} (c_t X_{c, h}^t)^\gamma \, dt + e^{-\rho T} \frac{1}{\gamma} (X_{c, h}^T)^\gamma \right]. \tag{1.3}
\]

Here \(A_T (\subset \mathcal{H}_T)\) is the space of admissible strategies defined later; see Section 4.

We will see that (IC) becomes a standard stochastic control problem with partial information. We regard the factor process \(Y\) as the state process, the price process \(S\) of the risky stocks as the observation process, and the consumption rate and investment policy can be considered as the control processes.

In Section 2 we reduce (IC) with partial information to a problem with full information. In order to do that we introduce an innovation process. We first derive the Kalman filter of the factor process, which is a time-inhomogeneous diffusion process. Note that the coefficients of (2.11) for the state estimate depend on the conditional variance \(\Pi_t\) of the filter which solves the Riccati equation (2.5). By using the innovation process and the Kalman filter, we are able to rewrite criterion (1.3). The Kalman filter becomes the state process. We can then reduce (IC) to a problem with full information. In this regard, we compare our approach with [41], in which the authors made a measure change before using the Kalman filter. Instead, we make the measure change after rewriting the dynamics of the risky assets using the Kalman filter and an innovation process. This seems to greatly simplify the calculation. An innovation process and filter are often used in the study of the problem for models with partial information; see, e.g. [3], [4], [23], [28], [30], [35], and [37]. For the convenience of the reader, we provide a brief discussion in Appendices A and B. For more detailed discussions, we refer the reader to [2], [11], and [38].

In Section 3, applying a dynamic programming approach, we derive a HJB equation. By a suitable transformation, the HJB equation turns out to be a linear partial differential equation. As a result, an explicit solution for the HJB equation is obtained by the solution of the time-inhomogeneous Riccati equation (3.4).
In Section 4 we prove the verification theorem. Indeed, the proof follows closely the arguments of Theorem 5.1 of Nagai [43]. The optimal value and the optimal strategy is obtained by the Kalman filter (2.3) and the solution of the time-inhomogeneous Riccati equation (3.4).

In Section 5, adopting the martingale method, we consider our consumption problem. First, we introduce the result for a general utility function. Next, we treat the power utility function in particular. Then we verify that the result obtained from the martingale method is the same as that obtained from the dynamic programming approach. In this particular case, we can apply the martingale method. We obtain an optimal strategy from the martingale representation of a random variable. To derive its feedback form, we still need to consider the associated partial differential equation. As another interesting observation, we know that the Merton consumption problem in an incomplete market with full information does not have an analytic (or explicit) solution; see [39]. For the same problem with partial information, an analytic solution may be obtained, as shown in this paper. Finally, we indicate an extension for future research, which is the consumption problem on the infinite time horizon.

2. Reduction to a corresponding control problem with full information

Let
\[ Z_i^t := \log S_i^t, \quad i = 1, \ldots, m, \]
with \( Z_t = (Z_1^t, \ldots, Z_m^t) \). Then \( Z \) solves
\[ dZ_t = (d + AY_t) dt + \Sigma dW_t, \tag{2.1} \]
with \( d = (d^i), d^i := a_i^i - \frac{1}{2}(\Sigma \Sigma^T)^{ii}. \)

In this section we compute the Kalman filter \( \hat{Y}_t \) of \( Y_t \) defined by
\[ \hat{Y}_t := \mathbb{E}[Y_t | G_t]. \]
In this case we consider the innovation process \( I_t \) defined by
\[ I_t := \int_0^t \{dZ_u - (d + A\hat{Y}_u) du\}. \tag{2.2} \]

Then we have the following results. For the convenience of the reader, we state the proof in Appendices A and B. See also [2], [11], and [38] for more detailed discussions.

**Lemma 2.1.** The innovation process \( I_t \) is a \( \mathcal{G}_t \)-Wiener process with covariance matrix \( \Sigma \Sigma^T \).

**Proof.** The proof can be found in Appendix A. \[ \square \]

**Lemma 2.2.** The Kalman filter \( \hat{Y}_t \) solves the stochastic differential equation
\[ d\hat{Y}_t = (b + B\hat{Y}_t) dt + \lambda(\Pi_t) dI_t, \quad \hat{Y}_0 = y, \tag{2.3} \]
where \( \lambda(\Pi_t) \) is defined by
\[ \lambda(\Pi_t) = (\Pi_t A^T + \Lambda \Sigma^T)(\Sigma \Sigma^T)^{-1}, \tag{2.4} \]
and \( \Pi_t = \mathbb{E}[(Y_t - \hat{Y}_t)(Y_t - \hat{Y}_t)^T | g_t] \) solves the Riccati equation
\[ \dot{\Pi}_t + (\Pi_t A^T + \Lambda \Sigma^T)(\Sigma \Sigma^T)^{-1}(A \Pi_t + \Sigma \Lambda^T) - \Lambda \Lambda^T - B\Pi_t - \Pi B^T = 0, \quad \Pi_0 = 0. \tag{2.5} \]

**Proof.** The proof can be found in Appendix B. \[ \square \]
Using (2.1) and (2.2), we have
\[ dI_t = A(Y_t - \hat{Y}_t) \, dt + \Sigma \, dW_t. \] (2.6)

Moreover, by (2.6), we have
\[ dZ_t = (d + A\hat{Y}_t) \, dt + dI_t. \] (2.7)

From (2.3) and (2.7), using \( \hat{Y}_t \) as a new factor process, we have the market with full information which has the filtration \( \{g_t\} \) and the dynamics of the risky assets become \( Z_t \) given in (2.7).

Then we can write (1.2) as
\[ dX_{t,c,h} = \left( r + h_t^\top (A\hat{Y}_t + a - r1) - c_t \right) \, dt + h_t^\top \, dI_t. \] (2.8)

Hence, we have
\[ (X_{t,c,h})^\gamma = x^\gamma \exp \left[ \gamma \int_0^t \{ \ell(\hat{Y}_u, h_u) - c_u \} \, du \right] M^h_t, \]
where \( \ell(y, h) \) and \( M^h_t \) are defined as
\[ \ell(y, h) := -\frac{1}{2}h^\top \Sigma \Sigma^\top h + r + h^\top (Ay + a - r1), \]
\[ M^h_t := \exp \left[ \gamma \int_0^t h_u^\top \, dI_u - \frac{\gamma^2}{2} \int_0^t h_u^\top \Sigma \Sigma^\top h_u \, du \right]. \]

Therefore, we can write (1.3) as
\[ J(x, y; c, h; T) = \frac{x^\gamma}{\gamma} \mathbb{E} \left[ \int_0^T c_t^\gamma \exp \left[ \int_0^t \left( \gamma \ell(\hat{Y}_u, h_u) - c_u \right) \, du \right] M^h_t \, dt \right. \]
\[ \left. + \exp \left[ \int_0^T \left( \gamma \ell(\hat{Y}_u, h_u) - c_u \right) \, du \right] M^h_T \right]. \] (2.9)

Let
\[ A_T^h := \{(c_t, h_t) \in \mathcal{H}_T; \, \mathbb{E}[M^h_T] = 1\}. \]

Then, for all \( h \in A_T^h \), we can define the probability measure \( \mathbb{P}^h \) on \((\Omega, \mathcal{F}, g_T)\) using
\[ \frac{d\mathbb{P}^h}{d\mathbb{P}} \bigg|_{g_T} = M^h_T. \] (2.10)

Under the probability measure \( \mathbb{P}^h \), we have
\[ d\hat{Y}_t = \{ b + B\hat{Y}_t + \gamma \lambda(\Pi_t) \Sigma \Sigma^\top h_t \} \, dt + \lambda(\Pi_t) \, dI_t^h, \quad \hat{Y}_0 = y, \] (2.11)

where \( I_t^h \) defined by \( I_t^h := I_t - \gamma \int_0^t \Sigma \Sigma^\top h_u \, du \) is a \( g_t \)-Wiener process with covariance \( \Sigma \Sigma^\top t \). Then (2.9) can be expressed as
\[ J(x, y; c, h; T) = \frac{x^\gamma}{\gamma} \mathbb{E}^h \left[ \int_0^T c_t^\gamma \exp \left[ \int_0^t \left( \gamma \ell(\hat{Y}_u, h_u) - c_u \right) \, du \right] \, dt \right. \]
\[ \left. + \exp \left[ \int_0^T \left( \gamma \ell(\hat{Y}_u, h_u) - c_u \right) \, du \right] \right]. \] (2.12)

We conclude that (IC) can be reduced to a stochastic control problem with full information, which is the problem of maximizing (2.12) subject to a system process \( \hat{Y}_t \) given by (2.11) on the filtered probability space \((\Omega, \mathcal{F}, g_T, \mathbb{P}^h)\).
3. The HJB equation and its explicit solution

In this section we will derive the HJB equation for the new control problem with full information obtained in Section 2. We will obtain an explicit solution for the HJB equation.

