Optimal Portfolios of Mean-Reverting Instruments[*]

Gordana Dmitrašinović-Vidović[†] and Antony Ware[‡]

Abstract. In this paper we investigate portfolios consisting of instruments whose logarithms are mean-reverting. Under the assumption that portfolios are constant, we derive analytic expressions for the expected wealth and the quantile-based risk measure capital at risk. Assuming that short-selling and borrowing are allowed, we then solve the problems of global minimum capital at risk and the problem of finding maximal wealth subject to constrained capital at risk. We illustrate these results with some numerical examples that show the strong effect of the mean-reversion rates on the portfolio choice.

Key words. mean-reverting process, portfolio optimization, quantile, capital at risk

AMS subject classifications. 91B28, 93E20

DOI. 10.1137/100787714

1. Introduction. Mean-reversion has received considerable attention in the financial world as a classic indicator of predictability in financial markets. In this paper we investigate portfolios consisting of mean-reverting instruments, more specifically, exponential Ornstein–Uhlenbeck processes, or one-factor Schwartz processes (see [22]) that belong to a broader class of processes that follow affine distributions. We note that affine distributions are often used to model the dynamics of commodity spot and futures markets, interest rates, and exchange rates (see [1], [6], [5], [17], [18], [20], [21], [22]). The problem of portfolio optimization in this setting has been investigated in a number of papers (see, for instance, [7], [3], [4], [23]). In most of the cited papers the authors analyze the optimal investment problem in a financial market where the risky asset follows the price dynamics of Schwartz and the risk preferences are described by some utility function. It has been claimed (see [4], [8], [15]) that the predictability of asset returns affects the choice of optimal portfolio and yields significant improvements in portfolio performance.

In our paper we utilize the risk reward approach, with risk measured by capital at risk (CaR), defined as the difference between the riskless wealth and the α-quantile, which was introduced in [14] and further investigated in [12], [9]. Under the assumption that portfolios are constant, we derive analytic expressions for the expected value and CaR of the corresponding wealth process. We then solve the problems of global minimum CaR, and the problem of finding maximal wealth subject to constrained CaR, assuming that short-selling of risky assets and borrowing of the riskless asset are unconstrained.

[†]The Mathematical and Computational Finance Laboratory, University of Calgary, 2500 University Drive NW, Calgary T2N 1N4, AB, Canada (gdmitras@ucalgary.ca).

[‡]Corresponding author. The Mathematical and Computational Finance Laboratory, University of Calgary, 2500 University Drive NW, Calgary, T2N 1N4, AB, Canada (aware@ucalgary.ca).

The outline of this paper is as follows. In section 2, we give the notation, the market setting, the definition of the portfolio process, and the dynamics of the wealth process and the assets' logarithms. In section 3, we derive analytic expressions for the log-wealth process, the expected value, and CaR of the wealth process. In section 4 we show that CaR is strongly quasi-convex as a function of the portfolio, which is an essential property for optimization since it guarantees the uniqueness of the optimal solution. We then find solutions to the global minimum CaR, and the maximal expected wealth subject to constrained CaR. Finding these solutions numerically is done in section 5 for some specific examples that show a strong effect of the mean-reversion rates on the portfolio choice. Finally, section 6 concludes the paper.

2. Preliminaries. In this section we give the market setting, define the portfolio and the wealth process, and investigate the dynamics of the underlying assets.

Throughout the paper we denote vectors and matrices by bold letters, their transposes by $(\cdot)'$, and the Euclidean norm of a matrix or vector by $\|\cdot\|$.

We assume that the following conditions are satisfied.

Assumption 2.1.
  (i) The securities are perfectly divisible.
  (ii) Negative positions in securities are possible.
  (iii) Rebalancing of the holdings does not lead to transaction costs.

In addition to the above assumptions, we make the following assumption.

Assumption 2.2.
  (i) $m + 1$ assets are traded continuously over a finite horizon $[0, T]$.
  (ii) m of these assets follow the mean-reverting dynamics given in the following system of stochastic differential equations:

$$(2.1) \qquad \frac{\mathrm{d}S_i(t)}{S_i(t)} = \beta_i(L_i - c_i \ln S_i(t))\mathrm{d}t + \sum_j \sigma_{ij}\mathrm{d}W_j(t), \quad S_i(0) > 0, \quad i = 1, \ldots, m.$$

We recall that this model is called a one-factor Schwartz model (see [22]). In this model the mean-reversion rate $\beta_i > 0$ and $L_i \in \mathbb{R}$, $i = 1, \ldots, m$. Further, $c_i = 0$ or 1, $i = 1, \ldots, m$, are scaling factors which allow investigating optimal portfolios in this setting, in the Black–Scholes setting, or in a combination of both. $\boldsymbol{W}(t) := \big(W_1(t), \ldots, W_m(t)\big)'$ is the m-dimensional vector of independent Brownian motions, and $\boldsymbol{\sigma} := (\sigma_{ij})$ is the volatility matrix.

  (iii) One of the assets is riskless, and its value $S_0(t)$, $t \geq 0$, is equal to

$$(2.2) \qquad\qquad\qquad\qquad\qquad S_0(t) = \mathrm{e}^{rt},$$

where $r > 0$ is the interest rate of the asset. Throughout this work, we assume that the holdings in this asset are unconstrained and, in particular, can be positive or negative.

  (iv) $\boldsymbol{\sigma}$ is invertible.

If we define

$$a_i = \beta_i L_i, \quad b_i = \beta_i c_i, \quad i = 1, \ldots, m, \quad \boldsymbol{\sigma_i} = (\sigma_{i1}, \ldots, \sigma_{im}),$$

the system (2.1) can be written as

$$(2.3) \qquad \frac{\mathrm{d}S_i(t)}{S_i(t)} = (a_i - b_i \ln S_i(t))\mathrm{d}t + \boldsymbol{\sigma_i}\mathrm{d}\boldsymbol{W}(t), \quad i = 1, \ldots, m.$$

At any time t, $N_i(t)$ shares are held in the asset $S_i(t)$, leading to the wealth $X^{\boldsymbol{\pi}}(t) = \sum_{i=0}^{m} N_i(t)S_i(t)$. The $(m+1)$-dimensional vector-valued function $\boldsymbol{N}(t) = (N_0(t), \ldots, N_m(t))'$ is called the *trading strategy*. We denote the fraction of the wealth $X^{\boldsymbol{\pi}}(t)$ invested in the risky asset $S_i(t)$ by

$$\pi_i(t) = \frac{N_i(t)S_i(t)}{X^{\boldsymbol{\pi}}(t)} \in \mathbb{R}, \quad i = 1, \ldots, m.$$

Assumption 2.3. In order to maintain tractability, we assume that $\pi_i(t) \equiv \pi$ is constant for each i, and we call $\boldsymbol{\pi} := (\pi_1, \ldots, \pi_m)' \in \mathbb{R}^m$ the *portfolio*. The fraction held in the bond is $\pi_0 = 1 - \boldsymbol{\pi}'\boldsymbol{1}$.

Remark 2.1. We note that the constant-portfolio assumption is somewhat restrictive and has been made in order to maintain tractability. The current setting differs from other contributions in the area, such as [7], [3], [4], [23], which do not make this assumption and which use utility optimization with a logarithmic or power utility function; here the portfolio is n-dimensional, and the optimization is in a risk-reward setting with the risk measured by CaR. The difficulties of such an approach have been pointed out by Emmer, Klüppelberg, and Korn, who state (see [14, p. 369]) that "determining the wealth process quantile for a general, random n-dimensional portfolio is nearly impossible." Emmer and Klüppelberg [13] investigate portfolio optimization problems under CaR, when prices follow exponential Lévy processes, but also assuming constant portfolios. They derive a weak limit law for CaR's approximation and numerical algorithms for the corresponding solutions, but no closed form solution.

Under the assumption that the trading strategy is *self-financing,* the wealth process (see [19]) follows the dynamics

$$\mathrm{d}X^{\boldsymbol{\pi}}(t) = \sum_{i=0}^{m} N_i(t)dS_i(t)$$

$$= N_0 S_0(t)rdt + \sum_{i=1}^{m} N_i(t)S_i(t)\left[(a_i - b_i \ln S_i(t))\mathrm{d}t + \boldsymbol{\sigma_i}\mathrm{d}\boldsymbol{W}(t)\right]$$

$$(2.4) \qquad = X^{\boldsymbol{\pi}}(t)\left(rdt + \sum_{i=1}^{m} \pi_i \left[(a_i - r - b_i \ln S_i(t))\mathrm{d}t + \boldsymbol{\sigma_i}\mathrm{d}\boldsymbol{W}(t)\right]\right), \quad X^{\boldsymbol{\pi}}(0) = X_0,$$

where $X_0 > 0$ is the initial wealth.

In order to facilitate further investigation, we define $Y_i(t) := \ln S_i(t)$, $i = 1, \ldots, m$. The characteristics of the processes $Y_i(t)$ are given in the following proposition.

