INVESTING FOR RETIREMENT: OPTIMAL CAPITAL GROWTH AND DYNAMIC ASSET ALLOCATION

Hans U. Gerber* and Elias S.W. Shiu†

ABSTRACT

A frequently quoted rule of thumb for allocating assets in a pension plan is that at any time 60% should be in stocks and 40% in bonds. How can these percentages be justified? What are the criteria under which this type of dynamic investment strategy is optimal, and what are the optimal strategies, if the criteria are modified?

Suppose that an investor has a certain decision horizon and has chosen an appropriate utility function for measuring his utility of wealth at that time. Then, maximizing the expected utility of wealth at the decision horizon leads to a rational compromise between risk and return. First, we consider the one-period model, in which arbitrary random payments, due at the end of the time period, are traded at the beginning of the interval and valued by a price density. To facilitate understanding of the optimal decision, we introduce the risk tolerance function that is associated with the utility function, and also the implied utility function, that is, the maximal expected utility considered a function of the initial wealth. Second, we consider continuous-time complete securities market models, in which rebalancing of the asset portfolios can take place dynamically over time. This seemingly more complex problem can be reduced to the first problem. The key is that the optimal investment strategy corresponds to the self-financing portfolio that replicates the optimal payoff in the first problem. If there are only two investment vehicles, a risky and a risk-free asset, then the optimal investment strategy is as follows: at any time, the amount invested in the risky asset must be the product of the current risk tolerance and the risk premium on the risky asset, divided by the square of the diffusion coefficient of the risky asset. This result can be restated as follows: the Merton ratio, which is the fraction of current wealth invested in the risky asset, must be the risk-neutral Esscher parameter divided by the elasticity, with respect to current wealth, of the expected marginal utility of optimal terminal wealth. In the more realistic case with more than one risky asset, equally explicit rules are given for the optimal investment strategy. For example, the ratios of the amounts invested in the different risky assets are constant in time; the ratios depend only on 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. Throughout the paper, explicit formulas are given for the cases of linear risk tolerance utility functions.

1. INTRODUCTION

For a pension plan, the investment of assets is a central and challenging problem. In a defined-benefit plan, the investment problem is of primary concern to the employer, and the awareness among the employees may be minimal. This situation has radically changed in the recent past. In order to reduce the financial risk and the regulatory burden, there has been a substantial shift toward defined-contribution plans such as 401(k) and 403(b) plans (Mitchell and Schieber 1998). In such a plan, the asset allocation problem is of crucial interest to the employee. Typically, the employee is offered a family of mutual funds to which he can allocate (and reallocate) the pension

*Hans U. Gerber, A.S.A., Ph.D., is Professor of Actuarial Science, Ecole des HEC, University of Lausanne, CH-1015 Lausanne, Switzerland, e-mail, hgerber@hec.unil.ch.

†Elias S.W. Shiu, A.S.A., Ph.D., is Principal Financial Group Professor of Actuarial Science, Department of Statistics and Actuarial Science, The University of Iowa, Iowa City, Iowa 52242-1409, e-mail, eshiu@stat.iowa.edu.

fund contributions. How should he decide? What are rational solutions of the asset allocation problem?

A proposal for reforming the U.S. Social Security system is to allow the taxpayer to invest a portion of his Social Security tax in the stock market or mutual funds. If this private account proposal is adopted by Congress, then every taxpayer will face an asset allocation problem. How should the decision be made?

If we present the asset allocation problem to an actuarial student or an MBA student, he might suggest Markowitz mean-variance analysis as a solution. Unfortunately, the Markowitz model is a one-period model. Investing for retirement is a long-term, multiperiod problem. As time passes, the employee can and will reallocate or rebalance his assets. Let us quote from a recent book (Luenberger 1998):

> Conclusions about multiperiod 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. (p. 417)

Other insightful remarks can be found in the articles (Samuelson 1990, 1997) by the Nobel laureate Paul Samuelson.

In this paper we treat the multiperiod, continuous-time investment problem, where reallocation or rebalancing of assets takes place dynamically over time. For a given decision horizon and a given criterion to assess the utility of wealth, what is the optimal dynamic investment strategy? In a defined-contribution plan, this problem is relevant to an individual employee; typically the decision horizon is the time of his retirement. In a defined-benefit plan, the problem concerns the employer—in particular, the pension fund manager.

Although the problem of optimal capital growth and dynamic asset allocation is normally not treated in MBA-level textbooks, expositions of it 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), Szegö and Shell (1972, Part 1), 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 pension fund, the first problem appears to be much more important. Including the second problem may add unnecessary complications and require heavy mathematical machinery such as dynamic programming in continuous time (stochastic control theory) and the Hamilton-Jacobi-Bellman (HJB) equation. Indeed, in 401(k) and 403(b) plans, consumption before retirement may lead to tax penalties.

The aim of this paper is to give 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, in which 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 the decision horizon). To understand the optimal decision, it is judicious to introduce the risk tolerance function that is associated with the utility function, and also 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 to 8. In Section 6, the investment vehicles are a riskless asset and a single risky asset. In Section 7, for greater realism the universe of investment vehicles is expanded to include more than one risky asset. The optimal dynamic investment strategy can be found through application of 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 parameters (a term coined to honor the Swedish actuary F. Esscher), and, as shown in Section 8, the results can be formulated in terms of the elasticity, with respect to the current wealth, of the expected marginal utility of optimal terminal wealth.

A frequently given rule of thumb for allocating assets in a pension fund is that at any time 60% should be in stocks and 40% in bonds; for example, this is the official investment strategy for the Social Choice Account of TIAA-CREF. A rationalization of this ‘‘constant mix’’ strategy is given at the end of Section 6.

2. SECURITIES MARKET MODELS WITH PRICE DENSITY

We consider an investor with wealth w at time 0. His goal is to invest in the securities market in order to optimize, in a certain sense, his wealth at a fixed future time $T (T > 0)$. Because T represents the strategic planning horizon of the investor, all discussions in this paper are restricted to the time interval $[0, T]$. We assume that there is no withdrawal or addition of funds between time 0 and time 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 $E[\Psi] = 1$. (In this paper, all expectations are taken with respect to ''real-world'' probabilities, which are assumed to be known.) For each random payment Y due at time T, its price, known at time 0 and payable 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 remains constant over 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 classic 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 paying no dividends between times 0 and T, and their derivative securities. For $0 \le t \le T$, let the price of a unit of the risky asset a time t be denoted as $S(t)$. Let

$$X(t) = \ln[S(t)/S(0)] \tag{2.1}$$

be the continuously compounded rate of return over the time interval $[0, t]$. As in the classic Black-Scholes model, we assume that $\{X(t)\}$ is a *Wiener process* (or *Brownian motion*) with ''real-world'' drift parameter μ and diffusion coefficient σ. That is, for each fixed t, $X(t)$ is a normal random variable with mean μt and variance $\sigma^2 t$. As pointed out in Gerber and Shiu (1994, 1996) and in Panjer et al. (1998, Section 10.9), there is a unique number h^*, called the *risk-neutral Esscher parameter,* such that

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

is the price density of the securities market. Considering $S(T)$ as a random payment, we have

$$S(0) = e^{-rT}E[S(T)\Psi] \tag{2.3}$$

or

$$1 = e^{-rT}\frac{E[e^{(1+h^*)X(T)}]}{E[e^{h^*X(T)}]},$$

from which, and from the moment-generating-function formula

$$E[e^{hX(t)}] = \exp[(h\mu + \tfrac{1}{2}h^2\sigma^2)t], \tag{2.4}$$

we obtain

$$h^* = \frac{r - \sigma^2/2 - \mu}{\sigma^2}. \tag{2.5}$$

Remarks

(i) With the notation

$$\mu^* = r - \frac{\sigma^2}{2} \tag{2.6}$$

(2.5) becomes

$$h^* = \frac{\mu^* - \mu}{\sigma^2}. \tag{2.7}$$

The number μ^* plays a major role in modern option-pricing theory. The random variable Ψ can be used as a Radon-Nikodym derivative, with which the probability measure is changed to the so-called *risk-neutral measure* or *equivalent martingale measure.* With respect to the risk-neutral measure, the process $\{X(t)\}$ is still a Wiener process, with the same diffusion coefficient σ but a changed drift parameter μ^*.

(ii) Many authors prefer to describe the price dynamics of the risky asset in terms of a stochastic differential equation, which in our notation is

$$\frac{dS(t)}{S(t)} = \left(\mu + \frac{\sigma^2}{2}\right)dt + \sigma dZ(t), \tag{2.8}$$

where $\{Z(t)\}$ is a standard Wiener process. These authors would use μ in place of our $(\mu + \sigma^2/2)$ term; with this μ, (2.5) would be

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

(iii) Define

$$Z^*(t) = Z(t) - \sigma h^* t. \tag{2.10}$$

Then $\{Z^*(t)\}$ is a standard Wiener process with respect to the risk-neutral measure. Equation (2.8) becomes

$$\frac{dS(t)}{S(t)} = rdt + \sigma dZ^*(t), \tag{2.11}$$

showing that, under the risk-neutral measure, the expected instantaneous rate of return of the risky asset (and hence every asset) is the risk-free rate r.