We first introduce the value functions

$$ V(t, x, y) = \sup_{(c, h) \in A^1_{t,T}} \gamma J(y; c, h; [t, T]) $$

or, equivalently,

$$ V(t, x, y) = \gamma v(t, y), $$

where

$$ v(t, y) := \sup_{(c, h) \in A^1_{t,T}} \gamma \hat{J}(y; c, h; [t, T]), \quad \gamma \in (0, 1), $$

$$ v(t, y) = \inf_{(c, h) \in A^1_{t,T}} \gamma \hat{J}(y; c, h; [t, T]), \quad \gamma \in (-\infty, 0). $$

Here $\hat{J}(y; c, h; [t, T])$ is defined by

$$ \hat{J}(y; c, h; [t, T]) = E^{h}_{t,y} \left[ \int_t^T c_S \exp \left[ \int_t^s \gamma \left( \ell(\hat{Y}_u, h_u) - c_u \right) - \rho \right] du \right] dt + \exp \left[ \int_t^T \gamma \left( \ell(\hat{Y}_u, h_u) - c_u \right) - \rho \right]. $$

We denote by $E^{h}_{t,y}$ the expectation with respect to a probability $P^h_{t,y}$ defined similarly to $P^h$, but with $\hat{Y}_t = y$. Also, $A^1_{t,T}$ is the restriction of the space $A^1_T$ on the time interval $[t, T]$. Then, using the dynamic programming approach (see [20, Chapter III.7]), we can formally deduce the following HJB equation for $v$ with the dynamic (2.11):

$$ \frac{\partial v}{\partial t} + \frac{1}{2} \text{tr} \left[ (\lambda(\Pi_t))^T \Sigma (\lambda(\Pi_t))^T D^2 v \right] + (b + By)^T Dv - \rho v + \gamma \sup_{c \geq 0} \left( -cv + \frac{c^\gamma}{\gamma} \right) $$

$$ + \gamma \sup_{h \in \mathbb{R}^m} \left\{ \ell(y, h) + h^T \Sigma (\lambda(\Pi_t))^T Du \right\} v $$

$$ = 0, \quad t < T, \quad v(T, y) = 1. \quad (3.1) $$

We provide a formal derivation of (3.1) in Appendix C. This holds for $\gamma \in (-\infty, 0) \cup (0, 1)$. We define $u(t, y)$ by the relation $v(t, y) := u(t, y)^1-\gamma$. Then $u$ satisfies

$$ \frac{\partial u}{\partial t} + \frac{1}{2} \text{tr} \left[ (\lambda(\Pi_t))^T \Sigma (\lambda(\Pi_t))^T D^2 u \right] - \frac{\rho}{1-\gamma} u + \gamma \sup_{c \geq 0} \left( -cu^{1-\gamma} + \frac{c^\gamma}{\gamma} \right) u^{\gamma} $$

$$ + \gamma \sup_{h \in \mathbb{R}^m} \left\{ \ell(y, h) + (1-\gamma)h^T \Sigma (\lambda(\Pi_t))^T Du \right\} u $$

$$ = 0, \quad t < T, \quad u(T, y) = 1. \quad (3.2) $$
Simplifying this equation, we have
\[
\frac{\partial u}{\partial t} + L_t u + 1 = 0, \quad t < T, \quad u(T, y) = 1. \tag{3.3}
\]

The supremum in (3.2) is attained by \((\tilde{c}(t, y), \tilde{h}(t, y))\):
\[
\tilde{c}(t, y) := \frac{1}{1 - \gamma} (\Sigma \Sigma^\top)^{-1} \left\{ Ay + a - r \mathbf{1} + (1 - \gamma) \Sigma \lambda(\Pi_t) \Sigma^\top \frac{Du}{u} \right\},
\]
\[
\tilde{h}(t, y) := \frac{1}{1 - \gamma} \left\{ b + \frac{\gamma}{1 - \gamma} \lambda(\Pi_t) (a - r \mathbf{1}) + \left\{ B + \frac{\gamma}{1 - \gamma} \lambda(\Pi_t) A \right\} y \right\}^\top Df + \frac{1}{1 - \gamma} \left\{ \frac{\gamma}{2(1 - \gamma)} (Ay + a - r \mathbf{1})^\top (\Sigma \Sigma^\top)^{-1} (Ay + a - r \mathbf{1}) + \gamma r - \rho \right\} f.
\]

Here \(L_t\) is the differential operator defined by
\[
L_t f := \frac{1}{2} \text{tr} [\lambda(\Pi_t) \Sigma \Sigma^\top \lambda(\Pi_t) \Sigma^\top D^2 f]
\]
\[
+ \left\{ b + \frac{\gamma}{1 - \gamma} \lambda(\Pi_t) (a - r \mathbf{1}) + \left\{ B + \frac{\gamma}{1 - \gamma} \lambda(\Pi_t) A \right\} y \right\}^\top Df
\]
\[
+ \frac{1}{1 - \gamma} \left\{ \frac{\gamma}{2(1 - \gamma)} (Ay + a - r \mathbf{1})^\top (\Sigma \Sigma^\top)^{-1} (Ay + a - r \mathbf{1}) + \gamma r - \rho \right\} f.
\]

We now consider the following time-inhomogeneous Riccati equation:
\[
\dot{U}(t) + U(t) K_2(t) U(t) + K_1(t)^\top U(t) + U(t) K_1(t) + K_0 = 0, \quad t < s, \quad U(s) = 0, \tag{3.4}
\]
where
\[
K_2(t) := \frac{1}{1 - \gamma} \lambda(\Pi_t) \Sigma \Sigma^\top \lambda(\Pi_t) \Sigma^\top \geq 0, \quad K_1(t) := B + \frac{\gamma}{1 - \gamma} \lambda(\Pi_t) A,
\]
\[
K_0 := \frac{\gamma}{1 - \gamma} A^\top (\Sigma \Sigma^\top)^{-1} A.
\]

We also use \(U(t) = U(t; s)\) for the dependence of \(U(t)\) on \(s \in [t, T]\). The term \(g(t) = g(t; s)\) is the solution of a linear differential equation:
\[
\dot{g}(t) + [K_1(t) + K_2(t) U(t)]^\top g(t) + U(t) b + \frac{\gamma}{1 - \gamma} (A^\top (\Sigma \Sigma^\top)^{-1} + U(t) \lambda(\Pi_t)) (a - r \mathbf{1}) = 0, \quad t < s, \quad g(s) = 0, \tag{3.5}
\]
and \(l(t) = l(t; s)\) is the solution of
\[
\dot{l}(t) + \frac{1}{2} \text{tr} [\lambda(\Pi_t) \Sigma \Sigma^\top \lambda(\Pi_t) U(t)] + \frac{1}{2} g(t)^\top \lambda(\Pi_t) \Sigma \Sigma^\top \lambda(\Pi_t) g(t) + b^\top g(t) + \gamma r
\]
\[
+ \frac{\gamma}{2(1 - \gamma)} (a - r \mathbf{1}) + \Sigma \Sigma^\top \lambda(\Pi_t) g(t))^\top (\Sigma \Sigma^\top)^{-1} (a - r \mathbf{1} + \Sigma \Sigma^\top \lambda(\Pi_t) g(t))
\]
\[
= 0, \quad t < s, \quad l(s) = 0. \tag{3.6}
\]

Then we have the following results.

