Brazilian Journal of Probability and Statistics
2015, Vol. 29, No. 3, 656–676
DOI: 10.1214/14-BJPS238

An optimal combination of risk-return and naive hedging

Wan-Yi Chiu

National United University

Abstract. Taking an approach contrary to the mean–variance portfolio, recent studies have appealed to an older wisdom, "the naive rule provides the best solution," to improve out-of-sample performance in portfolio selection. Naive diversification, which invests equally across risky assets, is such an example of this simple rule. Previous studies also show that a portfolio combining naive diversification with the mean–variance strategy based on minimizing expected quadratic utility losses may show strong out-of-sample performance. Using the mean squared error, this paper derives an optimal combination of nonstochastic allocation and the mean–variance portfolio. We find that this design is equivalent to the combination of the naive rule and mean–variance strategy based on minimizing the expected utility losses. As an application of this finding, we propose a regression-based combination of maximal risk-return hedging and naive hedging. Our illustration also shows out-of-sample performance of a combined hedging that is superior to that of other methods.

1 Introduction

It is widely known that Markowitz's mean–variance efficient portfolio (1952) is a constrained optimization problem of the first two statistical moments. The first moment is the expected value, which measures the expected excess return of a portfolio, and the second moment is the variance (or the standard deviation), which describes the risk of the portfolio's excess return. Specifically, the mean–variance optimization is a tradeoff between expected return and risk, where the mathematical formulation attempts to maximize the portfolio's expected return conditional on a given level of portfolio risk or, equivalently, minimize risk for a given expected return by efficiently choosing the allocations across various assets (Merton (1972), Jobson and Korkie (1980)).

Despite its theoretical importance in the financial industry, critics of the mean–variance efficient portfolios challenge whether it is an ideal investing strategy when investors do not accurately consider model uncertainty. The two most widely discussed model uncertainties in determining the optimal portfolio selection are parameter sensitivity and estimation errors. First, in particular, sensitivity due to changes in the equilibrium parameters results in mean–variance uncertainty (Best

Key words and phrases. Optimal hedging, risk-return hedging, naive hedging, naive rule, combined forecasts.

Received June 2013; accepted February 2014.

and Grauer (1991), Chopra and Ziemba (1993)). Second, the mean–variance efficient portfolio weights are expressed in terms of the expected return and the covariance matrix of assets, and these parameters are not directly observed. The investors must use the sample information to estimate assets' return and volatility, and such estimators often fail to consider estimation errors where the historical data were generated (Black and Litterman (1992), Kan and Zhou (2007)).

Many impetuses to translate the theoretical restrictions into a practical portfolio construction have been advocated by introducing constrained and unconstrained procedures to the literature. For example: (1) Previous studies have shown that instabilities in mean–variance analysis disappear when a regularizing constraint is incorporated into the optimization procedure (Green and Hollifield (1992), Jagannthan and Ma (2003), Wang (2005), DeMiguel et al. (2009a)). (2) A handful of Bayesian approaches or nonstandard Bayesian frameworks have attempted to incorporate expert knowledge into the data-generating process. These studies draw on the combination of investors' subjective opinion and the expected return vector as well as the covariance matrix in the long run (Jorion (1986), Black and Litterman (1992), MacKinlay and Pástor (2000)). (3) Some authors assert that the minimum-variance portfolio shows superior out-of-sample performance in the presence of estimation errors (Haugen and Baker (1991), Ledoit and Wolf (2004)). (4) Recently, Hirschberger et al. (2013) propose a posteriori way to balance a classical mean–variance optimization with a given nonstochastic portfolio. In this approach, we are allowed to include an additional criterion than mean–variance in the optimization, where we may calculate the entire set of theoretically attainable and optimal portfolios for investors, irrespective of their risk aversion.

Contrary to the optimal tangency portfolio, recent studies have proposed a rather simple alternative to the mean–variance efficient portfolios, so-called naive diversification, which invests equally across k risky assets. For example: (1) DeMiguel et al. (2009b) and Duchin and Levy (2009) report that naive diversification dominates mean–variance optimization in out-of-sample asset allocation tests. Some evidence indicates that a long estimation horizon is needed for the mean–variance efficient portfolio to outperform naive diversification. (2) However, Kirby and Ostdiek (2012) find that volatility timing and reward-to-risk timing mean–variance strategies outperform naive diversification in the presence of high transaction costs. (3) Specifically, Tu and Zhou (2011) characterize an optimal combination of the naive rule and the mean–variance efficient portfolio under the minimal expected loss of mean–variance utility criterion, which considers both estimation error (variance) and error resulting from model misspecification (bias). This combination is not theoretically unbiased, where the naive weight is obviously a biased estimator of the true tangency portfolio, but there is a low expected loss of mean–variance utility. Empirical studies show that this strategy may successfully improve a portfolio's out-of-sample performance.

In this article, we deliver a setting similar to but more general than that of Tu and Zhou (2011). Rather than deriving our combination coefficient in terms of the investor's utility function, we solve the optimization from a combined forecasts perspective (Bates and Granger (1969), Clemen (1989)). We use the mean squared error criterion, which allows us to derive an analytic expression for the optimal combination of any nonstochastic weight and mean–variance portfolio without considering the investor's utility preferences. Among the interesting properties of the constructed design are the following: (1) We obtain an analytic solution for the optimal combination between a nonstochastic allocation and the mean–variance efficient portfolio. From a multicriteria decision making perspective, it falls in the category of a priori methods (Hwang and Masud (1979)), where the risk aversion of the investor is given, and a single optimal portfolio for the investor is being computed. (2) The optimal design is equivalent to the Tu and Zhou's model. Moreover, naive diversification may be replaced by any nonstochastic allocation that is biased but does not depend on the sample information. (3) We use this result to obtain an optimal combination of risk-return hedging and naive hedging. The hedging strategies using Chiu's (2013) regression-based procedures are further analyzed. (4) We find that this optimal design appears to include as a special case the combination of the naive rule and the optimal mean–variance strategy based on minimizing the combined variance. (5) An example is presented to investigate the hedging effectiveness. We demonstrate that the combined rules not only have a significant impact in improving the mean–variance strategies but also outperform the naive portfolio in our studies.

This article is organized as follows. In Section 2, we briefly investigate the framework between the CARA utility and the mean–variance utility. According to this equivalence, we extend Tu and Zhou's results to a statistical perspective. In Section 3, we employ the optimal combination of an unbiased estimator and a biased estimator to derive the Tu–Zhou optimal combination coefficient under the mean squared errors criterion. Finally, a regression-based application of the optimal combination of risk-return hedging and naive hedging is presented. Section 4 concludes.

2 The preliminaries

2.1 CARA investor's portfolio choice

We consider a two-period framework populated by risk-averse investors. In the first period, a budget-constrained investor forms a portfolio by allocating her or his wealth into a risk-free asset with a rate r, Q_1 shares of risky asset 1 with a price $P_{1,t}$ per share and a payoff $P_{1,t+1}$ in the second period, Q_2 shares of risky asset 2 with a price $P_{2,t}$ per share and a payoff $P_{2,t+1}$ in the second period, and likewise for $(Q_3, P_{3,t}, P_{3,t+1}), \dots, (Q_k, P_{k,t}, P_{k,t+1})$.

The investor's first- and second-period wealth must satisfy the budget constraints,

$$
\begin{aligned}
W_t &= \sum_{i=1}^{k} Q_i P_{i,t} + \left(W_t - \sum_{i=1}^{k} Q_i P_{i,t} \right), \\
W_{t+1} &= \sum_{i=1}^{k} Q_i P_{i,t+1} + \left(W_t - \sum_{i=1}^{k} Q_i P_{i,t} \right)(1+r).
\end{aligned}
\tag{1}
$$

The constraints in equation (1) implies that the investor's excess return of portfolio w_{p} at $t+1$ is given by

$$
R_{\mathrm{p},t+1} = W_{t+1} - (1+r)W_t = \sum_{i=1}^{k} Q_i P_{i,t} \left(\frac{P_{i,t+1} - P_{i,t}}{P_{i,t}} - r \right) = w_{\mathrm{p}}' R_{t+1}, \tag{2}
$$

where $w_{\mathrm{p}} = (w_1, w_2, \ldots, w_k)$, and $R_{t+1} = (R_{1,t+1}, R_{2,t+1}, \ldots, R_{k,t+1})$ represent the portfolio allocation vector and excess returns vector, respectively, k risky assets at time $t+1$. Moreover, this portfolio has the expected excess return $E(R_{\mathrm{p},t+1})$ and the variance of the excess return $\mathrm{Var}(R_{\mathrm{p},t+1})$. For the investor, we further employ the assumption that the values of the k excess returns vector are multivariate and normally distributed random variables with unknown expected mean vector μ and covariance Σ. Moreover, we assume that the sample data of n periods excess returns $\{R_1, R_2, \ldots, R_n, R_{n+1}\}$ are independent and identically distributed over time. That is,

$$
R_1, R_2, \ldots, R_{n+1} \overset{\text{i.i.d.}}{\sim} \mathrm{MN}(\mu, \Sigma). \tag{3}
$$

Tu and Zhou (2004) show that to incorporate the uncertainty of a data-generating process into portfolio performance, the normality assumption works well in evaluating a mean–variance portfolio. Moreover, assume that the profit-maximizing investor has a negative exponential utility function (or, in terms of the constant absolute risk aversion utility, CARA) with a given risk aversion coefficient parameter $\gamma > 0$. In other words, the investor's utility function is of the form

$$
U(W_{t+1}) = -\exp(-\gamma W_{t+1}). \tag{4}
$$

Grossman and Stiglitz (1980) have motivated the exponentially normal framework to different applications in financial documents. Khoury and Martel (1985) and Kirilenko (2001), among others, use the model to evaluate an investment project given different types of investor behavior.

