The premium of dynamic trading

CHUN HUNG CHIU† and XUN YU ZHOU*†‡

†Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong
‡Nomura Centre for Mathematical Finance, and Oxford–Man Institute of Quantitative Finance, University of Oxford, 24–29 St Giles, Oxford OX1 3LB, UK

(Received 1 December 2008; in final form 4 January 2010)

It is well established that, in a market with inclusion of a risk-free asset, the single-period mean–variance efficient frontier is a straight line tangent to the risky region, a fact that is the very foundation of the classical CAPM. In this paper, it is shown that, in a continuous-time market where the risky prices are described by Itô processes and the investment opportunity set is deterministic (albeit time-varying), any efficient portfolio must involve allocation to the risk-free asset at any time. As a result, the dynamic mean–variance efficient frontier, although still a straight line, is strictly above the entire risky region. This in turn suggests a positive premium, in terms of the Sharpe ratio of the efficient frontier, arising from dynamic trading. Another implication is that the inclusion of a risk-free asset boosts the Sharpe ratio of the efficient frontier, which again contrasts sharply with the single-period case.

Keywords: Continuous time; Portfolio selection; Mean–variance efficiency; Sharpe ratio

1. Introduction

Given a single investment period and a market where there are a number of basic assets including a risk-free asset, Markowitz’s classical mean–variance theory (Markowitz 1952, 1987, Merton 1972) stipulates that, when the risk-free asset is available, the efficient frontier is a straight line tangent to the risky hyperbola $\mathcal{H}$ (see figure 1). Several fundamentally important conclusions have been drawn from this fact. (1) There is at least one portfolio (called a tangent portfolio or tangent fund), composed of the basic risky assets only, that is efficient. (2) If one uses the Sharpe ratio to measure the reward-to-risk, then the inclusion of a risk-free asset does not increase the highest Sharpe ratio $\mathcal{S}$ he could possibly achieve with the basic risky assets. (3) Any efficiency-seeker needs only to invest in the risk-free asset and the tangent fund with a suitable proportion in accordance with her/his risk taste; this observation is called the mutual fund theorem (Tobin 1958). (4) At demand-supply equilibrium the market portfolio is nothing else than the tangent portfolio, and the expected excess rate of return of any individual asset is linearly related to its beta, a.k.a. the Sharpe–Lintner–Mossin capital asset pricing model (abbreviated CAPM; see Sharpe 1964, Lintner 1965, and Mossin 1966).

Figure 1 visualizes some of the most important results in modern portfolio theory and asset pricing theory. The questions we would like to answer in this paper are the following. What is the corresponding figure for a dynamic market that allows continuous trading? What are the implications if the figure in the continuous-time setting

---

*Corresponding author. Email: zhouxy@maths.ox.ac.uk
§Any portfolio using the basic assets can be mapped onto the mean–standard deviation diagram, a two-dimensional diagram where the mean and standard deviation of the portfolio rate of return are used as the vertical and horizontal axes, respectively. The risky hyperbola is the left boundary of the region—the risky region (see section 2 for a precise definition of the risky region)—on the diagram representing all the possible portfolios formed from the basic risky assets.

*The Sharpe ratio is the ratio between the excess rate of return (over the risk-free rate) and the standard deviation of the return rate of a portfolio. Here the risk-free rate serves as a reference point or benchmark only; a portfolio’s Sharpe ratio is defined regardless of whether or not a risk-free asset is available for inclusion when constructing the portfolio.

||Although in this paper we work within the continuous-time framework, all the discussions and results carry over to the dynamic, discrete-time setting.
becomes different? To address these questions, one needs to solve the dynamic Markowitz problem first. Fortunately, the dynamic extension of the Markowitz model, especially in continuous time, has been studied extensively in recent years (see, e.g., Richardson 1989, Bajeux-Besnainous and Portait 1998, Li and Ng 2000, Zhou and Li 2000, and Lim 2004).

In many of the above-cited works on dynamic Markowitz’s problems, explicit, analytic forms of efficient portfolios have been obtained. In particular, if the investment opportunity set is deterministic (though possibly time-varying) and portfolios are unconstrained, then the efficient frontier is shown by Bajeux-Besnainous and Portait (1998) and Zhou and Li (2000) to remain a straight line in a continuous-time market.† Based on this result together with the explicitly derived efficient portfolios, we are going to show in this paper that the efficient frontier is, indeed, strictly above the risky region, as indicated by figure 2.

Notice that the risky region (the dark area) in figure 2 represents all the portfolios that could continuously rebalance among basic risky assets; therefore, it is a much expanded region than the one corresponding to static (buy-and-hold) portfolios involving no transactions between the initial and terminal times. Bajeux-Besnainous and Portait (1998, p. 83) published a figure showing that the dynamic efficient frontier is strictly above a risky region. The figure there and figure 2 here may appear similar at first glance, yet a fundamental difference between the two is that the risky region of Bajeux-Besnainous and Portait (1998) is the one spanned by buy-and-hold portfolios involving basic risky assets (hereafter referred to as the buy-and-hold risky region to be distinguished from the dynamic risky region we are dealing with in this paper). For reasons explained above, our (dynamic) risky region is much larger than the buy-and-hold risky region for the same continuous-time market (the latter is irrelevant to the dynamic economy anyway); therefore, our result is more powerful and more relevant to the dynamic setting. Indeed, the figure of Bajeux-Besnainous and Portait (1998) is nothing else than the (almost trivial) statement that a dynamic efficient portfolio is strictly better than any buy-and-hold portfolio, whereas our figure implies that a dynamic efficient portfolio is strictly better than any portfolio that is allowed to continuously switch among risky assets.‡

