Portfolio Selection in Stochastic Environments

Jun Liu
University of California, San Diego

In this article, I explicitly solve dynamic portfolio choice problems, up to the solution of an ordinary differential equation (ODE), when the asset returns are quadratic and the agent has a constant relative risk aversion (CRRA) coefficient. My solution includes as special cases many existing explicit solutions of dynamic portfolio choice problems. I also present three applications that are not in the literature. Application 1 is the bond portfolio selection problem when bond returns are described by "quadratic term structure models." Application 2 is the stock portfolio selection problem when stock return volatility is stochastic as in Heston model. Application 3 is a bond and stock portfolio selection problem when the interest rate is stochastic and stock returns display stochastic volatility. (*JEL* G11)

There is substantial evidence of time variation in interest rates, expected returns, and asset return volatilities. Interest rates change over time, and although expected stock returns are not directly observed, future stock returns seem to be predictable using term structure variables and scaled prices such as dividend yields.[1] Similarly, there is well-documented evidence of stochastic volatility,[2] whose existence is also supported by the "smile curve" of volatilities implied by option prices.

Therefore, any serious study of dynamic portfolio choice must take account of stochastic variation in investment opportunities. The seminal work of Merton (1971) establishes the framework for dynamic portfolio choice with stochastic variation in investment opportunities. The portfolio weights in Merton's framework are expressed in terms of the solution to a nonlinear partial differential equation (PDE), and because there is no closed-form solution of a nonlinear PDE in general, explicit portfolio weights are not available in general. There are approximate solutions to

I am grateful to Geert Bekaert, Darrell Duffie, and Ken Singleton for introducing me to portfolio choice problems and affine models, Andrew Ang and Jun Pan for extensive discussions, Michael Brennan, Campbell Harvey, the editor, John Hughes, and two anonymous referees for their comments and insights. I also thank Qiang Dai, Mark Ferguson, Mark Garmaise, Steven Grenadier, Harrison Hong, Ming Huang, and seminar participants at Barclays Global Investors, Columbia University, Massachusetts Institute of Technology, New York University, Stanford University, UCLA, University of Michigan, and University of Rochester for helpful comments. Address correspondence to Jun Liu, The Rady School of Management, UCSD, 9500 Gilman Dr, MC 0093, Pepper Canyon Hall, Room 320, San Diego, CA 92130-0093, or e-mail: junliu@ucsd.edu

[1] See for example, Fama and Schwert (1977) and Ferson and Harvey (1993).

[2] French, Schwert, and Stambaugh (1987) and Pagan and Schwert (1990) document stochastic volatility in asset returns.

doi:10.1093/rfs/hhl001 Advance Access publication May 15, 2006

Merton's problem such as the asymptotic solutions of Kogan and Uppal (2000) and log-linearization solution of Campbell and Viceira (2001) and Chacko and Viceira (2005). Cases for which exact solutions to Merton's problem have been obtained include Kim and Omberg (1996), Brennan (1998), Brennan and Xia (2000, 2001, 2002a,b), Sangvinatsos and Wachter (2005), and Wachter (2003). In all of these studies, asset returns are assumed to have constant volatility and the short rate and/or the risk premium follow Ornstein–Uhlenbeck processes, which allow these variables to take on negative values.

In this article, I derive explicit solutions of dynamic portfolio choice problems with what I term quadratic asset returns. In the context of dynamic portfolio choice, the investment opportunity set is described by four characteristics of the asset return dynamics: the short rate, the maximal squared-Sharpe ratio, the hedging covariance vector (the vector of covariances between the mean-variance efficient portfolio and the hedging portfolios), and the unspanned covariance matrix (the covariance matrix of the state variables that is not spanned by the covariance of the assets). Asset returns are "quadratic returns" when all four characteristics are quadratic functions of a "quadratic process," which is a Markovian diffusion processes whose drift and diffusion coefficients are quadratic functions of the processes themselves. Quadratic processes include both affine processes and the Ornstein–Uhlenbeck process as special cases. Quadratic processes have the same analytical tractability as, but are more flexible than, affine processes. Quadratic returns are common in the literature; examples include returns of zero-coupon bonds in affine term structure models [Duffie and Kan (1996)], stock returns with a risk premium that follows an Ornstein–Uhlenbeck process, and stock returns in Heston's (1993) model of stochastic volatility.

The main analytic results of this article are explicit solutions to dynamic portfolio choice problems when the asset returns are quadratic and the agent has a constant relative risk aversion (CRRA) utility function. The utility function is defined over intermediate consumption and terminal wealth when financial markets are complete and over terminal wealth only when financial markets are incomplete. This class of dynamic portfolio problems includes as special cases many of the dynamic portfolio problems for which explicit solutions have been derived previously. In addition, the class also accommodates models in which the short rate and/or the risk premium follow square-root processes that ensure that the short rate and/or the risk premium are positive, and models in which asset returns have stochastic volatility. I also prove a separation result, showing that the bond and stock portfolio selection problems can be decomposed into bonds-only and stock-only portfolio selection problems when the state variables that govern interest rates are independent of the state variables that govern only the stock returns.

I consider three applications. Application 1 is a bond portfolio selection problem. I first propose quadratic term structure models that nest both the quadratic gaussian term structure models [Constantinides (1992)] and the affine term structure models [Duffie and Kan (1996)] that have been used in many empirical and theoretical studies. The returns of the bonds are derived using the quadratic term structure model to rule out arbitrage opportunities between the bonds. The portfolio selection problem with these bond returns is then solved explicitly. The portfolio weight of a zero-coupon bond is derived in closed-form for the special case of the Cox–Ingersoll–Ross (CIR) term structure model. Application 2 is a stock portfolio choice problem when the stock return is described by Heston's (1993) stochastic volatility model. Application 3 is a portfolio choice problem with a stock and a bond in a stochastic interest rate–stochastic volatility model. The stock returns in the model have both interest rate risk as described by the CIR model and stochastic volatility as described by the Heston model. The bond and stock portfolio weights in this model are expressed in terms of the explicit solutions for the CIR and Heston models by an application of the separation result.

The explicit solutions allow us to show that dynamic portfolio weights have several properties that are different from static portfolio weights. First, the dynamic portfolio weight of a risky asset can be negative even if the risk premium is strictly positive. In the works of Kim and Omberg (1996) and Brennan and Xia (2001), dynamic portfolio weights can also be negative. However, one cannot determine whether the negative portfolio weight in their model is due to dynamic portfolio choice or the feature that the risk premium in their models is given by an Ornstein–Uhlenbeck process and thus can be negative. In Application 2, the risk premium is strictly positive, and hence, one can conclude that the negative portfolio weight is due to dynamic portfolio choice.

Second, the dynamic portfolio weight of a risky asset may not be decreasing in risk aversion even if the risk premium is strictly positive. Although Kogan and Uppal (2000) independently point out that a dynamic portfolio weight may increase with risk aversion, the example in this article is striking because the risk premium is strictly positive.

Third, I show that the ratio of bond to stock portfolio weights increases with risk aversion. This resolves an asset allocation puzzle identified by Canner, Mankiw, and Weil (1997). Brennan, Schwartz, and Lagnado (1997) were the first to study dynamic portfolio choice problems with stocks and bonds simultaneously. These problems have since been studied by Brennan and Xia (2000), Campbell and Viceira (2001), Wachter (2003), and Sangvinatsos and Wachter (2005) when the interest rate follows an Ornstein–Uhlenbeck process. Brennan and Xia (2000), Campbell and Viceira (2001), and Wachter (2003) also show that the puzzle of Canner, Mankiw, and Weil (1997) can be resolved in a dynamic portfolio choice context. The asset returns in their models are described by Ornstein–Uhlenbeck processes, whereas they are described by square-root processes in my examples.

Schroder and Skiadas (1999) characterize portfolio selection in complete markets under stochastic differential utility, which is the continuous-time analog of recursive utility and includes CRRA utility as a special case. They provide a general characterization of the solution to dynamic portfolio problems in the non-Markovian case, deriving explicit portfolio weights for CRRA utility when asset returns are affine and markets are complete. I obtain explicit solutions with CRRA utility over the terminal wealth for *incomplete* markets and when asset returns are quadratic, which, as mentioned earlier, include affine asset returns as a special case.

The remainder of the article is organized as follows. In Section 1, I set up the framework and introduce quadratic processes and quadratic returns. Then, assuming that the asset returns are quadratic, I explicitly derive the optimal consumption and portfolio weights when markets are complete and the optimal portfolio weights, assuming no intermediate consumption when markets are incomplete. In Section 2, I first show that under appropriate independence of state variables, bond and stock portfolio choice problems can be separated into bond-only and stock-only portfolio choice problems. I then define quadratic term structure models and explicitly solve bond portfolio selection problems when the term structure is quadratic. Next, I solve the stock portfolio problem when the stock return volatility is described by the Heston model. Finally, I derive in closed-form the stock and bond portfolio weights in a model with both stochastic interest rates and stochastic volatility, discuss properties of dynamic portfolio weights, and provide a potential resolution of the puzzle of Canner, Mankiw, and Weil. I conclude in Section 3. I leave calculation details to the Appendix.

1. Dynamic Portfolio Weights: General Results

1.1 The setup

I assume that the asset prices $P_t = (P_{0t}, P_{1t}, \ldots, P_{Mt})$ satisfy the following equations:

$$\frac{dP_{0t}}{P_{0t}} = r(X_t)dt,$$
$$\frac{dP_{it}}{P_{it}} = \mu_i(X_t)dt + \Sigma_i(X_t)dB_t,\ i = 1, \ldots, M,$$

where B_t is a standard M-dimensional Brownian motion. I further assume that $r(X_t)$, $\mu_i(X_t)$, and $\Sigma_i(X_t)$ are functions of an N-dimensional state variable vector, X_t, which follows a Markovian diffusion process,

$$dX_t = \mu^X dt + \Sigma^X dB_t^X, \tag{1}$$

where the drift coefficient (or "drift") μ^X and the diffusion coefficient[3] (or "diffusion") $\Sigma^X \Sigma^{X\top}$ are an $N \times 1$ vector function and an $N \times N$ matrix function of X_t, respectively, and B_t^X is a N-dimensional standard Brownian motion. Note that $\top$ denotes the transpose of a matrix.

I use $\mu(X_t)$ and $\Sigma(X_t)$ to denote $[\mu_1(X_t), \ldots, \mu_M(X_t)]$ and $[\Sigma_1(X_t), \ldots, \Sigma_M(X_t)]$, respectively. I assume that $\Sigma(X_t)$ is invertible almost surely. The correlation matrix between dB_t^X and dB_t is $\rho(X_t)dt$, where $\rho(X_t)$ is a matrix function of X_t with dimension $N \times M$. The state variable X could be the stochastic short rate as in the CIR term structure model, or predictors of stock returns such as dividend yields in predictability models, or the volatility process in the stochastic volatility models. I will later specify functional forms for $\mu^X(X)$, $\Sigma^X(X)$, $r(X)$, $\mu(X)$, $\Sigma(X)$, and $\rho(X)$ that allow for explicit solutions of the optimal consumption and portfolio rules. When there is no confusion, I will denote $\mu^X(X)$, $\Sigma^X(X)$, $r(X)$, $\mu(X)$, $\Sigma(X)$, and $\rho(X)$ by μ^X, Σ^X, r, μ, Σ, and ρ, respectively.

Following Merton (1971), I assume that (1) there are no transaction costs, taxes, or asset indivisibitility; (2) the agent is a price taker; (3) short sales of all assets, with full use of proceeds, are allowed; (4) and trading in assets takes place continuously in time.

The agent maximizes the following expected utility:

$$\max_{\{\phi_t\}_{t=0}^T, \{c_t\}_{t=0}^T} \mathrm{E}_0 \left[\int_0^T \alpha e^{-\beta t} \frac{C_t^{1-\gamma}}{1-\gamma} dt + (1-\alpha) e^{-\beta T} \frac{W_T^{1-\gamma}}{1-\gamma} \right], \tag{2}$$

where ϕ_t is the M-dimensional vector of the portfolio weights of risky assets, c_t is the consumption rate, and W_T is the value at T of a trading strategy that finances $\{c_t\}_{t=0}^T$,

$$dW_t = \left\{ W_t \left[\phi_t^\top (\mu - r) + r \right] - C_t \right\} dt + W_t \phi_t^\top \Sigma dB_t.$$

γ is the risk aversion coefficient (as well as being the inverse of the elasticity of intertemporal substitution) and β is the subjective discount

[3] Sometimes Σ^X instead of $\Sigma^X \Sigma^{X\top}$ is referred as the diffusion coefficient.

rate. The parameter α determines the relative importance of the intermediate consumption and the bequest. When $\alpha = 0$, expected utility only depends on the terminal wealth and the problem is called an asset allocation problem.

Following Merton (1971), I use the stochastic control approach to solve the problem. Let $J(t,W,X)$ denote indirect utility function. The principle of optimality leads to the following Hamilton–Jacobi–Bellman (HJB) equation [Merton (1971)] for J:

$$\begin{aligned}\max_{\phi,C}\bigg\{&\frac{\partial J}{\partial t}+\frac{1}{2}W^2\phi^\top\Sigma\Sigma^\top\phi J_{WW}+W\big[\phi^\top(\mu-r)+r\big]J_W-CJ_W\\&+W\phi^\top\Sigma\rho^\top\Sigma^{X\top}J_{WX}+\frac{1}{2}\mathrm{Tr}(\Sigma^X\Sigma^{X\top}J_{XX^\top})\\&+\mu^{X\top}J_X+\alpha e^{-\beta t}\frac{C^{1-\gamma}}{1-\gamma}\bigg\}=0,\end{aligned}\tag{3}$$

with boundary condition

$$J(T,W,X)=(1-\alpha)e^{-\beta T}\frac{W^{1-\gamma}}{1-\gamma},$$

where $\partial J/\partial t$, J_W, and J_X denote the derivatives of J with respect to t, W, and X, respectively. I use similar notation for higher derivatives and the derivatives of other functions.

