Optimality of Naive Investment Strategies in Dynamic Mean-Variance Optimization Problems with Multiple Priors

Yuki Shigeta

July, 2016

Optimality of Naive Investment Strategies in Dynamic Mean-Variance Optimization Problems with Multiple Priors*

Yuki Shigeta[†]

July 1, 2016

Abstract

We study dynamic mean-variance optimization problems with multiple priors. We introduce two types of multiple priors, the priors for expected returns and the priors for covariances. Our framework suggests that the global minimum-variance portfolio is optimal when the investor strongly doubts the correctness of the estimated expected returns, and the equally weighted portfolio is optimal when the investor strongly doubts the correctness of the estimated covariances. From the back tests, we find that for some data sets, the strategy that invests in the global minimum-variance portfolio or the equally weighted portfolio considering the market condition is more efficient than the other mean-variance efficient portfolios.

Key words: Robust mean-variance optimization; dynamic portfolio selections; naive diversification; global minimum-variance portfolio; mean-variance efficiency.

JEL Classification: G11

1 Introduction

The mean-variance analysis proposed by Markowitz (1952) and developed by Tobin (1958) and Merton (1972) tells us that the tangency portfolio combined with the risk-free asset is mean-variance efficient, which means that it delivers the highest Sharpe ratio among all portfolios. However, it is not always efficient in practice. For example, Jagannathan and Ma (2003) mention that the global minimum-variance portfolio delivers as large an out-of-sample Sharpe ratio as the other efficient portfolios when the estimated expected returns are used. DeMiguel, Garlappi, and Uppal (2009b) report that in their back tests, the equally weighted portfolio ($1/N$ portfolio) often has a higher Sharpe ratio than other optimal portfolios including the tangency portfolio.

Probable explanations of these practical pitfalls are difficulties in estimating expected returns and sensitivity of the tangency portfolio to the expected returns. Merton (1980) reports that the estimation of expected returns is more difficult than that of a covariance matrix of returns. Best and Grauer (1991) show that the mean-variance efficient portfolios are very sensitive to the values of the expected returns. Therefore, the estimation errors of expected returns have a large impact on optimal portfolio selections. Consequently, in practice, the global minimum-variance portfolio, which is not mean-variance efficient in theory, and the equally weighted portfolio, which seems to be naive, can become more efficient than the mean-variance efficient portfolios.

In this paper, we construct plausible mean-variance optimization problems with multiple priors and show that in limiting cases, the solutions to these problems include the equally weighted portfolio and the global minimum-variance portfolio. The mean-variance problems with multiple priors are robust to estimation errors of moments of asset returns. We consider two-dimensional multiple priors: the priors for expected returns and the priors for a covariance matrix of returns. In our framework, the naive and inefficient strategies such as the equally weighted portfolio and the global minimum-variance portfolio are characterized by an investor's suspicion of the estimated parameters. If an investor strongly doubts the estimated expected returns, her optimal portfolio becomes similar to the global minimum-variance portfolio. In contrast to that, it becomes similar to the equally weighted portfolio if an investor strongly doubts the estimated covariances.

Moreover, our framework is based on dynamic optimization problems and allows dependency among the asset returns over time. More specifically, the market model in our framework is the Markovian market model where a Markov process affects the distribution of returns. This includes various return models such as the factor pricing models (e.g., capital asset pricing model (CAPM) by Sharpe (1964) and the Fama-French three factor model by Fama and French (1993)), the Markov regime switching models (e.g., Hamilton (1989) and Ang and Bekaert (2002)) and so on. Furthermore, the dynamic approach justifies the investment strategies which seem myopic. For example, the strategy investing in the global minimum-variance portfolio during economic booming and investing in the equally weighted portfolio during economic recession can be justified by our framework.

In order to investigate the effects of the estimation errors and study the efficiency of the optimal portfolio in our framework, we conduct back tests using various data sets. Similar to previous research, our back tests demonstrate that the global minimum-variance portfolio is sometimes more efficient than the tangency portfolio. On the other hand, the investment strategy that invests in the equally weighted portfolio during the economic booming and the global minimum-variance portfolio during the economic recession is sometimes the most efficient in all the portfolios. This strategy seems to be naive. However, since it is characterized by the limiting cases in our framework, the reason why the investor chooses it is because of the strong suspicion of the estimated parameters.

In the investment theory, the various methods have been proposed to deal with the esti-

mation errors. They are classified into two major approaches: the Bayesian approach and the non-Bayesian approach. In this paper, we adopt the non-Bayesian approach, specifically, the max-min approach; maximizing an objective function with respect to portfolios *after* minimizing the objective function with respect to priors of return's distribution. The max-min problem is popular when dealing with the multiple prior optimization. Goldfarb and Iyengar (2003) solve the mean-variance portfolio selections with multiple priors under the factor pricing model. Garlappi, Uppal, and Wang (2007) characterize multiple priors for the expected returns by the confidence interval around the estimated expected returns and find that their multiple-priors optimal portfolios have high Sharpe ratios in their back tests. The max-min approach is economically axiomatized by Gilboa and Schmeidler (1989) in order to describe the ambiguity aversion of investors illustrated by Ellsberg (1961).

In addition, we use one concept from the Bayesian approach. The Bayesian approach is based on the Bayesian statistics, e.g., shrinkage estimators (Jobson and Korkie (1980) and Jorion (1986)), the Black-Litterman model (Black and Litterman (1990)) and the other methods such as Pástor and Stambaugh (2000). Ledoit and Wolf (2003) propose the shrinkage estimator to estimate a covariance matrix of asset returns when facing the small sample problem that is when the number of assets is larger than the number of observations. To compute the shrinkage estimator, they minimize the Frobenius norm of estimated errors of the covariance matrix. Using the idea of Ledoit and Wolf (2003), we introduce the Frobenius norm of estimation errors of a covariance matrix in order to deal with multiple priors.

The results obtained by the mean-variance optimization constraining portfolio norms are similar to the results by our framework. The mean-variance optimization constraining portfolio norms is proposed by DeMiguel, Garlappi, Nogales, and Uppal (2009a). Especially, their 2-norm-constrained portfolio is basically the same as the one of our limiting cases. The differences between our model and the one in DeMiguel et al. (2009a) are investment horizon and problem formulation. DeMiguel et al. (2009a) consider static portfolio selections such as

$$\min_u u'\widehat{\Sigma}u \quad \text{subject to } u'u \leq k,$$

where u is a portfolio vector and u' is its transpose. $\widehat{\Sigma}$ is an estimated covariance matrix. k represents the upper bound of the constraint. DeMiguel et al. (2009a) solve the above static minimum-variance problem since it is free from the problem of time-inconsistency[1]. The optimal 2-norm-constrained portfolio is based on the covariance estimator of Ledoit and Wolf (2004). If the upper boundary of the constraint is the reciprocal of the number of assets, then the optimal portfolio becomes the equally weighted portfolio.

In contrast to the above problem, we consider the mean-variance-style objective function such that

$$\mathrm{E}^{\theta,V}\left[\sum_{t=0}^{T-1}\delta^t\left(\mathrm{E}^{\theta,V}[X_{t+1}|\mathcal{F}_t^{R,F}] - \frac{\widehat{\gamma}_t}{2X_t}\mathrm{Var}^{\theta,V}(X_{t+1}|\mathcal{F}_t^{R,F})\right)\right],$$

where X_t is a wealth of the investor at time t and $\mathcal{F}_t^{R,F}$ is the information of the investor at time t. $\widehat{\gamma}_t/X_t$ represents the trade-off between the expected return and associated risks. The superscripts θ and V express multiple priors and δ is a constant discount rate. Hence, this objective function is an expected value of discounted sum of objectives in conditional mean-variance problems. The above objective function is proposed by Chen, Li, and Zhao (2014) to avoid the problem of time-inconsistency. Chen et al. (2014) consider this objective function with a unique prior, but we consider the objective function with multiple priors.

Although there are several differences, the results of Pflug, Pichler, and Wozabal (2012) are essentially the same as our results. Pflug et al. (2012) characterize a degree of ambiguity

---

[1]The standard dynamic programming procedure can not be applied to the simple mean-variance optimization. See Li and Ng (2000) for more details.

as the Kantorovich metric, and show that for various objective functions including a mean-standard-deviation objective function and for arbitrary distributions of asset returns, the investor chooses the equally weighted portfolio under high model ambiguity. In contrast to Pflug et al. (2012), we focus on mean-variance objective functions and Markovian return models. This allows us to use more simple and clear degrees of ambiguity than the Kantorovich metric, that is, we use a confidence interval around the estimated expected return vector and a relative error of the estimated covariance matrix. Furthermore, our optimization problems are dynamic optimization problems, whereas optimization problems in Pflug et al. (2012) are static optimization problems. Therefore, our models have different implications from the model in Pflug et al. (2012).

The rest of this paper is organized as follows. Section 2 formulates and solves the dynamic mean-variance optimization problem with multiple priors. Section 3 investigates limiting behaviors of the solutions. Section 4 conducts back tests and reports their results. Section 5 is the concluding section. The most proofs of theorems, lemmas and propositions in this paper are in Appendix.

2 Mean-Variance Portfolio Selections with Multiple Priors in a Markovian Market

First, we introduce the following notations.

- $\mathbb{R}^d$, the d-dimensional Euclidean space. Specifically, we write $\mathbb{R} := \mathbb{R}^1$ and $\mathbb{R}_+ := \{x \in \mathbb{R} \mid x \geq 0\}$.

- $\mathbb{R}^{m,n}$, a set of $m \times n$ matrices.

- x', the transpose of a vector or matrix x.

- $\mathbf{0}_d$, the d-dimensional real-valued vector whose all elements are 0.

- $\mathbf{1}_d$, the d-dimensional real-valued vector whose all elements are 1.

- I_d, the d-dimensional identity matrix.

- $\mathbb{1}_A(\omega)$, an indicator function. If $\omega \in A$, then $\mathbb{1}_A(\omega) = 1$. If $\omega \notin A$, then $\mathbb{1}_A(\omega) = 0$.

- $(x)^+ := \max\{x, 0\}$ for all $x \in \mathbb{R}$.

- $\|x\|$, a general expression of norms. Specifically, $\|x\|$ is the Euclidean norm of $x \in \mathbb{R}^d$ and $\|A\|$ is the Frobenius norm of $A \in \mathbb{R}^{m \times n}$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We consider optimal portfolio selections of an investor during T periods. There are one risk-free asset and d risky assets in the financial market. Moreover, we assume that a sequence of $\mathbb{R}^K$-valued random vectors, denoted by $(F_t)_{t=0}^T$, influences distributions of the return vectors of the risky assets: The return vector of the risky assets at time t, denoted by R_t, satisfies

$$R_t = \mu(F_t) + \sigma(F_t)\epsilon_t, \quad t \geq 0,$$

where $(\epsilon_t)_{t=1}^T$ is a sequence of mutually independent and d-dimensional standard normal random vectors, and $\mu : \mathbb{R}^K \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^K \to \mathbb{R}^{d \times d}$ are measurable functions. Furthermore,

we assume that F_t and ϵ_t are mutually independent for each $t = 1, \ldots, T$. Then, the conditional expectation and variance of R_t given by F_t are

$$\mathrm{E}[R_t \mid F_t] = \mu(F_t),$$
$$\mathrm{Var}(R_t \mid F_t) = \sigma(F_t)(\sigma(F_t))',$$

for all $t \geq 1$. Hence, the conditional mean and variance of the risky assets' returns are driven by the movement of $(F_t)_{t=0}^T$. For convenience, we write $R_0 = \mathbf{0}_d$. The risk-free rate is a constant over time and it is denoted by r_{f}, and $(R_t^e)_{t=1}^T$ is the excess return process, $R_t^e = R_t - r_{\mathrm{f}}\mathbf{1}_d$.

Denote by $\mathbb{F}^{R,F} = (\mathcal{F}_t^{R,F})_{t=0}^T$ and $\mathbb{F}^F = (\mathcal{F}_t^F)_{t=0}^T$ the filtrations generated by $(R_t, F_t)_{t=0}^T$ and $(F_t)_{t=0}^T$, respectively. We assume that an investor can observe the value of F_t and R_t at each $t \geq 0$, so her information at time t is represented by $\mathcal{F}_t^{R,F}$. In addition, we assume the following:

Assumption 1 The random process $(F_t)_{t=0}^T$ is a time-homogeneous Markov process with respect to $\mathbb{F}^{R,F}$.

Assumption 1 is important for our optimization. Many return models satisfy this assumption.

Example 2

1. *Factor pricing models.* Suppose that $(F_t)_{t=0}^T$ satisfies Assumption 1. Let $\mu(F) = r_{\mathrm{f}}\mathbf{1}_d + BF$ and $\sigma(F) = \overline{\sigma}$, where $B \in \mathbb{R}^{d,K}$ and $\overline{\sigma} \in \mathbb{R}^{d,d}$ are constant matrices. Then, the return process can be expressed as

$$R_t = r_{\mathrm{f}}\mathbf{1}_d + BF_t + \overline{\sigma}\epsilon_t, \quad t \geq 1.$$

The above return process is the one of factor pricing models such as the capital asset pricing model (CAPM) and the Fama-French three factor model.

2. *Stochastic volatility models.* Suppose that $K = d$ and that $(F_t)_{t=0}^T$ is specified by

$$\log(F_t) = m + B(\log(F_{t-1}) - m) + \eta_t, \quad t \geq 1,$$

where $m \in \mathbb{R}^d$ is a constant vector and $B \in \mathbb{R}^{d,d}$ is a constant diagonal matrix. $(\eta_t)_{t=1}^T$ is a sequence of d-dimensional i.i.d. random vectors that are independent of $(\epsilon_t)_{t=1}^T$. The return process is defined as

$$R_t = \overline{\mu} + \mathrm{diag}(\sqrt{F_t})C\epsilon_t, \quad t \geq 1,$$

where CC' is a constant correlation matrix. Then, $(F_t)_{t=0}^T$ satisfies Assumption 1.

3. *Markov regime-switching models.* Suppose that $(F_t)_{t=0}^T$ is a K-states, time-homogeneous Markov chain and that μ and σ are specified by

$$\mu(F_t) = \mu_i, \quad \sigma(F_t) = \sigma_i, \quad \text{if } F_t \text{ is } i\text{th state}.$$

In addition, $(F_t)_{t=0}^T$ and $(\epsilon_t)_{t=1}^T$ are mutually independent. Then, $(F_t)_{t=0}^T$ satisfies Assumption 1.

In usual portfolio optimization, researchers assume that an investor knows the conditional mean $\mu(F_t)$ and variance $\sigma(F_t)(\sigma(F_t))'$, but, the investor needs to estimate the moments in practice. In the robust optimization literature, researchers take into account the estimation error of these moments. Similar to the literature, we introduce two statistical error components. Let $(\theta_t)_{t=0}^{T-1}$ be a $\mathbb{R}^d$-valued random process of errors of mean and let $(V_t)_{t=0}^{T-1}$ be a $\mathbb{R}^{d,d}$-valued process of errors of variance. The one-period-ahead conditional mean and variance with respect to the probability measure representing the investor's belief, denoted by $\mathbb{P}^{\theta,V}$, can be expressed as

$$\mathrm{E}^{\theta,V}[R_{t+1}|\mathcal{F}_t^{R,F}] = \mathrm{E}[R_{t+1}|\mathcal{F}_t^{R,F}] + \theta_t, \quad \mathrm{Var}^{\theta,V}(R_{t+1}|\mathcal{F}_t^{R,F}) = \mathrm{Var}(R_{t+1}|\mathcal{F}_t^{R,F}) + V_t, \quad t \geq 0.$$

Hence, θ_t and V_t can be regarded as statistical errors of conditional mean and variance given $\mathcal{F}_t^{R,F}$ at each t. The assumptions regarding $(\theta_t)_{t=0}^{T-1}$ and $(V_t)_{t=0}^{T-1}$ are

Assumption 3

1. For each t, θ_t and V_t are $\mathcal{F}_t^{R,F}$-measurable.

2. For each t, $\mathrm{Var}(R_{t+1}|\mathcal{F}_t^{R,F}) + V_t$ and $\sigma(F_{t+1})(\sigma(F_{t+1}))' + V_t$ are positive-definite matrices.

From the definitions of $(\theta_t)_{t=0}^{T-1}$ and $(V_t)_{t=0}^{T-1}$, Assumption 3 is natural.

The important technical issue is whether the probability measure $\mathbb{P}^{\theta,V}$ exists or not. To construct the probability measure $\mathbb{P}^{\theta,V}$, we introduce the random variable,

$$Z_t = \prod_{s=0}^{t} \zeta_s, \quad t \geq 0,$$

where

$$\zeta_{t+1} = \sqrt{\frac{\det(\sigma(F_{t+1})(\sigma(F_{t+1}))')}{\det(\sigma(F_{t+1})(\sigma(F_{t+1}))' + V_t)}}$$
$$\times \exp\left\{ -\frac{1}{2}\theta_t'\Big(\sigma(F_{t+1})(\sigma(F_{t+1}))' + V_t\Big)^{-1}\theta_t + \theta_t'\Big(\sigma(F_{t+1})(\sigma(F_{t+1}))' + V_t\Big)^{-1}\sigma(F_{t+1})\epsilon_{t+1} \right.$$
$$\left. +\frac{1}{2}\epsilon_{t+1}'\Big(I_d - (\sigma(F_{t+1}))'\Big(\sigma(F_{t+1})(\sigma(F_{t+1}))' + V_t\Big)^{-1}\sigma(F_{t+1})\Big)\epsilon_{t+1} \right\}, \quad t \geq 1$$

and $\zeta_0 = 1$. Then, the following lemma holds.

Lemma 4 Under Assumptions 1 and 3, there exists a probability measure $\mathbb{P}^{\theta,V}$, such that

$$\mathbb{P}^{\theta,V}(A) = \mathrm{E}[\mathbb{1}_A Z_T], \quad A \in \mathcal{F}_T^{R,F}.$$

Under $\mathbb{P}^{\theta,V}$,

$$\mathrm{E}^{\theta,V}[R_{t+1}|\mathcal{F}_t^{R,F}] = \mathrm{E}[R_{t+1}|\mathcal{F}_t^{R,F}] + \theta_t,$$
$$\mathrm{Var}^{\theta,V}(R_{t+1}|\mathcal{F}_t^{R,F}) = \mathrm{Var}(R_{t+1}|\mathcal{F}_t^{R,F}) + V_t,$$

for all $t \geq 0$. Furthermore, the conditional distribution of F_{t+1} given $\mathcal{F}_t^{R,F}$ under $\mathbb{P}^{\theta,V}$ is the same as the conditional distribution of F_{t+1} given $\mathcal{F}_t^{R,F}$ under $\mathbb{P}$.

The proof of Lemma 4 is in Section A.1. Since $(F_t)_{t=0}^T$ is the time-homogeneous Markov process with respect to $\mathbb{F}^{R,F}$, we can express the conditional mean and variance of R_{t+1} as functions depending on F_t, that is,

$$m(F) := \mathrm{E}\left[R_{t+1}\middle|F_t = F\right] = \mathrm{E}\left[R_{t+1}\middle|\mathcal{F}_t^{R,F}\right],$$

$$A(F) := \mathrm{Var}(R_{t+1}\middle|F_t = F) = \mathrm{Var}(R_{t+1}\middle|\mathcal{F}_t^{R,F}),$$

for all $t \geq 0$. We denote the conditional expected excess return by $m^e(F) = m(F) - r_{\mathrm{f}}\mathbf{1}_d$.

Now, let us formulate the optimization problem of the investor. Let $(X_t)_{t=0}^T$ be a wealth process of the investor. We assume that X_t satisfies a self-financing constraint, that is,

$$X_{t+1} = \Big((1 + r_{\mathrm{f}}) + (R_{t+1}^e)'u_t\Big)X_t, \quad t \geq 0,$$

where $(u_t)_{t=0}^{T-1}$ is a sequence of d-dimensional portfolio vectors for risky assets. For each u_t, the ith element of u_t represents the weight of the wealth invested in the ith risky asset. The investor buys or sells $X_t(1 - (\mathbf{1}_d)'u_t)$ dollar's risk-free asset at each $t \geq 0$. The investor's objective function is

$$\mathrm{E}^{\theta,V}\left[\sum_{t=0}^{T-1}\delta^t\left(\mathrm{E}^{\theta,V}[X_{t+1}|\mathcal{F}_t^{R,F}] - \frac{\widehat{\gamma}_t}{2X_t}\mathrm{Var}^{\theta,V}(X_{t+1}|\mathcal{F}_t^{R,F})\right)\right], \tag{2.1}$$

where $\widehat{\gamma} := (\widehat{\gamma}_t)_{t=0}^{T-1}$ is a positive and $\mathbb{F}^F$-adapted process which represents the investor's trade-off between the expected return and associated risks. We emphasize that $\widehat{\gamma}$ is not only $\mathbb{F}^{R,F}$-adapted, but also $\mathbb{F}^F$-adapted. The $\mathbb{F}^F$-adaptedness of $\widehat{\gamma}$ is important in the derivation of the explicit form of the value function.

The trade-off parameter is divided by X_t whereas in the other standard problem, it is not divided. This parameterization is used by Björk, Murgoci, and Zhou (2014). They provide two natural interpretations of the parameterization. One is an adjustment of units. The unit of the conditional expected wealth is $(dollar)$, whereas the unit of the conditional variance is $(dollar)^2$. So the variance needs to be divided by wealth in order to measure the objective function in dollars. Another interpretation is to measure the objective function by the rate of return. We replace the wealth in the objective by the gross return of the wealth, namely, we consider

$$\mathrm{E}^{\theta,V}\left[\frac{X_{t+1}}{X_t}\middle|\mathcal{F}_t^{R,F}\right] - \frac{\widehat{\gamma}_t}{2}\mathrm{Var}^{\theta,V}\left(\frac{X_{t+1}}{X_t}\middle|\mathcal{F}_t^{R,F}\right)$$

$$= \frac{1}{X_t}\left\{\mathrm{E}^{\theta,V}[X_{t+1}|\mathcal{F}_t^{R,F}] - \frac{\widehat{\gamma}_t}{2X_t}\mathrm{Var}^{\theta,V}(X_{t+1}|\mathcal{F}_t^{R,F})\right\}.$$

Then, Björk et al. (2014) argue that the equilibrium in the above objective is the same as that in the following objective,

$$\mathrm{E}^{\theta,V}[X_{t+1}|\mathcal{F}_t^{R,F}] - \frac{\widehat{\gamma}_t}{2X_t}\mathrm{Var}^{\theta,V}(X_{t+1}|\mathcal{F}_t^{R,F}).$$

However, in our settings, the optimal portfolios are different between these two objectives. Our settings, therefore, should be interpreted as the adjustment of units.