Proposition 2.1. *The means and covariances of the normally distributed random variables* $Y_i(t)$, $i = 1, \ldots, m$, *are*

$$\mathbb{E}\left[Y_i(t)\right] = Y_i(0)e^{-b_i t} + \hat{a}_i \mathbb{E}\left(t, b_i\right),$$

$$\mathrm{Cov}[Y_i(t), Y_j(t)] = \boldsymbol{\sigma_i}\boldsymbol{\sigma_j}'\mathbb{E}\left(t, b_i + b_j\right),$$

where

$$\hat{a}_i = a_i - \frac{1}{2} \|\boldsymbol{\sigma_i}\|^2$$

and

(2.5)
$$\mathbb{E}(t,b) := \int_0^t \mathrm{e}^{-sb}\mathrm{d}s = \begin{cases} \frac{1-\mathrm{e}^{-tb}}{b} & \text{if } b \neq 0, \\ t & \text{if } b = 0. \end{cases}$$

Proof. Applying Itô's lemma to (2.3) we find that

(2.6)
$$\begin{aligned} dY_i(t) &= (a_i - b_i Y_i(t))\mathrm{d}t + \boldsymbol{\sigma_i}d\mathbf{W}(t) - \frac{1}{2}\|\boldsymbol{\sigma_i}\|^2\,\mathrm{d}t \\ &= (\hat{a}_i - b_i Y_i(t))\mathrm{d}t + \boldsymbol{\sigma_i}\mathrm{d}\mathbf{W}(t), \quad i = 1, \ldots, m. \end{aligned}$$

The substitution

$$Z_i(t) = \mathrm{e}^{b_i t}Y_i(t)$$

yields

$$\begin{aligned} dZ_i(t) &= b_i \mathrm{e}^{b_i t}Y_i(t)\mathrm{d}t + \hat{a}_i \mathrm{e}^{b_i t}\mathrm{d}t - b_i \mathrm{e}^{b_i t}Y_i(t)\mathrm{d}t + \mathrm{e}^{b_i t}\boldsymbol{\sigma_i}\mathrm{d}\mathbf{W}(t) \\ &= \mathrm{e}^{b_i t}\hat{a}_i\mathrm{d}t + \mathrm{e}^{b_i t}\boldsymbol{\sigma_i}\mathrm{d}\boldsymbol{W}(t). \end{aligned}$$

Thus

$$Z_i(t) = Z_i(0) + \hat{a}_i\mathbb{E}(t, -b_i) + \boldsymbol{\sigma_i}\int_0^t \mathrm{e}^{b_i u}\mathrm{d}\boldsymbol{W}(u).$$

This leads to

(2.7)
$$Y_i(t) = \mathrm{e}^{-b_i t}Y_i(0) + \hat{a}_i\mathbb{E}(t, b_i) + \mathrm{e}^{-b_i t}\boldsymbol{\sigma_i}\int_0^t \mathrm{e}^{b_i u}\mathrm{d}\boldsymbol{W}(u).$$

The expression for $\mathbb{E}[Y_i(t)]$ follows immediately from (2.7). To evaluate the covariances of $Y_i(t)$, we expand the expression for $\mathrm{Cov}[Y_i(t), Y_j(t)]$, make use of Itô's isometry, and find that

(2.8)
$$\begin{aligned} \mathrm{Cov}[Y_i(t), Y_j(t)] &= \mathbb{E}\left[\left(\mathrm{e}^{-b_i t}\boldsymbol{\sigma_i}\int_0^t \mathrm{e}^{b_i u}\mathrm{d}\boldsymbol{W}(u)\right)\left(\mathrm{e}^{-b_j t}\boldsymbol{\sigma_j}\int_0^t \mathrm{e}^{b_j u}\mathrm{d}\boldsymbol{W}(u)\right)\right] \\ &= \mathrm{e}^{-(b_i+b_j)t}\mathbb{E}\left[\left(\sum_k \sigma_{ik}\int_0^t \mathrm{e}^{b_i u}\mathrm{d}\boldsymbol{W}_k(u)\right)\left(\sum_l \sigma_{jl}\int_0^t \mathrm{e}^{b_j u}\mathrm{d}\boldsymbol{W}_l(u)\right)\right] \\ &= \mathrm{e}^{-(b_i+b_j)t}\left(\sum_k \sigma_{ik}\sigma_{jk}\int_0^t \mathrm{e}^{(b_i+b_j)u}\mathrm{d}u\right) \\ &= \boldsymbol{\sigma_i}\boldsymbol{\sigma_j}'\mathbb{E}(t, b_i + b_j), \end{aligned}$$

which completes the proof of the proposition. ∎

Proposition 2.2. *Let the matrix* $\mathbf{F}(t, \boldsymbol{b})$ *have the entries*

(2.9)
$$F_{ij}(t, \boldsymbol{b}) := \mathrm{Cov}[Y_i(t), Y_j(t)] = \boldsymbol{\sigma_i}\boldsymbol{\sigma_j}'\mathbb{E}(t, b_i + b_j), \quad i, j = 1, \ldots, m.$$

Then, for any $t > 0$, $\mathbf{F}(t, \boldsymbol{b})$ is positive-definite.

Proof. Let $\boldsymbol{x}$ be an arbitrary vector in $\mathbb{R}^m$, and define, for $s \in \mathbb{R}$, the vector $\boldsymbol{z}(s)$ as having entries

$$z_i(s) = x_i \mathrm{e}^{-sb_i}.$$

Then

$$
\begin{aligned}
\boldsymbol{x}' \boldsymbol{F}(t, \boldsymbol{b}) \boldsymbol{x} = \sum_{i,j} x_i x_j F_{ij}(t, \boldsymbol{b}) &= \sum_{i,j} x_i x_j \boldsymbol{\sigma_i} \boldsymbol{\sigma_j}' \mathbb{E}(t, b_i + b_j) \\
&= \sum_{i,j} x_i x_j \boldsymbol{\sigma_i} \boldsymbol{\sigma_j}' \int_{s=0}^{t} \mathrm{e}^{-s(b_i + b_j)} \mathrm{d}s \\
&= \int_{s=0}^{t} \sum_{i,j} x_i \mathrm{e}^{-sb_i} \boldsymbol{\sigma_i} x_j \mathrm{e}^{-sb_j} \boldsymbol{\sigma_j}' \mathrm{d}s \\
&= \int_{s=0}^{t} \|\boldsymbol{\sigma} \boldsymbol{z}(s)\|^2 \, \mathrm{d}s \geq 0.
\end{aligned}
$$

Since $\boldsymbol{z}(\cdot)$ is a continuous function and $\boldsymbol{\sigma}$ is invertible, this integral will be zero if and only if $\boldsymbol{z}(s) = \boldsymbol{0}$ for all $s \in [0, t]$, i.e., if $\boldsymbol{x} = \boldsymbol{0}$. ∎

Remark 2.2.

(i) We note that the market described above consists of processes that follow affine distributions, and contains as a special case (when $c_i = 0$, $i = 1, \ldots, m$) the setting of [9], [10], [12], and [14], where the assets are lognormal, in which case the matrix $\mathbf{F}(t, \boldsymbol{b})$ reduces to $\boldsymbol{\sigma}\boldsymbol{\sigma}'t$.

(ii) Affine processes have been used extensively in financial modeling. This is due to their tractability and flexibility in capturing a range of assets' dynamics. In particular, affine processes are used to model the term structure of interest rates (see [1], [16]). Further, energy spot and forward price dynamics are captured by mean-reversion models that track their tendency to revert to a price level determined by the cost of production (see [21], [22]). Also, mean-reverting behavior is observed in the real exchange rate series (see [18]).

3. Wealth and capital at risk. In this section we will use the results of Proposition 2.1 to determine the mean and CaR of the wealth process $X^{\boldsymbol{\pi}}(t)$. We recall that the CaR of $X^{\boldsymbol{\pi}}(t)$ was given in [14] by

$$X_0 \mathrm{e}^{rt} - q_x,$$

where q_x is the α-quantile of $X^{\boldsymbol{\pi}}(t)$, X_0 denotes the initial wealth invested in the portfolio, and r denotes the riskless rate of return.

In order to determine the CaR of the wealth process $X^{\boldsymbol{\pi}}(t)$, we first investigate the log-wealth process $H(t) := \ln X^{\pi}(t)$ and prove the following proposition.

Proposition 3.1. *The log-wealth process $H(t)$ follows the dynamics*

$$dH(t) = \mu dt + \sum_{i=1}^{m} \pi_i \left[-b_i Y_i(t) dt + \boldsymbol{\sigma_i} d\boldsymbol{W}(t) \right], \quad H(0) = \ln X_0,$$

where

(3.1)
$$\mu = r + \boldsymbol{\pi}'(\boldsymbol{a} - r\mathbf{1}) - \frac{1}{2}\left\|\boldsymbol{\pi}'\boldsymbol{\sigma}\right\|^2.$$

The mean and variance of $H(t)$ are given by

$$\mathbb{E}[H(t)] = H(0) + \mu t + \boldsymbol{\pi}'\mathcal{A}(t, \boldsymbol{b})$$

and

$$\mathbb{V}[H(t)] = \boldsymbol{\pi}'\mathbf{F}(t, \boldsymbol{b})\boldsymbol{\pi},$$

where

(3.2)
$$\mathcal{A}_i = \mathbb{E}\left(t, b_i\right)(\hat{a}_i - b_i Y_i(0)) - \hat{a}_i t, \quad i = 1, \ldots, m.$$

Proof. We can rewrite (2.4) in the following form:

$$dX^\pi(t) = X^\pi(t)\left(\left(\upsilon + \sum_{i=1}^m \pi_i(-b_i Y_i(t))\right)dt + \boldsymbol{\pi}'\boldsymbol{\sigma}d\boldsymbol{W}(t)\right),$$

where $\upsilon = r + \boldsymbol{\pi}'(\boldsymbol{a} - r\mathbf{1})$. Applying Itô's lemma to the process $H(t)$, we get

$$dH(t) = \left(\upsilon + \sum_{i=1}^m \pi_i(-b_i Y_i(t))\right)dt + \boldsymbol{\pi}'\boldsymbol{\sigma}d\boldsymbol{W}(t) - \frac{1}{2}\left\|\boldsymbol{\pi}'\boldsymbol{\sigma}\right\|^2 dt.$$

Using (3.1) we get that

$$dH(t) = \mu dt + \sum_{i=1}^m \pi_i\left[-b_i Y_i(t)dt + \boldsymbol{\sigma}_i d\boldsymbol{W}(t)\right].$$

