Mutual Fund Theorem for continuous time markets with random coefficients*

Nikolai Dokuchaev

Department of Mathematics & Statistics, Curtin University,

GPO Box U1987, Perth, 6845 Western Australia

September 15, 2013

Abstract

The optimal investment problem is studied for a continuous time incomplete market model. It is assumed that the risk-free rate, the appreciation rates and the volatility of the stocks are all random; they are independent from the driving Brownian motion, and they are currently observable. It is shown that some weakened version of Mutual Fund Theorem holds for this market for general class of utilities. It is shown that the supremum of expected utilities can be achieved on a sequence of strategies with a certain distribution of risky assets that does not depend on risk preferences described by different utilities.

Key words: optimal portfolio, Mutual Fund Theorem, continuous time market models.

Mathematical Subject Classification (2010): 91G10

1 Introduction

We study an optimal portfolio selection problem in a market model which consists of a risk–free bond or bank account and a finite number of risky stocks. The evolution of stock prices is described by Ito stochastic differential equations with the vector of the appreciation rates $a(t)$ and the volatility matrix $\sigma(t)$, while the bond price is exponentially increasing with a random risk free rate $r(t)$. A typical optimal portfolio selection problem is to find an investment strategy that maximizes $\mathbf{E}U(\widetilde{X}(T))$, where $\mathbf{E}$ denotes the mathematical expectation, $U(\cdot)$ is an utility function, $X(T)$ represents the wealth at final time T, and $\widetilde{X}(T) = \exp\left(-\int_0^T r(s)ds\right)X(T)$ is the discounted wealth. There are many works devoted to different modifications of this problem (see, e.g., Merton (1969) and review in Karatzas and Shreve (1998)).

*In press in "Theory and Decision". Web-published version: DOI 10.1007/s11238-013-9368-1

Dynamic portfolio selection problems are usually studied in the framework of stochastic control. To suggest a strategy, one needs to forecast future market scenarios (or the probability distributions, or the future distributions of $r(t)$, $a(t)$ and $\sigma(t)$). Unfortunately, the nature of financial markets is such that the choice of a hypothesis about the future distributions is always difficult to justify. In fact, it is still an open question if there is any useful information in the past prices that helps to predict the future. Respectively, there are serious reservations toward usual tools of stochastic control such as Dynamic Programming or Stochastic Maximum Principle that require knowledge of future values of the process $\mu(t) = (r(t), a(t), \sigma(t))$. It is why some special methods were developed for the financial models to deal with limited predictability.

One of this tools is the so-called Mutual Fund Theorem that says that the distribution of the risky assets in the optimal portfolio does not depend on forecast of future values of μ and on the investor's risk preferences (or utility function). This means that all rational investors may achieve optimality using the same mutual fund plus a saving account, and this mutual fund does not need to use the market forecast. Clearly, calculation of the optimal portfolio is easier in this case.

If Mutual Fund Theorem holds, then, for a typical model, portfolio stays on the *efficient frontier* even if there are errors in the forecast, i.e., it is optimal for some other risk preferences. This reduces the impact of forecast errors. This is another reason why it is important to know when Mutual Fund Theorem holds.

Mutual Fund Theorem was established first for the single period mean variance portfolio selection problem, i.e., for the problem with quadratic criterions. This result was a cornerstone of the modern portfolio theory. In particular, the Capital Assets Pricing Model (CAPM) is based on it. For the multi-period discrete time setting, some versions of Mutual Fund Theorem were obtained so far for problems with quadratic criterions only (Li and Ng (2010), Dokuchaev (2010a)). For the continuous time setting, Mutual Fund Theorem was obtained for portfolio selection problems with quadratic criterions as well as for more general utilities. In particular, Merton's optimal strategies for $U(x) = \delta^{-1}x^{\delta}$ and $U(x) = \log(x)$ are such that Mutual Fund Theorem holds for the case of random coefficients independent from the driving Brownian motion (Karatzas and Shreve (1998)). It is also known that Mutual Fund Theorem does not hold for power utilities in the presence of correlations; see, e.g., Brennan (1998), Feldman (2007). Khanna and Kulldorff (1999) proved that Mutual Fund Theorem theorem holds for a general utility function $U(x)$ for the case of non-random coefficient, and for a setting with consumption. Lim (2004,2005) and Lim and Zhou (2002) found some cases when Mutual Fund Theorem holds for problems with quadratic criterions. Dokuchaev and Haussmann (2001) found that Mutual Fund Theorem holds if the scalar value $\int_0^T |\theta(t)|^2 dt$ is non-random, where $\theta(t)$ is the market price of risk process. Schachermayer *et al* (2009) found sufficient conditions for Mutual Fund Theorem expressed via replicability of the European type claims $F(Z(T))$, where $F(\cdot)$ is a deterministic function and $Z(t)$ is the discounted wealth generated by the log-optimal discounted wealth process. The required replicability has to be achieved by trading of the log-optimal mutual fund with the discounted wealth $Z(t)$.

It can be summarized that Mutual Fund Theorem was established so far for the following continuous time optimal portfolio selection problems:

(i) For $U(x) \equiv \log(x)$ and general random coefficients (r, a, σ);

(ii) For $U(x) = \delta^{-1}x^{\delta}$, $\delta \neq 0$ and random coefficients (r, a, σ) being independent from the driving Brownian motions;

(iii) For problems with quadratic criterions;

(iv) For general utility and non-random coefficients (r, a, σ);

(v) For general utility when the integral $\int_0^T |\theta(t)|^2 dt$ is non-random;

(vi) For general utility when the claims $F(Z(T))$ can be replicated via trading of a mutual fund with the discounted wealth $Z(t)$, for deterministic functions F.

It can be noted that conditions (iv) implies (v), and (v) implies (vi).

Extension of Mutual Fund Theorem on problems (i)-(vi) was not trivial; it required significant efforts and variety of mathematical methods.

In this paper, we present one more case when Mutual Fund Theorem holds. More precisely, we found that it holds for general utility when the parameters $r(t)$, $a(t)$ and $\sigma(t)$ are all random, they are independent from the driving Brownian motion, and they are currently observable. It is an incomplete market; it is a case of "totally unhedgeable" coefficients, according to terms from Karatzas and Shreve (1998), Chapter 6. In fact, we found that only a weakened version of Mutual Fund Theorem holds: the supremum of expected utilities can be achieved on a sequence of strategies with a certain distribution of risky assets that does not depend on utility.

2 Definitions

We are given a standard probability space $(\Omega, \mathcal{F}, \mathbf{P})$, where $\Omega = \{\omega\}$ is a set of elementary events, $\mathcal{F}$ is a complete σ-algebra of events, and $\mathbf{P}$ is a probability measure that describes a prior probability distributions.

Market model

We consider a market model in a generalized Black-Scholes framework. We assume that the market consists of a risk free asset or bank account with price $B(t)$, $t \geq 0$, and n risky stocks with prices $S_i(t)$, $t \geq 0$, $i = 1, 2, \ldots, n$, where $n < +\infty$ is given.

We assume that

$$B(t) = B(0) \exp\Big(\int_0^t r(s)ds\Big), \tag{2.1}$$

where $r(t)$ is the random process of the risk-free interest rate (or the short rate). We assume that $B(0) = 1$. The process $B(t)$ will be used as numeraire.

The prices of the stocks evolve according to

$$dS_i(t) = S_i(t)\Big(a_i(t)dt + \sum_{j=1}^{n} \sigma_{ij}(t)dw_j(t)\Big), \quad t > 0, \tag{2.2}$$

where $w(\cdot) = (w_1(\cdot), \ldots, w_n(\cdot))$ is a standard Wiener process with independent components, $a_i(t)$ are the appreciation rates, and $\sigma_{ij}(t)$ are the volatility coefficients. The initial price $S_i(0) > 0$ is a given non-random constant.

We assume that $r(t)$, $a(t) \triangleq \{a_i(t)\}_{i=1}^n$, and $\sigma(t) \triangleq \{\sigma_{ij}(t)\}_{i,j=1}^n$ are currently observable uniformly bounded, measurable random processes In addition, we assume that the inverse matrix $\sigma(t)^{-1}$ is defined and bounded and $r(t) \geq 0$.

Let $\mathcal{F}_t$ be the filtration generated by all observable data. More precisely, it is the minimal filtration such that $(S(t), r(t), a(t), \sigma(t))$ is adapted to $\mathcal{F}_t$, where $S(t) \triangleq (S_1(t), \ldots, S_n(t))^\top$.

Set $\mu(t) \triangleq (r(t), \widetilde{a}(t), \sigma(t))$, where $\widetilde{a}(t) \triangleq a(t) - r(t)\mathbf{1}$ and $\mathbf{1} \triangleq (1, 1, \ldots, 1)^\top \in \mathbf{R}^n$. The process μ represents the vector of current market parameters.

We assume that the process $\mu(t)$ is independent from $w(\cdot)$.

Let

$$\widetilde{S}(t) = (\widetilde{S}_1(t), \ldots, \widetilde{S}_n(t))^\top \triangleq \exp\left(-\int_0^t r(s)ds\right) S(t).$$

Wealth and strategies

Let $X_0 > 0$ be the initial wealth at time $t_0 \in [0,T)$, and let $X(t)$ be the wealth at time $t > t_0$, $X(t_0) = X_0$. Let the process $\pi_0(t)$ represents the proportion of the wealth invested in the bond, $\pi_i(t)$ is the proportion of the wealth invested in the ith stock. In other words, the process $\pi_0(t)X(t)$ represents the dollar amount of the wealth invested in the bond, $\pi_i(t)X(t)$ is the dollar amount of the wealth invested in the ith stock, $\pi(t) = (\pi_1(t), \dots, \pi_n(t))^\top$, $t \geq 0$. We assume that

$$\pi_0(t) + \sum_{i=1}^{n} \pi_i(t) = 1. \tag{2.3}$$

The case of negative π_i is not excluded.