We can also interpret our objective function as an approximation of a certainty equivalent of a CRRA utility. Let

$$u_t(x) := \begin{cases} \dfrac{x^{1-\widehat{\gamma}_t} - 1}{1 - \widehat{\gamma}_t}, & \text{if } \widehat{\gamma}_t \neq 1, \\ \log x, & \text{if } \widehat{\gamma}_t = 1, \end{cases}$$

7

where $\widehat{\gamma}_t > 0$ is the constant relative risk aversion coefficient at time t. Let us denote by $\bar{c}_t(X)$ the certainty equivalent of a random variable X under the probability measure $\mathbb{P}^{\theta,V}$, that is

$$\bar{c}_t(X) := (u_t)^{-1}\Big(\mathrm{E}^{\theta,V}[u_t(X) \mid \mathcal{F}_t^{R,F}]\Big),$$

where $(u_t)^{-1}$ is the functional inverse of u_t. Then, it is well known that $\bar{c}_t(X)$ can be approximated as follows.

$$\bar{c}_t(X_{t+1}) \approx \mathrm{E}^{\theta,V}[X_{t+1} \mid \mathcal{F}_t^{R,F}] + \frac{1}{2}\frac{u_t''(\mathrm{E}^{\theta,V}[X_{t+1} \mid \mathcal{F}_t^{R,F}])}{u_t'(\mathrm{E}^{\theta,V}[X_{t+1} \mid \mathcal{F}_t^{R,F}])}\,\mathrm{Var}^{\theta,V}(X_{t+1} \mid \mathcal{F}_t^{R,F}),$$

where u_t' and u_t'' are the first and the second derivatives of u_t, respectively. Now, we assume $X_t \approx \mathrm{E}^{\theta,V}[X_{t+1} \mid \mathcal{F}_t^{R,F}]$. Then, since u_t is a CRRA utility, we have

$$\frac{u_t''(\mathrm{E}^{\theta,V}[X_{t+1} \mid \mathcal{F}_t^{R,F}])}{u_t'(\mathrm{E}^{\theta,V}[X_{t+1} \mid \mathcal{F}_t^{R,F}])} = -\frac{\widehat{\gamma}_t}{\mathrm{E}^{\theta,V}[X_{t+1} \mid \mathcal{F}_t^{R,F}]} \approx -\frac{\widehat{\gamma}_t}{X_t}.$$

Therefore, we have

$$\bar{c}_t(X_{t+1}) \approx \mathrm{E}^{\theta,V}[X_{t+1} \mid \mathcal{F}_t^{R,F}] - \frac{\widehat{\gamma}_t}{2X_t}\,\mathrm{Var}^{\theta,V}(X_{t+1} \mid \mathcal{F}_t^{R,F}).$$

Hence, the expectation of the discounted sum of the certainty equivalents can be approximated as follows:

$$\mathrm{E}^{\theta,V}\left[\sum_{t=0}^{T-1}\delta^t\bar{c}_t(X_{t+1})\right] \approx \mathrm{E}^{\theta,V}\left[\sum_{t=0}^{T-1}\delta^t\left(\mathrm{E}^{\theta,V}[X_{t+1}|\mathcal{F}_t^{R,F}] - \frac{\widehat{\gamma}_t}{2X_t}\,\mathrm{Var}^{\theta,V}(X_{t+1}|\mathcal{F}_t^{R,F})\right)\right].$$

The above approximation coincides with our objective function (2.1). Furthermore, the time varying $\widehat{\gamma}$ implies that the relative risk aversion coefficient varies over time.

Note that, we assume that X_t is almost surely strictly positive for all $t \geq 0$. Strictly speaking, it does not hold when permitting short selling, whereas it holds if short selling is not allowed. Nevertheless, we assume the strict positivity of X_t in all cases for mathematical convenience. Furthermore, the division of the trade-off parameter by X_t is a crucial assumption for the derivation of the explicit solution to our problem. Without it, we can not derive the explicit solution. We again refer to this assumption in the problem without the risk-free asset.

By the Markov property of $(F_t)_{t=0}^T$, the objective under $\mathbb{P}^{\theta,V}$ can be expressed as

$$\mathrm{E}^{\theta,V}\left[\sum_{t=0}^{T-1}\delta^t\left(\mathrm{E}^{\theta,V}[X_{t+1}|\mathcal{F}_t^{R,F}] - \frac{\widehat{\gamma}_t}{2X_t}\,\mathrm{Var}^{\theta,V}(X_{t+1}|\mathcal{F}_t^{R,F})\right)\right]$$

$$= \mathrm{E}^{\theta,V}\left[\sum_{t=0}^{T-1}\delta^t X_t c_t(F_t, \theta_t, V_t, u_t)\right],$$

where $c_t : \mathbb{R}^K \times \mathbb{R}^d \times \mathbb{R}^{d,d} \times \mathbb{R}^d \to \mathbb{R}$ is a measurable function such that

$$c_t(F, \theta, V, u) = 1 + r_{\mathrm{f}} + \Big(m^e(F) + \theta\Big)'u - \frac{\widehat{\gamma}_t}{2}u'(A(F) + V)u.$$

Now, we construct the optimization problem with multiple priors. The investor selects the portfolio maximizing the expected utility of the worst case. This means that the investor

faces on the following maximization problem:

$$V_T(x, f) = \max_{(u_t)_{t=0}^{T-1} \in \mathcal{A}_T^u} \min_{(\theta_t, V_t)_{t=0}^{T-1} \in \mathcal{A}_T^{\theta, V}} \mathrm{E}^{\theta, V} \left[\sum_{t=0}^{T-1} \delta^t X_t c_t(F_t, \theta_t, V_t, u_t) \right]$$

$$\text{subject to } X_{t+1} = \left(1 + r_\mathrm{f} + (R_{t+1}^e)' u_t \right) X_t, \quad t \geq 0, \tag{2.2}$$

$$X_0 = x, \quad F_0 = f,$$

where $\mathcal{A}_T^u$ is a set of admissible portfolios and $\mathcal{A}_T^{\theta, V}$ is a set of admissible errors, $(\theta_t)_{t=0}^{T-1}$ and $(V_t)_{t=0}^{T-1}$. The set of admissible portfolios can be expressed as

$$\mathcal{A}_T^u = \left\{ u = (u_t)_{t=0}^{T-1} \mid u \text{ is a } \mathbb{R}^d\text{-valued and } \mathbb{F}^{R,F}\text{-adapted process.} \right\}.$$

So, the investor is permitted to short sell. We later analyze the case when the investor is prohibited from short selling.

The set of admissible errors is characterized by two inequalities. For all $(\theta_t, V_t)_{t=0}^{T-1} \in \mathcal{A}_T^{\theta, V}$, $(\theta_t)_{t=0}^{T-1}$ is a $\mathbb{R}^d$-valued process and $(V_t)_{t=0}^{T-1}$ is an $\mathbb{R}^{d \times d}$-valued process. Furthermore, they satisfy Assumption 3 and

$$(\eta_t^\theta)^2 \geq \theta_t' \left(A(F_t) + V_t \right)^{-1} \theta_t, \tag{2.3}$$

$$(\eta_t^V)^2 \mathrm{tr}((A(F_t))^2) \geq \mathrm{tr}(V_t^2), \tag{2.4}$$

where $(\eta_t^\theta)_{t=0}^{T-1}$ and $(\eta_t^V)_{t=0}^{T-1}$ are positive $\mathbb{F}^F$-adopted processes and $\mathrm{tr}(A)$ is the trace of matrix A.

The inequality (2.3) represents a confidence interval in statistics. This idea is introduced by Garlappi et al. (2007). We can express $\theta_t'(A(F_t) + V_t)^{-1} \theta_t$ as

$$\theta_t' \left(\mathrm{Var}^{\theta, V}(R_{t+1} | \mathcal{F}_t^{R,F}) \right)^{-1} \theta_t =$$

$$\left(\mathrm{E}[R_{t+1} | \mathcal{F}_t^{R,F}] - \mathrm{E}^{\theta, V}[R_{t+1} | \mathcal{F}_t^{R,F}] \right)' \left(\mathrm{Var}^{\theta, V}(R_{t+1} | \mathcal{F}_t^{R,F}) \right)^{-1} \left(\mathrm{E}[R_{t+1} | \mathcal{F}_t^{R,F}] - \mathrm{E}^{\theta, V}[R_{t+1} | \mathcal{F}_t^{R,F}] \right),$$

so $\theta_t'(A(F_t) + V_t)^{-1} \theta_t$ is the F statistic of the estimated means in the null hypothesis of $\mathrm{E}[R_{t+1} | \mathcal{F}_t^{R,F}] = \mathrm{E}^{\theta, V}[R_{t+1} | \mathcal{F}_t^{R,F}]$.

Garlappi et al. (2007) characterize this constraint of $(\theta_t)_{t=0}^{T-1}$ as the confidence interval of the estimated mean vector. Their discussion is justified if the return vectors of the risky assets are i.i.d.. However, the return vectors in our model are *not* i.i.d. in general. Therefore, our constraint (2.3) does not represent the confidence interval. Since the financial models are usually estimated by using the maximum likelihood methods, one should use the inverse of the information matrix as the covariance matrix in our constraint (2.3) in order to characterize this constraint as the confidence interval. However, in this paper, we use this constraint, similar to Garlappi et al. (2007), to reduce the mathematical complexity.

The inequality (2.4) can be expressed as

$$\eta_t^V \geq \frac{\mathrm{tr}(V_t^2)}{\mathrm{tr}((A(F_t))^2)} = \frac{\| \mathrm{Var}^{\theta, V}(R_{t+1} | \mathcal{F}_t^{R,F}) - \mathrm{Var}(R_{t+1} | \mathcal{F}_t^{R,F}) \|^2}{\| \mathrm{Var}(R_{t+1} | \mathcal{F}_t^{R,F}) \|^2},$$

for all $t \geq 0$. Therefore, $(\eta_t^V)_{t=0}^{T-1}$ represents the upper boundary of the ratio of the least square error to the conditional variance. Ledoit and Wolf (2003) consider the shrinkage estimator minimizing the following objective:

$$\min_\alpha \mathrm{tr} \left((\widehat{\Sigma}(\alpha) - \Sigma)^2 \right) = \min_\alpha \| \widehat{\Sigma}(\alpha) - \Sigma \|^2,$$

where Σ is an actual covariance matrix of returns and $\widehat{\Sigma}(\alpha)$ is a shrunk estimated covariance matrix depending on the parameter α. We use the above objective in the inequality (2.4). Therefore, the inequality (2.4) represents the upper boundary of the objective in Ledoit and Wolf (2003).

As in the case of a unique prior, we can apply the dynamic programming procedure to the optimization problem (2.2).

Theorem 5 Let $(Y_t^T)_{t=0}^{T-1}$ be a sequence of random variables, such that

$$Y_{T-1}^T = \max_{u_{T-1} \in \mathbb{R}^d} \min_{(\theta_{T-1}, V_{T-1}) \in \mathcal{A}_T^{\theta,V}|_{T-1}^{T-1}} \left\{ c_{T-1}(F_{T-1}, \theta_{T-1}, V_{T-1}, u_{T-1}) \right\},$$

$$Y_t^T = \max_{u_t \in \mathbb{R}^d} \min_{(\theta_t, V_t) \mathcal{A}_T^{\theta,V}|_t^t} \left\{ c_t(F_t, \theta_t, V_t, u_t) + \delta \, \mathrm{E}^{\theta,V} \left[Y_{t+1}^T (1 + r_{\mathrm{f}} + (R_{t+1}^e)' u_t) \, \Big| \, \mathcal{F}_t^{R,S} \right] \right\},$$

for all $0 \leq t \leq T - 2$, where $\mathcal{A}_T^{\theta,V}|_t^t$ is a set of t-th elements in $\mathcal{A}_T^{\theta,V}$. Then, the above random sequence exists, and

$$V_T(x, f) = \mathrm{E}\left[Y_0^T \mid F_0 = f \right] x, \quad x \in \mathbb{R}, \ f \in \mathbb{R}^K.$$

Furthermore, the optimal portfolio process $(u_t^*)_{t=0}^{T-1}$ derived by the above dynamic programming is time-consistent.

The explicit solution is provided by the following proposition.

Proposition 6 Let

$$Y_T^T = 0,$$
$$m_t^e = \mathrm{E}\left[(1 + \delta Y_{t+1}^T)(\mu(F_{t+1}) - r_{\mathrm{f}} \mathbf{1}_d) | \mathcal{F}_t^{R,F} \right], \quad g_t^e = 1 + \delta \, \mathrm{E}\left[Y_{t+1}^T | \mathcal{F}_t^{R,F} \right],$$
$$a_t^e = (m_t^e)' \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d \right)^{-1} m_t^e,$$

for all $0 \leq t \leq T - 1$. Then, $(Y_t^T)_{t=0}^{T-1}$ in Theorem 5 can be expressed as

$$Y_t^T = (1 + r_{\mathrm{f}})g_t^e + \frac{((\sqrt{a_t^e} - |g_t^e| \eta_t^\theta)^+)^2}{2\widehat{\gamma}_t},$$

and the optimal portfolio is

$$u_t^* = \frac{1}{\widehat{\gamma}_t} \left(1 - \frac{|g_t^e| \eta_t^\theta}{\sqrt{a_t^e}} \right)^+ \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d \right)^{-1} m_t^e. \tag{2.5}$$

The proofs of Theorem 5 and Proposition 6 are in Section A.2. By the construction, $(Y_t^T)_{t=0}^{T-1}$ is $\mathbb{F}^F$-adapted. Furthermore, if $\widehat{\gamma}_t, \eta_t^\theta$ and η_t^V depend only on the value of F_t for all $0 \leq t \leq T - 1$, Y_t^T also depends only on the value of F_t. Therefore, we can easily compute it numerically using the Markov property of $(F_t)_{t=0}^T$.

The optimal portfolio (2.5) is similar to the ambiguity averse minimum-variance portfolio by Pınar (2014). However, the multiple priors for mean and variance affects the portfolio (2.5) whereas the ambiguity averse minimum-variance portfolio is only affected by the multiple priors for mean. Therefore, the ambiguity averse minimum-variance portfolio is a special case of our portfolios.

Note that, we obtain similar results to Proposition 6 even if r_{f} varies over time. Let $r_{\mathrm{f},t+1}$ be a time-varying risk-free rate at time $t + 1$ such that

$$r_{\mathrm{f},t+1} = h(F_t), \quad t \geq 0,$$

where h is a measurable function from $\mathbb{R}^K$ onto $\mathbb{R}$. The reason why the risk-free rate at time $t+1$ depends on the value of F_t is that the risk-free rate needs to be determined before investment decision. Then, the following portfolio u_t^{TV*} is optimal.

$$u_t^{TV*} = \frac{1}{\widehat{\gamma}_t}\left(1 - \frac{|g_t^{TVe}|\eta_t^\theta}{\sqrt{a_t^{TVe}}}\right)^+ \left(A(F_t) + \eta_t^V\|A(F_t)\|I_d\right)^{-1}m_t^{TVe},$$

where

$$Y_T^{TVT} = 0, \quad m_t^{TVe} = \mathrm{E}\left[(1 + \delta Y_{t+1}^{TVT})(\mu(F_{t+1}) - h(F_t)\mathbf{1}_d)|\mathcal{F}_t^{R,F}\right],$$

$$g_t^{TVe} = 1 + \delta\,\mathrm{E}\left[Y_{t+1}^{TVT}|\mathcal{F}_t^{R,F}\right], \quad a_t^{TVe} = (m_t^{TVe})'\left(A(F_t) + \eta_t^V\|A(F_t)\|I_d\right)^{-1}m_t^{TVe},$$

$$Y_t^{TVT} = (1 + h(F_t))g_t^{TVe} + \frac{((\sqrt{a_t^{TVe}} - |g_t^{TVe}|\eta_t^\theta)^+)^2}{2\widehat{\gamma}_t}.$$

It is clear that Y_t^{TVT} is $\mathcal{F}_t^{R,F}$-measurable for all $t \geq 0$. In addition, u^{TV*} is admissible. The proof is similar to that of Proposition 6, so we omit it. The above result indicates that if the movement of $r_{\mathrm{f},t}$ is independent of ϵ_t, then the optimal portfolio with the constant risk-free rate can be naturally extended into the optimal portfolio with the time-varying risk-free rate.

We next consider an optimization problem with multiple priors without the risk-free asset. In this case, the investor's wealth updating formula is

$$X_{t+1} = \left(1 + (R_{t+1})'u_t\right)X_t, \quad 0 \leq t \leq T-1,$$

where u_t satisfies $\mathbf{1}_d'u_t = 1$. The objective can be expressed as

$$\mathrm{E}^{\theta,V}\left[\sum_{t=0}^{T-1}\delta^t\left(\mathrm{E}^{\theta,V}[X_{t+1}|\mathcal{F}_t^{R,F}] - \frac{\widehat{\gamma}_t}{2X_t}\mathrm{Var}^{\theta,V}(X_{t+1}|\mathcal{F}_t^{R,F})\right)\right]$$

$$= \mathrm{E}^{\theta,V}\left[\sum_{t=0}^{T-1}\delta^t X_t\widehat{c}_t(F_t, \theta_t, V_t, u_t)\right],$$

where $\widehat{c}_t : \mathbb{R}^K \times \mathbb{R}^d \times \mathbb{R}^{d,d} \times \mathbb{R}^d \to \mathbb{R}$ is a measurable function such that

$$\widehat{c}_t(F, \theta, V, u) = 1 + \left(m(F) + \theta\right)'u - \frac{\widehat{\gamma}_t}{2}u'(A(F) + V)u.$$

Hence, the objective without the risk-free asset has the same form as the objective with the risk-free asset. This implies that Theorem 5 can be applied to the problem without the risk-free asset.

The optimization problem without the risk-free asset is formulated as

$$V_T^{worf}(x, f) = \max_{(u_t)_{t=0}^{T-1}\in\widetilde{\mathcal{A}}_T^u}\min_{(\theta_t, V_t)_{t=0}^{T-1}\in\mathcal{A}_T^{\theta,V}}\mathrm{E}^{\theta,V}\left[\sum_{t=0}^{T-1}\delta^t X_t\widehat{c}_t(F_t, \theta_t, V_t, u_t)\right]$$

$$\text{subject to } X_{t+1} = \left(1 + (R_{t+1})'u_t\right)X_t, \quad t \geq 0, \tag{2.6}$$

$$X_0 = x, \quad F_0 = f,$$

where the set of admissible portfolios $\widetilde{\mathcal{A}}_T^u$ is

$$\widetilde{\mathcal{A}}_T^u = \left\{u = (u_t)_{t=0}^{T-1} \mid u \text{ is a } \mathbb{R}^d\text{-valued and } \mathbb{F}^{R,F}\text{-adapted process and } \mathbf{1}_d'u_t = 1\right\}.$$

The following proposition provides the optimal portfolios and the value functions.

Proposition 7 For all $0 \le t \le T-1$, let

$$\widehat{Y}_T^T = 0,$$

$$g_t = 1 + \delta \, \mathrm{E}[\widehat{Y}_{t+1}^T \mid \mathcal{F}_t^{R,F}], \quad m_t = \mathrm{E}\left[(1 + \delta \widehat{Y}_{t+1}^T)\mu(F_{t+1}) \mid \mathcal{F}_t^{R,F}\right],$$

$$a_t^{worf} = m_t'\Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d\Big)^{-1} m_t, \quad c_t^{worf} = m_t'\Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d\Big)^{-1} \mathbf{1}_d,$$

$$b_t^{worf} = \mathbf{1}_d'\Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d\Big)^{-1} \mathbf{1}_d, \quad d_t^{worf} = a_t^{worf} b_t^{worf} - (c_t^{worf})^2,$$

$$\widehat{Y}_t^T = g_t + \frac{c_t^{worf} - (\widehat{\gamma}_t + |g_t|\eta_t^\theta/\psi_t^*)}{b_t^{worf}} + \frac{\widehat{\gamma}_t}{2}(\psi_t^*)^2,$$

where ψ_t^* is a unique positive solution of the polynomial equation.

$$b_t^{worf}(\psi_t^*)^2 = \frac{d_t^{worf}}{(\widehat{\gamma}_t + |g_t|\eta_t^\theta/\psi_t^*)^2} + 1. \tag{2.7}$$

Then, the value function is

$$V_T^{worf}(x, f) = \mathrm{E}[\widehat{Y}_0^T \mid F_0 = f]x, \quad (x, f) \in \mathbb{R}_+ \times \mathbb{R}^K,$$

and the time-consistent optimal portfolio is

$$u_t^* = \frac{1}{\widehat{\gamma}_t + |g_t|\eta_t^\theta/\psi_t^*}\Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d\Big)^{-1}\left(m_t - \frac{c_t^{worf} - (\widehat{\gamma}_t + |g_t|\eta_t^\theta/\psi_t^*)}{b_t^{worf}}\mathbf{1}_d\right),$$

for all $0 \le t \le T-1$.

The proof of Proposition 7 is in Section A.3. The equation (2.7) is essentially the same as the equation (A11) in Garlappi et al. (2007), so it must have a unique positive solution by Garlappi et al. (2007). We give the details of the proofs of the existence of the unique positive solution to the equation (2.7) in Section A.3.

The main difference between Garlappi et al. (2007) and our model is the covariance matrix. The optimal solution in Garlappi et al. (2007) can be regarded as the case when η_t^V is zero in our solution. In our case, the multiple priors for covariance affects the optimal portfolio.

The division of $\widehat{\gamma}_t$ by X_t has an important role in the problem without the risk-free asset. Without this division, we can not derive the explicit solution since the solution of the polynomial (2.7) depends on the value of X_t. This implies that the t-step value function $\widehat{Y}_t^T$, also depends on the value of X_t. Then, it is difficult to compute the conditional expected value of $\widehat{Y}_t^T$ given $\mathcal{F}_{t-1}^{R,F}$ under the probability measure $\mathbb{P}^{\theta,V}$. Hence, the trade-off parameter's division by the wealth is crucial.