Figure 2 reveals a major and surprising departure of a dynamic economy from a static one. Immediate are the following consequences. (1) No portfolio consisting of only risky assets could be efficient. In other words, any efficient portfolio must invest in the risk-free asset. (2) The efficient frontier line is pushed away from the (dynamic) risky region as a result of the availability of dynamic trading. We call this enhancement of the Sharpe ratio the premium of dynamic trading. (3) The inclusion of the risk-free asset in one’s portfolio indeed strictly increases the best Sharpe ratio achievable compared with the case when s/he only has risky assets at his/her disposal. (4) Efficient portfolios are no longer simple convex combinations of the risk-free asset and a risky fund; therefore, the mutual fund theorem (in the conventional sense) fails.§ (5) The market portfolio is no longer mean–variance efficient even under the supply–demand equilibrium, if the market portfolio is defined analogous

†The efficient frontier in the continuous-time setting is plotted on the mean–standard deviation diagram at the end of the investment horizon. We are interested in the terminal time only because the two criteria of a mean–variance model—mean and variance, that is—concern only the terminal payoff.

‡Only in one special case, i.e. when there is merely one risky asset, do the two figures coincide.

§Bajeux-Besnainous and Portait (1998) derive a so-called strong separation, which asserts that buy-and-hold strategies involving the bond and an appropriate dynamic strategy describe the entire efficient frontier. However, it is not mentioned there whether or not—the answer is no by virtue of our results—that particular dynamic strategy is a pure risky strategy as with a single-period model.
to that in the single-period setting. The CAPM in the present setting needs to be studied more carefully.

What is the cause of such a drastic change in the dynamic setting? An immediate answer might be that there are many more portfolios to choose from because of the possibility of dynamic trading, and hence the admissible region is much expanded than its single-period counterpart. While this is indeed a good point, it has yet to explain the puzzling phenomenon wholly: Why is the efficient frontier completely pushed away from the risky region rather than, say, the whole admissible region as a strict subset of the risky region? We defer the detailed discussions on this to section 4.

In deriving the aforementioned separation result, some properties of a dynamic efficient policy will be revealed, which are interesting in their own right. These properties, unique to the dynamic setting, include the facts that any efficient wealth process is capped by a deterministic upward curve, and the fact that transaction cost and consumption are not considered.

The remainder of the paper is organized as follows. In section 2, the market under consideration is described and the continuous-time mean–variance formulation is presented. In section 3, some known results on mean–variance efficient portfolios and frontier are highlighted. Section 4 is devoted to a detailed study of the strict separation between the frontier and the dynamic risky region, introducing the term of the premium of dynamic trading. Section 5 offers economical explanations on the premium. Section 6 concludes the paper. All proofs are deferred to the appendix.

2. A continuous-time mean–variance market

Throughout this paper ($\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t \in \mathbb{R}}$) is a fixed filtered complete probability space on which is defined a standard $\mathcal{F}_t$-adapted $m$-dimensional Brownian motion $\{W(t), t \geq 0\}$ with $W(0) = (W^1(0), \ldots, W^m(0)) \in \mathcal{W}$ and $W(0) = 0$. It is assumed that $\mathcal{F}_t = \sigma(W(s); s \leq t)$ augmented by all the $\mathcal{P}$-null sets. For a fixed $T > 0$, we denote by $L^2_T(0, T; \mathbb{R}^d)$ the set of all $\mathbb{R}^d$-valued, $\mathcal{F}_T$-progressively measurable stochastic processes $f(\cdot)$ with $E \int_0^T |f(t)|^2 dt < +\infty$. Also, we use $M'$ to denote the transpose of any vector or matrix $M$, and $\sigma_t$ to denote the standard deviation of a random variable $x(t)$.

Suppose there is a capital market in which $m + 1$ basic assets are being traded continuously. One of the assets is a risk-free bond whose value process $S_0(t)$ is subject to the following ordinary differential equation:

$$
\begin{align*}
dS_0(t) &= r(t)S_0(t) \, dt, \quad t \geq 0, \\
S_0(0) &= s_0 > 0
\end{align*}
$$

where $r(t) > 0$ is the interest rate. The other $m$ assets are risky stocks whose price processes $S_i(t), \ldots, S_m(t)$ satisfy the following stochastic differential equation (SDE):

$$
\begin{align*}
dS_i(t) &= S_i(t) \left[ \mu_i(t) \, dt + \sum_{j=1}^m \sigma_{ij}(t) \, dW^j(t) \right], \quad t \geq 0, \\
S_i(0) &= s_i > 0, \quad i = 1, 2, \ldots, m,
\end{align*}
$$

where $\mu_i(t) > 0$ is the appreciation rate, and $\sigma_{ij}(t)$ the volatility or dispersion rate of the stocks. Here the investment opportunity set $\{(r(t), \mu_i(t), \sigma_{ij}(t), t, j = 1, 2, \ldots, m)\}$ is deterministic (yet time-varying).