**Theorem 3.1.** If (3.4)–(3.6) have solutions \(U(t; s), g(t; s), \) and \(l(t; s)\), then \(\bar{u}(t, y)\) defined by
\[
\bar{u}(t, y) := \int_t^T \tilde{w}(t; s, y) \, ds + \tilde{w}(t; T, y) \tag{3.7}
\]
solves (3.3), where \( \tilde{w}(t; s, y) \), \( s \in [t, T] \), is defined by
\[
\tilde{w}(t; s, y) := \exp\left[ \frac{1}{1-\gamma} \left( \frac{1}{2} y^T U(t; s) y + g(t; s)^T y + l(t; s) - \rho (s-t) \right) \right].
\] (3.8)

In particular, corresponding to the value function \( V(t, x, y) \) for (IC), we define
\[
V(t, x, y) = \frac{x}{\gamma} \tilde{w}(t, y)^{1-\gamma}. \] (3.9)

Proof. A straightforward calculation shows that \( w(t, y) \equiv \tilde{w}(t; s, y) \), given by (3.8), solves
\[
\frac{\partial w}{\partial t} + L_t w = 0, \quad t < s, \quad w(s, y) = 1.
\]
Moreover, using an argument similar to that of Lemma 2 of Liu [39], we see that (3.7) solves (3.3). \( \Box \)

Remark 3.1. In the \( \gamma < 0 \) case, it is known that (3.4) has a unique solution (see Lemma 5.2 of [6, Chapter V], or Theorem 5.2 of [19, Chapter 5]).

4. Verification theorem

In this section we prove the verification theorem. In order to do so, we first rewrite (3.2). Define \( \eta \) by
\[
\eta(t, y) := (1-\gamma) \log \tilde{u}(t, y). \] (4.1)

Then, using (3.2) and Theorem 3.1, \( \eta \) solves
\[
\frac{\partial \eta}{\partial t} + \frac{1}{2} \text{tr} [\lambda(\Pi_t) \Sigma \Sigma^T \lambda(\Pi_t) D^2 \eta] + \frac{1}{2} (D\eta)^T \lambda(\Pi_t) \Sigma \Sigma^T \lambda(\Pi_t)^T D\eta + (b + By)^T D\eta - \rho
\]
\[
+ y \sup_{c \geq 0} \left( -c + \frac{e^\eta}{\gamma} \right) + \gamma \sup_{h \in \mathbb{R}^m} \{ \ell(y, h) + h^T \Sigma \Sigma^T \lambda(\Pi_t)^T D\eta \}
\]
\[
= 0, \quad t < T, \quad \eta(T, y) = 0. \] (4.2)

Next we define the space of admissible strategies
\[
\mathcal{A}_T := \{(c, h) \in \mathcal{A}_1^T; \ E^h[\overline{M}_T^h] = 1\},
\]
where \( \overline{M}_T^h \) is defined by
\[
\overline{M}_T^h := \exp\left[ \int_0^t (D\eta(u, \hat{Y}_u))^T \lambda(\Pi_u) dI_u^h
\]
\[
- \frac{1}{2} \int_0^t (D\eta(u, \hat{Y}_u))^T \lambda(\Pi_u) \Sigma \Sigma^T \lambda(\Pi_u)^T D\eta(u, \hat{Y}_u) du \right].
\]

Then we have the following theorem.

Theorem 4.1. Assume that (H) holds and (3.4)–(3.6) have solutions \( U(t) \), \( g(t) \), and \( \ell(t) \). Note that \( \eta(t, y) \) is given by (4.1). Define
\[
\tilde{c}(t, y) := e^{-\eta(t, y)/(1-\gamma)},
\]
\[
\tilde{h}(t, y) := \frac{1}{1-\gamma} \left( \Sigma \Sigma^T \right)^{-1} \left[ a - r \mathbf{1} + Ay + \Sigma \Sigma^T \lambda(\Pi_t)^T D\eta(t, y) \right].
\]
Then
\[(\hat{c}(t, \hat{Y}_t), \hat{h}(t, \hat{Y}_t)) \in A_T (4.3)\]
is an optimal strategy for (IC). Moreover, \(V(0, x, y) = \overline{V}(0, x, y)\). Here, from (3.9) and (4.1), \(V\) is given by
\[V(0, x, y) = x\gamma e^{\eta(0, y)}. (4.4)\]

**Proof.** We apply the idea of the proof of Theorem 5.1 of [43]. First, for any \((c_t, h_t) \in A_T\), we will show that
\[J(x, y; c, h; T) \leq V(0, x, y). (4.5)\]
Now, we consider the \(\gamma \in (0, 1)\) case. We apply Itô’s differential rule to \(\eta(t, \hat{Y}_t)\). Note that \(\hat{Y}_t\) is given in (2.11). Then we have
\[
\eta(t, \hat{Y}_t) - \eta(0, y) = \int_0^t (D\eta(u, \hat{Y}_u))^\top \lambda(\Pi_u) dI^h_u \\
+ \int_0^t \left[ \frac{\partial \eta}{\partial u}(u, \hat{Y}_u) + \frac{1}{2} tr(\lambda(\Pi_u)\Sigma^\top \lambda(\Pi_u)\Sigma^\top h_u) \\
+ (D\eta(u, \hat{Y}_u))^\top [b + B\hat{Y}_u + \gamma \lambda(\Pi_u)\Sigma^\top h_u] \right] du \\
\leq \int_0^t [\rho - \gamma (\ell(\hat{Y}_u, h_u) - c_u) - c_u e^{-\eta(u, \hat{Y}_u)}] du + \log M^h_t,
\]
where the second inequality follows from (4.2). Hence, we obtain
\[\gamma \int_0^t [\ell(\hat{Y}_u, h_u) - c_u] du - \rho t \leq \eta(0, y) - \eta(t, \hat{Y}_t) + \log \psi^c_t + \log M^h_t. (4.6)\]
where \(\psi^c_t\) is defined by
\[\psi^c_t := \exp \left[ - \int_0^t c_u^\gamma e^{-\eta(u, \hat{Y}_u)} du \right].\]
Moreover, using (2.12) and (4.6), we have
\[J(x, y; c, h; T) \leq x^\gamma e^{\eta(0, y)} E^h \left[ \int_0^T c^\gamma_t e^{-\eta(t, \hat{Y}_t)} \psi^c_t M^h_T dt + \psi^c_T M^h_T \right] \\
= V(0, x, y) E^h \left[ - \int_0^T d\psi^c_t + \psi^c_T \right] \\
= \overline{V}(0, x, y).
\]
Here \(E^b [\cdot]\) is the expectation with respect to the probability measure \(P^b\):
\[\frac{dP^b}{dP^h} \bigg|_{\mathcal{F}_t} = \frac{M^h_t}{M_t^h}.\]
Next, we consider the \(\gamma \in (-\infty, 0)\) case. In a similar way to the above, we have
\[\gamma \int_0^t [\ell(\hat{Y}_u, h_u) - c_u] du - \rho t \geq \eta(0, y) - \eta(t, \hat{Y}_t) + \log \psi^c_t + \log M^h_t.
\]
Hence, we have (4.5) again.
Next, we take \((\hat{c}_t, \hat{h}_t)\). By Theorem 3.1, we have \(|\hat{h}_t| \leq C(1 + |\hat{Y}_t|)\). Furthermore, following the arguments of Lemma 4.1.1 of [2], we have \(E[\hat{M}_T^\hat{h}] = 1\) and \(E[\hat{M}_T^{\hat{h}_T}] = 1\). Hence, we verify that \((\hat{c}_t, \hat{h}_t) \in A_T\). Letting \(\gamma \in (\infty, 0) \cup (0, 1)\), we will show that \(J(x, y, \hat{c}, \hat{h}; T) = \mathbb{V}(0, x, y)\).

