Rivista di matematica per le scienze economiche e sociali 20 (2) 1997, 133 – 151

ON OPTIMAL INVESTMENT STRATEGIES

HANS U. GERBER
École des Hautes Études Commerciales
Université de Lausanne

ELIAS S. W. SHIU
Department of Statistics and Actuarial Science
University of Iowa

Versione definitiva pervenuta il 16/02/99

Suppose an investor has a fixed decision horizon and an appropriate utility function for measuring his or her utility of wealth. If there are only two investment vehicles, a risky and a risk-free asset, then the optimal investment strategy is such that, at any time, the amount invested in the risky asset must be the product of his or her "current risk tolerance" and the risk premium on the risky asset, divided by the square of the diffusion coefficient of the risky asset. In the case of more than one risky asset, the optimal investment strategy is similar, with the ratios of the amounts invested in the different risky assets being constant over time.

KEYWORDS: Dynamic asset allocation, optimal capital growth, utility functions, risk tolerance, Esscher transform.

1. Introduction

This paper is about optimal investment strategies. A classical investment strategy is based on the Markowitz mean-variance analysis. Unfortunately, the Markowitz model is a one-period model. In contrast, investing (for example for retirement) may well be a long-term, multi-period problem. As time passes, the investor can and will re-allocate assets or rebalance his or her investment portfolio. As Luenberger (1998, p. 417) has noted, "conclusions about multi-period investment situations are not mere variations of single-period conclusions — rather they often *reverse* those earlier conclusions. This makes the subject exciting, both intellectually and in practice. Once the subtleties of multiperiod investment are understood, the reward in terms of enhanced investment performance can be substantial". Other insightful remarks can be found in articles by the Nobel laureate Paul Samuelson (1990, 1997).

This paper treats the multiperiod, continuous-time investment problem, where re-allocation or re-balancing of assets takes place dynamically. For a given decision horizon and a given criterion to assess the utility of wealth, what is the optimal dynamic investment strategy? Expositions of this problem can be found in advanced textbooks and research monographs such as Bicksler and Samuelson (1974, Part VI), Björk (1998), Duffie (1996, Chapter 9), Karatzas and Shreve (1998), Korn (1997), Malliaris and Brock (1982, Chapter 4), Merton (1990), Pliska (1997, Chapter 5), Sethi (1996) and Ziemba and Vickson (1975). Some early papers on this subject are Mossin (1968), Samuelson (1969) and Merton (1969), and recent surveys can be found in Hakansson (1987), Constantinides and Mallarias (1995), and Hakansson and Ziemba (1995). Several authors treat the problem of optimal investment simultaneously with the problem of optimal consumption. In the context of a saving process, the first problem appears to be much more important. Including the second problem may add unnecessary complications and require advanced mathematics such as dynamic programming in continuous time (stochastic control theory) and the Hamilton-Jacobi-Bellman (HJB) equation.

This paper gives a largely self-contained exposition on optimal capital growth and dynamic asset allocation. The bases are laid in Sections 2 to 5; we consider a one-period model, where arbitrary random payments, due at the terminal time (decision horizon), can be traded at time 0. For a given utility function, the investor seeks to maximize the expected utility of terminal wealth (wealth at decision horizon). To understand the optimal decision, it is judicious to introduce the risk tolerance function associated with the utility function, and the implied utility, that is the maximal expected utility of terminal wealth, considered as a function of the current wealth. Some very explicit results are obtained for utility functions with a linear risk tolerance function. The solution of the dynamic asset allocation problem is presented in Sections 6 and 7. In Section 6, the investment vehicles are riskless asset and a single risky asset. In Section 7, the universe of investment vehicles is enlarged to the more realistic case of more than one risky asset. The optimal dynamic investment strategy can be found with a simple idea: consider the self-financing portfolio that replicates the optimal terminal wealth. The key results can be formulated in terms of the risk-neutral Esscher parameter (a term coined to honor the Swedish actuary F. Esscher (1932)) and in terms of the elasticity, with respect to the current wealth, of the expected marginal utility of optimal terminal wealth.

2. Securities Market Models with Price Density

Consider an investor who is to invest an amount w in the securities market at time 0. The investor's planning horizon is T, a fixed positive number. There is no withdrawal or addition of funds between time 0 and time T. The goal is to optimize, in a certain sense, the value of the investor's investment holdings at time T. All discussions in this paper are restricted to the time interval $[0, T]$.

The securities market model under consideration is one in which the securities or assets are priced by a *price density* Ψ. The price density Ψ is a positive random variable with expectation $E[\Psi] = 1$. For a contingent payment Y due at time T, its price, known at time 0 and paid at time T, is the expectation $E[Y\Psi]$. In other words, the quantity $E[Y\Psi]$ is the time-0 *forward price* for the random payment Y due at time T. In this paper, we simply assume that the risk-free force of interest is constant through time and we denote it as r. Hence the time-0 price of Y is $e^{-rT}E[Y\Psi]$.

2.1. *A Single Risky Asset*

For a classical example of Ψ, consider a securities market model consisting of a risk-free asset (or risk-free bond) whose value accumulates at the risk-free rate r, a risky asset which pays no dividends between times 0 and T, and their derivative securities. (The assumption that the risky asset pays no dividends is not critical, because one can consider all dividends being reinvested in the asset as soon as they are received.) For $0 \le t \le T$, let the price of a unit of the risky asset at time t be denoted as $S(t)$. Assume that $\{S(t)\}$ is a *geometric Brownian motion*, i.e., it satisfies the stochastic differential equation

$$dS(t) = \mu S(t)dt + \sigma S(t)dZ(t), \tag{2.1}$$

where μ and σ are constants and $\{Z(t)\}$ is a standard *Wiener process* (or Brownian motion). As pointed out in Gerber and Shiu (1994, 1996), there is a unique number h^*, called the *risk-neutral Esscher parameter*, such that

$$\Psi = \frac{S(T)^{h^*}}{E\left[S(T)^{h^*}\right]} \tag{2.2}$$

is the price density of the securities market; see also Section 10.9 of Panjer *et al.* (1998).