The process $\widetilde{X}(t) \triangleq \exp\left(-\int_0^t r(s)ds\right) X(t)$ is called the discounted wealth.

Let $\mathbf{S}(t) \triangleq \operatorname{diag}(S_1(t), \dots, S_n(t))$ and $\widetilde{\mathbf{S}}(t) \triangleq \operatorname{diag}(\widetilde{S}_1(t), \dots, \widetilde{S}_n(t))$ be the diagonal matrices with the corresponding diagonal elements.

The portfolio is said to be self-financing, if

$$dX(t) = X(t)(\pi(t)^\top \mathbf{S}(t)^{-1} dS(t) + \pi_0(t)B(t)^{-1}dB(t)). \tag{2.4}$$

It follows that for such portfolios

$$d\widetilde{X}(t) = \widetilde{X}(t)\pi(t)^\top \widetilde{\mathbf{S}}(t)^{-1} d\widetilde{S}(t), \tag{2.5}$$

so π alone suffices to specify the portfolio for $t > t_0$, given some t_0 and $X(t_0)$. We denote the corresponding wealth by $X(t, t_0, X_0, \pi)$.

Let

$$\theta(t) \triangleq \sigma(t)^{-1}\widetilde{a}(t) \tag{2.6}$$

be the *risk premium process.*

Let $\Sigma(t_1, t_2)$ be the class of all $\mathcal{F}_t$-adapted processes $\pi(\cdot) = (\pi_1(\cdot), \dots, \pi_n(\cdot)) : [t_1, t_2] \times \Omega \to \mathbf{R}^n$ such that $\sup_{t,\omega} |\pi(t,\omega)| < +\infty$ and that if $\theta(t) = 0$ then $\pi(t) = 0$.

We shall consider classes $\Sigma(t_1, t_2)$ as classes of admissible strategies. For these strategies, $X(t) > 0$ for all t a.s.

Let $\Sigma_{MFT}(t_1, t_2)$ be the set of all strategies $\pi \in \Sigma(t_1, t_2)$ such that

$$\pi(t)^\top = \nu(t)\theta(t)^\top \sigma(t)^{-1},$$

where $\nu(t)$ is an one dimensional process adapted to $\mathcal{F}_t$.

For a given strategy $\pi \in \Sigma(0,T)$, we define $C_\pi \triangleq \sup_{t,\omega} |\sigma(t,\omega)^\top \pi(t,\omega)|$, and we denote by Σ_π the set of all strategies $\widetilde{\pi} \in \Sigma(0,T)$ such that $\sup_{t,\omega} |\sigma(t,\omega)^\top \widetilde{\pi}(t,\omega)| \leq C_\pi$,

3 The main result

Let $T > 0$ and $X_0 > 0$ be given. Let $U(\cdot) : (0, +\infty) \to \mathbf{R}$ be a non-decreasing on $(0, +\infty)$ right-continuous function such that there exist constant $C > 0$ and $N > 0$ such that

$$|U(x)| \le C \left(x^N + x^{-N}\right), \quad x > 0. \tag{3.1}$$

Let

$$J(\pi) \triangleq \mathbf{E} U(X(T, 0, X_0, \pi)).$$

We will study the problem

$$\text{Maximize} \quad J(\pi) \quad \text{over} \quad \pi(\cdot) \in \Sigma(0, T). \tag{3.2}$$

Note the class of admissible U is quite general. For instance, it includes functions $\log x$ and $\delta^{-1} x^\delta$ for $\delta < 1$, $\delta \ne 0$, as well as their linear combinations. Concavity of U, Inada conditions or asymptotic elasticity conditions are not required. However, typical utility functions satisfying these conditions are covered, as well as other functions such as right-continuous function $U(x) = \mathbb{I}_{\{x \ge c\}}$, where $\mathbb{I}$ is the indicator function, $c > 0$; these particular U are used for goal-achieving problems.

Theorem 3.1 *Mutual Fund Theorem holds in the following sense:*

$$\sup_{\pi \in \Sigma(0,T)} J(\pi) = \sup_{\pi \in \Sigma_{MFT}(0,T)} J(\pi). \tag{3.3}$$

Moreover, for any $\pi \in \Sigma(0, T)$, any admissible U, and any $\delta > 0$ there exists a strategy $\widehat{\pi} \in \Sigma_{MFT}(0, T) \cap \Sigma_\pi$ such that

$$J(\widehat{\pi}) \ge J(\pi) - \delta, \qquad . \tag{3.4}$$

Since the matrix σ^{-1} is bounded, it follows from the definition of Σ_π that there exits a constant $C_0 > 0$ that depends only on n and $\sigma(\cdot)$ such that

$$\sup_{t,\omega} |\widehat{\pi}(t, \omega)| \le C_0 \sup_{t,\omega} |\pi(t, \omega)|.$$

4 Proofs

In addition to the function U, we will be using the function $u(y) \triangleq U(e^y)$.

Definition 4.1 *Let $t_1, t_2 \in [0,T]$ be given, $t_1 < t_2$. We denote by $\Sigma_M(t_1,t_2)$ the class of all processes $\pi(\cdot) \in \Sigma(0,T)$ such that there exists a measurable function $u : \mathbf{R} \times [t_1,t_2] \times \Omega \to \mathbf{R}^n$ such that the following holds.*

(i) $\pi(t) = u(Y(t),t,\omega)$ for $t \in [t_1,t_2]$, where $Y(t) = \log \widetilde{X}(t)$, and where $\widetilde{X}(t)$ is the corresponding discounted wealth.

(ii) The random variable $u(y,t,\omega)$ is $\mathcal{F}_t^\mu$-measurable for all $(y,t) \in \mathbf{R} \times [t_1,t_2]$.

(iii) The function $u(y,t,\omega)$ is continuously differentiable in y for all t,ω, and there exists a constant $L > 0$ such that, for any $(y,t) \in \mathbf{R} \times [t_1,t_2]$,

$$|u(y_1,t,\omega) - u(y_2,t,\omega)| \le L|y_1 - y_2|, \quad |u(y,t,\omega)| \le L \quad a.s. \tag{4.1}$$

Lemma 4.1 *Let $u(y) = U(e^y)$ be bounded and continuous on $\mathbf{R}$ together with the derivatives $u'(y)$ and $u''(y)$. Let $\mu(t) = (r(t), \widetilde{a}(t), \sigma(t))$ be a non-random process. Then Mutual Fund Theorem holds in the following sense: for any $\pi \in \Sigma(0,T)$ and any $\delta > 0$, there exists a strategy $\widehat{\pi} \in \Sigma_{MFT}(0,T) \cap \Sigma_M(0,T) \cap \Sigma_\pi$ such that (3.4) hold and*

$$\widehat{\pi}(t,\omega)^\top = \nu(t,\omega)e(t)^\top \sigma(t)^{-1}$$

where $\nu(t,\omega)$ is a random scalar $\mathcal{F}_t$-adapted process such that $|\nu(t,\omega)| \le \sup_{t,\omega} |\sigma(t)\pi(t,\omega)|$, and where

$$e(t,\omega) = \frac{\theta(t)}{|\theta(t)|} \quad \text{if} \quad \theta(t) \neq 0, \qquad e(t,\omega) = 0 \quad \text{if} \quad \theta(t) = 0.$$

Proof of Lemma 4.1. Let $\pi \in \Sigma(0,T)$ and $\delta > 0$ be given. By the assumptions about $\Sigma(0,T)$, we have that $C_\pi = \sup_{t,\omega} |\sigma(t)^\top \pi(t,\omega)| < +\infty$. Clearly, the set Σ_π is convex.

Consider the optimal control problem with the controlled process $Y(t) \triangleq \log \widetilde{X}(t)$ and with Σ_π as the class of admissible strategies. By Theorem V.2.5(c) from Krylov (1980), p.225, we obtain that there exists a Markov strategy $\pi_M(t) \in \Sigma_M(0,T) \cap \Sigma_\pi$ such that $J(\pi_M) \ge J(\pi) - \delta$ and $\pi_M(t) = F(Y_M(t),t)$, where $F(x,t) : \mathbf{R} \times \mathbf{R} \to \mathbf{R}^n$ is a measurable bounded functions such that the derivative $\partial F(x,t)/\partial x$ is bounded. It follows that the solution of the closed equation for $Y_M(t) \triangleq \log \widetilde{X}(t,0,X_0,\pi_M)$ is a diffusion process.