(iv) The assumption that the risky asset pays no dividends is not critical, because we can consider all dividends being reinvested in the asset immediately. That is, we can let $S(t)$ denote the time-t price of the risky asset with all dividends reinvested.

We shall return to this geometric Brownian motion model for the risky asset in Subsection 4.4 and in Section 6. It turns out 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 $-h^*$. Note that the numerator of

$$-h^* = \frac{\mu - \mu^*}{\sigma^2} \tag{2.12}$$

is the *risk premium* on the risky asset, that is, the excess of the drift parameter over the one used for pricing, or equivalently, the excess of the expected instantaneous rate of return of the risky asset over the risk-free rate. In the finance literature, the quantity $-h^*\sigma$ is sometimes called the "market price of risk" (Baxter and Rennie 1996, p. 119; Hull 1993, p. 276); 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 $k = 1, 2, \ldots, n$, we let $S_k(t)$ denote the price of the k-th risky asset at time t, and write

$$X_k(t) = \ln[S_k(t)/S_k(0)].$$

We assume that $\{X_1(t), X_2(t), \ldots, X_n(t)\}$ is an n-dimensional Wiener process, with

$$\mathrm{E}[X_k(1)] = \mu_k \tag{2.13}$$

and

$$\mathrm{Cov}[X_i(1), X_j(1)] = \sigma_{ij}. \tag{2.14}$$

Under the assumption that the covariance matrix $\mathbf{C} = (\sigma_{ij})$ is nonsingular, Gerber and Shiu (1994, Section 7) have shown that there are n numbers $h_1^*, h_2^*, \ldots, h_n^*$ such that

$$\Psi = \frac{S_1(T)^{h_1^*}S_2(T)^{h_2^*} \cdots S_n(T)^{h_n^*}}{\mathrm{E}[S_1(T)^{h_1^*}S_2(T)^{h_2^*} \cdots S_n(T)^{h_n^*}]} \tag{2.15}$$

is the price density of the securities market. The n risk-neutral Esscher parameters $h_1^*, h_2^*, \ldots, h_n^*$ are determined by the n equations

$$S_k(0) = \mathrm{e}^{-rT}\mathrm{E}[S_k(T)\Psi], \quad k = 1, 2, \ldots, n, \tag{2.16}$$

which generalize (2.3).

Similar to (2.6), we define

$$\mu_k^* = r - \frac{\sigma_{kk}^2}{2}, \qquad k = 1, 2, \ldots, n. \tag{2.17}$$

Applying the moment-generating function formula for multivariate normal random variables to (2.16), we can show that $h_1^*, h_2^*, \ldots, h_n^*$ are the solution of the system of linear equations

$$\sum_{j=1}^{n} \sigma_{ij}h_j^* = \mu_i^* - \mu_i, \qquad i = 1, 2, \ldots, n. \tag{2.18}$$

Remark

Let

$$\mathbf{h}^* = (h_1^*, h_2^*, \ldots, h_n^*), \tag{2.19}$$

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n) \tag{2.20}$$

and

$$\boldsymbol{\mu}^* = (\mu_1^*, \mu_2^*, \ldots, \mu_n^*). \tag{2.21}$$

With these row vectors, we have

$$\mathbf{h}^* = (\boldsymbol{\mu}^* - \boldsymbol{\mu})\mathbf{C}^{-1}, \tag{2.22}$$

which generalizes (2.7).

3. CHARACTERIZING THE OPTIMAL TERMINAL WEALTH

If the investor (with wealth w at time 0) buys the time-T random payment Y in the securities market, his wealth at time T is

$$W(Y) = w\mathrm{e}^{rT} + Y - \mathrm{E}[Y\Psi]. \tag{3.1}$$

We note that, for each Y,

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

We shall use the term *terminal wealth* to describe such a random variable $W(Y)$, because T is the decision horizon for the investment problem. Note that "wealth" is a standard term in the financial literature; in a given application, there may be more appropriate expressions such as ''accumulation,'' ''net worth,'' or ''fortune.''

We assume that the investor uses a risk-averse utility function $u(x)$ for determining the *optimal* terminal wealth at time T. That is, there is a function $u(x)$, with $u'(x) > 0$ and $u''(x) < 0$, for which we are to determine the terminal wealth $W(Y)$ that maximizes the expectation $\mathrm{E}[u(W(Y))]$. Furthermore, we 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 this journal (Gerber and Pafumi 1998).

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

$$\mathrm{E}[W_T \Psi] = w e^{rT}. \tag{3.3}$$

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

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

that is, the random variable $u'(W_T)$, the marginal utility of the optimal terminal wealth, is proportional to the random variable Ψ, the price density. To see that (3.4) is a necessary condition for optimality, we use the following variational argument, which is becoming a standard tool in actuarial science (Deprez and Gerber 1985; Embrechts and Meister 1997, p. 23); see also Davis (1997) and Bingham and Kiesel (1998, Subsection 7.1.1). Let V be an arbitrary random payment at time T and ξ an arbitrary real number. Then

$$W_T + \xi V - \xi \mathrm{E}[V\Psi]$$

is another terminal wealth. Put

$$\phi(\xi) = \mathrm{E}[u(W_T + \xi V - \xi \mathrm{E}[V\Psi])].$$

The random variable W_T being optimal means

$$\mathrm{E}[u(W_T)] \geq \phi(\xi), \qquad \text{for all } \xi.$$

In other words, the function $\phi(\xi)$ has a maximum at $\xi = 0$. The first-order condition

$$\phi'(0) = 0$$

is

$$\mathrm{E}[(V - \mathrm{E}[V\Psi])u'(W_T)] = 0,$$

which can be rewritten as

$$\mathrm{E}[V\{u'(W_T) - \mathrm{E}[u'(W_T)]\Psi\}] = 0. \tag{3.5}$$

Since V is arbitrary, it follows that the expression within the braces in (3.5) must be zero. Hence we have (3.4).

We now show that (3.4) is a sufficient condition for optimality and also that W_T is unique. Because the utility function u is a strictly concave function, we have

$$u(y) < u(x) + u'(x)(y - x), \qquad x \neq y. \tag{3.6}$$

Suppose that W_T satisfies (3.4) [and (3.3)]. Any other terminal wealth can be written in the form

$$W_T + Y - \mathrm{E}[Y\Psi],$$

where Y is a non-degenerate random variable. It follows from (3.6) and (3.4) that

$$\begin{aligned} &u(W_T + Y - \mathrm{E}[Y\Psi]) \\ &\quad \leq u(W_T) + u'(W_T)(Y - \mathrm{E}[Y\Psi]) \\ &\quad = u(W_T) + \mathrm{E}[u'(W_T)]\Psi(Y - \mathrm{E}[Y\Psi]), \end{aligned}$$

with strict inequality holding with positive probability. Taking expectations, we obtain

$$\mathrm{E}[u(W_T + Y - \mathrm{E}[Y\Psi])] < \mathrm{E}[u(W_T)] + 0,$$

showing that W_T is optimal and that it is unique.

Condition (3.4) states that the random variable for the marginal utility of optimal terminal wealth, $u'(W_T)$, must be a multiple of the price density random variable Ψ. With the notation

$$m = m(w) = \mathrm{E}[u'(W_T)], \tag{3.7}$$

(3.4) becomes

$$u'(W_T) = m\Psi. \tag{3.8}$$

Because $u'' < 0$, the marginal utility function u' has an inverse, which we denote as v. (In Section 8.10 of Panjer et al. 1998, the function v is denoted as h.) That is,

$$v(u'(x)) = x. \tag{3.9}$$

Applying the function v to both sides of (3.8), we obtain

$$W_T = v(m\Psi). \tag{3.10}$$

By (3.3),

$$\mathrm{E}[\Psi\, v(m\Psi)] = w e^{rT}; \tag{3.11}$$

this formula establishes the functional relationship between m and w. The next section presents some important examples of the optimal terminal wealth random variable W_T. In each of these examples, (3.11) leads to an explicit expression for m, which can then be substituted in (3.10).

Remark

Condition (3.4) can also be obtained by the *method of Lagrange multipliers* (Panjer et al. 1998, Section 8.10). Consider the Lagrangian

$$L(W, \lambda) = \mathrm{E}[u(W)] - \lambda\{\mathrm{E}[W\,\Psi] - we^{rT}\}, \quad (3.12)$$

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

$$u'(W_T) = \hat{\lambda}\Psi, \quad (3.13)$$

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

$$\mathrm{E}[u'(W_T)] = \hat{\lambda}. \quad (3.14)$$

Substitution in (3.13) yields (3.4).

4. LRT UTILITY FUNCTIONS

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

$$\tau(x) = \frac{-u'(x)}{u''(x)} \quad (4.1)$$

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 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. We say two utility functions are *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 following three subsections, we give a formula for the utility function for a subclass and the corresponding formula for the optimal terminal wealth W_T.