Let

$$X(t) = ln[S(t)/S(0)] \tag{2.3}$$

be the continuously compounded rate of return over the time interval $[0, t]$. Then (2.2) can be rewritten as

$$\Psi = \frac{e^{h^* X(T)}}{E\left[e^{h^* X(T)}\right]}. \tag{2.4}$$

Differentiating (2.3) with respect to t using Itô's Lemma and applying (2.1) yields

$$dX(t) = (\mu - \sigma^2/2)dt + \sigma dZ(t), \tag{2.5}$$

or

$$X(t) = (\mu - \sigma^2/2)t + \sigma Z(t). \tag{2.6}$$

Considering $S(T)$ as a random payment, one sees that

$$S(0) = e^{-rT}E[S(T)\Psi], \tag{2.7}$$

or

$$1 = e^{-rT}E\left[e^{(1+h^*)X(T)}\right] / E\left[e^{h^* X(T)}\right]. \tag{2.8}$$

It follows from (2.8) and (2.6) (with $t = T$) that

$$h^* = (r - \mu)/\sigma^2. \tag{2.9}$$

We shall return to this model in Section 6 and show that the *optimal* investment strategy is one in which the amount invested in the risky asset divided by the investor's "current risk tolerance' is always kept at the constant level $-h^* = (\mu - r)/\sigma^2$. In finance, the quantity $-h^*\sigma = (\mu - r)/\sigma$ is usually called the "market price of risk" (Baxter and Rennie, 1996, p. 119); it is related to the Sharpe Ratio or the Sharpe Index (Sharpe, 1994; Luenberger, 1998, p. 187).

2.2. *Multiple Risky Assets*

For a more realistic securities market model, we consider one consisting of a risk-free asset accumulating at rate r, n risky assets which pay no dividends between times 0 and T, and their derivative securities. For $0 \le t \le T$ and $j = 1, 2, \ldots, n$, let $S_j(t)$ denote the price of the j-th risky asset at time t. Following the usual approach in the finance literature, assume that the dynamics of the asset prices are given by the following system of n stochastic differential equations

$$dS_j(t) = S_j(t)\left[\mu_j dt + \sum_{k=1}^{n} \sigma_{jk} dZ_k(t)\right], \qquad j = 1, 2, \ldots, n, \tag{2.10}$$

where μ_j and σ_{jk} are constants and $\{Z_1(t)\}, \{Z_2(t)\}, \ldots, \{Z_n(t)\}$ are n independent standard Wiener processes. Under the assumption that the matrix

$$C = (\sigma_{jk} \tag{2.11}$$

is nonsingular, there exist n unique numbers, $h_1^*, h_2^*, \ldots, h_n^*$, called risk-neutral Esscher parameters, such that

$$\Psi = \frac{S_1(T)^{h_1^*} S_2(t)^{h_2^*} \cdots S_n(T)^{h_n^*}}{E\left[S_1(T)^{h_1^*} S_2(T)^{h_2^*} \cdots S_n(T)^{h_n^*}\right]} \tag{2.12}$$

is the price density of the securities market. Gerber and Shiu (1994, Section 7) determine the n parameters using the n equations.

$$S_j(0) = e^{-rT} E[S_j(T)\Psi], \qquad j = 1, 2, \ldots, n. \tag{2.13}$$

Let

$$X_j(t) = ln[S_j(t)/S_j(0)], \qquad j = 1, 2, \ldots, n. \tag{2.14}$$

Then

$$dX_j(t) = \left[\mu_j - \frac{1}{2}\sum_{k=1}^{n} \sigma_{jk}^2\right] dt + \sum_{k=1}^{n} \sigma_{jk} dZ_k(t), \tag{2.15}$$

generalizing (2.5). Equation (2.15) also follows from (2.5) by identifying the Brownian motion $\sum_{k=1}^{n} \sigma_{jk} Z_k(t)$ in (2.15) with the Brownian motion $\sigma Z(t)$ in (2.1). Because

$$\begin{aligned}\mathrm{Cov}(X_i(t), X_j(t)) &= \mathrm{Cov}\left(\sum_{k=1}^{n} \sigma_{ik} dZ_k(t), \sum_{m=1}^{n} \sigma_{jm} dZ_m(t)\right) \\ &= t \sum_{k=1}^{n} \sigma_{ik}\sigma_{jk},\end{aligned}$$

the covariance matrix for $\{X_j(t)\}$ is $t\mathbf{C}\mathbf{C}^T$. For a simple formula of the risk-neutral Esscher parameters $\{\mathbf{h}_j^*\}$, define the column vector

$$\mathbf{h}^* = (h_1^*, h_2^*, \ldots, h_n^*)^T, \tag{2.16}$$

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)^T \tag{2.17}$$

and

$$\mathbf{1} = (1, 1, \ldots, 1)^T. \tag{2.18}$$

Then

$$\mathbf{h}^* = (\mathbf{C}\mathbf{C}^T)^{-1}(r\mathbf{1} - \boldsymbol{\mu}), \tag{2.19}$$

which generalizes (2.9).