According to assumption (4), the investor will determine her optimal portfolio allocation by maximizing the expected utility of the liquidation period net return,

$$
\max_{w} E\{-\exp(-\gamma \Delta W)\} = \max_{w} E\{-\exp[-\gamma (R_{\mathrm{p},t+1})]\},
$$

where $\Delta W = W_{t+1} - (1+r)W_t$ (positive or negative) is her net wealth increment in the form of excess return at the second period for some risk-aversion parameter $\gamma > 0$. Because the expected utility function is the moment-generating function of a multivariate normal random vector under assumption (3), the functional and normal assumptions jointly imply that the investor's expected utility is

$$E\big(U(W_{t+1})\big) = -\exp\left[-\gamma\big(E(R_{\mathrm{p},t+1})\big) + \frac{\gamma^2(\mathrm{Var}(R_{\mathrm{p},t+1}))}{2}\right].$$

Therefore, we may translate the CARA investor's portfolio choice into a mean–variance utility optimization as follows:

$$\begin{aligned}\max_{w_\mathrm{p}} U(w) &= \max_{w_\mathrm{p}}\left[E(R_{\mathrm{p},t+1}) - \frac{\gamma}{2}\mathrm{Var}(R_{\mathrm{p},t+1})\right]\\ &= \max_{w_\mathrm{p}}\left[w_\mathrm{p}'\mu - \frac{\gamma}{2}w_\mathrm{p}'\Sigma w_\mathrm{p}\right].\end{aligned} \tag{5}$$

The maximum mean–variance utility in model (5) is a standard criterion that Kan and Zhou (2007), DeMiguel et al. (2009b), Das et al. (2010), Tu and Zhou (2011), among others, use to establish an optimal portfolio rule. It is trivial to derive the optimal mean–variance portfolio

$$w_\mathrm{o} = \frac{\Sigma^{-1}\mu}{\gamma}. \tag{6}$$

Substituting equation (6) into equation (5), the expected mean–variance utility relative to w_o at time $t+1$ is reduced to the form

$$U(w_\mathrm{o}) = E\big(w_\mathrm{o}'R_{t+1}\big) - \frac{\gamma}{2}V\big(w_\mathrm{o}'R_{t+1}\big) = w_\mathrm{o}'\mu - \frac{\gamma}{2}w_\mathrm{o}'\Sigma w_\mathrm{o} = \frac{\mu'\Sigma^{-1}\mu}{2\gamma}, \tag{7}$$

where $\sqrt{\mu'\Sigma^{-1}\mu}$ is the Sharpe ratio relative to the optimal tangency portfolio.

Remark 1. It should be emphasized that the optimal portfolio w_o in equation (6) is selected from the capital market line and depends only on the tangency portfolio conditional on a given investor's risk aversion. Basically, our approaches take into account the following investment assumptions: the investor is unconstrained in her investment decision. That is, she has no short constraints, minimum or maximum investment requirements on single assets, and she can borrow arbitrarily at the risk-free rate.

It is widely known the risk and excess return measures used by equations (6) and (7) are based on expected values, indicating that they are statistical statements about the future (the expected vector of excess returns is explicit in equation (6), and implicit in the definition of the covariance matrix). Because these model parameters (μ, Σ) of the true optimal mean–variance portfolio are not directly observed, in practice, investors must substitute the equilibrium parameters based on historical measurements of asset return and volatility for these parameters. Traditionally, under the normality assumption, the maximum likelihood estimators of the return vector and sample covariance matrix (μ, Σ) are

$$\hat{\mu}' = \left(\frac{1}{n}\sum_{t=1}^{n} R_{1t}, \ldots, \frac{1}{n}\sum_{t=1}^{n} R_{kt}\right) \quad \text{and} \quad \tilde{\Sigma} = \frac{1}{n}\sum_{t=1}^{n}(R_t - \bar{R})(R_t - \bar{R})',$$

where $(R_{1t}, R_{2t}, \ldots, R_{kt})$ is the excess return vector of k risky assets at period t, $t = 1, 2, \ldots, n$. Therefore, a natural estimator of w_o is given by

$$\hat{w}_o = \frac{\tilde{\Sigma}^{-1}\hat{\mu}}{\gamma}. \tag{8}$$

The exact distribution of $\hat{w}_o$ is derived by Okhrin and Schmid (2006). Unfortunately, very often such estimates fail to consider model uncertainty relative to the historical data. Best and Grauer (1991) indicate that such estimation errors result in the estimator being substantially different from the true optimal portfolio w_o. With regard to the coefficient of the maximum likelihood estimator $(\hat{\mu}, \tilde{\Sigma})$, Kan and Zhou (2007) obtain an unbiased estimator of w_o as follows:

$$\hat{w}_u = \left(\frac{n}{n-k-2}\right)\frac{\tilde{\Sigma}^{-1}\hat{\mu}}{\gamma} = \left(\frac{n\tilde{\Sigma}^{-1}}{n-k-2}\right)\frac{\hat{\mu}}{\gamma} = \frac{\hat{\Sigma}^{-1}\hat{\mu}}{\gamma}. \tag{9}$$

Note that $\hat{w}_u$ performs rather better than $\hat{w}_o$. Moreover, some useful properties of $\hat{\mu}$, $\tilde{\Sigma}$, and $\hat{\Sigma}$ are cited in the following remark.

Remark 2. Under the multivariate normality:

1. It is well known that $\tilde{\Sigma}$ and $\hat{\mu}$ are independent. The sample mean and sample covariance have the following distribution:

$$\hat{\mu} \sim N\left(\mu, \frac{\Sigma}{n}\right) \quad \text{and} \quad \tilde{\Sigma} \sim \frac{W(n-1, \Sigma)}{n},$$

   where $W(n-1, \Sigma)$ denotes a Wishart distribution with $n-1$ degrees of freedom and covariance matrix Σ.
2. (Haff's identity (1979)) If $n \geq k+4$, then the expectation of $\Sigma\hat{\Sigma}^{-2}\Sigma$ is given by

$$E\big(\Sigma\hat{\Sigma}^{-2}\Sigma\big) = \alpha I_k, \tag{10}$$

   where $\alpha = \frac{(n-2)(n-k-2)}{(n-k-1)(n-k-4)}$ and I_k is the $k \times k$ identity matrix.
3. Kan and Zhou (2007) use these results to derive a closed-form solution for the out-of-sample performance of the optimal mean–variance portfolio. Equation (9) and the independence between $\tilde{\Sigma}$ and $\hat{\mu}$ imply that the expected return of portfolio $\hat{w}_u$ is

$$E(\hat{w}_u'\mu) = E\left(\frac{\hat{\mu}'\hat{\Sigma}^{-1}\mu}{\gamma}\right) = \frac{\mu'\Sigma^{-1}\mu}{\gamma}. \tag{11}$$

The corresponding variance of return relative to portfolio $\hat{w}_u$ follows equation (10):

$$E(\hat{w}_u' \Sigma \hat{w}_u) = E\left(\frac{\hat{\mu}' \Sigma^{-1/2} (\Sigma \hat{\Sigma}^{-2} \Sigma) \Sigma^{-1/2} \hat{\mu}}{\gamma^2}\right) = \frac{\alpha \mu' \Sigma^{-1} \mu}{\gamma^2}. \quad (12)$$

2.2 The combination of sophisticated strategies and naive diversification

For the sample-based strategies, Tu and Zhou (2011) present the following combination of two portfolio rules:

$$w_c = (1 - \beta) w_n + \beta \hat{w}_u,$$

where $\hat{w}_u$ is the unbiased estimator based on historical data in equation (9), w_n is the naive rule that invests equally across k risky assets, and β is the combination coefficient, $0 \leq \beta \leq 1$. The implied portfolio return of w_c at $t + 1$ is

$$R_{c,t+1} = w_c' R_{t+1}.$$

The multivariate normality assumptions of the excess returns of the k risky assets imply that the expected utility of $\hat{w}_c$ is given by

$$U(w_c) = w_c' \mu - \frac{\gamma}{2} w_c' \Sigma w_c,$$

where γ is the mean–variance investor's relative risk aversion coefficient. The objective is to find an optimal combination coefficient β such that the following expected loss is minimized:

$$L(w_o, w_c) = U(w_o) - E(U(w_c)), \quad (13)$$

where $U(w_o)$ is the expected utility of the equilibrium portfolio $w_o = \frac{\Sigma^{-1}\mu}{\gamma}$.

