

Continuous time mean-variance portfolio optimization with piecewise state-dependent risk aversion

Xiangyu Cui¹ · Lu Xu² · Yan Zeng³

Received: 21 November 2014 / Accepted: 20 October 2015 / Published online: 3 November 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract In this paper, a continuous time mean-variance portfolio optimization problem is considered within a game theoretic framework, where the risk aversion function is assumed to depend on the current wealth level and the discounted (preset) investment target. We derive the explicit time consistent investment policy, and find that if the current wealth level is less (larger) than the discounted investment target, the future wealth level along the time consistent investment policy is always less (larger) than the discounted investment target.

Keywords Piecewise linear risk aversion · Continuous time mean-variance model · Time consistent policy

1 Introduction

Markowitz's mean-variance model (see [10]) initiated the famous return-risk assets allocation framework. An investor who considers a mean-variance criterion seeks the

✉ Yan Zeng
zengy36@mail.sysu.edu.cn

Xiangyu Cui
cui.xiangyu@mail.shufe.edu.cn

Lu Xu
stsdn@163.com

¹ School of Statistics and Management, Shanghai Key Laboratory of Financial Information Technology, Shanghai University of Finance and Economics, Shanghai, China

² School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, China

³ Lingnan (University) College, Sun Yat-sen University, Guangzhou, China

best investment policy such that the expected value of the terminal wealth (i.e., return) is maximized and the variance of the terminal wealth (i.e., risk) is minimized, which can be formulated as follows,

$$\begin{aligned}(MV(\gamma)) : & \min \quad \text{Var}(X_1|X_0) - \gamma\mathbb{E}[X_1|X_0], \\(MV(\omega)) : & \max \quad \mathbb{E}[X_1|X_0] - \omega\text{Var}(X_1|X_0),\end{aligned}$$

where X_0 is the initial wealth level, X_1 is the terminal wealth level, $\gamma \geq 0$ and $\omega \geq 0$ are the trade-off parameters between two conflict objectives. We call γ and ω risk aversion parameters, which represent the risk aversion attitude of the investor. When $\gamma = 0$ or $\omega = +\infty$, the investor is totally risk averse.

Besides constant risk aversion parameter, there are several state-dependent risk aversion parameters. Björk et al. [4] and Wu [14] proposed, respectively, in continuous time setting and multi-period setting, that risk aversion parameter ω takes a fractional form of current wealth level X_t ,

$$\omega(X_t) = \frac{\omega}{X_t}, \quad (\omega \geq 0).$$

Hu et al. [7] proposed the risk aversion parameter γ takes a linear function of current wealth level X_t ,

$$\gamma(X_t) = \mu_1 X_t + \mu_2, \quad (\mu_1 \geq 0),$$

in continuous time. Cui et al. [6] proposed an extended piecewise linear risk aversion parameter in a multi-period setting as follows,

$$\gamma(X_t) = \begin{cases} \gamma^+(X_t - \rho_t^{-1}W), & \text{if } X_t \geq \rho_t^{-1}W, \\ -\gamma^-(X_t - \rho_t^{-1}W), & \text{if } X_t < \rho_t^{-1}W, \end{cases} \quad (\gamma^+ \geq 0, \gamma^- \geq 0),$$

where ρ_t^{-1} is the riskless discount factor and W is the preset investment target. For Björk et al.'s setting, due to the positiveness of controlled wealth level X_t in continuous time, the investor is always risk averse and the higher the wealth level, the lower the risk aversion. In Wu's setting, when the wealth level is negative, $\omega(X_t)$ is negative, i.e., the investor is risk seeking and tries to maximize both the expected value and the variance of the terminal wealth. In Hu et al.'s setting, when the wealth level is less than $-\mu_2/\mu_1$, $\gamma(X_t)$ becomes negative, i.e., the investor tries to minimize the variance as well as the expected value of the terminal wealth. In Cui et al.'s setting, the investor is always risk averse and may have different views with respect to the difference between current wealth level and discounted preset investment target. In this paper, we solve the continuous time mean-variance portfolio optimization problem with Cui et al. [6]'s piecewise linear risk aversion parameter.