Now, we can define the probability measure \(\mathbb{P}^{\hat{h}}\) (see (2.10)), and we see that under \(\mathbb{P}^{\hat{h}}\), \(\hat{Y}_t\) solves
\[
d\hat{Y}_t = \{b + B\hat{Y}_t + \gamma \lambda(\Pi_t)\Sigma^\top \hat{h}_t\} dt + \lambda(\Pi_t) dI_t^{\hat{h}}, \quad \hat{Y}_0 = y.
\]

Then we have
\[
\eta(t, \hat{Y}_t) - \eta(0, y) = \int_0^t (D\eta(u, \hat{Y}_u)) \lambda(\Pi_u) dI_u^{\hat{h}}
+ \int_0^t \left[ \frac{\partial\eta}{\partial u}(u, \hat{Y}_u) + \frac{1}{2} \text{tr}(\lambda(\Pi_u)\Sigma^\top \lambda(\Pi_u)^\top D^2 \eta(u, \hat{Y}_u))
+ (D\eta(u, \hat{Y}_u)) \{b + B\hat{Y}_u + \gamma \lambda(\Pi_u)\Sigma^\top \hat{h}_u\} \right] du
= \int_0^t \left[ \rho - \gamma \{\ell(\hat{Y}_u, \hat{h}_u) - \hat{c}_u\} - c_u^\gamma e^{-\eta(u, \hat{Y}_u)} \right] du + \log \hat{M}_t^{\hat{h}},
\]

where the second equality follows from the fact that \((\hat{c}, \hat{h})\) is the maximizer in (4.2). Hence, we have
\[
\gamma \int_0^t \{\ell(\hat{Y}_u, \hat{h}_u) - \hat{c}_u\} du - \rho t = \eta(0, y) - \eta(t, \hat{Y}_t) + \log \hat{M}_T^{\hat{h}}.
\]

Moreover, we have
\[
J(x, y, \hat{c}, \hat{h}; T) = \frac{\chi_T^\gamma}{\gamma} e^{\eta(0, y)} \mathbb{E}^{\hat{h}} \left[ \int_0^T \hat{c}_u^\gamma e^{-\eta(u, \hat{Y}_u)} \psi_T^\gamma \hat{M}_T^{\hat{h}} du + \psi_T^\gamma \hat{M}_T^{\hat{h}} \right]
= \mathbb{V}(0, x, y) \mathbb{E}^{\hat{h}} \left[ - \int_0^T d\psi_T^\gamma + \psi_T^\gamma \right]
= \mathbb{V}(0, x, y). \quad \Box
\]

5. The martingale method

In this section we consider (IC) by adopting the martingale method. In Subsection 5.1 we treat the general utility case. In Subsection 5.2 we treat the power utility case. In addition, we check that the optimal strategy obtained by the martingale method accords with the optimal strategy (4.3) by the dynamic programming approach. In Subsection 5.3 we introduce an infinite time horizon optimal consumption problem as an extension for future research. Actually, the problem on the infinite time horizon is not completely solved. We state the difficulty faced.

5.1. General utility

In this subsection we use the argument of Karatzas and Shreve [29, Chapter 3]. Use (2.3) and (2.8). Define
\[
H_t^0 := e^{-rt} M_t^0, \quad (5.1)
\]
An optimal consumption and investment problem

where $M^0_t$ is defined by

$$M^0_t := \exp \left[ -\int_0^t (a - r \mathbf{1} + A \tilde{Y}_u) \mathbf{\Sigma} \mathbf{\Sigma}^{-1} dI_u ight]$$

Using (2.8) and (5.1), we have

$$d(H^0_t X_{c,h}) = H^0_t dX_{c,h} + X_{c,h} dH^0_t + d\leftlangle H^0 \cdot, X_{c,h} \cdot \right\rangle.$$ 

Then, the left-hand side of (5.2) is nonnegative and a $\mathcal{F}_t$-local martingale. That is, it is a $\mathcal{F}_t$-supermartingale. Therefore, we have

$$E\left[ H^0_T X_{c,h} + \int_0^T H^0_t \bar{C}_t dI_t \right] \le x$$

for all $(c, h) \in \mathcal{H}_T$, where $\mathcal{H}_T$ is given in Section 1. Then, following similar arguments as in Theorem 3.5 of [29, Chapter 3], we have the following proposition.

**Proposition 5.1.** Assume that a contingent claim $\xi \in \mathcal{F}_T$, $\xi \ge 0$, and $\bar{C}_t \ge 0$ such that

$$E\left[ H^0_T \xi + \int_0^T H^0_u \bar{C}_u du \right] = x.$$ 

In addition, define

$$c_t := \frac{\bar{C}_t}{X_t},$$

where $X_t$ is defined by

$$X_t := \frac{1}{H^0_t} E\left[ H^0_T \xi + \int_0^T H^0_u \bar{C}_u du \mid \mathcal{F}_t \right].$$ 

Then there is an $h$ such that $(c, h) \in \mathcal{H}_T$ and $X^{c,h}_t = \xi$, $X^{c,h}_T = X_t$. Indeed, $h$ is given by

$$h_t := (\mathbf{\Sigma} \mathbf{\Sigma}^{-1})(a - r \mathbf{1} + A \tilde{Y}_t) + \frac{1}{H^0_t X_t} \psi_t,$$

where $\psi$ is a $\mathcal{F}_t$-progressively measurable integrand $\psi: [0, T] \times \Omega \rightarrow \mathbb{R}^m$ such that the following martingale representation theorem holds:

$$H^0_t X^{c,h}_t + \int_0^T H^0_t \bar{C}_t dt = x + \int_0^T \psi_u dI_u.$$
Let \( U_i : [0, \infty) \to \mathbb{R} \cup \{-\infty\} \) \((i = 1, 2)\) be the general utility functions. We assume that \( U_i \) is strictly increasing, strictly concave, and twice continuously differentiable. In addition, we assume that
\[
\lim_{x \to \infty} U_i'(x) = 0 \quad \text{and} \quad \lim_{x \downarrow 0} U_i'(x) = \infty.
\]
Moreover, \( I_i : (0, \infty) \to (0, \infty) \) denotes the inverse function of \( U_i' \).

Now we consider the following problem:

\[
(GU) \quad V_0(x, y) := \sup_{(c, h) \in \tilde{A}_T} \mathbb{E} \left[ \int_0^T e^{-\rho t} U_1(c_t X_{\tilde{c}, h_t}^c) \, dt + e^{-\rho T} U_2(X_T^{c, h}) \right], \quad \rho > 0,
\]

where \( \tilde{A}_T \) is the space of admissible strategies defined by
\[
\tilde{A}_T := \left\{ (c, h) \in \mathcal{H}_T; \mathbb{E} \left[ \int_0^T e^{-\rho t} U_1^- (c_t X_{\tilde{c}, h}^c) \, dt \right] < \infty, \mathbb{E} \left[ e^{-\rho T} U_2^- (X_T^{c, h}) \right] < \infty \right\}.
\]

Define
\[
\chi(k) := \mathbb{E} \left[ \int_0^T H_0^0 l_1(ke^{\rho t} H_0^0) \, dt + H_0^0 l_2(ke^{\rho T} H_0^0) \right].
\]

Then, following closely the arguments of Theorem 6.3 of [29, Chapter 3], we have the following proposition.

**Proposition 5.2.** Assume that \( \chi(k) < \infty \) for all \( k > 0 \). Define \( \tilde{k} \) such that \( \chi(\tilde{k}) = x \).