To facilitate further analysis, we reformulate some results from Tu–Zhou optimization, in which we replace w_n by using an arbitrary nonstochastic portfolio weight w_e. We thus form a general combination as follows:

$$w_c = (1 - \beta) w_e + \beta \hat{w}_u, \quad (14)$$

where w_e is any arbitrary nonstochastic portfolio weight that does not depend on the sample information. Consider the following theorem.

Theorem 1 (Some similar results of Tu and Zhou's model). *Under the minimal expected loss of quadratic utility* (13), *the optimal choice of the combination* (14) *is given by*

$$\beta_{TZ} = \frac{B(w_e)}{V(\hat{w}_u) + B(w_e)},$$

$$B(w_e) = E[(w_e - w_o)' \Sigma (w_e - w_o)] = w_e' \Sigma w_e + \frac{\mu' \Sigma^{-1} \mu}{\gamma^2} - \frac{2 w_e' \mu}{\gamma}, \quad (15)$$

$$V(\hat{w}_u) = E[(\hat{w}_u - w_o)' \Sigma (\hat{w}_u - w_o)] = \frac{(\alpha - 1) \mu' \Sigma^{-1} \mu}{\gamma^2},$$

where $B(w_{\mathrm{e}})$ and $V(\hat{w}_{\mathrm{u}})$ measure the bias of $w_{\mathrm{e}}'R_{t+1}$ and the variance of $\hat{w}_{\mathrm{u}}'R_{t+1}$, respectively. The corresponding optimal combination of the naive rule and optimal mean–variance portfolio is

$$w_{\mathrm{TZ}} = (1-\beta_{\mathrm{TZ}})w_{\mathrm{e}} + \beta_{\mathrm{TZ}}\hat{w}_{\mathrm{u}}.$$

The expected loss relative to w_{TZ} is shown as

$$L(w_{\mathrm{o}}, w_{\mathrm{TZ}}) = \left[\frac{V(\hat{w}_{\mathrm{u}})}{V(\hat{w}_{\mathrm{u}}) + B(w_{\mathrm{e}})}\right]^2 B(w_{\mathrm{e}}) + \left[\frac{B(w_{\mathrm{e}})}{V(\hat{w}_{\mathrm{u}}) + B(w_{\mathrm{e}})}\right]^2 V(\hat{w}_{\mathrm{u}}).$$

Moreover, the optimal w_{TZ} dominates both the naive rule and optimal mean–variance portfolio with probability one. That is,

$$L(w_{\mathrm{o}}, w_{\mathrm{TZ}}) \leq \min\big[L(w_{\mathrm{o}}, \hat{w}_{\mathrm{u}}), L(w_{\mathrm{o}}, w_{\mathrm{e}})\big]$$

with probability one.

Proof. It suffices to prove the optimality relative to β_{TZ} in equation (15). First, we derive the optimal combination coefficient. The expected loss of utility function is

$$\begin{aligned}
&E\big[L(w_{\mathrm{o}}, w_{\mathrm{c}})\big] \\
&\quad= \frac{\mu'\Sigma^{-1}\mu}{2\gamma} - E\big[(1-\beta)w_{\mathrm{e}}' + \beta\hat{w}_{\mathrm{u}}'\big]\mu \\
&\qquad+ \frac{\gamma}{2}E\big[\big((1-\beta)w_{\mathrm{e}} + \beta\hat{w}_{\mathrm{u}}\big)'\Sigma\big((1-\beta)w_{\mathrm{e}} + \beta\hat{w}_{\mathrm{u}}\big)\big] \\
&\quad= \frac{\mu'\Sigma^{-1}\mu}{2\gamma} - (1-\beta)w_{\mathrm{e}}'\mu - \beta\frac{\mu'\Sigma^{-1}\mu}{\gamma} \\
&\qquad+ \frac{2\gamma(1-\beta)\beta E[(w_{\mathrm{e}} - w_{\mathrm{o}})'\Sigma(\hat{w}_{\mathrm{u}} - w_{\mathrm{o}})]}{2} + \frac{2\gamma\beta E[(\hat{w}_{\mathrm{u}} - w_{\mathrm{o}})'\Sigma w_{\mathrm{o}}]}{2} \\
&\qquad+ \frac{2\gamma(1-\beta)(w_{\mathrm{e}} - w_{\mathrm{o}})'\Sigma w_{\mathrm{o}}}{2} + \frac{\gamma E[w_{\mathrm{o}}'\Sigma w_{\mathrm{o}}]}{2} \\
&\qquad+ \frac{\gamma(1-\beta)^2 E[(w_{\mathrm{e}} - w_{\mathrm{o}})'\Sigma(w_{\mathrm{e}} - w_{\mathrm{o}})]}{2} + \frac{\gamma\beta^2 E[(\hat{w}_{\mathrm{u}} - w_{\mathrm{o}})'\Sigma(\hat{w}_{\mathrm{u}} - w_{\mathrm{o}})]}{2} \\
&\quad= \frac{\gamma}{2}\big[(1-\beta)^2 E\big[(w_{\mathrm{e}} - w_{\mathrm{o}})'\Sigma(w_{\mathrm{e}} - w_{\mathrm{o}})\big] + \beta^2 E\big[(\hat{w}_{\mathrm{u}} - w_{\mathrm{o}})'\Sigma(\hat{w}_{\mathrm{u}} - w_{\mathrm{o}})\big]\big] \\
&\quad= \frac{\gamma}{2}\big[(1-\beta)^2 B(w_{\mathrm{e}}) + \beta^2 V(\hat{w}_{\mathrm{u}})\big].
\end{aligned}$$

Substituting the components of w_{o} and $\hat{w}_{\mathrm{u}}$ into the $B(w_{\mathrm{e}})$ and $V(\hat{w}_{\mathrm{u}})$ will lead to the following result when we determine the optimal combination coefficient:

$$B(w_{\mathrm{e}}) = w_{\mathrm{e}}'\Sigma w_{\mathrm{e}} - 2w_{\mathrm{e}}'\Sigma w_{\mathrm{o}} + w_{\mathrm{o}}'\Sigma w_{\mathrm{o}} = w_{\mathrm{e}}'\Sigma w_{\mathrm{e}} - \frac{2w_{\mathrm{e}}'\mu}{\gamma} + \frac{\mu'\Sigma^{-1}\mu}{\gamma^2},$$

and

$$V(\hat{w}_{\mathrm{u}}) = E(\hat{w}_{\mathrm{u}}' \Sigma \hat{w}_{\mathrm{u}}) - w_{\mathrm{o}}' \Sigma w_{\mathrm{o}} = \left[\frac{(n-2)(n-k-2)}{(n-k-1)(n-k-4)} - 1 \right] \frac{\mu' \Sigma^{-1} \mu}{\gamma^2}.$$

Second, take the first-order and second-order derivatives of $E[L(w_{\mathrm{o}}, w_{\mathrm{c}})]$ with respect to β; the first-order condition of optimality is

$$\frac{dE[L(w_{\mathrm{o}}, w_{\mathrm{c}})]}{d\beta} = \gamma\big[-(1-\beta)B(w_{\mathrm{e}}) + \beta \operatorname{Var}(\hat{w}_{\mathrm{u}})\big] = 0.$$

Therefore, the optimal combination coefficient of this optimization is

$$\beta_{\mathrm{TZ}} = \frac{B(w_{\mathrm{e}})}{V(\hat{w}_{\mathrm{u}}) + B(w_{\mathrm{e}})}.$$

In addition, the second-order condition of optimality holds, as

$$\frac{d^2 E[L(w_{\mathrm{o}}, w_{\mathrm{c}})]}{d\beta^2} = \gamma\big[B(w_{\mathrm{e}}) + V(\hat{w}_{\mathrm{u}})\big] > 0$$

with probability one. This completes the proof. □

There are some interesting implications in Theorem 1. First, Tu and Zhou adopt a minimal expected loss of mean–variance utility that consider both estimation error ($V(\hat{w}_{\mathrm{u}})$, the variance of the mean–variance portfolio) and error resulting from model misspecification ($B(w_{\mathrm{e}})$, the bias of the naive rule). Note that the combination coefficient β_{TZ} in equation (15) measures the trade-off between the bias and the variance. Second, this combination is not theoretically unbiased, as the naive weight is obviously a biased estimator of the true mean–variance efficient portfolio, but there is a low expected loss of mean–variance utility. Third, Tu and Zhou observe that the combination rule always converges and is theoretically designed to be better than either the naive rule or $\hat{w}_{\mathrm{u}}$.