4.1 Constant Risk Tolerance Utility Functions

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

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

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

and

$$v(x) = \frac{-\ln(x)}{a}, \qquad -\infty < x < \infty. \quad (4.4)$$

The risk tolerance function is

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

Here, (3.10) and (3.11) become

$$W_T = \frac{-[\ln(m) + \ln(\Psi)]}{a} \quad (4.6)$$

and

$$\frac{-\{\ln(m) + \mathrm{E}[\Psi\ln(\Psi)]\}}{a} = we^{rT}, \quad (4.7)$$

respectively. By eliminating the $\ln(m)$ term, we obtain

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

4.2 Decreasing LRT Utility Functions

These are the power utility functions with parameters c and s, $c > 0$, where c is the 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,

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

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

and

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

The risk tolerance function is

$$\tau(x) = \frac{s - x}{c}, \qquad x < s. \tag{4.12}$$

Here, (3.10) and (3.11) become

$$W_T = s - m^{1/c}\Psi^{1/c} \tag{4.13}$$

and

$$s - m^{1/c}\mathrm{E}[\Psi^{1+1/c}] = w\mathrm{e}^{rT}, \tag{4.14}$$

respectively. Hence

$$W_T = s - \frac{s - w\mathrm{e}^{rT}}{\mathrm{E}[\Psi^{1+1/c}]}\Psi^{1/c}, \qquad w < s\mathrm{e}^{-rT}. \tag{4.15}$$

Note that, if $w \geq s\mathrm{e}^{-rT}$, the investment problem has a trivial solution; the investor will achieve maximal satisfaction simply by investing all his 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 $c \neq 1$,

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

for $c = 1$,

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

Then

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

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

and

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

Hence

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

and

$$s + m^{-1/c}\mathrm{E}[\Psi^{1-1/c}] = w\,\mathrm{e}^{rT}, \tag{4.22}$$

from which we obtain

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

Note that the investment problem has no solution if $w < s\mathrm{e}^{-rT}$. 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.

Remark

Increasing LRT utility functions are of a particular interest in the context of *portfolio insurance*. Suppose that we are only to use investment strategies for which the experienced average rate of return is at least r_0 with certainty, $r_0 \leq r$. In other words, the terminal wealth is to be at least $s = w\mathrm{e}^{r_0 T}$. Then one way to treat this problem is to adopt an increasing LRT utility function with this value of s and an appropriately chosen value for the parameter c. More insights will be provided in Subsection 6.3.

4.4 Lognormal Price Density

Let us consider the situation where the price density Ψ is a lognormal random variable, that is, where $\ln(\Psi)$ has a normal distribution. Then the distribution of the optimal terminal wealth can readily be identified. From (4.8) it follows that W_T has normal distribution, if the utility function is exponential. From (4.15), we gather that W_T is the level of saturation minus a lognormal random variable in the case of a decreasing LRT utility function. From (4.23), we see that W_T is the minimal required wealth level plus a lognormal random variable in the case of an increasing LRT utility function.

Note that the price densities of both Subsections 2.1 and 2.2 have a lognormal distribution. In Subsection 2.1 we have

$$\Psi = \frac{\mathrm{e}^{h^* X(T)}}{\mathrm{E}[\mathrm{e}^{h^* X(T)}]}. \tag{4.24}$$

We can use (2.4) to evaluate the expectation in the denominator of (4.24). With the definition

$$\alpha = h^*\mu + \frac{(h^*\sigma)^2}{2}, \tag{4.25}$$

we have

$$\Psi = \mathrm{e}^{h^* X(T) - \alpha T}. \tag{4.26}$$

Then

$$\mathrm{E}[\ln \Psi] = (h^*\mu - \alpha)T \tag{4.27}$$

and

$$\mathrm{E}[\Psi \ln \Psi] = (h^*\mu^* - \alpha)T, \tag{4.28}$$

where μ^* is the risk-neutral drift of $\{X(t)\}$ defined by (2.6).

Consider first the case of an exponential utility function. The optimal terminal wealth, as given by (4.8), is normal. It follows from (4.8), (4.27), (4.28), and (2.7) that

$$\mathrm{E}[W_T] = w\mathrm{e}^{rT} + \frac{h^*(\mu^* - \mu)T}{a}$$
$$= w\mathrm{e}^{rT} + \frac{(h^*\sigma)^2 T}{a} \qquad (4.29)$$

and

$$\mathrm{Var}[W_T] = \frac{(h^*\sigma)^2 T}{a^2}. \qquad (4.30)$$

Thus, the higher the risk tolerance ($1/a$), the higher the expectation of terminal wealth is at the expense of a progressively increasing variance.

Next, consider the case of an increasing LRT utility function. It follows from (4.26) and (4.25) that the denominator in (4.23) is

$$\mathrm{E}[\Psi^{1-1/c}] = \exp\left[(h^*\sigma)^2 T\left(-\frac{1}{2c} + \frac{1}{2c^2}\right)\right]. \qquad (4.31)$$

Hence (4.23) becomes

$$W_T = s + (w\mathrm{e}^{rT} - s)\mathrm{e}^{\gamma T - (h^*/c)X(T)},\ w \geq s\mathrm{e}^{-rT}, \qquad (4.32)$$

where

$$\gamma = (\mathrm{h}^*\sigma)^2\left(\frac{1}{c} - \frac{1}{2c^2}\right) + \frac{h^*\mu}{c}. \qquad (4.33)$$

Thus $W_T = w\mathrm{e}^{rT}$ if $w = s\mathrm{e}^{-rT}$, and

$$W_T = s + \mathrm{e}^Z \qquad \text{if } w > s\mathrm{e}^{-rT}, \qquad (4.34)$$

where

$$Z = \ln(w\mathrm{e}^{rT} - s) + \gamma T - \frac{h^* X(T)}{c} \qquad (4.35)$$

is a normal random variable with

$$\mathrm{E}[Z] = \ln(w\mathrm{e}^{rT} - s) + (h^*\sigma)^2 T\left(\frac{1}{c} - \frac{1}{2c^2}\right) \qquad (4.36)$$

and

$$\mathrm{Var}[Z] = \left(\frac{h^*\sigma}{c}\right)^2 T. \qquad (4.37)$$

By using

$$\mathrm{E}[\mathrm{e}^Z] = \exp(\mathrm{E}[Z] + \frac{1}{2}\mathrm{Var}[Z])$$

and

$$\mathrm{Var}[\mathrm{e}^Z] = \mathrm{e}^{2\mathrm{E}[Z]+\mathrm{Var}[Z]}(\mathrm{e}^{\mathrm{Var}[Z]} - 1)$$

(Aitchison and Brown 1957), we find that

$$\mathrm{E}[W_T] = s + (w\mathrm{e}^{rT} - s)\exp\left[\frac{(h^*\sigma)^2 T}{c}\right], \quad w > s\mathrm{e}^{-rT}, \qquad (4.38)$$

and

$$\mathrm{Var}[W_T] = (w\mathrm{e}^{rT} - s)^2\left[\exp\left(\frac{3(h^*\sigma)^2 T}{c}\right) - \exp\left(\frac{2(h^*\sigma)^2 T}{c}\right)\right], \quad w > s\mathrm{e}^{-rT}. \qquad (4.39)$$

It follows from (4.38) and (4.39) that, as the minimal required wealth level s increases from 0 to $w\mathrm{e}^{rT}$, the mean of the optimal terminal wealth decreases linearly to $w\mathrm{e}^{rT}$ and the variance decreases quadratically to zero. Likewise, the mean and the variance of the optimal terminal wealth are decreasing functions of the parameter c. These monotonicity properties reflect the fact that the risk tolerance function $\tau(x)$, $x > s$, as given by (4.20), is a decreasing function in each of the parameters s and c.

For $0 < p < 1$, let ζ_p denote the $100p$-th percentile of the standard normal distribution. Then the $100p$-th percentile of the distribution of W_T is

$$s + \exp(\mathrm{E}[Z] + \zeta_p\sqrt{\mathrm{Var}[Z]}),\ w > s\mathrm{e}^{-rT}. \qquad (4.40)$$

In particular, the median of the distribution of W_T is

$$\mathrm{Mdn}(W_T)$$
$$= s + \mathrm{e}^{\mathrm{E}[Z]}$$
$$= s + (w\mathrm{e}^{rT} - s)\exp\left[(h^*\sigma)^2 T\left(\frac{1}{c} - \frac{1}{2c^2}\right)\right],\ w > s\mathrm{e}^{-rT}. \qquad (4.41)$$

Let us also examine Mode (W_T), the mode of the distribution of W_T. It follows from (4.34) and the formula for the mode of a lognormal distribution (Aitchison and Brown 1957, p. 15) that

$$\mathrm{Mode}(W_T) = s + \mathrm{e}^{\mathrm{E}(Z) - \mathrm{Var}(Z)}. \qquad (4.42)$$

Then, substituting according to (4.36) and (4.37), we obtain the formula

$$\text{Mode}(W_T) = s + (we^{rT} - s) \exp\left[(h^*\sigma)^2 T \left(\frac{1}{c} - \frac{3}{2c^2}\right)\right], \; w > se^{-rT}, \quad (4.43)$$

which is similar to (4.38) and (4.41). Comparing these three formulas, we see that, for $w > se^{-rT}$,

$$\text{Mode}(W_T) < \text{Mdn}(W_T) < we^{rT} < \text{E}[W_T], \; 0 < c < \tfrac{1}{2}$$

$$\text{Mode}(W_T) < we^{rT} < \text{Mdn}(W_T) < \text{E}[W_T], \quad \tfrac{1}{2} < c < \tfrac{3}{2}$$

$$we^{rT} < \text{Mode}(W_T) < \text{Mdn}(W_T) < \text{E}[W_T], \quad \tfrac{3}{2} < c.$$

As s increases to we^{rT}, W_T converges to the degenerate random variable with the constant value we^{rT}. We have already noted that the mean of W_T is always a decreasing function of s. In contrast, the median is an increasing function of s, if $c < 1/2$ (and a decreasing function of s, if $c > 1/2$), and the mode is an increasing function of s, if $c < 3/2$ (and a decreasing function of s, if $c > 3/2$).