Consider the dynamic mean-variance portfolio optimization problem at time 0,

$$(MV_0(\gamma(X_0))) : \quad \min_{\{u_t\}_{t \in [0, T]}} \quad \text{Var}(X_T|X_0) - \gamma(X_0)\mathbb{E}[X_T|X_0],$$

where u_t is the investment policy at time t and X_T is the terminal wealth level. At time 0, although $\gamma(X_0)$ is a known constant, the problem is hard to solve, due to the non smoothing property of the variance term, i.e., $\text{Var}(X_T|X_0) \neq \text{Var}(\text{Var}(X_T|X_t)|X_0)$. Li and Ng [9] and Zhou and Li [15] adopted the embedding scheme to derive the optimal mean-variance policy under a multi-period setting and a continuous time setting, respectively. We call the optimal mean-variance policy determined at time 0 *pre-committed optimal mean-variance policy*. However, for $t > 0$, the investor faces a truncated mean-variance portfolio optimization problem,

$$(MV_t(\gamma(X_t))) : \min_{\{u_s\}_{s \in [t, T]}} \text{Var}(X_T|X_t) - \gamma(X_t)\mathbb{E}[X_T|X_t],$$

whose short term optimal mean-variance policy is different from the pre-committed optimal mean-variance policy derived at time 0 in general (see Basak and Chabakauri [2], Cui et al. [5], Wang and Forsyth [13]). This phenomenon is called time inconsistency. In the language of dynamic programming, *Bellman’s principle of optimality* is not applicable in this dynamical return-risk portfolio selection model, as the global and local objectives are not consistent (see Artzner et al. [1], Cui et al. [5]). In the fields of dynamic risk measures and dynamic risk management, time consistency is a basic requirement of dynamic risk measures (see Rosazza Gianin [11], Artzner et al. [1] and Jobert and Rogers [8]).

How to resolve this inconsistency? Basak and Chabakauri [2] extended Strotz [12]’s proposal of strategy of consistent planning and reformulated the dynamic mean-variance model as an interpersonal game model where the investor optimally chooses the policy at any time t , on the premise that he or she has already decided his or her time-consistent policies in the future. For example, in multi-period setting, the investor at time t faces the following nested portfolio selection problem

$$(NMV_t(\gamma(X_t))) : \min_{u_t} \text{Var}(X_T|X_t) - \gamma(X_t)\mathbb{E}[X_T|X_t],$$

$$\text{s.t. } u_j \text{ solves } (NMV_j(\gamma(X_j))), \quad t < j \leq T,$$

with terminal period problem given as

$$(NMV_{T-1}(\gamma(X_T))) : \min_{u_{T-1}} \text{Var}(X_T|X_{T-1}) - \gamma(X_T)\mathbb{E}[X_T|X_{T-1}].$$

The subgame Nash equilibrium solution of the nested problem is called time consistent mean-variance policy, which can be derived by backward induction. Basak and Chabakauri [2] assumed that the investor has a constant risk aversion during the investment procedure. Björk et al. [4], Hu et al. [7], Wu [14] and Cui et al. [6] extended Basak and Chabakauri’s results by studying different state-dependent risk aversion parameters. For general time inconsistent control problems, Björk and Murgoci [3] proposed the time inconsistent stochastic control framework to derive the time consistent control.

In this paper, we aim to adopt a time inconsistent stochastic control framework to derive a time consistent mean-variance policy (subgame Nash equilibrium policy) for

continuous time mean-variance portfolio selection problems with the defined piecewise linear risk aversion parameter. Totally different from multi-period setting, we are exposed to a Hamilton–Jacobi–Bellman (HJB) system of equations with nonsmooth coefficients, which may not admit classical solution. To overcome the difficulty, we first study some important properties of the time consistent mean-variance policy and then solve the HJB system of equations in two separate domains. By combining the solutions in two domains, we derive the time consistent mean-variance policy. In Sect. 2, we formulate the continuous time mean-variance model with piecewise linear risk aversion. In Sect. 3, we obtain the explicit time consistent mean-variance policy by solving the extended HJB system.