Further, let us apply the idea of the proof of Theorem 1 from Khanna and Kulldorff (1999) adjusted to our case of the model without consumption. Let us select $\widehat{\pi}(t) \in \Sigma_{MFT}(0,T) \cap$

$\Sigma_M(0,T)$ such that $\widehat\pi(t)=f(\widehat Y(t),t)$, where $\widehat Y(t)\triangleq\log\widetilde X(t,0,X_0,\widehat\pi)$ and where the function $f(x,t):\mathbf{R}^2\to\mathbf{R}$ is defined as a solution of the finite dimensional maximization problem

$$\text{Maximize}\quad f^\top\widetilde a(t)\quad\text{over}\quad\{f\in\mathbf{R}^n:\quad |f^\top\sigma(t)|\le|F_M(x,t)^\top\sigma(t)|\}.$$

If $\theta(t)\neq0$ then the solution $f=f(x,t)$ is

$$f^\top=f(x,t)^\top=\nu(x,t)\theta(t)^\top\sigma(t)^{-1},\quad\text{where}\quad \nu(x,t)\triangleq\frac{|F_M(x,t)^\top\sigma(t)|}{|\theta(t)|}. \tag{4.2}$$

If $\theta(t)=0$ then, by the choice of $\Sigma(0,T)$, we have that $F_M(x,t)=0$, and the optimal vector is $f(x,t)=0$. Note the function $f(x,t)$ is bounded and satisfies Lipschitz condition in x uniformly in t.

We have that

$$\begin{aligned} d_t\widetilde X(t,0,X_0,\pi_M) &= \widetilde X(t,0,X_0,\pi_M)\pi_M(t)^\top\widetilde{\mathbf{S}}(t)^{-1}d\widetilde S(t)\\ &= \pi_M(s)^\top\widetilde a(t)dt+\pi_M(s)^\top\sigma(s)dw(s).\end{aligned}$$

Hence

$$dY_M(t)=\left(F_M(Y_M(t),t)^\top\widetilde a(t)-\frac12|F_M(Y_M(t),t)^\top\sigma(t)|^2\right)dt+F_M(Y_M(t),t)^\top\sigma(t)dw(t).$$

Let $\widehat Y(t)\triangleq\log\widetilde X(t,0,X_0,\widehat\pi)$. Similarly to the equation for $Y_M(t)$, we obtain that

$$d\widehat Y(t)=\left(f(\widehat Y(t),t)^\top\widetilde a(t)-\frac12|f(\widehat Y(t),t)^\top\sigma(t)|^2\right)dt+f(\widehat Y(t),t)^\top\sigma(t)dw(t).$$

Let

$$\xi(x,t)\triangleq f(x,t)^\top\widetilde a(t)-F_M(x,t)^\top\widetilde a(t).$$

By the choice of f, we have that $\xi(x,t)\ge0$. Hence

$$d\widehat Y(t)=\left(F_M(\widehat Y(t),t)^\top\widetilde a(t)+\xi(\widehat Y(t),t)-\frac12|f(\widehat Y(t),t)^\top\sigma(t)|^2\right)dt+f(\widehat Y(t),t)^\top\sigma(t)dw(t).$$

It follows that the Kolmogorov's equations for the distribution of $\widehat Y(t)$ have the same diffusion coefficient as for the distribution of $Y_M(t)$, and that the drift coefficient for the Kolmogorov's equations for the distribution of $\widehat Y(t)$ at any time is no less than the drift for the Kolmogorov's equation for $Y_M(t)$. It follows that $J(\widehat\pi)\ge J(\pi_M)\ge J(\pi)-\delta$.

By the selection of $\widehat\pi$, we have that $\pi\in\Sigma_\pi$. This completes the proof of Lemma 4.1. $\square$.

Let us consider now the case when the parameters are predicable on a some given finite horizon.

Lemma 4.2 *Let the function $u(y) = U(e^y)$ be bounded and continuous on $\mathbf{R}$ together with the derivatives $u'(y)$ and $u''(y)$. Let there exists a finite set $\{t_k\}_{k=0}^N$ such that $0 = t_0 < t_1 < ... < t_N = T$ and such that the values $\mu(t)|_{t\in[t_k,t_{k+1})}$ can be predicted at times t_k, meaning that $\mu(t)$ is $\mathcal{F}_{t_k}$-measurable for $t \in [t_k, t_{k+1})$, $k < N$. Then Mutual Fund Theorem holds in the following sense: for any $\pi \in \Sigma(0,T)$ and any $\delta > 0$, there exists a strategy $\widehat{\pi} \in \Sigma_{MFT}(0,T) \cap \Sigma_\pi$ such that (3.4) hold and*

$$\widehat{\pi}(t,\omega)^\top = \nu(t,\omega)e(t)^\top\sigma(t,\omega)^{-1}$$

where $\nu(t,\omega)$ is a random scalar $\mathcal{F}_t$-adapted process such that $|\nu(t,\omega)| \le \sup_{t,\omega}|\sigma(t)\pi(t,\omega)|$, and where

$$e(t,\omega) = \frac{\theta(t,\omega)}{|\theta(t,\omega)|} \quad \text{if} \quad \theta(t,\omega) \neq 0, \qquad e(t,\omega) = 0 \quad \text{if} \quad \theta(t,\omega) = 0.$$

Corollary 4.1 *Lemma 4.2 holds if there exists $\varepsilon > 0$ such that $\mu(t) = (r(t), \widetilde{a}(t), \sigma(t))$ is predictable with time horizon ε, meaning that $\mu(t+\tau)$ is $\mathcal{F}_t$-measurable for any $\tau \le \varepsilon$. Then Lemma 4.2 holds, i.e., the Mutual Fund Theorem holds in the sense of Lemma 4.2.*

Proof of Lemma 4.2. It suffices to prove that, for any strategy $\pi \in \Sigma(0,T)$ and any $\delta > 0$, there exists a strategy $\widehat{\pi} \in \Sigma_{MFT}(0,T)$ such that (3.4) holds. Let us construct $\widehat{\pi}$ as the following.

Let $\pi \in \Sigma(0,T)$ be given. Let $x(t)$ be the corresponding discounted wealth for the strategy π. Consider a strategy $\widehat{\pi} \in \Sigma_{MFT}(0,T)$ and a sequence of functions $V_k(x,\omega) : \mathbf{R} \times \Omega \to \mathbf{R}$ constructed recursively for $k = N-1, N-2, ..., 0$, such that the following holds.

(I) $\widehat{\pi} \in \cap_{k=1}^N \Sigma_M(t_{k-1}, t_k)$ and

$$\sup_{t,\omega}|\sigma(t,\omega)^\top\widehat{\pi}(x,t,\omega)| \le \sup_{t,\omega}|\sigma(t,\omega)^\top\pi(x,t,\omega)| \quad \text{a.s.}$$

(II) $V_N(x) = U(x)$.

(III) For $k = N-1, ..., 0$,

$$\mathbf{E}\{V_{k+1}(\widetilde{X}(t_{k+1}, t_k, x(t_k), \widehat{\pi}))|\mathcal{F}_{t_k}\} \ge \mathbf{E}\{V_{k+1}(x(t_{k+1}))|\mathcal{F}_{t_k}\} - \frac{\delta}{N} \quad \text{a.s.}$$

(IV) For $k = N, ..., 1$,

$$V_k(x) \triangleq \mathbf{E}\{U(\widetilde{X}(T, t_k, x, \widehat{\pi}))|\mathcal{F}_{t_k}\}.$$

The properties of this procedure are discussed in Propositions 4.1 and 4.2 below.

Proposition 4.1 *Under the assumptions of Lemma 4.2, the following holds.*

(i) The functions $V_{k+1}(x,\omega)$ are $\mathcal{F}_{t_k}$-measurable for all x, $k=0,1,...,N-1$.

(ii) The functions $V_k(x,\omega)$ are continuous in $x\in(0,+\infty)$ a.s.

(iii) The functions $V_k(x,\omega)$ are non-decreasing in $x\in(0,+\infty)$ a.s..

(iv) There exists a strategy $\widehat{\pi}$ and a sequence of functions $\{V_k\}$ with the properties (I)-(IV) listed above.

Proof of Proposition 4.1. In the following proof, we recall that $\widehat{\pi}$ and V_k are constructed recursively: for $k=N-1,N-2,...$, the process $\widehat{\pi}|_{t\in[t_k,t_{k+1}]}$ is selected using the value V_{k+1}. On the next step, V_k is defined using $\widehat{\pi}|_{t\in[t_k,t_{k+1}]}$.

Let us prove statement (i). By the requirement that $\widehat{\pi}\in\cap_{q=k+1}^{N-1}\Sigma_M(t_q,t_{q+1})$, it follows that the random variable $\widetilde{X}(T,t_{k+1},x,\widehat{\pi})$ is defined by x and by the set of all values $\{w(t)-w(t_{k+1}),\mu(t)\}_{t\in[t_{k+1},T]}$. The values $\mu(\cdot)|_{t\in[0,t_{k+1})}$ and $w(\cdot)|_{t\in[0,t_k)}$ are $\mathcal{F}_{t_k}$-measurable, and the process $\mu(\cdot)|_{t\in[t_{k+1},T]}$ is independent from $w(\cdot)$ the conditional probability space given $\mathcal{F}_{t_k}$. Hence the process $(w(\cdot)-w(t_{k+1}),\mu(\cdot))|_{t\in[t_{k+1},T]}$ is independent from $\{w(t),\mu(t)\}_{t\in[0,t_{k+1})}$ on the conditional probability space given $\mathcal{F}_{t_k}$. It follows that $V_{k+1}(x,\cdot)$ is independent from $\{w(t),\mu(t)\}_{t\in[t_k,t_{k+1}]}$ on the conditional probability space given $\mathcal{F}_{t_k}$. Then statement (i) follows.