We have noted that the mean of W_T is a decreasing function of c. This is in contrast to the median and the mode: by differentiating (4.41) and (4.43) with respect to c, we find that the median attains its maximum for $c = 1$, while the mode has its maximum for $c = 3$.

Remarks

(i) The quantity $(h^*\sigma)^2$, which appears frequently in this subsection, is the square of the "market price of risk" mentioned at the end of Subsection 2.1.

(ii) If $c = -h^* > 0$, then it follows from (4.33) and (2.9) that

$$\gamma = (h^*\sigma)^2 \left(\frac{-1}{h^*} - \frac{1}{2h^{*2}}\right) - \mu = -r, \quad (4.44)$$

and hence (4.32) becomes

$$W_T = s + (w - se^{-rT})e^{X(T)}. \quad (4.45)$$

It follows from (4.45) that an investor with this utility function can obtain his optimal terminal wealth simply by investing at time 0 the amount se^{-rT} in the riskless asset and his remaining initial wealth, $w - se^{-rT}$, in the risky asset. We shall return to this "buy-and-hold" investment strategy in Subsection 6.4, after we discuss dynamic investment strategies.

5. UTILITY OF THE INITIAL WEALTH

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

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

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 the function m defined by (3.7):

$$u_0'(w) = me^{rT}. \quad (5.2)$$

Equivalently,

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

To prove (5.2), we apply (3.10) to rewrite (5.1) as

$$u_0(w) = \text{E}[u(v(m\Psi))]. \quad (5.4)$$

By the chain rule,

$$\begin{aligned} u_0'(w) &= \frac{d}{dm}\,\text{E}[u(v(m\Psi))]\,\frac{dm}{dw} \\ &= \text{E}[u'(W_T)v'(m\Psi)\Psi] \Big/ \frac{dw}{dm}. \end{aligned} \quad (5.5)$$

It follows from (3.8) that

$$\text{E}[u'(W_T)v'(m\Psi)\Psi] = m\,\text{E}[v'(m\Psi)\Psi^2]. \quad (5.6)$$

By (3.11),

$$\frac{dw}{dm} = e^{-rT}\text{E}[v'(m\Psi)\Psi^2]. \quad (5.7)$$

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

Looking at (5.3), one might wonder if there is also an equality between the *random variables* $d/dw\; u(W_T)$ and $u'(W_T)e^{rT}$ in some situations—in other words, whether $d/dw\; W_T$ can be e^{rT}. By inspecting (4.8), (4.15), and (4.23), we conclude that this is true for the exponential utility functions, but not valid for the power utility functions.

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):

$$u_0''(w) = e^{rT} \frac{dm}{dw} = e^{rT} \Big/ \frac{dw}{dm}. \qquad (5.8)$$

Applying (5.7) to (5.8) yields

$$u_0''(w) = \frac{e^{2rT}}{E[v'(m\Psi)\Psi^2]}. \qquad (5.9)$$

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

Let

$$\tau_0(w) = -\frac{u_0'(w)}{u_0''(w)} \qquad (5.10)$$

be the risk tolerance function that is associated with the derived utility function $u_0(w)$. It follows from (5.2) and (5.8) that

$$\tau_0(w) = -m \frac{dw}{dm}. \qquad (5.11)$$

We shall come back to this formula in Section 6.

It turns out that, for the LRT utility functions,

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

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 $u'(we^{rT})$; that is, the ratio, $u_0'(w)/u'(we^{rT})$, does not depend on w. In view of (5.2), it is sufficient to check that the function $m(w)$ 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 results in the 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. We now consider a market in which securities can be traded at all times t, $0 \le t \le T$. Because it is assumed that the market is frictionless and there are no taxes, tradings will not add any cost. 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. Our goal is to determine this optimal dynamic investment strategy in such a securities market.

In this section, we consider the model of a single primitive risky asset that was discussed in Subsection 2.1. This is a model of a complete securities market. It follows from (3.10) and (4.26) that the optimal terminal wealth is

$$W_T = v(m\,\Psi) = v(m\, e^{h^*X(T)-\alpha T}). \qquad (6.1)$$

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 $S(t) = s$, $0 \le t < T$,

$$V(s, t) = e^{-r(T-t)}E[\pi(S(T))\,\Psi | S(t) = s]. \qquad (6.2)$$

Because the market is complete, the contingent claim can be replicated by a dynamic, self-financing portfolio. For $0 \le t \le T$, 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

$$\eta(s, t) = s \frac{\partial}{\partial s} V(s, t). \qquad (6.3)$$

The partial derivative $\partial/\partial s\, V(s, t)$ is called *delta* in the option-pricing literature.

Formula (6.3) is not directly applicable to determine the replicating portfolio of W_T. However, it can be applied to a related contingent claim defined by

$$\begin{aligned}\pi(S(T)) &= v(me^{-\alpha T}S(T)^{h^*}) \\ &= v(mS(0)^{h^*}\,\Psi). \qquad (6.4)\end{aligned}$$

Rewriting (3.11) as

$$e^{-rT}E[v(m\,\Psi)\,\Psi] = w = w(m), \qquad (6.5)$$

we see that the time-0 price of (6.4) must be

$$V(S(0), 0) = w(m\,S(0)^{h^*}). \qquad (6.6)$$

According to (6.3) and the chain rule, we find that

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

for the replicating portfolio of (6.4).

We are looking for 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 $W_t - \rho(W_t, t)$.) We note that the right-hand side of (6.4) reduces to that of (6.1), if we set $S(0) = 1$. Hence

$$\begin{aligned}\rho(w, 0) &= \eta(1, 0)\\ &= h^* \, m \, w'(m). \qquad (6.8)\end{aligned}$$

It now follows from (5.11) and (2.12) that

$$\rho(w, 0) = -h^*\tau_0(w) \qquad (6.9)$$

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

The extension of these formulas to time t, $0 < t < T$, is evident. Let $u_t(y)$ denote the conditional expected utility of optimal terminal wealth, given that $W_t = y$, and let $\tau_t(.)$ be the corresponding risk tolerance function. Hence the random variable $\tau_t(W_t)$ is the implied risk tolerance at time t, and we shall call it the *current risk tolerance*. Generalizing (6.9) and (6.10), we have

$$\rho(W_t, t) = -h^*\tau_t(W_t) \qquad (6.11)$$

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

Thus we have the simple investment rule that we 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.*

Let us also discuss the optimal investment in the risky asset as a fraction of the investor's current wealth at time t, W_t. Hence we define

$$M(W_t, t) = \frac{\rho(W_t, t)}{W_t}. \qquad (6.13)$$

We call (6.13) the *Merton ratio,* in honor of the Nobel laureate Robert C. Merton. By (6.8) we see that

$$M(w, 0) = \frac{h^* m \, w'(m)}{w}. \qquad (6.14)$$

This formula can be restated more elegantly in terms of the *elasticity* between the initial wealth w and the expected marginal utility of optimal terminal wealth m; see Section 8.

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

$$\begin{aligned}\rho(w, 0) &= -h^* e^{-rT}\tau(we^{rT})\\ &= \frac{\mu - \mu^*}{\sigma^2} e^{-rT}\tau(we^{rT}); \qquad (6.15)\end{aligned}$$

from this and (6.13) we obtain

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

Hence for the LRT utility functions, some very explicit formulas can be obtained, as we now show.

6.1 Constant Risk Tolerance Utility Functions

It follows from (4.5) that

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

and

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

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

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

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 his risk tolerance function is constant.

6.2 Decreasing LRT Utility Functions

It follows from (4.12) that

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

and

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

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

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

which says that the optimal amount invested in the risky asset is proportional to what the current wealth is short of the discounted level of saturation.

6.3 Increasing LRT Utility Functions

It follows from (4.20) that

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

and

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

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

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

Hence the optimal 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 be constrained to a terminal wealth of at least s, $0 < s \leq w\, e^{rT}$. Then $se^{-r(T-t)}$ is the portion of current wealth W_t that grows to s at time T with certainty. The complement, $W_t - se^{-r(T-t)}$, is the portion of current wealth that can be considered "free." Thus (6.25) tells us that, at any time, a constant proportion of the "free" wealth should be invested in the risky asset.

In the special case $s = 0$, (6.25) reduces to

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

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, $\mu - \mu^*$, 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%, which is the figure mentioned at the beginning of the Abstract and at the end of Section 1.

6.4 Optimality of Buy-and-Hold Strategies

Formula (4.45) implies that an investor, who has an increasing LRT utility function with $c = -h^*$, uses a buy-and-hold investment strategy. Here we show that the converse also holds. Consider an investor for whom a buy-and-hold investment strategy is optimal. At time 0, he invests the amount a, $a \leq w$, in the riskless asset and the complement, $w - a$, in the risky asset; he finds it optimal not to rebalance his portfolio. Hence

$$W_t = ae^{rt} + \rho(W_t, t), \qquad 0 \leq t < T. \quad (6.27)$$

We shall show that the investor has an increasing LRT utility function with $s = ae^{rT}$ and $c = -h^*$.