Then
\[
\tilde{C}_t := I_1(\tilde{k} e^{\rho t} H_0^0) \quad \text{and} \quad \tilde{\xi} := I_2(\tilde{k} e^{\rho T} H_0^0)
\]

lead to
\[
V_0(x, y) := \mathbb{E} \left[ \int_0^T e^{-\rho t} U_1(\tilde{C}_t) \, dt + e^{-\rho T} U_2(\tilde{\xi}) \right].
\]

In addition, the optimal strategy \((\tilde{c}_t, \tilde{h}_t) \in \tilde{A}_T\) is defined by
\[
\tilde{c}_t := \frac{\tilde{C}_t}{X_{\tilde{c}, h}^c} \quad \text{and} \quad \tilde{h}_t := (\Sigma \Sigma^T)^{-1} (a - r 1 + A \tilde{Y}_t) + \frac{1}{H_0^0 X_{\tilde{c}, h}^c} \psi_t,
\]

where \( \psi \) is a \( \mathcal{G}_t \)-progressively measurable integrand in the martingale representation
\[
H_0^0 X_{\tilde{c}, h}^c + \int_0^t H_0^0 \tilde{c}_u X_u^{\tilde{c}, h} \, du = x + \int_0^t \psi_u \, dI_u.
\]

Here \( X_{\tilde{c}, h}^c \) is the optimal wealth process defined by
\[
X_{\tilde{c}, h}^c := \frac{1}{H_0^0} \mathbb{E} \left[ H_0^0 \tilde{\xi} + \int_t^T H_0^0 \tilde{C}_u \, du \mid \mathcal{G}_t \right] \quad \text{and} \quad X_{T}^{c, h} = \tilde{\xi}.
\]

### 5.2. Power utility functions

In this subsection we consider the case of the power utility functions. Under this setting, we obtain the optimal wealth process, the optimal strategy, and the value function. We confirm that the same result as Theorem 4.1 is obtained.
An optimal consumption and investment problem

First, we consider

\[ U_1(x) = U_2(x) = \frac{x^\gamma}{\gamma} \quad \text{for } \gamma \in (-\infty, 0) \cup (0, 1). \]

Then we have

\[ I_1(y) = I_2(y) = y^{-1/(1-\gamma)}. \]

First, we observe that

\[
\chi(k) = \mathbb{E}\left[ \int_0^T H_0^T (k e^{\rho t} H_0^T) \frac{1}{1-\gamma} dt + H_0^T (k e^{\rho t} H_0^T) \frac{1}{1-\gamma} \right]
\]

\[
= k^{-1/(1-\gamma)} \mathbb{E}\left[ \int_0^T e^{-\rho t/(1-\gamma)} (H_0^T)^{\gamma/(1-\gamma)} dt + e^{-\rho T/(1-\gamma)} (H_0^T)^{-\gamma/(1-\gamma)} \right]
\]

\[
= k^{-1/(1-\gamma)} \mathbb{E}\left[ \int_0^T \hat{M}_t \exp\left[ \int_0^t \phi(\hat{Y}_u) \, du \right] dt + \hat{M}_T \exp\left[ \int_0^T \phi(\hat{Y}_t) \, dt \right] \right]. \quad (5.6)
\]

where \( \hat{M}_t \) and \( \phi(y) \) are defined as follows:

\[
\hat{M}_t := \exp\left[ \frac{\gamma}{1-\gamma} \int_0^t (a - r_1 + A\hat{Y}_u) (\Sigma \Sigma^T)^{-1} \, du - \frac{1}{2} \left( \frac{\gamma}{1-\gamma} \right)^2 \int_0^t (a - r_1 + A\hat{Y}_u) (\Sigma \Sigma^T)^{-1} (a - r_1 + A\hat{Y}_u) \, du \right],
\]

\[
\phi(y) := \frac{1}{1-\gamma} \left[ -\rho + \gamma r + \frac{\gamma}{2(1-\gamma)} (a - r_1 + Ay) (\Sigma \Sigma^T)^{-1} (a - r_1 + Ay) \right].
\]

Now, we define the probability measure \( \tilde{P} \) as

\[
\frac{d\tilde{P}}{dP} \bigg|_{\mathcal{G}_t} = \tilde{M}_t.
\]

Under the probability measure \( \tilde{P} \), \( \hat{Y}_t \) solves

\[
d\hat{Y}_t = \left\{ b + B\hat{Y}_t + \frac{\gamma}{1-\gamma} \lambda(\Pi_t)(a - r_1 + A\hat{Y}_t) \right\} dt + \lambda(\Pi_t) \, d\tilde{I}_t,
\]

where \( \tilde{I}_t \) is defined by

\[
\tilde{I}_t := I_t - \frac{\gamma}{1-\gamma} \int_0^t (a - r_1 + A\hat{Y}_u) \, du.
\]

Note that \( \tilde{I}_t \) is a \( \mathcal{G}_t \)-Wiener process with covariance matrix \( \Sigma \Sigma^T \) under \( \tilde{P} \) (see Lemma 2.1).

Then we have

\[
\chi(k) = k^{-1/(1-\gamma)} f(0, y),
\]

where \( f(t, y) \) is defined by

\[
f(t, y) = \tilde{E}\left[ \int_t^T \exp\left[ \int_t^s \phi(\hat{Y}_u) \, du \right] ds + \exp\left[ \int_t^T \phi(\hat{Y}_s) \, ds \right] \right],
\]

https://doi.org/10.1017/apr.2018.7 Published online by Cambridge University Press.
where \( \hat{\mathbb{E}}[\cdot] \) is the expectation with respect to the probability measure \( \hat{\mathbb{P}} \). Noting that \( f(t, y) \) solves (3.3), we see that

\[
f(t, y) = \hat{u}(t, y),
\]

where \( \hat{u} \) is given by (3.7). Hence, from (5.6), we have

\[
\chi(k) = k^{-1/(1-\gamma)} \hat{u}(0, y).
\]

Moreover, noting that \( x = \chi(\hat{k}) \), we have

\[
\hat{k} = x^{-(1-\gamma)}(0, y)^{1-\gamma}.
\] (5.7)

Then we have

\[
H_{\hat{t} \xi}^0 \hat{\xi} = H_{\hat{t} \xi}^0 J_2 \hat{\xi} = \hat{u}(0, y)^{-1} \exp \left[ \int_0^{\hat{t}} \phi(\hat{\gamma}) \, dt \right] H_\xi^0 - (1/(1-\gamma)) \hat{u}(0, y)^{1-\gamma} \exp \left[ \int_0^{\hat{t}} \phi(\hat{\gamma}) \, dt \right]
\] (5.8)

and, similarly,

\[
\hat{C}_t := x \hat{u}(0, y)^{-1} \hat{M}_t \exp \left[ \int_0^t \phi(\hat{\gamma}_u) \, du \right] \left( H_\xi^0 \right)^{-1}.
\] (5.9)

Using (5.8) and (5.9), we have

\[
H_{\hat{t} \xi}^0 \hat{X}_{\hat{t}}^{\hat{c}, \hat{h}} = x \hat{u}(0, y)^{-1} \mathbb{E} \left[ \int_0^\hat{t} \hat{M}_t \exp \left[ \int_0^t \phi(\hat{\gamma}_u) \, du \right] \right. \left. \exp \left[ \int_0^{\hat{t}} \phi(\hat{\gamma}_u) \, du \right] \right] \bigg| \mathcal{G}_\hat{t}
\]

\[
= x \hat{u}(0, y)^{-1} \hat{M}_t \exp \left[ \int_0^t \phi(\hat{\gamma}_u) \, du \right]
\]

\[
\times \mathbb{E} \left[ \int_0^\hat{t} \hat{M}_t \exp \left[ \int_0^t \phi(\hat{\gamma}_u) \, du \right] \right. \left. \exp \left[ \int_0^{\hat{t}} \phi(\hat{\gamma}_u) \, du \right] \right] \bigg| \mathcal{G}_\hat{t}
\]

\[
= x \hat{u}(0, y)^{-1} \hat{M}_t \exp \left[ \int_0^t \phi(\hat{\gamma}_u) \, du \right]
\]

\[
\times \hat{\mathbb{E}} \left[ \int_0^\hat{t} \exp \left[ \int_0^t \phi(\hat{\gamma}_u) \, du \right] \right. \left. \exp \left[ \int_0^{\hat{t}} \phi(\hat{\gamma}_u) \, du \right] \right] \bigg| \mathcal{G}_\hat{t}
\]

\[
= x \hat{u}(0, y)^{-1} \hat{M}_t \exp \left[ \int_0^t \phi(\hat{\gamma}_u) \, du \right] \left( \hat{u}(0, y)^{1-\gamma} \right) \exp \left[ \int_0^{\hat{t}} \phi(\hat{\gamma}_u) \, du \right]
\]

\[
= x \hat{u}(0, y)^{-1} \hat{M}_t \exp \left[ \int_0^t \phi(\hat{\gamma}_u) \, du \right] \hat{u}(t, \hat{\gamma}_t).
\] (5.10)

From (5.4), (5.9), and (5.10), we have

\[
\hat{c}_t := \hat{u}(t, \hat{\gamma}_t)^{-1}.
\] (5.11)