Now we introduce a short-selling constraint. Let $\widetilde{\mathcal{A}}_{+T}^u$ be a set of admissible portfolios with a short-selling constraint such that

$$\widetilde{\mathcal{A}}_{+T}^u = \Big\{u = (u_t)_{t=0}^{T-1} \mid u \text{ is a } \mathbb{R}^d\text{-valued and } \mathbb{F}^{R,F}\text{-adapted process, } \mathbf{1}_d' u_t = 1 \text{ and } u_t \in \mathbb{R}_+^d\Big\}.$$

Then, the dynamic optimization problem with a short-selling constraint can be expressed as

$$V_{+T}^{worf}(x, f) = \max_{(u_t)_{t=0}^{T-1} \in \widetilde{\mathcal{A}}_{+T}^u} \min_{(\theta_t, V_t)_{t=0}^{T-1} \in \mathcal{A}_T^{\theta,V}} \mathrm{E}^{\theta,V}\left[\sum_{t=0}^{T-1} \delta^t X_t \widehat{c}_t(F_t, \theta_t, V_t, u_t)\right]$$

$$\text{subject to } X_{t+1} = \Big(1 + (R_{t+1})' u_t\Big)X_t, \quad t \ge 0, \tag{2.8}$$

$$X_0 = x, \quad F_0 = f.$$

Then, the following proposition holds.

Proposition 8 For all $0 \le t \le T-1$, let

$$\widehat{Y}_T^{+T} = 0,$$

$$g_t^+ = 1 + \delta \, \mathrm{E}[\widehat{Y}_{t+1}^{+T} \mid \mathcal{F}_t^{R,F}], \quad m_t^+ = \mathrm{E}\left[(1 + \delta \widehat{Y}_{t+1}^{+T}) \mu(F_{t+1}) \mid \mathcal{F}_t^{R,F} \right],$$

$$\widehat{Y}_t^{+T} = \max_{u_t \in \overline{\mathcal{C}^+}} \min_{(\theta_t, V_t) \in \mathcal{A}_T^{\theta,V}|_t^t} \left\{ g_t^+ + \left(m_t^+ + g_t^+ \theta_t \right)' u_t - \frac{\widehat{\gamma}_t}{2} u_t' \left(A(F_t) + V_t \right) u_t \right\},$$

where $\overline{\mathcal{C}^+}$ is a set of portfolios such that

$$\overline{\mathcal{C}^+} = \left\{ \phi \in \mathbb{R}_+^d \ \middle| \ 1 = \mathbf{1}_d' \phi. \right\}.$$

Then, the value function is

$$V_{+T}^{worf}(x, f) = \mathrm{E}[\widehat{Y}_0^{+T} \mid F_0 = f]x,$$

and the optimal portfolio at time t is

$$u_t^* = \arg\max_{u_t \in \overline{\mathcal{C}^+}} \min_{(\theta_t, V_t) \in \mathcal{A}_T^{\theta,V}|_t^t} \left\{ g_t^+ + \left(m_t^+ + g_t^+ \theta_t \right)' u_t - \frac{\widehat{\gamma}_t}{2} u_t' \left(A(F_t) + V_t \right) u_t \right\}.$$

The proof of Proposition 8 is in Section A.3. The t stage problem,

$$\max_{u_t \in \overline{\mathcal{C}^+}} \min_{(\theta_t, V_t) \in \mathcal{A}_T^{\theta,V}|_t^t} \left\{ g_t^+ + \left(m_t^+ + g_t^+ \theta_t \right)' u_t - \frac{\widehat{\gamma}_t}{2} u_t' \left(A(F_t) + V_t \right) u_t \right\},$$

can be rewritten as follows,

$$\max_{u_t \in \overline{\mathcal{C}^+}} \min_{(\theta_t, V_t) \in \mathcal{A}_T^{\theta,V}|_t^t} \left\{ g_t^+ + \left(m_t^+ + g_t^+ \theta_t \right)' u_t - \frac{\widehat{\gamma}_t}{2} u_t' \left(A(F_t) + V_t \right) u_t \right\} =$$

$$\max_{u_t \in \overline{\mathcal{C}^+}} \left\{ g_t^+ + (m_t^+)' u_t - \frac{\widehat{\gamma}_t}{2} u_t' \left(A(F_t) + \eta_t \|A(F_t)\| I_d \right) u_t - |g_t^+| \eta_t^\theta \sqrt{u_t' \left(A(F_t) + \eta_t \|A(F_t)\| I_d \right) u_t} \right\}.$$

Therefore, the solution to the above problem exists since the above objective is continuous in u_t.

3 Limiting Behaviors

In this section, we study limiting behaviors of optimal portfolios. We have postulated that the parameters $(\widehat{\gamma}_t)_{t=0}^{T-1}, (\eta_t^\theta)_{t=0}^{T-1}$, and $(\eta_t^V)_{t=0}^{T-1}$ are non-negative. We therefore consider the behaviors of the optimal portfolios when these parameters go to infinity.

We first consider the optimal portfolio with the risk-free asset. By Proposition 6, the optimal portfolio with the risk-free asset is

$$u_t^* = \frac{1}{\widehat{\gamma}_t} \left(1 - \frac{|g_t^e| \eta_t^\theta}{\sqrt{a_t^e}} \right)^+ \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d \right)^{-1} m_t^e,$$

for all $t \ge 0$. Then, m_t^e and g_t^e are determined independently of the values of $\widehat{\gamma}_t, \eta_t^\theta$ and η_t^V. a_t^e can be expressed as

$$a_t^e = (m_t^e)' \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d \right)^{-1} m_t^e$$

$$= \frac{1}{\eta_t^V} (m_t^e)' \left(\frac{1}{\eta_t^V} A(F_t) + \|A(F_t)\| I_d \right)^{-1} m_t^e,$$

13

therefore, $a_t^e \to 0$ as $\eta_t^V \to \infty$. This implies that for any fixed $\widehat{\gamma}_t$ and η_t^θ,

$$\lim_{\eta_t^V \to \infty} u_t^* = \mathbf{0}_d.$$

So, if the investor has absolutely no confidence in the estimated conditional variance, she quits investing in the risky assets. Similarly, if $\eta_t^\theta \to \infty$, then

$$\lim_{\eta_t^\theta \to \infty} \left(1 - \frac{|g_t^e| \eta_t^\theta}{\sqrt{a_t^e}} \right)^+ = 0$$

for any fixed η_t^V. Therefore, u_t^* converges to $\mathbf{0}_d$ when η_t^θ tends to infinity. These behaviors are natural since the investor with strong doubts of the correctnesses of the risky assets' parameters prefers investing in the risk-free asset which yields a deterministic return r_f. Thus, the effects of multiple priors are similar to the risk-averse behavior if the risk-free asset is present. However, with the absence of risk-free asset, limiting portfolios are two famous portfolios, namely, the global minimum-variance portfolio and the equally weighted portfolio.

Proposition 9

1. *With the risk-free asset.* If $\widehat{\gamma}_t, \eta_t^\theta$ or η_t^V tends to infinity, then the investor invests all wealth in the risk-free asset.

2. *Without the risk-free asset.*

   (a) If $\widehat{\gamma}_t$ or η_t^θ tends to infinity and if η_t^V is fixed, then the investor invests in the global minimum-variance portfolio under $\mathbb{P}^{\theta,V}$, that is

   $$\lim_{\eta_t^\theta \text{ or } \widehat{\gamma}_t \to \infty} u_t^* = \frac{1}{\mathbf{1}_d' \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d \right)^{-1} \mathbf{1}_d} \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d \right)^{-1} \mathbf{1}_d,$$

   when permitting short selling. When short selling is not allowed, then the optimal portfolio again converges to the global minimum-variance portfolio.

   $$\lim_{\eta_t^\theta \to \infty} u_t^* = u_t^{+GMV} := \arg \min_{u \in \overline{\mathcal{C}^+}} \left\{ u' \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d \right) u \right\}.$$

   (b) If η_t^V tends to infinity, then the investor invests in the equally weighted portfolio whether short selling is allowed or not, that is,

   $$\lim_{\eta_t^V \to \infty} u_t^* = \frac{1}{d} \mathbf{1}_d.$$

Proof. By Proposition 7, the optimal portfolio without the risk-free asset is

$$u_t^* = \frac{1}{\widehat{\gamma}_t + |g_t| \eta_t^\theta / \psi_t^*} \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d \right)^{-1} \left(m_t - \frac{c_t^{worf} - (\widehat{\gamma}_t + |g_t| \eta_t^\theta / \psi_t^*)}{b_t^{worf}} \mathbf{1}_d \right),$$

where

$$a_t^{worf} = m_t' \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d \right)^{-1} m_t, \quad c_t^{worf} = m_t' \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d \right)^{-1} \mathbf{1}_d,$$

$$b_t^{worf} = \mathbf{1}_d' \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d \right)^{-1} \mathbf{1}_d, \quad d_t^{worf} = a_t^{worf} b_t^{worf} - (c_t^{worf})^2,$$

14

and ψ_t^* is the unique solution of the following quartic equation,

$$b_t^{worf}(\psi_t^*)^2 = \frac{d_t^{worf}}{(\widehat{\gamma}_t + |g_t|\eta_t^\theta/\psi_t^*)^2} + 1.$$

Similar to the case with the risk-free asset, m_t and g_t are determined independently of the values of $\widehat{\gamma}_t, \eta_t^\theta$ and η_t^V, so we can regard them as constants. To study the limiting behavior of u_t^*, it is important to use the quartic equation for ψ_t^*.

The quartic equation can be transformed as follows:

$$\left(\widehat{\gamma}_t + \frac{|g_t|\eta_t^\theta}{\psi_t^*}\right)^2 = \frac{d_t^{worf}}{b_t^{worf}(\psi_t^*)^2 - 1}. \tag{3.1}$$

The benefit of the above expression is that it separates the terms depending on η_t^θ and η_t^V. In addition,

$$\left(\widehat{\gamma}_t + \frac{|g_t|\eta_t^\theta}{\psi_t^*}\right)^2 = \frac{d_t^{worf}}{b_t^{worf}(\psi_t^*)^2 - 1} \geq \widehat{\gamma}_t^2 > 0, \tag{3.2}$$

for all $\widehat{\gamma}_t, \eta_t^\theta$, and η_t^V. For any fixed $\widehat{\gamma}_t$ and η_t^V, two limits of the equation (3.1) as $\eta_t^\theta \to \infty$ can be expected. One is some positive constant $C > 0$. Then, we have

$$\lim_{\eta_t^\theta \to \infty} \left(\widehat{\gamma}_t + \frac{|g_t|\eta_t^\theta}{\psi_t^*}\right)^2 = \lim_{\eta_t^\theta \to \infty} \frac{d_t^{worf}}{b_t^{worf}(\psi_t^*)^2 - 1} = C.$$

The right equality follows that

$$\lim_{\eta_t^\theta \to \infty} \psi_t^* = C_1$$

for some constant $C_1 > 0$. However, this implies

$$\lim_{\eta_t^\theta \to \infty} \left(\widehat{\gamma}_t + \frac{|g_t|\eta_t^\theta}{\psi_t^*}\right)^2 = \infty.$$

Hence it is a contradiction. Therefore, the equation (3.1) diverges when η_t^θ tends to infinity. Indeed,

$$\lim_{\eta_t^\theta \to \infty} \frac{d_t^{worf}}{b_t^{worf}(\psi_t^*)^2 - 1} = \infty$$

implies that $(\psi_t^*)^2 \to 1/b_t^{worf}$. Then,

$$\lim_{\eta_t^\theta \to \infty} \left(\widehat{\gamma}_t + \frac{|g_t|\eta_t^\theta}{\psi_t^*}\right)^2 = \infty.$$

This is consistent. Therefore, the limiting portfolio with respect to η_t^θ is

$$\lim_{\eta_t^\theta \to \infty} u_t^* = \frac{1}{b_t^{worf}} \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d\right)^{-1} \mathbf{1}_d.$$

This limiting portfolio is the global minimum-variance portfolio under $\mathbb{P}^{\theta,V}$.

We consider the limiting behavior with respect to η_t^V. d_t^{worf} and b_t^{worf} can be expressed as

$$d_t^{worf} = \frac{1}{(\eta_t^V)^2} d_t^\eta, \quad \lim_{\eta_t^V \to \infty} d_t^\eta = d_0 > 0,$$

$$b_t^{worf} = \frac{1}{\eta_t^V} b_t^\eta, \quad \lim_{\eta_t^V \to \infty} b_t^\eta = b_0 > 0.$$

15

Then, the RHS in the equation (3.1) becomes

$$\frac{d_t^{worf}}{b_t^{worf}(\psi_t^*)^2 - 1} = \frac{d_t^\eta}{b_t^\eta \eta_t^V \left((\psi_t^*)^2 - \eta_t^V/b_t^\eta\right)}.$$

We first assume that the LHS in the equation (3.1) diverges, that is,

$$\lim_{\eta_t^V \to \infty} \left(\widehat{\gamma}_t + \frac{|g_t|\eta_t^\theta}{\psi_t^*}\right)^2 = \infty.$$

This implies that ψ_t^* converges to 0 when $\eta_t^V \to \infty$. However, for the RHS in the equation (3.1),

$$\lim_{\eta_t^V \to \infty} \frac{d_t^\eta}{b_t^\eta \eta_t^V \left((\psi_t^*)^2 - \eta_t^V/b_t^\eta\right)} = 0.$$

This is a contradiction. It follows that there exists some positive constant $C_2 > 0$ such that

$$\lim_{\eta_t^V \to \infty} \frac{d_t^\eta}{b_t^\eta \eta_t^V \left((\psi_t^*)^2 - \eta_t^V/b_t^\eta\right)} = C_2.$$

Then, $\eta_t^V \left((\psi_t^*)^2 - \eta_t^V/b_t^\eta\right) \to C_3$ for some positive constant $C_3 > 0$ as $\eta_t^V \to \infty$. This implies that ψ_t^* diverges. Then the LHS in the equation (3.1) also converges to $\widehat{\gamma}_t^2$, so it is consistent with our expectation.

By the definition, it holds that

$$\lim_{\eta_t^V \to \infty} \frac{c_t^{worf}}{b_t^{worf}} = C_4,$$

for some constant C_4. Furthermore, we have

$$\lim_{\eta_t^V \to \infty} \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d\right)^{-1} x = \mathbf{0}_d,$$

for any $x \in \mathbb{R}^d$. Moreover, $\widehat{\gamma}_t + |g_t|\eta_t^\theta/\psi_t^*$ converges to $\widehat{\gamma}_t$ as $\eta_t^V \to \infty$. Hence,

$$\lim_{\eta_t^V \to \infty} \frac{1}{\widehat{\gamma}_t + |g_t|\eta_t^\theta/\psi_t^*} \left(A(F_t) + \eta_t^V \|A(F_t)\| I_d\right)^{-1} \left(m_t - \frac{c_t^{worf}}{b_t^{worf}} \mathbf{1}_d\right) = \mathbf{0}_d.$$

On the other hand,

$$\lim_{\eta_t^V \to \infty} \frac{\left(A(F_t) + \eta_t^V \|A(F_t)\| I_d\right)^{-1} \mathbf{1}_d}{b_t^{worf}} = \lim_{\eta_t^V \to \infty} \frac{\left(A(F_t)/\eta_t^V + \|A(F_t)\| I_d\right)^{-1} \mathbf{1}_d}{\mathbf{1}_d' \left(A(F_t)/\eta_t^V + \|A(F_t)\| I_d\right)^{-1} \mathbf{1}_d} = \frac{1}{d} \mathbf{1}_d.$$

Finally, it holds that

$$\lim_{\eta_t^V \to \infty} u_t^* = \frac{1}{d} \mathbf{1}_d.$$

Hence, the limiting portfolio as $\eta_t^V \to \infty$ is the equally weighted portfolio.

In the case when short selling is not allowed, it holds that $u_t^* \to \mathbf{1}_d/d$ as $\eta_t^V \to \infty$ since $\mathbf{1}_d/d$ is an interior point of $\overline{\mathcal{C}^+}$. When $\eta_t^\theta \to \infty$,

$$\lim_{\eta_t^\theta \to \infty} u_t^* = u_t^{+GMV} = \arg\min_{u \in \mathcal{C}^+} \left\{u'\left(A(F_t) + \eta_t^V \|A(F_t)\| I_d\right)u\right\},$$

16

since the objective function is proportional to

$$\frac{1}{\eta_t^\theta}(m_t^+)'u_t - \frac{\widehat{\gamma}_t}{2\eta_t^\theta}u_t'\Big(A(F_t) + \eta_t^V\|A(F_t)\|I_d\Big)u_t - |g_t^+|\sqrt{u_t'\Big(A(F_t) + \eta_t^V\|A(F_t)\|I_d\Big)u_t}$$

and the effects of the first two terms become small when η_t^θ becomes large. $\qquad\qquad\square$

Interestingly, the optimal portfolio u_t^* also converges to the equally weighted portfolio when $\eta_t^V \to \infty$ and $\eta_t^\theta = 0$. This means that the investor chooses the equally weighted portfolio under the strong uncertainty for variance even if she believes that the errors of the estimated conditional expected returns do not exist. Therefore, we conclude that the uncertainty about variances has a stronger impact on the investor's portfolio selection than the uncertainty about means.

However, the above result slightly differs from natural intuition. The readers may think that the investor tends to choose the asset having the highest expected return if the degree of suspicion for variances η_t^V, is sufficiently large. This question can be resolved through the inequality constraint (2.3). If η_t^V is sufficiently large, then the error of variance V_t, is also large. This implies that

$$\theta_t'\Big(A(F_t) + V_t\Big)^{-1}\theta_t$$

is very close to 0. Then, the mean errors θ_t in the investor's belief can take an arbitrary value if $\eta^V \to \infty$. Therefore, the investor's worst scenario is that the means take large negative values which are the same across all risky assets. Hence, the investor chooses the equally weighted portfolio if $\eta^V \to \infty$.

Without the risk-free asset, the limiting case $\eta_t^\theta \to \infty$ is the same as the special case of the 2-norm-constrained optimal portfolio in DeMiguel et al. (2009a). However, our framework admits dynamic changes of covariance, so the conditional covariance matrix appears in the limiting portfolio, whereas the unconditional covariance matrix appears in DeMiguel et al. (2009a) since they consider the static optimization.

Unfortunately, without the risk-free asset, the limiting strategies are not applicable straightforwardly since the value functions diverge. These limiting portfolios are only admissible in the initial period of the investment horizon. However, if η_t^θ and η_t^V are sufficiently large, then the optimal portfolios are very similar to the limiting portfolios, so we can use the optimal portfolios with large η_t^θ and η_t^V as proxies of the limiting portfolios. In Section 4, we study the cases of large η_t^θ and η_t^V.

Remark 10 As mentioned below, in a limiting case, Pflug et al. (2012) reach essentially the same conclusion as us, although there are several differences. Pflug et al. (2012) show that as a model uncertainty measured by the Kantorovich metric increases, an optimal portfolio in a mean-standard-deviation optimization problem converges to the equally weighted portfolio. Specifically, Pflug et al. (2012) consider the following problem:[2]

$$\max_u \min_{\mathbb{Q}:\, d_2(\mathbb{P},\mathbb{Q})\leq\kappa}\left\{\mathrm{E}^{\mathbb{Q}}[X] - \frac{\gamma}{2}\sqrt{\mathrm{Var}^{\mathbb{Q}}(X)}\right\},$$

$$\text{subject to } X = R'u, \qquad\qquad\qquad (3.3)$$

$$1 = \mathbf{1}_d'u,$$

where d_2 is the Kantorovich metric with order 2, and R is a risky assets' return vector. γ is a non-negative trade-off parameter between returns and risks. $\mathrm{E}^{\mathbb{Q}}$ and $\mathrm{Var}^{\mathbb{Q}}$ are expectation and

---

[2]We modify the original problem in Pflug et al. (2012) for this to take the same form as in this paper. It can be easily seen that the problem (3.3) is equivalent to the original problem in Pflug et al. (2012).

variance operators under a probability measure $\mathbb{Q}$, respectively. κ is a non-negative constant that represents a degree of uncertainty. Pflug et al. (2012) show that the solution to the problem (3.3) converges to the equally weighted portfolio as $\kappa \to \infty$.

One of the differences between our model and the model in Pflug et al. (2012) is the measure of uncertainty. Pflug et al. (2012) adopts the Kantorovich metric, while it is a mathematically sophisticated concept of a distance among probability measures, it is usually hard to compute, and it is not usually used in practice. In contrast, our measures of uncertainty are a confidence interval around means and a relative error of covariances, which are widely used in practice and easy to compute.

Another difference between our model and the model in Pflug et al. (2012) is the objective function. Let us assume $T = 1$ in our model, so we consider a static problem. In order to apply the approach of Pflug et al. (2012) to our objective function, we need to replace the term of the standard deviation in (3.3) to the variance, that is

$$\max_{u} \min_{\mathbb{Q}:\, d_2(\mathbb{P}, \mathbb{Q}) \leq \kappa} \left\{ \mathrm{E}^{\mathbb{Q}}[X] - \frac{\gamma}{2} \mathrm{Var}^{\mathbb{Q}}(X) \right\},$$
$$\text{subject to } X = R'u, \tag{3.4}$$
$$1 = \mathbf{1}_d' u.$$

However, since the objective function in the problem (3.4) does not satisfy the assumptions in Proposition 1 in Pflug et al. (2012), we can not use the results of Pflug et al. (2012). Therefore, the solution to the problem (3.4) may not converge to the equally weighted portfolio. For details about this discussion, we also refer to Wozabal (2014).

Considering a unique prior case (i.e., $\kappa = 0$), we can easily show that the solution to the problem (3.3) is proportional to the solution to the problem (3.4). However, the trade-off parameter γ has different implications in these problems. As seen in Section 2, in our model, we can regard γ as a coefficient of (absolute) risk aversion. On the other hand, economic implication of γ in the problem (3.3) is not clear.

Finally, we adopt dynamic mean-variance optimization, whereas the model in Pflug et al. (2012) is static. Therefore, our model allows an investment strategy that changes portfolios in reaction to the state variable F_t. In Section 4, we will see that this investment strategy often performs well in practice.

4 Optimality of Naive Investment Strategies in Back Tests

In this section, we conduct back tests of dynamic optimization with multiple priors. We use the following three data sets.

1. **International Equity Indexes**. The four MSCI indexes: US, Japan, UK, and Germany. Each index is month-end US-dollar valued. The data source is the Thomson Reuters Datastream.

2. **Industry Indexes**. The five monthly returns of industry indexes in the US stock market: Consumer, Manufacture, HiTechnology, Health, and Others. The data source is the Kenneth French's web site.