Because of (6.27), (6.11) is now

$$W_t - ae^{rt} = -h^* \tau_t(W_t), \qquad 0 \leq t < T. \quad (6.28)$$

In the limit $t \to T$, (6.28) yields

$$W_T - ae^{rT} = -h^* \tau(W_T).$$

We note that W_T can take on any value between ae^{rT} and ∞. Hence it follows that

$$\tau(x) = \frac{x - ae^{rT}}{-h^*} \qquad \text{for } x \geq ae^{rT}. \quad (6.29)$$

Comparing (6.29) with (4.20) proves our claim.

Remark

The buy-and-hold investment strategies are very special and can be viewed in a more general framework. In Subsections 6.1 to 6.3 we have seen that, for the LRT utility functions, the optimal amount invested in the risky asset at any time t is a linear function of the investor's current wealth W_t. Conversely, consider an investor whose optimal amount in the risky asset at any time is a linear function of his current wealth, that is, there exist two functions $a(t)$ and $b(t)$ such that

$$\rho(W_t, t) = a(t) + b(t)W_t, \qquad 0 \leq t < T.$$

Furthermore, assume that, as $t \to T$, the function $a(t)$ has a limit which we denote as a $a(T)$, and $b(t)$ has limit $b(T)$. Then it follows from (6.11) that

$$\tau(x) = \frac{a(T) + b(T)x}{-h},$$

which means that the investor has an LRT utility function.

7. OPTIMAL DYNAMIC INVESTMENT STRATEGIES—MULTIPLE RISKY ASSETS

We are now ready to generalize the results in the last section to the case of multiple primitive risky assets (the model in Subsection 2.2), which is more relevant in practice. Here, the price density can be written as

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

(As in the last section, the value of α is not relevant in the following; it is the number such that $E[\Psi] = 1$.)

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 the amount invested in risky asset k in the replicating portfolio of W_T at time t. Generalizing (6.13), let

$$M_k(W_t, t) = \frac{\rho_k(W_t, t)}{W_t} \tag{7.2}$$

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

Formula (6.3) can be generalized as follows. Consider a contingent claim with payoff

$$\pi(S_1(T), S_2(T), \ldots, S_n(T)) \tag{7.3}$$

at time T for a 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 (7.3), given that $S_j(t) = s_j$, $j = 1, 2, \ldots, n$. Then it is known (Gerber and Shiu 1996, Formula 8.35) that

$$\eta_k = s_k \frac{\partial V}{\partial s_k}. \tag{7.4}$$

Let us consider the payoff

$$\begin{aligned}
&\pi(S_1(T), S_2(T), \ldots, S_n(T)) \\
&= v\left(m\, e^{-\alpha T} \prod_{j=1}^{n} S_j(T)^{h_j^*}\right) \\
&= v\left(m \left[\prod_{j=1}^{n} S_j(0)^{h_j^*}\right] \Psi\right).
\end{aligned} \tag{7.5}$$

Because $w = w(m)$ is the time-0 price of $W_T = v(m\,\Psi)$, the time-0 price of (7.5) is

$$V(s_1, s_2, \ldots, s_n, 0) = w\left(m \prod_{j=1}^{n} s_j^{h_j^*}\right), \tag{7.6}$$

where $s_j = S_j(0)$, $j = 1, 2, \ldots, n$. Hence, by (7.6) and the chain rule,

$$\begin{aligned}
&\eta_k(s_1, s_2, \ldots, s_n, 0) \\
&= w'\left(m \prod_{j=1}^{n} s_j^{h_j^*}\right) m\, h_k^* \prod_{j=1}^{n} s_j^{h_j^*}.
\end{aligned} \tag{7.7}$$

Because

$$\rho_k(w, 0) = \eta_k(1, 1, \ldots, 1, 0),$$

it follows from (7.7) and (5.11) that

$$\rho_k(w, 0) = -h_k^* \tau_0(w). \tag{7.8}$$

Hence

$$\rho_k(W_t, t) = -h_k^*\, \tau_t(W_t), \tag{7.9}$$

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

$$\frac{\rho_k(W_t, t)}{\sum_{j=1}^{n} \rho_j(W_t, t)} = \frac{h_k^*}{\sum_{j=1}^{n} h_j^*}, \tag{7.10}$$

which depends only on the risk-neutral Esscher parameters and remains 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 a risk-free bond fund, and the other 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 current risk tolerance is always kept at the constant level

$$-\sum_{j=1}^{n} h_j^*. \qquad (7.11)$$

Remarks

(i) The continuously compounded rate of return over the time interval $(0, t)$ of the risky-asset mutual fund is

$$X(t) = \sum_{k=1}^{n} q_k X_k(t). \qquad (7.12)$$

The stochastic process $\{X(t)\}$ is a Wiener process with drift and diffusion parameters

$$\mu = \sum_{k=1}^{n} q_k \mu_k \qquad (7.13)$$

and

$$\sigma^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} q_i \sigma_{ij} q_j. \qquad (7.14)$$

(ii) An elegant geometric derivation of the "mutual fund" theorem for the classical one-period model can be found in Section 6.9 of Luenberger (1998). In the one-period model, the relative weights of the risky assets in the risky mutual fund are proportional to the v_i's that are determined by Equation (6.10) of Luenberger (1998); the v_i's should be compared to the risk-neutral Esscher parameters h_i^*'s, which are determined by Equation (2.22). However, note the differences between these two equations: our (2.22) involves the covariances between $X_i(1)$ and $X_j(1)$, while Luenberger's (6.10) is in terms of the covariances between $\exp[X_i(1)]$ and $\exp[X_j(1)]$; furthermore, in (2.22) there is a difference of instantaneous rates of return, while in (6.10) the difference is of annual rates.

8. ELASTICITY

Many actuaries would have encountered the concept of *elasticity* in an economics course. From (6.14) we saw that the Merton ratio is the product of the risk-neutral Esscher parameter h^* with

$$\frac{m\, w'(m)}{w}, \qquad (8.1)$$

which turns out to be the elasticity of the initial wealth w with respect to the expected marginal utility of optimal terminal wealth m.

Here is a brief review of elasticity. Consider a differentiable function $y = f(x)$. Its derivative $y' = dy/dx$ is the limit of the ratio of the *absolute changes,*

$$\frac{\Delta y}{\Delta x}. \qquad (8.2)$$

In economic applications it is often more meaningful to discuss the ratio of the *relative changes,*

$$\left(\frac{\Delta y}{y}\right) \Big/ \left(\frac{\Delta x}{x}\right). \qquad (8.3)$$

Its limit $(\Delta x \to 0)$ is called the elasticity of y with respect to x and denoted by the symbol $\varepsilon_x(y)$. From (8.3), we find the formula

$$\varepsilon_x(y) = \frac{xy'}{y} = x(\ln y)'. \qquad (8.4)$$

Rewriting (8.4) as

$$\varepsilon_x(y) = \frac{y'}{y/x}, \qquad (8.5)$$

we can interpret the elasticity as the ratio of the slope of the tangent line at (x, y) divided by the slope of the ray that connects the origin with (x, y). Furthermore, interchanging the roles of x and y in (8.3), and taking the limit, we see that

$$\varepsilon_y(x) = \frac{1}{\varepsilon_x(y)}, \qquad (8.6)$$

that is, the elasticity of the inverse function is the reciprocal of the original elasticity.

From (6.14), (8.4), and (8.5) a key result follows:

$$M(w, 0) = h^* \varepsilon_m(w) = \frac{h^*}{\varepsilon_w(m)}. \qquad (8.7)$$

Similarly, it follows from (7.8) that

$$M_k(w, 0) = h_k^* \varepsilon_m(w) = \frac{h_k^*}{\varepsilon_w(m)}, \quad k = 1, 2, \ldots, n. \qquad (8.8)$$

In the case of LRT utility functions, there are explicit expressions for the elasticities. For the constant risk tolerance utility functions, we obtain from (8.7) and (6.18) the formula

$$\varepsilon_w(m) = -a\, w\, e^{rT}. \qquad (8.9)$$

For the decreasing LRT utility functions, we have from (8.7) and (6.21):

$$\varepsilon_w(m) = \frac{-cw}{se^{-rT} - w}, \qquad w < se^{-rT}. \qquad (8.10)$$

For the increasing LRT utility functions, we have from (8.7) and (6.24):

$$\varepsilon_w(m) = \frac{-cw}{w - se^{-rT}}, \qquad w > se^{-rT}; \tag{8.11}$$

in particular, with $s = 0$, (8.11) simplifies as

$$\varepsilon_w(m) = -c. \tag{8.12}$$

We end this section with two remarks on elasticity in actuarial science. The *Macaulay duration* of a fixed-income security in immunization theory is the elasticity of the price of the security with respect to the discount factor v (Panjer et al. 1998, p. 101). A recent paper (Lemaire 1998) in this journal has a subsection entitled "the elasticity of the mean stationary premium with respect to the claim frequency." If we assume that the severity is independent of the claim frequency, the theoretical premium is proportional to the claim frequency. Hence its elasticity with respect to the claim frequency is the constant 1. Therefore, the efficiency of a Bonus-Malus system can be measured by how close the elasticity of the mean stationary premium with respect to the claim frequency is to the constant 1.