In the finance literature (Baxter and Rennie, 1996, p. 188; Björk, 1998; Karatzas and Shreve, 1998), the elements of the vector

$$\mathbf{m} = \mathbf{C}^{-1}(\boldsymbol{\mu} - r\mathbf{1}) \tag{2.20}$$

are called market prices of risk. Hence the risk-neutral Esscher parameters and market prices of risk are related by formula

$$\mathbf{h}^* = -(\mathbf{C}^T)^{-1}\mathbf{m}. \tag{2.21}$$

3. Characterizing the Optimal Terminal Wealth

An investor, who has the amount w to invest at time 0 and buys the time-T random payment Y in the securities market, will have

$$W(Y) = we^{rT} + Y - E[Y\Psi] \tag{3.1}$$

at time T. Note that, for each Y,

$$E[W(Y)\Psi] = we^{rT}. \tag{3.2}$$

The term "terminal wealth" is used to describe the random variable $W(Y)$, because T is the decision horizon for the investment problem and "wealth" is a standard term in the financial literature. There are other appropriate substitutes for "wealth", such as "accumulation value", "accumulation" and "fortune".

Assume the investor uses a risk-averse *utility function* $u(x)$ for determining the *optimal* terminal wealth. That is, there is a function $u(x)$, with $u'(x) > 0$ and $u''(x) < 0$, for which the terminal wealth $W(Y)$ that maximizes the expectation $E[u(W(Y))]$ is to be determined. Furthermore, assume that the decreasing function $u'(x)$ varies between $+\infty$ and 0, as x varies in the domain of definition; this property is satisfied by many important utility functions. A recent survey on utility theory can be found in Gerber and Pafumi (1998).

Let W_T denote the optimal terminal wealth $W(Y)$. It follows from (3.2) that

$$E[W_T\Psi] = we^{rT}. \tag{3.3}$$

Subject to (3.3), the random variable W_T can be characterized by

$$u'(W_T) = E[u'(W_T)]\Psi. \tag{3.4}$$

To derive (3.4), consider the Lagrangian

$$L(W, \lambda) = E[u(W)] - \lambda\{E[W\Psi] - we^{rT}\}, \tag{3.5}$$

where λ denotes the Lagrange multiplier and W the wealth random variable. The "gradient" of L with respect to W is

$$u'(W) - \lambda\Psi$$

(an explanation can be found in Deprez and Gerber (1985)). The optimality condition is that the gradient vanishes. Hence

$$u'(W_T) = \hat{\lambda}\Psi, \tag{3.6}$$

where $\hat{\lambda}$ is the optimal value of the Lagrange multiplier. Taking expectations yields

$$E[u'(W_T)] = \hat{\lambda} \tag{3.7}$$

Thus the optimal Lagrange multiplier is the expected marginal utility of optimal terminal wealth. Substitution of (3.7) in (3.6) yields (3.4). For another derivation of (3.4), see Gerber and Shiu (1998).

Because $u'' < 0$, the marginal utility function u' has an inverse function. Let it be denoted as v, i.e., $v(u'(x)) = x$. (In Section 8.10 of Panjer *et al.* (1998), the function v is denoted as h.) Applying the function v to both sides of (3.6) gives

$$W_T = v(\hat{\lambda}\Psi). \tag{3.8}$$

By (3.3),

$$E[\Psi v(\hat{\lambda}\Psi)] = we^{rT} \tag{3.9}$$

This formula establishes the functional relationship between $\hat{\lambda}$ and w. The next section presents some important examples of the optimal terminal wealth random variable W_T. In each of these examples, (3.9) leads to an explicit expression for λ, which can then be substituted in (3.8).

4. LRT Utility Functions

For a twice-differentiable utility $u(x)$, the function

$$\tau(x) = -u'(x)/u''(x) \tag{4.1}$$

is called the *risk tolerance* function (Panjer *et al.* 1998, p. 161). The assumption that u is a risk-averse utility function ($u' > 0$ and $u'' < 0$) implies that $\tau(x)$ is strictly positive (in the domain of definition of u). Throughout this paper, we shall illustrate the theory with *linear risk tolerance* (LRT) utility functions. If τ is a constant or a linear function of x (restricted to the domain of definition of u), then the utility function u is a member of the LRT class. Because the reciprocal of the risk tolerance function is called the (Arrow-Pratt) absolute risk aversion (ARA) function, the LRT class of utility functions is also called the *hyperbolic absolute risk aversion* (HARA) class of utility functions.

The LRT utility functions can be classified into three subclasses, depending on whether $\tau(x)$ is constant, linear and decreasing, or linear and increasing. Two utility functions are said to be *equivalent* if they have the same risk tolerance function. In other words, two utility functions are equivalent if one is an affine function of the other. In each of the next three subsections, a formula is given for the utility functions (up to equivalence) of a subclass, followed by a formula for the optimal terminal wealth W_T.

4.1. *Constant Risk Tolerance Utility Functions*

These are the exponential function with parameter $a > 0$, $1/a$ being the constant level of risk tolerance. Up to equivalence,

$$u(x) = -e^{-ax}/a, \qquad -\infty < x < \infty, \tag{4.2}$$

$$u'(x) = e^{-ax}, \qquad -\infty < x < \infty, \tag{4.3}$$

and

$$v(x) = (-1/a)ln(x), \quad -\infty < x < \infty. \tag{4.4}$$

The risk tolerance function is

$$\tau(x) = 1/a, \quad -\infty < x < \infty. \tag{4.5}$$

Here, (3.8) and (3.9) become

$$W_T = -[ln(\lambda) + ln(\Psi)]/a \tag{4.6}$$

and

$$-\{ln(\lambda) + E[\Psi ln(\Psi)]\}/a = we^{rT}, \tag{4.7}$$

respectively. By eliminating the $ln(\hat{\lambda})$ term,

$$W_T = we^{rT} + \{E[\Psi ln(\Psi)] - ln(\Psi)\}/a. \quad (4.8)$$

4.2. *Decreasing LRT Utility Functions*

These are the power utility functions with parameters c and $s, c > 0$, where c is power of the marginal utility ($-1/c$ is the slope of the risk tolerance function) and s is the finite level of saturation or maximal satisfaction. Up to equivalence,

$$u(x) = -(s-x)^{c+1}/(c+1), \qquad x < s, \quad (4.9)$$

$$u'(x) = (s-x)^c, \qquad x < s, \quad (4.10)$$

and

$$v(x) = s - x^{1/c}, \qquad x < s. \quad (4.11)$$

The risk tolerance function is

$$\tau(x) = (s-x)/c, \qquad x < s. \quad (4.12)$$

Here, (3.8) and (3.9) become

$$W_T = s - \hat{\lambda}^{1/c}\Psi^{1/c} \quad (4.13)$$

and

$$s - \hat{\lambda}^{1/c}E[\Psi^{1+1/c}] = we^{rT}, \quad (4.14)$$

respectively. Hence

$$W_T = s - \frac{s - we^{rT}}{E[\Psi^{1+1/c}]}\Psi^{1/c}, \qquad w < se^{-rT}. \quad (4.15)$$