Define the excess rate of return vector $B(t) := (\mu_1(t) - r(t), \ldots, \mu_m(t) - r(t))^\top$ and the covariance matrix $\sigma(t) := (\sigma_{ij}(t))_{m \times m}$. We assume that $B(t)$ is a continuous function of $t$, and

$$
\sigma(t)\sigma(t) \geq \delta I, \quad \forall t \geq 0,
$$

for some $\delta > 0$. These conditions ensure that the market is arbitrage-free and complete.

Consider an agent, with an initial endowment $x_0 > 0$ and an investment horizon $[0, T]$, whose total wealth at time $t \geq 0$ is denoted by $x(t)$. Assume that the trading of shares is self-financed and takes place continuously, and that transaction cost and consumption are not considered. Then $x(t)$ satisfies (see, e.g., Karatzas and Shreve (1999))

$$
\begin{align*}
dx(t) &= \left[ r(t)x(t) + B(t)^\top \pi(t) \right] \, dt + \pi(t)^\top \sigma(t) \, dW(t), \\
x(0) &= x_0,
\end{align*}
$$

where $\pi(t) := (\pi_1(t), \ldots, \pi_m(t))^\top$, with $\pi_i(t), \quad i = 0, 1, 2, \ldots, m$, denoting the total market value of the agent’s wealth in the $i$th asset (in particular, $\pi_0(t)$ is the wealth invested in the bond at $t$). We call $\pi(\cdot)$ a portfolio of the agent, which is a stochastic process. A portfolio $\pi(\cdot)$ is said to be admissible if $\pi(\cdot) \in L^2_T(0, T; \mathbb{R}^m)$ and the SDE (4) has a unique solution $x(\cdot)$ corresponding to $\pi(\cdot)$. In this case, we refer to $(x(\cdot), \pi(\cdot))$ as an admissible (wealth–portfolio) pair. An admissible portfolio is also interchangeably referred to as an (admissible) asset.

With an admissible portfolio $\pi(\cdot)$ the corresponding wealth trajectory, $x(\cdot)$ is completely determined via the
SDE (4). As a result, \( \pi_0(t) \) of the allocation to the bond, is derived as \( \pi_0(t) = x(t) = \sum_{i=1}^{\infty} \pi^i(t) \). This is the reason why we should not include \( \pi_0(t) \) in defining a portfolio \( \pi(t) \). An admissible portfolio with \( P[\pi_0(t) = 0] = 1 \), a.e. \( t \in [0, T] \) is called a pure risky portfolio.

As with the single-period case a mean–standard deviation diagram (hereafter referred to as the diagram) is a two-dimensional diagram where the mean and standard deviation are used as vertical and horizontal axes, respectively. For any admissible wealth–portfolio pair satisfying (4), define \( R(t) := (x(t) - x_0)/x_0 \), i.e. the corresponding return rate at \( t \). The set of all points \( (\sigma(R(T)), E[R(T)]) \) on the diagram, where \( R(T) \) is the return rate of an admissible portfolio at \( T \), is called the admissible region. The subset of the admissible region corresponding to all the pure risky portfolios is called the (dynamic) risky region.

The agent’s objective is to find an admissible portfolio \( \pi(t) \), among all such admissible portfolios, where their expected terminal wealth \( E[X(T)] = z \), where \( z \in \mathbb{R} \) is given a priori, so that the risk measured by the variance of the terminal wealth, \( \text{Var}(X(T)) := E[(X(T) - E[X(T)])^2] = E[(x(T) - z)^2] \), is minimized. Geometrically, the problem is to locate the left boundary of the admissible region. Mathematically, we have the following formulation.

**Definition 2.1:** Fix the initial wealth \( x_0 \) and the terminal time \( T \). The mean–variance portfolio selection problem is formulated as a constrained stochastic optimization problem parameterized by \( z \geq x_0 e_0^T \sigma(t) dt \):

\[
\begin{align*}
\text{minimize} & \quad J_{MV}(\pi(.)) := E[(x(T) - z)^2], \\
\text{subject to} & \quad x(0) = x_0, \quad E[X(T)] = z, \quad (x(.), \pi(.)) \text{admissible}
\end{align*}
\]

Moreover, the problem is called feasible (with respect to \( z \)) if there is at least one admissible portfolio satisfying \( E[X(T)] = z \). An optimal portfolio to (6), if it ever exists, is called an efficient portfolio with respect to \( z \), and the corresponding point on the diagram is called an efficient point. The set of all the efficient points (with different values of \( z \)) is called the efficient frontier (at \( T \)).

In the preceding definition, the parameter \( z \) is restricted to be no less than \( x_0 e_0^T \sigma(t) dt \), the risk-free terminal payoff. Hence, as standard with the single-period case, we are interested only in the non-satiation portion of the minimum-variance set, or the upper portion of the left boundary of the admissible region.

3. **Efficient portfolios and frontier**

In this section, we highlight some existing results on the explicit solution to the mean–variance portfolio selection problem (6).

Denote the risk premium function

\[
\theta(t) \equiv (\theta_1(t), \ldots, \theta_m(t)) := B(t) (\sigma(t))^{-1}.
\]

The following result gives a complete solution to problem (6).