3 Main results

In this section, we briefly introduce a statistical perspective of Tu and Zhou's model. In combined forecast theory, two standard criteria of determining the optimal combination coefficient are the local minimum-variance (LMV) criterion and the mean squared error (MSE) criterion. The fundamental difference between the MSE criterion and the LMV criterion is that the MSE criterion highlights the comovement between the combined portfolio and the true mean–variance portfolio compared to the LMV criterion, which focuses solely on risk reduction.

3.1 The optimal combinations under the MSE and LMV criteria

With respect to the linear combination $w_c = (1-\beta)w_e + \beta\hat{w}_u$, we now introduce a statistical perspective of Tu and Zhou's model under minimizing the MSE criterion. The mean squared error of w_c, denoted by $\mathrm{MSE}(w_c)$, is defined as

$$\begin{aligned}\mathrm{MSE}(w_c) &= E\big[(1-\beta)w_e' R_{n+1} + \beta\hat{w}_u' R_{n+1} - w_o' R_{n+1}\big]^2 \\ &= E\big[\beta(\hat{w}_u - w_e)' R_{n+1} + (w_e - w_o)' R_{n+1}\big]^2 \\ &= \beta^2 E\big[(\hat{w}_u - w_e)'\Sigma(\hat{w}_u - w_e)\big] + (w_e - w_o)'\Sigma(w_e - w_o) \\ &\quad + 2\beta E\big[(\hat{w}_u - w_e)'\Sigma(w_e - w_o)\big].\end{aligned} \tag{16}$$

For the optimization (16) to exist, a sufficient condition that describes the permissible combination coefficient must hold. The following theorem shows that the equivalence between the economic view (the minimal expected loss of the mean–variance utility) and the statistical perspective (the minimal mean squared errors criterion).

Theorem 2 (The equivalence between the minimal expected loss of the mean–variance utility and the minimal mean squared errors criterion).

$$\beta_{\mathrm{MSE}} = \frac{E(w_e'\Sigma w_e) - E(\hat{w}_e'\Sigma\hat{w}_u) + [E(\hat{w}_u'\Sigma w_o) - E(w_e'\Sigma w_o)]}{E(w_e'\Sigma w_e) + E(\hat{w}_u'\Sigma\hat{w}_u) - 2E(w_e'\Sigma\hat{w}_u)} = \beta_{\mathrm{TZ}}.$$

Proof. The first-order condition of optimality in equation (16) is

$$2\beta E\big[(\hat{w}_u - w_e)'\Sigma(\hat{w}_u - w_e)\big] + 2E\big[(\hat{w}_u - w_e)'\Sigma(w_e - w_o)\big] = 0.$$

To solve this equation, we obtain that the optimal combination coefficient of equation (16) is

$$\begin{aligned}\beta_{\mathrm{MSE}} &= -\frac{E[(\hat{w}_u - w_e)'\Sigma(w_e - w_o)]}{E[(\hat{w}_u - w_e)'\Sigma(\hat{w}_u - w_e)]} \\ &= \frac{E(w_e'\Sigma w_e) - E(w_e'\Sigma\hat{w}_u) + [E(\hat{w}_u'\Sigma w_o) - E(w_e'\Sigma w_o)]}{E(w_e'\Sigma w_e) + E(\hat{w}_u'\Sigma\hat{w}_u) - 2E(w_e'\Sigma\hat{w}_u)} \\ &= \frac{w_e'\Sigma w_e - w_e'\mu/\gamma + (E(\hat{w}_u'\mu) - E(w_e'\mu))/\gamma}{w_e'\Sigma w_e + \alpha\mu'\Sigma\mu/\gamma^2 - 2w_e'\mu/\gamma} \\ &= \frac{w_e'\Sigma w_e + \mu'\Sigma\mu/\gamma^2 - 2w_e'\mu/\gamma}{w_e'\Sigma w_e + \alpha\mu'\Sigma\mu/\gamma^2 - 2w_e'\mu/\gamma} \\ &= \beta_{\mathrm{TZ}}.\end{aligned}$$

This completes the proof. □

Under the minimum-variance criterion of a combined portfolio, the objective is to find an optimal combination coefficient β such that the following variance is minimized:

$$\begin{aligned}\mathrm{Var}(w_c' R_{n+1}) &= (1-\beta)^2 E(w_e' \Sigma w_e) \\ &\quad + \beta^2 E(\hat{w}_u' \Sigma \hat{w}_u) + 2\beta(1-\beta) E(w_e' \Sigma \hat{w}_u).\end{aligned} \tag{17}$$

Theorem 2 provides a simple way to reach the optimal combination coefficient of the optimization (17) as follows.

Theorem 3. *The optimal combination coefficient related to the minimum-variance optimization* (17) *is given by*

$$\begin{aligned}\beta_{\mathrm{LMV}} &= \frac{E(w_e' \Sigma w_e) - E(w_e' \Sigma \hat{w}_u)}{E(w_e' \Sigma w_e) + E(\hat{w}_u' \Sigma \hat{w}_u) - 2E(w_e' \Sigma \hat{w}_u)} \\ &= \frac{w_e' \Sigma w_e - w_e' \mu/\gamma}{w_e' \Sigma w_e + \alpha \mu' \Sigma \mu/\gamma^2 - 2 w_e' \mu/\gamma}.\end{aligned} \tag{18}$$

Note that the LMV combined portfolio ignores the difference between the return of $\hat{w}_u$ and the return of w_e, $E(\hat{w}_u' \mu) - E(w_e' \mu)$. We find that the optimal design using the MSE criterion appears to include as a special case the combination of the naive rule and the optimal mean–variance strategy based on minimizing the combined variance.

Because μ and Σ in equations (15) and (18) are not observable, to estimate β_{LTV} and β_{MSE}, we replace them by the MLE sample counterpart, $\hat{\mu}$ and $\tilde{\Sigma}$. Consider the following theorem.

Theorem 4 (The estimated combination coefficient under the LMV and MSE criteria). *The estimated optimal combination coefficient related to the* β_{MSE} *is given by*

$$\hat{\beta}_{\mathrm{MSE}} = \frac{\gamma^2 w_e' \tilde{\Sigma} w_e + \hat{\mu}' \tilde{\Sigma}^{-1} \hat{\mu} - 2\gamma w_e' \hat{\mu}}{\gamma^2 w_e' \tilde{\Sigma} w_e + \alpha \hat{\mu}' \tilde{\Sigma}^{-1} \hat{\mu} - 2\gamma w_e' \hat{\mu}}. \tag{19}$$

The estimated optimal combination of the naive rule and optimal mean–variance portfolio is

$$\hat{w}_{\mathrm{MSE}} = (1 - \hat{\beta}_{\mathrm{MSE}}) w_e + \hat{\beta}_{\mathrm{MSE}} \hat{w}_u. \tag{20}$$

Similarly, the estimated optimal combination coefficient related to the β_{LMV} *is given by*

$$\hat{\beta}_{\mathrm{LMV}} = \frac{\gamma^2 w_e' \tilde{\Sigma} w_e - \gamma w_e' \hat{\mu}}{\gamma^2 w_e' \tilde{\Sigma} w_e + \alpha \hat{\mu}' \tilde{\Sigma}^{-1} \hat{\mu} - 2\gamma w_e' \hat{\mu}}. \tag{21}$$

The estimated optimal combination of the naive rule and optimal mean–variance portfolio is

$$\hat{w}_{\mathrm{LMV}} = (1 - \hat{\beta}_{\mathrm{LMV}})w_{\mathrm{e}} + \hat{\beta}_{\mathrm{LMV}}\hat{w}_{\mathrm{u}}. \tag{22}$$