From (2.6), we have

$$-b_i Y_i(t)dt + \boldsymbol{\sigma}_i d\boldsymbol{W}(t) = dY_i(t) - \hat{a}_i dt$$

so that we can write

$$dH(t) = \mu dt + \sum_{i=1}^m \pi_i\left[dY_i(t) - \hat{a}_i dt\right].$$

Thus

$$dH(t) = (\mu - \boldsymbol{\pi}'\hat{\boldsymbol{a}})dt + \boldsymbol{\pi}'d\boldsymbol{Y}(t),$$

and so

$$H(t) = H(0) + (\mu - \boldsymbol{\pi}'\hat{\boldsymbol{a}})t + \boldsymbol{\pi}'\left(\boldsymbol{Y}(t) - \boldsymbol{Y}(0)\right).$$

From (2.7) we get that

$$Y_i(t) - Y_i(0) = \left(\hat{a}_i - b_i Y_i(0)\right)\mathbb{E}\left(t, b_i\right) + e^{-b_i t}\boldsymbol{\sigma}_i \int_0^t e^{b_i u}d\boldsymbol{W}(u),$$

so that we can write

$$(3.3) \quad H(t) = H(0) + (\mu - \boldsymbol{\pi}'\hat{\boldsymbol{a}})t + \sum_{i=1}^{m} \pi_i \mathbb{E}(t, b_i)(\hat{a}_i - b_i Y_i(0)) + \sum_{i=1}^{m} \pi_i \mathrm{e}^{-b_i t} \boldsymbol{\sigma_i} \int_0^t \mathrm{e}^{b_i u} \mathrm{d}\boldsymbol{W}(u).$$

Thus

$$\mathbb{E}[H(t)] = H(0) + (\mu - \boldsymbol{\pi}'\hat{\boldsymbol{a}})t + \sum_{i=1}^{m} \pi_i \mathbb{E}(t, b_i)\big(\hat{a}_i - b_i Y_i(0)\big)$$

$$= H(0) + \mu t + \boldsymbol{\pi}'\boldsymbol{\mathcal{A}}(t, \boldsymbol{b}),$$

where $\boldsymbol{\mathcal{A}}(t, \boldsymbol{b})$ is defined in (3.2). Since

$$\mathbb{V}[H(t)] = \mathbb{E}[H(t)^2] - \mathbb{E}[H(t)]^2$$

$$= \mathbb{E}\left[\left(\sum_{i=1}^{m} \pi_i \mathrm{e}^{-b_i t} \boldsymbol{\sigma_i} \int_0^t \mathrm{e}^{b_i u} \mathrm{d}\boldsymbol{W}(u)\right)^2\right],$$

from the evaluation of $F_{ij}(t, \boldsymbol{b})$ in (2.8) we get that

$$\mathbb{V}[H(t)] = \boldsymbol{\pi}'\boldsymbol{F}(t, \boldsymbol{b})\boldsymbol{\pi},$$

with matrix $\boldsymbol{F}$ given by (2.9), which completes the proof of the proposition. ■

We now turn our attention to determining the α-quantile of $H(t)$. Given that the α-quantile q_u of a normal random variable $U \sim N\left(\mu, \sigma^2\right)$ is

$$q_u = \mu - |z_\alpha| \, \sigma,$$

where z_α denotes the corresponding α-quantile of the standard normal distribution, the α-quantile of the log-wealth process $H(t)$ is

$$q_h = \mathbb{E}[H(t)] - |z_\alpha| \, \sqrt{\mathbb{V}[H(t)]} = H(0) + rt + \boldsymbol{\pi}' \left[(\boldsymbol{a} - r\boldsymbol{1})t + \boldsymbol{\mathcal{A}}(t, \boldsymbol{b})\right]$$

$$- \frac{t}{2} \left\|\boldsymbol{\pi}'\boldsymbol{\sigma}\right\|^2 - |z_\alpha| \, \sqrt{\boldsymbol{\pi}'\boldsymbol{F}(t, \boldsymbol{b})\boldsymbol{\pi}}.$$

If we define

$$(3.4) \qquad\qquad \boldsymbol{g}(t) := (\boldsymbol{a} - r\boldsymbol{1})t + \boldsymbol{\mathcal{A}}(t, \boldsymbol{b}) \quad \text{and}$$

$$f(\boldsymbol{\pi}, t) := \boldsymbol{\pi}'\boldsymbol{g}(t) - \frac{t}{2} \left\|\boldsymbol{\pi}'\boldsymbol{\sigma}\right\|^2 - |z_\alpha| \, \sqrt{\boldsymbol{\pi}'\mathbf{F}(t, \boldsymbol{b})\boldsymbol{\pi}},$$

then the α-quantile of the log-wealth process $H(t)$ can be written as

$$q_h = H(0) + rt + f(\boldsymbol{\pi}, t),$$

so that the α-quantile of the wealth process $X^\pi(t)$ is given by

$$q_x = X_0 \mathrm{e}^{rt} \mathrm{e}^{f(\boldsymbol{\pi}, t)}.$$

Since CaR is the difference between the riskless return and the α-quantile of the risky portfolio, we have the following corollary.

Corollary 3.2. *The capital at risk of the wealth process $X^\pi(t)$ is given by the formula*

$$\mathrm{CaR}(\boldsymbol{\pi}, t) = X_0 \mathrm{e}^{rt} \left(1 - \mathrm{e}^{f(\boldsymbol{\pi}, t)} \right).$$

Remark 3.1. When $c_i = 0$, i.e., $b_i = 0$, we get that $\mathcal{A}_i = 0$, $i = 1, \ldots, m$, so that $\boldsymbol{g}(t) = (\boldsymbol{a} - r\boldsymbol{1})t$, which is actually the risk premium vector. Also, in this case we get that $\boldsymbol{F}(t, \boldsymbol{b}) = \boldsymbol{\sigma}\boldsymbol{\sigma}'t$, so that the expression for CaR becomes the expression from [9], or [5], in the Black–Scholes setting with constant coefficients.

The expected wealth formula is given in the following proposition.

Proposition 3.3. *The expected value of the wealth process $X^\pi(t)$ is*

$$\mathbb{E}[X^\pi(t)] = X_0 \exp\left(rt + \boldsymbol{\pi}'\boldsymbol{g}(t) - \frac{1}{2}\boldsymbol{\pi}'\widetilde{\boldsymbol{F}}(t, \boldsymbol{b})\boldsymbol{\pi} \right),$$

where $\widetilde{\boldsymbol{F}}(t, \boldsymbol{b}) := \boldsymbol{F}(t, \boldsymbol{0}) - \boldsymbol{F}(t, \boldsymbol{b})$.

Proof. Since $H(t) = \ln X^{\boldsymbol{\pi}}(t)$ is normally distributed, we have

$$\mathbb{E}[X^{\boldsymbol{\pi}}(t)] = \mathbb{E}\left[\mathrm{e}^{H(t)} \right] = \exp\left(\mathbb{E}[H(t)] + \frac{1}{2}\mathbb{V}[H(t)] \right).$$

Substituting for the mean and variance of $H(t)$ from Proposition 3.1 gives

$$\mathbb{E}[X^{\boldsymbol{\pi}}(t)] = X_0 \exp\left(rt + \boldsymbol{\pi}'\boldsymbol{g}(t) - \frac{t}{2}\left\| \boldsymbol{\pi}'\boldsymbol{\sigma} \right\|^2 + \frac{1}{2}\boldsymbol{\pi}'\boldsymbol{F}(t, \boldsymbol{b})\boldsymbol{\pi} \right).$$

Noting that $t\left\| \boldsymbol{\pi}'\boldsymbol{\sigma} \right\|^2 = \boldsymbol{\pi}'\boldsymbol{F}(t, \boldsymbol{0})\boldsymbol{\pi}$ yields the result. ■

Remark 3.2. Since the matrix $\widetilde{\boldsymbol{F}}(t, \boldsymbol{b})$ is the difference between two positive-definite matrices, it will not in general be positive-definite. The implications of this will be illustrated in numerical examples in subsection 5.2.

In the following section we will look at the problems of a global minimum of the portfolio CaR and the maximal expected wealth subject to a constrained CaR.

4. Minimal capital at risk. In this section we first prove that CaR is a strongly quasi-convex function of $\boldsymbol{\pi}$, which guarantees the uniqueness of the global minimum. We then find the global minimum of CaR and investigate the problem of the maximal expected wealth subject to a constrained CaR.

4.1. Quasi-convexity of CaR. We recall that a function $\psi : \mathbb{R}^m \to \mathbb{R}$ is strongly quasi-convex if

$$\psi(\lambda\boldsymbol{\pi} + (1 - \lambda)\boldsymbol{\xi}) < \max\{\psi(\boldsymbol{\pi}), \psi(\boldsymbol{\xi})\}$$

for all $\boldsymbol{\pi}, \boldsymbol{\xi} \in \mathbb{R}^m$ for which $\boldsymbol{\pi} \neq \boldsymbol{\xi}$, and for all $\lambda \in (0, 1)$. We now prove that CaR has this important property.