2 Portfolio optimization formulation

Our market setting is a standard Black–Scholes model, which includes a risky asset (such as a stock) and a riskless asset (such as a bank account). Denoting the stock price by S_t and the bank account by B_t , the dynamics of S_t and B_t are as follows,

$$\begin{cases} dS_t = \mu S_t dt + \sigma S_t dW_t, \\ S_0 = s_0, \\ dB_t = r B_t dt, \\ B_0 = b_0, \end{cases}$$

where $r > 0$ is the interest rate of bank account, μ is the appreciation rate of the stock, $\sigma > 0$ is the volatility or dispersion rate of the stock and W_t is a standard Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}_T, \{\mathcal{F}_t\}_{t \in [0, T]}, P)$. We assume that r, μ, σ are constants. In the analysis below, we will study self-financing portfolios (without consumption) consisting of the risky stock and the bank account. Denoting the dollar value invested in the risky asset at time t by u_t , the value of the portfolio at time t , X_t^u , is given by

$$\begin{cases} dX_t^u = [rX_t^u + (\mu - r)u_t]dt + \sigma u_t dW_t, \\ X_0^u = x_0, \end{cases} \quad (1)$$

where x_0 is the initial wealth level.

At time t , the investor faces a mean-variance portfolio selection problem,

$$\begin{aligned} (MV_t(\gamma(x))) : \quad & \min_{\{u_s\}_{s \in [t, T]}} \text{Var}_{t,x}(X_T^u) - \gamma(x)\mathbb{E}_{t,x}[X_T^u] \\ & \text{s.t.} \quad \begin{cases} dX_s^u = [rX_s^u + (\mu - r)u_s]dt + \sigma u_s dW_s, \\ X_t^u = x, \end{cases} \end{aligned}$$

where $\text{Var}_{t,x}(X_T^u) = \text{Var}(X_T^u | X_t^u = x)$, $\mathbb{E}_{t,x}[X_T^u] = \mathbb{E}[X_T^u | X_t^u = x]$. The risk aversion parameter of the investor is assumed to be

$$\gamma(x) = \begin{cases} \gamma^+(x - \rho_t^{-1}W), & \text{if } x \geq \rho_t^{-1}W, \\ -\gamma^-(x - \rho_t^{-1}W), & \text{if } x < \rho_t^{-1}W, \end{cases}$$

where W is the investment target set by the investor, $\rho_t^{-1} = e^{-r(T-t)}$ is the riskless discount factor from current time t to terminal time T , $\gamma^+ \neq 0$ and $\gamma^- \neq 0$ are the constant risk aversion coefficients. The different signs of γ^+ and γ^- respond different risk attitudes of the investor. In the case of $\gamma^+ > 0$ and $\gamma^- > 0$, if current wealth level is larger than the discounted investment target, the investor becomes less risk averse along with the increase of current wealth level; if current wealth level is less than the discounted investment target, the investor becomes less risk averse along with the decrease of current wealth level.

Setting $Y_t^u = X_t^u - \rho_t^{-1}W$, we have

$$\begin{aligned} dY_t^u &= dX_t^u - r\rho_t^{-1}Wdt \\ &= [r(X_t^u - \rho_t^{-1}W) + (\mu - r)u_t]dt + \sigma u_t dW_t \\ &= [rY_t^u + (\mu - r)u_t]dt + \sigma u_t dW_t. \end{aligned}$$