9. 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 elasticity. For example, the optimal 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 risky assets is constant in time. 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 the books cited in Section 1 and research papers such as Bajeux-Besnainou and Portait (1998), Boyle and Lin (1997), Boyle and Yang (1997), Browne (1998, 1999), Cairns (2000), Cox and Huang (1989), Dybvig (1988), Huang and Zariphopoulou (1999), Karatzas, Lehoczky, Sethi, and Shreve (1986), Khanna and Kulldorff (1999), and Pliska (1986), and their references. Also, many articles on optimal investment strategies have recently been published in the journal *Mathematical Finance*. Finally, a referee has kindly pointed out to us that human capital has important implications for deciding asset allocation in a defined-contribution plan and that a seminal paper in this direction is Bodie, Merton, and Samuelson (1992).

ACKNOWLEDGMENT

We thank the anonymous referees for their valuable comments. Elias Shiu gratefully acknowledges the support from the Principal Financial Group Foundation. The 1999 Edward A. Lew Award has provided us much impetus for developing this paper. We thank the Committee on Knowledge Extension Research for this great honor. We have always admired Ed Lew for his long-time contributions to actuarial research efforts.

REFERENCES

AITCHISON, J., AND BROWN, J.A.C. 1957. *The Lognormal Distribution: With Special Reference to Its Use in Economics.* Cambridge, U.K.: Cambridge University Press.

BAJEUX-BESNAINOU, I., AND PORTAIT, R. 1998. "Dynamic Asset Allocation in a Mean-Variance Framework," *Management Science* 44(11):S79–S94.

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

BICKSLER, J.L., AND SAMUELSON, P.A., EDS. 1974. *Investment Portfolio Decision-Making.* Lexington, Mass.: Lexington Books.

BINGHAM, N.H., AND KIESEL, R. 1998. *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives.* London: Springer.

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

BODIE, Z., MERTON, R.C., AND SAMUELSON, W.F. 1992. "Labor Supply Flexibility and Portfolio Choice in a Life Cycle Model," *Journal of Economic Dynamics and Control* 16: 427–49.

BOYLE, P.P., AND LIN, X. 1997. "Optimal Portfolio Selection with Transaction Costs," *North American Actuarial Journal* 1 (2):27–39.

BOYLE, P.P., AND YANG, H. 1997. "Asset Allocation with Time Variation in Expected Returns," *Insurance: Mathematics and Economics* 21:201–18.

BROWNE, S. 1998. "The Return on Investment from Proportional Portfolio Strategies," *Advances in Applied Probability* 30: 216–38.

BROWNE, S. 1999. "Beating a Moving Target: Optimal Portfolio Strategies for Outperforming a Stochastic Benchmark," *Finance and Stochastics* 3:275–94.

CAIRNS, A.G. 2000. "Some Notes on the Dynamics and Optimal Control of Stochastic Pension Fund Models in Continuous Time," to appear in *ASTIN BULLETIN 30*.

CONSTANTINIDES, G.M., AND MALLARIAS, A.G. 1995. "Portfolio Theory," in *Handbooks in Operations Research and Management Science,* Vol. 9, *Finance,* edited by R.A. Jarrow, V. Maksimovic, and W.T. Ziemba, 1–30. Amsterdam: Elsevier.

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

DAVIS, M.H.A. 1997. "Option Pricing in Incomplete Markets," in *Mathematics of Derivative Securities,* edited by M.A.H. Dempster and S. Pliska, 216–26. Cambridge: Cambridge University Press.

DEMPSTER, M.A.H., AND PLISKA, S., EDS. 1997. *Mathematics of Derivative Securities.* Cambridge: Cambridge University Press.

DEPREZ, O., AND GERBER, H.U. 1985. "On Convex Principles of Premium Calculation," *Insurance: Mathematics and Economics* 4:179–89.

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

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

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

EMBRECHTS, P., AND MEISTER, S. 1997. "Pricing Insurance Derivatives: The Case of CAT Futures," in *Securization of Insurance Risk: The 1995 Bowles Symposium,* SOA Monograph M-FI97-1, 15–26. Schaumburg, Ill.: Society of Actuaries.

ESSCHER, F. 1932. "On the Probability Function in the Collective Theory of Risk," *Skandinavisk Aktuarietidskrift* 15: 175–95.

GERBER, H.U., AND PAFUMI, G. 1988. "Utility Functions: From Risk Theory to Finance," *North American Actuarial Journal* 2 (3):74–100.

GERBER, H.U., AND SHIU, E.S.W. 1994. "Option Pricing by Esscher Transforms," *Transactions of the Society of Actuaries* 46:99–191.

GERBER, H.U., AND SHIU, E.S.W. 1996. "Actuarial Bridges to Dynamic Hedging and Option Pricing," *Insurance: Mathematics and Economics* 18:183–218.

GERBER, H.U., AND SHIU, E.S.W. 1997. "On Optimal Investment Strategies," *Rivista di Matematica per le Scienze Economiche e Sociali* 20:133–51.

HAKANSSON, N.H. 1987. "Portfolio Analysis," in *The New Palgrave: A Dictionary of Economics,* Vol. 3, edited by J. Eatwell, M. Milgate, and P. Newman, 917–20. London:; Macmillan.

HAKANSSON N.H., AND ZIEMBA, W.T. 1995. "Capital Growth Theory," in *Handbooks in Operations Research and Management Science,* Vol. 9, *Finance,* edited by R.A. Jarrow, V. Maksimovic, and W.T. Ziemba, 65–86. Amsterdam: Elsevier.

HUANG, C.-F., AND ZARIPHOPOULOU, T. 1999. "Turnpike Behavior of Long-Term Investments," *Finance and Stochastics* 3: 15–34.

HULL, J. 1993. *Options, Futures, and Other Derivative Securities,* 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.

KARATZAS, I., LEHOCZKY, J.P., SETHI, S.P., AND SHREVE, S.E. 1986. "Explicit Solution of a General Consumption/Investment Problem," *Mathematics of Operations Research* 11:261–94.

KARATZAS, I., AND SHREVE, S.E. 1998. *Methods of Mathematical Finance.* New York: Springer.

KHANNA, A., AND KULLDORFF, M. 1999. "A Generalization of the Mutual Fund Theorem," *Finance and Stochastics* 3:167–85.

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

LEMAIRE, J. 1998. "Bonus-Malus Systems: The European and Asian Approach to Merit-Rating," *North American Actuarial Journal* 2 (1):26–47.

LUENBERGER, D.G. 1998. *Investment Science.* New York: Oxford University Press.

MALLIARIS, A.G., AND BROCK, W.A. 1982. *Stochastic Methods in Economics and Finance.* Amsterdam: North-Holland.

MERTON, R.C. 1969. "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case," *Review of Economics and Statistics* 51:247–57.

MERTON, R.C. 1990. *Continuous-Time Finance.* Oxford: Blackwell.

MITCHELL, O.S., AND SCHIEBER, S.J., EDS. 1998. *Living with Defined Contribution Pensions: Remaking Responsibility for Retirement.* Philadelphia: University of Pennsylvania Press.

MOSSIN, J. 1968. "Optimal Multiperiod Portfolio Policies," *Journal of Business* 41:215–29.

PANJER, H.H. ED., ET AL. 1998. *Financial Economics: With Applications to Investments, Insurance, and Pensions.* Schaumburg, Ill.: The Actuarial Foundation.

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

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

SAMUELSON, P.A. 1969. "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics* 51:239–46.

SAMUELSON, P.A. 1990. "Asset Allocation Could Be Dangerous to Your Health," *Journal of Portfolio Management* (Spring):5–8.

SAMUELSON, P.A. 1997. "How Best to Flip-Flop if You Must: Integer Dynamic Programming for Either-Or," *Journal of Risk and Uncertainty* 15:183–90.

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

SHARPE, W.F. 1994. "The Sharpe Ratio," *Journal of Portfolio Management* (Fall):49–58.

SZEGÖ, G.P., AND SHELL, K., EDS. 1972. *Mathematical Methods in Investment and Finance:* Amsterdam: North-Holland.

ZIEMBA, W.T., AND VICKSON, R.G., EDS. 1975. *Stochastic Optimization Models in Finance*. New York: Academic Press.

DISCUSSIONS

PHELIM P. BOYLE*

The asset allocation decision is important for investment managers of insurance companies and pension plans. The optimal decision will depend on the nature of the underlying assets and liabilities as well as the objective of the decision maker. In 1969 Robert Merton provided an elegant solution to this problem using stochastic dynamic programming. Merton imposed strong assumptions on the stock return distribution and the decision maker's preferences to obtain tractable solutions. Hans Gerber and Elias Shiu have provided us with a much more accessible route to one of Merton's key results. The Gerber-Shiu path permits us to travel lighter, because we do not need to carry as much heavy technical equipment. This is a well-written paper that derives a classic result in a user-friendly fashion. It is a pleasure to read.