Let us prove statement (ii). Let

$$Y(t,x)=\log X(t,t_{k+1},x,\widehat{\pi}),\quad t\geq t_{k+1},\ x>0.$$

By Ito formula, the equation for $Y(t,x)$ is

$$d_tY(t,x)=\pi(t)^\top\widetilde{a}(t)dt-\frac{1}{2}|\widehat{\pi}(t)^\top\sigma(t)|^2dt+\widehat{\pi}(t)^\top\sigma(t)dw(t).$$

It follows from the properties of $\widehat{\pi}|_{[t_{k+1},T]}$ that the equation for $Y(t,x)$ can be rewritten as

$$\begin{aligned}&d_tY(t,x)=f(Y(t,x),t,\omega)dt+b(Y(t,x),t,\omega)dw(t),\quad t>t_{k+1},\\&Y(t_{k+1},x)=\log x,\end{aligned}$$

where $f:\mathbf{R}\times[t_{k+1},T]\times\Omega\to\mathbf{R}$ and $b:\mathbf{R}\times[t_{k+1},T]\times\Omega\to\mathbf{R}^{1\times n}$ are bounded measurable functions such that the functions $f(y,t,\omega)$ and $b(y,t,\omega)$ are $\mathcal{F}_t^\mu$-adapted for any $(y,t)\in$

$\mathbf{R} \times [t_{k+1}, T]$ and such that there exists a constant $\widehat{C} > 0$ such that

$$|f(y_1,t,\omega) - f(y_2,t,\omega)| + |b(y_1,t,\omega) - b(y_2,t,\omega)| \le \widehat{C}(|y_1 - y_2|),$$
$$|f(y,t,\omega)| + |b(y,t,\omega)| \le \widehat{C} \qquad \forall (y,t) \in \mathbf{R} \times [t_1, t_2] \quad \text{a.s.}$$

By the definitions, $V_{k+1}(x)$ can be represented as

$$V_{k+1}(x) = \mathbf{E}\Big\{ u(Y(T,x)) \Big| \mathcal{F}_{t_{k+1}} \Big\}, \tag{4.3}$$

where $u(y) = U(e^y)$ is a bounded and continuous function.

To prove the continuity in x of $V_{k+1}(x,\omega)$ it suffices to prove that, for a.e. ω, for any sequence $x_i \to x$ there is a subsequence converging to x for which $V_{k+1}(x_i,\omega) \to V_{k+1}(x,\omega)$.

Let $\{x_i\}$ be a sequence converging to x. By Theorem II.8.1 from Krylov (1980), p.102, it follows that if $x_i \to x$ then

$$\mathbf{E}\{|Y(T,x_i) - Y(T,x)|^2 \,|\mathcal{F}_{t_{k+1}}\} \to 0 \quad \text{as} \quad i \to +\infty.$$

It follows that, with probability 1 on the conditional probability space given $\mathcal{F}_{t_{k+1}}$, there exists a subsequence $\{x_i\}$ converging to x for which

$$Y(T,x_i) \to Y(T,x) \quad \text{as} \quad i \to +\infty.$$

By Lebesgue's Dominated Convergence Theorem, this subsequence $\{x_i\}$ is such that

$$\mathbf{E}\{|u(Y(T,x_i) - u(Y(T,x)|^2 \,|\mathcal{F}_{t_{k+1}}\} \to 0 \quad \text{as} \quad x_i \to x. \tag{4.4}$$

By (4.3) and (4.4), it follows that

$$V_{k+1}(x_i,\omega) \to V_{k+1}(x,\omega) \quad \text{as} \quad x_i \to 0. \tag{4.5}$$

Then the statement (ii) follows.

Let us prove statement (iii). We will use the (4.3) and the process $Y(x,t)$ introduced above. It follows from Theorem 2.8.4 from Krylov (1980) and from the corresponding proof that, for a given $x > 0$,

$$\frac{dY}{dx}(x,T) = \lim_{\varepsilon \to 0} \varepsilon^{-1}(Y(T,x+\varepsilon) - Y(T,x)) = \xi(T,x),$$

where the limit exists in $L_2(\Omega)$ and where the process $\xi(t,x)$ is such that

$$d_t\xi(t,x) = \frac{\partial f}{\partial y}(Y(t,x),t)\xi(t,x)dt + \frac{\partial b}{\partial y}(Y(t,x),t)\xi(t,x)dw(t), \qquad t \ge t_{k+1},$$
$$\xi(t_{k+1},x) = x^{-1}.$$

Clearly, $\xi(T,x) \geq 0$. By the assumptions, the function $u(y)$ in (4.3) is nondecreasing and absolutely continuous. Then statement (iii) follows.

Let us prove statement (iv). For $k = N-1, N-2, ..., 0$, the existence of the corresponding $\widehat{\pi}|_{[t_k,t_{k+1}]}$ follows from Lemma 4.1 applied on the time interval $[t_k, t_{k+1}]$ and on the conditional probability space given $\mathcal{F}_{t_k}$. Statements (i)-(iii) ensure applicability of this lemma for each step $k = N-1, N-2, ..0$. Then statement (iv) follows. This completes the proof of Proposition 4.1. □

The following Proposition establishes an analog of Markov property.

Proposition 4.2 *Under the assumptions of Lemma 4.2, let* $m \in \{0, .., N-1\}$, *and let*

$$\alpha \triangleq x(t_{m+1}), \qquad \beta \triangleq \widetilde{X}(t_{m+1}, t_m, x(t_m), \widehat{\pi}).$$

Then

$$\begin{aligned} \mathbf{E}\{V_{m+1}(\alpha)|\mathcal{F}_{t_m}\} &= \mathbf{E}\{U(\widetilde{X}(T, t_{m+1}, \alpha, \widehat{\pi}))|\mathcal{F}_{t_m}\}, \\ \mathbf{E}\{V_{m+1}(\beta)|\mathcal{F}_{t_m}\} &= \mathbf{E}\{U(\widetilde{X}(T, t_{m+1}, \beta, \widehat{\pi}))|\mathcal{F}_{t_m}\}. \end{aligned} \tag{4.6}$$

Proof. By the definitions,

$$V_{m+1}(x) = \mathbf{E}\{U(\widetilde{X}(T, t_{m+1}, x, \widehat{\pi}))|\mathcal{F}_{t_{m+1}}\}. \tag{4.7}$$

We have that α and β depends only on $\{w(t), \mu(t)\}_{t<t_{m+1}}$ on the conditional probability space given $\mathcal{F}_{t_m}$. Since $\mu(t)$ is $\mathcal{F}_{t_m}$-measurable for all $t < t_{m+1}$, we have that α and β depend only on $\{w(t)\}_{t<t_{m+1}}$ on the conditional probability space given $\mathcal{F}_{t_m}$. In addition, the process $\{\mu(t)\}_{t\geq t_{m+1}}$ is independent from $w(\cdot)$. Therefore, α and β are independent from the process $\{w(t) - w(t_{m+1}), \mu(t)\}_{t\geq t_{m+1}}$ on the conditional probability space given $\mathcal{F}_{t_m}$. On the other hand, the process $\widehat{\pi}(t)|_{t\geq t_{m+1}}$ is adapted to the filtration generated by $\{w(t) - w(t_{m+1}), \mu(t)\}_{t\geq t_{m+1}}$ on the conditional probability space given $\mathcal{F}_{t_m}$ and given the initial value of the discounted wealth $\widetilde{X}(t_{m+1})$. By the version of the Markov property described in Theorem II.9.4 from Krylov (1980), p.113, and applied on the conditional space given $\mathcal{F}_{t_m}$ and on the time interval $[t_{m+1}, T]$, we have that (4.6) follows from (4.7). This completes the proof of Proposition 4.2. □

We now in the position to complete the proof of Lemma 4.2.

Clearly, it suffices to prove that, for any $m \in \{0, 1, ..., N-1\}$,

$$\mathbf{E}\{U(\widetilde{X}(T, t_m, x(t_m), \widehat{\pi}))|\mathcal{F}_{t_m}\} \geq \mathbf{E}\{U(x(T))|\mathcal{F}_{t_m}\} - \frac{N-m}{N}\delta. \tag{4.8}$$

We will use mathematical induction with decreasing m.

First, (4.8) holds for $m = N-1$ by Lemma 4.1 applied on the time interval $[t_{N-1}, T]$ and on the corresponding conditional probability space given $\mathcal{F}_{t_{N-1}}$. It suffices to prove that if (4.8) holds for some $m+1 \le N$ then it implies that it holds for m.