3. **Size- and Value-Sorted Portfolios**. The six monthly returns of the 2×3 size- and book-to-market-sorted portfolios in the US stock market by Fama and French (1993). The data source is the Kenneth French's web site.

We also use monthly returns of 90-day US treasury bill from the Thomson Reuters Datastream as risk-free rates. Each dataset consists of monthly returns from January 1975 to December 2014. In order to focus on the diversification within the risky assets, we consider the asset allocations without the risk-free asset and compare out-of-sample Sharpe ratios.

4.1 Methodology of the Back Tests

In each dataset, we assume that the return vector is modeled by a two-state, recursive and time-homogeneous Markov regime-switching model, such that

$$R_{t+1} = \mu(F_{t+1}) + \sigma(F_{t+1})\epsilon_{t+1}, \quad t \geq 0,$$

where $(F_t)_{t=0}^T$ is a two-state, recursive and time-homogeneous Markov chain. $(F_t)_{t=0}^T$ and $(\epsilon_t)_{t=1}^T$ are mutually independent. The transition probability matrix of $(F_t)_{t=0}^T$ is constant over time by the time-homogeneity of $(F_t)_{t=0}^T$. This model satisfies Assumption 1.

We assume that one of the states of F_t is a "good state" and another state is a "bad state". Whether the market condition is good or bad is determined by the values of conditional means and variances of the return vectors. Consider the state in which more than half of means and reciprocals of variances are larger than those in another state. We call this state the "good state" and call the other state the "bad state". If these values are the same, as the good state we choose the state in which the sum of the marginal conditional means is larger than that in the other state.

We consider two types of investors. The first type assumes that the returns of the indexes are *not* regime switching. This type of investor always considers the mean and variance of the returns are constants. We call this type "IID". The second type assumes that the returns are driven by the Markov chain $(F_t)_{t=0}^T$. Hence, this type of investor assumes that the conditional mean and variance of the returns vary over the time. We call this type "RS". The IID and RS investors compute their portfolios using the dynamic optimization without the risk-free asset proposed in Propositions 7 and 8.

As for the preference parameters, $\widehat{\gamma}_t$ and δ are constant over time and states. We set $\widehat{\gamma}_t = 1$ and $\delta = 0.99$. To study the effect of the multiple priors, we consider six different values of ($\eta_t^\theta = 0, 1, 2, 3, 4, 5$) and five different values of ($\eta_t^V = 0.0, 0.5, 1.0, 3.0, 5.0$). These parameters are fixed over time and states. Furthermore, we also consider two different cases: the one in which short selling is permitted and the one in which it is not permitted. Therefore, we compute $6 \times 5 \times 2$ portfolios for each investor type. The large η_t^θ and η_t^V represent strong degrees of suspicion of investors, so we can compare the investors with different levels of suspicion. As benchmarks, we also consider the equally weighted portfolio (EW), the portfolio maximizing the single-period empirical Sharpe ratio (max SR), and the single-period global minimum-variance portfolio (GMV). In addition, we use the value-weighted portfolio (VW) as one of the benchmarks for the industry indexes and the size- and value-sorted portfolios data set[3]. Note that the IID investor's max SR at time t maximizes the following objective function:

$$f_{\text{IID}}(u) := \frac{\overline{\mu}'u - r_{\text{f},t}}{\sqrt{u'\overline{\Sigma}u}},$$

where $\overline{\mu}$ and $\overline{\Sigma}$ are sample mean and variance of returns and $r_{\text{f},t}$ is the risk-free rate at time t. On the other hand, the RS investor's max SR at time t with the state $F_t = i$ maximizes the following objective function,

$$f_{\text{RS}}(u) := \frac{\overline{\mu}_i'u - r_{\text{f},t}}{\sqrt{u'(\overline{\sigma}_i\overline{\sigma}_i')u}},$$

---

[3]Unfortunately, we can not obtain the data of the market values of the US dollar-based MSCI indexes. So, we compute the value-weighted portfolio only for the abovementioned two data sets

where $\overline{\mu}_i$ and $\overline{\sigma}_i\overline{\sigma}_i'$ are estimated mean and variance of returns at state i. By the above definition, max SR delivers the largest Sharpe ratio without estimation errors.

Furthermore, we consider the following six extreme strategies. In these six strategies, η_t^θ and η_t^V take different values in the different states of F_t. The first strategy is "GMV in the good state and EW in the bad state" (GMV-EW). Under this strategy, the investor invests in the global minimum-variance portfolio in the state of good market condition and in the equally weighted portfolio in the state of bad market condition. This strategy can be interpreted as the optimal portfolio when $\eta_t^V = 0$ and η_t^θ tends to infinity in the good state and when η_t^V tends to infinity in the bad state. The second strategy is "EW in the good state and No Error in the bad state" (EW-NE). Under this strategy, the investor invests in the equally weighted portfolio in the good state and in the optimal portfolio of Proposition 7 and 8 with $\eta_t^\theta = \eta_t^V = 0$ in the bad state. This strategy can be interpreted as the optimal portfolio when $\eta_t^V = 0$ and η_t^θ tends to infinity in the good state and when η_t^θ and η_t^V are 0 in the bad state. The third strategy is "GMV in the good state and No Error in the bad state" (GMV-NE). Under this strategy, the investor invests in the global minimum-variance portfolio in the good state and in the optimal portfolio with $\eta_t^\theta = \eta_t^V = 0$ in the bad state. This strategy can be interpreted as the optimal portfolio when $\eta_t^V = 0$ and η_t^θ tends to infinity in the good state and when η_t^θ and η_t^V are 0 in the bad state. The rest of the strategies are the reverse strategies of the above three strategies "EW in the good state and GMV in the bad state" (EW-GMV), "No Error in the good state and EW in the bad state" (NE-EW), and "No Error in the good state and GMV in the bad state" (NE-GMV). The abbreviations and the six extreme strategies are summarized in Table 1.

In the back tests, each investor needs to estimate the distribution parameters from the data. The rolling window of estimation is fixed to 240 months. At each time, the IID investor computes the sample mean and variance of the data over the past 240 months from current time. Similarly, the RS investor estimates the distribution parameters using the data over the past 240 months from current time. The RS investor estimates the parameters by the EM algorithm proposed by Hamilton (1990). At the time when t months have passed after the start of investment, the investors compute the optimal plans of the portfolios for $240 - t$ periods and invest in their optimal portfolios in the initial period. For example, suppose that the investors would like to decide the portfolios in January 2000. They first compute the optimal portfolio plans for 180 months, from January 2000 to December 2014. Then, they invest in the portfolios in the initial period in January 2000.

However, the RS investors can not determine their portfolios since the state variable F in actual data is not observable. To determine their portfolios, they regard the state having the highest conditional probability as the current state, that is, the current state at time t is the following i_t,

$$i_t = \arg\max_{i \in \{1,2\}} \mathbb{P}(F_t \text{ is in the } i \text{ th state.} \mid R_1, R_2, \ldots, R_t),$$

where R_s is the return vector of the indexes at time $s = 1, \ldots, t$.

The optimality of the portfolios derived in Proposition 7 and 8 are not guaranteed in the above rolling-window approach. However, in general, investors usually choose their portfolios based on the latest information. Moreover, the existing literature (e.g., Garlappi et al. (2007) and DeMiguel et al. (2009b)) adopts the rolling-window approach. For these reasons, we also use this approach.

4.2 Results of the Back Tests

Tables 2, 3, and 4 display the out-of-sample Sharpe ratios obtained from the back tests. In all data sets, the extreme strategies tend to deliver larger Sharpe ratios.

In the international diversification (Table 2), max SR has larger Sharpe ratios than the other typical portfolios, GMV and EW, in all cases (whether the investor type is IID or RS and with or without permitting short selling). When permitting short selling, max SR of the IID investor has the largest Sharpe ratio (0.1014) except for the extreme strategies. However, the IID investor's optimal portfolio with $\eta_t^\theta = 0.0$ and $\eta_t^V = 5.0$ when not permitting short selling delivers the largest Sharpe ratio (0.1081) among all portfolios. Also, the Sharpe ratios when short selling is not permitted tend to be larger than those when it is permitted. This is consistent with the well-known results of Jagannathan and Ma (2003), which states that a short-selling constraint can improve the investment performance.

According to Table 2, the two extreme strategies, GMV-EW and GMV-NE, have good performances whether short selling is permitted or not. GMV-EW with short selling delivers the largest Sharpe ratio (0.1032) among the portfolios with short selling. On the other hand, GMV-NE without short selling has the second largest Sharpe ration (0.1078) among all portfolios. Furthermore, the Sharpe ratios of GMV-EW without short selling and GMV-NE with short selling, 0.0922 and 0.0891 respectively, are not small. Among the typical portfolios (max SR, GMV, and EW) and the typical extreme strategies (GMV-EW and EW-GMV), GMV-EW when permitting short selling has the largest Sharpe ratio. Taking into account the simplicity of computation, GMV-EW when permitting short selling works effectively.

Table 3 reports the results of the industry indexes data set. According to Table 3, GMV has larger Sharpe ratio than max SR, EW and VW in all cases. Precisely, the IID investor's GMV with short selling has the largest Sharpe ratio among all portfolios. Among portfolios without short selling, the RS investor's GMV has the largest Sharpe ratio. However, the two extreme strategies, GMV-EW and EW-GMV, are not so bad either. Both of GMV-EW and EW-GMV defeat max SR, EW, and VW regardless of whether short selling is permitted or not. Comparing GMV-EW and EW-GMV, EW-GMV performs better than GMV-EW. When short selling is permitted, the Sharpe ratio of GMV-EW is 0.2234, whereas the Sharpe ratio of EW-GMV is 0.2564. Without short selling, the Sharpe ratio of GMV-EW is 0.2281, whereas the Sharpe ratio of EW-GMV is 0.2473.

Table 4 reports the results of the size- and value-sorted portfolios data set. The IID investor's optimal portfolio with $\eta_t^\theta = 1.0$ and $\eta^V = 0.0$ when permitting short selling has the largest Sharpe ratio (0.3406) among all portfolios. On the other hand, similar to the industry indexes data set, GMV has a larger Sharpe ratio than max SR, EW, and VW in all cases. Furthermore, regardless of whether permitting short selling or not, GMV-EW defeats max SR, EW, VW, and GMV, except for the IID investor's GMV. Precisely, GMV-EW has the largest Sharpe ratio (0.2336) among portfolios without short selling.

In summary, GMV-EW and GMV tend to perform well in all data sets. DeMiguel et al. (2009b) report relatively good performances of GMV, so our results are consistent with their results. Unlike DeMiguel et al. (2009b), EW is less efficient in our data sets. However, GMV-EW delivers larger Sharpe ratio in the international indexes data set when short selling is permitted and in the size- and value-sorted portfolios data set when short selling is not permitted. Furthermore, GMV-EW defeats max SR in the industry indexes data set regardless of whether permitting short selling or not. Therefore, EW is efficient under particular situations and GMV-EW effectively uses the efficiency of EW.

Note that the IID investor's max SRs deliver larger Sharpe ratios than the RS investor's max SRs in all cases. One explanation of this result is the difference between the objective functions. The RS investor's max SR maximizes the conditional Sharpe ratio whereas the IID investor's max SR maximizes the unconditional Sharpe ratio. Since our back tests compute out-of-sample Shape ratios, which are sample analogs of the unconditional Sharpe ratio, it is not surprising that the IID investor's max SRs perform better than the RS investor's max SRs.

As seen in Section 3, the extreme strategies can not be justified by our framework since the value functions also diverge. However, optimal portfolios with sufficiently large η_t^θ and η_t^V will work as extreme strategies. We show behaviors of the optimal portfolios with sufficiently large η_t^θ and η_t^V.

We consider the following metrics from the optimal portfolios of EW and GMV:

$$\|u_t^*\|_{GMV} := \|u_t^* - u_t^{GMV}\|, \quad \|u_t^*\|_{EW} := \|u_t^* - \mathbf{1}_d/d\|,$$

where u_t^* is the optimal portfolio at time t. These metrics represent the root square errors from EW and GMV, so we call these the portfolio errors from EW and GMV. Figure 1 shows the portfolio errors of the international indexes data set from EW and GMV in the case when short selling is permitted. From the upper figures in Figure 1, we see that the high η_t^θ reduces the portfolio errors from GMV for both IID and RS investors. Similarly, the lower figures in Figure 1 show that the high η_t^V reduces the portfolio errors from EW for the both investors[4]. This observation is consistent with our theoretical results in Section 3.

Figure 1 suggests that the optimal portfolios with sufficiently large η_t^θ and η_t^V can be good proxies for the extreme strategies. We try to approximate GMV-EW and EW-GMV by these optimal portfolios. We denote these approximated portfolios by mGMV-EW and mEW-GMV. The first letter m means *mimicking* extreme strategy.

Table 5 reports the Sharpe ratios of the optimal portfolios with sufficiently large η_t^θ and η_t^V. The second, third, fourth, and fifth columns are the specific values of η_t^θ and η_t^V of the mimicking strategies. In Table 5, the averages of the root square errors of the mimicking strategies from the (original) extreme strategies are sufficiently small in all the data sets. This implies that the mimicking strategies are good proxies for the extreme strategies. Consequently, the mimicking strategies' Sharpe ratios are close to the original extreme strategies' Sharpe ratios. Hence, we conclude that the investors can asymptotically justify the extreme strategies and exploit the portfolio selections that are extremely robust to the estimation errors.

GMV-EW and GMV-EW are very simple strategies, which means investing in the global minimum-variance portfolio or the equally weighted portfolio with respect to the market condition. These strategies seem to be naive and ad hoc, but they are one of the results of the plausible portfolio optimization; For GMV-EW, the investor assumes that the estimates of the conditional mean in economic booming and the conditional variance in recession are not credible at all, whereas for EW-GMV, she assumes that the estimated variances in economic booming and the estimated means in recession are not credible. Therefore, the investor chooses GMV-EW or EW-GMV as the extremely robust portfolio to estimation errors.

5 Conclusion

In this paper, we derive the optimal portfolios in the dynamic mean-variance problems with multiple priors. Furthermore, we show that the optimal portfolios include the equally weighted portfolio and the global minimum-variance portfolio in the limiting cases.

In the back tests, we find that the extreme strategies, especially GMV-EW tends to be relatively mean-variance efficient in the various data sets. In addition, the portfolios with sufficiently large η_t^θ and η_t^V can be good proxies for the extreme strategies. Therefore, we can mathematically justify the extreme strategies although the extreme strategies seem to be naive; the reason why the investors choose the extreme strategies is that the investors have strong suspicion of the estimated expected return and covariance.

---

[4]In the lower figures in Figure 1, we fix $\eta_t^\theta = 5$. Since the portfolio errors from EW with $\eta_t^\theta = 1$ is larger than those with $\eta_t^\theta = 5$, we do not report the results of $\eta_t^\theta = 1$

Our analysis has an important implication about asset pricing. As seen above, the standard mean-variance analysis does not work in practice, neither does the capital asset pricing model. However, according to our framework, two mean-variance inefficient in theory, and naive portfolios, the global minimum portfolio and the equally weighted portfolio, can be justified as rational selections of the investors. Thus, it is possible that the two portfolios develop asset pricing models. Then, our framework can be used to analyze the rational reasoning behind the investors' choices of these portfolios.

A   Appendix

A.1   Proof of Lemma 4

Proof of Lemma 4. We first prove that

$$
\mathrm{E}\left[\exp\{u'R_{t+1}\}\zeta_{t+1}|\mathcal{F}_t^{R,F}, F_{t+1}\right] = \exp\left\{u'\Big(\mu(F_{t+1}) + \theta_t\Big) + \frac{1}{2}u'(\sigma(F_{t+1})(\sigma(F_{t+1}))' + V_t)u\right\}
$$
(A.1)

for all $t \geq 0$ and $u \in \mathbb{R}^d$. To simplify the notations, we write $\mu_{t+1} = \mu(F_{t+1})$ and $\sigma_{t+1} = \sigma(F_{t+1})$. For any fixed constant vector $u \in \mathbb{R}^d$, we have

$$
\mathrm{E}\left[\exp\{u'R_{t+1}\}\zeta_{t+1}|\mathcal{F}_t^{R,F}, F_{t+1}\right]
$$

$$
= \int_{\mathbb{R}^d} \frac{1}{(2\pi)^{d/2}} \frac{1}{\sqrt{\det(\sigma_{t+1}\sigma'_{t+1} + V_t)}} \exp\left\{-\frac{1}{2}\theta'_t\Big(\sigma_{t+1}\sigma'_{t+1} + V_t\Big)^{-1}\theta_t + u'\mu_{t+1}\right.
$$

$$
+ \Big(\big(\sigma_{t+1}\sigma'_{t+1} + V_t\big)^{-1}\theta_t + u\Big)'\sigma_{t+1}\epsilon_{t+1} + \frac{1}{2}\epsilon'_{t+1}\Big(I_d - \sigma'_{t+1}\big(\sigma_{t+1}\sigma'_{t+1} + V_t\big)^{-1}\sigma_{t+1}\Big)\epsilon_{t+1}
$$

$$
\left. -\frac{1}{2}\epsilon'_{t+1}\epsilon_{t+1}\right\} \mathrm{d}\epsilon_{t+1}
$$

$$
= \exp\left\{u'(\mu_{t+1} + \theta_t) + \frac{1}{2}u'\Big(\sigma_{t+1}\sigma'_{t+1} + V_t\Big)u\right\}
$$

$$
\times \int_{\mathbb{R}^d} \frac{1}{(2\pi)^{d/2}} \frac{1}{\sqrt{\det(\sigma_{t+1}\sigma'_{t+1} + V_t)}}
$$

$$
\times \exp\left\{-\frac{1}{2}\Big(\sigma_{t+1}\epsilon_{t+1} - mean_{t+1}\Big)'\Big(\sigma_{t+1}\sigma'_{t+1} + V_t\Big)^{-1}\Big(\sigma_{t+1}\epsilon_{t+1} - mean_{t+1}\Big)\right\} \mathrm{d}\epsilon_{t+1},
$$

where $mean_{t+1} = \theta_t + \Big(\sigma_{t+1}\sigma'_{t+1} + V_t\Big)u$. Then, the above integral is equal to 1 and the equation (A.1) holds.

If $u = \mathbf{0}_d$, we have

$$
\mathrm{E}\left[\zeta_{t+1}|\mathcal{F}_t^{R,F}, F_{t+1}\right] = \mathrm{E}\left[\exp\{\mathbf{0}'_d R_{t+1}\}\zeta_{t+1}|\mathcal{F}_t^{R,F}, F_{t+1}\right] = 1.
$$

Then, the process $(Z_t)_{t=0}^T$ is a martingale with respect to $\mathbb{F}^{R,F}$ since

$$
\mathrm{E}\left[Z_{t+1}\Big|\mathcal{F}_t^{R,F}\right] = \mathrm{E}\left[Z_t \mathrm{E}\left[\zeta_{t+1}\Big|\mathcal{F}_t^{R,F}, F_{t+1}\right]\Big|\mathcal{F}_t^{R,F}\right] = Z_t,
$$

for all $t \geq 0$. Moreover, $\mathrm{E}[Z_t] = \mathrm{E}[Z_0] = 1$. These equalities imply that we can define the probability measure $\mathbb{P}^{\theta,V}$ such that

$$
\mathbb{P}^{\theta,V}(A) = \mathrm{E}[\mathbb{1}_A Z_T], \quad A \in \mathcal{F}_T^{R,F}.
$$

Let $\sqrt{-1}$ be an imaginary unit. For any fixed $u \in \mathbb{R}^K$, by the Bayes rule, it holds that

$$\mathrm{E}^{\theta,V}\left[\exp\{\sqrt{-1}u'F_{t+1}\}\Big|\mathcal{F}_t^{R,F}\right] = \mathrm{E}\left[\exp\{\sqrt{-1}u'F_{t+1}\}\zeta_{t+1}\Big|\mathcal{F}_t^{R,F}\right]$$

$$= \mathrm{E}\left[\exp\{\sqrt{-1}u'F_{t+1}\}\,\mathrm{E}\left[\zeta_{t+1}\Big|\mathcal{F}_t^{R,F}, F_{t+1}\right]\Big|\mathcal{F}_t^{R,F}\right]$$

$$= \mathrm{E}\left[\exp\{\sqrt{-1}u'F_{t+1}\}\Big|\mathcal{F}_t^{R,F}\right].$$

This implies that the conditional distribution of F_{t+1} given $\mathcal{F}_t^{R,F}$ under $\mathbb{P}^{\theta,V}$ is the same as the conditional distribution of F_{t+1} given $\mathcal{F}_t^{R,F}$ under $\mathbb{P}$. It follows that

$$\mathrm{E}^{\theta,V}\left[R_{t+1}\Big|\mathcal{F}_t^{R,F}\right] = \mathrm{E}^{\theta,V}\left[\mathrm{E}^{\theta,V}\left[R_{t+1}\Big|\mathcal{F}_t^{R,F}, F_{t+1}\right]\Big|\mathcal{F}_t^{R,F}\right]$$

$$= \mathrm{E}^{\theta,V}\left[\mu(F_{t+1})\Big|\mathcal{F}_t^{R,F}\right] + \theta_t = \mathrm{E}\left[\mu(F_{t+1})\Big|\mathcal{F}_t^{R,F}\right] + \theta_t$$

$$= \mathrm{E}\left[R_{t+1}\Big|\mathcal{F}_t^{R,F}\right] + \theta_t,$$

for all $t \geq 0$. Similarly, it holds that

$$\mathrm{Var}^{\theta,V}(R_{t+1}|\mathcal{F}_t^{R,F}) = \mathrm{Var}^{\theta,V}(\mathrm{E}^{\theta,V}[R_{t+1}|\mathcal{F}_t^{R,F}, F_{t+1}]|\mathcal{F}_t^{R,F})$$

$$+ \mathrm{E}^{\theta,V}[\mathrm{Var}^{\theta,V}(R_{t+1}|\mathcal{F}_t^{R,F}, F_{t+1})|\mathcal{F}_t^{R,F}]$$

$$= \mathrm{Var}^{\theta,V}(\mu(F_{t+1})|\mathcal{F}_t^{R,F}) + \mathrm{E}^{\theta,V}[\sigma(F_{t+1})(\sigma(F_{t+1}))'|\mathcal{F}_t^{R,F}] + V_t$$

$$= \mathrm{Var}(\mu(F_{t+1})|\mathcal{F}_t^{R,F}) + \mathrm{E}[\sigma(F_{t+1})(\sigma(F_{t+1}))'|\mathcal{F}_t^{R,F}] + V_t$$

$$= \mathrm{Var}(R_{t+1}|\mathcal{F}_t^{R,F}) + V_t,$$

for all $t \geq 0$. $\qquad\square$

A.2 Proofs of Theorem 5 and Proposition 6

To prove Theorem 5, we need to show two lemmas. The objective function in the problem (2.2) can be written as

$$\mathrm{E}^{\theta,V}\left[\sum_{t=0}^{T-1}\delta^t X_t c_t(F_t, \theta_t, V_t, u_t)\right] = \mathrm{E}\left[\sum_{t=0}^{T-1}\delta^t X_t c_t(F_t, \theta_t, V_t, u_t)Z_t\right]$$

$$= \mathrm{E}\left[\sum_{t=0}^{T-1}\delta^t X_t c_t(F_t, \theta_t, V_t, u_t)\prod_{s=0}^{t}\zeta_s\right].$$

Let ρ_t be a mapping such that

$$\rho_t(W) = \delta\,\mathrm{E}\left[W\zeta_{t+1}\Big|\mathcal{F}_t^{R,S}\right], \quad 0 \leq t \leq T-1.$$

Then, we can express the objective as

$$J_T(x, f, \theta, V, u) := \mathrm{E}^{\theta,V}\left[\sum_{t=0}^{T-1}\delta^t X_t c_t(F_t, \theta_t, V_t, u_t)\right]$$

$$= xc_0(f, \theta_0, V_0, u_0) + \rho_0\Big(X_1 c_1(F_1, \theta_1, V_1, u_1) + \rho_1\Big(\cdots$$

$$+ \rho_{T-2}\Big(X_{T-1}c_{T-1}(F_{T-1}, \theta_{T-1}, V_{T-1}, u_{T-1})\Big)\cdots\Big)\Big),$$

for all $u \in \mathcal{A}_T^u$ and $(\theta, V) = (\theta_t, V_t)_{t=0}^{T-1} \in \mathcal{A}_T^{\theta,V}$, where $X_0 = x$ and $F_0 = f$. The above recursive structure is important for the optimization. To simplify the notations, we write

$$c_t^{u,\theta,V} = X_t c_t(F_t, \theta_t, V_t, u_t), \quad 0 \le t \le T-1.$$

We define $\rho_{t,r-1}$ as

$$\rho_{t,r-1}(W_t, \dots, W_r) = W_t + \rho_t \Big(W_{t+1} + \rho_{t+1} \big(\cdots + \rho_{r-1}(W_r) \big) \Big).$$

Then, for all $u \in \mathcal{A}_T^u$ and $(\theta, V) \in \mathcal{A}^{\theta,V}$, we have

$$J_T(x, f, \theta, V, u) = \rho_{0,t-1} \Big(c_0^{u,\theta,V}, \dots, c_{t-1}^{u,\theta,V}, \rho_{t,T-2} \big(c_t^{u,\theta,V}, \dots, c_{t-2}^{u,\theta,V}, c_{T-1}^{u,\theta,V} \big) \Big),$$

for all $0 \le t \le T-1$. We first prove that the order of the minimization with respect to (θ, V) can be exchangeable. The proof of the following lemma is essentially the same as the proof of Lemma 1 in Chen et al. (2014).