The Gerber-Shiu paper relies on a number of fundamental ideas in modern financial economics. It is gratifying to see these ideas gradually becoming more popular in actuarial literature. Their paper deals with concepts such as no arbitrage, complete markets, risk neutral valuation, and dynamic hedging. On the surface, it seems as if the natural probability measure, or P measure, is playing the leading role. However, the risk neutral measure, or Q measure, is doing a lot of the heavy work behind the scenes. The distinction between these two measures is not always clearly appreciated in actuarial circles.

Gerber and Shiu aim to provide a largely self-contained exposition on optimal capital growth and dynamic asset allocation and in this they succeed. They start with a simple model in which an investor seeks to maximize end-of-period expected utility. The available investments consist of a single stock with lognormal returns and a riskless bond that earns a constant rate r. When the utility function belongs to a certain class with the special property of linear risk tolerance, the optimal investment in the risky asset has a very simple form. The key idea is to first find the investor's optimal end-of-period wealth. Given the maintained assumptions of no arbitrage and market completeness we can find the portfolio of assets that perfectly replicates this optimal wealth. Gerber and Shiu formulate their main result in terms of the risk-neutral Esscher parameters. As they note in the remark at the end of Section 3 of the paper, the result can also be derived in an alternative way. This alternative method was developed by a number of authors, including Pliska (1982), Karatzas, Lehoczky and Shreve (1987) and Cox and Huang (1989). John Cox and Chi-fu Huang presented a preliminary version of the approach in a seminar at Carnegie Mellon University in 1983; their paper was subsequently published in the *Journal of Economic Theory* in 1989. There is no agreed-upon name for this approach. It is known by different names, including the risk-neutral computational approach, the martingale approach, and the Cox-Huang approach.

Gerber and Shiu concentrate on the Esscher approach to the problem, and they provide references to the alternative Cox-Huang approach. The Esscher approach has deep roots in classic risk theory, and it no doubt provides helpful intuition for some actuaries. However, it is also useful to analyze this problem in the Cox-Huang-Pliska (CHP) framework, because this is the approach used in the finance literature and it facilitates communication with nonactuaries if we speak their language. A clear introduction to the CHP approach in the discrete-time setting is provided in Pliska's (1997) textbook.

The CHP method uses two important concepts in modern financial economics: no arbitrage and complete markets. Their method provides a very powerful tool for finding the optimal investment strategy for an agent who values both consumption and final wealth. The classic approach to this problem is to find the investment strategy that maximizes the agent's expected utility. The CHP approach uses a direct method to find the optimal attainable consumption and terminal wealth. In this approach the no-arbitrage condition is brought right into the objective function. This can be seen from Equation (3.12) in the Gerber-Shiu paper, which reads:

$$L(W, \lambda) = \mathrm{E}_P[u(W)] - \lambda\{\mathrm{E}_P[W\Psi] - we^{rT}\}. \quad (1)$$

The expectations are taken under the P measure, which in this example corresponds to the decision maker's subjective probability measure. Since the

*Phelim P. Boyle, F.C.I.A., F.I.A., Ph.D., holds the J. Page Wadsworth Chair of Finance at the University of Waterloo, Waterloo, Ontario N2L 3G1, Canada, e-mail, pboyle@uwaterloo.ca.

state price, Ψ, enables us to switch from the P measure to the risk-neutral measure, the constraint can be written as

$$w = e^{-rT} E_Q[W]. \quad (2)$$

This last result exemplifies one of the key valuation results in modern finance. It states that the current price of a security equals its expected discounted value under the risk-neutral measure. If we define the money market account by $B(t) = \exp[rt]$, then Equation (2) can be written as

$$\frac{w}{B(0)} = E_Q\left[\frac{W(T)}{B(T)}\right].$$

Hence the normalized wealth process is a martingale under Q. (Recall that the time-zero wealth is equal to w.)

Knowing the optimal payoffs, we can find the trading strategy that replicates it, as we assume a complete market. Recall that we can replicate any derivative security in a complete market. This replicating portfolio yields the trading strategy that produces the desired payoff pattern. To show the existence of the replicating portfolio, we can use a powerful result known as the Martingale Representation Theorem. The textbook by Dothan provides an example of this result in a discrete-time setting. In continuous time the representation formula for the optimal portfolio was derived by Ocone and Karatzas (1991) using the Clark-Ocone formula. In this representation the optimal portfolio is expressed in terms of expectations of random variables involving the so-called Malliavin derivatives of the coefficients of the model. Until recently this result was of only theoretical interest. However Detemple, Garcia, and Rindisbacher (1999) have developed new techniques that enable us to obtain numerical solutions to a broad range of portfolio allocation problems.

This new approach of Detemple, Garcia, and Rindisbacher extends the range of applicability of the CHP approach. It means that the CHP methodology can be applied under much more general assumptions than those under which the Merton result is obtained. Recall that the classic result was derived using stationary lognormal stock returns and constant interest rates. We know that asset returns are not strictly lognormal and that the distribution is not stationary over time. Furthermore, interest rates are stochastic and the market price of risk also varies over time. Detemple, Garcia, and Rindisbacher have implemented their model and calibrated it to U.S. data for the 1965–1996 period. They use a one-factor Cox-Ingersoll-Ross model for the interest rate process and a mean reverting process for the market price of risk. They examine different assumptions with regard to the stock price process and estimate the parameters using historical data. The critical factors in the optimal asset allocation decision are the risk aversion and the investment horizon of the investor. Detemple, Garcia, and Rindisbacher state

> For example, our results show that in the short run, risk averse investors want to reduce their stock demand relative to logarithmic investors in order to hedge against the market price of risk, while in the long run they will increase it to hedge against the interest rate risk.

The results of the analysis carried out by Detemple, Garcia, and Rindisbacher differ from those of the classic Merton model. Nevertheless, it still makes sense to study models with strong and perhaps unrealistic assumptions that lead to elegant closed-form solutions. The Merton solution provides a useful benchmark. It also serves as a useful base camp from which we can tackle higher and more difficult peaks. Our thanks are due to Gerber and Shiu for stimulating interest in this important topic.

REFERENCES

COX, J., AND HUANG, C.-F. 1989. "Optimal Consumption and Portfolio Policies When Asset Prices Follow a Diffusion Process," *Journal of Economic Theory* 49:33–83.

DETEMPLE, JEROME, GARCIA, RENE, AND RINDISBACHER, MARCEL. 1999. "A Monte Carlo Method for Optimal Portfolios." Working Paper, Rotman School of Management, University of Toronto, Canada.

DOTHAN, M.U. 1990. *Prices in Financial Markets.* Oxford: Oxford University Press.

KARATZAS, I., LEHOCZKY, J. AND SHREVE, S. 1987. "Optimal Portfolio and Consumption Decisions for a Small Investor on a Finite Horizon," *SIAM Journal of Control and Optimization* 25:1157–86.

PLISKA, S.R. 1982. "A Discrete Time Stochastic Decision Model," in *Advances in Filtering and Optimal Stochastic Control,* edited by E.H. Fleming and L.G. Gorostiza, 290–304. Lecture Notes in Control and Information Sciences, 42. New York: Springer Verlag.

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

GÉRARD PAFUMI*

The purpose of this discussion is to illustrate Section 4.4 of the paper ("Lognormal Price Density"). In particular, I consider the case of an increasing LRT utility function (see Section 4.3). Under an increase of initial wealth, an investor with such a utility function increases the dollar amount invested in the risky asset, which is consistent with certain empirical observations. Another advantage of this type of utility function is that it allows explicit calculations about the distribution of the optimal terminal wealth W_T.

Increasing LRT utility functions are the power utility functions with two parameters c and s, $c > 0$, where $1/c$ is the slope of the risk tolerance function and s is the minimal requirement of the terminal wealth. Hence, if we assume an increasing LRT utility function and no transaction costs, we obtain a two-parameter family for the distribution of the optimal terminal wealth (see Formulas (4.32) through (4.35)). For illustrative purposes, we consider the three cases $c = 0.25, 1, 2$ combined with the three cases $s = 80, 90, 95$. We assume

$$\mu - \mu^* = 0.05,$$
$$\sigma = 0.25,$$
$$T = 4$$

and that

$$we^{4r} = 100.$$

This way s and W_T can be interpreted as percentages of the risk-free terminal wealth.