J is conjectured to have the form:

$$J(t,W,X)=e^{-\beta t}\frac{W^{1-\gamma}}{1-\gamma}\big[f(X,t)\big]^\gamma.\tag{4}$$

Under this conjecture, the optimal consumption and the optimal portfolio weights are given by

$$C^*=\alpha^{\frac{1}{\gamma}}Wf^{-1},\tag{5}$$

$$\phi^*=\frac{1}{\gamma}(\Sigma\Sigma^\top)^{-1}\left[(\mu-r)+\gamma\Sigma\rho^\top\Sigma^{X\top}\frac{\partial\ln f}{\partial X}\right].\tag{6}$$

The first term in expression (6) of ϕ^*, $\frac{1}{\gamma}(\Sigma\Sigma^\top)^{-1}(\mu-r)$, is the vector of the mean-variance efficient portfolio weights. It is also called the myopic demand because this is the vector of portfolio weights for an agent who has only a single period objective or a very short investment horizon.

The second term in Equation (6) is the intertemporal hedging demand, which is determined by the covariance $\Sigma^{\top -1}\rho^{\top}\Sigma_X^{\top}$ and the indirect utility function J. The term $\Sigma^{\top -1}\rho^{\top}\Sigma_X^{\top}$ selects the portfolios that have the maximum correlation with the state variable X. The factor $\frac{\partial}{\partial X}\ln f$ measures the sensitivity of the indirect utility function to the opportunity set and summarizes the agent's attitude toward changes in the state variable X.

Substituting the conjecture of J and the resulting optimal policies into Equation (3), we obtain

$$\begin{aligned}&\frac{\partial f}{\partial t}+\frac{1}{2}\mathrm{Tr}\left(\Sigma^X\Sigma^{X\top}f_{XX^{\top}}\right)+\left[\mu^X+\frac{1-\gamma}{\gamma}\Sigma^X\rho\Sigma^{-1}(\mu-r)\right]^{\top}f_X\\&\quad+\frac{1}{2f}(\gamma-1)f_X^{\top}\left(\Sigma^X\Sigma^{X\top}-\Sigma^X\rho\rho^{\top}\Sigma^{X\top}\right)f_X\\&\quad+\left[\frac{1-\gamma}{2\gamma^2}(\mu-r)^{\top}(\Sigma\Sigma^{\top})^{-1}(\mu-r)+\frac{1-\gamma}{\gamma}r-\beta\right]f+\alpha^{\frac{1}{\gamma}}=0,\end{aligned}\tag{7}$$

with condition $f(T,X)=(1-\alpha)^{\frac{1}{\gamma}}$.

The left-hand side of the above PDE is, up to a multiplicative factor γJ, the instantaneous expected change in the indirect utility function $J(t,W,X)=\frac{W^{1-\gamma}}{1-\gamma}f^{\gamma}(t,X)$. By Ito's lemma, $\frac{\partial f}{\partial t}+\frac{1}{2}\mathrm{Tr}(\Sigma^X\Sigma^{X\top}f_{XX^{\top}})$ $+\mu^{X\top}f_X+\frac{1}{2f}(\gamma-1)f_X^{\top}\Sigma^X\Sigma^{X\top}f_X$ is the expected change of $f^{\gamma}(t,X)$. The $\alpha^{\frac{1}{\gamma}}$ and $-\beta f$ terms are due to intermediate consumption and subjective time discount, respectively. The term $\left[\frac{1-\gamma}{2\gamma^2}(\mu-r)^{\top}(\Sigma\Sigma^{\top})^{-1}(\mu-r)\right.$ $\left.+\frac{1-\gamma}{\gamma}r\right]f$ is due to the expected change in $W^{1-\gamma}$ from the return of the myopic component of the optimal portfolio. Note that $(\mu-r)^{\top}(\Sigma\Sigma^{\top})^{-1}(\mu-r)$ is the maximal squared-Sharpe ratio. The term $\frac{1-\gamma}{\gamma}\left[\Sigma^X\rho\Sigma^{-1}(\mu-r)\right]^{\top}f_X$ is due to the expected change in $W^{1-\gamma}$ from the return of the intertemporal hedging component of the optimal portfolio and the covariance between changes in $W^{1-\gamma}$ and in f^{γ}. Also note that the $N\times 1$ vector $\Sigma^X\rho\Sigma^{-1}(\mu-r)$ is the covariance between the shocks $\Sigma^X dB_t^X$ to the state variable X and the shock $[\Sigma^{-1}(\mu-r)]^{\top}dB_t$ to the return on the myopic component of the optimal portfolio; I will refer it as the hedging covariance vector. Finally, $-\frac{1}{2f}(\gamma-1)f_X\Sigma^X\rho\rho^{\top}\Sigma^{X\top}f_X$ is due to the variance of the return of the intertemporal hedging component of the optimal portfolio.

We can decompose the Brownian motions of the state variable X into the Brownian motion dB_t of the assets and Brownian motions that are independent of dB_t:

$$dB_t^X = \rho dB_t + \hat{\rho} d\hat{B}_t^\top,$$

where $\rho\hat{\rho} = 0$. Therefore, the covariance matrix $\Sigma^X\Sigma^{X\top}$ of the state variable vector X may be written as

$$\Sigma^X\Sigma^{X\top} = \Sigma^X\rho\rho^\top\Sigma^{X\top} + \Sigma^X\hat{\rho}\hat{\rho}^\top\Sigma^{X\top}.$$

The first term, $\Sigma^X\rho\rho^\top\Sigma^{X\top}$, is associated with the component of dX that can be replicated by the Brownian motions of the asset prices. The second component, $\Sigma^X\hat{\rho}\hat{\rho}^\top\Sigma^{X\top}$, is associated with $d\hat{B}_t$, the component of dX that is independent of the Brownian motion of the asset prices. Therefore, I refer to $\Sigma^X\rho\rho^\top\Sigma^{X\top}$ and $\Sigma^X\hat{\rho}\hat{\rho}^\top\Sigma^{X\top} \equiv \Sigma^X\Sigma^{X\top} - \Sigma^X\rho\rho^\top\Sigma^{X\top}$ as the spanned and unspanned (state variable) covariance matrices, respectively. When $\hat{\rho} = 0$, dB_t^X is replicated by dB_t, so that markets are complete.

The assumption of complete markets is strong and is not satisfied in many specifications of the stochastic environments that have been analyzed. For example, markets are incomplete if stock returns exhibit stochastic volatility and there are no derivatives included in the set of securities, or if stock returns are predictable and the uncertainties associated with the predictors of stock returns cannot be replicated using existing securities.

For the dynamic portfolio choice problems that are considered in this article, solving PDE (7) is equivalent to solving the following PDE:[4]

$$\begin{aligned}
&\frac{\partial\hat{f}}{\partial t} + \frac{1}{2}\mathrm{Tr}\left(\Sigma^X\Sigma^{X\top}\hat{f}_{XX^\top}\right) + \left[\mu^X + \frac{1-\gamma}{\gamma}\Sigma^X\rho\Sigma^{-1}(\mu - r)\right]^\top \hat{f}_X \\
&\quad + \frac{1}{2\hat{f}}(\gamma - 1)\hat{f}_X^\top\left(\Sigma^X\Sigma^{X\top} - \Sigma^X\rho\rho^\top\Sigma^{X\top}\right)\hat{f}_X \\
&\quad + \left[\frac{1-\gamma}{2\gamma^2}(\mu - r)^\top(\Sigma\Sigma^\top)^{-1}(\mu - r) + \frac{1-\gamma}{\gamma}r - \beta\right]\hat{f} = 0, \qquad (8)
\end{aligned}$$

with condition $\hat{f}(T,X) = 1$. When $\alpha = 0$, the two PDEs are the same and $f = \hat{f}$. When $\alpha > 0$, I will show that if markets are complete, f can be expressed in terms of $\hat{f}$. I would point out here that the nonlinear term in the two PDEs drops out when markets are complete, which we will assume when we have intermediate consumption ($\alpha > 0$).

The key observation of this article is that the function $\hat{f}(t,X) = e^{c(t)+d(t)^\top X + \frac{1}{2}X^\top\eta^\top Q(t)\eta X}$ is the solution of PDE (8), if all coefficients

[4] This is proved in the Appendix.

of the PDE are quadratic in ηX and linear in X with some parameter restrictions, where η is a constant $N_1 \times N$ matrix and thus ηX are N_1 linear combinations of X and $c(t)$, $d(t)$, and $Q(t)$ are a scalar, an N-dimensional vector, and an $N_1 \times N_1$ matrix functions of t, respectively. When $\hat{f}$ is substituted into the PDE, with terms removed by the parameter restrictions, the terms on the left-hand side of the PDE will be at most quadratic in ηX and linear in X, up to a factor of $\hat{f}$. Because the equation holds for all X, the coefficients of these terms have to be zero, which leads to ordinary differential equations (ODEs) for $c(t)$, $d(t)$, and $Q(t)$. PDE (8) is thus solved when these ODEs are solved.

By inspection of PDE (8), its coefficients are quadratic in ηX and linear in X if the following expressions are quadratic in ηX and linear in X: the drift and the diffusion coefficient of the state variables μ^X and $\Sigma^X \Sigma^{X\top}$, the short rate r, the maximal squared-Sharpe ratio $(\mu - r)^\top (\Sigma \Sigma^\top)^{-1} (\mu - r)$, the hedging covariance vector $\Sigma^X \rho \Sigma^{-1} (\mu - r)$, and the unspanned covariance matrix $\Sigma^X \Sigma^{X\top} - \Sigma^X \rho \rho^\top \Sigma^{X\top}$. These are conditions, together with additional parameter restrictions, that I will impose next.

1.2 Quadratic processes

In order to obtain an analytical solution to the optimal portfolio problem, I begin by imposing restrictions on the dynamics of the state variable vector X. In particular, I assume that the state vector X_t has drift and diffusion coefficients that are quadratic functions of itself,

$$\mu^X = k - KX + \frac{1}{2} X^\top \eta^\top \cdot K_2 \cdot \eta X, \tag{9}$$

$$\Sigma^X \Sigma^{X\top} = h_0 + h_1 \cdot X + X^\top \eta^\top \cdot h_2 \cdot \eta X, \tag{10}$$

where k is an $N \times 1$ constant vector, K and h_0 are $N \times N$ constant matrices, $K_2 = \{K_{2k}^{ij}, i, j = 1, \dots, N_1, k = 1, \dots, N\}$ is a constant tensor with three indices (two upper indices and one lower index), $h_1 = \{h_{1jk}^i, i, j, k = 1, \dots, N\}$ is a constant tensor with three indices (one upper index and two lower indices), and $h_2 = \{h_{2kl}^{ij}, i, j = 1, \dots, N_1, k, l = 1, \dots, N\}$ is a constant tensor with four indices (two upper indices and two lower indices). If $\mathcal{M}$ is a matrix of appropriate dimension, then $K_2 \cdot \mathcal{M}$ denotes contraction of upper indices while $K_2 \mathcal{M}$ denotes contraction of lower indices.[5] For example, if $\mathcal{M}$ is a vector, then,

[5] Tensor is a straightforward generalization of vector and matrix. While a vector and a matrix have a collection of numbers with 1 or 2 indices, a tensor of rank n is a collection of numbers with n indices. So a vector and a matrix are a rank 1 tensor and a rank 2 tensor, respectively. $K_2 = \{K_{2k}^{ij}\}$ has three indices i, j, and k and thus is a tensor of rank 3. In this article, the indices have two ranges: N and N_1. The

$$(K_2 \cdot \mathcal{M})^i_k = \sum_{j=1}^{N_1} K^{ij}_{2k} \mathcal{M}_j,$$

(note that $\mathcal{M}$ has dimension $N_1 \times 1$ in this case) and

$$(K_2 \mathcal{M})^{ij} = \sum_{k=1}^{N} K^{ij}_{2k} \mathcal{M}_k,$$

(note that $\mathcal{M}$ has dimension $N \times 1$ in this case). Therefore, $X^\top \eta^\top \cdot K_2 \cdot \eta X = (\eta X)^\top \cdot K_2 \cdot \eta X$ is a vector of dimension $N \times 1$ whose k-th component is

$$(X^\top \eta^\top \cdot K_2 \cdot \eta X)_k = \sum_{i,\, j=1}^{N_1} K^{ij}_{2k} (\eta X)_i (\eta X)_j.$$

Also, $h_1 \cdot X$ is an $N \times N$ matrix whose (k,l) element is

$$(h_1 \cdot X)_{kl} = \sum_{i=1}^{N} h^i_{1kl} X_i.$$

Finally, $X^\top \eta^\top \cdot h_2 \cdot \eta X = (\eta X)^\top \cdot h_2 \cdot \eta X$ is an $N \times N$ matrix whose (k,l) element is

$$(X^\top \eta^\top \cdot h_2 \cdot \eta X)_{kl} = \sum_{i,j=1}^{N_1} h^i_{2kl} (\eta X)_i (\eta X)_j.$$

I impose the following additional restrictions on parameters:

$$K^\top \eta^\top = \eta^\top \hat{K}, \quad K_2 \eta^\top = 0, \quad h_1 \eta^\top = 0, \quad h_2 \eta^\top = 0. \tag{11}$$