Furthermore,

$$\begin{aligned} \text{Var}_{t,x}(X_T^u) &= \text{Var}_{t,y}(Y_T^u), \\ \mathbb{E}_{t,x}[X_T^u] &= \mathbb{E}_{t,y}[Y_T^u] + \rho_t^{-1}W, \end{aligned}$$

with $y = x - \rho_t^{-1}W$. Thus, problem $(MV_t(\gamma(x)))$ is equivalent to

$$\begin{aligned} (MV_t(\gamma(y))) : \quad & \min_{\{u_s\}_{s \in [t,T]}} \text{Var}_{t,y}(Y_T^u) - \gamma(y)\mathbb{E}_{t,y}[Y_T^u] \\ \text{s.t.} \quad & \begin{cases} dY_s^u = [rY_s^u + (\mu - r)u_s]dt + \sigma u_s dW_s, \\ Y_t^u = y, \end{cases} \end{aligned} \tag{2}$$

where

$$\gamma(y) = \begin{cases} \gamma^+y, & \text{if } y \geq 0, \\ -\gamma^-y, & \text{if } y < 0. \end{cases}$$

3 Time consistent mean-variance policy

We define the objective function as

$$J(t, y, u) = \text{Var}_{t,y}(Y_T^u) - \gamma(y)\mathbb{E}_{t,y}[Y_T^u] = E_{t,y}[F(y, Y_T^u)] + G(y, \mathbb{E}_{t,y}[Y_T^u]), \tag{3}$$

where the two new functions $F(y, z) = z^2 - \gamma(y)z$ and $G(y, z) = -z^2$. We try to derive the time consistent investment policy (i.e., the subgame Nash equilibrium policy) of $J(t, y, u)$, \hat{u} , which is rigorously defined in [4] and restated in the following definition.

Definition 1 (Björk et al. [4]) Given a policy \hat{u} , construct a policy u_h by

$$u_h(s, y) = \begin{cases} u, & \text{for } t \leq s < t + h, \\ \hat{u}(s, y), & \text{for } t + h \leq s \leq T, \end{cases}$$

where $u \in \mathbb{R}$, $h > 0$ and $(t, y) \in [0, T] \times \mathbb{R}$ are arbitrarily chosen. If

$$\limsup_{h \rightarrow 0} \frac{J(t, y, \hat{u}) - J(t, y, u_h)}{h} \leq 0,$$

for all $u \in \mathbb{R}$ and $(t, y) \in [0, T] \times \mathbb{R}$, we say that \hat{u} is an equilibrium policy. And the equilibrium value function V is defined by

$$V(t, y) = J(t, y, \hat{u}).$$

Following the time inconsistent stochastic control approach proposed in Björk and Murgoci [3] and Björk et al. [4], the extended HJB system of equations for the subgame Nash equilibrium problem takes the following form:

$$\begin{aligned} & \inf_{u \in \mathbb{R}} \{(\mathbf{A}^u V)(t, y) - (\mathbf{A}^u f)(t, y, y) + (\mathbf{A}^u f^y)(t, y) \\ & - \mathbf{A}^u(G \diamond g)(t, y) + (\mathbf{H}^u g)(t, y)\} = 0, \quad 0 \leq t \leq T, \\ & \mathbf{A}^{\hat{u}} f^z(t, y) = 0, \quad 0 \leq t \leq T, \\ & \mathbf{A}^{\hat{u}} g(t, y) = 0, \quad 0 \leq t \leq T, \\ & V(T, y) = F(y, y) + G(y, y), \\ & f(T, y, z) = F(z, y), \\ & g(T, y) = y, \end{aligned}$$

where the infinitesimal operator \mathbf{A}^u and notations f^x , $G \diamond g$, $\mathbf{H}^u g$ are defined by

$$\begin{aligned} \mathbf{A}^u &= \frac{\partial}{\partial t} + [ry + (\mu - r)u] \frac{\partial}{\partial y} + \frac{1}{2} \sigma^2 u^2 \frac{\partial^2}{\partial y^2}, \\ f^z(t, y) &= f(t, y, z), \\ (G \diamond g)(t, y) &= G(y, g(t, y)), \\ \mathbf{H}^u g(t, y) &= \frac{\partial G}{\partial z}(y, z) \cdot \mathbf{A}^u g(t, y). \end{aligned}$$