Figures 1, 2, and 3 show the probability density function of the optimal terminal wealth for $c = 0.25, 1, 2$, respectively. Remember that for increasing LRT utility functions, the risk tolerance function $\tau(x)$ is given by

$$\tau(x) = \frac{x - s}{c}, \qquad x > s.$$

Hence, for a fixed wealth level x, $\tau(x)$ is a decreasing function of both c and s. Observe that for a fixed value of s, the less risk tolerant an investor is (that is, the higher the value of c), the more symmetrical the distribution of optimal terminal wealth W_T, as can be seen by comparing Figures 1, 2, and 3. This can be verified by looking at Formulas (4.34) and (4.37),

Figure 1
P.D.F. of the Optimal Terminal Wealth for c = 0.25

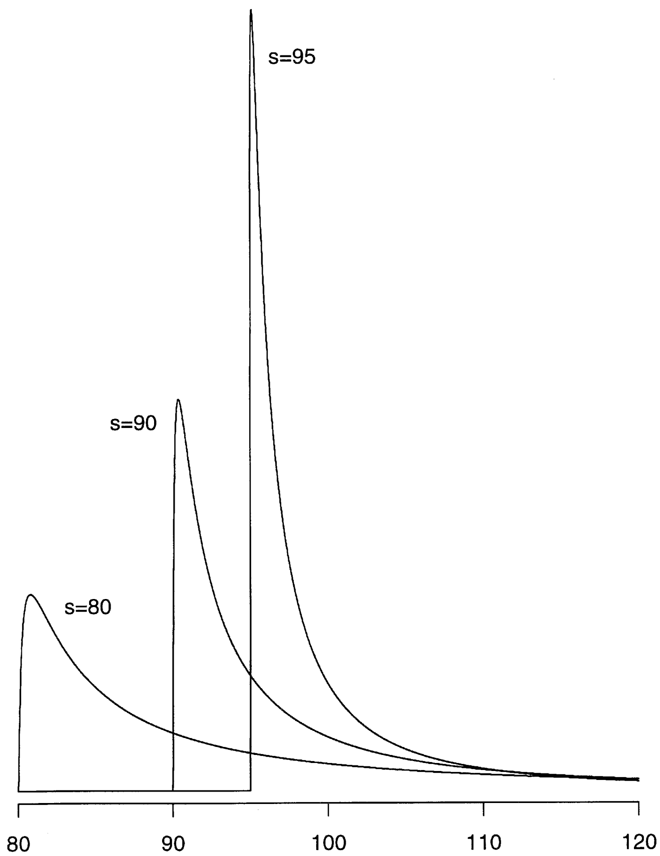
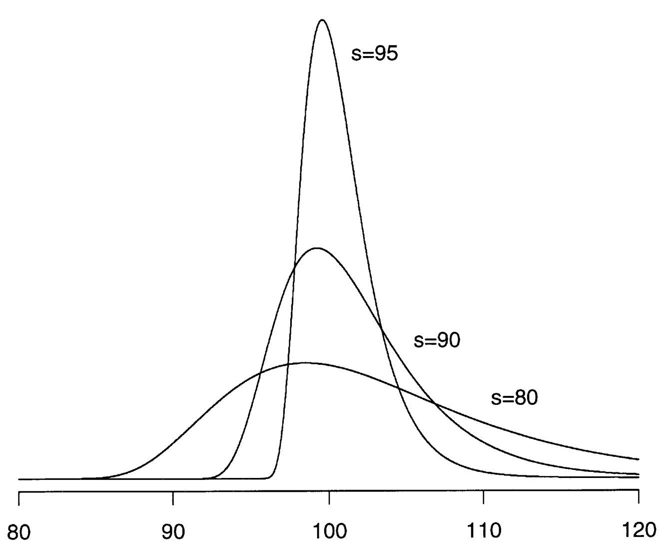
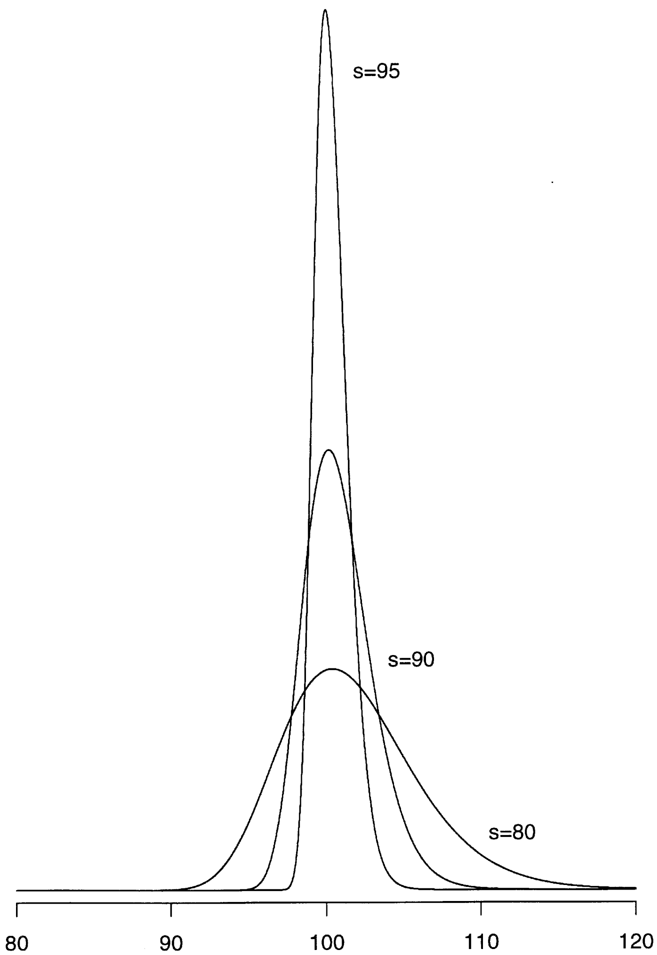


Figure 2
P.D.F. of the Optimal Terminal Wealth for c = 1



*Gérard Pafumi, Ph.D., is a risk analyst with a bank. He can be reached at Petit-St-Jean 5, CH-1003 Lausanne, Switzerland, e-mail, gpafumi@hec.unil.ch

Figure 3
P.D.F. of the Optimal Terminal Wealth for $c = 2$



which enable us to obtain the index of skewness of the distribution of W_T:

$$\gamma(W_T) = \frac{E[(W_T - E[W_T])^3]}{(\text{Var}[W_T])^{3/2}}$$

$$= \frac{e^{(h^*\sigma/c)^2T}(e^{2(h^*\sigma/c)^2T} - 3) + 2}{(e^{(h^*\sigma/c)^2T} - 1)^{3/2}}.$$

Table 1 exhibits the standard deviation and the index of skewness of the optimal terminal wealth.

The 5th and 95th percentile of the distribution of W_T are shown in Table 2. Observe that, for a given value of c, as s increases the difference between the 95th and the 5th percentile of the distribution of W_T diminishes. This illustrates the fact that as s increases to 100, W_T converges to the degenerate random variable with the constant value 100. Table 3 shows the mode, median, and mean of the distribution of W_T. One can verify the inequalities obtained by the authors for those characteristics in Section 4.4.

Consider now Formula (6.13) for the optimal investment in the risky asset as a fraction of the investor's current wealth at time t (Merton ratio). In the case of increasing LRT utility functions, the Merton ratio is given by Formula (6.25), which tells us that the optimal amount invested in the risky asset is proportional to the excess of current wealth over the discounted value of the minimal required terminal wealth. The proportionality factor is given by

$$\frac{\mu - \mu^*}{\sigma c}.$$

For the values of μ, σ, and T we consider that here this factor amounts to 0.4 for $c = 2$, 0.8 for $c = 1$, and 3.2 for $c = 0.25$. In this last case, in which the investor is the most risk tolerant, this means that the investor borrows money to invest it (in addition to its initial wealth w) in the risky investment.

AUTHORS' REPLY

HANS U. GERBER AND ELIAS S.W. SHIU

We are grateful to the discussants, who have provided valuable complements to our paper. Professor Boyle,

Table 1
Standard Deviation/Index of Skewness of the Distribution of the Optimal Terminal Wealth

$s\backslash c$	0.25	1	2
80	131.0/51.601	9.8/1.322	4.4/0.614
90	65.5/51.601	4.9/1.322	2.2/0.614
95	32.8/51.601	2.4/1.322	1.1/0.614

Table 2
The 5th and 95th Percentile of the Distribution of the Optimal Terminal Wealth

$s\backslash c$	0.25	1	2
80	80.8–226.6	91.2–131.8	95.3–109.5
90	90.4–163.3	95.6–110.9	97.6–104.8
95	95.2–131.6	97.8–105.5	98.8–102.4

Table 3
Mode/Median/Mean of the Distribution of the Optimal Terminal Wealth

$s\backslash c$	0.25	1	2
80	80.8/90.5/117.9	98.5/101.7/103.5	100.4/101.2/101.7
90	90.4/95.3/109.0	99.2/100.8/101.7	100.2/100.6/100.8
95	95.2/97.6/104.5	99.6/100.4/100.9	100.1/100.3/100.4

an authority in the field, has put our paper in an appropriate perspective. We agree with Professor Boyle who points out that a careful distinction should be made between the Q-measure (the risk-neutral or equivalent martingale measure) and the P-measure (the natural or physical measure). Some authors use the attributes *subjective* for the Q-measure and *objective* for the P-measure. This is an unfortunate terminology. Because the prices are the same for all agents, the underlying Q-measure is the same and it can be directly observed. Hence it merits the attribute *objective.* In contrast, ''the'' P-measure may very well vary between the different agents, depending on their information and beliefs. Thus it is the P-measure that deserves the attribute *subjective.* One could go even one philosophical step further and question the existence of a P-measure.

Dr. Pafumi provides an excellent illustration of what the choice is all about. In practice, a utility function may be an abstract concept for some investors. Alternatively, they could be confronted with the family of possible distributions of terminal wealth, which can be well explained by charts and figures such as in Dr. Pafumi's discussion. Based on this information, the investor may reach a conscious decision. Finally, it may be useful to point out that, for an insurer selling equity-indexed annuities, the value of s in an increasing LRT utility function can be set to be the minimum interest-rate guarantee required by the non-forfeiture law.

Additional discussions on this paper can be submitted until October 1, 2000. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.