Note that, if $w \geq e^{-rT}s$, the investment problem has a trivial solution; the investor can achieve maximal satisfaction simply by investing all funds in the risk-free asset.

4.3. *Increasing LRT Utility Functions*

These are the power utility functions with parameters c and $s, c > 0$, where $-c$ is the power of the marginal utility ($1/c$ is the slope of the risk tolerance function) and s is the minimal requirement of the terminal wealth. Up to equivalence and for $c \neq 1$,

$$u(x) = (x-s)^{1-c}/(1-c), \qquad x > s; \quad (4.16)$$

for $c = 1$,

$$u(x) = ln(x - s), \quad x > s. \tag{4.17}$$

Then

$$u'(x) = (x - s)^{-c}, \quad x > s. \tag{4.18}$$

$$\rho(x) = s + x^{-1/c}, \quad x > s, \tag{4.19}$$

and

$$\tau(x) = (x - s)/c, \quad x > s. \tag{4.20}$$

Hence

$$W_T = s + \hat{\lambda}^{-1/c}\Psi^{-1/c} \tag{4.21}$$

and

$$s + \hat{\lambda}^{-1/c}E[\Psi^{1-1/c}] = we^{rT}, \tag{4.22}$$

from which it follows that

$$W_T = s + \frac{we^{rT} - s}{E[\Psi^{1-1/c}]}\Psi^{-1/c}, \quad w \geq se^{-rT}. \tag{4.23}$$

Note that the investment problem has no solution if $w < e^{-rT}s$. The investor views the situation as hopeless: for any random payment Y, the terminal wealth (3.1) will be below the minimal required wealth with positive probability.

5. Utility of the Initial Wealth

The maximal expected utility of terminal wealth is a function of the initial wealth w. Let

$$u_0(w) = E[u(W_T)]. \tag{5.1}$$

In the literature, this is called a derived utility function, an implied utility function, an indirect utility function, or an induced utility function. The following relation shows that it is intimately connected with $\hat{\lambda}$:

$$u_0'(w) = \hat{\lambda}e^{rT}, \tag{5.2}$$

which, by (3.7), is equivalent to

$$\frac{d}{dw}E[u(W_T)] = E[u'(W_T)]e^{rT}. \tag{5.3}$$

To prove (5.2), we apply (3.8) to rewrite (5.1) as

$$u_0(w) = E[u(\rho(\hat{\lambda}\Psi))]. \tag{5.4}$$

By the chain rule,

$$\begin{aligned} u_0'(w) &= \frac{d}{d\hat\lambda} E[u(v(\hat\lambda\Psi))] \frac{d\hat\lambda}{dw} \\ &= E[u'(W_T)v'(\hat\lambda\Psi)\Psi] / \frac{dw}{d\hat\lambda}. \end{aligned} \tag{5.5}$$

Now, it follows from (3.6) that

$$E[u'(W_T)v'(\hat\lambda\Psi)\Psi] = \hat\lambda\, E[v'(\hat\lambda\Psi)\Psi^2]. \tag{5.6}$$

Also, by (3.9),

$$\frac{dw}{d\hat\lambda} = e^{-rT} E[v'(\hat\lambda\Psi)\Psi^2]. \tag{5.7}$$

Applying (5.6) and (5.7) to the right-hand side of (5.5) yields (5.2).

The function $u_0(w)$ does have the two properties of a risk-averse utility function: $u_0' > 0$ and $u_0'' < 0$. It follows from (5.2) and (3.7) that the first derivative u_0' has the same sign as u', which is positive. To check the sign of the second derivative u_0'', we differentiate (5.2):

$$u_0''(w) = e^{rT} \frac{d\hat\lambda}{dw} = e^{rT} / \frac{dw}{d\hat\lambda}. \tag{5.8}$$

Applying (5.7) to (5.8) yields

$$u_0''(w) = e^{2rT} / E[v'(\hat\lambda\Psi)\Psi^2]. \tag{5.9}$$

Because u'' is negative, it follows that v' is negative. Hence (5.9) shows that u_0'' is negative.

Let

$$\tau_0(w) = -u_0'(w)/u_0''(w) \tag{5.10}$$

be the risk tolerence function that is associated to the derived utility function $u_0(w)$. It follows form (5.2) and (5.8) that

$$\tau_0(w) = -\hat\lambda \frac{dw}{d\hat\lambda}. \tag{5.11}$$

which is needed to derive the important formulas (6.9) and (7.10).

It turns out that, for the LRT utility functions,

$$\tau_0(w) = e^{-rT} \tau(we^{rT}). \tag{5.12}$$

In other words, $u_0(w)$ is equivalent to $u(we^{rT})$. This equivalence is verified by showing that, in each of the three subclasses of LRT utility functions, $u_0'(w)$ is proportional to $u'(we^{rT})$. In view of (5.2), it is sufficient to check that $\hat\lambda$, as a function of w, is proportional to $u'(we^{rT})$. This can be done using formulas (4.7), (4.14) and (4.22), respectively.