Given $F(y, z) = z^2 - \gamma(y)z$ and $G(y, z) = -z^2$, the HJB system of equations can be reduce into the following,

$$\begin{aligned} V_t(t, y) + \inf_{u \in \mathbb{R}} & \left\{ [ry + (\mu - r)u](V_y(t, y) - f_z(t, y, y)) \right. \\ & \left. + \frac{1}{2} \sigma^2 u^2 (V_{yy}(t, y) - f_{zz}(t, y, y)) \right\} \end{aligned}$$

$$-2f_{yz}(t, y, y) + 2(g_y(t, y))^2 \Big\} = 0, \quad 0 \leq t \leq T, \tag{4}$$

$$f_t(t, y, z) + [ry + (\mu - r)\hat{u}]f_y(t, y, z) + \frac{1}{2}\sigma^2\hat{u}^2 f_{yy}(t, y, z) = 0, \quad 0 \leq t \leq T, \tag{5}$$

$$g_t(t, y) + [ry + (\mu - r)\hat{u}]g_y(t, y) + \frac{1}{2}\sigma^2\hat{u}^2 g_{yy}(t, y) = 0, \quad 0 \leq t \leq T, \tag{6}$$

$$V(T, y) = -\gamma(y)y, \tag{7}$$

$$f(T, y, z) = y^2 - \gamma(z)y, \tag{8}$$

$$g(T, y) = y, \tag{9}$$

where $V_t, V_y, V_{yy}, f_z, f_{zz}, f_{yz}, g_y$ are corresponding partial derivatives of $V(t, y)$ and $f(t, y, z)$.

However, as the risk version parameter $\gamma(y)$ is a nonsmooth function, there does not exist classical solution of the HJB system (4)–(9). To overcome this difficulty, we first investigate the properties of time consistent mean-variance policy.

Proposition 1 *The time consistent mean-variance policy has the following properties:*

1. Whenever the state $Y_t^u = 0$, time consistent mean-variance policy over $[t, T]$ is

$$\hat{u}(s, Y_s^u = 0) = 0, \quad \text{for } t \leq s \leq T.$$

2. When $Y_0^u \leq 0$ (or ≥ 0), the state at time t along time consistent mean-variance policy is always nonpositive (or nonnegative), i.e., $Y_t^u \leq 0$ (or ≥ 0).

Proof When the state $Y_s^u = 0$ at future time s ($t \leq s \leq T$), the investor becomes totally risk averse. The objective function of the investor at time s can attain its minimum value 0 by investing all the wealth in the riskless asset for time interval $[s, T]$, which implies that the optimal decisions of the investor at all future time instances are consistent. On the other hand, along the proposed time consistent mean-variance policy $\hat{u}(s, Y_s^u) = 0$ (i.e., investing all the wealth in the riskless asset), the state $Y_s^u = 0$ for $t \leq s \leq T$. Therefore, the investor at all future time instances will become totally risk averse and would like to insist on investing all the wealth in the riskless asset, which implies that the subgame Nash equilibrium policy is just to invest all the wealth in the riskless asset.

For the second property, as the state process $\{Y_t^u\}_{t \in [0, T]}$ has continuous paths, Y_t^u should touch zero before changing its sign, and remain to be zero due to the first property. □

With the help of Proposition 1, we can solve the HJB system (4)–(9) in two domains, $\{(t, y) | y \geq 0\}$ and $\{(t, y) | y \leq 0\}$. In each domain, the risk aversion parameter $\gamma(\cdot)$ becomes a smooth function now. Our main result is given in the following Theorem.