3.2 An application: The optimal hedging strategies

Naive hedging is the most intuitive strategy in the futures markets and provides a benchmark against which more sophisticated models may be compared. The so-called naive hedging is defined as a portfolio for which an investor takes an equal but opposite position of spot assets in futures contracts. That is, the hedging ratio of the naive hedging is defined as $h = -1$. In this section, we propose an optimal combination of naive hedging and mean–variance efficient hedging (or the maximal risk-return hedging) as follows:

$$w_{\mathrm{c}} = (1-\beta)w_{\mathrm{e}} + \beta\hat{w}_{\mathrm{u}} = (1-\beta)\begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix} + \beta\begin{bmatrix} \hat{w}_{\mathrm{f}} \\ \hat{w}_{\mathrm{s}} \end{bmatrix}. \tag{23}$$

Or, the optimal combination may be directly implemented in the Tu and Zhou's framework by taking a negative returns for the futures contracts:

$$w_{\mathrm{c}} = (1-\beta)w_{\mathrm{e}} + \beta\hat{w}_{\mathrm{u}} = \frac{(1-\beta)}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \beta\begin{bmatrix} -\hat{w}_{\mathrm{f}} \\ \hat{w}_{\mathrm{s}} \end{bmatrix}. \tag{24}$$

Because the optimal portfolio weight vector w_{c} is unique, the relative magnitude of portfolio weight vector w_{e} does not affect the optimal combination coefficient β.

The statistical inferences of the hedging effectiveness of this optimal combination will be made using Chiu's (2013) regression-based approach. Assume that the investor considers a hedged portfolio by holding w_{f} units futures contracts and a spot position of w_{s} units. Let the in-sample data be substituted into an unrestricted OLS regression as follows:

$$\underset{(\ell)}{\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}} = \underset{(R)}{\begin{bmatrix} r_{\mathrm{f}1} & r_{\mathrm{s}1} \\ r_{\mathrm{f}2} & r_{\mathrm{s}2} \\ \vdots & \vdots \\ r_{\mathrm{f}n} & r_{\mathrm{s}n} \end{bmatrix}} \underset{(w)}{\begin{bmatrix} w_{\mathrm{f}} \\ w_{\mathrm{s}} \end{bmatrix}} + \underset{(\varepsilon)}{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}}, \tag{25}$$

where $w = [w_{\mathrm{f}} \;\; w_{\mathrm{s}}]'$ represents the portfolio weight vector, $\ell = (1, \ldots, 1)'$ is the $n \times 1$ vector of ones, the n observations of the hedged residuals are contained in the $n \times 1$ vector ε and the observations of futures excess returns and spot excess returns are contained in the $n \times 2$ matrix R.

For the estimation procedures of this regression-based model, Chiu (2013) shows that the unscaled OLS estimator of the portfolio weight, the unrestricted sum of squared residuals, and the corresponding optimal mean–variance hedging ratio are

$$\tilde{w}_{\text{OLS}} = \begin{bmatrix} \tilde{w}_{\text{f}} \\ \tilde{w}_{\text{s}} \end{bmatrix} = \frac{\tilde{\Sigma}^{-1}\hat{\mu}}{1+\hat{\mu}'\tilde{\Sigma}^{-1}\hat{\mu}}, \qquad SSR_{\text{mv}} = \frac{n}{1+\hat{\mu}'\tilde{\Sigma}^{-1}\hat{\mu}}, \qquad \hat{h}_{\text{mv}} = \frac{\tilde{w}_{\text{f}}}{\tilde{w}_{\text{s}}}, \tag{26}$$

where $\hat{\mu} = \left[\begin{smallmatrix} \bar{r}_{\text{f}} \\ \bar{r}_{\text{s}} \end{smallmatrix}\right]$ and $\tilde{\Sigma} = \left[\begin{smallmatrix} s_{\text{f}}^2 & s_{\text{sf}} \\ s_{\text{sf}} & s_{\text{s}}^2 \end{smallmatrix}\right]$ represent the vector of sample means of excess returns and the sample covariance matrix between excess returns, respectively. Note that the OLS estimator, $\tilde{w}_{\text{OLS}}$, is independent of the investor's CARA parameter. If an investor employs the naive rule where $w_{\text{f}} = -w_{\text{s}}$, the restricted in-sample regression of the naive hedging may be represented by

$$\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} r_{\text{f1}} - r_{\text{s1}} \\ r_{\text{f2}} - r_{\text{s2}} \\ \vdots \\ r_{\text{f}n} - r_{\text{s}n} \end{bmatrix} w_{\text{f}} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}. \tag{27}$$

The sum of squared residuals of the naive rule is formed as

$$SSR_{\text{nr}} = \frac{n}{1+n(\bar{r}_{\text{f}}-\bar{r}_{\text{s}})^2/\sum_{i=1}^{n}(r_{\text{f}i}-r_{\text{s}i})^2} = \frac{n[(s_{\text{f}}^2 - 2s_{\text{fs}} + s_{\text{s}}^2) + (\bar{r}_{\text{f}}-\bar{r}_{\text{s}})^2]}{(s_{\text{f}}^2 - 2s_{\text{fs}} + s_{\text{s}}^2) + 2(\bar{r}_{\text{f}}-\bar{r}_{\text{s}})^2}.$$

This finding implies that

$$w_{\text{e}}'\tilde{\Sigma}w_{\text{e}} = \left(s_{\text{f}}^2 - 2s_{\text{fs}} + s_{\text{s}}^2\right) = \frac{(\bar{r}_{\text{f}}-\bar{r}_{\text{s}})^2(n-2SSR_{\text{nr}})}{SSR_{\text{nr}}-n}. \tag{28}$$

As a finding of previous results, we now proceed to construct the optimal combination between the risk-return hedging and naive hedging. The corresponding hedging ratio is described as the following theorem.

Theorem 5. *The regression-based optimal combination coefficient of β_{MSE} is given by*

$$\begin{aligned} \hat{\beta}_{\text{MSE}} &= \big(\gamma^2(\bar{r}_{\text{f}}-\bar{r}_{\text{s}})^2(n-2SSR_{\text{nr}})SSR_{\text{mv}} - (n-SSR_{\text{mv}})(n-SSR_{\text{nr}}) \\ &\quad - 2\gamma(\bar{r}_{\text{f}}-\bar{r}_{\text{s}})(n-SSR_{\text{nr}})SSR_{\text{mv}}\big)/\big(\gamma^2(\bar{r}_{\text{f}}-\bar{r}_{\text{s}})^2(n-2SSR_{\text{nr}})SSR_{\text{mv}} \\ &\quad - \alpha(n-SSR_{\text{mv}})(n-SSR_{\text{nr}}) - 2\gamma(\bar{r}_{\text{f}}-\bar{r}_{\text{s}})(n-SSR_{\text{nr}})SSR_{\text{mv}}\big). \end{aligned} \tag{29}$$

The corresponding optimal combination of the naive rule and optimal mean–variance portfolio is

$$\hat{w}_{\text{MSE}} = (1-\hat{\beta}_{\text{MSE}})w_{\text{e}} + \hat{\beta}_{\text{MSE}}\hat{w}_{\text{u}}.$$

The hedging ratio relative to $\hat{w}_{\mathrm{MSE}}$ *is shown as*

$$\hat{h}_{\mathrm{MSE}} = \frac{(\hat{\beta}_{\mathrm{MSE}} - 1) + \hat{\beta}_{\mathrm{MSE}}\hat{w}_{\mathrm{f}}}{(1 - \hat{\beta}_{\mathrm{MSE}}) + \hat{\beta}_{\mathrm{MSE}}\hat{w}_{\mathrm{s}}} = \frac{(1 - 1/\hat{\beta}_{\mathrm{MSE}})1/\hat{w}_{\mathrm{s}} + \tilde{h}_{\mathrm{mv}}}{(1/\hat{\beta}_{\mathrm{MSE}} - 1)1/\hat{w}_{\mathrm{s}} + 1}. \tag{30}$$