The parameter restrictions specified in Equation (11) ensure that ηX is a multivariate Ornstein–Uhlenbeck process. To understand this, multiply both sides of Equation (1) to obtain the following equation satisfied by ηX:

indices with the range of N (or N_1) will be denoted as lower indices (or upper indices). More precisely, for a fixed i and j, $\{K^{ij}_{2k},\ k = 1, \ldots, N\}$ is a column vector of dimension N, and for a fixed k, $\{K^{ij}_{2k},\ i, j = 1, \ldots, N_1$ is a matrix of dimension $N_1 \times N_1$. Similarly, η is a tensor with one upper index and one lower index; for a fixed i, $\{\eta^i_k,\ k = 1, \ldots, N\}$ is a row vector of dimension N, and for a fixed k, $\{\eta^i_k,\ i = 1, \ldots, N_1\}$ is a column vector of dimension N_1. Products without "·" (or with "·") are just matrix multiplications involving lower indices (or upper indices), as explained in the examples below.

$$d(\eta X) = \eta\mu^X dt + \eta\Sigma^X dB_t^X.$$

The parameter restriction implies that

$$\eta\mu^X = \eta k - \hat{K}^\top \eta X$$

noting that $\eta K X = \hat{K}^\top \eta X$ because $K^\top \eta^\top = \eta^\top \hat{K}$ and $\eta(X^\top \eta^\top \cdot K_2 \cdot \eta X) = 0$ because $K_2 \eta^\top = 0$. Furthermore,

$$\eta\Sigma^X \Sigma^{X\top} \eta^\top = \eta h_0 \eta^\top$$

noting that $\eta h_1 \cdot X\eta^\top = 0$ because $h_1\eta^\top = 0$ and $\eta(X^\top\eta^\top \cdot h_2 \cdot \eta X)\eta = 0$ because $h_2\eta^\top = 0$. Therefore,

$$d(\eta X) = \eta k - \hat{K}^\top (\eta X) dt + \Sigma^{\eta X} dB_t^X,$$

where $\Sigma^{\eta X}$ is a constant matrix, which satisfies $\Sigma^{\eta X}\Sigma^{\eta X\top} = \eta h_0 \eta^\top$.

Noting that the PDE for the indirect utility is very similar to the PDE satisfied by the price of a zero-coupon bond, we can view these conditions in the context of term structure models. Constantinides (1992) shows that zero-coupon bond yields are quadratic functions of an Ornstein–Uhlenbeck process if the short rate itself is a quadratic function of the Ornstein–Uhlenbeck process. It turns out that if quadratic functions of an Ornstein–Uhlenbeck process are added to the drift and diffusion of an affine process and the short rate is an affine functions of the modified affine process and a quadratic function of the Ornstein–Uhlenbeck process, zero-coupon bond yields will also be an affine functions of the modified affine process and a quadratic function of the Ornstein–Uhlenbeck process. A generic quadratic process is just a collection of the modified affine process and the Ornstein–Uhlenbeck process.

To understand the restrictions specified in Equation (11) in more technical detail, let us substitute $\hat{f} = e^{c(t)+d(t)^\top X + \frac{1}{2} X^\top \eta^\top Q(t) \eta X}$ into PDE (8). Note that

$$\begin{aligned}
\mu^{X^\top} \hat{f}_X &= \left(k - KX + \frac{1}{2} X^\top \eta^\top \cdot K_2 \cdot \eta X\right)(d + \eta^\top Q \eta X)\hat{f} \\
&= \left(k - KX + \frac{1}{2} X^\top \eta^\top \cdot K_2 \cdot \eta X\right) d\hat{f} + k^\top \eta^\top Q \eta X \hat{f} \\
&\quad - X^\top K^\top \eta^\top Q \eta X \hat{f} + \frac{1}{2} X^\top \eta^\top \cdot (K_2 \eta^\top Q \eta X) \cdot \eta X \hat{f}.
\end{aligned}$$

All except the last two terms have the quadratic form. The term $\frac{1}{2}X^\top\eta^\top \cdot (K_2\eta^\top Q\eta X) \cdot \eta X$ is cubic in X; thus, it must be set to zero, which implies the restriction $K_2\eta^\top = 0$. For $X^\top K^\top \eta^\top Q\eta X$ to be quadratic in ηX, we must have $K^\top\eta^\top = \eta^\top\hat{K}$. Similarly, for the terms from $\mathrm{Tr}(\Sigma^X\Sigma^{X\top}f_{XX^\top})$ to have the quadratic form, we must have the restrictions $h_1\eta^\top = 0$ and $h_2\eta^\top = 0$.

Definition 1. *A quadratic diffusion process ("quadratic process") is a vector of Markovian diffusion processes X_t that satisfies Equations (9)–(11).*

I now present some special cases of quadratic diffusion processes.

1. If $\eta = 0$, which is equivalent to $K_2 = 0$ and $h_2 = 0$, quadratic terms will be absent from both drift and diffusion coefficients and the quadratic diffusion process reduces to the affine processes of Duffie and Kan (1996).
2. If $\eta = I$, $h_2 = 0$, and $K_2 = 0$, quadratic terms will be also absent from both drift and diffusion coefficients and quadratic processes become multivariate Ornstein–Uhlenbeck processes, which are a special case of affine processes.
3. The following is an example of a quadratic process that extends affine process (it is an affine process when the quadratic terms in the drift of X_{2t} and drift and diffusion coefficients of X_{3t} are absent.).

$$
\begin{aligned}
dX_{1t} &= (k_1 - K_{11}X_{1t})dt + \sigma_1 dB_{1t};\\
dX_{2t} &= \left[k_2 - K_{22}X_{2t} + (X_{1t} - \bar{X}_{12})^2\right]dt + \sigma_2\sqrt{X_{2t}}dB_{2t};\\
dX_{3t} &= \left(k_3 - K_{33}X_{3t} - K_{12}X_{2t} - K_{13}X_{3t} + X_{1t}^2\right)dt\\
&\quad + \sigma_{13}dB_{1t} + \sigma_{23}\sqrt{X_{2t}}dB_{2t}\\
&\quad + \sigma_3\sqrt{1 + M_1X_{2t} + M_2(X_{1t} - \bar{X}_{13})^2}dB_{3t},
\end{aligned}
\tag{12}
$$

where dB_{1t}, dB_{2t}, and dB_{3t} are independent Brownian motions. Note that the squares of X_{1t} appear in the drifts of X_{2t} and X_{3t} and in the diffusion of X_{3t}.

A most general quadratic process X_t can be partitioned into three processes $X_t = \{X_{1t}, X_{2t}, X_{3t}\}$, up to an affine transformation. The first process, X_{1t}, is a multivariate Ornstein–Uhlenbeck process. The second process, X_{2t}, is a generalization of a multivariate correlated square-root processes, whose diffusion is a linear function of itself (the same as a square-root process) and whose drift can depend on a positive quadratic functions of X_{1t} and a linear function of X_{2t}. The third process, X_{3t}, has a drift that depends quadratically on X_{1t} and linearly on all three processes, a diffusion term with a coefficient that is a positive quadratic function of X_{1t} and a linear function of X_{2t}, and, in addition, the diffusion terms of X_{1t} and X_{2t}.

1.3 Quadratic return

I now specify the remaining determinants of the dynamic portfolio choice problem, namely the short rate r, the maximal squared-Sharpe ratio $(\mu - r)^\top (\Sigma\Sigma^\top)(\mu - r)$, the hedging covariance vector $\Sigma^X \rho \Sigma^{-1}(\mu - r)$, and the unspanned covariance $\Sigma^X \rho\rho^\top \Sigma^{X\top} - \Sigma^X \Sigma^{X\top}$. Recall that together with drift and diffusion coefficients of the state variables that have been specified above, these terms appear in the coefficients of PDE (7) and thus determine the dynamic behavior of the optimal portfolio weights. I require that all these terms also be quadratic functions of the state variables with parameter restrictions exactly like those for μ^X and $\Sigma^X \Sigma^{X\top}$. Specifically, I assume that the asset returns satisfy the following conditions:

$$r = \delta_0 + \delta_1 X + \frac{1}{2} X^\top \eta^\top \delta_2 \eta X, \tag{13}$$

$$(\mu - r)^\top \left(\Sigma\Sigma^\top\right)^{-1} (\mu - r) = H_0 + H_1 X + \frac{1}{2} X^\top \eta^\top H_2 \eta X, \tag{14}$$

$$\Sigma^X \rho \Sigma^{-1}(\mu - r) = g_0 + g_1 X + \frac{1}{2} X^\top \eta^\top \cdot g_2 \cdot \eta X, \tag{15}$$

$$\Sigma^X \rho\rho^\top \Sigma^{X\top} - \Sigma^X \Sigma^{X\top} = l_0 + l_1 \cdot X + X^\top \eta^\top \cdot l_2 \cdot \eta X, \tag{16}$$

with restrictions

$$g_1^\top \eta^\top = \eta^\top \hat{g}_1, \quad g_2^\top \eta^\top = 0, \quad l_1 \eta^\top = 0, \quad l_2 \eta^\top = 0, \tag{17}$$

where δ_0 is a constant, δ_1 is a constant vector of dimension N, δ_2 is a constant matrix of dimension $N_1 \times N_1$, H_0 is a constant, H_1 is a constant vector of dimension N, H_2 is a constant matrix of dimension $N_1 \times N_1$, g_0 is a constant vector of dimension N, g_1 is a constant matrix of dimension $N \times N$, g_2 is a constant tensor with three indices (two upper indices running from 1 to N_1 and one lower index running from 1 to N), $\hat{g}_1$ is a constant matrix of dimension $N_1 \times N_1$, l_0 is an $N \times N$ constant matrix, l_1 is a constant tensor with three indices (with one upper index and two lower indices all running from 1 to N), and l_2 is a constant tensor with four indices (with two upper indices running from 1 to N_1 and two lower indices running from 1 to N). The tensor product is computed using the same rules I gave earlier, so $X^\top \eta^\top \cdot g_2 \cdot \eta X$ is an $N \times 1$ dimensional vector with its k-th element given by

$$(X^\top \eta^\top \cdot g_2 \cdot \eta X)_k = \sum_{i,j=1}^{N_1} g_{2k}^{ij} (\eta X)_i (\eta X)_j;$$

$l_1 \cdot X$ is an $N \times N$ matrix with its (k,l) element given by

$$(X^\top \eta^\top \cdot l_1 \cdot \eta X)_{kl} = \sum_{i=1}^{N} l_{1kl}^{ij} X_i;$$

and $(X^\top \eta^\top \cdot l_2 \cdot \eta X)$ is also an $N \times N$ matrix with its (k,l) element given by

$$(X^\top \eta^\top \cdot l_2 \cdot \eta X)_{kl} = \sum_{i,j=1}^{N_1} l_{2kl}^{ij} (\eta X)_i (\eta X)_j.$$

The first restriction is not empty because $\hat{g}_1$ may not exist.

Conditions (13)–(16) together with parameter restrictions (17) require that the short rate, the maximal squared-Sharpe ratio, the hedging covariance vector, and the unspanned covariance matrix are quadratic in ηX and linear in X. The parameter restrictions (17) are required for the same reasons as restrictions in Equation (11). Computing the indirect utility function is equivalent to computing a zero-coupon bond price under a measure where $\Sigma^X \rho \Sigma^{-1}(\mu - r)$ is added to the drift of X and $\Sigma^X \rho \rho^\top \Sigma^{X\top} - \Sigma^X \Sigma^{X\top}$ is added to diffusion of X. The parameter restrictions ensure that ηX is an Ornstein–Uhlenbeck process under this measure.

These conditions rule out dependence such as cubic and/or higher order dependence on ηX and quadratic and/or higher dependence on components of X other than ηX. They are imposed for tractability and, in general, lead to less flexibility in data fitting. Nevertheless, as we will show shortly, they are flexible enough to accommodate as special cases widely-used models of asset returns.

Definition 2. *The returns of the assets are quadratic returns if the state variable vector X is a quadratic process and Equations (13)–(17) are satisfied.*

The following are special cases of quadratic returns.

1. Asset returns in models with a constant investment opportunity set: when the short rate, the risk premium, and the volatility are all constant.

2. Stock returns with a constant interest rate, a constant volatility of asset returns, and a risk premium that follows an Ornstein–Uhlenbeck process. This is the return specification that is widely used in the literature.
3. Returns of zero-coupon bonds in quadratic gaussian term structure models [Constantinides (1992)] and in affine term structure models [Duffie and Kan (1996)].

Other examples of quadratic returns will be given later in this article.

As I show in the next subsection, when asset returns are quadratic, we will be able to derive an explicit solution to PDE (7) and thus to the optimal portfolio problem.

1.4 Optimal policies

The complete market case and incomplete market case are treated separately. The agent's utility is defined over both intermediate consumption and terminal wealth when markets are complete. When markets are incomplete, explicit solutions can be derived only if the agent's utility is defined over the terminal wealth.