Theorem 1 *Under piecewise linear state-dependent risk aversion framework, the time consistent mean-variance policy is given by*

$$\hat{u}(t, y) = \begin{cases} -k^+(t)y, & \text{if } y \geq 0, \\ -k^-(t)y, & \text{if } y \leq 0, \end{cases} \tag{10}$$

where

$$k^+(t) = \frac{\mu - r}{\sigma^2} \cdot \frac{2c^+(t) - \gamma^+b^+(t) - 2(b^+(t))^2}{2c^+(t)}, \tag{11}$$

$$k^-(t) = \frac{\mu - r}{\sigma^2} \cdot \frac{2c^-(t) + \gamma^-b^-(t) - 2(b^-(t))^2}{2c^-(t)}. \tag{12}$$

The parameters $b^+(t)$, $c^+(t)$, $b^-(t)$, $c^-(t)$ solve the following system of ordinary differential equations,

$$\begin{aligned} \dot{c}^+(t) + 2[r - (\mu - r)k^+(t)]c^+(t) + \sigma^2(k^+(t))^2c^+(t) &= 0, \\ c^+(T) &= 1, \\ \dot{b}^+(t) + [r - (\mu - r)k^+(t)]b^+(t) &= 0, \\ b^+(T) &= 1, \\ \dot{c}^-(t) + 2[r - (\mu - r)k^-(t)]c^-(t) + \sigma^2(k^-(t))^2c^-(t) &= 0, \\ c^-(T) &= 1, \\ \dot{b}^-(t) + [r - (\mu - r)k^-(t)]b^-(t) &= 0, \\ b^-(T) &= 1, \end{aligned}$$

where $\dot{c}^+(t)$, $\dot{b}^+(t)$, $\dot{c}^-(t)$ and $\dot{b}^-(t)$ are the first order derivatives with respect to time t .

Proof We have the probabilistic interpretations of $f(t, y, z)$, $g(t, y)$ and $V(t, y)$ as follows,

$$\begin{aligned} f(t, y, z) &= \mathbb{E}_{t,y}[F(y, Y_T^{\hat{u}})] = \mathbb{E}_{t,y}[(Y_T^{\hat{u}})^2] - \gamma(z)\mathbb{E}_{t,y}[Y_T^{\hat{u}}], \\ g(t, y) &= \mathbb{E}_{t,y}[Y_T^{\hat{u}}], \\ V(t, y) &= \mathbb{E}_{t,y}[F(y, Y_T^{\hat{u}})] + G(y, g(t, y)) = f(t, y, y) - g^2(t, y). \end{aligned}$$

Then, HJB equation (4) can be reduced into

$$\begin{aligned} f_t(t, y, y) - 2g(t, y)g_t(t, y) + \inf_{u \in \mathbb{R}} \left\{ [ry + (\mu - r)u](f_y(t, y, y) - 2g(t, y)g_y(t, y)) \right. \\ \left. + \frac{1}{2}\sigma^2u^2(f_{yy}(t, y, y) - 2g(t, y)g_{yy}(t, y)) \right\} = 0, \end{aligned}$$

which implies

$$\hat{u}(t, y) = -\frac{\mu - r}{\sigma^2} \frac{f_y(t, y, y) - 2g(t, y)g_y(t, y)}{f_{yy}(t, y, y) - 2g(t, y)g_{yy}(t, y)}.$$

For $\{(t, y) | y \geq 0\}$, due to the second property in Proposition 1, $z = Y_T^{\hat{u}}$ has the same sign as y . We prove that the solution set of HJB system is

$$V(t, y) = [c^+(t) - \gamma^+b^+(t) - (b^+(t))^2]y^2,$$

$$f(t, y, z) = c^+(t)y^2 - \gamma^+b^+(t)yz,$$

$$g(t, y) = b^+(t)y.$$

It is easy to check that setting $c^+(T) = b^+(T) = 1$, $V(T, y)$, $f(T, y, z)$ and $g(T, y)$ satisfies terminal conditions (7)–(9). Furthermore,

$$g_t(t, y) = \dot{b}^+(t)y, \quad g_y(t, y) = b^+(t), \quad g_{yy}(t, y) = 0,$$

$$f_t(t, y, z) = \dot{c}^+(t)y^2 - \gamma^+\dot{b}^+(t)yz, \quad f_y(t, y, z) = 2c^+(t)y - \gamma^+b^+(t)z,$$

$$f_{yy}(t, y, z) = 2c^+(t).$$

The time consistent policy is given by

$$\hat{u}(t, y) = -\frac{\mu - r}{\sigma^2} \frac{2c^+(t) - \gamma^+b^+(t) - 2(b^+(t))^2}{2c^+(t)} y = -k^+(t)y.$$