Similarly, the regression-based optimal combination coefficient of β_{LMV} *is given by*

$$\begin{aligned}\hat{\beta}_{\mathrm{LMV}} & \\ &= \big(\gamma^2(\bar{r}_{\mathrm{f}} - \bar{r}_{\mathrm{s}})^2(n - 2SSR_{\mathrm{nr}})SSR_{\mathrm{mv}} - \gamma(\bar{r}_{\mathrm{f}} - \bar{r}_{\mathrm{s}})(n - SSR_{\mathrm{nr}})SSR_{\mathrm{mv}}\big) \\ &\quad /\big(\gamma^2(\bar{r}_{\mathrm{f}} - \bar{r}_{\mathrm{s}})^2(n - 2SSR_{\mathrm{nr}})SSR_{\mathrm{mv}} - \alpha(n - SSR_{\mathrm{mv}})(n - SSR_{\mathrm{nr}}) \\ &\quad - 2\gamma(\bar{r}_{\mathrm{f}} - \bar{r}_{\mathrm{s}})(n - SSR_{\mathrm{nr}})SSR_{\mathrm{mv}}\big).\end{aligned} \tag{31}$$

The corresponding optimal combination of the naive rule and optimal mean–variance portfolio is

$$\hat{w}_{\mathrm{LMV}} = (1 - \hat{\beta}_{\mathrm{LMV}})w_{\mathrm{e}} + \hat{\beta}_{\mathrm{LMV}}\hat{w}_{\mathrm{u}}.$$

The hedging ratio relative to $\hat{w}_{\mathrm{LMV}}$ *is shown as*

$$\hat{h}_{\mathrm{LMV}} = \frac{(\hat{\beta}_{\mathrm{LMV}} - 1) + \hat{\beta}_{\mathrm{LMV}}\hat{w}_{\mathrm{f}}}{(1 - \hat{\beta}_{\mathrm{LMV}}) + \hat{\beta}_{\mathrm{LMV}}\hat{w}_{\mathrm{s}}} = \frac{(1 - 1/\hat{\beta}_{\mathrm{LMV}})1/\hat{w}_{\mathrm{s}} + \tilde{h}_{\mathrm{mv}}}{(1/\hat{\beta}_{\mathrm{LMV}} - 1)1/\hat{w}_{\mathrm{s}} + 1}. \tag{32}$$

Proof. To obtain an unbiased estimator of the optimal mean–variance portfolio w_{o}, we must adjust the coefficient of the OLS weight, which may be accomplished by combining equations (9) and (26)

$$\begin{aligned}\hat{w}_{\mathrm{u}} &= \frac{\hat{\Sigma}^{-1}\hat{\mu}}{\gamma} = \left[\frac{n(1 + \hat{\mu}'\tilde{\Sigma}^{-1}\hat{\mu})}{(n - k - 2)\gamma}\right]\left[\frac{\tilde{\Sigma}^{-1}\bar{r}}{1 + \hat{\mu}'\tilde{\Sigma}^{-1}\hat{\mu}}\right] \\ &= \left[\frac{n^2}{\gamma(n - k - 2)SSR_{\mathrm{mv}}}\right]\tilde{w}_{\mathrm{OLS}}.\end{aligned} \tag{33}$$

In addition, we compute the Sharpe ratio $\hat{\mu}'\tilde{\Sigma}^{-1}\hat{\mu}$ using equation (26)

$$\hat{\mu}'\tilde{\Sigma}^{-1}\hat{\mu} = \frac{n - SSR_{\mathrm{mv}}}{SSR_{\mathrm{mv}}}.$$

For the case of minimizing the expected loss of mean–variance utility, the optimal combination coefficient is given by

$$\begin{aligned}\hat{\beta}_{\mathrm{TZ}} = &\left(\frac{\gamma^2(\bar{r}_{\mathrm{f}} - \bar{r}_{\mathrm{s}})^2(n - 2SSR_{\mathrm{nr}})}{SSR_{\mathrm{nr}} - n} + \frac{n - SSR_{\mathrm{mv}}}{SSR_{\mathrm{mv}}} - 2\gamma(\bar{r}_{\mathrm{s}} - \bar{r}_{\mathrm{f}})\right) \\ &\Big/\left(\frac{\gamma^2(\bar{r}_{\mathrm{f}} - \bar{r}_{\mathrm{s}})^2(n - 2SSR_{\mathrm{nr}})}{SSR_{\mathrm{nr}} - n} + \alpha\frac{n - SSR_{\mathrm{mv}}}{SSR_{\mathrm{mv}}} - 2\gamma(\bar{r}_{\mathrm{s}} - \bar{r}_{\mathrm{f}})\right)\end{aligned}$$

$$\begin{aligned}&= \big(\gamma^2(\bar{r}_{\mathrm{f}}-\bar{r}_{\mathrm{s}})^2(n-2SSR_{\mathrm{nr}})SSR_{\mathrm{mv}}-(n-SSR_{\mathrm{mv}})(n-SSR_{\mathrm{nr}})\\ &\quad -2\gamma(\bar{r}_{\mathrm{f}}-\bar{r}_{\mathrm{s}})(n-SSR_{\mathrm{nr}})SSR_{\mathrm{mv}}\big)/\big(\gamma^2(\bar{r}_{\mathrm{f}}-\bar{r}_{\mathrm{s}})^2(n-2SSR_{\mathrm{nr}})SSR_{\mathrm{mv}}\\ &\quad -\alpha(n-SSR_{\mathrm{mv}})(n-SSR_{\mathrm{nr}})-2\gamma(\bar{r}_{\mathrm{f}}-\bar{r}_{\mathrm{s}})(n-SSR_{\mathrm{nr}})SSR_{\mathrm{mv}}\big).\end{aligned}$$

The Tu–Zhou combined portfolio is thus determined as

$$\hat{w}_{\mathrm{MSE}} = (1-\hat{\beta}_{\mathrm{MSE}})\begin{bmatrix}-1\\1\end{bmatrix}+\hat{\beta}_{\mathrm{MSE}}\begin{bmatrix}\hat{w}_{\mathrm{f}}\\ \hat{w}_{\mathrm{s}}\end{bmatrix}.$$

Therefore, the hedging ratio relative to $\hat{w}_{\mathrm{MSE}}$ is shown as

$$\hat{h}_{\mathrm{MSE}} = \frac{(\hat{\beta}_{\mathrm{MSE}}-1)+\hat{\beta}_{\mathrm{MSE}}\hat{w}_{\mathrm{f}}}{(1-\hat{\beta}_{\mathrm{MSE}})+\hat{\beta}_{\mathrm{MSE}}\hat{w}_{\mathrm{s}}} = \frac{(1-1/\hat{\beta}_{\mathrm{MSE}})1/\hat{w}_{\mathrm{s}}+\tilde{h}_{\mathrm{mv}}}{(1/\hat{\beta}_{\mathrm{MSE}}-1)1/\hat{w}_{\mathrm{s}}+1}.$$

Similarly, to slightly modify the coefficient, we obtain the resulting properties of the minimum-variance (LMV) hedging. This completes the proof. □

Investors may wish to assess whether the out-of-sample performance relative to the hedging policy $\hat{h}_{\mathrm{MSE}}$ (or $\hat{h}_{\mathrm{LMV}}$) is effective. The null hypothesis is thus stated as follows:

$$H_0\colon\ h=\hat{h}_{\mathrm{MSE}} \qquad (\text{or } H_0\colon\ h=\hat{h}_{\mathrm{LMV}}). \tag{34}$$

For the testing procedures of this regression-based model, the restricted out-of-sample regression under the null hypothesis $H_0\colon\ h=\hat{h}_{\mathrm{MSE}}$, is thus reduced into a simple regression as

$$\begin{bmatrix}1\\1\\ \vdots\\1\end{bmatrix} = \begin{bmatrix}\hat{h}_{\mathrm{MSE}}\cdot r_{\mathrm{f}(n+1)}+r_{\mathrm{s}(n+1)}\\ \hat{h}_{\mathrm{MSE}}\cdot r_{\mathrm{f}(n+2)}+r_{\mathrm{s}(n+2)}\\ \vdots\\ \hat{h}_{\mathrm{MSE}}\cdot r_{\mathrm{f}(n+m)}+r_{\mathrm{s}(n+m)}\end{bmatrix} w_{\mathrm{s}} + \begin{bmatrix}\varepsilon_{n+1}\\ \varepsilon_{n+2}\\ \vdots\\ \varepsilon_{n+m}\end{bmatrix}, \tag{35}$$

where the m observations of excess returns are contained in the $m \times 1$ matrix. Similarly, the ex post evaluation of the mean–variance efficient portfolio may be conducted following the same procedures. The unrestricted out-of-sample regression of the hedged portfolio is represented by

$$\begin{bmatrix}1\\1\\ \vdots\\1\end{bmatrix} = \begin{bmatrix}r_{\mathrm{f}(n+1)} & r_{\mathrm{s}(n+1)}\\ r_{\mathrm{f}(n+1)} & r_{\mathrm{s}(n+2)}\\ \vdots & \vdots\\ r_{\mathrm{f}(n+m)} & r_{\mathrm{s}(n+m)}\end{bmatrix}\begin{bmatrix}w_{\mathrm{f}}\\ w_{\mathrm{s}}\end{bmatrix} + \begin{bmatrix}\varepsilon_{n+1}\\ \varepsilon_{n+2}\\ \vdots\\ \varepsilon_{n+m}\end{bmatrix}. \tag{36}$$

Under the multivariate normality assumption (3), Chiu (2013) shows that assessing the hedging effectiveness (36) is possible using the following F(OLS)-statistic:

$$F = \frac{(SSR_{\mathrm{r}}-SSR_{\mathrm{u}})}{SSR_{\mathrm{u}}/(m-2)} \sim F(1,m-2), \tag{37}$$

where SSR_{r} and SSR_{u} denote the restricted sum of squared residuals (35) and the unrestricted sum of squared residuals (36), respectively. The test statistic F is distributed as a central $F(1, m-2)$ with 1 and $m-2$ degrees of freedom. Similarly, to slightly modify the coefficient, we obtain the testing properties of the minimum-variance (LMV) hedging.