1.4.1 Complete markets. When there is intermediate consumption, we need markets to be complete to obtain an explicit solution to the portfolio choice problem.[6]

Let $c(t)$ be a scalar function of t, $d(t)$ an $N \times 1$ vector function of t, and $Q(t)$ an $N_1 \times N_1$ matrix function of time t. I assume that the functions $c(t)$, $d(t)$, and $Q(t)$ satisfy the following Riccati Equations:

$$\frac{d}{dt}c + \left(k + \frac{1-\gamma}{\gamma}g_0\right)^{\top} d + \frac{1}{2}d^{\top}[h_0 + (1-\gamma)l_0]d + \frac{1}{2}\mathrm{Tr}\left(h_0\eta Q\eta^{\top}\right) + \frac{1-\gamma}{2\gamma^2}H_0 + \frac{1-\gamma}{\gamma}\delta_0 - \beta = 0, \tag{18}$$

$$\frac{d}{dt}d + \left(-K + \frac{1-\gamma}{\gamma}g_1\right)^{\top} d + \frac{1}{2}d^{\top}[h_1 + (1-\gamma)l_1]d + \eta^{\top}Q\eta\left[h_0 + (1-\gamma)l_0\right]d + \eta^{\top}Q\eta\left(k + \frac{1-\gamma}{\gamma}g_0\right) + \frac{1-\gamma}{2\gamma^2}H_1 + \frac{1-\gamma}{\gamma}\delta_1 = 0, \tag{19}$$

[6] Technically, PDE (7) can be converted to PDE (8) only when markets are complete, as shown in the proof of Proposition 1 in the Appendix.

$$\frac{d}{dt}Q + \left(-\hat{K} + \frac{1-\gamma}{\gamma}\hat{g}_1\right)^{\top} Q + Q\left(-\hat{K} + \frac{1-\gamma}{\gamma}\hat{g}_1\right) + Q\eta^{\top}[h_0 + (1-\gamma)l_0]\eta Q$$
$$+ d^{\top}[h_2 + (1-\gamma)l_2]d + \frac{1-\gamma}{\gamma}g_2^{\top}d + \frac{1-\gamma}{2\gamma^2}H_2 + \frac{1-\gamma}{\gamma}\delta_2 = 0, \tag{20}$$

with initial conditions $c(T) = 0$, $d(T) = 0$, and $Q(T) = 0$. These equations are well known and they have well-defined solutions. Next, let us define the following function f

$$f(t,X) = \alpha^{\frac{1}{\gamma}} \int_t^T e^{c(u)+d(u)^{\top}X + \frac{1}{2}X^{\top}\eta^{\top}Q(u)\eta X} du + (1-\alpha)^{\frac{1}{\gamma}} e^{c(t)+d(t)^{\top}X+\frac{1}{2}X^{\top}\eta^{\top}Q(t)\eta X}.$$

Proposition 1. *If returns are quadratic and markets are complete* ($l_0 = l_1 = l_2 = 0$), *then the optimal consumption policy* C^* *is given by*

$$C^* = \alpha^{\frac{1}{\gamma}} W f^{-1}$$

and the optimal portfolio choice ϕ^* *is given by*

$$\phi^* = \frac{1}{\gamma}(\Sigma\Sigma^{\top})^{-1}(\mu - r) + (\Sigma^{\top})^{-1}\rho^{\top}\Sigma^{X\top}\frac{\partial \ln f}{\partial X}.$$

In particular, for an asset allocation problem ($\alpha = 0$), $f = e^{c(t)+d(t)^{\top}X + \frac{1}{2}X^{\top}\eta^{\top}Q(t)\eta X}$ *and*

$$\phi^* = \frac{1}{\gamma}(\Sigma\Sigma^{\top})^{-1}(\mu - r) + (\Sigma^{\top})^{-1}\rho^{\top}\Sigma^{X\top}(d + \eta^{\top}Q\eta X).$$

Note that I suppressed the argument t of functions $d(t)$ and $Q(t)$ in the above equation. From here on, I suppress the argument t of functions of t when there is no confusion.

The first component of ϕ^* is the myopic demand and the second component is the intertemporal hedging demand. The intertemporal hedging demand depends on γ [through the functions $c(t)$ and $d(t)$] and T.

Note that in ODE (19)–(20), if the inhomogeneous terms $\frac{1-\gamma}{2\gamma^2}H_0 + \frac{1-\gamma}{\gamma}\delta_0$, $\frac{1-\gamma}{2\gamma^2}H_1 + \frac{1-\gamma}{\gamma}\delta_1$, and $\frac{1-\gamma}{2\gamma^2}H_2 + \frac{1-\gamma}{\gamma}\delta_2$ are zero, then $c + (T-t)\beta$, d, and $Q(t)$ will be zero. For example, all three inhomogeneous terms are zero for logarithmic utility ($\gamma = 1$), so $c(t) + (T-t)\beta = d(t) = Q(t) = 0$, and there is no intertemporal hedging demand in this case, which is the well known result that a logarithmic utility maximizer behaves like a myopic agent.

The more interesting case is for very conservative agents ($\gamma \to \infty$). In this case, the nonzero inhomogeneous terms are δ_0, δ_1, and δ_2, terms that are associated with the short rate. Thus, the intertemporal hedging demand of an infinitely risk-averse agent is determined only by the term structure. This is intuitively clear. The infinitely risk-averse agent only wants to hold cash, but the return from cash is also risky because the short rate changes. Therefore, the agent will hedge against the short rate risk by investing in bonds; these are exactly the intertemporal hedging demands represented by the $\delta_i (i = 0,1,2)$ terms. Campbell and Viceira (2001) and Wachter (2003) obtain a similar result when the returns are described by Ornstein–Uhlenbeck processes.

This case has important implications. First, the contribution to the intertemporal hedging demand is qualitatively different for the risks associated with the risk rate and for those not associated with the interest rate; specifically, only the interest rate risk contributes to the intertemporal hedging demand of an infinitely risk averse agent. Second, bond portfolio weights may increase with the risk aversion due to its intertemporal hedging component for the interest rate risk, whereas stock portfolio weights in general decrease with the risk aversion. It follows that, in principle, the ratio of stock to bond portfolio weights may decrease with risk aversion.

For the case of complete markets, Schroder and Skiadas (1999) characterize the solution to optimal portfolio selection and consumption problems when the agent has stochastic differential utility and derive the explicit solution for CRRA utility when the asset returns are described by affine processes, which are nested by the quadratic returns considered in this article.

1.4.2 Incomplete markets. For asset allocation problems without intermediate consumption [$\alpha = 0$ in Equation (2)], explicit solutions can be obtained without the condition that markets are complete.

Proposition 2. *Assume that returns are quadratic. Then the optimal portfolio weight is given by*

$$\phi^* = \frac{1}{\gamma}(\Sigma\Sigma^\top)^{-1}(\mu - r) + (\Sigma\Sigma^\top)^{-1}\Sigma\rho^\top\Sigma^{X\top}(d + Q\eta X). \tag{21}$$

The intertemporal hedging demand depends on γ (through the function d) and the investment horizon T.

Note that the function $f = e^{c(t)+d(t)^\top X + \frac{1}{2}X^\top \eta^\top Q(t)\eta X}$ and the intertemporal hedging demand and the optimal portfolio weights do not depend on c and therefore do not depend on k, g_0, h_0, l_0, H_0, or δ_0. The reason for this is that these parameters do not characterize the changes in the stochastic environment. When there is intermediate consumption, $\alpha \neq 0$, the portfolio weights depend on c and thus the parameters, k, g_0, h_0, l_0, and H_0, because of intertemporal substitution.

The following are special cases of the general results presented above. First, suppose that $\eta = I$, $K_2 = 0$, $h_1 = 0$, $h_2 = 0$, $\delta_2 = 0$, and $g_0 = 0$. Under these restrictions, the state variable vector X is a multivariate Ornstein–Uhlenbeck process, both the short rate and the risk premium are affine functions of X, and the return volatility is constant. Many existing explicit solutions of dynamic portfolio problems belong to this case; for example, Kim and Omberg (1996), Brennan (1998), Brennan and Xia (2000, 2001, 2002a,b), Wachter (2003), and Sangvinatsos and Wachter (2005). Second, suppose that $\eta = 0$. In this case, the short rate, the instantaneous Sharpe ratio, the hedging covariance vector, and the unspanned covariance matrix are all affine functions of an affine process X. Third, Schroder and Skiadas (2003, 2005) characterize the solution to optimal portfolio problems in a general non-Markovian setting. For the special case of nonstochastic r, they derive a similar ODE, but they do not solve for nor analyze the optimal policies.

1.5 A separation result

In this subsection, I prove a separation result that provides conditions under which dynamic portfolio choice problems with bonds and stocks can be decomposed into separate dynamic portfolio choice problems with bonds only and stocks only.

Suppose that the process X can be partitioned into two independent processes X_1 and X_2, so that

$$\mu^X = \begin{pmatrix} \mu^{X_1} \\ \mu^{X_2} \end{pmatrix}, \tag{22}$$

$$\Sigma^X = \begin{pmatrix} \Sigma^{X_1} & 0 \\ 0 & \Sigma^{X_2} \end{pmatrix}. \tag{23}$$

We also assume that

$$r = r_1 + r_2, \tag{24}$$

$$\Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ \nu\Sigma_1 & \Sigma_2 \end{pmatrix}, \tag{25}$$

$$\Sigma^{-1}(\mu - r) = \begin{pmatrix} \Sigma_1^{-1}(\mu_1 - r_1) \\ \Sigma_2^{-1}(\mu_2 - r_2) \end{pmatrix}, \tag{26}$$

$$\rho = \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix}, \tag{27}$$

where Σ^{X_i}, μ^{X_i}, r_i, and $\Sigma_i^{-1}(\mu_i - r)$ depend only on X_i and ν is a constant matrix. Note that if M risky assets are partitioned into M_1 assets and M_2 assets with $M = M_1 + M_2$, then Σ_1^{-1} denotes the inverse of Σ_1 as an $M_1 \times M_1$ matrix, and similarly for Σ_2^{-1}.

Equation (24) states that the short rate can be written as the sum of two functions that each only depends on X_1 or X_2, respectively. Equation (26) states that the risky assets can be partitioned into two classes and the market price of risk associated with the first class depends only on X_1 and the market price of risk associated with the second class depends only on X_2. Equation (25) states that the Brownian motions dB of the risky assets can be partitioned in such a way that the Brownian motion dB_1 appears only in the return of the first class risky assets and the Brownian motions of the second class risky assets have both dB_1 and dB_2. Equation (27) states that the Brownian motion dB^{X_1} of X_1 is correlated only with the Brownian motion dB_1 of the first class risky assets and the Brownian motion dB^{X_2} of X_2 is correlated only with the Brownian motion dB_2. Note that, even though the state variables are independent, the bond and stock returns can be correlated; for example, a stock return can have both interest risk and stochastic volatility.

Although conditions (22)–(27) might look contrived and restrictive, the following arguments suggest that they arise naturally in arbitrage-free models with bonds and stocks simultaneously. The partition of risky assets is natural: bonds are in the first class and stocks are in the second. The first class of state variables are state variables that appear in the drift and diffusion of the bond returns; thus, conditions (22)–(23) are satisfied. It is possible that the state variables of the first class consist of all the state variables. In this case, $r = r_1$ and $r_2 = 0$ so Equation (24) is satisfied, Σ has the form of Equation (25), $\Sigma^{-1}(\mu - r)$ has the form of Equation (26), and ρ has the form of Equation (27). Conditions (25)–(26) also require that ν be a constant and both Σ_2 and $\Sigma_2^{-1}(\mu_2 - r_2)$ depend only on X_2, which is restrictive.

Asset return models with both bonds and stocks that satisfy the partition conditions (22)–(27) are given by Brennan and Xia (2000), Mamaysky (2002), and Sangvinatsos and Wachter (2005). This partition decomposes the dynamic portfolio problem with stocks and bonds into separate bond-only and stock-only portfolio problems, as shown in the proposition below. In this case, the stock portfolio weight is the same as if there were no interest rate risk and the bond portfolio weight has a term that is proportional to the stock portfolio weight. The interest rate risk associated with this term offsets the interest rate risk of the stock in the optimal portfolio.

Proposition 3. (Separation Theorem) *Assume conditions (22)–(27) hold. Then*

$$\phi^* = \begin{pmatrix} \phi_1^* - \nu^\top \phi_2^* \\ \phi_2^* \end{pmatrix},$$

with

$$\phi_i^* = \frac{1}{\gamma}\left[\Sigma_i^{-1\top}\Sigma_i^{-1}(\mu_i - r_i) + (\Sigma_i^X \rho_i \Sigma_i^{-1})^\top \frac{\partial}{\partial X_i}\ln f\right], \quad i = 1,2.$$

The function $f(t,X)$ is given by

$$f(t,X) = \alpha^{\frac{1}{\gamma}} \int_t^T f_1(u,X_1) f_2(u,X_2) du + (1-\alpha)^{\frac{1}{\gamma}} f(t,X_1) f(t,X_2),$$

with functions $f_i(t,X_i)$ satisfying

$$\begin{aligned}\frac{\partial f_i}{\partial t} &+ \frac{1}{2}\mathrm{Tr}\left(\Sigma^{X_i}\Sigma^{X_i\top} f_{iX_iX_i^\top}\right) + \left[\mu^{X_i} + \frac{1-\gamma}{\gamma}\Sigma^{X_i}\rho_i\Sigma_i^{-1}(\mu_i - r_i)\right]^\top f_{X_i} \\ &+ \frac{1}{2f_i}(\gamma - 1) f_{X_i}^\top \left(\Sigma^{X_i}\Sigma^{X_i\top} - \Sigma^{X_i}\rho_i\rho_i^\top\Sigma^{X_i\top}\right) f_{X_i} \\ &+ \left[\frac{1-\gamma}{2\gamma^2}(\mu_i - r_i)^\top (\Sigma_i\Sigma_i^\top)^{-1}(\mu_i - r_i) + \frac{1-\gamma}{\gamma} r_i\right] f_i = 0,\end{aligned}$$

with the initial condition $f_i(T,X_i) = 1$, $i = 1,2$. When there is no intermediate consumption, $\alpha = 0$, the portfolio weights can be characterized in more detail:

$$\phi^* = \begin{pmatrix} \phi_1^* - \nu^\top \phi_2^* \\ \phi_2^* \end{pmatrix},$$

where ϕ_1^* is the vector of optimal portfolio weights for risky assets when there are only assets from the first class and only X_1 risk is present (similarly for ϕ_2^*):

$$\phi_i^* = \frac{1}{\gamma}\left[\Sigma_i^{-1\top}\Sigma_i^{-1}(\mu_i - r_i) + (\Sigma_i^X \rho_i \Sigma_i^{-1})^\top \frac{\partial}{\partial X_i}\ln f_i\right],\ i = 1,\ 2.$$

Note that the form of Σ means that there are X_1 risks in the second class of assets and these risks are proportional to those in the first class, with proportionality constant ν. The corollary implies that in this case, the X_1 risks in the second class of assets are hedged by the first class of assets.