Theorem 4.1. *The CaR is a strongly quasi-convex function of the portfolio.*

Proof. Suppose that $X^{\pi}(t)$ and $X^{\xi}(t)$ are two wealth processes defined by portfolios $\boldsymbol{\pi}, \boldsymbol{\xi} \in \mathbb{R}^m$, $\boldsymbol{\pi} \neq \boldsymbol{\xi}$, with the same initial wealth $X^{\pi}(0) = X^{\xi}(0) = X_0$, and suppose that

$$(4.1) \qquad \text{CaR}(\boldsymbol{\pi}, T) \geq \text{CaR}(\boldsymbol{\xi}, T).$$

This implies that

$$(4.2) \qquad f(\boldsymbol{\pi}, T) \leq f(\boldsymbol{\xi}, T).$$

If we define the matrix $\boldsymbol{A}(t, \boldsymbol{b})$ to be the unique positive-definite square root of $\boldsymbol{F}(t, \boldsymbol{b})$, i.e.,

$$(4.3) \qquad \boldsymbol{A}(t, \boldsymbol{b}) = \boldsymbol{F}(t, \boldsymbol{b})^{\frac{1}{2}},$$

we can write (4.2) as

$$\boldsymbol{\pi}' \boldsymbol{g}(T) - \frac{T}{2} \left\| \boldsymbol{\sigma}' \boldsymbol{\pi} \right\|^2 - |z_\alpha| \left\| \boldsymbol{A} \boldsymbol{\pi} \right\| \leq \boldsymbol{\xi}' \boldsymbol{g}(T) - \frac{T}{2} \left\| \boldsymbol{\sigma}' \boldsymbol{\xi} \right\|^2 - |z_\alpha| \left\| \boldsymbol{A} \boldsymbol{\xi} \right\|,$$

which is equivalent to

$$(4.4) \qquad (\boldsymbol{\xi} - \boldsymbol{\pi})' \boldsymbol{g}(T) \geq \frac{T}{2} \left(\left\| \boldsymbol{\sigma}' \boldsymbol{\xi} \right\|^2 - \left\| \boldsymbol{\sigma}' \boldsymbol{\pi} \right\|^2 \right) + |z_\alpha| \left(\left\| \boldsymbol{A} \boldsymbol{\xi} \right\| - \left\| \boldsymbol{A} \boldsymbol{\pi} \right\| \right).$$

We now have to prove that, for $\lambda \in (0, 1)$,

$$(4.5) \qquad \text{CaR}(\lambda \boldsymbol{\pi} + (1 - \lambda) \boldsymbol{\xi}, T) < \text{CaR}(\boldsymbol{\pi}, T)$$

under condition (4.1). Inequality (4.5) is equivalent to

$$(4.6) \qquad f(\lambda \boldsymbol{\pi} + (1 - \lambda) \boldsymbol{\xi}, T) > f(\boldsymbol{\pi}, T),$$

i.e.,

$$(\lambda \boldsymbol{\pi} + (1 - \lambda) \boldsymbol{\xi})' \boldsymbol{g}(T) - \frac{T}{2} \left\| \boldsymbol{\sigma}'(\lambda \boldsymbol{\pi} + (1 - \lambda) \boldsymbol{\xi}) \right\|^2 - |z_\alpha| \left\| \boldsymbol{A}(\lambda \boldsymbol{\pi} + (1 - \lambda) \boldsymbol{\xi}) \right\|$$
$$> \boldsymbol{\pi}' \boldsymbol{g}(T) - \frac{T}{2} \left\| \boldsymbol{\sigma}' \boldsymbol{\pi} \right\|^2 - |z_\alpha| \left\| \boldsymbol{A} \boldsymbol{\pi} \right\|.$$

Note that, from Proposition 2.2, it follows that the matrix $\boldsymbol{A}(t, \boldsymbol{b})$ is positive-definite, and, since $\boldsymbol{\sigma}$ is invertible, we have that the function $f(\boldsymbol{\pi}, T)$ is strictly concave. This implies that

$$f(\lambda \pi + (1 - \lambda) \xi, T) > \lambda f(\pi, T) + (1 - \lambda) f(\xi, T) \quad \forall \lambda \in (0, 1),$$

or, under condition (4.2),

$$f(\lambda \pi + (1 - \lambda) \xi, T) > f(\pi, T).$$

Since this holds for all $\lambda \in (0, 1)$, (4.6) is true, and the theorem is proved. ∎

This theorem has an immediate, important consequence. Namely, from Theorem 3.5.9 in [2], if a function $\psi : U \subset \mathbb{R}^m \to \mathbb{R}$ is strongly quasi-convex, then its local minimum is its unique global minimum. Therefore, the following corollary is true.

Corollary 4.2. *If* $\text{CaR}(\boldsymbol{\pi}, T)$ *has a local minimum at* $\boldsymbol{\pi}^* \in \mathcal{Q}$, *then* $\boldsymbol{\pi}^*$ *is its global minimum.*

4.2. Global minimum CaR. We will now find the global minimum of $\mathrm{CaR}(\boldsymbol{\pi}, T)$ at the time horizon T. Note that

$$
\text{(4.7)} \qquad \operatorname*{argmin}_{\boldsymbol{\pi}} \mathrm{CaR}(\boldsymbol{\pi}, T) = \operatorname*{argmax}_{\boldsymbol{\pi}} f(\boldsymbol{\pi}, T)
$$

$$
= \operatorname*{argmax}_{\boldsymbol{\pi}} \boldsymbol{\pi}' \boldsymbol{g}(T) - \frac{T}{2} \left\| \boldsymbol{\pi}' \boldsymbol{\sigma} \right\|^2 - |z_\alpha| \sqrt{\boldsymbol{\pi}' \boldsymbol{F} \boldsymbol{\pi}}.
$$

To solve the above problem we prove the following theorem.

Theorem 4.3. (1) *If* $\boldsymbol{g}(T)' \boldsymbol{F}^{-1} \boldsymbol{g}(T) > |z_\alpha|^2$, *the optimal solution of problem* (4.7) *is equal to*

$$
\boldsymbol{\pi}^* = \left(T \boldsymbol{\sigma} \boldsymbol{\sigma}' + \frac{|z_\alpha|}{\lambda^*} \boldsymbol{F} \right)^{-1} \boldsymbol{g}(T),
$$

where λ^* *is the unique positive solution of the equation*

$$
\left\| \boldsymbol{A} \left(\lambda T \boldsymbol{\sigma} \boldsymbol{\sigma}' + |z_\alpha| \, \boldsymbol{F} \right)^{-1} \boldsymbol{g}(T) \right\| = 1
$$

and $\boldsymbol{A}$ *is defined in* (4.3).

(2) *If* $\boldsymbol{g}(T)' \boldsymbol{F}^{-1} \boldsymbol{g}(T) \le |z_\alpha|^2$, *the optimal solution of problem* (4.7) *is* $\boldsymbol{\pi} = \boldsymbol{0}$.

Proof. The critical points of $f(\boldsymbol{\pi}, T)$ are $\boldsymbol{\pi} = \boldsymbol{0}$ and the point at which the gradient is zero, i.e., the solution of the equation

$$
\text{(4.8)} \qquad \boldsymbol{g}(T) - T \boldsymbol{\sigma} \boldsymbol{\sigma}' \boldsymbol{\pi} - \frac{|z_\alpha|}{\sqrt{\boldsymbol{\pi}' \boldsymbol{F} \boldsymbol{\pi}}} \boldsymbol{F} \boldsymbol{\pi} = \boldsymbol{0}.
$$

The solution of (4.8) must satisfy

$$
\boldsymbol{\pi} = \left(T \boldsymbol{\sigma} \boldsymbol{\sigma}' + \frac{|z_\alpha|}{\sqrt{\boldsymbol{\pi}' \boldsymbol{F} \boldsymbol{\pi}}} \boldsymbol{F} \right)^{-1} \boldsymbol{g}(T).
$$

If we define $\lambda := \sqrt{\boldsymbol{\pi}' \boldsymbol{F} \boldsymbol{\pi}}$, we get

$$
\boldsymbol{\pi}' \boldsymbol{F} \boldsymbol{\pi} = \lambda^2 = \boldsymbol{g}(T)' \left(T \boldsymbol{\sigma} \boldsymbol{\sigma}' + \frac{|z_\alpha|}{\lambda} \boldsymbol{F} \right)^{-1} \boldsymbol{F} \left(T \boldsymbol{\sigma} \boldsymbol{\sigma}' + \frac{|z_\alpha|}{\lambda} \boldsymbol{F} \right)^{-1} \boldsymbol{g}(T)
$$

$$
= \lambda^2 \boldsymbol{g}(T)' \left(\lambda T \boldsymbol{\sigma} \boldsymbol{\sigma}' + |z_\alpha| \, \boldsymbol{F} \right)^{-1} \boldsymbol{F} \left(\lambda T \boldsymbol{\sigma} \boldsymbol{\sigma}' + |z_\alpha| \, \boldsymbol{F} \right)^{-1} \boldsymbol{g}(T).
$$

This implies that

$$
\text{(4.9)} \qquad \boldsymbol{g}(T)' \left(\lambda T \boldsymbol{\sigma} \boldsymbol{\sigma}' + |z_\alpha| \, \boldsymbol{F} \right)^{-1} \boldsymbol{F} \left(\lambda T \boldsymbol{\sigma} \boldsymbol{\sigma}' + |z_\alpha| \, \boldsymbol{F} \right)^{-1} \boldsymbol{g}(T) = 1.
$$

Using (4.3) we can write (4.9) in the following way:

$$
\left\| \boldsymbol{A} \left(\lambda T \boldsymbol{\sigma} \boldsymbol{\sigma}' + |z_\alpha| \, \boldsymbol{F} \right)^{-1} \boldsymbol{g}(T) \right\| = 1.
$$