**Theorem 3.1:** If \( \sum_{i=1}^{\infty} \int_0^T |\mu_i(t) - r(t)| dt \neq 0 \), then problem (6) is feasible for every \( z \geq x_0 e_0^T \sigma(t) dt \). Moreover, the efficient portfolio corresponding to each given \( z \geq x_0 e_0^T \sigma(t) dt \) can be uniquely represented as

\[
\pi^*(t) = (\pi^*_1(t), \ldots, \pi^*_m(t)) = -[\sigma(t)\sigma(t)]^{-1} B(t) x^*(t) T \text{d} \varphi(t)
\]

where \( x^*(t) \) is the corresponding wealth process and

\[
\gamma = \frac{z - x_0 e_0^T \sigma(t) dt}{1 - e^{-\int_0^T \sigma(t)^2 dt}} > 0.
\]

Moreover, the corresponding minimum variance can be expressed as

\[
\text{Var}(X(T)) = \frac{1}{e_0^T \sigma(t) dt} \left[ z - x_0 e_0^T \sigma(t) dt \right] \leq 1.
\]

Define \( R^*(t) := (x^*(t) - x_0)/x_0 \), the return rate of an efficient strategy at time \( t \). Then by virtue of (10) we immediately obtain the efficient frontier.

**Theorem 3.2:** The efficient frontier is

\[
E[R^*(T)] = R_f(T) + \sqrt{e_0^T \sigma(t)^2 dt - 1} \sigma_{R^*(T)},
\]

where \( R_f(T) := e_0^T \sigma(t) dt - 1 \) is the risk-free return rate over the entire horizon \( [0, T] \).

Hence, the efficient frontier is a straight line which is the continuous-time analog to the capital market line of the classical single-period model (see, e.g., Sharpe 1964).

4. **Premium of dynamic trading**

As discussed in the introduction the efficient frontier in the single-period case is tangent to the risky region (see figure 1). Consequently, there is a portfolio consisting of only the risky assets, i.e. the tangent portfolio, whose Sharpe ratio is the same as that of any efficient portfolio. In particular, if there is only one risky asset (e.g., in a Black–Scholes market), then the Sharpe ratios of an efficient portfolio and the risky asset coincide. However, \( \dagger \)

\( \dagger \)The fact that in the continuous-time setting the efficient frontier is still a straight line may seem, at first sight, to be a routine (and, shall we say, boring) extension of the single-period case. However, as will be discussed in the sequel, this fact is not to be taken lightly.
The premium of dynamic trading

119

this is no longer true in the continuous-time setting, as shown by the following theorem.

Theorem 4.1: In a Black–Scholes market where there is one risk-free asset with a short rate \( r > 0 \) and one risky asset with an appreciation rate \( \mu > r \) and a volatility rate \( \sigma > 0 \), the Sharpe ratio of any mean–variance efficient portfolio is always strictly greater than that of the risky asset.

This result was first observed by Richardson (1989) via simulation, and proved by Bajeux-Besnainous and Portait (1998, pp. 87–88). We will supply a proof in the appendix not only for the convenience of the reader, but, more importantly, for understanding why the proof falls apart when there are multiple risky assets and the risky region contains all the continuously traded (instead of buy-and-hold) portfolios.

Theorem 4.1 is exemplified by the following example.

Example 4.2: Consider a Black–Scholes market where there is one risky asset (the stock) with \( \mu = 0.12 \) and \( \sigma = 0.15 \), and a risk-free rate (continuously compounding) \( r = 0.06 \). Assume an investment period \( T = 1 \) (year). Then the risk-free return rate over one year is \( e^{0.06} - 1 = 0.0618 \). According to (12) the Sharpe ratio of the efficient frontier is

\[
\sqrt{e^{\frac{1}{2}(\mu-r)^2 dt}} - 1 = \sqrt{e^{0.12 - 0.06} / 0.15^2} - 1 = 0.4165.
\]

On the other hand, to determine the position of the stock in the diagram we calculate the return rate of the stock

\[
E[R(1)] = e^\mu - 1 = 0.1275,
\]

and its standard deviation

\[
\sigma[R(1)] = e^\mu \sqrt{e^{\delta^2}} - 1 = e^{0.12} \sqrt{e^{0.15^2}} - 1 = 0.1701.
\]

Hence the Sharpe ratio of the stock is

\[
(0.1275 - 0.0618)/0.1701 = 0.3862 < 0.4165.
\]

This means that the efficient frontier lies much above the stock. More precisely, the inclusion of the risk-free security increases the Sharpe ratio by approximately 7.85%.

Theorem 4.1 demonstrates an intriguing difference between the single-period (static) and continuous-time (dynamic) cases, at least for the Black–Scholes market: the efficient frontier is strictly separated from the risky region in the latter case. But the result is really not that surprising if one thinks a little deeper: the phenomenon can be explained by the special structure of the Black–Scholes market, in particular the availability of only one risky asset. The portfolio consisting of the risky asset is essentially a buy-and-hold strategy owing to the lack of another risky asset, therefore it is not taking advantage of dynamic trading as enjoyed by an efficient portfolio which could continuously adjust the weight between the risk-free and risky assets. This is why the risky asset underperforms—in terms of the Sharpe ratio—the efficient portfolio.

In the case of multiple risky assets, the same argument applies, implying that the efficient frontier is strictly separated from the buy-and-hold risky region, as portrayed by figure 1 of Bajeux-Besnainous and Portait (1998).