Set

\[
\hat{M}_t := H_{\hat{t} \xi}^0 \hat{X}_{\hat{t}}^{\hat{c}, \hat{h}} + \int_0^\hat{t} H_{\hat{t} \xi}^0 \hat{c}_u \hat{X}_{\hat{t}}^{\hat{c}, \hat{h}} \, du.
\]
Using (5.10) and (5.11), we have
\[
\tilde{M}_t = x \tilde{u}(0, y)^{-1} \left\{ \tilde{M}_t \exp \left[ \int_0^t \phi(\tilde{Y}_a) \, du \right] \tilde{u}(t, \tilde{Y}_t) + \int_0^t \tilde{M}_s \exp \left[ \int_0^s \phi(\tilde{Y}_a) \, ds \right] ds \right\}.
\]

Here, we see that
\[
d\tilde{M}_t = \frac{\gamma}{1 - \gamma} \tilde{M}_t (a - r 1 + A \tilde{Y}_t)^\top (\Sigma \Sigma^\top)^{-1} \, dI_t, \\
d\tilde{u}(t, \tilde{Y}_t) = \left\{ \frac{\partial \tilde{u}(t, \tilde{Y}_t)}{\partial t} + \mathcal{L}_u(t, \tilde{Y}_t) + 1 \right\} dt \\
\]
\[
\quad + \left\{ \tilde{u}(t, \tilde{Y}_t) \frac{\gamma}{1 - \gamma} \tilde{M}_t (a - r 1 + A \tilde{Y}_t)^\top (\Sigma \Sigma^\top)^{-1} + (D \tilde{u}(t, \tilde{Y}_t))^\top (\Sigma \Sigma^\top)^{-1} \right\} \, dI_t
\]
\[
= H^0_1 X_t^{\tilde{h}_t} \left\{ \frac{\gamma}{1 - \gamma} (\Sigma \Sigma^\top)^{-1} (a - r 1 + A \tilde{Y}_t) + (\Sigma \Sigma^\top)^{-1} \frac{D \tilde{u}(t, \tilde{Y}_t)}{\tilde{u}(t, \tilde{Y}_t)} \right\} \, dI_t,
\]
where the last equality follows from the fact that \( \tilde{u} \) solves (3.3) and (5.10). From (5.3), we see that
\[
\psi_t = H^0_1 X_t^{\tilde{h}_t} \left\{ \frac{\gamma}{1 - \gamma} (\Sigma \Sigma^\top)^{-1} (a - r 1 + A \tilde{Y}_t) + (\Sigma \Sigma^\top)^{-1} \frac{D \tilde{u}(t, \tilde{Y}_t)}{\tilde{u}(t, \tilde{Y}_t)} \right\}.
\]

Using (5.5), we have
\[
\tilde{h}_t = \frac{1}{1 - \gamma} (\Sigma \Sigma^\top)^{-1} \left\{ a - r 1 + A \tilde{Y}_t + (1 - \gamma) \Sigma \Sigma^\top \lambda (\Pi_t)^\top \frac{D \tilde{u}(t, \tilde{Y}_t)}{\tilde{u}(t, \tilde{Y}_t)} \right\}, \quad (5.12)
\]
Finally, we have
\[
V_0(x, y) = E \left[ \int_0^T e^{-\rho t} U_1(\tilde{c}_t X_t^{\tilde{h}_t}, \tilde{c}_t) \, dt + e^{-\rho T} U_2(X_T^{\tilde{h}_t}) \right] \\
= E \left[ \int_0^T e^{-\rho t} U_1(\tilde{I}_1(\tilde{k}e^{\rho t} H^0_1)) \, dt + e^{-\rho T} U_2(I_2(\tilde{k}e^{\rho T} H^0_T)) \right].
\]

Note that
\[
U_i(I_i(y)) = \frac{1}{\gamma} (y^{1/(1 - \gamma)})^\gamma = \frac{1}{\gamma} y^{-\gamma/(1 - \gamma)}.
\]
Therefore, we have
\[
V_0(x, y) = \frac{1}{\gamma} E \left[ \int_0^T e^{-\rho t} (\tilde{k}e^{\rho t} H^0_1)^{-\gamma/(1 - \gamma)} \, dt + e^{-\rho T} (\tilde{k}e^{\rho T} H^0_T)^{-\gamma/(1 - \gamma)} \right] \\
= \frac{\tilde{k}^{-\gamma/(1 - \gamma)}}{\gamma} E \left[ \int_0^T e^{-\rho t/(1 - \gamma)} (H^0_1)^{-\gamma/(1 - \gamma)} \, dt + e^{-\rho T/(1 - \gamma)} (H^0_T)^{-\gamma/(1 - \gamma)} \right] \\
= \frac{\tilde{k}^{-\gamma/(1 - \gamma)}}{\gamma} \chi(\tilde{k}).
\]
\[\tilde{u}(0, y)^{1-\gamma}.\] (5.13)

Here, the third equality follows from (5.6) and the last equality follows from (5.7).

**Remark 5.1.** Recall that
\[\tilde{u}(t, y) = \exp \left[\eta(t, y) \frac{1}{1-\gamma}\right];\]
see (4.1). Then we see that \(\tilde{c}_t = \tilde{c}_t\) and \(\tilde{h}_t = \tilde{h}_t\) hold. Here \(\tilde{c}_t\) and \(\tilde{h}_t\) are given by (5.11) and (5.12), respectively. Also, \(\hat{c}_t\) and \(\hat{h}_t\) are given by (4.3). Moreover, we also see that \(V_0(x, y) = \tilde{V}(0, x, y)\) holds, where \(V_0(x, y)\) and \(\tilde{V}(0, x, y)\) are given by (5.13) and (4.4), respectively. Hence, we can check that the results obtained using the martingale method are in agreement with those obtained using the dynamic programming approach.

### 5.3. Infinite time horizon problem

As an extension for future research, we introduce a consumption problem on the infinite time horizon. That is, using the general utility function \(U_1\) given in Subsection 5.1, we consider the following:
\[V_\infty(x, y) := \sup_{(c, h) \in \tilde{A}} \mathbb{E} \left[ \int_0^\infty e^{-\rho t} U_1(c_t X_{c, h}^t) \, dt \right], \quad \rho > 0,\]
where \(\tilde{A}\) is the space of admissible strategies defined by:
\[\tilde{A} := \{(c, h) \in \mathcal{H}_T; \mathbb{E} \left[ \int_0^T e^{-\rho t} U_1^{-1}(c_t X_{c, h}^t) \, dt \right] < \infty \text{ for each } T > 0 \}.\]

Define
\[\chi_\infty(k) := \mathbb{E} \left[ \int_0^\infty H_t^0 I_1(ke^{\rho t} H_t^0) \, dt \right].\]
Then we have the following proposition.

**Proposition 5.3.** Assume that
\[\chi_\infty(k) < \infty.\] (5.14)

Define \(\tilde{k}\) such that \(\chi_\infty(\tilde{k}) = x.\) Then we have
\[V_\infty(x, y) := \mathbb{E} \left[ \int_0^\infty e^{-\rho t} U_1(\tilde{C}_t(\infty)) \, dt \right].\]

Here \(\tilde{C}_t(\infty)\) is defined by
\[\tilde{C}_t(\infty) := I_1(\tilde{k} e^{\rho t} H_t^0),\]
where \(I_1\) is given in Subsection 5.1. Also, \((\tilde{c}_t(\infty), \tilde{h}_t(\infty)) \in \tilde{A}_T\) is the optimal strategy. Here \(\tilde{c}_t(\infty)\) is defined by
\[\tilde{c}_t(\infty) := \frac{\tilde{C}_t(\infty)}{X_{c, h}^t(\infty), \tilde{k}(\infty)},\]
where \( \tilde{X}_t^{(\infty)}, \tilde{h}_t^{(\infty)} \) is the optimal wealth process defined by

\[
\tilde{X}_t^{(\infty)}, \tilde{h}_t^{(\infty)} := \frac{1}{H_0} \mathbb{E} \left[ \int_t^\infty H_u^0 \tilde{C}_u^{(\infty)} \, du \mid \xi_t \right].
\]  
(5.15)