6. Optimal Dynamic Investment Strategies — A Single Risky Asset

The result in previous sections are obtained under the assumption of a securities market in which random payments due at time T can be traded or contracted at time 0. If such a market does not exist, it may be possible to create the random payments in a synthetic way by dynamically trading the primitive securities. Now consider a market in which securities can be traded without transaction costs or taxes at all times $t, 0 \leq t \leq T$. In a complete securities market, each contingent claim or random payment can be replicated by a self-financing, dynamically adjusted portfolio. In particular, the optimal terminal wealth W_T can be replicated, starting with the amount w at time 0. The goal is to determine this optimal dynamic investment strategy.

Consider the model of one primitive risky asset that was discussed in Subsection 2.1. It follows from (2.3) and (2.4) that the price density can be written as

$$\Psi = S(0)^{-h^*} e^{-\alpha T} S(T)^{h^*}, \tag{6.2}$$

where α is a constant such that $E[\Psi] = 1$. By (3.8), the optimal terminal wealth is

$$W_T = v(\hat{\lambda} S(0)^{-h^*} e^{-\alpha T} S(T)^{h^*}). \tag{6.3}$$

As in Section 10.6 of Panjer *et al.* (1998), let us consider a contingent claim at time T, which is a function of the risky asset price at time $T, S(T)$, but does not otherwise depend on the asset prices before time T. That is, the contingent payment is $\pi(S(T))$, where the function $\pi(\cdot)$ is called a *payoff* function. Let $V(s,t)$ denote the price for the contingent claim at time t, given that $S(t) = s, 0 \leq t \leq T$. Because the market is complete, the contingent claim can be replicated by a dynamic, self-financing portfolio. For $0 \leq t \leq T$, let $\eta(S(t), t)$ be the amount in the replicating portfolio invested in the risky asset at time t. [Hence $V(S(t), t) - \eta(S(t), t)$ is the amount in the replicating portfolio invested in the risk-free asset at time t]. It can be shown (Baxter and Rennie 1996, p. 95; Dothan 1990, p. 228; Gerber and Shiu 1996, Formula 7.22; Panjer *et al.* 1998, Formula 10.6.6) that

$$\eta(s,t) = s \frac{\partial}{\partial s} V(s,t). \tag{6.4}$$

[The partial derivative $\frac{\partial}{\partial s} V(s,t)$ is called *delta* in the option-pricing literature].

Because W_T depends on both $S(T)$ and $S(0)$, formula (6.4) is not directly applicable to determine the replicating portfolio of W_T. However, it can be applied to a related contingent claim defined by

$$\pi(S(T)) = v(\hat{\lambda} e^{-\alpha T} S(T)^{h^*}). \tag{6.5}$$

With $\hat{\lambda}$ viewed as a parameter, both (6.3) and (6.5) are actually two related families of payoffs. Formula (6.5) is (6.3) with $\hat{\lambda}$ replaced by $\hat{\lambda} S(0)^{h^*}$. Thus, because $w = w(\hat{\lambda})$ is the time-0 price for (6.3) as a function of the parameter $\hat{\lambda}$, the time-0 price of (6.5) must be

$$V(S(0), 0) = w(\hat{\lambda} S(0)^{h^*}). \tag{6.6}$$

According to (6.4) and the chain rule,

$$\begin{aligned} \eta(s,0) &= s w'(\lambda s^{h^*})\lambda h^* s^{h^*-1} \\ &= h^* \lambda s^{h^*} w'(\lambda s^{h^*}) \end{aligned} \tag{6.7}$$

for the replicating portfolio of (6.5).

The goal is to find the replicating portfolio of W_T. For $0 \le t < T$, let W_t denote the value at time t of the replicating portfolio. In other words, by using the optimal strategy, the investor's initial investment of $W_0 = w$ accumulates to wealth W_t at time t. Let $\rho(W_t, t)$ be the amount in the replicating portfolio invested in the risky asset at time t. [The amount invested in the risk-free asset at time t is $W_t - \rho(W_t, t)$.] Note that the right-hand sides of (6.3) and (6.5) are identical if

$$s = S(0) = 1.$$

Hence it suffices to make this substitution in (6.7) to obtain

$$\rho(w, 0) = h^* \lambda w'(\lambda). \tag{6.8}$$

Furthermore, it follows from (5.11) and (2.9) that

$$\rho(w, 0) = -h^* \tau_0(w) \tag{6.9}$$

$$= \frac{\mu - r^*}{\sigma^2} \tau_0(w). \tag{6.10}$$

The extension of these formulas to time $t, 0 < t < T$, is evident. Let $w_t(w)$ denote the conditional expected utility of optimal terminal wealth, given that $W_t = w$, and let $\tau_t(w)$ be the corresponding risk tolerance function Hence the random variable $\tau_t(W_t)$ is the implied risk tolerance at time t, and shall be called the *current risk tolerance*.

Generalizing (6.9) and (6.10),

$$\rho(W_t, t) = - h^* \tau_t(W_t) \tag{6.11}$$

$$= \frac{\mu - r}{\sigma^2} \tau_t(W_t), \qquad 0 \le t < T. \tag{6.12}$$

This provides the simple investment rule mentioned at the end of Subsection 2.1: **the optimal amount invested in the risky asset is the product of the current risk tolerance and the risk premium on the risky asset, divided by the square of the diffusion coefficient.**

Consider next the optimal investment in the risky asset as a fraction of the investor's current wealth at time t, W_t. Hence define

$$M(W_t, t) = \rho(W_t, t)/W_t, \tag{6.13}$$

which is called the *Merton ratio*, in honor of the Nobel laureate Robert C. Merton. By (6.8),

$$M(w, 0) = h^* \lambda w'(\lambda)/w. \tag{6.14}$$

Note that the quantity $\hat{\lambda} w'(\hat{\lambda})/w$ is the *elasticity* of the initial wealth w with respect to the expected marginal utility of optimal terminal wealth $\hat{\lambda}$. For further discussion on this point, see Section 8 of Gerber and Shiu (1998).