2. Dynamic Portfolio Weights: Applications

In this section, I consider three applications. In Application 1, I first explicitly solve the bond portfolio selection problem for quadratic term structure models. For the special case of a one-factor CIR model, I derive a closed-form portfolio weight for a zero-coupon bond. Financial markets are complete in this case, and agent's utility depends on both intermediate consumption and the terminal wealth. In Application 2, I solve a stock portfolio selection problem, assuming that the stock return is described by Heston's stochastic volatility model. In this case, financial markets are incomplete and there is no intermediate consumption. In Application 3, I present a model with a bond and a stock in which the bond return is given by the CIR model and the stock returns have exposure to both interest rate risk and additional stochastic volatility risk as described by the Heston model. This model shares many features with several empirical models of stock returns that have both interest rate risk and stock volatility risk. The closed-form solutions of the bond and stock portfolio weights are obtained by an application of the separating proposition using the explicit bond and stock portfolio weights of the CIR and Heston models, respectively. In this case, markets are incomplete and there is no intermediate consumption.

2.1 Bond portfolio selection

Application 1 is an application of Proposition 1. In this case, the risks are interest rate risks, the risky assets are zero-coupon bonds, and financial markets are complete. This application is important for the following reasons. First, bond portfolio selection problems are important in themselves and are one component of the bond and stock portfolio selection problem. Second, the solution is empirically relevant because quadratic term structure models nest both the quadratic gaussian term structure models [Constantinides (1992)] and the affine term structure models [Duffie and Kan (1996)], both of which are used in many empirical studies. Finally, dynamic portfolio choice effects are likely to be strongest in bonds.

2.1.1 Quadratic term structure models. I assume that the interest rate dynamics are determined by an N^r-dimensional quadratic process X^r and denote its Brownian motion by dB_t^r and its parameters by k^r, K^r, η^r, K_2^r, h_0^r, h_1^r, h_2^r, and $\hat{K}^r$. Let λ^I denote a constant $N^r \times 1$ vector and λ_1^I a constant $N_1^r \times 1$ vector.

Definition 3. *Quadratic term structure models are term structure models in which 1) the short rate r is a quadratic function of a quadratic process* X^r, $r = \delta_0 + \delta_1 X^r + \frac{1}{2} X^{r\top} \eta^{r\top} \delta_2 \eta^r X^r$ *and 2) the market price of risk is* $(\lambda^I + \lambda_1^I \eta X^r)^\top \Sigma^{X^r}$, *where* λ_1^I *satisfies* $h_1^r \lambda_1^I \eta = 0$ *and* $h_2^r \lambda_1^I \eta = 0$.

For zero-coupon bond yields to be quadratic function of a quadratic process, we need X^r to be a quadratic process under the risk-neutral measure. The pricing kernel of quadratic term structure models can be obtained by using the standard formula that links pricing kernels to the short rate and the market price of risk:

$$\exp\left(-\int_0^t r_u du\right) \exp\left(\int_0^t (\lambda^I + \lambda_1^I \eta^r X_u^r)^\top \Sigma_u^{X^r} dB_u^r \right.$$
$$\left. - \frac{1}{2}(\lambda^{I\top} \Sigma_u^{X^r} \Sigma_u^{X^r\top} \lambda^I + X_u^{r\top} \eta^{r\top} \lambda_1^{I\top} h_0^r \lambda_1^I \eta^r X_u^r) du \right).$$

The second factor in the above pricing kernel is the Radon–Nykodym derivative that links the physical measure to the risk-neutral measure. Using Girsanov's theorem, one can easily check that the restrictions $h_1^r \lambda_1^I \eta = 0$ and $h_2^r \lambda_1^I \eta = 0$ imply that X^r is a quadratic process under the risk-neutral measure.

As mentioned earlier, quadratic term structure models nest both the quadratic gaussian term structure models ($K_2^r = 0$ and $h_1^r = h_2^r = 0$) of Constantinides (1992) and affine term structure models ($\eta^r = 0$) of Duffie and Kan (1996). The above pricing kernel for the special case of the CIR model is derived from an equilibrium model [CIR (1985)]. The main advantage of specifying a pricing kernel is that the bond returns derived from it will be free of arbitrage.[7]

There are $N^r + 1$ assets: an instantaneously riskless asset with an instantaneously riskless return of r_t and N^r bonds. The price of a zero-coupon bond is similar to that in the works of Constantinides (1992) and Duffie and Kan (1996):

[7] The bond returns are determined from an expectation hypothesis of Brennan, Schwartz and Lagnado (1997).

$$P^r(t,T_1) = \exp\left(a(t) + b(t)^\top X_t^r + \frac{1}{2} X_t^{r\top} \eta^{r\top} G(t) \eta^r X_t^r\right),$$

where $a(t)$ is a scalar function of t,$b(t)$ is an $N \times 1$ vector function of t, and $G(t)$ is an $N_1^r \times N_1^r$ matrix function of time t. They depend on maturity $\tau_1 = T_1 - t$, where T_1 is the vector of maturity dates.

For simplicity, I choose N^r zero-coupon bonds with N^r different maturity dates $T_1 > 0$, so T_1 is a constant N^r vector. Because there are N^r bonds for N^r sources of uncertainties, markets are complete. Note that as time changes, the maturities of the bonds $T_1 - t$ change, but not the bonds themselves (the maturity dates T_1 do not change). This is different from investing in constant maturity bonds as described by Campbell and Viceira (2001). If we let $T_1 - t$ be constant so that $T_1 = \tau_1 + t$ with constant τ_1, then our following observations apply also to the constant maturity case discussed by Campbell and Viceira (2001).

One can easily check that the bond price P^r satisfies

$$\begin{aligned}\frac{dP^r(t,T_1)}{P^r(t,T_1)} = &\left[r_t + (b + G\eta^r X^r)^\top \Sigma^{X^r} \Sigma^{X^r\top} (\lambda^I + \lambda_1^I \eta^r X^r)\right] dt \\ &+ (b + G\eta^r X^r)^\top \Sigma^{X^r} dB_t^r.\end{aligned}$$

In terms of notation in Section 1, one can verify that

$$\Sigma^{-1}(\mu - r) = \Sigma^{X^r\top}(\lambda^I + \lambda_1^I \eta^r X^r),\ \rho = I.$$

Therefore,

$$\begin{aligned}
r &= \delta_0 + \delta_1 X^r + \frac{1}{2} X^{r\top} \eta^{r\top} \delta_2 \eta^r X^r, \\
(\mu - r)^\top (\Sigma\Sigma^\top)^{-1} (\mu - r) &= \lambda^{\top I} \Sigma^{X^r} \Sigma^{X^r\top} \lambda^I + 2\lambda^{I\top} h_0^r \lambda_1^I \eta^r X^r \\
&\quad + X^{r\top} \eta^{r\top} \lambda_1^{I\top} h_0^r \lambda_1^I \eta^r X^r, \\
\Sigma^X \rho \Sigma^{-1} (\mu - r) &= \Sigma^r \Sigma^{r\top} \lambda^I + h_0^r \lambda_1^I \eta^r X^r, \\
\Sigma^X \rho\rho^\top \Sigma^{X\top} &= \Sigma^X \Sigma^{X\top}, \\
h_1 \eta^\top &= 0, \\
g_1^{r\top} \eta^{r\top} &= \eta^{r\top} \lambda_1^{I\top} h_0^r \eta^{r\top} (\text{so } \hat{g}_1^r = \lambda_1^{I\top} h_0^r \eta^{r\top}),
\end{aligned}$$

implying that Condition 1 is satisfied. Let c_r, d_r, and Q_r be a scalar, an $N_1^r \times 1$ vector, and an $N_1^r \times N_1^r$ matrix functions of time t, respectively. Equations (18)–(20) become

$$\frac{d}{dt}c_r + \left(k^r + \frac{1-\gamma}{\gamma}h_0^r\lambda^I\right)^\top d_r + \frac{1}{2}d^\top h_0^r d + \frac{1}{2}\mathrm{Tr}\left(h_0^r\eta^r Q_r\eta^{r\top}\right)$$
$$+\frac{1-\gamma}{2\gamma^2}\lambda^{I\top}h_0^r\lambda^I + \frac{1-\gamma}{\gamma}\delta_0 - \beta = 0, \tag{28}$$

$$\frac{d}{dt}d_r + \left[-K^r + \frac{1-\gamma}{\gamma}(h_1^r\lambda^I + h_0^r\lambda_1^I\eta^r)\right]^\top d_r + \frac{1}{2}d_r^\top h_1^r d_r + \eta^{r\top}Q_r\eta^r h_0^r d_r$$
$$+\eta^{r\top}Q_r\eta^r\left(k^r + \frac{1-\gamma}{\gamma}h_0^r\lambda^I\right) + \frac{1-\gamma}{\gamma^2}\eta^{r\top}\lambda_1^{I\top}h_0^r\lambda_I$$
$$+\frac{1-\gamma}{2\gamma^2}\lambda^{I\top}h_1^r\lambda^I + \frac{1-\gamma}{\gamma}\delta_1 = 0, \tag{29}$$

$$\frac{d}{dt}Q_r + \left(-\hat{K}^r + \frac{1-\gamma}{\gamma}\lambda_1^{I\top}h_0^r\eta^{r\top}\right)^\top Q_r + Q_r\left(-\hat{K}^r + \frac{1-\gamma}{\gamma}\lambda_1^{I\top}h_0^r\eta^\top\right)$$
$$+Q^r\eta^{r\top}h_0^r\eta^r Q^r + d_r^\top h_2^r d_r + K_2^{r\top}d_r$$
$$+\frac{1-\gamma}{2\gamma^2}\left(\lambda^{I\top}h_2^r\lambda^I + \lambda_1^{I\top}h_0^r\lambda_1^I\right) + \frac{1-\gamma}{\gamma}\delta_2 = 0. \tag{30}$$

Note that a subscript r (for short rate r) is added to the functions c, d, and Q. The optimal consumption and the optimal portfolio weights can be obtained by a straightforward application of Proposition 1.

Corollary 1. *Under the quadratic term structure model, the optimal consumption and optimal portfolio weights are*

$$C^* = \alpha^{\frac{1}{\gamma}}Wf_r^{-1},$$
$$\phi_r^* = \left(b + G\eta^r X^r\right)^{-1}\left[\frac{1}{\gamma}(\lambda^I + \lambda_1^I\eta^r X^r) + \frac{\partial \ln f_r}{\partial X^r}\right],$$

with

$$f_r = \alpha^{\frac{1}{\gamma}}\int_t^T e^{c_r(u)+d_r(u)^\top X^r + \frac{1}{2}X^{r\top}\eta^{r\top}Q_r(u)\eta^r X^r}$$
$$du + (1-\alpha)^{\frac{1}{\gamma}}e^{c_r(t)+d(t)_r^\top X^r + \frac{1}{2}X^{r\top}\eta^{r\top}Q_r(t)\eta^r X^r}.$$

Note that $(b + G\eta^r X^r)^{-1}$ exists if the bonds are not instantaneously redundant.

For the asset allocation problem ($\alpha = 0$),

$$\phi_r^* = b^{-1}\left(\frac{1}{\gamma}\lambda^I + d_r + \eta^{r\top} Q_r \eta^r X^r\right).$$

Furthermore, for the special case of an affine term structure model ($\eta^r = 0$), the optimal portfolio choice is given by

$$\phi_r^* = b^{-1}\left(\frac{1}{\gamma}\lambda^I + d_r\right).$$

In this case, the portfolio weights are independent of the state variable X^r, so there is no market timing issue and how much to invest in different bonds will not depend on the realization of the short rate.

Note that the above ODEs [Equations (28)–(30)] for functions c_r, d_r, and Q_r, and thus the functions c_r, d_r, and Q_r themselves, are independent of the type or the maturity of the bonds. This is because markets are complete. More generally, we can choose as risky assets any N^r term structure instruments, such as coupon bonds or bond derivatives. The only advantage of using zero-coupon bonds as the risky assets is that the matrix Σ is known (up to the solution to an ODE). For other term structure instruments, Σ may not be explicitly given, but Equations (28)–(30) remain the same and the optimal utility level will be the same as long as the number of the bonds is the same as the number of the state variables.