By the definitions and by the induction assumption that (4.8) holds with m replaced by $m+1$, we obtain that

$$\mathbf{E}\{U(x(T))|\mathcal{F}_{t_{m+1}}\} \le \mathbf{E}\{U(\widetilde{X}(T, t_{m+1}, x(t_{m+1}), \widehat{\pi}))|\mathcal{F}_{t_{m+1}}\} + \frac{N-m-1}{N}\delta. \tag{4.9}$$

Hence

$$\mathbf{E}\{U(x(T))|\mathcal{F}_{t_m}\} \le \mathbf{E}\{U(\widetilde{X}(T, t_{m+1}, x(t_{m+1}), \widehat{\pi}))|\mathcal{F}_{t_m}\} + \frac{N-m-1}{N}\delta$$

Let α and β be defined by (4.6). By Proposition 4.2 for $\alpha = x(t_{m+1})$, it follows that

$$\mathbf{E}\{U(x(T))|\mathcal{F}_{t_m}\} \le \mathbf{E}\{V_{m+1}(x(t_{m+1}))|\mathcal{F}_{t_m}\} + \frac{N-m-1}{N}\delta. \tag{4.10}$$

Further, by the choice of $\widehat{\pi}|_{t\in[t_m,t_{m+1}]}$, we obtain that

$$\mathbf{E}\{V_{m+1}(x(t_{m+1}))|\mathcal{F}_{t_m}\} \le \mathbf{E}\{V_{m+1}(\widetilde{X}(t_{m+1}, t_m, x(t_m), \widehat{\pi}))|\mathcal{F}_{t_m}\} + \frac{\delta}{N}.$$

By (4.10), it follows that

$$\begin{aligned}\mathbf{E}\{U(x(T))|\mathcal{F}_{t_m}\} \le \mathbf{E}\{V_{m+1}(\widetilde{X}(t_{m+1}, t_m, x(t_m), \widehat{\pi}))|\mathcal{F}_{t_m}\} + \frac{\delta}{N} + \frac{N-m-1}{N}\delta \\ = \frac{N-m}{N}\delta. \end{aligned}\tag{4.11}$$

By the definitions,

$$\begin{aligned}\mathbf{E}\{U(\widetilde{X}(T, t_{m+1}, \alpha, \widehat{\pi}))|\mathcal{F}_{t_m}\} &= \mathbf{E}\{U(\widetilde{X}(T, t_{m+1}, \widetilde{X}(t_{m+1}, t_m, x(t_m), \widehat{\pi}), \widehat{\pi}))|\mathcal{F}_{t_m}\} \\ &= \mathbf{E}\{U(\widetilde{X}(T, t_m, x(t_m), \widehat{\pi}))|\mathcal{F}_{t_m}\}. \end{aligned}\tag{4.12}$$

By Proposition 4.2 for β, we have

$$\mathbf{E}\{V_{m+1}(\beta)|\mathcal{F}_{t_m}\} = \mathbf{E}\{U(\widetilde{X}(T, t_m, x(t_m), \widehat{\pi}))|\mathcal{F}_{t_m}\}. \tag{4.13}$$

By (4.10) and (4.6), it follows that

$$\mathbf{E}\{U(\widetilde{X}(T, t_m, x(t_m), \pi))|\mathcal{F}_{t_m}\} \le \mathbf{E}\{U(\widetilde{X}(T, t_m, x(t_m), \widehat{\pi}))|\mathcal{F}_{t_m}\} - \frac{N-m}{N}\delta.$$

Since it holds for any $\pi \in \Sigma(t_m, T)$, it follows that Lemma 4.2 holds. □

Lemma 4.3 *Theorem 3.1 holds for the case when the function $u(y) = U(e^y)$ is bounded and continuous on* $\mathbf{R}$ *together with the derivatives $u'(y)$ and $u''(y)$. In addition, the statement of Lemma 4.2 holds in this case.*

Proof. Let $t \vee s = \max(t, s)$,

$$r_\varepsilon(t) \triangleq \frac{1}{\varepsilon}\int_{(t-2\varepsilon)\vee 0}^{(t-\varepsilon)\vee 0} r(s)ds, \quad a_\varepsilon(t) \triangleq \frac{1}{\varepsilon}\int_{(t-2\varepsilon)\vee 0}^{(t-\varepsilon)\vee 0} a(s)ds, \quad \sigma_\varepsilon(t) \triangleq \frac{1}{\varepsilon}\int_{(t-2\varepsilon)\vee 0}^{(t-\varepsilon)\vee 0} \sigma(s)ds,$$

and let

$$\mu_\varepsilon(t) \triangleq (r_\varepsilon(t), \widetilde{a}_\varepsilon(t), \sigma_\varepsilon(t)), \quad \widetilde{a}_\varepsilon(t) \triangleq a_\varepsilon(t) - r_\varepsilon(t), \quad \theta_\varepsilon(t) \triangleq \sigma_\varepsilon(t)^{-1}\widetilde{a}_\varepsilon(t).$$

Consider a sequence $\varepsilon = \varepsilon_N = 1/N \to 0$, $N = 1, 2, ...$ For every $\varepsilon = \varepsilon_i$, consider a finite sequences of times $\{t_j\}_{j=0}^N$ such that $t_{k+1} = t_k + \varepsilon$.

Let $\mathcal{F}_t^{\mu,\varepsilon}$ be the filtration generated by $\mu_\varepsilon(t)$ and let $\mathcal{F}_t^\varepsilon$ be the filtration generated by $(\mu_\varepsilon(t), w(t))$.

Let $\widetilde{\Sigma}(0,T)$ be the class of all $\mathcal{F}_t^\varepsilon$-adapted processes $\pi(\cdot) = (\pi_1(\cdot), \ldots, \pi_n(\cdot)) : [0,T] \times \Omega \to \mathbf{R}^n$ such that $\sup_{t,\omega} |\pi(t,\omega)| < +\infty$ and that if $\theta_\varepsilon(t) = 0$ then $\pi(t) = 0$.

Further, let $\Sigma_{\varepsilon,MFT}(0,T)$ denote the set of strategies from $\Sigma_\varepsilon(0,T)$ that have the form $\pi(t) = \nu(t)\sigma_\varepsilon(t)^{-1}\theta_\varepsilon(t)$, where $\nu_\varepsilon(t)$ is an one dimensional process adapted to $\mathcal{F}_t^\varepsilon$.

For $\varepsilon > 0$, let

$$J_\varepsilon(\pi) \triangleq \mathbf{E}U(\widetilde{X}_\varepsilon(T, 0, X_0, \pi)),$$

where $\widetilde{X}_\varepsilon(T, 0, X_0, \pi)$ is the discounted wealth for the model with μ replaced by $\mu = \mu_\varepsilon$ for the strategy π given that $\widetilde{X}(0) = X_0$. The case of $\varepsilon = 0$ corresponds to the original model; in this case, the discounted wealth is denoted as $\widetilde{X}(T, 0, X_0, \pi)$.

Note that the market models with $\mu = \mu_\varepsilon$ are such that the assumptions of Lemma 4.2 are satisfied for $\varepsilon > 0$.

Let $\delta > 0$ be given. Let $\pi \in \Sigma(0,T)$ be such that

$$J(\pi) \geq \inf_{\pi \in \Sigma(0,T)} J(\pi) - \frac{\delta}{4}.$$

Let $\widetilde{X}(t) = \widetilde{X}(T, 0, X_0, \pi)$. By the choice of $\Sigma(0,T)$, we have that $C_\pi \triangleq \sup_{t,\omega} |\pi(t,\omega)| < +\infty$. Let

$$\pi_\varepsilon(t) \triangleq \frac{1}{\varepsilon}\int_{(t-2\varepsilon)\vee 0}^{(t-\varepsilon)\vee 0} \pi(s)ds.$$

Clearly, $\pi_\varepsilon \in \Sigma_\varepsilon(0,T)$. By Lemma 3 from Shilov and Gurevich (1967), Chapter IV, Section 5, it follows that

$$\mu_\varepsilon \to \mu, \quad \pi_\varepsilon \to \pi \quad \text{as} \quad \varepsilon \to 0+ \quad \text{a.e. on} \quad [0,T]\times\Omega.$$

We have that

$$\begin{aligned}\widetilde{X}(T,0,X_0,\pi) &= X_0 + \int_0^T \widetilde{X}(t,0,X_0,\pi)\pi(t)^\top \widetilde{\mathbf{S}}(t)^{-1} d\widetilde{S}(t)\\ &= X_0 \exp\left[\int_0^T \pi(t)^\top \widetilde{a}(t)dt - \frac{1}{2}\int_0^T |\pi(t)^\top\sigma(t)|^2 dt + \int_0^T \pi(t)^\top \sigma(t)dw(t)\right]. \end{aligned} \tag{4.14}$$

Similarly,

$$\begin{aligned}&\widetilde{X}_\varepsilon(T,0,X_0,\pi_\varepsilon)\\ &= X_0 \exp\left[\int_0^T \pi_\varepsilon(t)^\top \widetilde{a}_\varepsilon(t)dt - \frac{1}{2}\int_0^T |\pi_\varepsilon(t)^\top\sigma_\varepsilon(t)|^2 dt + \int_0^T \pi_\varepsilon(t)^\top \sigma_\varepsilon(t)dw(t)\right]. \end{aligned} \tag{4.15}$$

Let $Y_{\varepsilon,\varepsilon}(t) \triangleq \log \widetilde{X}_\varepsilon(t,0,X_0,\pi_\varepsilon)$ and $Y(t) \triangleq \log \widetilde{X}(t,0,X_0,\pi)$.