It follows from (5.12) that, for the LRT utility functions, (6.9) and (6.10) can be written as

$$\rho(w,0) = -h^* e^{-rT} \tau(we^{eT}) = \frac{\mu - r}{\sigma^2} e^{-rT} \tau(we^{rT}). \qquad (6.15)$$

Applying (6.15) to (6.13) yields

$$M(w,0) = -h^* e^{-rT} \tau(we^{rT})/w = \frac{\mu - r}{\sigma^2 w} e^{-rT} \tau(we^{rT}). \qquad (6.16)$$

Hence for the LRT utility functions, some very explicit formulas can be obtained, as shown in the next three subsections.

6.1. *Constant Risk Tolerance Utility Functions*

It follows from (4.5) that

$$\rho(w,0) = -h^* e^{-rT}/a \qquad (6.17)$$

and

$$M(w,0) = -h^* e^{-rT}/(aw). \qquad (6.18)$$

For $0 \le t < T$, (6.17) is generalized to

$$\rho(W_t,t) = -h^* e^{-r(T-t)}/a = \frac{\mu - r^*}{\sigma^2 a} e^{-r(T-t)}. \qquad (6.19)$$

Thus for an exponential utility function, the product of the amount invested in the risky asset and the interest accumulation factor to time T is the constant $-h^*/a$. In particular, the amount invested in the risky asset is independent of the investor's wealth, reflecting the fact that the risk tolerance function is constant.

6.2. *Decreasing* LRT *Utility Functions*

It follows from (4.12) that

$$\rho(w,0) = -(h^*/c)(se^{-rT} - w), \quad w < se^{-rT}, \qquad (6.20)$$

and

$$M(w,0) = -(h^*/c)(se^{-rT} - w)/w, \quad w < se^{-rT}. \qquad (6.21)$$

For $0 \leq t < T$, (6.21) is generalized to

$$
\begin{aligned}
M(W_t, t) &= -(h^*/c)(se^{-r(T-t)} - W_t)/W_t \\
&= \frac{\mu - r^*}{\sigma^2 c} \frac{se^{-r(T-t)} - W_t}{W_t}, \quad W_t < se^{-r(T-t)},
\end{aligned} \tag{6.22}
$$

which shows that the optimal amount invested in the risky asset is proportional to the difference between the discounted level of saturation and the current wealth.

6.3. *Increasing* LRT *Utility Functions*

It follows from (4.20) that

$$
\rho(w, 0) = -(h^*/c)(w - se^{-rT}), \quad w \geq se^{-rT}, \tag{6.23}
$$

and

$$
M(w, 0) = -(h^*/c)(w - se^{-rT})/w, \quad w \geq se^{-rT} \tag{6.24}
$$

For $0 \leq t < T$, (6.24) is generalized to

$$
\begin{aligned}
M(W_t, t) &= -(h^*/c)(W_t - se^{-r(T-t)}/W_t \\
&= \frac{\mu - r^*}{\sigma^2 c} \frac{W_t - se^{-r(T-t)}}{W_t}, \quad W_t \geq se^{-r(T-t)}.
\end{aligned} \tag{6.25}
$$

Hence optimal the amount invested in the risky asset is proportional to the excess of current wealth over the discounted value of the minimal required terminal wealth.

An investor may have the constraint that the terminal wealth must be at least s, $s > 0$. Then $se^{-r(T-t)}$ is the portion of current wealth W_t that grows to s at time T with certainty. Then complement, $W_t - se^{-r(T-t)}$, is the portion of current wealth that can considered as "free". Thus (6.25) shows that, at any time, a constant proportion of the "free" wealth should be invested in the risky asset.

In the special case $s = 0$, (6.23) reduces to

$$
M(W_t, t) = -h^*/c = \frac{\mu - r^*}{\sigma^2 c}, \tag{6.26}
$$

which is a constant. Thus, at any time $t, 0 \leq t < T$, a constant proportion of the current wealth should be invested in the risky asset. This formula, not the more general formula (6.13), is usually called the Merton ratio in the literature (Panjer *et al.*, 1998, p. 141). Consider an investor who has an increasing LRT utility function with $c = 4/3$ and $s = 0$. If the risk premium on the risky asset, $\mu - r^*$, is 5% and the volatility of the risky asset as measured by σ is 25%, then the Merton ratio (6.26) for this investor is 60%.

7. Optimal Dynamic Investment Strategies — Multiple Risky Assets

The results in the last section can be generalized to the case of multiple primitive risky assets (the model in Subsection 2.2), which is more relevant in practice. Similar to (6.2), the price density can be written as

$$\Psi = e^{-\alpha T} \prod_{k=1}^{n} [S_k(T)/S_k(0)]^{h_k^*}. \tag{7.1}$$

(As in the last section, the value of α is not relevant in the following. It is such that $E[\Psi] = 1$.) By (3.8) and (7.1), the optimal terminal wealth is now

$$W_T = v\left(\lambda e^{-\alpha T} \prod_{k=1}^{n} [S_k(T)/S_k(0)]^{h_k^*}\right). \tag{7.3}$$

Again, let $W_0 = w$, and for $0 < t < T$, let W_t denote the value of the replicating portfolio for W_T. For $0 \le t < T$ and $k = 1, 2, \ldots, n$, let $\rho_k(W_t, t)$ be amount invested in risky asset k in the replicating portfolio of W_T at time t. Generalizing (6.13), let

$$M_k(W_t, t) = \rho_k(W_t, t)/W_t \tag{7.4}$$

be the Merton ratio for risky asset k at time $t, k = 1, 2, \ldots, n$.