So, what if now we have multiple risky assets and the risky region is generated by all the dynamically changing risky portfolios? The dynamic risky region is much larger; would there be a chance that it is large enough to touch the efficient frontier?

The answer to the last question is no: We are to establish the strict separation for a general continuous-time market involving multiple stocks and time-varying market parameters. Note that we are no longer able to prove this using direct calculation as in the proof of theorem 4.1 (see appendix), because it is not possible, at least for us, to obtain the expression for the (dynamic) risky region. Instead, we will derive a property of an efficient portfolio which implies the said strict separation. In doing so we need other properties that are interesting in their own right.

Theorem 4.3: Let \( x(t) \) be the wealth process under the efficient portfolio corresponding to \( z \geq x_0 e^{\int_0^T r(s) ds} \). Then

\[
P\left\{ x(t) = x_0 e^{\int_0^T r(s) ds}\right\} = 1.
\]

Moreover, the above inequality is strict if and only if \( z > x_0 e^{\int_0^T r(s) ds} \).

This theorem implies that, with probability 1, the wealth under an efficient strategy must be capped at any time by the present value of \( \gamma \), which is a deterministic constant depending only on the target \( z \). In particular, with probability 1 the terminal wealth will never exceed \( \gamma \). Notice that such a property is unavailable in the single-period case.

Theorem 4.4: Let \( \pi(t) \) be an efficient portfolio corresponding to \( z > x_0 e^{\int_0^T r(s) ds} \). If at \( t_0 \in [0,T] \), \( \mu(t_0) \neq r(t_0) \) for some i, then

\[
P[\pi[t_0] \neq 0] = 1.
\]

This result suggests that any efficient strategy (other than the risk-free one) invests in at least one risky asset whenever any one of the risky appreciation rates is different from the risk-free rate.

Notwithstanding the preceding result, any efficient strategy must also invest in the risk-free asset at any time, as shown by the following theorem.

Theorem 4.5: Let \( \pi(t) \) be an efficient portfolio corresponding to \( z \geq x_0 e^{\int_0^T r(s) ds} \). Then we must have

\[
P[\pi[t_0] \neq 0] > 0, \quad \forall t \in [0,T],
\]

where \( \pi[t_0] \) is the allocation to the bond at \( t \).

We know that the efficient frontier is a straight line which by definition must lie above the risky hyperbola. Now, theorem 4.5 excludes the possibility that the former
intersects with the latter (because any efficient portfolio must invest in the risk-free asset at any time). In other words, the efficient frontier line is strictly separated from the dynamic risky region, and hence must lie strictly above the risky hyperbola, as indicated by figure 2.

We can now formally state the following main result of this paper.

**Theorem 4.6:** The Sharpe ratio of any continuous-time mean–variance efficient portfolio is always strictly greater than that of any admissible portfolio consisting of only the basic risky assets.

One implication of theorem 4.6 is that the availability of dynamic trading helps increase the Sharpe ratio of the efficient frontier compared with static trading. In the case of example 4.2 (although this example is too simple to be a really good one), where all the market data are quite typical, the Sharpe ratio of an efficient frontier is 0.4165 compared with 0.3862 of the stock— an increase of nearly 8% by dynamic trading. We call this ‘the premium of dynamic trading’.

Another more intriguing (albeit delicate) implication of theorem 4.6 is that the availability of a risk-free asset also helps increase the Sharpe ratio of one’s portfolios in the continuous-time setting. This, again, is in sharp contrast with the single-period case. To elaborate, consider first the single-period case. If there is no risk-free asset available, then the portfolio that produces the highest Sharpe ratio is the tangent portfolio. Hence, the availability of an additional risk-free asset does not yield a Sharpe ratio higher than that of the tangent portfolio (see figure 1). However, in the continuous-time setting, the newly discovered strict separation of the efficient frontier from the risky hyperbola suggests that a higher Sharpe ratio is achieved when the risk-free asset is available for trading.

5. Cause of the premium

But what is the fundamental reason underlying such a great difference between the continuous-time case and the single-period case? This is best explained† by first recalling the reason for the tangent line in the single-period case. Consider the risk-free asset and any risky asset, which are two points on the diagram. These two assets are combined to form a portfolio using a weight of \( \alpha \) for the risk-free asset, where \( \alpha \in \mathbb{R} \). It is easy to show (see, e.g., Luenberger 1998, p. 165) that all such portfolios are on the straight line connecting the original two assets (see figure 3). For each risky asset there is such a straight line; thus the admissible region (the one including both risk-free and risky assets) is a triangularly shaped region with the upper boundary being the straight line tangent to the risky hyperbola‡ (see figure 4).