Clearly,

$$\mathbf{E}|Y_{\varepsilon,\varepsilon}(T) - Y(T)|^2 \to 0 \quad \text{as} \quad \varepsilon \to 0. \tag{4.16}$$

It follows that there exists a subsequence $\{\varepsilon\} = \{\varepsilon_i\}$ such that

$$Y_{\varepsilon,\varepsilon}(T) \to Y(T) \quad \text{a.s.} \quad \text{as} \quad \varepsilon = \varepsilon_i \to 0. \tag{4.17}$$

By Lebesgue's Dominated Convergence Theorem, this subsequence $\{\varepsilon\} = \{\varepsilon_i\}$ is such that

$$\mathbf{E}|U(\widetilde{X}_\varepsilon(T,0,X_0,\pi_\varepsilon)) - U(\widetilde{X}(T,0,X_0,\pi))|^2 \to 0 \quad \text{as} \quad \varepsilon = \varepsilon_i \to 0. \tag{4.18}$$

It follows that

$$\mathbf{E}U(\widetilde{X}_\varepsilon(T,0,X_0,\pi_\varepsilon)) \to \mathbf{E}U(\widetilde{X}(T,0,X_0,\pi)) \quad \text{as} \quad \varepsilon = \varepsilon_i \to 0. \tag{4.19}$$

In other words,

$$J_\varepsilon(\pi_\varepsilon) = \mathbf{E}U(\widetilde{X}_\varepsilon(T,0,X_0,\pi_\varepsilon)) \to \mathbf{E}U(\widetilde{X}(T,0,X_0,\pi)) = J(\pi) \quad \text{as} \quad \varepsilon = \varepsilon_i \to 0. \tag{4.20}$$

It follows that there exists $N_1 > 0$ such that, for every $i \ge N_1$,

$$J_\varepsilon(\pi_\varepsilon) \ge J(\pi) - \frac{\delta}{4}, \quad \varepsilon = \varepsilon_i.$$

Let $\pi_{\varepsilon,\varepsilon} \in \Sigma_{\varepsilon,MFT}(0,T)$ be the strategy defined in Lemma 4.2 as a strategy such that $\nu_\varepsilon(t)$ is $\mathcal{F}_t^\varepsilon$-adapted process and

$$J_\varepsilon(\pi_{\varepsilon,\varepsilon}) \geq J_\varepsilon(\pi_\varepsilon) - \frac{\delta}{4}.$$

Following the proof of Lemma 4.1 we obtain similarly to (4.2) that, if $\theta(t) \neq 0$, then

$$\pi_{\varepsilon,\varepsilon}(t)^\top = \nu_\varepsilon(t) e_\varepsilon(t)^\top \sigma_\varepsilon(t)^{-1}, \quad \text{where} \quad e_\varepsilon(t) = \frac{\theta_\varepsilon(t)}{|\theta_\varepsilon(t)|}, \tag{4.21}$$

and where $\nu_\varepsilon(t) = \nu_\varepsilon(t,\omega)$ is a $\mathcal{F}_t^\varepsilon$-adapted one-dimensional process such that $|\nu_\varepsilon(t,\omega)| \leq \sup_{t,\omega} |\sigma_\varepsilon(t,\omega)^\top \pi(t,\omega)| \leq C_\pi$. If $\theta(t) = 0$ then $\widehat{\pi}_{\varepsilon,\varepsilon}(t) = 0$.

It follows that

$$\sup_{t,\omega,\varepsilon} |\sigma_\varepsilon(t,\omega)^\top \pi_{\varepsilon,\varepsilon}(t,\omega)| \leq C_\pi. \tag{4.22}$$

Let

$$\pi_{\varepsilon,0}(t)^\top \triangleq \nu_\varepsilon(t) e(t)^\top \sigma(t)^{-1} \quad \text{if} \quad \theta(t) \neq 0,\ \theta_\varepsilon(t) \neq 0,$$
$$\pi_{\varepsilon,0}(t) = 0 \quad \text{if} \quad \theta(t) = 0 \quad \text{or} \quad \theta_\varepsilon(t) = 0.$$

It follows that, if $\theta(t) \neq 0$, $\theta_\varepsilon(t) \neq 0$

$$\sup_{t,\omega,\varepsilon} |\sigma(t,\omega)^\top \pi_{\varepsilon,0}(t,\omega)| \leq C_\pi. \tag{4.23}$$

The equations for $\widetilde{X}_\varepsilon(T,0,X_0,\pi_{\varepsilon,\varepsilon})$ and $\widetilde{X}_\varepsilon(T,0,X_0,\pi_{\varepsilon,0})$ are similar to equations (4.14)-(4.15). Clearly, $\pi_{\varepsilon,\varepsilon}(t,\omega) - \pi_{\varepsilon,0}(t,\omega) \to 0$ a.e.. Using (4.22)-(4.23), we obtain that $\mathbf{E}|\log \widetilde{X}_\varepsilon(T,0,X_0,\pi_{\varepsilon,\varepsilon}) - \log \widetilde{X}(T,0,X_0,\pi_{\varepsilon,0})|^2 \to 0$ as $\varepsilon \to 0$. It follows that there exists another subsequence $\{\varepsilon\} = \{\varepsilon_i\}$ (a subsequence of the subsequence from (4.17)) such that $\varepsilon_i \to 0$ and $\log \widetilde{X}_\varepsilon(T,0,X_0,\pi_{\varepsilon,\varepsilon}) - \log \widetilde{X}(T,0,X_0,\pi_{\varepsilon,0}) \to 0$ a.s. as $\varepsilon = \varepsilon_i \to 0$. Similarly to (4.19)-(4.20), we obtain that this subsequence $\{\varepsilon\} = \{\varepsilon_i\}$ is such that

$$J_\varepsilon(\pi_{\varepsilon,\varepsilon}) - J_\varepsilon(\pi_{\varepsilon,0}) = \mathbf{E}U(\widetilde{X}_\varepsilon(T,0,X_0,\pi_{\varepsilon,\varepsilon})) - \mathbf{E}U(\widetilde{X}(T,0,X_0,\pi_{\varepsilon,0})) \to 0$$

as $\varepsilon = \varepsilon_i \to 0$. It follows that there exists $N > N_1 > 0$ such that, for every $i \geq N$,

$$J_0(\pi_{\varepsilon,0}) \geq J_\varepsilon(\pi_{\varepsilon,\varepsilon}) - \frac{\delta}{4}, \quad \varepsilon = \varepsilon_i.$$

Finally, we obtain that

$$J_0(\pi_{\varepsilon,0}) \geq J_\varepsilon(\pi_{\varepsilon,\varepsilon}) - \frac{\delta}{4} \geq J_\varepsilon(\pi_\varepsilon) - \frac{\delta}{2} \geq J_0(\pi) - \frac{3\delta}{4}, \quad \varepsilon = \varepsilon_i.$$

This completes the proof of Lemma 4.3. $\square$

Lemma 4.4 *If the functions U is bounded, then Theorem 3.1 holds and Lemma 4.2 holds.*

Proof. It suffices to show that, for any $\delta > 0$ and $\pi \in \Sigma(0,T)$, there exists $\widehat{\pi} \in \Sigma_{MFT}(0,T) \cap \Sigma_\pi$ such that (3.4) holds.

For $\varepsilon > 0$, let $u_\varepsilon(y)$ be a function that is continuous and bounded $\mathbf{R}$ together with the derivatives $u_\varepsilon'(y)$ and $u_\varepsilon''(y)$ and such that $\sup_y |u(y) - u_\varepsilon(y)| \le \varepsilon$. Let $U_\varepsilon(x) = u_\varepsilon(\log x)$. It follows that $\sup_{x>0} |U(x) - U_\varepsilon(x)| \le \varepsilon$. Let $\delta > 0$ and $\pi \in \Sigma(0,T)$ be given. Let $\mathcal{J}_\varepsilon(\pi) = \mathbf{E}U_\varepsilon(\widetilde{X}(T,0,X_0,\pi))$.

Clearly, there exists $\varepsilon > 0$ such that $|J(\pi) - \mathcal{J}_\varepsilon(\pi)| \le \delta/3$ for any π. (It suffices to define $\widetilde{U}_\varepsilon$ as the convolution of U with appropriate convolution kernels such as $\bar{k}_\varepsilon(x) = \varepsilon^{-1}\bar{k}(x/\varepsilon)$, where $\bar{k}(x)$ is the density for the standard normal distribution or some other appropriate smoothing kernel). It follows that $J(\pi) \ge \mathcal{J}_\varepsilon(\pi) - \delta/3$ for any π.

By Lemma 4.3, there exists $\widehat{\pi} \in \Sigma_{MFT}(0,T) \cap \Sigma_\pi$ such that $\mathcal{J}_\varepsilon(\widehat{\pi}) \ge \mathcal{J}_\varepsilon(\pi) - \delta/3$. We have that

$$J(\widehat{\pi}) \ge \mathcal{J}_\varepsilon(\widehat{\pi}) - \delta/3 \ge \mathcal{J}_\varepsilon(\pi) - 2\delta/3 \ge J(\pi) - \delta.$$

Then the proof follows. □

Lemma 4.5 *Theorem 3.1 holds for the case when* $\sup_{x>0} U(x) < +\infty$ *is bounded.*

Proof of Lemma 4.5. It suffices to show that, for any $\delta > 0$ and $\pi \in \Sigma(0,T)$, there exists $\widehat{\pi} \in \Sigma_{MFT}(0,T) \cap \Sigma_\pi$ such that (3.4) holds.

Consider a sequence of positive integers $\{K\}$ such that $K \to +\infty$. Let $U^{(K)}(x)$ be defined by (3.1) with U replaced by $\max(-K, U(x))$. Let $\widetilde{\mathcal{J}}_K(\pi) = \mathbf{E}U^{(K)}(\widetilde{X}(T,0,X_0,\pi))$.

Without loss of generality, we assume that $U(1) = 0$. In this case, there exists $N > 0$, $C > 0$ such that $-C/x^N \le U(x) \le 0$ for $x \in (0,1)$.

Let $V_K(\widetilde{\pi})$ be the event $\{\widetilde{X}(T,0,X_0,\widetilde{\pi}) < (C/K)^{1/N}\}$. It follows that $\{U(\widetilde{X}(T,0,X_0,\pi)) < -K\} \subseteq V_K(\widetilde{\pi})$.