3.3 An illustration

In this section, we perform a comparative out-of-sample performance analysis of the optimal combination approaches derived in the previous section against the ex ante hedging ratio ($\hat{h}_{\mathrm{mv}}$) and naive hedging ratio ($h = -1$) via the F-test in equation (37). The data employed in this illustration comprise 149 weekly observations on the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) and stock index futures contract (TAIEX Futures) over the period 2010–2012 which includes some important financial crises. Days corresponding to Taiwan public holidays are removed from the series to avoid the computation of zero returns. The futures contract "TAIEX Futures" is actively traded, and the underlying investment asset is the TAIEX, issued by the Taiwan Stock Exchange Corporation (TWSE). Data sources are the weekly percentage returns from the TEJ database, which is maintained by the Taiwan Economic Journal.

For further statistical inferences, we divide our sample into two subsamples: the in-sample part which includes January 2010–December 2011 (with 100 weekly observations), is used to generate the initial expected returns and covariance estimation; and the out-of-sample part, which contains January 2012–December 2012 (with 49 weekly observations), is then used as the ex post to assess the hedging effectiveness of the estimation. Both the in-sample estimation and out-of-sample testing are performed using weekly closing prices of TAIEX Futures and TAIEX. Data sources are the weekly percentage returns from the TEJ database, which is maintained by the Taiwan Economic Journal.

The statistics of the in-sample estimation are computed according to the unrestricted regression (25) and restricted regression (27). We first generate the estimation of expected returns and covariance matrix which are displayed in panel A of Table 1. Similarly, the summary statistics and regression results for the out-of-sample testing (January 2012–December 2012, with 49 weekly observations) are reported in panel B of Table 1.

In our case study of the out-of-sample performance, we consider eight portfolios whose hedging ratios are in the second column of panel B of Table 1. Three of these portfolios illustrate the case when the Tu–Zhou optimal MSE combination is applied to pool the risk-return hedging and naive rule (with different risk aversion coefficients $\gamma = 0.5$, $\gamma = 1$, and $\gamma = 2$). Three of these portfolios illustrate the case when the minimum-variance (LMV) combination is used to pool the risk-return hedging and naive rule (with different risk aversion coefficients ($\gamma = 0.5$, $\gamma = 1$ and $\gamma = 2$)), while the other two portfolios illustrate the ex ante optimal

Table 1 *Summary statistics* (*January* 2010–*December* 2012)

Panel A: The in-sample estimation (January 2010–December 2011)

Mean return	Covariance matrix	Ex ante optimal weight
$\hat{\mu}_{\text{in}} = \begin{bmatrix} -0.0962 \\ -0.1255 \end{bmatrix}$	$\tilde{\Sigma}_{\text{in}} = \begin{bmatrix} 8.10 & 7.18 \\ 7.18 & 6.85 \end{bmatrix}$	$\tilde{w}_{\text{OLS}} = \begin{bmatrix} 0.0601 \\ -0.0812 \end{bmatrix}$
Regression (25): $SSR_{\text{mv}} = 99.55965$		Regression (27): $SSR_{\text{nr}} = 99.856778$

Panel B: The out-of-sample testing (January 2012–December 2012)

Mean return	Covariance matrix	Ex ante optimal weight
$\hat{\mu}_{\text{out}} = \begin{bmatrix} 0.19309 \\ 0.28224 \end{bmatrix}$	$\tilde{\Sigma}_{\text{out}} = \begin{bmatrix} 6.25 & 5.13 \\ 5.13 & 4.65 \end{bmatrix}$	$\tilde{w}_{\text{out}} = \begin{bmatrix} -0.18680 \\ 0.26405 \end{bmatrix}$

Models	$\hat{h}$	SSR	F-value	Sharpe ratio	Deduction
Method $A_{(\gamma=1)}$	−0.7399	47.136160	0.02058	0.196811	−0.005713
Method $A_{(\gamma=2)}$	−0.7402	47.136570	0.02099	0.196789	−0.005820
Method $A_{(\gamma=1/2)}$	−0.7394	47.135516	0.01994	0.196847	−0.005528
Method $B_{(\gamma=1)}$	−0.2870	47.893367	0.77593	0.150448	−0.239934
Method $B_{(\gamma=2)}$	−0.3391	47.815384	0.69814	0.155786	−0.212968
Method $B_{(\gamma=1/2)}$	−0.1641	48.039512	0.92172	0.139949	−0.292976
Method C	−0.7405	47.137027	0.02145	0.196764	−0.005947
Method D	−1.0000	48.411588	1.29290	0.109116	−0.448744
Method E	−0.7074	47.115529		0.197941	

Method $A_{(\gamma=1)}$: The Tu–Zhou (MSE) optimal combination (regression (35), H_0: $h = -0.7399$).
Method $A_{(\gamma=2)}$: The Tu–Zhou (MSE) optimal combination (regression (35), H_0: $h = -0.7402$).
Method $A_{(\gamma=1/2)}$: The Tu–Zhou (MSE) optimal combination (regression (35), H_0: $h = -0.7394$).
Method $B_{(\gamma=1)}$: The minimum-variance (LMV) combination (regression (35), H_0: $h = -0.2870$).
Method $B_{(\gamma=2)}$: The minimum-variance (LMV) combination (regression (35), H_0: $h = -0.3391$).
Method $B_{(\gamma=1/2)}$: The minimum-variance (LMV) combination (regression (35), H_0: $h = -0.1641$).
Method C: The ex ante mean–variance portfolio (regression (35), H_0: $h = -0.7405$).
Method D: The naive hedging strategy (regression (35), H_0: $h = -1$).
Method E: The benchmark, the ex post optimal mean–variance portfolio (regression (36)).
The F-value is computed according to the F-test of equation (37).

risk-return hedging and the naive diversification, respectively. We use the ex post optimal risk-return hedging (method E) in Table 1 as a benchmark for comparison. The regression analysis is based on comparing equation (35) against equation (36).

We summarize some important results of our discussions as follows:

1. The sums of the squared residuals of the ex ante and ex post optimal risk-return hedging are 47.1370 and 47.1155, respectively. Using equation (35), we obtain that the F-value (p-value) of testing the naive hedging ratio H_0: $h = -1$ is

given by

$$F = \frac{(SSR_r - SSR_u)}{SSR_u/(n-2)} = \frac{(48.4116 - 47.1155)}{47.1155/47} = 1.2929\ (0.2613).$$

Similarly, the F-value (p-value) of testing the ex ante optimal risk-return hedging ratio H_0: $h = \hat{h}_{\mathrm{LMV}} = -0.7405$ is

$$F = \frac{(47.1370 - 47.1155)}{47.1155/47} = 0.0215\ (0.8841).$$

Note that the risk tolerance affects the optimal risk-return portfolio weights but does not affect the hedging ratio. That is, the optimal hedging ratio depends on $\hat{\Sigma}^{-1}\hat{\mu}$ only.