Formula (6.4) can be generalized as follows. Consider a contingent claim with payoff

$$\pi(S_1(T), S_2(T), \ldots, S_n(T)) \tag{7.5}$$

at time T, for some payoff function π. For $0 \le t < T$, let $V(s_1, s_2, \ldots, s_n, t)$ denote its price at time t, and for $k = 1, 2, \ldots, n$, let $\eta_k(s_1, s_2, \ldots, s_n, t)$ denote the amount of risky asset k in the replicating portfolio for (7.5), given that $S_j(t) = s_j, j = 1, 2, \ldots, n$. Then it is known (Gerber and Shiu 1996, Formula 8.35) that

$$\eta_k = s_k \frac{\partial V}{\partial s_k}. \tag{7.6}$$

In particular, consider the payoff

$$\pi(S_1(T), S_2(T), \ldots, S_n(T)) = v\left(\hat{\lambda} e^{-\alpha T} \prod_{j=1}^{n} S_j(T)^{h_k^*}\right). \tag{7.7}$$

Again, with $\hat{\lambda}$ viewed as a parameter, this is a family of payoffs. Because $w = w(\hat{\lambda})$ is the time-0 price of (7.3), we gather that the time-0 price of (7.7) is

$$V(s_1, s_2, \ldots, s_n, 0) = w\left(\hat{\lambda} \prod_{j=1}^{n} s_j^{h_k^*}\right). \tag{7.8}$$

Hence, by (7.6) and the chain rule,

$$\eta_k(s_1, s_2, \ldots, s_n, 0) = w'\left(\hat{\lambda}\prod_{j=1}^{n} s_j^{h_k^*}\right)\hat{\lambda}h_k^*\prod_{j=1}^{n} s_j^{h_k^*}. \tag{7.9}$$

To obtain $\rho_k(w, 0)$, it suffices to set $s_1 = s_2 = \ldots = s_n = 1$. Thus

$$\rho_k(w, 0) = \hat{\lambda}h_k^* w'(\hat{\lambda}) = -h_k^*\tau_0(w), \tag{7.10}$$

Hence

$$\rho_k(W_t, t) = -h_k^*\tau_t(W_t), \tag{7.11}$$

generalizing (6.12).

Note that the amount invested in risky asset k, as a fraction of the total amount invested in all risky assets, is

$$\frac{\rho_k(W_t, t)}{\sum_{j=1}^{n}\rho_j(W_t, t)} = \frac{h_k^*}{\sum_{j=1}^{n} h_j^*}, \tag{7.12}$$

which depends only on the risk-neutral Esscher parameters and is constant, say q_k, at all times. Hence we have a "mutual fund" theorem: for any risk-averse investor, the risky asset portion of his optimal investment portfolio is of identical composition. That is, each investor only needs to invest in (or borrow from) the same two mutual funds — one is a risk-free bond fund while the other is a risky-asset mutual fund whose portfolio mix is continuously adjusted so that at all times the fraction of its value invested in risky asset k is q_k. The investor's exposure or amount of investment in the risky-asset mutual fund divided by his or her "current risk tolerance" is always kept at the constant level

$$-\sum_{j=1}^{n} h_j^*. \tag{7.13}$$

8. Concluding Remarks

For the investor who wants to maximize the expected utility of terminal wealth, some general and simple rules have been established. In the case where the investment vehicles are a risk-free asset and a risky asset, the amount invested in the risky asset should at any time be the product of the current risk tolerance and the risk premium on the risky asset, divided by the square of the diffusion coefficient. It is natural to formulate this result in terms of the risk-neutral Esscher parameter, and also in terms of the elasticity. For example, the amount invested in the risky asset, as a fraction of the total assets, should at any time be the risk-neutral Esscher parameter divided by the elasticity, with respect to current wealth, of the expected marginal utility of terminal wealth. More general, but similarly transparent, results are found for the optimal strategy in the more realistic case, where the investment vehicles comprise more than one risky asset. It is shown that the ratio of the amounts invested in the various risky assets are constant in time and are proportional to the risk-neutral Esscher parameters. Hence the risky assets can be replaced by a single mutual fund with the right asset mix. In this sense, the case of multiple risky assets can be reduced to the case of a single risky asset. This paper can be refined and extended in various directions. The interested reader may want to consult research papers such as Boyle and Lin (1997), Boyle and Yang (1997), Browne (1998), Cox and Huang (1989), Dybvig (1988), Dybvig (1988), Karatzas, Lehoczky, Sethi and Shreve (1986) and Pliska (1986) and their references. Also, many recent articles on optimal investment strategies are published in the journal *Mathematical Finance*.

Acknowledgments

The authors thank Professor Arnold Shapiro for his comments. Elias Shiu gratefully acknowledges the support from the Principal Financial Group Foundation.

REFERENCES

[1] BAXTER M., RENNIE R., *Financial Calculus: An Introduction to Derivative Pricing*, Cambridge University Press, Cambridge, U.K., 1996.

[2] BICKSLER J. L., SAMUELSON P. A., editors, *Investment Portfolio Decision-Making*, Lexington Books, Lexington, Mass., 1974.

[3] BJÖRK T., *Arbitrage Theory in Continuous Time*, Oxford University Press, New York, 1998.

[4] BOYLE P. P., LIN X., *Optimal Portfolio Selection with Transaction Costs*, North American Actuarial Journal, 1(2), 1997, 27-39.

[5] BOYLE P.P., YANG H., *Asset Allocation with Time Variation in Expected Returns*, Insurance: Mathematics and Economics, 21, 1997, 201-218.

[6] BROWNE S., *The Return on Investment from Proportional Portfolio Strategies*, Advances in Applied Probability, 30, 1998, 216-238.

[7] CONSTANTINIDES G. M., MALLARIAS A. G., *Portfolio Theory*, in *Handbooks in Operations Research and Management Science*, Finance, edited by Jarrow R.A., Maksimovic V. and Ziemba W. T., Elsevier, Amsterdam, 9, 1995, 1-30.