Proposition 4.3 *There exists a constant $c > 0$ such that, for all $\widetilde{\pi} \in \Sigma_\pi$,*

$$\mathbf{P}(V_K(\pi)) \le \frac{c}{\psi(K)} e^{-c\psi(K)^2}.$$

Proof of Proposition 4.3. Let $\widetilde{\pi} \in \Sigma_\pi$ be given. For $\varepsilon \in (0,1)$, set $\widetilde{\pi}_\varepsilon(t,\omega) = \widetilde{\pi}(t,\omega)$ if $|\widetilde{\pi}(t)| > \varepsilon$, and $\widetilde{\pi}_\varepsilon(t,\omega) = (\varepsilon, 0, 0, .., 0)^\top$ if $|\widetilde{\pi}(t)| \le \varepsilon$. Let $Y_\varepsilon(t) \triangleq \log \widetilde{X}(t,0,X_0,\widetilde{\pi}_\varepsilon)$. We have that

$$\mathbf{P}(V_K(\widetilde{\pi}_\varepsilon)) \le \mathbf{P}(\widetilde{X}(T,0,X_0,\widetilde{\pi}_\varepsilon) < (C/K)^{1/N}) = \mathbf{P}(Y_\varepsilon(t) < N^{-1}\log C - N^{-1}\log K).$$

It can be seen that $Y_\varepsilon(t) = Z_\varepsilon(t) + M_\varepsilon(t)$, where

$$Z_\varepsilon(t) = \log X_0 + \int_0^t \widetilde{\pi}_\varepsilon(s)^\top \widetilde{a}(s)ds - \frac{1}{2}\int_0^t |\widetilde{\pi}_\varepsilon(s)^\top \sigma(s)|^2 ds,$$

$$M_\varepsilon(t) = \int_0^t \widetilde{\pi}_\varepsilon(s)^\top \sigma(s)dw(s).$$

Let $\xi_\varepsilon(t)^\top = \widetilde{\pi}_\varepsilon(t)^\top \sigma(t)$. We have that

$$Z_\varepsilon(t) = \log X_0 + \int_0^t \xi_\varepsilon(s)^\top \sigma(s)^{-1}\widetilde{a}(s)ds - \frac{1}{2}\int_0^t |\xi_\varepsilon(s)|^2 ds,$$

$$M_\varepsilon(t) = \int_0^t \xi_\varepsilon(s)^\top dw(s).$$

Clearly, there exist $C_Z > 0$ such that $|Z_\varepsilon(T)| \le C_Z$ for all $\widetilde{\pi} \in \Sigma_\pi$ and all $\varepsilon \in (0,1)$. It follows that

$$\begin{aligned} &\mathbf{P}(V_K(\widetilde{\pi}_\varepsilon)) \le \mathbf{P}(Z_\varepsilon(T) + M_\varepsilon(T) < N^{-1}\log C - N^{-1}\log K) \\ &\le \mathbf{P}(M_\varepsilon(T) < -Z_\varepsilon(T) + N^{-1}\log C - N^{-1}\log K) \le \mathbf{P}(M_\varepsilon(T) < -\psi(K)), \end{aligned} \tag{4.24}$$

where

$$\psi(K) \triangleq C_Z - N^{-1}\log C + N^{-1}\log K.$$

We assume that $K > 0$ is large enough such that $\psi(K) > 0$.

We have that there exists a constant $k > 0$ such that $|\xi_\varepsilon(t,\omega)| \ge k\varepsilon$ for all t, ω, and that $M_\varepsilon(t)$ is a martingale with quadratic variation process

$$[M_\varepsilon](t) \triangleq \int_0^t |\xi_\varepsilon(s)|^2 ds.$$

Let $\rho_\varepsilon(t) \triangleq [M_\varepsilon]^{-1}(t)$ be the inverse function, i.e.,

$$\rho_\varepsilon(t) = \inf\{s \ge 0 : \ [M_\varepsilon](s) = t\}, \qquad t = \int_0^{\rho_\varepsilon(t)} |\xi_\varepsilon(s)|^2 ds.$$

Note that $\rho_\varepsilon(0) = 0$, and the process $\rho_\varepsilon(t)$ is strictly increasing in t.

Let $\tau = \tau(\pi_\varepsilon) = [M_\varepsilon](T)$ and $\mathcal{T} = \sup_{\varepsilon\in(0,1)} \sup_{\widetilde{\pi}\in\Sigma_\pi} \sup_{\omega\in\Omega} [M_\varepsilon](T)$. Note that $\mathcal{T}$ is non-random and τ is measurable with respect to $\mathcal{F}_T^\mu$, where F_t^μ is the filtration generated by the process of market parameters $\mu(t)$.

By Dambis–Dubins–Schwarz Theorem (see, e.g., Revuz and Yor (1999)), the process $W_\varepsilon(t) \triangleq M_\varepsilon(\rho_\varepsilon(t))$ is a Wiener process for t on the conditional probability space given μ, and $M_\varepsilon(t) = W([M_\varepsilon](t))$. In particular, $M_\varepsilon(T) = W_\varepsilon(\tau)$. Hence

$$\mathbf{P}(V_K(\pi_e)) \le \mathbf{P}(M_\varepsilon(T) < -\psi(K)) = \mathbf{P}(W_\varepsilon(\tau) < -\psi(K)).$$

Consider events $A_t = A_{t,\varepsilon} = \{W_\varepsilon(t) < -\psi(K)\}$. We have that

$$\mathbf{P}(V_K(\pi_\varepsilon)) \le \mathbf{P}(A_\tau) = \frac{\mathbf{P}(A_\mathcal{T})\mathbf{P}(A_\tau|A_\mathcal{T})}{\mathbf{P}(A_\mathcal{T}|A_\tau)} \le \frac{\mathbf{P}(A_\mathcal{T})}{\mathbf{P}(A_\mathcal{T}|A_\tau)}. \tag{4.25}$$

It is known that that $\mathbf{P}(\zeta < -s) \le \frac{1}{s2\pi}e^{-s^2/2}$, if ζ is a Gaussian random variable such that $\mathbf{E}\zeta = 0$ and $\operatorname{Var}\zeta = 1$. It follows that $\mathbf{P}(\zeta < -s) \le \frac{v}{s2\pi}e^{-s^2/(2v^2)}$, if ζ is a Gaussian random variable such that $\mathbf{E}\zeta = 0$ and $\operatorname{Var}\zeta = v^2 > 0$. We have that $W_\varepsilon(\mathcal{T})$ is a Gaussian random variable such that $\mathbf{E}W_\varepsilon(\mathcal{T}) = 0$ and $\operatorname{Var}W_\varepsilon(\mathcal{T}) = \mathcal{T}$. Hence

$$\mathbf{P}(A_\mathcal{T}) \le \frac{\sqrt{\mathcal{T}}}{\psi(K)2\pi}e^{-\psi(K)^2/(2\mathcal{T})}. \tag{4.26}$$

Further, we have that $W_\varepsilon(\mathcal{T}) = W_\varepsilon(\tau) + \Delta$, where $\Delta = W_\varepsilon(\mathcal{T}) - W_\varepsilon(\tau)$. By the definitions, $\mathbf{P}(A_\mathcal{T}|A_\tau) \ge \mathbf{P}(\Delta < 0|A_\tau)$. Since $\Delta = W_\varepsilon(\mathcal{T}) - W_\varepsilon(\tau)$ is independent from $W(\tau)$ on the conditional probability space given μ, we have that $\mathbf{P}(\Delta < 0|A_\tau, \mu) = 1/2$ for any μ, any π, and any $\varepsilon \in (0,1)$. It follows that $\mathbf{P}(\Delta < 0|A_\tau) = 1/2$. By (4.25),(4.26), the statement of proposition follows for $\widetilde{\pi} = \widetilde{\pi}_\varepsilon$ for any $\varepsilon \in (0,1)$. Similarly to (4.18), we obtain that $\widetilde{X}(T,0,X_0,\widetilde{\pi}_\varepsilon) \to \widetilde{X}(T,0,X_0,\widetilde{\pi})$ in probability as $\varepsilon \to 0$ and therefore $\mathbf{P}(V_K(\widetilde{\pi}_e)) \to \mathbf{P}(V_K(\widetilde{\pi}))$ as $\varepsilon \to 0$. This completes the proof of Proposition 4.3. □

We now are in position to complete the proof of Lemma 4.5. Let $\mathbb{I}$ denote the indicator function of an event. We have that

$$J(\widetilde{\pi}) - \widetilde{\mathcal{J}}_K(\widetilde{\pi}) = \mathbf{E}\mathbb{I}_{V_K(\widetilde{\pi})}(U(\widetilde{X}(T,0,X_0,\widetilde{\pi}) - K).$$

By Hölder inequality,

$$|J(\widetilde{\pi}) - \widetilde{\mathcal{J}}_K(\widetilde{\pi})| \le \left[\mathbf{E}\mathbb{I}_{V_K(\widetilde{\pi})}\right]^{1/2}\left[\mathbf{E}\left(U(\widetilde{X}(T,0,X_0,\widetilde{\pi}) - K)\right)^2\right]^{1/2}.$$

It follows that, by the assumptions on U, there there exist $m > 0$, $C_i > 0$ such that, for all $\widetilde{\pi} \in \Sigma_\pi$,

$$|J(\widetilde{\pi}) - \widetilde{\mathcal{J}}_K(\widetilde{\pi})| \le C_0(\mathbf{P}(V_K(\widetilde{\pi})))^{1/2}\mathbf{E}(|K| + |\widetilde{X}(T,0,X_0,\widetilde{\pi})|^m) \le C_1(\mathbf{P}(V_K(\widetilde{\pi})))^{1/2}(|K| + C_2).$$

The second inequality here we obtain using that $\sup_{\widetilde{\pi}\in\Sigma_\pi}\mathbf{E}|\widetilde{X}(T,0,X_0,\widetilde{\pi})|^m < +\infty$.