We now define the function

$$
\text{(4.10)} \qquad \omega(\lambda) = \left\| \boldsymbol{A} \left(\lambda T \boldsymbol{\sigma} \boldsymbol{\sigma}' + |z_\alpha| \, \boldsymbol{F} \right)^{-1} \boldsymbol{g}(T) \right\|^2
$$

and prove that $\omega(\lambda)$ is strictly decreasing in λ. Note that, since $\boldsymbol{F} = \boldsymbol{A}\boldsymbol{A}'$,

$$\omega(\lambda) = \left\| \left(\lambda T \left(\boldsymbol{A}' \right)^{-1} \boldsymbol{\sigma}\boldsymbol{\sigma}' \boldsymbol{A}^{-1} + |z_\alpha| \boldsymbol{I} \right)^{-1} \left(\boldsymbol{A}' \right)^{-1} \boldsymbol{g}(T) \right\|^2,$$

where $\boldsymbol{I}$ denotes the identity matrix. Since $\left(\boldsymbol{A}' \right)^{-1} \boldsymbol{\sigma}\boldsymbol{\sigma}' \boldsymbol{A}^{-1}$ is symmetric positive-definite, it has an eigenvalue decomposition

$$\left(\boldsymbol{A}' \right)^{-1} \boldsymbol{\sigma}\boldsymbol{\sigma}' \boldsymbol{A}^{-1} = \boldsymbol{P}' \boldsymbol{D} \boldsymbol{P},$$

where the diagonal elements of $\boldsymbol{D}$ are positive and $\boldsymbol{P}$ is an orthonormal matrix. Writing $\boldsymbol{u} = \left(\boldsymbol{A}' \right)^{-1} \boldsymbol{g}(T)$, we find then that

$$\begin{aligned}
\omega(\lambda) &= \left\| \left(\lambda T \boldsymbol{P}' \boldsymbol{D} \boldsymbol{P} + |z_\alpha| \boldsymbol{I} \right)^{-1} \boldsymbol{u} \right\|^2 \\
&= \left\| \boldsymbol{P}' \left(\lambda T \boldsymbol{D} + |z_\alpha| \boldsymbol{I} \right)^{-1} \boldsymbol{P} \boldsymbol{u} \right\|^2 \\
&= \boldsymbol{u}' \boldsymbol{P}' \left(\lambda T \boldsymbol{D} + |z_\alpha| \boldsymbol{I} \right)^{-1} \boldsymbol{P} \boldsymbol{P}' \left(\lambda T \boldsymbol{D} + |z_\alpha| \boldsymbol{I} \right)^{-1} \boldsymbol{P} \boldsymbol{u} \\
&= \boldsymbol{u}' \boldsymbol{P}' \left(\lambda T \boldsymbol{D} + |z_\alpha| \boldsymbol{I} \right)^{-1} \left(\lambda T \boldsymbol{D} + |z_\alpha| \boldsymbol{I} \right)^{-1} \boldsymbol{P} \boldsymbol{u} \\
&= \boldsymbol{u}' \boldsymbol{P}' \left(\lambda T \boldsymbol{D} + |z_\alpha| \boldsymbol{I} \right)^{-2} \boldsymbol{P} \boldsymbol{u}.
\end{aligned}$$

For the sake of simplicity we define the vector $\boldsymbol{v} = \boldsymbol{P}\boldsymbol{u}$ and the matrix $\boldsymbol{U} = \left(\lambda T \boldsymbol{D} + |z_\alpha| \boldsymbol{I} \right)^{-1}$. Then the above equation can be written as

$$\omega(\lambda) = \boldsymbol{v}' \boldsymbol{U}^2 \boldsymbol{v}.$$

$\boldsymbol{U}$ is a diagonal matrix with entries

$$U_i(\lambda) = \frac{1}{\lambda T d_i + |z_\alpha|},$$

where d_i are the diagonal elements of $\boldsymbol{D}$ and are positive, so that $U_i(\lambda)$ is a decreasing function of λ for $\lambda > 0$ and for all $i = 1, \ldots, m$. Therefore,

$$\omega(\lambda) = \sum_{i=1}^{m} \frac{v_i^2}{\left(\lambda T d_i + |z_\alpha| \right)^2},$$

so that $\omega(\lambda)$ is a decreasing function of λ for $\lambda > 0$. From (4.10) we have that

$$\omega(0) = \frac{\boldsymbol{g}(T)' \boldsymbol{F}^{-1} \boldsymbol{g}(T)}{|z_\alpha|^2}.$$

We now have two cases:

(1) Suppose that $\boldsymbol{g}(T)' \boldsymbol{F}^{-1} \boldsymbol{g}(T) > |z_\alpha|^2$. In this case the equation $\omega(\lambda) = 1$ has a unique positive solution λ^*, and the optimum $\boldsymbol{\pi}$ is given by

$$\boldsymbol{\pi}^* = \left(T \boldsymbol{\sigma}\boldsymbol{\sigma}' + \frac{|z_\alpha|}{\lambda^*} \boldsymbol{F} \right)^{-1} \boldsymbol{g}(T).$$

(2) If $\boldsymbol{g}(T)' \boldsymbol{F}^{-1} \boldsymbol{g}(T) \leq |z_\alpha|^2$, the equation $\omega(\lambda) = 1$ has no positive solution, so that the optimum portfolio is $\boldsymbol{\pi}^* = \boldsymbol{0}$, another critical point of function $f(\boldsymbol{\pi}, T)$. ∎

Remark 4.1.

(i) When $c_i = 0$, i.e., $b_i = 0$, we find that $\boldsymbol{g}(T)'\boldsymbol{F}^{-1}\boldsymbol{g}(T) = \boldsymbol{g}(T)'(\boldsymbol{\sigma}\boldsymbol{\sigma}')^{-1}\boldsymbol{g}(T) = \left\|\boldsymbol{g}(T)'\boldsymbol{\sigma}^{-1}\right\|^2$, where $\boldsymbol{g}(T)$ is the risk premium vector and $\boldsymbol{\theta}(T) := \boldsymbol{\sigma}^{-1}\boldsymbol{g}(T)$ is the market price of risk. Therefore, the condition $\boldsymbol{g}(T)'\boldsymbol{F}^{-1}\boldsymbol{g}(T) > |z_\alpha|^2$ reduces to the condition $\left\|\boldsymbol{g}(T)'\boldsymbol{\sigma}^{-1}\right\|^2 > |z_\alpha|^2$, or $\left\|\boldsymbol{\theta}(T)\right\|^2 > |z_\alpha|^2$. We note that the last condition is the criterion under which we invest in the risky assets under minimal CaR in the Black–Scholes setting (see [9], [14]).

(ii) We note that the optimal strategy given in Theorem 4.3 depends on $S_i(0)$ through the vector $\boldsymbol{g}(T)$, which in turn involves $\boldsymbol{\mathcal{A}}$, defined in (3.2). In our previous studies (see [9], [10], [12]) or in [14], where the optimization was done in the Black–Scholes setting with respect to CaR or conditional CaR, this was not the case. We believe that the time inconsistency in the current setting is due to a combination of the presence of mean-reversion and the restriction to constant portfolios.

(iii) When $m = 1$, i.e., the portfolio consists of one risky and one riskless asset, we can find the exact analytic solution to problem (4.7), which will be shown in the following lemma.

Lemma 4.4. In the case when $m = 1$, the optimal solution of problem (4.7) is equal to

$$(4.11) \qquad \pi^* = \begin{cases} \frac{g(T) - \text{sign}(g(T))\|z_\alpha\|\sigma\sqrt{\mathbb{E}(T,2b)}}{\sigma^2 T} & \text{if } \|g(T)\| > \|z_\alpha\|\sigma\sqrt{\mathbb{E}(T,2b)}, \\ 0 & \text{if } \|g(T)\| \le |z_\alpha|\sigma\sqrt{\mathbb{E}(T,2b)}. \end{cases}$$

Proof. We note that in the one-dimensional case problem (4.7) reduces to the problem

$$(4.12) \qquad \underset{\pi}{\text{argmax}}\, f(\pi,T) = \pi g(T) - \frac{T}{2}(\pi\sigma)^2 - |z_\alpha|\,\|\pi\|\,\sigma\sqrt{\mathbb{E}(T,2b)}, \quad \text{with}$$

$$g(T) = \left(\frac{1}{2}\sigma^2 - r\right)T + \mathbb{E}(T,b)\left(a - \frac{1}{2}\sigma^2 - bY_0\right).$$

Clearly, as in the m-dimensional case, the critical points of $f(\pi,T)$ are $\pi = 0$ and the point at which $f'(\pi,T) = 0$. From the expression

$$f'(\pi,T) = \begin{cases} g(T) - T\pi\sigma^2 - |z_\alpha|\sigma\sqrt{\mathbb{E}(T,2b)} & \text{if } \pi > 0, \\ g(T) - T\pi\sigma^2 + |z_\alpha|\sigma\sqrt{\mathbb{E}(T,2b)} & \text{if } \pi < 0, \end{cases}$$

we get the following cases.

(a) If $g(T) > |z_\alpha|\sigma\sqrt{\mathbb{E}(T,2b)}$, or if $g(T) < -|z_\alpha|\sigma\sqrt{\mathbb{E}(T,2b)}$, the optimal solution to problem (4.12) is

$$\pi^* = \frac{g(T) - |z_\alpha|\sigma\sqrt{\mathbb{E}(T,2b)}}{\sigma^2 T} > 0 \quad \text{or} \quad \pi^* = \frac{g(T) + |z_\alpha|\sigma\sqrt{\mathbb{E}(T,2b)}}{\sigma^2 T} < 0.$$

(b) If $|z_\alpha|\sigma\sqrt{\mathbb{E}(T,2b)} \le g(T) \le |z_\alpha|\sigma\sqrt{\mathbb{E}(T,2b)}$, the optimal solution to problem (4.12) is $\pi^* = 0$, so that the proof of the lemma is complete. ∎

Remark 4.2.

(i) The condition $\frac{g(T)^2}{\sigma^2 \mathbb{E}(T,2b)} > |z_\alpha|^2$ corresponds to the condition $\boldsymbol{g}(T)'\boldsymbol{F}^{-1}\boldsymbol{g}(T) > |z_\alpha|^2$ from Theorem 4.3.