If $\gamma = \infty$, one can show that

$$d(T - t) = b(T - t),$$

in which case the portfolio weight becomes

$$\phi^* = b(T_1 - t)^{-1} b(T - t).$$

If we choose the first zero-coupon bond to have the same maturity as the investment horizon, $T_1 = (T,\ldots)^\top$, then we get

$$\phi^* = (1,0,\ldots,0)^\top.$$

Therefore, an infinitely risk-averse agent puts all of his wealth in the zero-coupon bond with the maturity that is the same as the investment horizon. Intuitively, an infinitely risk-averse agent with the end of period utility at time T wants to consume exactly the same amount for different realizations of the state at T, and the zero-coupon bond with the same maturity as the investment horizon delivers exactly that. Campbell and Viceira (2001) and Wachter (2003) obtain a similar result for the special case when the term structure model is affine.[8]

2.1.2 Bond allocation problem in the CIR model. I now derive the explicit bond portfolio weight for an asset allocation problem ($\alpha = 0$) when the short rate is described by the CIR model. In the one factor ($N^r = 1$) CIR model, the interest rate state variable X_t^r is the short rate itself, $X_t^r = r_t$, which satisfies

$$dr_t = (k^r - K^r r_t)dt + \sigma^r \sqrt{r_t} dB_t^r,$$

where k^r, K^r, and σ^r are constants. The price $P(t,T_1)$ of a zero-coupon bond with maturity $\tau_1 = T_1 - t$ is given by [Cox, Ingersoll, and Ross(1985)]

$$P(t,T_1) = \exp\big(a(t) + b(t)r\big),$$

with

$$a(t) = \frac{2k^r}{\sigma^{r2}} \ln\left(\frac{2\tilde{\xi} \exp\left(\frac{(K^r + \lambda^I \sigma^{r2} + \tilde{\xi})\tau_1}{2}\right)}{(K^r + \lambda^I \sigma^{r2} + \tilde{\xi})\left[\exp\left(\tilde{\xi}\tau_1\right) - 1\right] + 2\tilde{\xi}}\right),$$

$$b(t) = -\frac{2\left[\exp\left(\tilde{\xi}\tau_1\right) - 1\right]}{(K^r + \lambda^I \sigma^{r2} + \tilde{\xi})\left[\exp\left(\tilde{\xi}\tau_1\right) - 1\right] + 2\tilde{\xi}},$$

where $\tilde{\xi} = \sqrt{(K^r + \lambda^I \sigma^{r2})^2 + 2\sigma^{r2}}$ and λ^I is the price of interest risk.[9] Note that $b < 0$; that is, the bond prices vary inversely with the short rate.

The return at time t of a zero-coupon bond $P^r(t,T_1)$ maturing at time T_1 satisfies the following equation:

$$\frac{dP^r(t,T_1)}{P^r(t,T_1)} = \left(r_t + b\lambda^I \sigma^{r2} r_t\right)dt + b\sigma^r \sqrt{r_t} dB_t^r.$$

This is derived by Cox, Ingersoll, and Ross (1985). The bond return has a risk premium $b(t)\lambda_1^I r_t$ that changes with time t both implicitly (through the dependence on r_t) and explicitly (through the dependence on b). As

[8] I remark that the affine case of bond portfolio problems has been extended to the case with an infinite number of factors by Collin-Dufresne and Goldstein (2003). Bond-only portfolio selection problems have also been studied by Campbell and Viceira (2001), Kargin (2003), and Tehranchi and Ringer (2004).

[9] A notational clarification: the price of the interest rate risk λ^I is λ/σ^r given in the study of Cox, Ingersoll, and Ross (1985).

pointed out by Cox, Ingersoll, and Ross, $\lambda^I < 0$, if the bond risk premium is positive.

Note that the optimal portfolio weight does not depend on c_r when $\alpha = 0$, as we pointed out before, and $Q_r = 0$. We only need to solve for function d of Equation (29), which becomes

$$\frac{d}{dt}d_r - \left(K^r - \frac{1-\gamma}{\gamma}\sigma^{r2}\lambda^I\right)d_r + \frac{1}{2}\sigma^{r2}d_r^2 + \frac{1-\gamma}{2\gamma^2}\sigma^{r2}\lambda^{I2} + \frac{1-\gamma}{\gamma} = 0, \quad (31)$$

with initial conditions $d_r(T) = 0$.

Define $\delta_r = -\left[\frac{1-\gamma}{2\gamma^2}(\lambda^I)^2\sigma^{r2} + \frac{1-\gamma}{\gamma}\right]$, $\tilde{K}^r = K^r - \frac{1-\gamma}{\gamma}\lambda^I\sigma^{r2}$,
$\xi^r = \sqrt{\tilde{K}^{r2} + 2\delta_r\sigma^{r2}}$, and $\zeta^r = -i\xi^r$.

Corollary 2. *The function $d_r(t)$ is given by*

$$d_r(t) = \begin{cases} -\dfrac{2\left[\exp\left(\xi^r\tau\right)-1\right]}{(\tilde{K}^r+\xi^r)\left[\exp\left(\xi^r\tau\right)-1\right]+2\xi^r}\delta_r, if & \xi^{r2} \geq 0; \\ -\dfrac{2}{\tilde{K}^r+\zeta^r\frac{\cos(\zeta^r\tau/2)}{\sin(\zeta^r\tau/2)}}\delta_r, if & \zeta^{r2} > 0. \end{cases}$$

The optimal bond portfolio weight is given by

$$\phi_r^* = \frac{1}{\gamma}b^{-1}(\lambda^I + \gamma d_r).$$

From the above formula, one can verify that the bond portfolio weight is always positive and decreases monotonically to 1 as risk aversion increases to ∞, if $2K^r\lambda^I\sigma^{r2} + (\lambda^I\sigma^{r2})^2 + 2\sigma^{r2} \leq 0$. More interestingly, the bond portfolio weight is negative for small risk aversion and increases monotonically to 1 as the risk aversion increases to ∞ if $2K^r\lambda^I\sigma^{r2} + (\lambda^I\sigma^{r2})^2 + 2\sigma^{r2} > 0$, even though the risk premium is strictly positive. This happens only in a dynamic setting. In a static choice setting, a bond is just another risky asset like a stock and the holding of both bonds and stocks goes to zero as the risk aversion increases. In a dynamic setting, when the interest rate is stochastic, bonds have an extra function. As the risk aversion increases, the agent wants to hold more of the instantaneously riskless asset, which has risks in the future when the short rate changes. In this case, more bonds need to be held to hedge the short rate risk.

It is also interesting that, in this case, the optimal terminal wealth W_T^* for $\gamma = \infty$ is a constant and thus riskless, but it has the highest exposure to dB^r_t risk (because ϕ^*_r is the highest for $\gamma = \infty$) among the optimal wealth levels for different risk aversion γ.

2.2 Stock allocation in a stochastic volatility model

Application 2 is an application of Proposition 2. I explicitly solve a stock portfolio selection problem. I assume that there are two assets: a riskless asset with a constant return r and a risky asset which is the stock. The stock return volatility is described by the Heston (1993) model. The markets are incomplete in this case and there is no intermediate consumption. I solve for the optimal stock portfolio weight.

The stock price P^s in the Heston (1993) model satisfies

$$\frac{dP^s_t}{P^s_t} = (r + \lambda_s V_t)dt + \sqrt{V_t}dB_t,$$

where the volatility V_t is a square-root process

$$dV_t = (k_v - K_v V_t)dt + \sigma_v \sqrt{V_t}dB^v_t.$$

The correlation between B_t and B^v_t is a constant ρ_v.[10] Because there are two sources of risk, dB_t and dB^v_t, and there is only one risky asset, the markets are incomplete.

In terms of the notation of Section 1,

$$\Sigma^X = \sigma_v \sqrt{V}, \quad \Sigma = \sqrt{V}, \quad \rho = \rho_v, \quad \mu - r = \lambda_s V.$$

These specifications imply

$$\begin{aligned}
&r = \delta_0, \\
&(\mu - r)^\top (\Sigma\Sigma^\top)^{-1}(\mu - r) = \lambda_s^2 V, \\
&\Sigma^X \rho \Sigma^{-1}(\mu - r) = \rho_v \sigma_v \lambda_s V, \\
&\Sigma^X(\rho\rho^\top - I)\Sigma^{X\top} = -(1 - \rho_v^2)\sigma_v^2 V.
\end{aligned}$$

Note that the optimal portfolio weight does not depend on c when $\alpha = 0$, as we pointed out before, and $Q = 0$; we only need to solve for function d of Equation (19), which becomes

[10] This model captures a number of features of stock returns and is applied by Heston (1993), Bates (2000), and Pan (2002) to empirical data. When $\rho_v < 0$, this model produces an asymmetric smile curve [Bates (2000) and Pan (2002)].

$$\frac{d}{dt}d_v - \left(K_v - \frac{1-\gamma}{\gamma}\lambda_s\sigma_v\rho_v\right)d_v + \frac{\sigma_v^2}{2}\left[1-(1-\gamma)(1-\rho_v^2)\right]d_v^2 + \frac{1-\gamma}{2\gamma^2}\lambda_s^2 = 0, \tag{32}$$

with initial conditions $d_v(T) = 0$. Note that a subscript v (for volatility) is added to function d.

Let us define $\delta_v - \frac{1-\gamma}{2\gamma^2}\lambda_s^2$, $\tilde{K}_v = K_v - \frac{1-\gamma}{\gamma}\lambda_s\sigma_v\rho_v$, $\xi_v = \sqrt{\tilde{K}_v^2 + 2\delta_v[\rho_v^2 + \gamma(1-\rho_v^2)]\sigma_v^2}$, and $\zeta_v = -i\xi_v$.

Corollary 3. *The function $d_v(t)$ is given by*

$$d_v(t) = \begin{cases} -\dfrac{2\left[\exp\left(\xi_v\tau\right)-1\right]}{(\tilde{K}_v+\xi_v)\left[\exp\left(\xi_v\tau\right)-1\right]+2\xi_v}\delta_v, \ if\ \xi_v^2 \ \geq \ 0; \\ -\dfrac{2}{\tilde{K}_v+\zeta_v\frac{\cos(\zeta_v\tau/2)}{\sin(\zeta_v\tau/2)}}\delta_v, \ if\ \zeta_v^2 \geq 0. \end{cases}$$

The optimal stock portfolio weight ϕ_s^* is given by

$$\phi_s^* = \frac{1}{\gamma}\lambda_s + \rho_v\sigma_v d_v.$$

From the above formula, one can verify that the stock portfolio weight is always positive and decreases monotonically to 0 as the risk aversion increases to ∞, if $2K_v\rho_v\lambda_s\sigma_v + (\lambda_s\sigma_v)^2 \leq 0$. More notably, the stock portfolio weight is negative for small risk aversion and reaches a maximum value and then decreases to 0 as the risk aversion increases to ∞, if $2K_v\rho_v\lambda_s\sigma_v + (\lambda_s\sigma_v)^2 < 0$, even though the risk premium is strictly positive.

In static portfolio choice, a risk-averse agent will always hold a positive amount of a risky asset with a positive risk premium. This property requires only risk aversion and a positive risk premium and thus holds quite generally. Because of this property, some researchers have termed the fact that many agents do not hold stocks even though the market risk premium is positive, the nonmarket participation puzzle.

However, this property need not hold in dynamic portfolio choice theory; as we pointed earlier, the stock portfolio is negative for small risk aversion in the stochastic volatility model. Why an agent with small (but positive) risk aversion would short a risky asset with a positive premium can be understood as follows. The utility function of this agent, $\frac{W^{1-\gamma}}{1-\gamma}$, is bounded from below and unbounded from above for $\gamma < 1$, implying that a positive return followed by a high Sharpe ratio (good future opportunity) yields a big utility gain and a negative return followed by a low Sharpe ratio (bad future opportunity) does not yield a big utility loss. In our model, a positive return is more likely to be followed by a low Sharpe ratio and a negative return is more likely to be followed by a high Sharpe ratio because the correlation between the return and the Sharpe ratio is negative (for both a bond and a stock). The agent will hold the risky asset later when the horizon is short and the myopic demand dominates. By shorting the risky assets now, the agent effectively creates a trading strategy that has the pattern of high returns more likely being followed by a high Sharpe ratio.

In the works of Kim and Omberg (1996) and Brennan and Xia (2001), the agent also shorts stock. However, the risk premium in their models can be negative; therefore, it is not possible to determine whether the shorting in their models is due to the potential negative risk premium or dynamic choice.

Also, in static choice settings, a more risk-averse agent will always hold less of a risky asset with a positive risk premium. This property too can be proved under quite general conditions. On the basis of this property, an agent's holding of risky assets is often used as a proxy for the agent's risk aversion because the risk aversion coefficient is difficult to measure; in industry practice and often in academic studies, the agents' risk aversion coefficients are inferred from their holdings of stocks.

However, as pointed out earlier, the stock portfolio weight in the stochastic volatility model is a nonmonotonic function of the risk aversion. To explain this phenomenon, first note that the myopic component decreases when the risk aversion increases. However, the effect of increasing risk aversion on the intertemporal hedging component is not clear. On one hand, the smaller myopic amount implies a smaller amount to be hedged and therefore a smaller intertemporal hedging component; on the other hand, a more risk-averse agent values more the mean-reversion effect of the risky asset, which will lead to a larger intertemporal hedging component. If the latter effect is big enough, the portfolio weight will increase with risk aversion and a risk-averse agent may want to short the risky asset with a positive risk premium.

In the works of Kim and Omberg (1996) and Brennan and Xia (2001), the stock holdings are also nonmonotonic in γ. However, the risk premium in their models is not strictly positive; therefore, it is not possible to determine whether this is due to a negative risk premium or dynamic choice. For example, the mean-variance portfolio weight, $\mu - r/\gamma\sigma^2$, is increasing in γ if $\mu - r < 0$.