Then, the HJB equations (4)–(6) become,

$$\dot{c}^+y^2 - \gamma^+\dot{b}^+y^2 - 2b^+\dot{b}^+y^2 + [ry + (\mu - r)\hat{u}][2c^+y - \gamma^+b^+y - 2(b^+)^2y] + \sigma^2\hat{u}^2c^+ = 0,$$

$$\dot{c}^+y^2 - \gamma^+\dot{b}^+yz + [ry + (\mu - r)\hat{u}][2c^+y - \gamma^+b^+z] + \sigma^2\hat{u}^2c^+ = 0,$$

$$\dot{b}^+y + [ry + (\mu - r)\hat{u}]b^+ = 0,$$

which implies

$$\dot{c}^+ + 2[r - (\mu - r)k^+]c^+ + \sigma^2(k^+)^2c^+ = 0,$$

$$\dot{b}^+ + [r - (\mu - r)k^+]b^+ = 0.$$

Here we omit the arguments of functions $c^+(t)$, $b^+(t)$ and $k^+(t)$.

For $\{(t, y) \mid y \leq 0\}$, due to the second property in Proposition 1, $z = Y_T^{\hat{u}}$ has the same sign as y . We prove that the solution set of HJB system is

$$V(t, y) = [c^-(t) + \gamma^-b^-(t) - (b^-(t))^2]y^2,$$

$$f(t, y, z) = c^-(t)y^2 + \gamma^-b^-(t)yz,$$

$$g(t, y) = b^-(t)y.$$

It is easy to check that setting $c^-(T) = b^-(T) = 1$, $V(T, y)$, $f(T, y, z)$ and $g(T, y)$ satisfies terminal conditions (7)–(9). Furthermore,

$$g_t(t, y) = \dot{b}^-(t)y, \quad g_y(t, y) = b^-(t), \quad g_{yy}(t, y) = 0,$$

$$f_y(t, y, z) = 2c^-(t)y + \gamma^-b^-(t)z, \quad f_t(t, y, z) = \dot{c}^-(t)y^2 + \gamma^-\dot{b}^-(t)yz,$$

$$f_{yy}(t, y, z) = 2c^-(t).$$

The time-consistent policy is given by

$$\hat{u}(t, y) = -\frac{\mu - r}{\sigma^2} \frac{2c^-(t) + \gamma^- b^-(t) - 2(b^-(t))^2}{2c^-(t)} y = -k^-(t)y.$$

Then, the HJB equations (4) - (6) become,

$$\begin{aligned} \dot{c}^- y^2 + \gamma^- \dot{b}^- y^2 - 2b^- \dot{b}^- y^2 + [ry + (\mu - r)\hat{u}][2c^- y \\ + \gamma^- b^- y - 2(b^-)^2 y] + \sigma^2 \hat{u}^2 c^- = 0, \\ \dot{c}^- y^2 + \gamma^- \dot{b}^- yz + [ry + (\mu - r)\hat{u}][2c^- y + \gamma^- b^- z] + \sigma^2 \hat{u}^2 c^- = 0, \\ \dot{b}^- y + [ry + (\mu - r)\hat{u}]b^- = 0, \end{aligned}$$

which implies

$$\begin{aligned} \dot{c}^- + 2[r - (\mu - r)k^-]c^- + \sigma^2 (k^-)^2 c^- = 0, \\ \dot{b}^- + [r - (\mu - r)k^-]b^- = 0. \end{aligned}$$

Here we also omit the arguments of functions $c^-(t)$, $b^-(t)$ and $k^-(t)$. \square

Remark 1 Theorem 1 has shown that in both domains $\{(t, y)|y \geq 0\}$ and $\{(t, y)|y \leq 0\}$, the time consistent mean-variance policy has the same form as the one in Björk et al. [4]'s paper. It is an extension of [4]'s result.