(ii) When the risk premium $g(T)$ decreases, the optimal strategy changes from investing in the risky asset ($\pi^* > 0$), to investing everything in the riskless asset, and finally to short-selling the risky asset.

In section 5 we will illustrate the results of the above theorem.

4.3. CaR-constrained wealth maximization. We now turn our attention to solving the following optimization problem:

$$(4.13) \qquad \max_{\boldsymbol{\pi}\in\mathbb{R}^m} E\left[X^{\boldsymbol{\pi}}(T)\right] \quad \text{subject to} \quad \text{CaR}(\boldsymbol{\pi},T) \leq C.$$

Problem (4.13) can be rewritten in the following form:

$$(4.14) \qquad \max_{\boldsymbol{\pi}\in\mathbb{R}^m} X_0 e^{rT} e^{\boldsymbol{\pi}'\boldsymbol{g}(T)-\frac{1}{2}\boldsymbol{\pi}'\widetilde{\boldsymbol{F}}(T,\boldsymbol{b})\boldsymbol{\pi}} \quad \text{subject to} \quad X_0 e^{rT}\left(1 - e^{f(\boldsymbol{\pi},T)}\right) \leq C,$$

where $f(\boldsymbol{\pi},T)$ and $\widetilde{\boldsymbol{F}}(T,\boldsymbol{b})$ are defined in (3.4) and Proposition 3.3. In the following proposition we will show that the above problem, under appropriate conditions for C, has an optimal solution.

Proposition 4.5.

(i) *If the constant C satisfies the conditions*

$$\begin{aligned} 0 \leq C < X_0 R\left(T\right) & \quad \textit{if } \boldsymbol{g}(T)'\boldsymbol{F}^{-1}\boldsymbol{g}(T) \leq |z_\alpha|^2\,, \\ (4.15) \qquad \text{CaR}(\boldsymbol{\pi}^*,T) \leq C < X_0 R(T) & \quad \textit{if } \boldsymbol{g}(T)'\boldsymbol{F}^{-1}\boldsymbol{g}(T) > |z_\alpha|^2\,, \end{aligned}$$

where $\boldsymbol{\pi}^*$ is given in (4.11), then problem (4.14) has an optimal solution.

(ii) *If the matrix $\widetilde{\boldsymbol{F}}(T,\boldsymbol{b})$ is strictly positive-definite, the solution is unique.*

Proof.

(i) Problem (4.14) is equivalent to the problem

$$(4.16) \qquad \max_{\boldsymbol{\pi}\in\mathbb{R}^m} \boldsymbol{\pi}'\boldsymbol{g}(T) - \frac{1}{2}\boldsymbol{\pi}'\widetilde{\boldsymbol{F}}(T,\boldsymbol{b})\boldsymbol{\pi} \quad \text{subject to} \quad e^{f(\boldsymbol{\pi},T)} \geq 1 - \frac{C}{X_0 e^{rT}}.$$

Note that, under conditions (4.15), the constraint in (4.13) is well defined. Problem (4.16) is equivalent to

$$\max_{\boldsymbol{\pi}\in\mathbb{R}^m}\{-h(\boldsymbol{\pi},T)\}$$
$$\text{subject to} \quad f(\boldsymbol{\pi},T) \geq c,$$

where

$$(4.17) \qquad c = \ln\left(1 - \frac{C}{X_0 e^{rT}}\right) \quad \text{and} \quad h(\boldsymbol{\pi},T) = \frac{1}{2}\boldsymbol{\pi}'\widetilde{\boldsymbol{F}}(T,\boldsymbol{b})\boldsymbol{\pi} - \boldsymbol{\pi}'\boldsymbol{g}(T),$$

or to the problem

$$(4.18) \qquad \begin{aligned} & \min_{\boldsymbol{\pi}\in\mathbb{R}^m} h(\boldsymbol{\pi},T) \\ & \text{subject to} \quad -f(\boldsymbol{\pi},T) + c \leq 0. \end{aligned}$$

Problem (4.18) consists of minimizing a continuous function over the set

$$Q = \{\boldsymbol{\pi} \in \mathbb{R}^m \mid \mathrm{CaR}(\boldsymbol{\pi}, T) \leq C\}.$$

We first note that the set Q under conditions (4.15) is nonempty and compact. Since $h(\boldsymbol{\pi}, T)$ is a continuous function over a compact set, it achieves both its absolute minimum and maximum over the set Q, so that the existence of an optimal solution is proved.

(ii) To prove the uniqueness of the optimal solution under the condition that $h(\boldsymbol{\pi}, T)$ is positive-definite, we first note that the set Q is convex, since $\mathrm{CaR}(\boldsymbol{\pi}, T)$ is a strongly quasi-convex function of $h(\boldsymbol{\pi}, T)$ (see Theorem 3.5.2 in [2]). It is easy to show (see Exercise 4.15 in [2]) that the absolute minimum of problem (4.18) is achieved either at the unique critical point of the function $h(\boldsymbol{\pi}, T)$, i.e., the point $\boldsymbol{\pi}^{**} = \widetilde{\boldsymbol{F}}(T, \boldsymbol{b})^{-1} \boldsymbol{g}(T)$, or at the boundary of the set Q which we denote by ∂Q. We now consider the following two cases.

(1) If $\boldsymbol{\pi}^{**}$, i.e., the global minimum of $h(\boldsymbol{\pi}, T)$, or the global maximum of $E\left[X^{\boldsymbol{\pi}}(T)\right]$ belongs to Q, then it is also its unique constrained minimum, and the proof is complete.

(2) If $\boldsymbol{\pi}^{**} \notin Q$, then the constrained minimum of $h(\boldsymbol{\pi}, T)$ belongs to ∂Q. Suppose that there exist two solutions $\boldsymbol{\pi}^1$ and $\boldsymbol{\pi}^2$ that satisfy $h(\boldsymbol{\pi}^1, T) = h(\boldsymbol{\pi}^2, T) = \min_{\boldsymbol{\pi} \in \mathbb{Q}} h(\boldsymbol{\pi}, T)$ and $\boldsymbol{\pi}^1, \boldsymbol{\pi}^2 \in \partial Q$. Since Q is convex, we have that $\frac{\boldsymbol{\pi}^1 + \boldsymbol{\pi}^2}{2} \in Q$ and, from the strict convexity of $h(\boldsymbol{\pi}, T)$, that

$$h\left(\frac{\boldsymbol{\pi}^1 + \boldsymbol{\pi}^2}{2}, T\right) < \frac{1}{2}\left(h(\boldsymbol{\pi}^1, T) + h(\boldsymbol{\pi}^2, T)\right) = \min_{\boldsymbol{\pi} \in \mathbb{Q}} hf(\boldsymbol{\pi}, T).$$

This is obviously a contradiction, so the proof of Proposition 4.5 is thus complete. ■

However, due to the complexity of the constraint in problem (4.18), we are only able to find an optimal solution numerically, which will be illustrated in the next section.

5. Numerical results. In this section we illustrate the results of Theorem 4.3 and Proposition 4.5 in a series of numerical experiments. The parameters used in these experiments were chosen for illustrative purposes only. We have not addressed in this paper the issue of calibration of model parameters from market data. Obtaining precise values for parameters such as β or L for the various assets is an important problem, and in practice one must live with a level of uncertainty in the parameter values. This is an issue we hope to address in future work.

5.1. Minimal CaR: One-dimensional case. We begin by considering the case where the portfolio consists of one risky asset in addition to the riskless asset, with their dynamics given in (2.1) and (2.2) for $m = 1$.

We first give a series of examples illustrating how the mean-reversion parameter β affects the optimal portfolio in the case of one risky asset. The results are shown in Figure 1. The basic setting used here has $T = 2$, $S(0) = 5$, $X_0 = 100$, $r = 0.02$, $\alpha = 0.05$, and $L = \log 7$. Thus the risky asset is undervalued at the outset, in the sense that it is reverting to a long-run mean that is higher than its current value. We note that the above data are simulated using the Euler scheme with daily granularity.