By Proposition 4.3, we have

$$|J(\widetilde{\pi}) - \widetilde{\mathcal{J}}_K(\widetilde{\pi})| \le C_1(\mathbf{P}(V_K(\pi)))^{1/2}(|K| + C_2) \le \frac{C_1(K + C_2)c^{1/2}}{\psi(K)^{1/2}}e^{-c\psi(K)^2/2} \to 0 \quad \text{as} \quad K \to +\infty.$$

Further, assume that $\inf_{t,\omega} |\xi(t,\omega)| = 0$. there exists $K > 0$ such that $|J(\widetilde{\pi}) - \widetilde{\mathcal{J}}_K(\widetilde{\pi})| \le \delta/3$ for any $\pi \in \Sigma_{MFT}(0,T) \cap \Sigma_\pi$. It follows that $J(\widetilde{\pi}) \ge \widetilde{\mathcal{J}}_K(\widetilde{\pi}) - \delta/3$ for any $\widetilde{\pi} \in \Sigma_{MFT}(0,T) \cap \Sigma_\pi$.

By Lemma 4.3, there exists $\widehat{\pi} \in \Sigma_{MFT}(0,T) \cap \Sigma_\pi$ such that $\widetilde{\mathcal{J}}_K(\widehat{\pi}) \ge \widetilde{\mathcal{J}}_K(\pi) - \delta/3$. We have that

$$J(\widehat{\pi}) \ge \widetilde{\mathcal{J}}_K(\widehat{\pi}) - \delta/3 \ge \widetilde{\mathcal{J}}_K(\pi) - 2\delta/3 \ge J(\pi) - \delta,$$

where $C_i > 0$ are independent from $\widetilde{\pi} \in \Sigma_\pi$ Then the proof of Lemma 4.5 follows. □

Proof of Theorem 3.1. It suffices to show that, for any $\delta > 0$ and $\pi \in \Sigma(0,T)$, there exists $\widehat{\pi} \in \Sigma_{MFT}(0,T) \cap \Sigma_\pi$ such that (3.4) hold.

For $K > 0$, let $U^{(K)}(x)$ be defined by (3.1) with U replaced by $\min(U(x), K)$. Let $\widehat{\mathcal{J}}_K(\pi) = \mathbf{E}U^{(K)}(\widetilde{X}(T, 0, X_0, \pi))$.

Let $\delta > 0$ and $\pi \in \Sigma(0,T)$ be given. Clearly, there exists $K > 0$ such that $\widehat{\mathcal{J}}_K(\pi) \ge J(\pi) - \delta/2$. By Lemma 4.3, there exists $\widehat{\pi} \in \Sigma_{MFT}(0,T) \cap \Sigma_\pi$ such that $\widehat{\mathcal{J}}_K(\widehat{\pi}) \ge \widehat{\mathcal{J}}_K(\pi) - \delta/2$. In addition, we have that $J(\widehat{\pi}) \ge \widehat{\mathcal{J}}_K(\widehat{\pi})$ for large enough K. For these K, we have that

$$J(\widehat{\pi}) \ge \widehat{\mathcal{J}}_K(\widehat{\pi}) \ge \widehat{\mathcal{J}}_K(\pi) - \delta/2 \ge J(\pi) - \delta.$$

Then the proof follows. □

5 Discussion and comments

(i) In fact, the model in Lemma 4.2 is quite reasonable itself, since it is natural to assume some stability and predictability of the parameters of the distributions. There are many well developed methods that may help to forecast the market parameters on a small enough horizon $\varepsilon > 0$; in particular, a frequency criterion of predictability on a finite horizon can be found in Dokuchaev (2010b).

(ii) In our setting, we assumed that the admissible strategies are such that if $\theta(t) = 0$ then $\pi(t) = 0$. Without this restriction, the presented version of Mutual Fund Theorem does not necessary hold for the given class of utilities. For instance, consider a convex function $U(x) = x^2$ and $\theta(t) \equiv 0$ (the case that is not excluded). Then the only strategy $\pi \in \Sigma_{MFT}$ is zero. However, this strategy is outperformed by any non-trivial strategy.

(iii) It can be seen from the construction of the suboptimal strategies in the proof that, without some special assumptions about evolution of $\mu(t)$, these strategies cannot be represented as $\pi(t) = f(X(t), S(t), \mu(t), t)$, where f is a deterministic function. This means that dynamic programming method cannot be applied directly to this model.

(iv) Theorem 3.1 represents a weakened version of Mutual Fund Theorem since it states only suboptimality of the strategies from the required class. A stronger version of this theorem is known for many special cases. In particular, there are stronger versions of Lemma 4.1; see, e.g., Khanna and Kulldorff (1999), Dokuchaev and Haussmann (2001), Schachermayer *et al* (2009). Let us explain why these versions of Lemma 4.1 cannot be applied in our proof.

Khanna and Kulldorff (1999) proved that any Markov strategy can be outperformed by a strategy from from a class similar to Σ_{MFT}. Our setting with random parameters requires to cover strategies that are not necessary Markov.

Schachermayer *et al* (2009)) found that the Mutual Fund Theorem holds for a market where claims $F(Z(T))$ can be replicated via trading of a mutual fund with the discounted price $Z(t)$ for deterministic functions $F : \mathbf{R} \to \mathbf{R}$. Here $Z(t)$ is the log-optimal discounted wealth such that

$$dZ(t) = Z(t)\theta(t)^\top \sigma(t)^{-1}\mathbf{S}(t)^{-1}dS(t), \quad Z(0) = 1.$$

In a similar framework, Dokuchaev and Haussmann (2001) found that Mutual Fund Theorem holds when the scalar value $\int_0^T |\theta(t)|^2 dt$ is non-random (Dokuchaev and Haussmann (2001), Lemma 4.1). This condition leads to the replicability of the claims $F(Z(T))$. However, these results cannot replace Lemma 4.1, because they require certain special properties for U. We have to apply this lemma to U replaced by V_m in the proof of Lemma 4.2. If we assume the required properties for U, it is not clear if these properties will be transferred to V_m.

(v) It can be seen from the proofs that, for a general case of random $\mu(t)$, the suboptimal terminal discounted wealth cannot be presented as $F(Z(T))$ for a deterministic function F. Respectively, these cases cannot be covered by the the method based on the replication of these claims (Schachermayer *et al* (2009), Dokuchaev and Haussmann (2001)).

Acknowledgments

This work was supported by ARC grant of Australia DP120100928 to the author.

References

Brennan, M.J. (1998). The role of learning in dynamic portfolio decisions. *European Finance Review* **1**, 295–306.

Dokuchaev, N., and Haussmann, U. (2001). Optimal portfolio selection and compression in an incomplete market. *Quantitative Finance*, **1**, 336-345.

Dokuchaev, N. (2010a). Mean variance and goal achieving portfolio for discrete-time market with currently observable source of correlations. *ESAIM: Control, Optimisation and Calculus of Variations*, in press.

Dokuchaev, N. (2010b). Predictability on finite horizon for processes with exponential decrease of energy on higher frequencies. *Signal processing* **90** Issue 2, 696–701.

Feldman, D. (2007). Incomplete Information Equilibria: Separation Theorems and Other Myths. *Annals of Operations Research* **151**, 119–149.

Karatzas, I., and Shreve, S.E. (1998). *Methods of Mathematical Finance*, Springer-Verlag, New York.

Khanna A.,, Kulldorff M. (1999). A generalization of the mutual fund theorem. *Finance and Stochastics* **3**, 167185 (1999).

Krylov, N.V. (1980). *Controlled diffusion processes.* Shpringer-Verlag.

Li, D., and Ng, W.L. (2000). Optimal portfolio selection: multi-period mean–variance optimization. *Mathematical Finance* **10** (3), 387–406.

Lim, A. (2004). Quadratic hedging and mean-variance portfolio selection with random parameters in an incomplete market. *Mathematics of Operations Research* **29**, iss.1, 132 - 161.

Lim, A. (2005). Mean-variance hedging when there are jumps. *SIAM Journal of Control and Optimization* **44**, 1893–1922.

Lim, A., and Zhou, X.Y. (2002). Mean-variance portfolio selection with random parameters in a complete market. *Mathematics of Operations Research* **27**, iss.1, 101 - 120.

Merton R. (1969). Lifetime portfolio selection under uncertainty: the continuous-time case. *Review of Economics and Statistics* **51**, 247–257.

Revuz, D., and Yor, M., 1999, *Continuous Martingales and Brownian Motion.* Springer-Verlag: New York.

Schachermayer, W., Srbu, M., Taflin, E. (2009). In which financial markets do mutual fund theorems hold true? *Finance and Stochastics* **13**, 49–77.

Shilov G.E, Gurevich B.L (1967). *Integral, measure and derivative: a unified approach.* Nauka, Moscow.