Moreover, \( \tilde{h}_t^{(\infty)} \) is determined by the following:

\[
\mathbb{E} \left[ \int_0^\infty H_u^0 \tilde{C}_u^{(\infty)} \, du \mid \xi_t \right] = \int_0^t H_u^0 \tilde{X}_u^{(\infty)}, \tilde{h}_u^{(\infty)} \left( \tilde{h}_u^{(\infty)} - (\Sigma \Sigma^\top)^{-1} (a - r \mathbf{I} + A \hat{Y}_u) \right)^\top dI_u.
\]  
(5.16)

**Remark 5.2.** As in (5.2), we have

\[
H_0^t \tilde{X}_t^{(\infty)}, \tilde{h}_t^{(\infty)} + \int_0^t H_u^0 \tilde{C}_u^{(\infty)} \, du = x + \int_0^t H_u^0 \tilde{X}_u^{(\infty)}, \tilde{h}_u^{(\infty)} \left( \tilde{h}_u^{(\infty)} - (\Sigma \Sigma^\top)^{-1} (a - r \mathbf{I} + A \hat{Y}_u) \right)^\top dI_u.
\]  

From this and (5.15), we obtain (5.16).

In particular, we are interested in the case of the power utility:

\[
U_1(c) = \frac{1}{\gamma} c^\gamma \quad \text{for} \quad \gamma \in (-\infty, 0) \cup (0, 1).
\]

Note that we need to check condition (5.14) in order to use Proposition 5.3.

Then, in a similar way to Subsection 5.2, we have

\[
\chi_{\infty}(k) = k^{-1/(1-\gamma)} \int_0^\infty \exp \left[ -\frac{1}{1-\gamma} \left\{ \frac{1}{2} y^\top U(0; t) y + y^\top g(0; t) + l(0; t) - \rho t \right\} \right] \, dt,
\]  
(5.17)

where \( U, g, \) and \( l \) solve (3.4)–(3.6), respectively. We guess that we may need the conditions of the discount factor \( \rho \) and the HARA parameter \( \gamma \) satisfying (5.14). In future work, we aim to investigate the asymptotic behaviors of \( U, g, \) and \( l \). One possible direction is to follow the ideas of Nagai and Peng [44]; we can assume the stability of the matrix

\[
G := B - \Lambda \Sigma^\top (\Sigma \Sigma^\top)^{-1} A.
\]

Then we see that \( \Pi_t \to \hat{\Pi} \) as \( t \to \infty \), where \( \hat{\Pi} \) is a solution of the algebraic Riccati equation

\[
G \hat{\Pi} + \hat{\Pi} G^\top + \hat{\Pi} A^\top (\Sigma \Sigma^\top)^{-1} A \hat{\Pi} + \Lambda (I - \Sigma^\top (\Sigma \Sigma^\top)^{-1} \Sigma) \Lambda^\top = 0.
\]

In addition, the stability of the filter can be proved. In this case, when time is large, we may replace the equation for the factor process (2.3) by a time-homogeneous equation

\[
d\hat{Y}_t = (b + B \hat{Y}_t) \, dt + \lambda(\hat{\Pi}) \, dI_t.
\]

We expect that similar results to [26] and [43] may be obtained. From (5.17), we will find the conditions of \( \rho \) and \( \gamma \) such that \( \chi_{\infty}(k) \) is finite. This will be discussed further elsewhere.
Appendix A.

Proof of Lemma 2.1. We apply the approach of Proposition 1.2.3 of [2]. For $s \leq t$, we have

\[
\mathbb{E}[I_t - I_s \mid g_s] = \mathbb{E}
\left[
Z_t - Z_s - \int_s^t (d + A\hat{Y}_u) \, du \mid g_s
\right]
= \mathbb{E}
\left[
\int_s^t \Sigma \, dW_u + \int_s^t (d + AY_u) \, du - \int_s^t (d + A\hat{Y}_u) \, du \mid g_s
\right]
= \mathbb{E}
\left[
\int_s^t \Sigma \, dW_u \mid \mathcal{F}_s
\right] + \int_s^t A(Y_u - \mathbb{E}[Y_u \mid g_u]) \, du \mid g_s
= 0.
\]

Moreover, noting that

\[
d(I_t - I_s)(I_t - I_s)^\top = (I_t - I_s)d(I_t - I_s)^\top + (I_t - I_s)(I_t - I_s)^\top + \Sigma \Sigma^\top du,
\]
we have

\[
\mathbb{E}(I_t - I_s)(I_t - I_s)^\top \mid g_s = \Sigma \Sigma^\top (t - s).
\]

Appendix B.

Proof of Lemma 2.2. We apply the approach of Theorem 22.1.9 of [11]. Define

\[
M_t := \hat{Y}_t - \hat{Y}_0 - \int_0^t \mathbb{E}(b + BY_u \mid g_u) \, du = \hat{Y}_t - y - \int_0^t (b + B\hat{Y}_u) \, du. \tag{B.1}
\]

Then we see that

\[
\mathbb{E}[\hat{Y}_t - \hat{Y}_s \mid g_s] = \mathbb{E}[Y_t - Y_s \mid g_s]
= \mathbb{E}
\left[
\int_s^t (b + BY_u) \, du + \int_s^t \Lambda \, dW_u \mid g_s
\right]
= \mathbb{E}
\left[
\int_s^t (b + BE[Y_u \mid g_u]) \, du + \mathbb{E}
\left[
\int_s^t \Lambda \, dW_u \mid \mathcal{F}_s
\right] \mid g_s
\right]
= \mathbb{E}
\left[
\int_s^t (b + B\hat{Y}_u) \, du \mid g_s
\right],
\]
and that $M_t$ is a $\mathcal{G}_t$-martingale. Hence, there exists a $\mathcal{G}_t$-progressively process $\beta_u \in \mathbb{R}^{n \times (n + m)}$ such that

\[
M_t := \int_0^t \beta_u \, dv_u,
\]
where $v_t$ is defined by

\[
v_t := \Sigma^\top (\Sigma \Sigma^\top)^{-1} I_t.
\]

Then, by (B.1), we have

\[
d\hat{Y}_t = (b + B\hat{Y}_t) \, dt + \beta_t \, dv_t. \tag{B.2}
\]

Using (2.1) and (B.2), we have

\[
d(Y_t Z_t^\top) = dN_t + H_t \, dt.
\]
An optimal consumption and investment problem

where

\[ N_t := \int_0^t \{ Y_u (\Sigma dW_u) + (\Lambda dW_u)Z_u^T \}, \]

\[ H_t := \int_0^t \{ Y_u (d + AY_u) + (b + BY_u)Z_u^T + \Lambda \Sigma^T \} du. \]

Note that \( N_t \) is an \( \mathcal{F}_t \)-local martingale.

Using a similar argument as in Appendix A,

\[ \hat{Y}_t Z_t - \int_0^t \hat{H}_u du \]

is \( \mathcal{G}_t \)-martingale. That is, since \( Z_t \) is \( \mathcal{G}_t \)-measurable,

\[ \hat{Y}_t Z_t - \int_0^t \hat{H}_u du \]

is also \( \mathcal{G}_t \)-martingale. (B.3)

On the other hand, using (B.2) and recalling that

\[ dZ_t = (a + A\hat{Y}_t) dt + \Sigma d\nu_t, \]

we see that

\[ \hat{Y}_t Z_t - \int_0^t \{ \hat{Y}_u (d + A\hat{Y}_u) + (b + BY_u)Z_u^T + \beta_u \Sigma^T \} du \]

is also \( \mathcal{G}_t \)-martingale. Since the two decompositions (B.3) and (B.4) of the semimartingale \( \hat{Y}_t Z_t \) should be the same, we have

\[ \beta_t \Sigma^T := \hat{Y}_t \hat{Y}_t^T A^T - \hat{Y}_t \hat{Y}_t^T A^T + \Lambda \Sigma^T = \Pi A^T + \Lambda \Sigma^T, \]

where \( \Pi_t = \hat{Y}_t \hat{Y}_t^T - \hat{Y}_t \hat{Y}_t^T = \mathbb{E}[(Y_t - \hat{Y}_t)(Y_t - \hat{Y}_t)^T \mid \mathcal{G}_t] \). Hence, by (B.2) and (B.5), we obtain (2.3).