Lemma 11 Fix any $u \in \mathcal{A}_T^u$. Then, it holds that

$$\min_{(\theta_r, V_r)_{r=t}^{T-1} \in \mathcal{A}_T^{\theta,V}|_t^{T-1}} \rho_{t,T-1}(c_t^{u,\theta,V}, \dots, c_{T-1}^{u,\theta,V})$$
$$= \min_{(\theta_r, V_r)_{r=t}^{s-1} \in \mathcal{A}_T^{\theta,V}|_t^{s-1}} \rho_{t,s-1} \Big(c_0^{u,\theta,V}, \dots, c_{s-1}^{u,\theta,V}, \min_{(\theta_r, V_r)_{r=t}^{s-1} \in \mathcal{A}_T^{\theta,V}|_s^{T-1}} \rho_{s,T-2} \big(c_s^{u,\theta,V}, \dots, c_{T-1}^{u,\theta,V} \big) \Big),$$

$$\text{(A.2)}$$

for all $0 \le t \le s \le T-1$, where $\mathcal{A}_T^{\theta,V}|_t^r$ is a set of $(\theta_s, V_s)_{s=t}^r$ which are components of $(\theta, V) \in \mathcal{A}_T^{\theta,V}$ from the time t to r.

Proof. Fix any $u \in \mathcal{A}_T^u, x \in \mathbb{R}$ and $f \in \mathbb{R}^K$. Fix any $0 \le t \le s \le T-1$. We define $(\widetilde{\theta}, \widetilde{V})$ as

$$(\widetilde{\theta}, \widetilde{V}) = (\widetilde{\theta}_r, \widetilde{V}_r)_{r=s}^{T-1} \in \arg\min_{(\theta_r, V_r)_{r=s}^{T-1} \in \mathcal{A}_T^{\theta,V}|_s^{T-1}} \rho_{s,T-2} \Big(c_s^{u,\theta,V}, \dots, c_{T-2}^{u,\theta,V}, c_{T-1}^{u,\theta,V} \Big).$$

Then, for every $(\theta, V) \in \mathcal{A}_T^{\theta,V}$, it holds that

$$\rho_{s,T-2} \Big(c_s^{u,\theta,V}, \dots, c_{T-2}^{u,\theta,V}, c_{T-1}^{u,\theta,V} \Big) \ge \rho_{s,T-2} \Big(c_s^{u,\widetilde{\theta},\widetilde{V}}, \dots, c_{T-2}^{u,\widetilde{\theta},\widetilde{V}}, c_{T-1}^{u,\widetilde{\theta},\widetilde{V}} \Big).$$

Since $\rho_r(W) \le \rho_r(W')$ for all r if $W \le W'$, it holds that

$$\rho_{t,s-1} \Big(W_t, \dots, W_{s-1}, W_s \Big) \le \rho_{t,s-1} \Big(W_t, \dots, W_{s-1}, W' \Big),$$

if $W_s \le W'$. Therefore, we have

$$\rho_{t,T-1}(c_t^{u,\theta,V}, \dots, c_{T-1}^{u,\theta,V}) \ge \rho_{t,s-1} \Big(c_t^{u,\theta,V}, \dots, c_{s-1}^{u,\theta,V}, \rho_{s,T-2} \big(c_s^{u,\widetilde{\theta},\widetilde{V}}, \dots, c_{T-2}^{u,\widetilde{\theta},\widetilde{V}}, c_{T-1}^{u,\widetilde{\theta},\widetilde{V}} \big) \Big).$$

Minimizing the above inequality, we obtain

$$\min_{(\theta_r, V_r)_{r=t}^{T-1} \in \mathcal{A}_T^{\theta,V}|_t^{T-1}} \rho_{t,T-1}(c_t^{u,\theta,V}, \dots, c_{T-1}^{u,\theta,V})$$
$$\ge \min_{(\theta_r, V_r)_{r=t}^{s-1} \in \mathcal{A}_T^{\theta,V}|_t^{s-1}} \rho_{t,s-1} \Big(c_t^{u,\theta,V}, \dots, c_{s-1}^{u,\theta,V}, \rho_{s,T-2} \big(c_s^{u,\widetilde{\theta},\widetilde{V}}, \dots, c_{T-2}^{u,\widetilde{\theta},\widetilde{V}}, c_{T-1}^{u,\widetilde{\theta},\widetilde{V}} \big) \Big)$$
$$= \min_{(\theta_r, V_r)_{r=t}^{s-1} \in \mathcal{A}_T^{\theta,V}|_t^{s-1}} \rho_{t,s-1} \Big(c_t^{u,\theta,V}, \dots, c_{s-1}^{u,\theta,V}, \min_{(\theta_r, V_r)_{r=s}^{T-1} \in \mathcal{A}_T^{\theta,V}|_s^{T-1}} \rho_{s,T-2} \big(c_s^{u,\theta,V}, \dots, c_{T-1}^{u,\theta,V} \big) \Big).$$

We denote by $(\widehat{\theta}, \widehat{V})$ the following minimizer:

$$(\widehat{\theta}, \widehat{V}) := (\widehat{\theta}_r, \widehat{V}_r)_{r=t}^{s-1} \in \arg \min_{(\theta_r, V_r)_{r=t}^{s-1} \in \mathcal{A}_T^{\theta,V}|_t^{s-1}} \rho_{t,s-1}\Big(c_t^{u,\theta,V}, \dots, c_{s-1}^{u,\theta,V}, \rho_{s,T-2}\big(c_s^{u,\widetilde{\theta},\widetilde{V}}, \dots, c_{T-1}^{u,\widetilde{\theta},\widetilde{V}}\big)\Big).$$

Then, $(\overline{\theta}, \overline{V}) := ((\widehat{\theta}_t \widehat{V}_t), \dots, (\widehat{\theta}_{s-1}, \widehat{V}_{s-1}), (\widetilde{\theta}_s, \widetilde{V}_s), \dots, (\widetilde{\theta}_{T-1}, \widetilde{V}_{T-1}))$ is in $\mathcal{A}_T^{\theta,V}|_t^{T-1}$. It follows that

$$\min_{(\theta_r, V_r)_{r=t}^{T-1} \in \mathcal{A}_T^{\theta,V}|_t^{T-1}} \rho_{t,T-1}(c_t^{u,\theta,V}, \dots, c_{T-1}^{u,\theta,V})$$

$$\leq \rho_{t,T-1}(c_t^{u,\overline{\theta},\overline{V}}, \dots, c_{T-1}^{u,\overline{\theta},\overline{V}}) = \rho_{t,s-1}\Big(c_t^{u,\widehat{\theta},\widehat{V}}, \dots, c_{s-1}^{u,\widehat{\theta},\widehat{V}}, \rho_{s,T-2}\big(c_s^{u,\widetilde{\theta},\widetilde{V}}, \dots, c_{T-2}^{u,\widetilde{\theta},\widetilde{V}}, c_{T-1}^{u,\widetilde{\theta},\widetilde{V}}\big)\Big)$$

$$= \min_{(\theta_r, V_r)_{r=t}^{s-1} \in \mathcal{A}_T^{\theta,V}|_t^{s-1}} \rho_{t,s-1}\Big(c_t^{u,\theta,V}, \dots, c_{s-1}^{u,\theta,V}, \min_{(\theta_r, V_r)_{r=s}^{T-1} \in \mathcal{A}_T^{\theta,V}|_s^{T-1}} \rho_{s,T-2}\big(c_s^{u,\theta,V}, \dots, c_{T-1}^{u,\theta,V}\big)\Big).$$

Therefore, we can conclude that the equality (A.2) holds. □

By Lemma 11, we can minimize the objective function with respect to (θ, V), iteratively. Let

$$\overline{\rho}_{T-1}(u_{T-1}) = \min_{(\theta_{T-1}, V_{T-1}) \in \mathcal{A}_T^{\theta,V}|_{T-1}^{T-1}} c_{T-1}^{u,\theta,V},$$

$$\overline{\rho}_t(W, u_t) = \min_{(\theta_t, V_t) \in \mathcal{A}_T^{\theta,V}|_t^t} \left\{ c_t^{u,\theta,V} + \delta\, \mathrm{E}\left[W \zeta_{t+1} | \mathcal{F}_t^{R,F}\right] \right\}, \quad 0 \leq t \leq T-2,$$

$$\overline{\rho}_{t,s-1}(W, (u_r)_{r=t}^{s-1}) = \overline{\rho}_t(\overline{\rho}_{t+1}(\cdots, \overline{\rho}_{s-1}(W, u_{s-1})\cdots, u_{t+1}), u_t)$$

Then, by Lemma 11, we derive that

$$J_T^*(x, f, u) := \min_{(\theta, V) \in \mathcal{A}_T^{\theta,V}} J_T(x, f, \theta, V, u) = \overline{\rho}_{0,T-1}(\overline{\rho}_{T-1}(u_{T-1}), (u_t)_{t=0}^{T-2}),$$

for all $u = (u_t)_{t=0}^{T-1} \in \mathcal{A}_T^u$.

The investor needs to maximize $J_T^*(x, f, u)$ over $u \in \mathcal{A}_T^u$. In the proof of Lemma 11, the monotonicity of ρ_t plays a key role in the exchangeability of the minimization with respect to (θ, V). Similarly, the monotonicity of $\overline{\rho}_t$ is important to the the order of the maximization with respect to u. $\overline{\rho}_t$ satisfies the monotonicity, that is, for all $y \in \mathbb{R}^d$ and $0 \leq t \leq T-1$, $\overline{\rho}_t(W, y) \leq \overline{\rho}_t(W', y)$ holds if $W \leq W'$. This implies that for any $u \in \mathcal{A}_T^u$,

$$\overline{\rho}_{t,r-1}(W_r, (u_s)_{s=t}^{r-1}) \leq \overline{\rho}_{t,r-1}(W', (u_s)_{s=t}^{r-1}),$$

if $W_r \leq W'$. Therefore, using the same argument as in Lemma 11, we can prove the following lemma.

Lemma 12 For all $0 \leq t \leq s \leq T-1$, it holds that

$$\max_{(u_r)_{r=t}^{T-1} \in \mathcal{A}_T^u|_t^{T-1}} \overline{\rho}_{t,T-1}(\overline{\rho}_{T-1}(u_{T-1}), (u_r)_{r=t}^{T-2})$$

$$= \max_{(u_r)_{r=t}^{s-1} \in \mathcal{A}_T^u|_t^{s-1}} \overline{\rho}_{t,s-1}\Big(\max_{(u_r)_{r=s}^{T-1} \in \mathcal{A}_T^u|_s^{T-1}} \overline{\rho}_{s,T-2}\big(\rho_{T-1}(u_{T-1}), (u_r)_{r=s}^{T-2}\big), (u_r)_{r=t}^{s-1}\Big),$$

where $\mathcal{A}_T^u|_t^r$ is a set of $(u_s)_{s=t}^r$ which are components of $u \in \mathcal{A}_T^u$ from the time t to r.

Using Lemma 12, the dynamic programming of the multiple priors optimization can be justified. Now, we prove Theorem 5 and Proposition 6.

Proof of Theorem 5 and Proposition 6. Consider the time $T-1$ problem

$$\max_{u_{T-1}\in\mathcal{A}_T^u|_{T-1}^{T-1}} \overline{\rho}_{T-1}(u_{T-1})$$

$$= \max_{u_{T-1}\in\mathcal{A}_T^u|_{T-1}^{T-1}} \min_{(\theta_{T-1},V_{T-1})\in\mathcal{A}_T^{\theta,V}|_{T-1}^{T-1}} c_{T-1}^{u,\theta,V}$$

$$= \max_{u_{T-1}\in\mathcal{A}_T^u|_{T-1}^{T-1}} \min_{(\theta_{T-1},V_{T-1})\in\mathcal{A}_T^{\theta,V}|_{T-1}^{T-1}} X_{T-1}c_{T-1}(F_{T-1},\theta_{T-1},V_{T-1},u_{T-1})$$

$$= \max_{u_{T-1}\in\mathcal{A}_T^u|_{T-1}^{T-1}} \min_{(\theta_{T-1},V_{T-1})\in\mathcal{A}_T^{\theta,V}|_{T-1}^{T-1}} X_{T-1}\Big\{1 + r_{\mathrm{f}} + (m^e(F_{T-1}) + \theta_{T-1})'u_{T-1}$$

$$- \frac{\widehat{\gamma}_{T-1}}{2}u'_{T-1}\Big(A(F_{T-1}) + V_{T-1}\Big)u_{T-1}\Big\}. \quad (A.3)$$

Then, the minimization problem in the problem (A.3) is reduced to the following problem,

$$\min_{(\theta_{T-1},V_{T-1})\in\mathcal{A}_T^{\theta,V}|_{T-1}^{T-1}} \Big\{\theta'_{T-1}u_{T-1} - \frac{\widehat{\gamma}_{T-1}}{2}u'_{T-1}V_{T-1}u_{T-1}\Big\}. \quad (A.4)$$

The Lagrange function of the problem (A.4) is

$$\mathcal{L}^{\theta,V} = \theta'_{T-1}u_{T-1} - \frac{\widehat{\gamma}_{T-1}}{2}u'_{T-1}V_{T-1}u_{T-1} - \lambda_\theta\Big((\eta_{T-1}^\theta)^2 - \theta'_{T-1}(A(F_{T-1}) + V_{T-1})^{-1}\theta_{T-1}\Big)$$

$$- \lambda_V\Big((\eta_{T-1}^V)^2\|A(F_{T-1})\|^2 - \|V_{T-1}\|^2\Big),$$

where λ_θ and λ_V are Lagrange multipliers. We first consider the case when $u_{T-1} \neq \mathbf{0}_d$. The first order condition for θ_{T-1} is

$$u_{T-1} + 2\lambda_\theta(A(F_{T-1}) + V_{T-1})^{-1}\theta_{T-1} = \mathbf{0}_d.$$

Hence, we have

$$\theta_{T-1} = -\frac{1}{2\lambda_\theta}(A(F_{T-1}) + V_{T-1})u_{T-1}.$$

Since the constraint for θ_{T-1} is binded, it holds that

$$(\eta_{T-1}^\theta)^2 = \theta'_{T-1}(A(F_{T-1}) + V_{T-1})^{-1}\theta_{T-1} = \frac{1}{4\lambda_\theta^2}u'_{T-1}(A(F_{T-1}) + V_{T-1})u_{T-1}.$$

Hence,

$$\lambda_\theta = \frac{1}{2\eta_{T-1}^\theta}\sqrt{u'_{T-1}(A(F_{T-1}) + V_{T-1})u_{T-1}},$$

and

$$\theta_{T-1} = -\frac{\eta_{T-1}^\theta}{\sqrt{u'_{T-1}(A(F_{T-1}) + V_{T-1})u_{T-1}}}(A(F_{T-1}) + V_{T-1})u_{T-1}.$$

We denote the $i \times j$th element of V_{T-1} by $v_{T-1}^{i,j}$. Then, the first order condition for $v_{T-1}^{i,j}$ is

$$- \widehat{\gamma}_{T-1}u_{T-1}^i u_{T-1}^j + 4\lambda_V v_{T-1}^{i,j}$$

$$= \mathrm{tr}\left\{(A(F_{T-1}) + V_{T-1})^{-1}\theta_{T-1}\theta'_{T-1}(A(F_{T-1}) + V_{T-1})^{-1}N^{i,j}\right\}, \text{ if } i \neq j,$$

$$- \frac{\widehat{\gamma}_{T-1}}{2}(u_{T-1}^i)^2 + 2\lambda_V v_{T-1}^{i,j}$$

$$= \mathrm{tr}\left\{(A(F_{T-1}) + V_{T-1})^{-1}\theta_{T-1}\theta'_{T-1}(A(F_{T-1}) + V_{T-1})^{-1}N^{i,i}\right\}, \text{ if } i = j,$$

27

where $N^{i,j}$ is a d-dimensional matrix whose $i \times j$th element and $j \times i$th element are 1 and the other elements are zero. Substituting θ_{T-1} into $(A(F_{T-1})+V_{T-1})^{-1}\theta_{T-1}\theta'_{T-1}(A(F_{T-1})+V_{T-1})^{-1}$, we have

$$(A(F_{T-1}) + V_{T-1})^{-1}\theta_{T-1}\theta'_{T-1}(A(F_{T-1}) + V_{T-1})^{-1}$$
$$= \frac{(\eta^\theta_{T-1})^2}{u'_{T-1}(A(F_{T-1}) + V_{T-1})u_{T-1}}u_{T-1}u'_{T-1}.$$

Hence, it holds that

$$\text{tr}\left\{(A(F_{T-1}) + V_{T-1})^{-1}\theta_{T-1}\theta'_{T-1}(A(F_{T-1}) + V_{T-1})^{-1}N^{i,j}\right\}$$
$$= \begin{cases} \frac{2(\eta^\theta_{T-1})^2}{u'_{T-1}(A(F_{T-1})+V_{T-1})u_{T-1}}u^i_{T-1}u^j_{T-1}, & \text{if } i \neq j \\ \frac{(\eta^\theta_{T-1})^2}{u'_{T-1}(A(F_{T-1})+V_{T-1})u_{T-1}}(u^i_{T-1})^2, & \text{if } i = j \end{cases}$$

It follows that

$$-\widehat{\gamma}_{T-1}u^i_{T-1}u^j_{T-1} + 4\lambda_V v^{i,j}_{T-1} = \frac{2(\eta^\theta_{T-1})^2}{u'_{T-1}(A(F_{T-1}) + V_{T-1})u_{T-1}}u^i_{T-1}u^j_{T-1},$$

for all i and j. Using the matrix notation, we have

$$-\widehat{\gamma}_{T-1}u_{T-1}u'_{T-1} + 4\lambda_V V_{T-1} = \frac{2(\eta^\theta_{T-1})^2}{u'_{T-1}(A(F_{T-1}) + V_{T-1})u_{T-1}}u_{T-1}u'_{T-1}.$$

Hence,

$$V_{T-1} = \frac{1}{4\lambda_V}\left(\widehat{\gamma}_{T-1} + \frac{2(\eta^\theta_{T-1})^2}{u'_{T-1}(A(F_{T-1}) + V_{T-1})u_{T-1}}\right)u_{T-1}u'_{T-1}.$$

Since the constraint for V_{T-1} is also binded, it holds that

$$(\eta^V_{T-1})^2\|A(F_{T-1})\|^2$$
$$= \|V_{T-1}\|^2 = \frac{1}{16\lambda_V^2}\left(\widehat{\gamma}_{T-1} + \frac{2(\eta^\theta_{T-1})^2}{u'_{T-1}(A(F_{T-1}) + V_{T-1})u_{T-1}}\right)^2\|u_{T-1}\|^4.$$

Therefore, the Lagrange multiplier λ_V is

$$\lambda_V = \frac{1}{4\eta^V_{T-1}\|A(F_{T-1})\|}\left(\widehat{\gamma}_{T-1} + \frac{2(\eta^\theta_{T-1})^2}{u'_{T-1}(A(F_{T-1}) + V_{T-1})u_{T-1}}\right)\|u_{T-1}\|^2.$$

Hence, we have

$$V_{T-1} = \frac{\eta^V_{T-1}\|A(F_{T-1})\|}{\|u_{T-1}\|^2}u_{T-1}u'_{T-1}.$$

Then,

$$(A(F_{T-1}) + V_{T-1})u_{T-1} = (A(F_{T-1}) + \eta^V_{T-1}\|A(F_{T-1})\|I_d)u_{T-1}$$

and

$$\theta_{T-1} = -\frac{\eta^\theta_{T-1}(A(F_{T-1}) + \eta^V_{T-1}\|A(F_{T-1})\|I_d)u_{T-1}}{\sqrt{u'_{T-1}(A(F_{T-1}) + \eta^V_{T-1}\|A(F_{T-1})\|I_d)u_{T-1}}}.$$

If $u_{T-1} = \mathbf{0}_d$, then every $(\theta_{T-1}, V_{T-1}) \in \mathcal{A}^{\theta,V}_T|^{T-1}_{T-1}$ is a solution. In this case, we choose $\theta_{T-1} = \mathbf{0}_d$ and $V_{T-1} = \mathbf{O}_d$ as the solution. In all cases, (θ_{T-1}, V_{T-1}) is in $\mathcal{A}^{\theta,V}_T|^{T-1}_{T-1}$.