We see how risk, as measured by CaR, depends heavily on the speed of reversion, i.e., β. When $\beta = 1$, the minimal CaR is obtained for $\pi = 0$; as β increases beyond 2, the optimal CaR
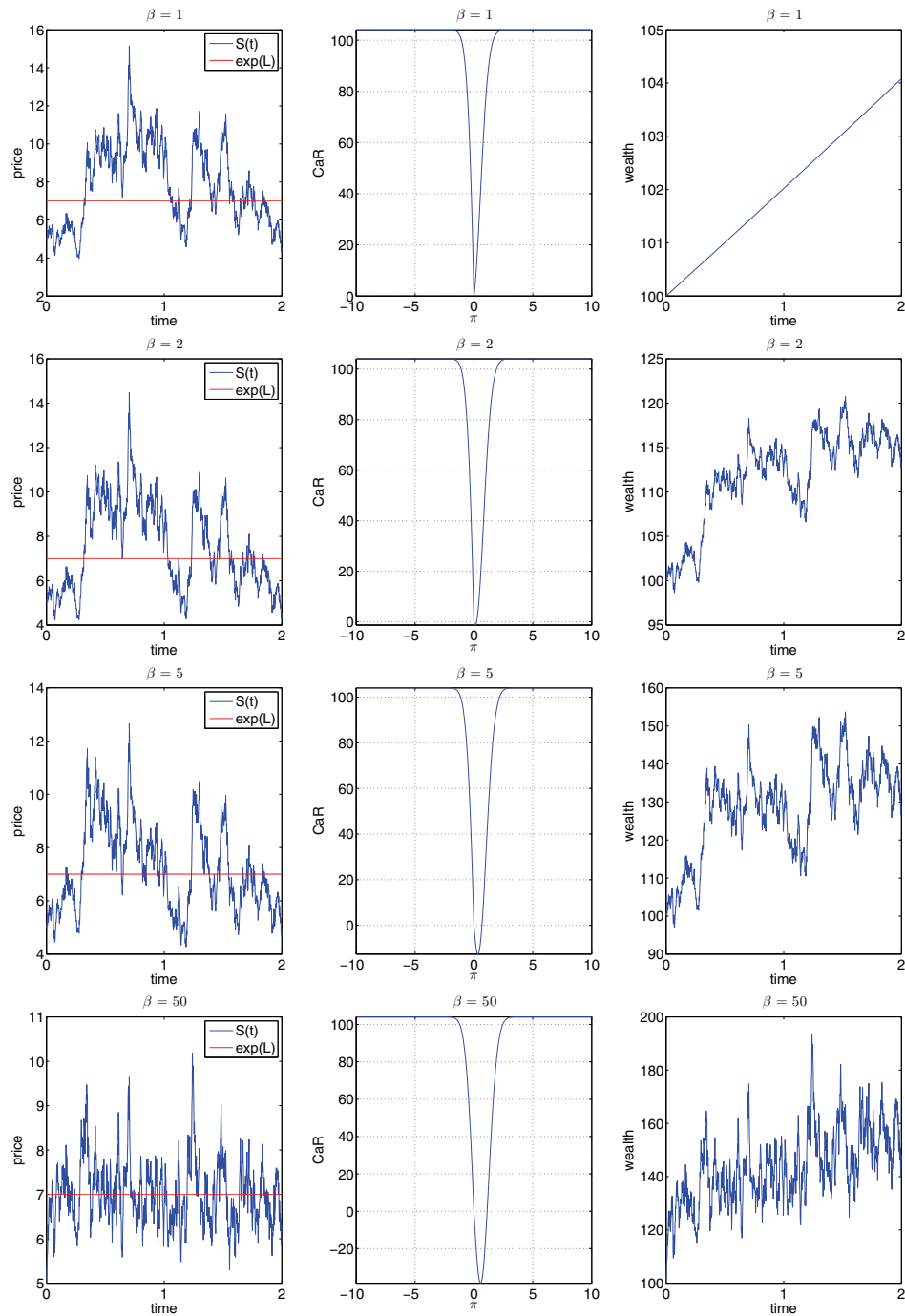
Figure 1. *Typical asset paths (left), 2-year CaR (middle), and wealth process for the minimal-CaR strategy (right) for the one-dimensional asset model with parameters $L = \log 7$, $c = 1$, $S(0) = 5$, $\sigma = 1.0$, $T = 2$, $X_0 = 100$, and for $\beta = 1, 2, 5$, and 50.*
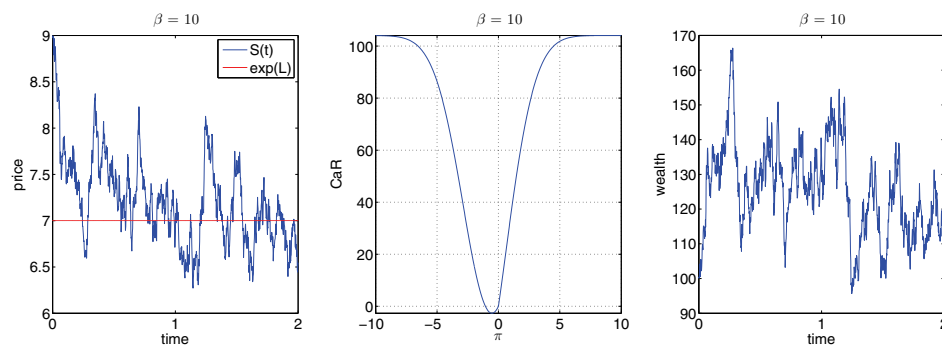
Figure 2. *Typical asset path (left), 2-year CaR (middle), and wealth process for the minimal-CaR strategy (right) for the one-dimensional asset model with parameters: $L = \log 7$, $c = 1$, $S(0) = 9$, $\sigma = 0.3$, $T = 2$, $X_0 = 100$, and $\beta = 10$.*

is obtained for progressively higher values of π (tending to a limit somewhere between 0 and 2 as $\beta \to \infty$). We also see in Figure 1, in the third column, the wealth process corresponding to the minimal-CaR strategy. Note the increased rate of growth in the wealth as β increases.

This behavior is dependent on the relation between $\exp(L)$ and $S(0)$. As noted in Remark 4.2(ii), negative values of the optimal π can arise when $g(T)$ is negative. This will occur, for instance, when $e^L < S(0)e^{(r-\sigma^2/2)T}$ and β is sufficiently high, as illustrated in Figure 2.

5.2. CaR-constrained optimal wealth: Two-dimensional case. In this section we illustrate what can happen when we seek to maximize expected wealth subject to constrained CaR when there are two mean-reverting assets in addition to the riskless asset. The process parameters are given by

L	0.5	2.0
β	4.0	0.1
$Y(0)$	1	1

and the covariance matrix is

$$\sigma\sigma' = \begin{bmatrix} 0.99 & -0.9 \\ -0.9 & 0.83 \end{bmatrix}.$$

We consider a time horizon of $T = 1$ and a riskless interest rate of $r = 0.02$. The resulting expected wealth surface is shown (on a logarithmic scale) as a function of (π_1, π_2) in Figure 3. Superimposed on the surface is the boundary of the portion of the surface attainable when the CaR is greater than 0.1: the corresponding region in (π_1, π_2)-space is shown at the base of the figure. What can be observed from Figure 3 is the fact that the logarithm of the expected wealth is neither convex nor concave (see Remark 3.2). The resulting possible nonuniqueness of the CaR-constrained optimal expected wealth can also be seen: in this setting it is quite possible that there might be two portfolios satisfying the constraint with identical expected wealth.

5.3. Minimal CaR: Three-dimensional case. We now look at the case where the portfolio consists of three risky assets and the riskless asset and find the optimal strategies that minimize CaR.
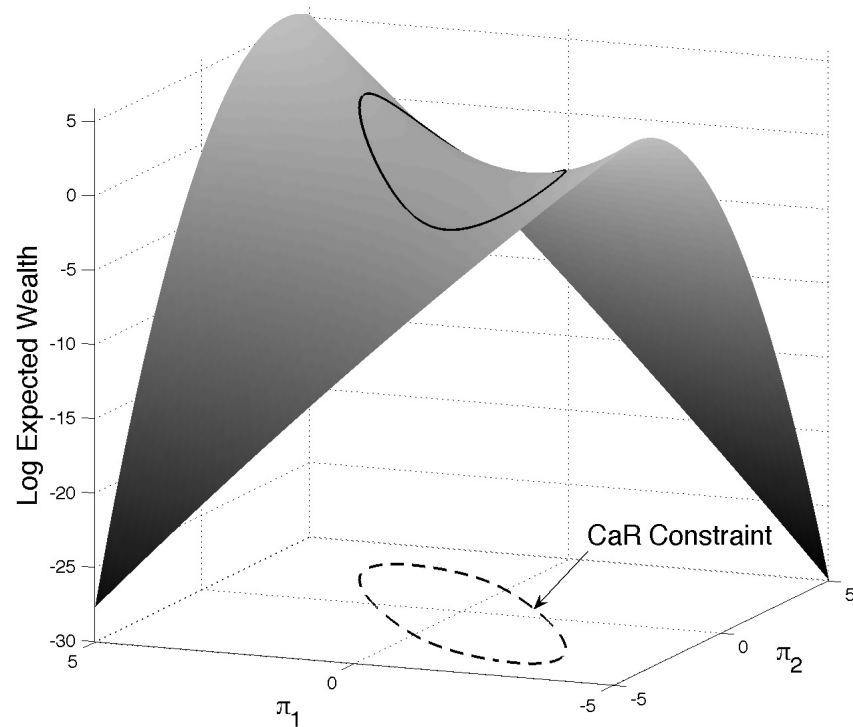
Figure 3. *The logarithm of the expected wealth as a function of (π_1, π_2), with the boundary of the region satisfying the constraint* $\mathrm{CaR} \geq 0.1$ *shown at the base of the figure and superimposed on the surface.*

We recall that the variance-covariance matrix associated with the stock returns, which we denote by $\mathbf{\Gamma}$, is equal to $\boldsymbol{\sigma\sigma}'$. We also recall that $\mathbf{\Gamma}$ can be decomposed as

$$(5.1) \qquad \mathbf{\Gamma} = \boldsymbol{\nu\rho\nu},$$

where $\boldsymbol{\rho}$ is the corresponding correlation matrix and $\boldsymbol{\nu}$ is a diagonal matrix with entries given by the individual stock return standard deviations. Therefore, from (5.1) we get

$$(5.2) \qquad \mathbf{\Gamma} = \boldsymbol{\sigma\sigma}' = \boldsymbol{\nu\rho\nu}.$$