Next, we will obtain (2.5). Note that, we also have

\[ \Pi_t = \mathbb{E}[(Y_t - \hat{Y}_t)(Y_t - \hat{Y}_t)^T] \]

since \( Y \) and \( Z \) are Gaussian processes, hence, \( Y_t - \hat{Y}_t \) is independent of \( \mathcal{G}_t \). Using (1.1), (2.3), and (2.6), we have

\[ d(Y_t - \hat{Y}_t) = (B - \lambda (\Pi_t) A)(Y_t - \hat{Y}_t) dt + (\Lambda - \lambda (\Pi_t) \Sigma) dW_t. \]

Then we have

\[ d(Y_t - \hat{Y}_t)(Y_t - \hat{Y}_t)^T = (\Lambda - \lambda (\Pi_t) \Sigma)(\Lambda - \lambda (\Pi_t) \Sigma)^T dt \]

\[ + ((Y_t - \hat{Y}_t)(B - \lambda (\Pi_t) A)(Y_t - \hat{Y}_t) dt + (\Lambda - \lambda (\Pi_t) \Sigma) dW_t)^T \]

\[ + [(B - \lambda (\Pi_t) A)(Y_t - \hat{Y}_t) dt + (\Lambda - \lambda (\Pi_t) \Sigma) dW_t](Y_t - \hat{Y}_t)^T. \]

Integrating this on \([0, t], \) and taking the expectation for both sides, we have, by (B.6),

\[ \Pi_t = \int_0^t [(B - \lambda (\Pi_u) A)\Pi_u + \Pi_u (B - \lambda (\Pi_u) A)^T + (\Lambda - \lambda (\Pi_u) \Sigma)(\Lambda - \lambda (\Pi_u) \Sigma)^T] du \]

\[ = \int_0^t [-\lambda (\Pi_u) \Sigma^T \lambda (\Pi_u)^T + \Lambda \Lambda^T + B \Pi_u + \Pi_u B^T] du. \]

For the last relation, we use (2.4). Hence, we see that \( \Pi_t \) is the solution of (2.5).
Appendix C. A formal derivation of (3.1)

For the derivation of (3.1), we consider only \( \gamma \in (0, 1) \). The derivation for \( \gamma \in (-\infty, 0) \) is similar.

Let \( 0 < t < T \) and \( \delta > 0 \) such that \( t + \delta < T \). By the dynamic programming principle (see [20, Theorem III.7]), we have

\[
v(t, y) = \sup_{(c, h) \in A_{t, t+\delta}} \mathbb{E}_{t, y} \left[ \int_t^{t+\delta} c_s^\gamma \exp \left[ \int_t^s [\gamma (\ell(\hat{Y}_u, h_u) - c_u) - \rho] \, du \right] \, ds \\
+ \exp \left[ \int_t^{t+\delta} [\gamma (\ell(\hat{Y}_u, h_u) - c_u) - \rho] \, du \right] v(t + \delta, \hat{Y}_{t+\delta}) \right].
\]

(C.1)

Here \( \hat{Y}_t \) is governed by (2.11). Assume that \( v(\cdot, \cdot) \) is \( C^2 \). We can apply Itô’s formula:

\[
dv(s, \hat{Y}_s) = \left\{ \frac{\partial v}{\partial s}(s, \hat{Y}_s) + \frac{1}{2} \text{tr}(\lambda(\Pi_s) \Sigma \Sigma^\top \lambda(\Pi_s)^\top D^2v(s, \hat{Y}_s)) \right\} ds + (Dv(s, \hat{Y}_s))^\top \lambda(\Pi_s) dI_h^s.
\]

Then with \( \hat{Y}_t = y \), we have

\[
\exp \left[ \int_t^{t+\delta} [\gamma (\ell(\hat{Y}_u, h_u) - c_u) - \rho] \, du \right] v(t + \delta, \hat{Y}_{t+\delta})
= v(t, y) + \int_t^{t+\delta} \exp \left[ \int_t^s [\gamma (\ell(\hat{Y}_u, h_u) - c_u) - \rho] \, du \right] \left\{ \frac{\partial v}{\partial s}(s, \hat{Y}_s) + \frac{1}{2} \text{tr}(\lambda(\Pi_s) \Sigma \Sigma^\top \lambda(\Pi_s)^\top D^2v(s, \hat{Y}_s)) \right\} ds \\
+ \left[ \gamma (\ell(\hat{Y}_s, h_u) - c_u) - \rho \right] v(s, \hat{Y}_s) \right\} ds + \int_t^{t+\delta} \exp \left[ \int_t^s [\gamma (\ell(\hat{Y}_u, h_u) - c_u) - \rho] \, du \right] (Dv(s, \hat{Y}_s))^\top \lambda(\Pi_s) dI_h^s.
\]
Assuming some growth condition on the derivatives of $v$ and also $h$ such that

$$
\mathbb{E}_{t,y}^b \left[ \int_t^{t+\delta} \exp\left[ \int_t^s [y(\hat{\ell}(\hat{Y}_u, h_u) - c_u) - \rho] \, du \right] (Dv(s, \hat{Y}_s))^\top \gamma(\Pi_s) \, dt \right] = 0,
$$

then

$$
\mathbb{E}_{t,y}^b \left[ \int_t^{t+\delta} \exp\left[ \int_t^s [y(\hat{\ell}(\hat{Y}_u, h_u) - c_u) - \rho] \, du \right] v(t + \delta, \hat{Y}_{t+\delta}) \right]
\begin{align*}
&= v(t, y) + \mathbb{E}_{t,y}^b \left[ \int_t^{t+\delta} \exp\left[ \int_t^s [y(\hat{\ell}(\hat{Y}_u, h_u) - c_u) - \rho] \, du \right] \\
&\quad \times \left\{ \frac{\partial v}{\partial s}(s, \hat{Y}_s) + \frac{1}{2} \text{tr}(\lambda(\Pi_s) \Sigma \Sigma^\top \lambda(\Pi_s) \, D^2 v(s, \hat{Y}_s)) \\
&\quad + (b + B\hat{Y}_s + \gamma \lambda(\Pi_s) \Sigma \Sigma^\top h_s)^\top Dv(s, \hat{Y}_s) \\
&\quad + [y(\hat{\ell}(\hat{Y}_u, h_u) - c_u) - \rho]v(s, \hat{Y}_s) \right\} ds \right].
\end{align*}
$$

Hence, (C.1) becomes

$$
\sup_{(c,h) \in A_i} \mathbb{E}_{t,y}^b \left[ \int_t^{t+\delta} c^y \exp\left[ \int_t^s [y(\hat{\ell}(\hat{Y}_u, h_u) - c_u) - \rho] \, du \right] \\
+ \exp\left[ \int_t^s [y(\hat{\ell}(\hat{Y}_u, h_u) - c_u) - \rho] \, du \right] \\
\times \left\{ \frac{\partial v}{\partial s}(s, \hat{Y}_s) + \frac{1}{2} \text{tr}(\lambda(\Pi_s) \Sigma \Sigma^\top \lambda(\Pi_s) \, D^2 v(s, \hat{Y}_s)) \\
+ (b + B\hat{Y}_s + \gamma \lambda(\Pi_s) \Sigma \Sigma^\top h_s)^\top Dv(s, \hat{Y}_s) \\
+ [y(\hat{\ell}(\hat{Y}_u, h_u) - c_u) - \rho]v(s, \hat{Y}_s) \right\} ds \right] = 0.
$$

Dividing by $\delta$ and letting $\delta \to 0$, we can formally derive (3.1).

**Acknowledgements**

The authors would like to thank the anonymous referees for helpful comments and suggestions. The research work of Hiroaki Hata is supported by Grant-in-Aid for Young Scientists (B) number 15K17584, Japan Society for the Promotion of Science. The research work of Shuenn-Jyi Sheu is supported by MOST 105-2115-M-008-001 and NCTS, Taiwan.

**References**


An optimal consumption and investment problem