One striking property of the optimal stock portfolio weight is that it is independent of the variance V. One might expect the agent to hold more stocks when the volatility is low and less when the volatility is high. However, this is only true if the risk premium is independent of the volatility. In general, the risk premium depends on the volatility at least for the market portfolio. In our model, the risk premium is proportional to the conditional variance. Therefore, when the variance is high, the risk premium is also high in such a way that the myopic demand is independent of the stochastic variance. The independence of the intertemporal hedging demand on the variance is a special feature of our model.[11]

Note that the Sharpe ratio in my model is $\lambda_s\sqrt{V}$, which is increasing in volatility. Evidence for this feature is provided by Campbell and Cochrane (1999). Chacko and Viceira (2005) derive approximate portfolio weights in a stochastic volatility model that is similar to the Heston model. The risk premium is constant in their model, and thus, the Sharpe ratio decreases with volatility. If the risk premium is proportional to $\sqrt{V}$ instead of V,[12] the instantaneous Sharpe ratio is constant and the intertemporal hedging component will be zero. The myopic component will depend on V and is given by

$$\phi_s^* = \frac{\lambda_s}{\gamma} V^{\frac{-1}{2}}.$$

In this case, the compensation (risk premium) for volatility is not "attractive" enough (the Sharpe ratio is constant instead of increasing in V), so the portfolio weight is decreasing in volatility.

2.3 Bond and stock allocation in a stochastic interest rate–stochastic volatility model

Application 3 is an application of Proposition 3. I derive the bond and stock portfolio weights in a model with stochastic volatility and a stochastic short rate by an application of Proposition 3 using the results of Corollaries 2 and 3. There are three assets in this case: an instantaneously riskless asset with an instantaneously riskless return r_t, a zero-coupon bond, and a stock. In this case, the markets are incomplete and there is no intermediate consumption. The bond and stock portfolio weights are derived in closed-form and their properties are analyzed.

The short rate r_t is given by the CIR model as in Section 2.1.2. The zero-coupon bond is the same as one in Section 2.1.2. The stock price P_t^s satisfies

---

[11] We could easily incorporate V_t dependence in the portfolio weight by specifying $dP_t^s/P_t^s = (r + \varphi(V_t)\lambda_s V_t)dt + \varphi(V_t)\sqrt{V_t}dB_t$, where $\varphi(V_t)$ is any positive function of V_t. Under this specification, one can easily check that the indirect utility f is the same as obtained above. Therefore, the optimal portfolio weight is $\phi_s^* = \varphi^{-1}\left(\frac{1}{\gamma}\lambda_s + \rho_v\sigma_v d_v\right)$.

[12] Cochrane and Sá-Requejo (2001) use this model to study option prices.

$$\frac{dP_t^s}{P_t^s} = (r_t + \lambda_s V_t + \nu\sigma^{r2}\lambda^I r_t)dt + \sqrt{V_t}dB_t + \nu\sigma^r\sqrt{r_t}dB_t^r,$$

where as before the volatility V_t satisfies

$$dV_t = (k_v - K_v V_t)dt + \sigma_v\sqrt{V_t}dB_t^v.$$

I assume that there is no correlation between the Brownian motions B_t^r and B_t^v and between B_t^r and B_t. The correlation between B_t and B_t^v is ρ_v. The "shock" to the stock return is the sum of two contributions: $\sqrt{V_t}dB_t$ and $\nu\sigma\sqrt{r_t}dB_t^r$. The volatility of the stock return is $\sqrt{V_t + \nu^2\sigma^2 r_t}$ and so depends on V_t as well as on the short rate r_t.[13] The correlation between the shock to the stock return and the shock to the short rate is time varying and given by $\frac{\nu\sigma\sqrt{r_t}}{\sqrt{V_t+\nu^2\sigma^2 r_t}}$, implying that the magnitude of the correlation is high at low V_t's and high r_t's. Note that because bond returns are negatively correlated with the short rate, the correlation between bond and stock returns is positive for $\nu < 0$. The correlation between the "shocks" to the stock returns and the "shocks" to the stochastic volatility V_t is time varying and given by $\frac{\rho_v\sqrt{V_t}}{\sqrt{V_t+\nu^2\sigma^2 r_t}}$. The magnitude of the correlation is decreasing in r_t but increasing in V_t. Both dB_t and dB_t^r risks are compensated in terms of the risk premium. The Sharpe ratio is time varying and is given by $\frac{\lambda_s V_t+\nu\sigma^2\lambda^I r_t}{\sqrt{V_t+\nu^2\sigma^2 r_t}}$. For $\nu < 0$, the Sharpe ratio will be high at high V_t's and low r_t's.

This model is the continuous-time version of the model estimated by Scruggs (1998). The risk premium of the stock is similar to that assumed by French, Schwert, and Stambaugh (1987), Harvey (1989), and Glosten, Jagannathan, and Runkle (1993). Fama and Schwert (1977) assume that the expected stock returns depend linearly on nominal T-bill yields. For short horizons, the interest rate can be treated as a constant; in this case, the above model for stock returns reduces to Heston's (1993) stochastic volatility model.

Note that with the markets for the interest rate risk complete, the absence of arbitrage dictates the risk premium associated with the dB_t^r risk in the stock return. In other words, the premium $\nu\sigma^2\lambda^I r_t$ of the risk $\nu\sigma\sqrt{r_t}dB_t^r$ is determined by the bond-pricing kernel (when markets for interest rate risk are complete, the pricing kernel for the

[13] Glosten, Jagannathan, and Runkle (1993) find evidence suggesting that the conditional volatility of stock returns is predicted by the short rate.

interest rate risks is uniquely determined; thus is the market price of dB_t^r risks).

In terms of the notation of Section 1, we have

$$\Sigma^X = \begin{pmatrix} \sigma^r\sqrt{r} & 0 \\ 0 & \sigma_v\sqrt{V} \end{pmatrix}, \ \mu^X = \begin{pmatrix} k^r - K^r r \\ k_v - K_v V \end{pmatrix},$$

and

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & \rho_v \end{pmatrix}, \ \Sigma = \begin{pmatrix} b\sigma^r\sqrt{r} & 0 \\ \nu\sigma^r\sqrt{r} & \sqrt{V} \end{pmatrix}, \ \Sigma^{-1}(\mu - r) = \begin{pmatrix} \lambda^I\sigma^r\sqrt{r} \\ \lambda_s\sqrt{V} \end{pmatrix}.$$

Therefore, the condition of Proposition 3 holds and we have the following corollary.

Corollary 4. *The optimal portfolio weight is given by*

$$\begin{pmatrix} \phi_r^* - \nu b^{-1}\phi_s^* \\ \phi_s^* \end{pmatrix}, \tag{33}$$

where $\phi_r^* = b^{-1}(\frac{\lambda^I}{\gamma} + d_r)$ *is the bond portfolio weight in the CIR model and* $\phi_s^* = \frac{1}{\gamma}\lambda_s + \rho_v\sigma_v d_v$ *is the stock portfolio weight in the Heston model.*

The proof is a straightforward application of Proposition 3 and Corollaries 2 and 3.

The stock portfolio weight is the same as that in the Heston model because the interest rate risk is completely hedged by the bond and the interest rate can be treated as a constant as far as stock portfolio selection is concerned. The second term in the bond portfolio weight offsets the interest rate risk exposure of the stock precisely. When $\nu = 0$, the bond portfolio weight is the one for the bond-only portfolio selection problem.

3. Conclusion

I explicitly solve dynamic portfolio choice problems, up to the solution of a system of ODEs when asset returns are quadratic. The solutions include as special cases many existing explicit solutions to dynamic portfolio choice problems and can be used to study agents' portfolio choice when interest rates, expected returns, and volatilities of returns are stochastic.

I consider the following three applications: Application 1 is a bond portfolio choice with quadratic term structure models; Application 2 is a stock portfolio choice with Heston's stochastic volatility model; and

Application 3 is a bond and stock portfolio choice when both the interest rate and the stock volatility are stochastic.

Using explicit portfolio weights, I show that the properties of dynamic portfolio choice can be quite different from those of static portfolio choice. For example, a risk-averse agent may short a risky asset with a strictly positive risk premium, and a more risk-averse agent may hold more of it. Explanations for these departures from static portfolio choice theory can ultimately be traced to the rebalancing that occurs under dynamic portfolio choice.

Appendix

I will first prove two lemmas. Define operator $\mathcal{L}$ on any function f by

$$\begin{aligned}\mathcal{L}f = &\left[k - KX + \frac{1}{2}X^\top\eta^\top \cdot K_2 \cdot \eta X + \frac{1-\gamma}{\gamma}\left(g_0 + g_1 X + \frac{1}{2}X^\top\eta^\top \cdot g_2 \cdot \eta X\right)\right]^\top f_X \\ &+ \frac{1}{2}\mathrm{Tr}\left((h_0 + h_1 \cdot X + X^\top\eta^\top \cdot h_2 \cdot \eta X) f_{XX^\top}\right) \\ &+ \frac{1-\gamma}{\gamma}\left[\frac{1}{2\gamma}\left(H_0 + H_1 X + \frac{1}{2}X^\top\eta^\top H_2 \eta X\right) + \left(\delta_0 + \delta_1 X + \frac{1}{2}X^\top\eta^\top\delta_2\eta X\right)\right] f - \beta f.\end{aligned}$$

Lemma 1. *Let* $\hat{f}(t,X) = e^{c(t)+d(t)^\top X + \frac{1}{2}X^\top\eta^\top Q\eta X}$. *If c, d, and Q satisfy Equations (18), (19), and (20), respectively, with* $l_0 = l_1 = l_2 = 0$, *then* $\hat{f}$ *satisfies*

$$\frac{\partial \hat{f}}{\partial t} + \mathcal{L}\hat{f} = 0, \tag{A1}$$

and $\hat{f}(T,X) = 1$.

Proof. First, noting that $h_1\eta^\top = 0$, $h_2\eta^\top = 0$ and

$$\hat{f}_{XX^\top} = \left(\eta^\top Q\eta + (d + \eta^\top Q\eta X)(d + \eta^\top Q\eta X)^\top\right)\hat{f},$$

we get

$$\begin{aligned}&\mathrm{Tr}\left((h_0 + h_1 \cdot X + X^\top\eta^\top \cdot h_2 \cdot \eta X)\hat{f}_{XX^\top}\right) \\ &= \mathrm{Tr}\left(h_0\eta^\top Q\eta\right)\hat{f} + X^\top\eta^\top Q\eta h_0\eta^\top Q\eta X\hat{f} + d^\top\left(h_0 + h_1 \cdot X + X^\top\eta^\top \cdot h_2 \cdot \eta X\right)d\hat{f} \\ &+ \left(d^\top h_0\eta^\top Q\eta X + X^\top\eta^\top Q\eta h_0 d\right)\hat{f}.\end{aligned} \tag{A2}$$

Second, noting that $K\eta^\top = \eta^\top\hat{K}$, $K_2\eta^\top = 0$, $g_1^\top\eta^\top = \eta^\top\hat{g}_1$, $g_2^\top\eta^\top = 0$, and

$$\hat{f}_X = (d + \eta^\top Q\eta X)\hat{f},$$

we get

$$\begin{aligned}&\left[k - KX + \frac{1}{2}X^\top \eta^\top \cdot K_2 \cdot \eta X + \frac{1-\gamma}{\gamma}\left(g_0 + g_1 X + \frac{1}{2}X^\top \eta^\top \cdot g_2 \cdot \eta X\right)\right]^\top \hat{f}_X \\ &= \left[k - KX + \frac{1}{2}X^\top \eta^\top \cdot K_2 \cdot \eta X + \frac{1-\gamma}{\gamma}\left(g_0 + g_1 X + \frac{1}{2}X^\top \eta^\top \cdot g_2 \cdot \eta X\right)\right]^\top d\hat{f} \\ &+ \left(k + \frac{1-\gamma}{\gamma} g_0\right)^\top \eta^\top Q \eta X \hat{f} - X^\top \eta^\top \hat{K} Q \eta X \hat{f} + \frac{1-\gamma}{\gamma} X^\top \eta^\top \hat{g}_1 Q \eta X \hat{f}. \end{aligned} \quad \text{(A3)}$$

Finally,

$$\frac{\partial \hat{f}}{\partial t} = \left(\frac{d}{dt}c + \frac{d}{dt}d^\top X + \frac{1}{2}X^\top \eta^\top \frac{d}{dt} Q \eta X\right)\hat{f}. \quad \text{(A4)}$$

Substituting Equations (A2)–(A4) into PDE (A1), the PDE's left-hand side has three classes of terms: the first class is independent of X, the second class is linear in X, and the third class is in the quadratic in ηX. For the equation to be true for all X, the three classes of term have to be zero separately, which leads to Equations (18)–(20). It is obvious that $\hat{f}(T,X) = 1$ since $c(T) = 0$, $d(T) = 0$, and $Q(T) = 0$. QED

Note that restrictions $h_1 \eta^\top = 0$, $h_2 \eta^\top = 0$, $K\eta^\top = \eta^\top \hat{K}$, $K_2 \eta^\top = 0$, $g_1^\top \eta^\top = \eta^\top \hat{g}_1$, and $g_2^\top \eta^\top = 0$ are essential for the left-hand side of Equation (A1) to be quadratic in ηX and linear in X.

Lemma 2. *Suppose that*

$$\frac{\partial \hat{f}}{\partial t} + \mathcal{L}\hat{f} = 0,$$

and $\hat{f}(T,X) = 1$. *Then the function f defined by*

$$f(t,X) = \alpha^{\frac{1}{\gamma}} \int_t^T \hat{f}(u,X) du + (1-\alpha)^{\frac{1}{\alpha}} \hat{f}(t,X)$$

satisfies

$$\frac{\partial f}{\partial t} + \mathcal{L}f + \alpha^{\frac{1}{\gamma}} = 0$$

and $f(T,X) = (1-\alpha)^{\frac{1}{\gamma}}$.