28

Substituting θ_{T-1} and V_{T-1} into the original time $T-1$ optimization problem (A.3), we obtain

$$
\max_{u_{T-1} \in \mathcal{A}_T^u|_{T-1}^{T-1}} X_{T-1}\Big\{ 1 + r_{\rm f} + (m^e(F_{T-1}))' u_{T-1}
$$

$$
- \frac{\widehat{\gamma}_{T-1}}{2} u'_{T-1}\Big(A(F_{T-1}) + \eta_{T-1}^V \|A(F_{T-1})\| I_d \Big) u_{T-1} \tag{A.5}
$$

$$
- \eta_{T-1}^\theta \sqrt{ u'_{T-1}\Big(A(F_{T-1}) + \eta_{T-1}^V \|A(F_{T-1})\| I_d \Big) u_{T-1}} \Big\}.
$$

We first consider the following case.

$$
a_{T-1}^e := (m^e(F_{T-1}))'\Big(A(F_{T-1}) + \eta_{T-1}^V \|A(F_{T-1})\| I_d \Big)^{-1} m^e(F_{T-1}) > (\eta_{T-1}^\theta)^2. \tag{A.6}
$$

The first order condition of this problem (A.5) is

$$
m^e(F_{T-1}) - \left(\widehat{\gamma}_{T-1} + \frac{\eta_{T-1}^\theta}{\psi_{T-1}} \right)\Big(A(F_{T-1}) + \eta_{T-1}^V \|A(F_{T-1})\| I_d \Big) u_{T-1} = \mathbf{0}_d,
$$

where

$$
\psi_{T-1} = \sqrt{ u'_{T-1}\Big(A(F_{T-1}) + \eta_{T-1}^V \|A(F_{T-1})\| I_d \Big) u_{T-1}}.
$$

Hence, we have

$$
u_{T-1} = \left(\widehat{\gamma}_{T-1} + \frac{\eta_{T-1}^\theta}{\psi_{T-1}} \right)^{-1}\Big(A(F_{T-1}) + \eta_{T-1}^V \|A(F_{T-1})\| I_d \Big)^{-1} m^e(F_{T-1}).
$$

Furthermore, we also have

$$
\psi_{T-1}^2 = u'_{T-1}\Big(A(F_{T-1}) + \eta_{T-1}^V \|A(F_{T-1})\| I_d \Big) u_{T-1}
$$

$$
= \left(\widehat{\gamma}_{T-1} + \frac{\eta_{T-1}^\theta}{\psi_{T-1}} \right)^{-2} a_{T-1}^e.
$$

Thus, we obtain the following quadratic equation for ψ_{T-1}.

$$
\widehat{\gamma}_{T-1}^2 \psi_{T-1}^2 + 2\widehat{\gamma}_{T-1}\eta_{T-1}^\theta \psi_{T-1} + (\eta_{T-1}^\theta)^2 - a_{T-1}^e = 0.
$$

By the inequality (A.6), the above equation has a unique positive solution such that

$$
\psi_{T-1} = \frac{\sqrt{a_{T-1}^e} - \eta_{T-1}^\theta}{\widehat{\gamma}_{T-1}}.
$$

Hence, the optimal portfolio at time $T-1$ is

$$
u_{T-1} = \frac{1}{\widehat{\gamma}_{T-1}} \left(1 - \frac{\eta_{T-1}^\theta}{\sqrt{a_{T-1}^e}} \right)\Big(A(F_{T-1}) + \eta_{T-1}^V \|A(F_{T-1})\| I_d \Big)^{-1} m^e(F_{T-1}).
$$

Let $K_{T-1} = A(F_{T-1}) + \eta_{T-1}^V \|A(F_{T-1})\| I_d$. If the inequality (A.6) does not hold, then the Cauchy-Schwartz inequality leads to

$$
\eta_{T-1}^\theta \sqrt{ u'_{T-1} K_{T-1} u_{T-1}} \geq \sqrt{a_{T-1}^e} \sqrt{u'_{T-1} K_{T-1} u_{T-1}}
$$

$$
= \sqrt{(m^e(F_{T-1}))' K_{T-1}^{-1} m^e(F_{T-1})} \sqrt{u'_{T-1} K_{T-1}(K_{T-1})^{-1} K_{T-1} u_{T-1}}
$$

$$
\geq \left| (m^e(F_{T-1}))'(K_{T-1})^{-1} K_{T-1} u_{T-1} \right| = \left| (m^e(F_{T-1}))' u_{T-1} \right|.
$$

Hence,

$$\max_{u_{T-1}\in\mathcal{A}^u_T|^{T-1}_{T-1}} X_{T-1}\Big\{1+r_{\mathrm{f}}+(m^e(F_{T-1}))'u_{T-1}$$

$$-\frac{\widehat{\gamma}_{T-1}}{2}u'_{T-1}\Big(A(F_{T-1})+\eta^V_{T-1}\|A(F_{T-1})\|I_d\Big)u_{T-1}$$

$$-\eta^\theta_{T-1}\sqrt{u'_{T-1}\Big(A(F_{T-1})+\eta^V_{T-1}\|A(F_{T-1})\|I_d\Big)u_{T-1}}\Big\}$$

$$\leq\max_{u_{T-1}\in\mathcal{A}^u_T|^{T-1}_{T-1}} X_{T-1}\Big\{1+r_{\mathrm{f}}+(m^e(F_{T-1}))'u_{T-1}-\Big|(m^e(F_{T-1}))'u_{T-1}\Big|\Big\}$$

The solution of the later problem is an arbitrarily vector $u_{T-1}\in\mathcal{A}^u_T|^{T-1}_{T-1}$, such that $(m^e(F_{T-1}))'u_{T-1}\geq 0$. We choose $u_{T-1}=\mathbf{0}_d$, and then the optimal value of the later problem becomes $X_{T-1}(1+r_{\mathrm{f}})$. Moreover, we have

$$1+r_{\mathrm{f}}+(m^e(F_{T-1}))'u_{T-1}-\frac{\widehat{\gamma}_{T-1}}{2}u'_{T-1}\Big(A(F_{T-1})+\eta^V_{T-1}\|A(F_{T-1})\|I_d\Big)u_{T-1}$$

$$-\eta^\theta_{T-1}\sqrt{u'_{T-1}\Big(A(F_{T-1})+\eta^V_{T-1}\|A(F_{T-1})\|I_d\Big)u_{T-1}}$$

$$=1+r_{\mathrm{f}},$$

if $u_{T-1}=\mathbf{0}_d$. Therefore, $u_{T-1}=\mathbf{0}_d$ is also a solution of the former problem.

In each case, we can express the optimal portfolio at time $T-1$ as

$$u_{T-1}=\frac{1}{\widehat{\gamma}_{T-1}}\left(1-\frac{\eta^\theta_{T-1}}{\sqrt{a^e_{T-1}}}\right)^+\Big(A(F_{T-1})+\eta^V_{T-1}\|A(F_{T-1})\|I_d\Big)^{-1}m^e(F_{T-1}).$$

It is clear that $u_{T-1}\in\mathcal{A}^u_T|^{T-1}_{T-1}$. In addition, the value function is

$$\max_{u_{T-1}\in\mathcal{A}^u_T|^T_T}\overline{\rho}_{T-1}(u_{T-1})=Y^T_{T-1}X_{T-1}$$

$$Y^T_{T-1}=\max_{u_{T-1}\in\mathbb{R}^d}\min_{(\theta_{T-1},V_{T-1})\in\mathcal{A}^{\theta,V}_T|^{T-1}_{T-1}}\{c_{T-1}(F_{T-1},\theta_{T-1},V_{T-1},u_{T-1})\}$$

$$=1+r_{\mathrm{f}}+\frac{\Big((\sqrt{a^e_{T-1}}-\eta^\theta_{T-1})^+\Big)^2}{2\widehat{\gamma}_{T-1}}.$$

By the definition, Y^T_{T-1} is $\mathcal{F}^F_{T-1}$-measurable.

Now, we assume that

$$\max_{(u_r)^{T-1}_{r=t+1}\in\mathcal{A}^u_T|^{T-1}_{t+1}}\min_{(\theta_r,V_r)^{T-1}_{r=t+1}\in\mathcal{A}^{\theta,V}_T|^{T-1}_{t+1}}\rho_{t+1,T-1}(c^{u,\theta,V}_{t+1},\ldots,c^{u,\theta,V}_{T-1})=Y^T_{t+1}X_{t+1},$$

for some $0\leq t\leq T-2$. Then, we have

$$\mathrm{E}^{\theta,V}[Y^T_{t+1}X_{t+1}|\mathcal{F}^{R,F}_t]$$

$$=\mathrm{E}^{\theta,V}[Y^T_{t+1}(1+r_{\mathrm{f}}+(R^e_{t+1})'u_t)X_t|\mathcal{F}^{R,F}_t]$$

$$=(1+r_{\mathrm{f}})\mathrm{E}^{\theta,V}[Y^T_{t+1}|\mathcal{F}^{R,F}_t]X_t+\Big(\mathrm{E}^{\theta,V}[Y^T_{t+1}R_{t+1}|\mathcal{F}^{R,F}_t]\Big)'u_tX_t$$

$$=(1+r_{\mathrm{f}})\mathrm{E}[Y^T_{t+1}|\mathcal{F}^{R,F}_t]X_t+\Big(\mathrm{E}[Y^T_{t+1}\mu^e(F_{t+1})|\mathcal{F}^{R,F}_t]+\mathrm{E}[Y^T_{t+1}|\mathcal{F}^{R,F}_t]\theta_t\Big)'u_tX_t,$$

where we have used Lemma 4 and $\mathcal{F}_{t+1}^F$-measurability of Y_{t+1}^T. Therefore, the objective in the time t problem is

$$c_t^{u,\theta,V} + \delta\, \mathrm{E}^{\theta,V}[Y_{t+1}^T X_{t+1}|\mathcal{F}_t^{R,F}] = X_t \left\{ (1+r_f)g_t^e + (m_t^e + g_t^e\theta_t)'u_t - \frac{\widehat{\gamma}_t}{2}u_t'(A_t+V_t)u_t \right\}.$$

Hence, the objective function in the time t problem has the same form as the objective function in the problem (A.3). Therefore, the error of the conditional mean is

$$\theta_t = -\frac{\mathrm{sgn}(g_t^e)\eta_t^\theta(A(F_t) + \eta_t^V\|A(F_t)\|I_d)u_t}{\sqrt{u_t'(A(F_t) + \eta_t^V\|A(F_t)\|I_d)u_t}},$$

where sgn is a sign function such that

$$\mathrm{sgn}(x) = \begin{cases} 1, & \text{if } x > 0, \\ -1, & \text{if } x < 0, \\ 0, & \text{if } x = 0. \end{cases}$$

The error of the conditional variance is

$$V_t = \frac{\eta_t^V\|A(F_t)\|}{\|u_t\|^2}u_t u_t'.$$

The optimal portfolio at time t is

$$u_t = \frac{1}{\widehat{\gamma}_t}\left(1 - \frac{|g_t^e|\eta_t^\theta}{\sqrt{a_t^e}}\right)^+ \left(A(F_t) + \eta_t^V\|A(F_t)\|I_d\right)^{-1} m_t^e.$$

It is clear that $u_t \in \mathcal{A}_T^u|_t$ and $(\theta_t, V_t) \in \mathcal{A}_T^{\theta,V}|_t$. The value function is

$$Y_t^T X_t = \max_{(u_r)_{r=t}^{T-1}\in\mathcal{A}_T^u|_t^{T-1}} \min_{(\theta_r,V_r)_{r=t}^{T-1}\in\mathcal{A}_T^{\theta,V}|_t^{T-1}} \rho_{t,T-1}(c_t^{u,\theta,V},\ldots,c_{T-1}^{u,\theta,V}),$$

$$Y_t^T = (1+r_f)g_t^e + \frac{\left((\sqrt{a_t^e} - |g_t^e|\eta_t^\theta)^+\right)^2}{2\widehat{\gamma}_t}.$$

By construction, Y_t^T is $\mathcal{F}_t^F$-measurable.

For any $0 \le t \le T-1$, by the mathematical induction, the value function at time t can be expressed as

$$Y_t^T X_t = \max_{(u_r)_{r=t}^{T-1}\in\mathcal{A}_T^u|_t^{T-1}} \min_{(\theta_r,V_r)_{r=t}^{T-1}\in\mathcal{A}_T^{\theta,V}|_t^{T-1}} \rho_{t,T-1}(c_t^{u,\theta,V},\ldots,c_{T-1}^{u,\theta,V}),$$

where

$$Y_t^T = (1+r_f)g_t^e + \frac{\left((\sqrt{a_t^e} - |g_t^e|\eta_t^\theta)^+\right)^2}{2\widehat{\gamma}_t}.$$

The optimal portfolio at time t is

$$u_t^* = \frac{1}{\widehat{\gamma}_t}\left(1 - \frac{|g_t^e|\eta_t^\theta}{\sqrt{a_t^e}}\right)^+ \left(A(F_t) + \eta_t^V\|A(F_t)\|I_d\right)^{-1} m_t^e.$$

Also, it is obvious that the optimal $(u_t^*)_{t=0}^{T-1}$ is time-consistent. $\qquad\square$

A.3 Proofs of Proposition 7 and 8

Proof of Proposition 7. By Theorem 5, we can use the following dynamic programming procedure.

$$\widehat{Y}_{T-1}^T = \max_{u_{T-1}\in\widetilde{\mathcal{A}}_T^u|_{T-1}^{T-1}} \min_{(\theta_{T-1},V_{T-1})\in\mathcal{A}_T^{\theta,V}|_{T-1}^{T-1}} \left\{ \widehat{c}_{T-1}(F_{T-1},\theta_{T-1},V_{T-1},u_{T-1}) \right\},$$

$$\widehat{Y}_t^T = \max_{u_t\in\widetilde{\mathcal{A}}_T^u|_t^t} \min_{(\theta_t,V_t)\mathcal{A}_T^{\theta,V}|_t^t} \left\{ \widehat{c}_t(F_t,\theta_t,V_t,u_t) + \delta\,\mathrm{E}^{\theta,V}\left[Y_{t+1}^T(1+(R_{t+1})'u_t) \;\middle|\; \mathcal{F}_t^{R,S} \right] \right\}.$$

Consider the time $T-1$ problem.

$$\widehat{Y}_{T-1}^T = \max_{u_{T-1}\in\widetilde{\mathcal{A}}_T^u|_{T-1}^{T-1}} \min_{(\theta_{T-1},V_{T-1})\in\mathcal{A}_T^{\theta,V}|_{T-1}^{T-1}} \left\{ 1 + (m(F_{T-1}) + \theta_{T-1})'u_{T-1} \right.$$
$$\left. - \frac{\widehat{\gamma}_{T-1}}{2}u_{T-1}'\left(A(F_{T-1}) + V_{T-1} \right)u_{T-1} \right\}. \tag{A.7}$$

The inner minimization problem in the problem (A.7) can be reduced to the following:

$$\min_{(\theta_{T-1},V_{T-1})\in\mathcal{A}_T^{\theta,V}|_{T-1}^{T-1}} \left\{ \theta_{T-1}'u_{T-1} - \frac{\widehat{\gamma}_{T-1}}{2}u_{T-1}'V_{T-1}u_{T-1} \right\}.$$

The above problem is the same as the problem (A.4) in the proof of Theorem 5. Therefore, the solutions are

$$\theta_{T-1} = -\frac{\eta_{T-1}^\theta(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d)u_{T-1}}{\sqrt{u_{T-1}'(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d)u_{T-1}}},$$

$$V_{T-1} = \frac{\eta_{T-1}^V\|A(F_{T-1})\|}{\|u_{T-1}\|^2}u_{T-1}u_{T-1}'.$$

Unlike Theorem 5, we do not need to consider the case $u_{T-1} = \mathbf{0}_d$ since $\mathbf{0}_d \notin \widetilde{\mathcal{A}}_T^u|_{T-1}^{T-1}$. Therefore, the time $T-1$ problem is

$$\widehat{Y}_{T-1}^T = \max_{u_{T-1}\in\widetilde{\mathcal{A}}_T^u|_{T-1}^{T-1}} \left\{ 1 + (m(F_{T-1}))'u_{T-1} - \frac{\widehat{\gamma}_{T-1}}{2}u_{T-1}'\left(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d \right)u_{T-1} \right.$$
$$\left. - \eta_{T-1}^\theta\sqrt{u_{T-1}'(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d)u_{T-1}} \right\}.$$

The Lagrange function of the above problem is

$$\mathcal{L}^u = 1 + (m(F_{T-1}))'u_{T-1} - \frac{\widehat{\gamma}_{T-1}}{2}u_{T-1}'\left(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d \right)u_{T-1}$$
$$- \eta_{T-1}^\theta\sqrt{u_{T-1}'(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d)u_{T-1}} + \lambda_u(1 - \mathbf{1}_d'u_{T-1}),$$

where λ_u is the Lagrange multiplier. The first order condition is

$$m(F_{T-1}) - \lambda_u\mathbf{1}_d - \left(\widehat{\gamma}_{T-1} + \frac{\eta_{T-1}^\theta}{\psi_{T-1}^*} \right)\left(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d \right)u_{T-1} = \mathbf{0}_d,$$

where

$$\psi_{T-1}^* = \sqrt{u_{T-1}'(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d)u_{T-1}}.$$

Hence,

$$u_{T-1} = \left(\widehat{\gamma}_{T-1} + \frac{\eta^{\theta}_{T-1}}{\psi^{*}_{T-1}}\right)^{-1} \left(A(F_{T-1}) + \eta^{V}_{T-1}\|A(F_{T-1})\|I_d\right)^{-1} \left(m(F_{T-1}) - \lambda_u \mathbf{1}_d\right).$$

Since the constraint is binded, we have

$$1 = \mathbf{1}'_d u_{T-1} = \left(\widehat{\gamma}_{T-1} + \frac{\eta^{\theta}_{T-1}}{\psi^{*}_{T-1}}\right)^{-1} \left(c^{worf}_{T-1} - \lambda_u b^{worf}_{T-1}\right).$$

Therefore, we obtain

$$\lambda_u = \frac{1}{b^{worf}_{T-1}} \left(c^{worf}_{T-1} - \left(\widehat{\gamma}_{T-1} + \frac{\eta^{\theta}_{T-1}}{\psi^{*}_{T-1}}\right)\right).$$

The optimal portfolio is

$$u_{T-1} = \frac{1}{\widehat{\gamma}_{T-1} + \eta^{\theta}_{T-1}/\psi^{*}_{T-1}} \left(A(F_{T-1}) + \eta^{V}_{T-1}\|A(F_{T-1})\|I_d\right)^{-1}$$
$$\times \left(m(F_{T-1}) - \frac{c^{worf}_{T-1} - (\widehat{\gamma}_{T-1} + \eta^{\theta}_{T-1}/\psi^{*}_{T-1})}{b^{worf}_{T-1}} \mathbf{1}_d\right).$$

Then,

$$b^{worf}_{T-1}(\psi^{*}_{T-1})^2$$
$$= b^{worf}_{T-1} u'_{T-1}(A(F_{T-1}) + \eta^{V}_{T-1}\|A(F_{T-1})\|I_d)u_{T-1}$$
$$= \frac{b^{worf}_{T-1}}{\left(\widehat{\gamma}_{T-1} + \frac{\eta^{\theta}_{T-1}}{\psi^{*}_{T-1}}\right)^2} \left(a^{worf}_{T-1} - 2\frac{c^{worf}_{T-1} - \left(\widehat{\gamma}_{T-1} + \frac{\eta^{\theta}_{T-1}}{\psi^{*}_{T-1}}\right)}{b^{worf}_{T-1}} c^{worf}_{T-1} + \frac{\left(c^{worf}_{T-1} - \left(\widehat{\gamma}_{T-1} + \frac{\eta^{\theta}_{T-1}}{\psi^{*}_{T-1}}\right)\right)^2}{b^{worf}_{T-1}}\right)$$
$$= \frac{b^{worf}_{T-1}}{\left(\widehat{\gamma}_{T-1} + \frac{\eta^{\theta}_{T-1}}{\psi^{*}_{T-1}}\right)^2} \left(a^{worf}_{T-1} - \frac{(c^{worf}_{T-1})^2}{b^{worf}_{T-1}} + \frac{\left(\widehat{\gamma}_{T-1} + \frac{\eta^{\theta}_{T-1}}{\psi^{*}_{T-1}}\right)^2}{b^{worf}_{T-1}}\right)$$
$$= \frac{a^{worf}_{T-1} b^{worf}_{T-1} - (c^{worf}_{T-1})^2}{\left(\widehat{\gamma}_{T-1} + \frac{\eta^{\theta}_{T-1}}{\psi^{*}_{T-1}}\right)^2} + 1 = \frac{d^{worf}_{T-1}}{\left(\widehat{\gamma}_{T-1} + \frac{\eta^{\theta}_{T-1}}{\psi^{*}_{T-1}}\right)^2} + 1.$$

Therefore, we can consider the following polynomial equation for ψ^{*}_{T-1}:

$$b^{worf}_{T-1}(\psi^{*}_{T-1})^2 = \frac{d^{worf}_{T-1}}{\left(\widehat{\gamma}_{T-1} + \eta^{\theta}_{T-1}/\psi^{*}_{T-1}\right)^2} + 1. \tag{A.8}$$