Remark 2 When substituting the time consistent mean-variance policy \hat{u} back to the dynamics of state Y_t^u in (2), we can see that Y_t^u is a geometric Brownian motion, which implies $Y_s^u > 0$ (or < 0) for $s \in (t, T]$, if and only if $Y_t^u = y > 0$ (or < 0). It means that if the current wealth level is less than (or larger than) the discounted investment target, the future wealth level along the time consistent mean-variance policy is always less than (or larger than) the discounted investment target.

Acknowledgments This work is partially supported by grants of the National Natural Science Foundation of China (Nos. 71201094, 71201173, 71571195), Innovation Project Research Fund of SHUFE (No. CXJJ-2014-463), Program for Changjiang Scholars and Innovative Research Team in SUFE (No. IRT13077), Guangdong Natural Science for Research Team (2014A030312003), Guangdong Natural Science Funds for Distinguished Young Scholar (No. 2015A030306040) and Natural Science Foundation of Guangdong Province of China (No.S2013010011959).

References

1. Artzner, P., Delbaen, F., Eber, J.M., Heath, D., Ku, H.: Coherent multiperiod risk adjusted values and Bellman's principle. *Ann. Oper. Res.* **152**, 5–22 (2007)
2. Basak, S., Chabakauri, G.: Dynamic mean-variance asset allocation. *Rev. Fin. Stud.* **23**, 2970–3016 (2010)
3. Björk, T., Murgoci, A.: A general theory of markovian time inconsistent stochastic control problems. Working paper, Stockholm School of Economics and Copenhagen Business School (2010)
4. Björk, T., Murgoci, A., Zhou, X.Y.: Mean-variance portfolio optimization with state-dependent risk aversion. *Math. Fin.* **24**, 1–24 (2014)
5. Cui, X.Y., Li, D., Wang, S.Y., Zhu, S.S.: Better than dynamic mean-variance: time inconsistency and free cash flow stream. *Math. Fin.* **22**, 346–378 (2012)

6. Cui, X.Y., Li, D., Li, X., Shi, Y.: Time-consistent behaviour portfolio policy for mean-variance formulation, available at SSRN (2014). <http://ssrn.com/abstract=2480299>
7. Hu, Y., Jin, H.Q., Zhou, X.Y.: Time-inconsistent stochastic linear-quadratic control. *SIAM J. Control Optim.* **50**, 1548–1572 (2012)
8. Jobert, A., Rogers, L.C.: Valuations and dynamic convex risk measures. *Math. Fin.* **18**, 1–22 (2008)
9. Li, D., Ng, W.L.: Optimal dynamic portfolio selection: multi-period mean-variance formulation. *Math. Fin.* **10**, 387–406 (2000)
10. Markowitz, H.M.: Portfolio selection. *J. Fin.* **7**, 77–91 (1952)
11. Rosazza Gianin, E.: Risk measures via g -expectations. *Insur. Math. Econ.* **39**, 19–34 (2006)
12. Strotz, R.: Myopia, inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* **23**, 165–180 (1955)
13. Wang, J., Forsyth, P.A.: Continuous time mean variance asset allocation, a time-consistent strategy. *Eur. J. Oper. Res.* **209**, 184–201 (2011)
14. Wu, H.L.: Time-consistent strategies for a multiperiod mean-variance portfolio selection problem. *J. Appl. Math.* **2013**, 13 (2013)
15. Zhou, X.Y., Li, D.: Continuous-time mean-variance portfolio selection: a stochastic LQ framework. *Appl. Math. Optim.* **42**, 19–33 (2000)