[8] COX J. C., HUANG C. F., *Optimum Consumption and Portfolio Policies When Asset Prices Follow a Diffusion Process*, Journal of Economic Theory, 49, 1989, 33-83.

[9] DEPREZ O., GERBER H. U., *On Convex Principles of Premium Calculation*, Insurance: Mathematics and Economics, 4, 1985, 179-89.

[10] DOTHAN M. U., *Prices in Financial Markets*, Oxford University Press, New York, 1990.

[11] DUFFIE D., *Dynamic Asset Pricing Theory*, 2nd ed. Princeton University Press, Princeton, 1996.

[12] DYBVIG P. H., *Inefficient Dynamic Portfolio Strategies or How to Throw Away a Million Dollars in the Stock Market*, Review of Financial Studies, 1: 1988, 67-88.

[13] ESSCHER F., *On the Probability Function in the Collective Theory of Risk*, Skandinavisk Aktuarietidskrift, 15, 1932, 175-195.

[14] GERBER H. U., PAFUMI G., *Utility Functions: From Risk Theory to Finance*, North American Actuarial Journal, 2(3), 1998, 74-91.

[15] GERBER H. U., SHIU E. S. W., *Option Pricing by Esscher Transforms*. Transactions of the Society of Actuaries, 46, 1994, 99-140, Discussions 141-191.

[16] GERBER H. U., SHIU E. S. W., *Actuarial Bridges to Dynamic Hedging and Option Pricing*, Insurance: Mathematics and Economics, 18, 1996, 183-218.

[17] GERBER H. U., SHIU E. S. W., *Investing for Retirement: Optimal Capital Growth and Dynamic Asset Allocation*, Working paper, 1998.

[18] HAKANSSON N. H., *Portfolio Analysis*, in *The New Palgrave: a Dictionary of Economics*, edited by J. Eatwell, M. Milgate and P. Newman, Macmillan, London, 3, 1987, 917-920.

[19] HAKANSSON N. H., ZIEMBA W. T., *Capital Growth Theory*, in *Handbooks in Operations Research and Management Science*, Finance, edited by R. A. Jarrow, V. Maksimovic and W. T. Ziemba, Elsevier, Amsterdam, 9, 1995, 65-86.

[20] KARATZAS I., LEHOCZKY J. P., SHREVE S. E., *Explicit Solution of a General Consumption/Investment Problem*, Mathematics of Operations Research, 11, 1986, 261-294.

[21] KARATZAS I., SHREVE S. E., *Methods of Mathematical Finance*, Springer, New York, 1998.

[22] KORN R., *Optimal Portfolios: Stochastic Models for Optimal Investment and Risk Management in Continuous time*, World Scientific, Singapore, 1997.

[23] LUENBERGER D. G., *Investment Science*, Oxford University Press, New York, 1998.

[24] MERTON R. C., *Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case*, Review of Economics and Statistics, 51, 1969, 247-57. Reprinted as Chapter 4 in Merton 1990.

[25] MERTON R. C., *Continuous-Time Finance*, Blackwell, Oxford, 1990.

[26] MOSSIN J., *Optimal Multiperiod Portfolio Policies*, Journal of Business, 41, 1968, 215-229.

[27] PANJER H. H., BOYLE P. P., COX S. H., DUFRESNE D., GERBER H. U., MUELLER H. H., PEDERSEN H. W., PLISKA S. R., SHERRIS M., SHIU E. S., TAN K. S., *Financial Economics: With Applications to Investment, Insurance, and Pensions*, The Actuarial Foundation, Schaumburg, Ill., 1998.

[28] PLISKA S. R., *A Stochastic Calculus Model of Continuous Time Trading: Optimal Portfolios*, Mathematics of Operations Research. 11, 1986, 371-84.

[29] PLISKA S. R., *Introduction to Mathematical Finance: Discrete Time Models*, Blackwell, Oxford, 1997.

[30] SAMUELSON P. A., *Lifetime Portfolio Selection by Dynamic Stochastic Programming*, Review of Economics and Statistics, 51, 1969, 239-46.

[31] SAMUELSON P. A., *Asset Allocation Could Be Dangerous to Your Health*, Journal of Portfolio Management, Spring issue, 1990, 5-8.

[32] SAMUELSON P. A., *How Best to Flip-Flop If You Must: Integer Dynamic Programming for Either-Or*, Journal of Risk and Uncertainty, 15, 1997. 183-190.

[33] SETHI S. P., *Optimal Consumption and Investment with Bankruptcy*, Kluwer, Hingham, Mass., 1996.

[34] SHARPE W. F., *The Sharpe Ratio*, Journal of Portfolio Management (Fall issue), 1994, 49-58.

[35] ZIEMBA W. T., VICKSON R. G., (editors), *Stochastic Optimization Models in Finance*, Academic Press, New York, 1975.

Sulle strategie di investimento ottimali

RIASSUNTO

Si consideri un investitore con un orizzonte decisionale dato e funzione d'utilità definita sulla sua ricchezza. Se le attività disponibili sono solo due, una rischiosa e l'altra no, allora la strategia d'investimento ottima è tale che in ogni istante l'ammontare investito nell'attività rischiosa deve essere il prodotto della sua "tolleranza al rischio attuale" e del premio al rischio per ogni attività, diviso per il quadrato del coefficiente di diffusione dell'attività rischiosa. Nel caso di più attività rischiose, la strategia d'investimento ottima è simile e mantiene i rapporti delle quantità investite nelle diverse attività costanti nel tempo.

Corresponding author:
HANS U. GERBER
École des Hautes Études Commerciales
Università di Losanna, Svizzera
CH-1015 Losanna
Svizzera
e-mail: hgerber@hec.unil.ch