Expanding it, we obtain

$$q(\psi^{*}_{T-1}) := b^{worf}_{T-1}\widehat{\gamma}^2_{T-1}(\psi^{*}_{T-1})^4 + 2b^{worf}_{T-1}\widehat{\gamma}_{T-1}\eta^{\theta}_{T-1}(\psi^{*}_{T-1})^3$$
$$+ \left(b^{worf}_{T-1}(\eta^{\theta}_{T-1})^2 - (d^{worf}_{T-1} + \widehat{\gamma}^2_{T-1})\right)(\psi^{*}_{T-1})^2 \tag{A.9}$$
$$- 2\widehat{\gamma}_{T-1}\eta^{\theta}_{T-1}\psi^{*}_{T-1} - (\eta^{\theta}_{T-1})^2 = 0.$$

The equation $q(\psi^{*}_{T-1}) = 0$ is essentially the same as the equation (A11) in Garlappi et al. (2007), therefore it has a unique positive solution by the discussion in Garlappi et al. (2007). We show the existence of the unique positive solution to the equation (A.9). Since the equation

33

(A.9) is a quartic equation, it has four solutions. Since $q(0) = -(\eta_{T-1}^{\theta})^2 < 0$ and $b_{T-1}^{worf}\widehat{\gamma}_{T-1}^2 > 0$, the equation (A.9) has at least two real solutions: the one is a positive solution and the one is a negative solution. We show that the positive solution is unique. Consider the following quadratic equation.

$$\frac{\mathrm{d}^2 q(\psi_{T-1}^*)}{\mathrm{d}(\psi_{T-1}^*)^2} = 12 b_{T-1}^{worf}\widehat{\gamma}_{T-1}^2 (\psi_{T-1}^*)^2 + 12 b_{T-1}^{worf}\widehat{\gamma}_{T-1}\eta_{T-1}^{\theta}\psi_{T-1}^*$$
$$+ 2\Big(b_{T-1}^{worf}(\eta_{T-1}^{\theta})^2 - (d_{T-1}^{worf} + \widehat{\gamma}_{T-1}^2)\Big) = 0. \tag{A.10}$$

The discriminant of the quadratic equation (A.10) is

$$D = 36(b_{T-1}^{worf})^2\widehat{\gamma}_{T-1}^2(\eta_{T-1}^{\theta})^2 - 24 b_{T-1}^{worf}\widehat{\gamma}_{T-1}^2\Big(b_{T-1}^{worf}(\eta_{T-1}^{\theta})^2 - (d_{T-1}^{worf} + \widehat{\gamma}_{T-1}^2)\Big)$$
$$= 12(b_{T-1}^{worf})^2\widehat{\gamma}_{T-1}^2(\eta_{T-1}^{\theta})^2 + 24 b_{T-1}^{worf}\widehat{\gamma}_{T-1}^2(d_{T-1}^{worf} + \widehat{\gamma}_{T-1}^2) > 0.$$

Hence, the equation (A.10) has two real solutions. Moreover, we have

$$\frac{\mathrm{d}^2 q(\psi_{T-1}^*)}{\mathrm{d}(\psi_{T-1}^*)^2} = 12 b_{T-1}^{worf}\widehat{\gamma}_{T-1}^2 \left(\psi_{T-1}^* + \frac{\eta_{T-1}^{\theta}}{2\widehat{\gamma}_{T-1}}\right)^2 - \Big(b_{T-1}^{worf}(\eta_{T-1}^{\theta})^2 + d_{T-1}^{worf} + \widehat{\gamma}_{T-1}^2\Big).$$

Since $\eta_{T-1}^{\theta}/(2\widehat{\gamma}_{T-1}) > 0$, the equation (A.10) has at least one negative solution. Next, consider the first derivative of q such that

$$\frac{\mathrm{d}q(\psi_{T-1}^*)}{\mathrm{d}\psi_{T-1}^*} = 4 b_{T-1}^{worf}\widehat{\gamma}_{T-1}^2(\psi_{T-1}^*)^3 + 6 b_{T-1}^{worf}\widehat{\gamma}_{T-1}\eta_{T-1}(\psi_{T-1}^*)^2$$
$$+ 2\Big(b_{T-1}^{worf}(\eta_{T-1}^{\theta})^2 - (d_{T-1}^{worf} + \widehat{\gamma}_{T-1}^2)\Big)\psi_{T-1}^* - 2\widehat{\gamma}_{T-1}\eta_{T-1}^{\theta} = 0. \tag{A.11}$$

Then, the cubic equation (A.11) has at least one positive solution since $4 b_{T-1}^{worf}\widehat{\gamma}_{T-1}^2 > 0$ and $-2\widehat{\gamma}_{T-1}\eta_{T-1}^{\theta} < 0$. However, since the equation (A.10) has at least one negative solution, we conclude that the equation (A.11) has a unique positive solution. Now, let us consider the quartic equation (A.9) again. Since $q(0) = -(\eta_{T-1}^{\theta})^2 < 0$ and $b_{T-1}^{worf}\widehat{\gamma}_{T-1}^2 > 0$ hold, and since the equation (A.11) has a unique positive solution, the equation (A.9) has at most one local minimum on $\psi_{T-1}^* > 0$. Furthermore, we also have

$$\frac{\mathrm{d}q(\psi_{T-1}^*)}{\mathrm{d}\psi_{T-1}^*}\bigg|_{\psi_{T-1}^*=0} = -2\widehat{\gamma}_{T-1}\eta_{T-1}^{\theta} < 0.$$

These facts imply that the equation (A.9) has only one positive solution. Therefore, the positive solution of the equation (A.9) is unique. We also write ψ_{T-1}^* as the positive solution of the equation (A.9). Then, the optimal portfolio is

$$u_{T-1} = \frac{1}{\widehat{\gamma}_{T-1} + \eta_{T-1}^{\theta}/\psi_{T-1}^*}\Big(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d\Big)^{-1}$$
$$\times \left(m(F_{T-1}) - \frac{c_{T-1}^{worf} - (\widehat{\gamma}_{T-1} + \eta_{T-1}^{\theta}/\psi_{T-1}^*)}{b_{T-1}^{worf}}\mathbf{1}_d\right)$$
$$= \frac{1}{\widehat{\gamma}_{T-1} + |g_{T-1}|\eta_{T-1}^{\theta}/\psi_{T-1}^*}\Big(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d\Big)^{-1}$$
$$\times \left(m_{T-1} - \frac{c_{T-1}^{worf} - (\widehat{\gamma}_{T-1} + |g_{T-1}|\eta_{T-1}^{\theta}/\psi_{T-1}^*)}{b_{T-1}^{worf}}\mathbf{1}_d\right),$$

where we have used $g_{T-1} = 1$ and $m_{T-1} = m(F_{T-1})$. Since the coefficients of the quartic equation (A.9) are $\mathcal{F}_{T-1}^{R,F}$-measurable, its solution ψ_{T-1}^* is also $\mathcal{F}_{T-1}^{R,F}$-measurable. Therefore, u_{T-1} is $\mathcal{F}_{T-1}^{R,F}$-measurable. The value function at time $T-1$ is

$$
\begin{aligned}
\widehat{Y}_{T-1}^T &= 1 + (m(F_{T-1}))' u_{T-1} - \frac{\widehat{\gamma}_{T-1}}{2}(\psi_{T-1}^*)^2 - \eta_{T-1}^\theta \psi_{T-1}^* \\
&= g_{T-1} + \frac{1}{\widehat{\gamma}_{T-1} + |g_{T-1}|\eta_{T-1}^\theta / \psi_{T-1}^*} \left(a_{T-1}^{worf} - \frac{c_{T-1}^{worf} - (\widehat{\gamma}_{T-1} + |g_{T-1}|\eta_{T-1}^\theta / \psi_{T-1}^*)}{b_{T-1}^{worf}} c_{T-1}^{worf} \right) \\
&\quad - \frac{\widehat{\gamma}_{T-1}}{2}(\psi_{T-1}^*)^2 - |g_{T-1}|\eta_{T-1}^\theta \psi_{T-1}^* \\
&= g_{T-1} + \frac{c_{T-1}^{worf}}{b_{T-1}^{worf}} + \frac{d_{T-1}^{worf}}{b_{T-1}^{worf}(\widehat{\gamma}_{T-1} + |g_{T-1}|\eta_{T-1}^\theta / \psi_{T-1}^*)} - \frac{\widehat{\gamma}_{T-1}}{2}(\psi_{T-1}^*)^2 - |g_{T-1}|\eta_{T-1}^\theta \psi_{T-1}^*,
\end{aligned}
$$

where we have used $g_{T-1} = 1$. Using the equation (A.8), we have

$$
\begin{aligned}
&\frac{d_{T-1}^{worf}}{b_{T-1}^{worf}(\widehat{\gamma}_{T-1} + |g_{T-1}|\eta_{T-1}^\theta / \psi_{T-1}^*)} - \frac{\widehat{\gamma}_{T-1}}{2}(\psi_{T-1}^*)^2 - |g_{T-1}|\eta_{T-1}^\theta \psi_{T-1}^* \\
&= \left((\psi_{T-1}^*)^2 - \frac{1}{b_{T-1}^{worf}} \right) \left(\widehat{\gamma}_{T-1} + \frac{|g_{T-1}|\eta_{T-1}^\theta}{\psi_{T-1}^*} \right) - \left(\widehat{\gamma}_{T-1} + \frac{|g_{T-1}|\eta_{T-1}^\theta}{\psi_{T-1}^*} \right) (\psi_{T-1}^*)^2 \\
&\quad + \frac{\widehat{\gamma}_{T-1}}{2}(\psi_{T-1}^*)^2 \\
&= -\frac{\widehat{\gamma}_{T-1} + |g_{T-1}|\eta_{T-1}^\theta / \psi_{T-1}^*}{b_{T-1}^{worf}} + \frac{\widehat{\gamma}_{T-1}}{2}(\psi_{T-1}^*)^2.
\end{aligned}
$$

Hence, the value function is

$$
\widehat{Y}_{T-1}^T = g_{T-1} + \frac{c_{T-1}^{worf} - (\widehat{\gamma}_{T-1} + |g_{T-1}|\eta_{T-1}^\theta / \psi_{T-1}^*)}{b_{T-1}^{worf}} + \frac{\widehat{\gamma}_{T-1}}{2}(\psi_{T-1}^*)^2
$$

and it is $\mathcal{F}_{T-1}^F$-measurable.

Now, we assume that $\widehat{Y}_{t+1}^T$ is $\mathcal{F}_{t+1}^F$-measurable for some $0 \leq t \leq T-2$. Then, we have

$$
\begin{aligned}
&\widehat{c}_t(F_t, \theta_t, V_t, u_t) + \delta \, \mathrm{E}^{\theta,V} \left[Y_{t+1}^T (1 + (R_{t+1})' u_t) \,\Big|\, \mathcal{F}_t^{R,S} \right] \\
&= g_t + (m_t + g_t \theta_t)' u_t - \frac{\widehat{\gamma}_t}{2} u_t' \Big(A(F_t) + V_t \Big) u_t.
\end{aligned}
$$

Then, the minimization problem for (θ_t, V_t) can be reduced to

$$
\min_{(\theta_t, V_t) \in \mathcal{A}_T^{\theta,V}|_t} \left\{ g_t \theta_t' u_t - \frac{\widehat{\gamma}_t}{2} u_t' V_t u_t \right\}.
$$

Similarly to the time $T-1$ problem, the solution of the above problem is

$$
\begin{aligned}
\theta_t &= -\frac{\mathrm{sgn}(g_t)\eta_t^\theta (A(F_t) + \eta_t^V \|A(F_t)\| I_d) u_t}{\sqrt{u_t'(A(F_t) + \eta_t^V \|A(F_t)\| I_d) u_t}}, \\
V_t &= \frac{\eta_t^V \|A(F_t)\|}{\|u_t\|^2} u_t u_t'.
\end{aligned}
$$

Therefore, the maximization problem for u_t is

$$\max_{u_t \in \widetilde{\mathcal{A}}_T|_t^t} \left\{ g_t + m_t' u_t - \frac{\widehat{\gamma}_t}{2} u_t' \Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d \Big) u_t \right.$$
$$\left. - |g_t| \eta_t^\theta \sqrt{u_t' \Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d \Big) u_t} \right\}.$$

Since the above problem has the same form as the time $T - 1$ problem, the optimal portfolio is

$$u_t = \frac{1}{\widehat{\gamma}_t + |g_t| \eta_t^\theta / \psi_t^*} \Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d \Big)^{-1} \left(m_t - \frac{c_t^{worf} - (\widehat{\gamma}_t + |g_t| \eta_t^\theta / \psi_t^*)}{b_t^{worf}} \mathbf{1}_d \right),$$

where ψ_t^* is the unique positive solution of the following quartic equation:

$$b_t^{worf} (\psi_t^*)^2 = \frac{d_t^{worf}}{(\widehat{\gamma}_t + |g_t| \eta_t^\theta / \psi_t^*)^2} + 1.$$

The value function is

$$\widehat{Y}_t^T = g_t + \frac{c_t^{worf} - (\widehat{\gamma}_t + |g_t| \eta_t^\theta / \psi_t^*)}{b_t^{worf}} + \frac{\widehat{\gamma}_t}{2} (\psi_t^*)^2,$$

and it is $\mathcal{F}_t^F$-measurable.

By the mathematical induction, we therefore conclude that the optimal portfolio and that the value function is determined as in Proposition 7. □

Proof of Proposition 8. In this case, we can also apply the dynamic programming procedure. Consider the time $T - 1$ problem,

$$\widehat{Y}_{T-1}^{+T}$$
$$= \max_{u_{T-1} \in \widetilde{\mathcal{A}}_{+T}^u|_{T-1}^{T-1}} \min_{(\theta_{T-1}, V_{T-1}) \in \mathcal{A}_T^{\theta,V}|_{T-1}^{T-1}} \left\{ 1 + \Big(m(F_{T-1}) + \theta_{T-1} \Big)' u_{T-1} \right.$$
$$\left. - \frac{\widehat{\gamma}_{T-1}}{2} u_{T-1}' \Big(A(F_{T-1}) + V_{T-1} \Big) u_{T-1} \right\}.$$

Since $g_{T-1}^+ = 1$ and $m_{T-1}^+ = m(F_{T-1})$ by the definition, the objective function of the time $T - 1$ problem can be expressed as

$$g_{T-1}^+ + \Big(m_{T-1}^+ + g_{T-1}^+ \theta_{T-1} \Big)' u_{T-1} - \frac{\widehat{\gamma}_{T-1}}{2} u_{T-1}' \Big(A(F_{T-1}) + V_{T-1} \Big) u_{T-1}.$$

Hence, it suffices to consider the following optimization problem.

$$\max_{u_{T-1} \in \widetilde{\mathcal{A}}_{+T}^u|_{T-1}^{T-1}} \min_{(\theta_{T-1}, V_{T-1}) \in \mathcal{A}_T^{\theta,V}|_{T-1}^{T-1}} \left\{ g_{T-1}^+ + \Big(m_{T-1}^+ + g_{T-1}^+ \theta_{T-1} \Big)' u_{T-1} \right.$$
$$\left. - \frac{\widehat{\gamma}_{T-1}}{2} u_{T-1}' \Big(A(F_{T-1}) + V_{T-1} \Big) u_{T-1} \right\}.$$

Solving the inner minimization problem, we can express the above problem as

$$\max_{u_{T-1} \in \widetilde{\mathcal{A}}_{+T}^u|_{T-1}^{T-1}} \left\{ g_{T-1}^+ + (m_{T-1}^+)' u_{T-1} - \frac{\widehat{\gamma}_{T-1}}{2} u_{T-1}' \Big(A(F_{T-1}) + \eta_{T-1}^V \|A(F_{T-1})\| I_d \Big) u_{T-1} \right.$$
$$\left. - |g_{T-1}^+| \eta_{T-1}^\theta \sqrt{u_{T-1}' \Big(A(F_{T-1}) + \eta_{T-1}^V \|A(F_{T-1})\| I_d \Big) u_{T-1}} \right\}. \quad \text{(A.12)}$$

Now, we consider the auxiliary problem of (A.12), whose feasible set of portfolios is $\overline{\mathcal{C}^+}$. If we regard $\overline{\mathcal{C}^+}$ as a correspondence from $\mathbb{R}^K$ to $\mathbb{R}^d$, then its graph is

$$Gr(\overline{\mathcal{C}^+}) := \{(f,\phi) \in \mathbb{R}^K \times \mathbb{R}^d \mid \phi \in \overline{\mathcal{C}^+}\} = \mathbb{R}^K \times \overline{\mathcal{C}^+}.$$

It is clear that $Gr(\overline{\mathcal{C}^+})$ is a Borel subset on $\mathbb{R}^K \times \mathbb{R}^d$. This implies that $\overline{\mathcal{C}^+}$ is Borel measurable. Furthermore, the objective function of the problem (A.12) is continuous in u_{T-1} and also $\mathcal{F}_{T-1}^F$-measurable for any fixed u_{T-1}. Moreover, $\overline{\mathcal{C}^+}$ is a non-empty compact subset of $\mathbb{R}^d$ endowed with the Euclidean topology. Hence, by the measurable selection theorem (see Appendix D in Hernández-Lerma and Lasserre (1996)), there exists an $\mathcal{F}_{T-1}^F$-measurable random vector $u_{T-1}^* \in \overline{\mathcal{C}^+}$ such that

$$
\begin{aligned}
g_{T-1}^+ &+ (m_{T-1}^+)'u_{T-1}^* - \frac{\widehat{\gamma}_{T-1}}{2}(u_{T-1}^*)'\Big(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d\Big)u_{T-1}^* \\
&\qquad - |g_{T-1}^+|\eta_{T-1}^\theta\sqrt{(u_{T-1}^*)'\Big(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d\Big)u_{T-1}^*} \\
&= \max_{u_{T-1}\in\overline{\mathcal{C}^+}}\Bigg\{ g_{T-1}^+ + (m_{T-1}^+)'u_{T-1} - \frac{\widehat{\gamma}_{T-1}}{2}u_{T-1}'\Big(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d\Big)u_{T-1} \\
&\qquad - |g_{T-1}^+|\eta_{T-1}^\theta\sqrt{u_{T-1}'\Big(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d\Big)u_{T-1}}\Bigg\}.
\end{aligned}
$$

Moreover, since $u_{T-1}^* \in \widetilde{\mathcal{A}}_{+T}^u|_{T-1}^{T-1}$ and $\widetilde{\mathcal{A}}_{+T}^u|_{T-1}^{T-1} \subseteq \overline{\mathcal{C}^+}$, we conclude that

$$
\begin{aligned}
g_{T-1}^+ &+ (m_{T-1}^+)'u_{T-1}^* - \frac{\widehat{\gamma}_{T-1}}{2}(u_{T-1}^*)'\Big(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d\Big)u_{T-1}^* \\
&\qquad - |g_{T-1}^+|\eta_{T-1}^\theta\sqrt{(u_{T-1}^*)'\Big(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d\Big)u_{T-1}^*} \\
&= \max_{u_{T-1}\in\widetilde{\mathcal{A}}_{+T}^u|_{T-1}^{T-1}}\Bigg\{ g_{T-1}^+ + (m_{T-1}^+)'u_{T-1} - \frac{\widehat{\gamma}_{T-1}}{2}u_{T-1}'\Big(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d\Big)u_{T-1} \\
&\qquad\qquad - |g_{T-1}^+|\eta_{T-1}^\theta\sqrt{u_{T-1}'\Big(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d\Big)u_{T-1}}\Bigg\} \\
&= \max_{u_{T-1}\in\overline{\mathcal{C}^+}}\Bigg\{ g_{T-1}^+ + (m_{T-1}^+)'u_{T-1} - \frac{\widehat{\gamma}_{T-1}}{2}u_{T-1}'\Big(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d\Big)u_{T-1} \\
&\qquad\qquad - |g_{T-1}^+|\eta_{T-1}^\theta\sqrt{u_{T-1}'\Big(A(F_{T-1}) + \eta_{T-1}^V\|A(F_{T-1})\|I_d\Big)u_{T-1}}\Bigg\} \\
&= \widehat{Y}_{T-1}^{+T}.
\end{aligned}
$$

This implies that u_{T-1}^* is the maximizer of the problem (A.12). Moreover, $\widehat{Y}_{T-1}^{+T}$ is $\mathcal{F}_{T-1}^F$-measurable by the measurable selection theorem.

Now, we hypothesize that at some time $t+1$, $0 \leq t+1 \leq T-1$, $\widehat{Y}_{t+1}^{+T}$ is $\mathcal{F}_{t+1}^S$-measurable.

Then, we have

$$
\begin{aligned}
\widehat{Y}_t^{+T} &:= \max_{u_t \in \widetilde{\mathcal{A}}_{+T}^u|_t} \min_{(\theta_t, V_t) \in \mathcal{A}_T^{\theta,V}|_t} \left\{ 1 + (m(F_t) + \theta_t)' u_t - \frac{\widehat{\gamma}_t}{2} u_t' \Big(A(F_t) + V_t \Big) u_t \right. \\
&\qquad\qquad\qquad\qquad \left. + \delta \operatorname{E}^{\theta,V} \left[\widehat{Y}_{t+1}^{+T} \mid \mathcal{F}_t^{R,F} \right] \right\} \\
&= \max_{u_t \in \widetilde{\mathcal{A}}_{+T}^u|_t} \min_{(\theta_t, V_t) \in \mathcal{A}_T^{\theta,V}|_t} \left\{ g_t^+ + \Big(m_t^+ + g_t^+ \theta_t \Big)' u_t - \frac{\widehat{\gamma}_t}{2} u_t' \Big(A(F_t) + V_t \Big) u_t \right\} \\
&= \max_{u_t \in \widehat{\mathcal{A}}_{+T}^u|_t} \left\{ g_t^+ + (m_t^+)' u_t - \frac{\widehat{\gamma}_t}{2} u_t' \Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d \Big) u_t \right. \\
&\qquad\qquad\qquad\qquad \left. - |g_t^+| \eta_t^\theta \sqrt{u_t' \Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d \Big) u_t} \right\}. \qquad (A.13)
\end{aligned}
$$

The problem (A.13) has the same form as the problem (A.12), and $m_t^+, g_t^+, \widehat{\gamma}_t, A(F_t), \eta_t^\theta$ and η_t^V are $\mathcal{F}_t^F$-measurable. Therefore, we can use the measurable selection theorem in the same manner as we did in the case of the problem (A.12). By the measurable selection theorem, there exists an $\mathcal{F}_t^F$-measurable random vector $u_t^* \in \overline{\mathcal{C}^+}$ such that

$$
\begin{aligned}
g_t^+ &+ (m_t^+)' u_t^* - \frac{\widehat{\gamma}_t}{2} (u_t^*)' \Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d \Big) u_t^* \\
&\qquad\qquad\qquad - |g_t^+| \eta_t^\theta \sqrt{(u_t^*)' \Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d \Big) u_t^*} \\
&= \max_{u_t \in \widetilde{\mathcal{A}}_{+T}^u|_t} \left\{ g_t^+ + (m_t^+)' u_t - \frac{\widehat{\gamma}_t}{2} u_t' \Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d \Big) u_t \right. \\
&\qquad\qquad\qquad\qquad \left. - |g_t^+| \eta_t^\theta \sqrt{u_t' \Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d \Big) u_t} \right\} \\
&= \max_{u_t \in \mathcal{C}^+} \left\{ g_t^+ + (m_t^+)' u_t - \frac{\widehat{\gamma}_t}{2} u_t' \Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d \Big) u_t \right. \\
&\qquad\qquad\qquad\qquad \left. - |g_t^+| \eta_t^\theta \sqrt{u_t' \Big(A(F_t) + \eta_t^V \|A(F_t)\| I_d \Big) u_t} \right\} = \widehat{Y}_t^{+T}.
\end{aligned}
$$

Moreover, $\widehat{Y}_t^{+T}$ is $\mathcal{F}_t^F$-measurable.