Now, in the continuous-time case, let us also start with two assets, a risk-free asset represented by the portfolio \( \pi_f(\cdot) \equiv 0 \), and a risky asset represented by \( \pi(\cdot) \neq \pi_f(\cdot) \). We use these two assets to form new portfolios of the form \( \pi_{\alpha}(t) = [1 - \alpha(t)]\pi_f(t) + \alpha(t)\pi(t) \equiv \alpha(t)\pi(t) \), where \( \alpha(\cdot) \) is any progressively measurable process so long as \( \pi_{\alpha}(\cdot) \) is admissible. Let \( x(\cdot) \) be the wealth process corresponding to \( \pi(\cdot) \). If \( \alpha(\cdot) \equiv \alpha \in \mathbb{R} \), that is the corresponding \( \pi_{\alpha}(\cdot) \) is a buy and hold combination of the two, then it is immediate from the linearity of the wealth equation (4) that \( \alpha x(t) \) is the wealth process corresponding to \( \pi_{\alpha}(\cdot) \) whose terminal \((\sigma, \tilde{r})\) value lies on the straight line connecting those of the two original assets. However, in general, \( \alpha(\cdot) \) can be any appropriate stochastic process (in other words, the new portfolio \( \pi_{\alpha}(\cdot) \) is, in general, a dynamic combination of the two assets); therefore, \( \alpha(\cdot)x(\cdot) \) is no longer necessarily the wealth process under \( \pi_{\alpha}(\cdot) \) or it may not lie on the straight line connecting the two original assets. As a consequence, all the new portfolios (with all the possible processes \( \alpha(\cdot) \))

†The discussions in this section apply, mutatis mutandis, also to a dynamic, discrete-time model.
‡This also suggests that the efficient frontier should be a straight line.
formed from the original two assets generate a much larger region as indicated in figure 5.

To discuss the shape of the admissible region in the continuous-time case (to be precise we are only interested in the upper left boundary of the admissible region), we first construct the dynamic risky region defined by the risky assets only, which is the shaded region in figure 6. Next, for each portfolio in this region we make combinations with the risk-free asset. As discussed above these new portfolios form a solid two-dimensional region with its upper left boundary being a curved line, in general (which could be a straight line in some special cases—such as the Black–Scholes case). There is such a curved line corresponding to every asset in the risky region. The envelope of these lines forms the upper left boundary of the entire admissible region (containing both the risk-free and risky assets), which is the efficient frontier we are seeking. It is these curved lines that push the efficient frontier away from the risky region, and hence the surprising phenomenon stipulated by theorem 4.5.

Last, but not least, the fact that in continuous time the mean–variance efficient frontier is still a straight line, derived by Bajeux-Besnainous and Portait (1998) and Zhou and Li (2000), is no longer a mere routine extension of its single-period counterpart, and should not be taken lightly. In fact, the efficient frontier is the envelope of infinitely many curved lines, which turn out to be a straight line. This is quite an unusual coincidence.

6. Concluding remarks

This paper has disclosed some rather unexpected phenomena associated with a continuous-time mean–variance market, suggesting that continuous-time financial models have more complex, sometimes strikingly different, structures and properties than its single-period counterpart.

One should appreciate that continuous time is probably a closer representation of the real-world investment today than its discrete-time counterpart, not to mention its analytical tractability which enables us to elicit important economic insights from results that are often explicit. On the other hand, in the realm of continuous-time asset allocation and asset pricing the literature has been dominated by the expected utility maximization (EUM) models. However, “few if any agents know their utility functions; nor do the functions which financial engineers and financial economists find analytically convenient necessarily represent a particular agent’s attitude towards risk and return” (Markowitz and Zhou 2004). It is fair to say that mean–variance (including the related Sharpe ratio) remains nowadays one of the most commonly used measures to assess the performance of fund managers. Theoretically, while mean–variance has some inherent drawbacks such as the sensitivity of the final solutions to the investment opportunity set, and the inconsistency with the dynamic programming principle in the continuous-time setting, recent studies have revealed some redeeming qualities of continuous-time mean–variance efficient policies. For example, it is shown by Li and Zhou (2006) that, under the same setting as in this paper, a mean–variance efficient portfolio realizes the (discounted) targeted return on or before the terminal date with a probability greater than 0.8072. This number is universal irrespective of the opportunity set, the targeted return, and the investment horizon. This, together with the new findings in this paper, suggests that it is necessary and important to revisit mean–variance models, in terms of both asset allocation and asset pricing, for dynamic markets.

In this paper there are several assumptions on the underlying model, such as free of transaction cost, market completeness, and deterministic investment opportunity set. Mean–variance models with proportional transaction costs have recently been studied by Dai et al. (2010). It would therefore be interesting to extend the results of this paper to the case with transaction costs. On the other hand, we believe that our results can be extended to the 
incomplete market case, noting the results of Lim (2004) on the mean–variance model in an incomplete market. The case with stochastic investment opportunity, however, is still open, as theorems 4.3–4.5 may no longer hold true.

Acknowledgements

This research was supported by RGC Earmarked grants CUHK 4175/03E and CUHK 418605, the Croucher Senior Research Fellowship, and an Oxford start-up fund. We are grateful to Harry Markowitz for sending us the book Markowitz (1987) and alerting us to the results on possible inefficiency of the market in the single-period setting. This paper benefited from comments of the participants at various conferences where previous versions of this paper were presented, and from the comments of Hanqing Jin, Dilip Madan, Stan Pliska, and Jianming Xia.