Although, theoretically, we assume that the vector of independent Brownian motions $\boldsymbol{W}(t)$ is $\{\mathcal{F}_t\}_{t \in [0,T]}$ adapted, i.e., known at time $t \in [0, T]$, it is a common practice that we observe only $\mathbf{\Gamma}$ or, equivalently, $\boldsymbol{\rho}$ and $\boldsymbol{\nu}$, but not $\boldsymbol{\sigma}$. From (5.2) we see that this leads to a nonunique decomposition of $\mathbf{\Gamma}$ into the product $\boldsymbol{\sigma\sigma}'$. Despite this fact, the expressions for $\mathrm{CaR}(\boldsymbol{\pi}, t)$ and $\mathbb{E}[X^\pi(t)]$ are uniquely determined, since all the terms that involve $\boldsymbol{\sigma}$, i.e., $\|\boldsymbol{\pi}'\boldsymbol{\sigma}\|^2$ and $\boldsymbol{F}(t, \boldsymbol{b})$, can be expressed as

$$\|\boldsymbol{\pi}'\boldsymbol{\sigma}\|^2 = \boldsymbol{\pi}'\boldsymbol{\sigma\sigma}'\boldsymbol{\pi} = \boldsymbol{\pi}'\mathbf{\Gamma}\boldsymbol{\pi} \quad \text{and}$$
$$\boldsymbol{F}(t, \boldsymbol{b}) = \boldsymbol{\sigma\sigma}' \centerdot \mathbb{E}(t, \boldsymbol{b}) = \mathbf{\Gamma} \centerdot \mathbb{E}(t, \boldsymbol{b}),$$
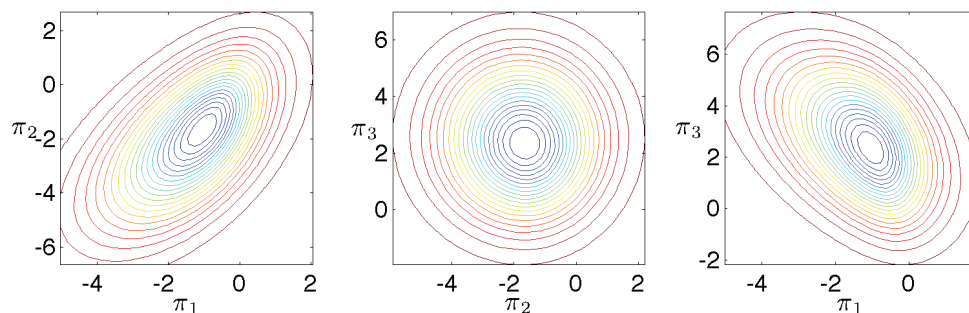
Figure 4. *CaR as a function of π_1, π_2, and π_3. In this setting CaR ranges from $\$-85.998$ to $\$104.081$. In each plot, the missing component of π is held constant at its optimal value.*

where the matrix $\mathbb{E}(t, \boldsymbol{b})$ has the entries equal to $\mathbb{E}(t, b_i + b_j)$ defined in (2.5) and the symbol "$\boldsymbol{\cdot}$" stands for the Hadamard product of two matrices.

We now give a brief illustration of the dependence of CaR on the portfolio $\boldsymbol{\pi}$ in the case of three risky assets.

We assume that the correlation matrix is

$$\boldsymbol{\rho} = \begin{bmatrix} 1.0 & -0.6 & 0.5 \\ -0.6 & 1.0 & 0 \\ 0.5 & 0 & 1 \end{bmatrix}.$$

The final time $T = 1$, $X_0 = 100$, the risk-free interest rate $r = 0.02$, and $\alpha = 0.05$; the other coefficients, and the resulting optimal strategy (π_1, π_2, π_3), are given by

$\boldsymbol{\pi}$	-1.02	-1.65	2.35
$\boldsymbol{L}$	0.05	0.50	1.50
$\boldsymbol{\beta}$	1	2	5
$\boldsymbol{Y(0)}$	1	1	1

.

Thus we see that $\pi_0 = 1.32$. The dependence of CaR on $\boldsymbol{\pi}$ is illustrated in Figure 4, where we see three cross-sectional contour plots showing CaR as a function of, in turn, (π_1, π_2) (keeping π_3 at its optimal value), (π_2, π_3) (keeping π_1 at its optimal value), and (π_1, π_3) (keeping π_2 at its optimal value). The effect of the negative correlation between S_1 and S_2, for example, is to distort the CaR function as shown in the first plot, and the zero correlation between S_2 and S_3 leads to a more symmetrical dependence of CaR on (π_2, π_3).

6. Conclusion. In this work we investigated portfolio selection in the market consisting of instruments that follow exponential Ornstein–Uhlenbeck distributions, instruments that follow log-normal distributions, or a combination of both. We considered the market setting with constant coefficients and investigated constant portfolios over a finite time horizon $[0, T]$. We note that keeping portfolios constant assumes continuous trading over the given time horizon.

Our approach to portfolio optimization was the risk reward approach, with risk measured by capital at risk and reward by the expected return of the wealth process. We derived analytic expressions for the wealth process, its expected value, and CaR.

After deriving all necessary expressions we solved the problem of finding the global minimum CaR and proved that it is unique since capital at risk is a strongly quasi-convex function of the portfolio. This solution is semianalytic in the sense that there is a scalar which has to be found numerically, while the solution to a one-dimensional case is purely analytic. We further developed optimal strategies that maximize the expected wealth under constrained CaR. Finding optimal strategies required solving nonlinear systems of equations so that the portfolio weights could be found only numerically.

Finally, we provided some numerical examples which illustrate that the presence and strength of mean-reversion in an asset model have a significant effect on optimal portfolio management.

There are several possible directions for future research. One important task will be to find optimal strategies in the case of time-dependent portfolios, or portfolios that depend on asset prices. Another interesting question is the asymptotic behavior of optimal portfolios over large time horizons. In this case, the dependence of the portfolio on the initial asset prices diminishes as the time horizon increases. Also, as mentioned in section 5, the calibration of model parameters is an important practical problem. Precise values are impossible to determine, and so the question arises of portfolio optimization in the context of uncertain model parameters.

One further possible future direction is to investigate the problem of portfolio optimization where the constraint is conditional CaR and compare the results with the results obtained in this paper. Finally, another interesting problem would be to extend the results in this paper to the problem of portfolio selection where the portfolio consists of more general instruments, such as instruments that are modeled using Lévy processes.

REFERENCES

[1] P. Balduzzi, S. Das, and S. Foresi, *The central tendency: A second factor in bond yields*, Rev. Econ. Statist., 80 (1998), pp. 62–72.

[2] M. S. Bazaraa and C. M. Shetty, *Nonlinear Programming*, John Wiley & Sons, New York, 1979.

[3] F. E. Benth and K. H. Karlsen, *A note on Merton's portfolio selection problem for the Schwartz mean-reversion model*, Stoch. Anal. Appl., 23 (2005), pp. 687–704.

[4] M. J. Brennan, E. S. Schwartz, and R. Lagnado, *Strategic asset allocation*, J. Econom. Dynam. Control, 21 (1997), pp. 1377–1403.

[5] P. Cheridito, D. Filipović, and R. L. Kimmel, *Market price of risk specifications for affine models: Theory and evidence*, J. Financial Econ., 83 (2007), pp. 123–170.

[6] L. Clewlow and C. Strickland, *Valuing Energy Options in a One Factor Model Fitted to Forward Prices*, Research Paper Series, Quantitative Finance Research Centre, University of Technology, Sydney, 1999.

[7] A. D'Aspremont, *Identifying small mean-reverting portfolios*, Quant. Finance, 11 (2011), pp. 351–364.

[8] W. F. M. De Bondt and R. H. Thaler, *Anomalies: A mean-reverting walk down Wall Street*, J. Econom. Perspectives, 3 (1989), pp. 189–202.

[9] G. Dmitrašinović-Vidović, A. Lari-Lavassani, X. Li, and T. Ware, *Dynamic portfolio selection under Capital-at-Risk under no short-selling constraints*, Int. J. Theor. Appl. Finance, 14 (2011), pp. 1–21.

[10] G. DMITRAŠINOVIĆ-VIDOVIĆ, A. LARI-LAVASSANI, X. LI, AND A. WARE, *Continuous time portfolio selection under conditional capital at risk*, J. Probab. Stat., 2010 (2010), 976371.

[11] G. DMITRAŠINOVIĆ-VIDOVIĆ, *Portfolio Selection Under Downside Risk Measures*, Ph.D. Thesis, University of Calgary, Calgary, Canada, 2004.

[12] G. DMITRAŠINOVIĆ-VIDOVIĆ AND A. WARE, *Asymptotic behaviour of mean-quantile efficient portfolios*, Finance Stoch., 10 (2006), pp. 529–551.

[13] S. EMMER AND C. KLÜPPELBERG, *Optimal portfolios when stock prices follow an exponential Lévy process*, Finance Stoch., 8 (2004), pp. 17–44.

[14] S. EMMER, C. KLÜPPELBERG, AND R. KORN, *Optimal portfolios with bounded capital at risk*, Math. Finance, 11 (2001), pp. 365–384.

[15] E. F. FAMA AND K. R. FRENCH, *Business conditions and expected returns on stocks and bonds*, J. Financial Econ., 25 (1989), pp. 23–49.

[16] D. FILIPOVIĆ, *A general characterization of one factor affine term structure models*, Finance Stoch., 5 (2001), pp. 389–412.

[17] D. FILIPOVIĆ, D. DUFFIE, AND W. SCHACHERMAYER, *Affine processes and applications in finance*, Ann. Appl. Probab., 13 (2003), pp. 984–1053.

[18] L. A. GIL-ALANA, *Mean reversion in the real exchange rates*, Econom. Lett., 69 (2000), pp. 285–288.

[19] R. KORN AND E. KORN, *Option Pricing and Portfolio Optimization*, Grad. Stud. Math. 31, AMS, Providence, RI, 2000.

[20] A. LARI-LAVASSANI, A. A. SADEGHI, AND T. WARE, *Modeling and Implementing Mean Reverting Price Processes in Energy Markets*, Electronic Publications of the International Energy Credit Association (www.ieca.net), Mt. Laurel, NJ, 2001.

[21] D. PILIPOVIĆ, *Energy Risk: Valuing and Managing Energy Derivatives*, McGraw-Hill Professional, New York, 2007.

[22] E. S. SCHWARTZ, *The stochastic behavior of commodity prices: Implications for valuation and hedging*, J. Finance, 52 (1997), pp. 923–973.

[23] J. A. WACHTER, *Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets*, J. Financial Quant. Anal., 37 (2002), pp. 63–91.