Proof. It is obvious that $f(T,X) = (1-\alpha)^{\frac{1}{\gamma}}$. Furthermore,

$$\begin{aligned}\frac{\partial f}{\partial t} + \mathcal{L}f &= -\alpha^{\frac{1}{\gamma}} \hat{f}(t,X) + \alpha^{\frac{1}{\gamma}} \int_t^T \mathcal{L}\hat{f}(u,X) du = -\alpha^{\frac{1}{\gamma}} \hat{f}(t,X) - \alpha^{\frac{1}{\gamma}} \int_t^T \frac{\partial \hat{f}}{\partial u} du \\ &= -\alpha^{\frac{1}{\gamma}} \hat{f}(t,X) - \alpha^{\frac{1}{\gamma}} \left[\hat{f}(T,X) - \hat{f}(t,X)\right] = -\alpha^{\frac{1}{\gamma}}, \end{aligned}$$

where the assumption $\frac{\partial \hat{f}}{\partial t} + \mathcal{L}\hat{f} = 0$ is used in the first and second equality. QED

Proof of Proposition 1. Using the specification of μ^X and Σ^X and Condition 1, the PDE (7) can be written as

$$\frac{\partial f}{\partial t} + \mathcal{L}f + \alpha^{\frac{1}{\gamma}} = 0.$$

Therefore, f specified in the proposition is the solution to PDE (7) by Lemmas 1 and 2. The optimal consumption and portfolio weights are obtained using Equations (5) and (6). QED

Proof of Proposition 2. The PDE satisfied by f in this case is in the same form (with different coefficients) as the PDE for $\hat{f}$. So the proof is exactly the same. The optimal portfolio weight is obtained using Equations (5) and (6). QED

Lemma 3. *Suppose that the state variable X can be partitioned into X_1 and X_2 and the assets are also partitioned into two classes, such that $r = r_1 + r_2$,*

$$\Sigma^X = \begin{pmatrix} \Sigma^{X_1} & 0 \\ 0 & \Sigma^{X_2} \end{pmatrix},\ \mu^X = \begin{pmatrix} \mu^{X_1} \\ \mu^{X_2} \end{pmatrix},\ \rho = \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix},\ \Sigma^{-1}(\mu - r) = \begin{pmatrix} \Sigma_1^{-1}(\mu_1 - r_1) \\ \Sigma_2^{-1}(\mu_2 - r_2) \end{pmatrix},$$

where r_i, Σ^{X_i}, μ^{X_i}, and $\Sigma_i^{-1}(\mu_i - r)$ depend only on X_i. Then the solution to PDE

$$\begin{aligned}
&\frac{\partial f}{\partial t} + \frac{1}{2}\mathrm{Tr}(\Sigma^X \Sigma^{X\top} f_{XX^\top}) + \left[\mu^X + \frac{1-\gamma}{\gamma}\Sigma^X \rho \Sigma^{-1}(\mu - r)\right]^\top f_X \\
&\quad + \frac{1}{2f}(\gamma - 1) f_X^\top \left(\Sigma^X \Sigma^{X\top} - \Sigma^X \rho\rho^\top \Sigma^{X\top}\right) f_X \\
&\quad + \left[\frac{1-\gamma}{2\gamma^2}(\mu - r)^\top (\Sigma\Sigma^\top)^{-1}(\mu - r) + \frac{1-\gamma}{\gamma} r\right] f = 0,
\end{aligned}$$

with the initial condition $f(T,X) = 1$ is given by

$$f(X,t) = f_1(X_1,t) f_2(X_2,t)$$

where $f_i(X_1,t)$ satisfies:

$$\begin{aligned}
&\frac{\partial f_i}{\partial t} + \frac{1}{2}\mathrm{Tr}\left(\Sigma^{X_i} \Sigma^{X_i\top} f_{iX_iX_i^\top}\right) + \left[\mu^{X_i} + \frac{1-\gamma}{\gamma}\Sigma^{X_i} \rho_i \Sigma_i^{-1}(\mu_i - r_i)\right]^\top f_{X_i} \\
&\quad + \frac{1}{2f_i}(\gamma - 1) f_{X_i}^\top \left(\Sigma^{X_i} \Sigma^{X_i\top} - \Sigma^{X_i} \rho_i \rho_i^\top \Sigma^{X_i\top}\right) f_{X_i} \\
&\quad + \left[\frac{1-\gamma}{2\gamma^2}(\mu_i - r_i)^\top (\Sigma_i \Sigma_i^\top)^{-1}(\mu_i - r_i) + \frac{1-\gamma}{\gamma} r_i\right] f_i = 0,
\end{aligned} \tag{A5}$$

with the initial condition $f_i(T,X_i) = 1$, $i = 1,2$.

Proof. It is straightforward to verify that the PDE for f is satisfied with the assumed form of $f = f_1 f_2$. QED

Proof of Proposition 3. The functional form of f is obtained as an application of Lemma 3 and Proposition 1. Noting that

$$\Sigma^{-1\top} = \begin{pmatrix} \Sigma_1^{-1\top} & -\nu^\top \Sigma_2^{-1\top} \\ 0 & \Sigma_2^{-1\top} \end{pmatrix},$$

the optimal portfolio weight is obtained. QED

Proof of Corollary 2. The proof is exactly the same as that of Corollary 3 below.

Proof of Corollary 3. The Equation (32) can be written as

$$\frac{d}{dt} d_v - \tilde{K}_v d_v + \frac{1}{2}\left[\rho_v^2 + \gamma(1-\rho_v^2)\right]\sigma_v^2 d_v^2 - \delta_v = 0,$$

$d_v(T) = 0$. Define $D(t) = \frac{d_v(t)}{\delta_v}$ and $\hat{\sigma}^2 = \delta_v[\rho_v^2 + \gamma(1-\rho_v^2)]\sigma_v^2$. The function $D(t)$ satisfies the following ODE

$$\frac{d}{dt} D - \tilde{K}_v D + \frac{1}{2}\hat{\sigma}^2 D^2 - 1 = 0,$$

$D(T) = 0$. This ODE is solved by Cox, Ingersoll, and Ross (1985) as

$$D(t) = -\frac{2[\exp(\xi_v \tau) - 1]}{(\tilde{K}_v + \xi_v)[\exp(\xi_v \tau) - 1] + 2\xi_v},$$

with $\tau = T - t$ and $\xi_v = \sqrt{\tilde{K}_v^2 + 2\hat{\sigma}^2}$. Substituting the definition of $D(t)$ and $\hat{\sigma}^2$, we get

$$d(t) = -\frac{2[\exp(\xi_v \tau) - 1]}{(\tilde{K}_v + \xi_v)[\exp(\xi_v \tau) - 1] + 2\xi_v}\delta_1,$$

where $\xi_v = \sqrt{\tilde{K}_v^2 + 2[\rho_v^2 + \gamma(1-\rho_v^2)]\sigma_v^2}$.

When $\tilde{K}_v^2 + 2[\rho_v^2 + \gamma(1-\rho_v^2)]\sigma_v^2 < 0$, ξ_v is imaginary, although d_v is still real. It is more convenient to express d in terms of real variables. To accomplish this, we define $\zeta_v = -i\xi_v$, which is real when ξ_v is imaginary. In terms of ζ_v, the function $d_v(t)$ can be expressed as

$$\begin{aligned} d_v(t) &= -\frac{2[\exp(i\zeta_v \tau) - 1]}{(\tilde{K}_v + i\zeta_v)[\exp(i\zeta_v \tau) - 1] + 2i\zeta_v}\delta_v \\ &= -\frac{2\sin(\zeta_v \tau/2)}{\tilde{K}_v \sin(\zeta_v \tau/2) + \zeta_v \cos(\zeta_v \tau/2)}\delta_v = -\frac{2}{\tilde{K}_v + \zeta_v \frac{\cos(\zeta_v \tau/2)}{\sin(\zeta_v \tau/2)}}\delta_v. \end{aligned}$$

From the last equality, the function d_v is manifestly real even when ξ_v is imaginary. Note that $d_v \to +\infty$ at a finite τ in this case. QED

References

Bates, D. S., 2000, “Post-’87 Crash Fears in S&P 500 Futures Options,” *Journal of Econometrics*, 94, 181–238.

Brennan, M. J., 1998, “The Role of Learning in Dynamic Portfolio Decisions?” *European Finance Review*, 1, 295–306.

Brennan, M. J., E. S. Schwartz, and R. Lagnado, 1997, “Strategic Asset Allocation,” *Journal of Economic Dynamics and Control*, 21, 1377–1403.

Brennan, M. J., and Y. Xia, 2000, "Stochastic Interest Rates and Bond-Stock Mix," *European Finance Review*, 4, 197–210.

Brennan, M. J., and Y. Xia, 2001, "Assessing Asset Pricing Anomalies," *Review of Financial Studies*, 14, 905–942.

Brennan, M. J., and Y. Xia, 2002a, "Dynamic Asset Allocation under Inflation," *Journal of Finance*, 57, 1201–1238.

Brennan, M. J., and Y. Xia, 2002b, "Optimal Investment Policies: Fixed Mix and Dynamic Stochastic Control Models," working paper, UCLA.

Campbell, J. Y., Cochrane, J. H., 1999, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

Campbell, J. Y., and L. M. Viceira, 2001, "Who Should Buy Long-Term Bonds?" *The American Economic Review*, 91, 99–127.

Canner, N., N. G. Mankiw, and D. N. Weil, 1997, "An Asset Allocation Puzzle," *The American Economic Review*, 87, 181–191.

Chacko, G., and L. M. Viceira, 2005, "Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets," *Review of Financial Studies*, 18, 1369–1402.

Cochrane, J. H., and J. Saá-Requejo, 2001, "Beyond Arbitrage: 'Good-Deal' Asset Price Bounds in Incomplete Markets," *Journal of Political Economy*, 108, 79–119.

Collin-Dufresne, P., and R. Goldstein, 2003, "Generalizing the Affine Framework to HJM and Random Field Models," working paper, Univeristy of California.

Constantinides, G. M., 1992, "A Theory of the Nominal Term Structure of Interest Rates," *Review of Financial Studies*, 5, 531–552.

Cox, J. C., J. E. Ingersoll, and S. A. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385–407.

Duffie, D., and R. Kan, 1996, "A Yield-Factor Model of Interest Rates," *Mathematical Finance*, 6, 379–406.

Fama, E. F., and G. W. Schwert, 1977, "Asset Returns and Inflation," *Journal of Financial Economics*, 5, 115–146.

Ferson, W. E., and C. R. Harvey, 1993, "The Risk and Predictability of International Equity Returns," *Review of Financial Studies*, 6, 527–566.

French, K. R., G. W. Schwert, and R. F. Stambaugh, 1987, "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19, 3–29.

Glosten, L. R., R. Jagannathan, and D. E. Runkle, 1993, "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48, 1779–1801.

Harvey, C. R., 1989, "Time-Varying Conditional Covariances in Tests of Asset Pricing Models," *Journal of Financial Economics*, 24, 289–317.

Heston, S. L., 1993, "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327–343.

Kargin, V., 2003, "Portfolio Management for a Random Field of Bond Returns," working paper, Cornerston Research.

Kim, T. S., and E. Omberg, 1996, "Dynamic Nonmyopic Portfolio Behavior," *Review of Financial Studies*, 9, 141–161.

Kogan, L., and R. Uppal, 2000, "Risk Aversion and Optimal Portfolio Policies in Partial and General Equilibrium Economies," working paper, UBC Commerce.

Mamaysky, H., 2002, "Market Prices of Risk and Return Predictability in a Joint Stock-Bond Pricing Model," working paper, Yale School of Management.

Merton, R. C., 1969, "Lifetime Portfolio Selection: The Continuous-Time Case," *Review of Economics and Statistics*, 51, 247–257.

Merton, R. C., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 3, 373–413.

Pagan, A. R., and G. W. Schwert, 1990, "Alternative Models for Conditional Stock Volatility," *Journal of Econometrics*, 45, 267–290.

Pan, J., 2002, "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study," *Journal of Financial Economics*, 63, 3–50.

Sangvinatsos, A., and J. Wachter, 2005, "Does the Failure of the Expectations Hypothesis Matter for Long-Term Investors?" *Journal of Finance*, 60, 179–230.

Schroder, M., and C. Skiadas, 1999, "Optimal Consumption and Portfolio Selection with Stochastic Differential Utility," *Journal of Economic Theory*, 89, 68–126.

Schroder, M., and C. Skiadas, 2003, "Optimal Lifetime Consumption-Portfolio Strategies under Trading Constraints and Generalized Recursive Preferences," *Stochastic Processes and their Applications*, 108, 155–202.

Schroder, M., and C. Skiadas, 2005, "Lifetime Consumption-Portfolio Choice under Trading Constraints, Recursive Preferences, and Nontradeable Income," *Stochastic Processes and their Applications*, 115, 130.

Scruggs, J. T., 1998, "Resolving the Puzzling Intertemporal Relation between the Market; Risk Premium and Conditional Market Variance: A Two-Factor Approach," *Journal of Finance*, 53, 575–603.

Tehranchi, M., and N. Ringer, 2004, "Optimal Portfolio Choice in the Bond Market," working paper, University of Texas.

Wachter, J., 2003, "Risk Aversion and Allocation to Long Term Bonds," *Journal of Economic Theory*, 112, 325–333.