References

Bajjouk-Benaimous, L. and Portait, R., Dynamic asset allocation in a mean–variance framework. Mgmt Sci., 1998, 44, 79–95.
Dai, M., Xu, Z. and Zhou, X.Y., Continuous-time mean–variance portfolio selection with proportional transaction costs. SIAM J. Financial Math., 2010, 1, 96–125.
Karatzas, I. and Shreve, S.E., Methods of Mathematical Finance, 1999 (Springer: New York).
Li, D. and Ng, W.L., Optimal dynamic portfolio selection: Multi-period mean–variance formulation. Math. Finance, 2000, 10, 387–406.
Li, X. and Zhou, X.Y., Continuous-time mean–variance efficiency: The 80% rule. Ann. Appl. Probab., 2006, 16, 1751–1763.
Lim, A.E.B., Quadratic hedging and mean–variance portfolio selection with random parameters in an incomplete market. Math. Oper. Res., 2004, 29, 132–161.
Lim, A.E.B. and Zhou, X.Y., Mean–variance portfolio selection with random parameters in a complete market. Math. Oper. Res., 2002, 27, 101–120.
Lintner, J., The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Rev. Econ. Statist., 1965, 47, 13–37.
Luenger, D.G., Investment Science, 1998 (Oxford University Press: New York).
Markowitz, H., Portfolio selection. J. Finance, 1952, 7, 77–91.
Markowitz, H., Mean–Variance Analysis in Portfolio Choice and Capital Markets, 1987 (Frank J. Fabozzi Associates: New Hope, PA).
Markowitz, H. and Zhou, X.Y., Private communication, 2004.
Merton, R.C., An analytic derivation of the efficient frontier. J. Financial Quant. Anal., 1972, 7, 1851–1872.
Mossin, J., Equilibrium in a capital asset market. Econometrica, 1966, 34, 768–783.
Richardson, H.R., A minimum variance result in continuous trading portfolio optimization. Mgmt Sci., 1989, 25, 1045–1055.
Sharpe, W.F., Capital asset prices: A theory of market equilibrium under conditions of risk. J. Finance, 1964, 19, 425–442.
Tobin, J., Liquidity preference as behavior towards risk. Rev. Econ. Stud., 1958, 25, 65–86.

Zhou, X.Y. and Li, D., Continuous-time mean–variance portfolio selection: A stochastic LQ framework. Appl. Math. Optimiz., 2000, 42, 19–33.

Appendix A

Proof (proof of theorem 3.1): The assertion concerning the feasibility of the problem is an immediate consequence of Lim and Zhou (2002, corollary 5.1). The efficient feedback strategies (8) and expression (10) were derived by Zhou and Li (2000, equation (5.12) and theorem 6.1). Finally, that γ > 0 is seen from the facts that x0 ≥ 0, z ≥ x0 e∫tθ(s) ds and |θ(i)| > 0 (due to B(i) ≠ 0) a.e. t ∈ [0, T]. □

Proof (proof of theorem 4.1): To prove this theorem we need the following technical lemma.

Lemma A.1: For any b > 0

\[
(e^{bx} - 1)(e^{b/x} - 1) \geq (e^b - 1)^2, \quad \forall x > 0. \tag{A1}
\]

The case when x = 1 is trivial. Hence due to symmetry we need only show (A1) for any 0 < x < 1.

Let f(x) = (e^{bx} - 1)(e^{b/x} - 1), 0 < x ≤ 1. Then

\[
f'(x) = be^{b(x) + b/x} \left(1 - \frac{1}{x^2} + \frac{1}{x^2} e^{-bx} - e^{-b/x}\right).
\]

Denote

\[
g(x) = 1 - \frac{1}{x^2} + \frac{1}{x^2} e^{-bx} - e^{-b/x}, \quad 0 < x ≤ 1.
\]

We have

\[
g'(x) = \left(2 - 2e^{-bx} - bxe^{-bx} - bxe^{-b/x}\right)
\]

\[
\geq \frac{2}{x^3}[\frac{1}{x} - (1 + bx)e^{-bx}] > 0, \quad \forall 0 < x ≤ 1.
\]

So g(x) < g(1) = 0 ∀0 < x < 1, or f'(x) < 0 ∀0 < x < 1. This leads to f(x) ≥ f(1) = (e^b - 1)^2. □

Proof: We now prove theorem 4.1, which is to show that

\[
\frac{E[R^*(T)] - R_f(T)}{\sigma_R(T)} > \frac{E[R(T)] - R_f(T)}{\sigma_R(T)}, \tag{A2}
\]

where R^*(T) is the rate of return of any efficient portfolio, R_f(T) is the risk-free rate of return, and R(T) is the rate of return of the risky asset, over [0, T]. Note E[R^*(T)] = e^{\mu_T} - 1, \quad \sigma^2_R(T) = e^{(\mu_T + \sigma_T^2)T} - e^{2\mu_T}, \quad \text{and} \quad R_f(T) = e^{\mu_T} - 1. Therefore, according to equation (12), equation (A2) is equivalent to

\[
e^{(\mu - \gamma)T} - 1 > \frac{(e^{\mu_T} - e^{\gamma T})^2}{e^{2\mu_T + \sigma_T^2 T} - e^{2\mu_T}}. \tag{A3}
\]

Now,

\[
e^{(\mu - \gamma)T} - 1 - \frac{(e^{\mu_T} - e^{\gamma T})^2}{e^{2\mu_T + \sigma_T^2 T} - e^{2\mu_T}}
\]

\[
= \frac{e^{2\mu_T}(e^{\gamma T} - 1)(e^{(\mu - \gamma)T} - 1) - e^{2\gamma T}(e^{(\mu - \gamma)T} - 1)^2}{e^{2\mu_T + \sigma_T^2 T} - e^{2\mu_T}}.
\]
The premium of dynamic trading