By the mathematical induction, $\widehat{Y}_t^{+T}$ is $\mathcal{F}_t^F$-measurable for all $0 \le t \le T - 1$. Therefore, the value function of the original optimization problem is

$$
V_{+T}^{worf}(x, f) = \operatorname{E}[\widehat{Y}_t^{+T} \mid F_0 = f] x,
$$

and the optimal portfolio at time t is u_t^*. □

References

Ang, A., and G. Bekaert. 2002. International Asset Allocation with Regime Shifts. *Review of Financial Studies* 15:1137–1187. URL `http://dx.doi.org/10.1093/rfs/15.4.1137`.

Best, M. J., and R. R. Grauer. 1991. On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results. *Review of Financial Studies* 4:315–342. URL `http://dx.doi.org/10.1093/rfs/4.2.315`.

Björk, T., A. Murgoci, and X. Y. Zhou. 2014. Mean-Variance Portfolio Optimization with State-Dependent Risk Aversion. *Mathematical Finance* 24:1–24. URL `http://dx.doi.org/10.1111/j.1467-9965.2011.00515.x`.

Black, F., and R. Litterman. 1990. Asset Allocation: Combining Investor Views with Market Equilibrium. *Discussion Paper, Goldman, Sachs & Co* .

Chen, Z. P., G. Li, and Y. G. Zhao. 2014. Time-Consistent Investment Policies in Markovian Markets: A Case of Mean-Variance Analysis. *Journal of Economic Dynamics and Control* 40:293–316. URL `http://dx.doi.org/10.1016/j.jedc.2014.01.011`.

DeMiguel, V., L. Garlappi, F. J. Nogales, and R. Uppal. 2009a. A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms. *Management Science* 55:798–812. URL `http://dx.doi.org/10.1287/mnsc.1080.0986`.

DeMiguel, V., L. Garlappi, and R. Uppal. 2009b. Optimal Versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy? *Review of Financial Studies* 22:1915–1953. URL `http://dx.doi.org/10.1093/rfs/hhm075`.

Ellsberg, D. 1961. Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics* 75:643–669. URL `http://dx.doi.org/10.2307/1884324`.

Fama, E. F., and K. R. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33:3–56. URL `http://dx.doi.org/10.1016/0304-405X(93)90023-5`.

Garlappi, L., R. Uppal, and T. Wang. 2007. Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach. *Review of Financial Studies* 20:41–81. URL `http://dx.doi.org/10.1093/rfs/hhl003`.

Gilboa, I., and D. Schmeidler. 1989. Maxmin Expected Utility with Non-Unique Prior. *Journal of Mathematical Economics* 18:141–153. URL `http://dx.doi.org/10.1016/0304-4068(89)90018-9`.

Goldfarb, D., and G. Iyengar. 2003. Robust Portfolio Selection Problems. *Mathematics of Operations Research* 28:1–38. URL `http://dx.doi.org/10.1287/moor.28.1.1.14260`.

Hamilton, J. D. 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica* 57:357–384. URL `http://dx.doi.org/10.2307/1912559`.

Hamilton, J. D. 1990. Analysis of Time Series Subject to Changes in Regime. *Journal of Econometrics* 45:39–70. URL `http://dx.doi.org/10.1016/0304-4076(90)90093-9`.

Hernández-Lerma, O., and J. B. Lasserre. 1996. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, vol. 30. Springer Science & Business Media.

Jagannathan, R., and T. Ma. 2003. Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *The Journal of Finance* 58:1651–1683. URL `http://dx.doi.org/10.1111/1540-6261.00580`.

Jobson, J. D., and B. Korkie. 1980. Estimation for Markowitz Efficient Portfolios. *Journal of the American Statistical Association* 75:544–554. URL `http://dx.doi.org/10.1080/01621459.1980.10477507`.

Jorion, P. 1986. Bayes-Stein Estimation for Portfolio Analysis. *Journal of Financial and Quantitative Analysis* 21:279–292. URL `http://dx.doi.org/10.2307/2331042`.

Ledoit, O., and M. Wolf. 2003. Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection. *Journal of Empirical Finance* 10:603–621. URL `http://dx.doi.org/10.1016/S0927-5398(03)00007-0`.

Ledoit, O., and M. Wolf. 2004. A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices. *Journal of Multivariate Analysis* 88:365–411. URL `http://dx.doi.org/10.1016/S0047-259X(03)00096-4`.

Li, D., and W.-L. Ng. 2000. Optimal Dynamic Portfolio Selection: Multi-period Mean-Variance Formulation. *Mathematical Finance* 10:387–406. URL `http://dx.doi.org/10.1111/1467-9965.00100`.

Markowitz, H. 1952. Portfolio Selection. *The Journal of Finance* 7:77–91. URL `http://dx.doi.org/10.1111/j.1540-6261.1952.tb01525.x`.

Merton, R. C. 1972. An Analytic Derivation of the Efficient Portfolio Frontier. *Journal of Financial and Quantitative Analysis* 7:1851–1872. URL `http://dx.doi.org/10.2307/2329621`.

Merton, R. C. 1980. On Estimating the Expected Return on the Market: An Exploratory Investigation. *Journal of Financial Economics* 8:323–361. URL `http://dx.doi.org/10.1016/0304-405X(80)90007-0`.

Pástor, L., and R. F. Stambaugh. 2000. Comparing Asset Pricing Models: An Investment Perspective. *Journal of Financial Economics* 56:335–381. URL `http://dx.doi.org/10.1016/S0304-405X(00)00044-1`.

Pflug, G. C., A. Pichler, and D. Wozabal. 2012. The 1/N Investment Strategy Is Optimal under High Model Ambiguity. *Journal of Banking and Finance* 36:410–417. URL `http://dx.doi.org/10.1016/j.jbankfin.2011.07.018`.

Pınar, M. Ç. 2014. Equilibrium in an Ambiguity-Averse Mean-Variance Investors Market. *European Journal of Operational Research* 237:957–965. URL `http://dx.doi.org/10.1016/j.ejor.2014.02.016`.

Sharpe, W. F. 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance* 19:425–442. URL `http://dx.doi.org/10.1111/j.1540-6261.1964.tb02865.x`.

Tobin, J. 1958. Liquidity Preference As Behavior Towards Risk. *The Review of Economic Studies* 25:65–86. URL `http://dx.doi.org/10.2307/2296205`.

Wozabal, D. 2014. Robustifying Convex Risk Measures for Linear Portfolios: A Nonparametric Approach. *Operations Research* 62:1302–1315. URL `http://dx.doi.org/10.1287/opre.2014.1323`.

Table 1. The Abbreviations of Typical Portfolios and Six Extreme Strategies. Panel A displays the abbreviations of typical portfolios, and Panel B displays the details concerning the extreme strategies. The first column in Panel B lists the abbreviations of the extreme strategies. The second and third columns represent portfolios in different states. The forth and fifth columns display behaviors of parameters which justify the extreme strategies.

Panel A

Abbreviation	Portfolio
max SR	The portfolio maximizing the empirical Sharpe ratio
GMV	The single-period global minimum-variance portfolio
EW	The equally weighted portfolio
VW	The value-weighted portfolio
NE (No Error)	The no error portfolio (always $\eta_t^\theta = \eta_t^V = 0$)

Panel B

Abbreviation	Portfolio		Parameters	
	Good State	Bad State	Good State	Bad State
GMV-EW	GMV	EW	$\eta_t^V = 0$ and $\eta_t^\theta \to \infty$	$\eta_t^V \to \infty$
EW-NE	EW	No Error	$\eta_t^V \to \infty$	$\eta_t^\theta = \eta_t^V = 0$
GMV-NE	GMV	No Error	$\eta_t^V = 0$ and $\eta_t^\theta \to \infty$	$\eta_t^\theta = \eta_t^V = 0$
EW-GMV	EW	GMV	$\eta_t^V \to \infty$	$\eta_t^V = 0$ and $\eta_t^\theta \to \infty$
NE-EW	No Error	EW	$\eta_t^\theta = \eta_t^V = 0$	$\eta_t^V \to \infty$
NE-GMV	No Error	GMV	$\eta_t^\theta = \eta_t^V = 0$	$\eta_t^V = 0$ and $\eta_t^\theta \to \infty$

Table 2. International Diversification $(d = 4)$**.** This table reports the Sharpe ratios of various strategies obtained from the back test of the international asset allocations. "IID" represents the investor who believes that the number of market states is 1. "RS" represents the investor who believes that the number of market states is 2. The abbreviations of portfolios are listed in Table 1. In the case of $\eta_t^\theta = 0$ with short selling, we can not compute the optimal portfolios since they diverge. Therefore, we use the portfolios of $\eta_t^\theta = 1$ and $\eta_t^V = 0$ as the No Error portfolios in the extreme strategies.

Short selling is permitted.

IID

max SR		0.0	1.0	η_t^θ 2.0	3.0	4.0	5.0	GMV ∞
0.1014	0.0	-	0.0954	0.0941	0.0936	0.0934	0.0933	0.0927
	0.5	-	0.0878	0.0872	0.0869	0.0868	0.0867	0.0864
η_t^V	1.0	-	0.0849	0.0843	0.0842	0.0841	0.0840	0.0838
	3.0	-	0.0818	0.0815	0.0814	0.0814	0.0813	0.0812
	5.0	-	0.0810	0.0808	0.0807	0.0807	0.0807	0.0806

RS

max SR		0.0	1.0	η_t^θ 2.0	3.0	4.0	5.0	GMV ∞
0.0889	0.0	-	0.0837	0.0855	0.0859	0.0860	0.0861	0.0863
	0.5	-	0.0856	0.0852	0.0850	0.0849	0.0849	0.0847
η_t^V	1.0	-	0.0835	0.0831	0.0830	0.0829	0.0829	0.0827
	3.0	-	0.0813	0.0811	0.0810	0.0809	0.0809	0.0808
	5.0	-	0.0808	0.0805	0.0805	0.0804	0.0804	0.0803

Short selling is not permitted.

IID

max SR		0.0	1.0	η_t^θ 2.0	3.0	4.0	5.0	GMV ∞
0.1014	0.0	0.1025	0.1028	0.1021	0.1018	0.1017	0.1016	0.1012
	0.5	0.0896	0.0878	0.0872	0.0869	0.0868	0.0867	0.0864
η_t^V	1.0	0.1078	0.0849	0.0843	0.0842	0.0841	0.0840	0.0838
	3.0	0.1019	0.0818	0.0815	0.0814	0.0814	0.0813	0.0812
	5.0	0.1081	0.0810	0.0808	0.0807	0.0807	0.0807	0.0806

RS

max SR		0.0	1.0	η_t^θ 2.0	3.0	4.0	5.0	GMV ∞
0.0889	0.0	0.0597	0.0973	0.0952	0.0942	0.0937	0.0934	0.0920
	0.5	0.0674	0.0856	0.0852	0.0850	0.0849	0.0849	0.0847
η_t^V	1.0	0.0583	0.0835	0.0831	0.0830	0.0829	0.0829	0.0827
	3.0	0.0577	0.0813	0.0811	0.0810	0.0809	0.0809	0.0808
	5.0	0.0570	0.0808	0.0805	0.0805	0.0804	0.0804	0.0803

EW	Extreme Strategies					
0.0795	Short selling is permitted.			Short selling is not permitted.		
	GMV-EW	EW-NE	GMV-NE	GMV-EW	EW-NE	GMV-NE
	0.1032	0.0664	0.0891	0.0922	0.0949	0.1078
	EW-GMV	NE-EW	NE-GMV	EW-GMV	NE-EW	NE-GMV
	0.0636	0.0976	0.0809	0.0783	0.0457	0.0433

Table 3. Diversification across Industries ($d = 5$). This table reports the Sharpe ratios of various strategies obtained from the back test of the asset allocations across industries. "IID" represents the investor who believes that the number of market states is 1. "RS" represents the investor who believes that the number of market states is 2. The abbreviations of portfolios are listed in Table 1. In the case of $\eta_t^\theta = 0$ with short selling, we can not compute the optimal portfolios since they diverge. Therefore, we use the portfolios of $\eta_t^\theta = 1$ and $\eta_t^V = 0$ as the No Error portfolios in the extreme strategies.

Short selling is permitted.

IID

max SR					η_t^θ			GMV
		0.0	1.0	2.0	3.0	4.0	5.0	∞
0.2210	0.0	-	0.2732	0.2791	0.2808	0.2817	0.2821	0.2839
	0.5	-	0.2213	0.2220	0.2222	0.2224	0.2224	0.2227
η_t^V	1.0	-	0.2165	0.2169	0.2170	0.2171	0.2171	0.2173
	3.0	-	0.2127	0.2129	0.2129	0.2130	0.2130	0.2131
	5.0	-	0.2119	0.2120	0.2121	0.2121	0.2121	0.2121

RS

max SR					η_t^θ			GMV
		0.0	1.0	2.0	3.0	4.0	5.0	∞
0.1968	0.0	-	0.1912	0.2524	0.2601	0.2632	0.2649	0.2707
	0.5	-	0.2238	0.2269	0.2278	0.2283	0.2286	0.2297
η_t^V	1.0	-	0.2181	0.2202	0.2209	0.2212	0.2214	0.2222
	3.0	-	0.2129	0.2141	0.2145	0.2147	0.2148	0.2153
	5.0	-	0.2118	0.2127	0.2130	0.2131	0.2132	0.2136

Short selling is not permitted.

IID

max SR					η_t^θ			GMV
		0.0	1.0	2.0	3.0	4.0	5.0	∞
0.2210	0.0	0.1822	0.2496	0.2547	0.2561	0.2570	0.2576	0.2596
	0.5	0.1643	0.2213	0.2220	0.2222	0.2224	0.2224	0.2227
η_t^V	1.0	0.1724	0.2165	0.2169	0.2170	0.2171	0.2171	0.2173
	3.0	0.1783	0.2127	0.2129	0.2129	0.2130	0.2130	0.2131
	5.0	0.1487	0.2119	0.2120	0.2121	0.2121	0.2121	0.2121

RS

max SR					η_t^θ			GMV
		0.0	1.0	2.0	3.0	4.0	5.0	∞
0.1968	0.0	0.0660	0.2394	0.2502	0.2543	0.2573	0.2592	0.2664
	0.5	0.0796	0.2239	0.2269	0.2278	0.2283	0.2286	0.2297
η_t^V	1.0	0.0769	0.2181	0.2202	0.2209	0.2212	0.2214	0.2222
	3.0	0.0653	0.2129	0.2141	0.2145	0.2147	0.2148	0.2153
	5.0	0.0841	0.2118	0.2127	0.2130	0.2131	0.2132	0.2136

	Extreme Strategies						
EW							
0.0795	Short selling is permitted.			Short selling is not permitted.			
	GMV-EW	EW-NE	GMV-NE	GMV-EW	EW-NE	GMV-NE	
VW	0.2234	0.2457	0.2592	0.2281	0.1199	0.1344	
0.1825	EW-GMV	NE-EW	NE-GMV	EW-GMV	NE-EW	NE-GMV	
	0.2564	0.1657	0.1964	0.2473	0.1410	0.1678	

Table 4. Diversification among Sizes and Values ($d = 6$). This table reports the Sharpe ratios of various strategies obtained from the back test of the asset allocations among the 2×3 size- and book-to-market-sorted portfolios by Fama and French (1993). "IID" represents the investor who believes that the number of market states is 1. "RS" represents the investor who believes that the number of market states is 2. The abbreviations of portfolios are listed in Table 1. In the case of $\eta_t^\theta = 0$ with short selling, we can not compute the optimal portfolios since they diverge. Therefore, we use the portfolios of $\eta_t^\theta = 1$ and $\eta_t^V = 0$ as the No Error portfolios in the extreme strategies.

Short selling is permitted.

IID

max SR		η_t^θ 0.0	1.0	2.0	3.0	4.0	5.0	GMV ∞
0.2139	0.0	-	0.3406	0.3293	0.3243	0.3215	0.3197	0.3122
	0.5	-	0.2237	0.2228	0.2224	0.2223	0.2222	0.2218
η_t^V	1.0	-	0.2176	0.2170	0.2168	0.2167	0.2166	0.2164
	3.0	-	0.2126	0.2123	0.2122	0.2122	0.2121	0.2120
	5.0	-	0.2115	0.2113	0.2112	0.2111	0.2111	0.2110

RS

max SR		η_t^θ 0.0	1.0	2.0	3.0	4.0	5.0	GMV ∞
0.2081	0.0	-	0.2389	0.2473	0.2486	0.2489	0.2491	0.2488
	0.5	-	0.2184	0.2195	0.2198	0.2200	0.2200	0.2203
η_t^V	1.0	-	0.2147	0.2154	0.2156	0.2157	0.2158	0.2160
	3.0	-	0.2113	0.2117	0.2118	0.2119	0.2119	0.2120
	5.0	-	0.2105	0.2108	0.2109	0.2109	0.2110	0.2110

Short selling is not permitted.

IID

max SR		η_t^θ 0.0	1.0	2.0	3.0	4.0	5.0	GMV ∞
0.2139	0.0	0.2202	0.2260	0.2263	0.2255	0.2253	0.2257	0.2256
	0.5	0.2289	0.2237	0.2228	0.2224	0.2223	0.2222	0.2218
η_t^V	1.0	0.2207	0.2176	0.2170	0.2168	0.2167	0.2166	0.2164
	3.0	0.2271	0.2126	0.2123	0.2122	0.2122	0.2121	0.2120
	5.0	0.2291	0.2115	0.2113	0.2112	0.2111	0.2111	0.2110

RS

max SR		η_t^θ 0.0	1.0	2.0	3.0	4.0	5.0	GMV ∞
0.2081	0.0	0.1791	0.2101	0.2133	0.2146	0.2155	0.2158	0.2175
	0.5	0.1874	0.2184	0.2195	0.2198	0.2199	0.2200	0.2203
η_t^V	1.0	0.1949	0.2147	0.2154	0.2156	0.2157	0.2158	0.2160
	3.0	0.1856	0.2113	0.2117	0.2118	0.2119	0.2119	0.2120
	5.0	0.1999	0.2105	0.2108	0.2109	0.2109	0.2110	0.2110

Extreme Strategies

EW 0.0795	Short selling is permitted.			Short selling is not permitted.		
	GMV-EW	EW-NE	GMV-NE	GMV-EW	EW-NE	GMV-NE
VW 0.2010	0.2763	0.1697	0.2281	0.2336	0.1790	0.1979
	EW-GMV	NE-EW	NE-GMV	EW-GMV	NE-EW	NE-GMV
	0.1876	0.2861	0.2592	0.1939	0.2069	0.1920

Table 5. The Results of the Mimicking Strategies' Sharpe Ratios Obtained from the Back Tests. The second and third columns display parameters in the good state, and the fourth and fifth columns display parameters in the bad state. The sixth column shows whether short selling is permitted or not. The seventh column reports the Sharpe ratios of the mimicking extreme strategies. The eighth column reports the Sharpe ratios of the original extreme strategies. The ninth column reports the average root square errors between the original extreme strategies and the mimicking strategies. The extreme strategies are listed in Table 1.

International Indexes

	Good State		Bad State		Short Selling	Sharpe Ratio	Original Sharpe Ratio	Ave. RSE
	η_t^θ	η_t^V	η_t^θ	η_t^V				
mGMV-EW	30	0	10	5	Yes	0.1038	0.1032	0.0116
					No	0.0929	0.0922	0.0102
mEW-GMV	10	5	30	0	Yes	0.0640	0.0636	0.0116
					No	0.0788	0.0783	0.0112

Industry Indexes

	Good State		Bad State		Short Selling	Sharpe Ratio	Original Sharpe Ratio	Ave. RSE
	η_t^θ	η_t^V	η_t^θ	η_t^V				
mGMV-EW	30	0	10	5	Yes	0.2252	0.2234	0.0134
					No	0.2303	0.2281	0.0089
mEW-GMV	10	5	30	0	Yes	0.2567	0.2564	0.0095
					No	0.2470	0.2473	0.0079

Size- and Value-Sorted Portfolios

	Good State		Bad State		Short Selling	Sharpe Ratio	Original Sharpe Ratio	Ave. RSE
	η_t^θ	η_t^V	η_t^θ	η_t^V				
mGMV-EW	30	0	10	5	Yes	0.2776	0.2763	0.0248
					No	0.2343	0.2336	0.0118
mEW-GMV	10	5	30	0	Yes	0.1882	0.1876	0.0199
					No	0.1945	0.1939	0.0117

Figure 1. Portfolio Errors of International Indexes Data Set When Short Selling Is Permitted. The following four figures display portfolio errors $\|u_t^*\|_{GMV}$ and $\|u_t^*\|_{EW}$ at each time t. The upper two figures show the portfolio errors from GMV, $\|u_t^*\|_{GMV}$, with $\eta_t^V = 0$. The lower two figures represent the portfolio errors from EW, $\|u_t^*\|_{EW}$, with $\eta_t^\theta = 5$. The left two figures represent the portfolio errors of the IID investor and the right two figures show those of the RS investor.

Portfolio Errors between GMV and the Optimal Portfolio

Portfolio Errors between EW and the Optimal Portfolio