2. As mentioned previously, any estimation of the true optimal mean–variance portfolio must consider the CARA investor's risk aversion coefficient. We now turn our focus to the adjustment and combination of the mean–variance portfolio. For example, we first assume that the investor chooses his risk aversion parameter $\gamma = 1$. Therefore, by using equations (20), (22) and (33), we obtain the resulting portfolio weights:

$$\hat{w}_u = \begin{bmatrix} 0.0629 \\ -0.0849 \end{bmatrix}, \qquad \hat{w}_{\mathrm{MSE}} = \begin{bmatrix} 0.0627 \\ -0.0847 \end{bmatrix}, \qquad \hat{w}_{\mathrm{LMV}} = \begin{bmatrix} 0.0084 \\ -0.0293 \end{bmatrix}.$$

$$(\hat{\beta} = 1) \qquad (\hat{\beta}_{\mathrm{MSE}} = 0.9998) \qquad (\hat{\beta}_{\mathrm{LMV}} = 0.9488)$$

$$(\hat{h}_{\mathrm{mv}} = -0.7405) \qquad (\hat{h}_{\mathrm{MSE}} = -0.7399) \qquad (\hat{h}_{\mathrm{LMV}} = -0.2870).$$

In this case, $\gamma = 1$, we test the hedging effectiveness of an MSE hedging ratio H_0: $h = \hat{h}_{\mathrm{MSE}} = -0.7399$ via the F-statistic:

$$F = \frac{(47.1361 - 47.1155)}{47.1155/47} = 0.02058\ (0.8865).$$

The F-value indicates that the hedging effectiveness at the hedging ratio $\hat{h}_{\mathrm{MSE}}$ is not significantly different from the ex post optimal risk-return hedging with a p-value 88.65%. That is, the MSE hedging ratio -0.7399 (with $\gamma = 1$) is an excellent estimate of the ex post optimal risk-return hedging. In addition, the F-value (p-value) of testing the hedging effectiveness of a LMV hedging ratio H_0: $h = \hat{h}_{\mathrm{LMV}} = -0.2870$ is:

$$F = \frac{(47.8934 - 47.1155)}{47.1155/47} = 0.7759\ (0.3829).$$

This shows that the hedging effectiveness at the hedging ratio $\hat{h}_{\mathrm{LMV}}$ is also not significantly different from the ex post optimal risk-return hedging with a p-value 38.29%. However, in terms of the Sharpe ratios, the MSE hedging is obviously superior to the LMV hedging.

3. Moreover, we also use two risk aversion coefficients ($\gamma = 0.5$ and $\gamma = 2$) as a simple robustness check. It should be emphasized that the Sharpe ratios of the MSE hedgings with respect to $\gamma = 0.5$, $\gamma = 1$, and $\gamma = 2$ are very close to the Sharpe ratios of the ex post optimal risk-return hedging. Finally, Table 1 presents the summary statistics for the analyses of interest. For comparison, the hedging techniques in panel B present decreases of 0.0578%, 0.0582%, 0.055%, 23.99%, 21.30%, 29.30%, 0.0595% and 44.87% in the Sharpe ratio from the ex post optimal mean–variance efficient portfolio. It is evident that the Sharpe ratios for both the MSE criterion and the LMV criterion are better than the naive hedging in our studies. This indicates that all hedging techniques outperform the naive hedging. Specifically, our illustration also shows out-of-sample performance of a combined MSE hedging that is superior to that of other methods. In fact, all of the MSE combination coefficients seem to be close to the ex ante optimal hedging. This means that the out-of-sample performance of the combined MSE hedging strategies is similar to the out-of-sample performance of the ex ante mean–variance portfolio.

4 Conclusion

It is imperative to understand that no estimation of the mean–variance efficient portfolio produces a consistent out-of-sample portfolio evaluation. Regardless of the decision made, the investor is always giving something up in return. An example of a difficult trade-off is the choice between the estimation error (the variance of the mean–variance portfolio) and the simplicity of using the naive rule at the cost of resulting in model misspecification (the bias of the naive rule). For this reason, an intuitive approach to minimizing risk is to construct a combination between the mean–variance portfolio and the naive rule. According to this view, we analytically derive an optimal combination of the unbiased mean–variance efficient portfolio and the naive rule while minimizing the mean squared error criterion. The presented combining estimators considerably reduce the out-of-sample variance of the portfolio return compared to the mean–variance estimator. An empirical study of the combination of optimal risk-return hedging and the naive hedging demonstrates the superiority of this approach over naive hedging and ex ante mean–variance estimator. To incorporate the Bayesian framework and posteriori methods into the combination procedures is for further studies.

Acknowledgments

I am grateful to the editor and referees for helpful comments and suggestions. The views in this paper are solely the responsibility of the author.

References

Bates, J. M. and Granger, C. W. J. (1969). The combination of forecasts. *Operational Research Quarterly* **20**, 451–468.

Best, M. J. and Grauer, R. R. (1991). On the sensitivity of mean-variance portfolios to change in asset means: Some analytical and computation results, and the structure of asset expected returns. *Review of Financial Studies* **4**, 315–342.

Black, F. and Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal* **48**, 28–43.

Chiu, W. Y. (2013). A simple test of optimal hedging policy. *Statistics and Probability Letters* **83**, 1062–1070. MR3041376

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting* **5**, 559–583.

Chopra, V. K. and Ziemba, W. T. (1993). The effects of errors in means, variance, and covariance on optimal portfolio choice. *Journal of Portfolio Management* **19**, 6–11.

Das, S., Markowitz, H., Scheidm, J. and Statman, M. (2010). Portfolio optimization with mental accounts. *Journal of Financial and Quantitative Analysis* **45**, 311–334.

DeMiguel, V., Garlappi, L., Nogales, F. and Uppal, R. (2009a). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management science* **55**, 798–812.

DeMiguel, V., Garlappi, L. and Uppal, R. (2009b). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies* **22**, 1915–1953.

Duchin, R. and Levy, H. (2009). Markowitz versus the Talmudic portfolio diversification strategies. *Journal of Portfolio Management* **35**, 71–74.

Green, R. C. and Hollifield, B. (1992). When will mean-variance efficient portfolios be well diversified? *Journal of Finance* **47**, 1785–1295.

Grossman, S. A. and Stiglitz, J. E. (1980). On the impossibility of information efficient markets. *American Economic Review* **70**, 393–408.

Haff, L. R. (1979). An identity for the Wishart distribution with applications. *Journal of Multivariate Analysis* **9**, 531–544.

Haugen, R. A. and Baker, N. L. (1991). The efficient market inefficiency of capitalization-weighted stock portfolios. *Journal of Portfolio Management* **17**, 35–40.

Hirschberger, M., Steuer, R. E., Utz, S., Wimmer, M. and Qi, Y. (2013). Computing the nondominated surface in tri-criterion portfolio selection. *Operations Research* **61**, 169–183. MR3042749

Hwang, C. L. and Masud, A. S. M. (1979). *Multiple Objective Decision Making-Methods and Applications*. Berlin: Springer. MR0567888

Jagannthan, R. and Ma, T. (2003). Risk reduction in large portfolio: Why imposing the wrong constraints helps. *Journal of Finance* **58**, 1651–1683.

Jobson, J. D. and Korkie, B. (1980). Estimation for Markowitz efficient portfolios. *Journal of the American Statistical Association* **56**, 544–554. MR0590686

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis* **21**, 279–292.

Kan, R. and Zhou, G. (2007). Optimal portfolio choice with parameters uncertainty. *Journal of Financial and Quantitative Analysis* **42**, 621–656.

Khoury, N. T. and Martel, J. M. (1985). Optimal futures hedging in the presence of asymmetric information. *Journal of Futures Markets* **5**, 595–605.

Kirby, C. and Ostdiek, B. (2012). It's all in the timing: Simple active portfolio strategies that outperform naive diversification. *Journal of Financial and Quantitative Analysis* **47**, 437–467.

Kirilenko, A. A. (2001). Valuation and control in venture finance. *Journal of Finance* **56**, 565–587.

Ledoit, O. and Wolf, M. (2004). Honey, I shrink the sample covariance. *Journal of Portfolio Management* **30**, 110–119.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance* **7**, 79–91.

MacKinlay, A. C. and Pástor, L. (2000). Asset pricing models: Implication for expected returns and portfolio selection. *Review of Financial Studies* **13**, 883–916.

Merton, R. C. (1972). An analytic derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis* **7**, 1851–1872.

Okhrin, Y. and Schmid, W. (2006). Distributional properties of portfolio weights. *Journal of Econometrics* **134**, 235–256. MR2328322

Tu, J. and Zhou, G. (2004). Data-generating process uncertainty: What difference does it make in portfolio decisions? *Journal of Financial Economics* **72**, 385–421.

Tu, J. and Zhou, G. (2011). Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics* **99**, 204–215.

Wang, Z. (2005). A shrinkage approach to model uncertainty and asset allocation. *Review of Financial Studies* **18**, 673–705.

Department of Finance
National United University
1, Lienda, Miaoli 36003
Taiwan, R.O.C.
E-mail: wychiu@nuu.edu.tw