Considering the numerator of the above and noting \( \mu > r \), we have

\[
e^{2\mu T}(e^{\sigma^2 T} - (e^{(\mu-r)T} - 1)^2) > e^{2\mu T}(e^{\sigma^2 T} - (e^{(\mu-r)T} - 1)^2)
\]

Letting \((\mu - r)T = b\) and \(\sigma^2 = (\mu - r)x\) where \(x > 0\), we have

\[
= (e^{bx} - 1)(e^{b/x} - 1) - (e^b - 1)^2 \geq 0,
\]

by virtue of lemma A.1. The proof is complete. \(\square\)

**Proof (proof of theorem 4.3):** Set \(y(t) := x^*(t) - \gamma e^{-\int_0^t \sigma(s)ds} \). Using the wealth equation (4) that \(x^*(\cdot)\) satisfies and applying (8), we deduce

\[
dy(t) = [r(t) - \theta(t)\sigma(t)]y(t) dt - \theta(t) y(t) dW(t),
\]

\[
y(0) = x_0 - e^{-\int_0^T \theta(s) dt} \leq 0.
\]

The above equation has a unique solution,

\[
y(t) = y(0) \exp \left\{ \int_0^t \left[ r(s) - \frac{3}{2} \theta(s)^2 \right] ds - \int_0^t \theta(s) dW(s) \right\} \leq 0,
\]

and the inequality is strict if and only if \(y(0) < 0\), or \(z > x_0 e^{\int_0^T \theta(s) ds} \).

**Proof (proof of theorem 4.4):** Since \(x^*(t) \neq \gamma e^{-\int_0^t \sigma(s)ds} \), the result follows directly from theorem 4.3 and (8). \(\square\)

**Proof (proof of theorem 4.5):** If \(z = x_0 e^{\int_0^T \theta(s) ds} \), then the theorem holds trivially. So we assume that \(z > x_0 e^{\int_0^T \theta(s) ds} \).

According to theorem 3.1, \(\pi^*(\cdot)\) can be written as

\[
\pi^*(t) = (\pi^*_1(t), \ldots, \pi^*_m(t))^T = - [\sigma(t) \sigma(t)']^{-1} B(t) \left[ x^*(t) - \gamma e^{-\int_0^t \sigma(s) ds} \right], \quad (A4)
\]

for some \(\gamma > 0\). If (17) is not true, then there is \(\tilde{t} \in (0, T)\) such that \(P[\pi^*_1(\tilde{t}) = 0] = 1\), or

\[
P\left\{ x^*(\tilde{t}) = \sum_{i=1}^m \pi^*_i(\tilde{t}) \right\} = 1. \quad (A5)
\]

Summing all the components of \(\pi^*(t)\) in (A4) at \(t = \tilde{t}\), we have

\[
x^*(\tilde{t}) = \sum_{i=1}^m \pi^*_i(\tilde{t}) = - \eta(\tilde{t}) \left[ x^*(\tilde{t}) - \gamma e^{-\int_{\tilde{t}}^T \sigma(s) ds} \right], \quad (A6)
\]

where \(\eta(t) := e^{[\sigma(t) \sigma(t)]^{-1} B(t)}\) with \(e = (1, 1, \ldots, 1)'\). Note that \(\eta(t) \neq -1 \quad \forall t \in [0, T]\), otherwise (A6) would yield \(y = 0\), contradicting (9). Hence, it follows from (A6) that

\[
x^*(\tilde{t}) = \frac{\gamma \eta(\tilde{t}) e^{-\int_{\tilde{t}}^T \sigma(s) ds}}{1 + \eta(\tilde{t})}. \quad (A7)
\]

In other words, the wealth at time \(\tilde{t}\) is a deterministic quantity.

However, the wealth equation (4) that \(x^*(\cdot)\) satisfies up to time \(\tilde{t}\) can be rewritten as

\[
dx^*(t) = \left[ r(t)x^*(t) + \theta(t)\sigma(t)x^*(t) \right] dt + \left( \sigma(t) \sigma(t) x^*(t) \right)' dW(t), \quad t \in [0, \tilde{t}],
\]

\[
x^*(\tilde{t}) = \frac{\gamma \eta(\tilde{t}) e^{-\int_{\tilde{t}}^T \sigma(s) ds}}{1 + \eta(\tilde{t})}. \quad (A8)
\]

The above is a linear backward stochastic differential equation (BSDE) with deterministic linear coefficients as well as a deterministic terminal condition. Hence by the uniqueness of its solution we must have \(\sigma(t) \pi^*(t) = 0 \) a.s., a.e. \( t \in [0, \tilde{t}] \). Appealing to (3), we conclude that \(\pi^*(t) = 0 \) a.s., a.e. \( t \in [0, \tilde{t}] \). This in turn implies \( B(t) = 0 \quad \forall t \in [0, \tilde{t}] \), in view of (A4), theorem 4.3, and the fact that \( B(t) \) is continuous in \( t \). Therefore, \(\eta(t) = 0\), \( \forall t \in [0, \tilde{t}] \), and hence \(x^*(\tilde{t}) = 0\). Again by the uniqueness of the solution to BSDE (A8), \(x_0 = 0\), which is a contradiction. \(\square\)