Journal of Banking & Finance 40 (2014) 154–164

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Riskiness-minimizing spot-futures hedge ratio ☆

Yi-Ting Chen [a], Keng-Yu Ho [b,*], Larry Y. Tzeng [b,c]

[a] Institute of Economics, Academia Sinica, Taiwan
[b] Department of Finance, National Taiwan University, Taiwan
[c] Risk and Insurance Research Center, National Chengchi University, Taiwan

ARTICLE INFO

Article history:
Received 9 March 2013
Accepted 23 November 2013
Available online 7 December 2013

JEL classification:
C58
G11
G32

Keywords:
Riskiness index
Optimal hedge ratio
Method-of-moments

ABSTRACT

In this paper, we propose a new spot-futures hedging method that determines the optimal hedge ratio by minimizing the riskiness of hedged portfolio returns, where the riskiness is measured by the index of Aumann and Serrano (2008). Unlike the risk measurements widely used in the literature, the riskiness index employed in our method satisfies monotonicity with respect to stochastic dominance. We also provide an empirical example to demonstrate how to estimate and test this optimal hedge ratio in equity data by the method-of-moments.

1. Introduction

One of the important purposes for financial derivatives, such as futures, is hedging. The fundamental concept of hedging is to combine investments in the spot market and futures market and to form a portfolio that eliminates the risk of the investment value. The optimal hedge ratio has been well-studied in the literature; it depends on the particular objective function to be optimized. Chen et al. (2003) provide a broad review on different theoretical approaches to the optimal hedge ratio. They suggest that, in general, no single optimal hedge ratio is superior to others.

Numerous papers have provided the solution for the optimal hedge ratio by minimizing alternative risk measures. Early studies, such as Ederington (1979), use variance as the risk measure. Howard and D'Antonio (1984) further propose a Sharpe ratio-based hedge ratio. More recently, the mean-Gini coefficient (Cheung et al., 1990; Kolb and Okunev, 1993; Lien and Luo, 1993), generalized semi-variance (GSV) (De Jong et al., 1997; Lien and Tse, 1998), and mean-GSV (Chen et al., 2001) have also been considered as risk measures in the studies on optimal hedging. Brooks et al. (2002) highlight the importance of allowing the optimal hedge ratio to be time-varying and asymmetric based on the value at risk (VaR).

Although the prior literature has documented many ingenious findings, the risk measures employed by most papers do not satisfy the monotonicity with respect to stochastic dominance[1]: if the return distribution of one asset (lottery) first- or second-order stochastically dominates that of the other asset (lottery), then the riskiness index of the former asset (lottery) is smaller than that of the latter. Recently, Aumann and Serrano (2008, A&S hereafter) have proposed an economic index of riskiness.[2] It is the reciprocal of the absolute risk aversion parameter of a constant risk-averse individual who is indifferent in taking or not taking the lottery. They show that the index follows two axioms: duality and positive homogeneity. Specifically, any index satisfying the duality and positive homogeneity axioms is a multiple of their riskiness index. More importantly,

☆ We thank Carol Alexander (Editor) and Rudi Zagst (Associate Editor) and especially the valuable comments and suggestions from the two anonymous referees that lead to substantial improvements in this paper.

* Corresponding author. Tel.: +886 2 3366 1094.
E-mail addresses: ytchen@sinica.edu.tw (Y.-T. Chen), kengyuho@ntu.edu.tw (K.-Y. Ho), tzeng@ntu.edu.tw (L.Y. Tzeng).

[1] We provide an example to demonstrate that variance does not satisfy the monotonicity with respect to stochastic dominance. Assume that there are two lotteries, A and B. Let $A = (0.1, 0.9; -1, 11)$ and $B = (0.1, 0.9; 10, -1)$, $\mathrm{Var}(A) = 12.96 > \mathrm{Var}(B) = 10.89$. However, A first-order stochastic dominates B, because the cumulative distribution function of B is always higher than that of A. Since lower order stochastic dominance implies higher order stochastic dominance, A second-order stochastic dominates B.

[2] In responding to A&S, Foster and Hart (2009) develop an alternative operational riskiness measure that only focuses on the wealth level. In addition, Bali et al. (2011) extend A&S and Foster and Hart (2009) by proposing a generalized measure of riskiness that nests the above two measures.

0378-4266/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jbankfin.2013.11.038

they also show that the index satisfies monotonicity with respect to stochastic dominance.

Following A&S, some papers apply the riskiness measure to various issues in finance. For example, Homm and Pigorsch (2012) propose a new performance measure based on the A&S riskiness index that generalizes the Sharpe ratio. They also provide empirical examinations from mutual fund data. In addition, Schreiber (2012) develops a theoretical framework on an equivalent index to measure relative risk, complementing the A&S riskiness index that measures absolute risk.

Our study, motivated by the riskiness index introduced by A&S, is different from the above two applications, because we estimate the optimal hedge ratio based on the A&S riskiness index. As their riskiness index is consistent with the concept of stochastic dominance and the theoretical axioms, it is intriguing and crucial to re-investigate the optimal hedge ratio by applying the new risk measurement. The optimal hedge ratio minimizing A&S riskiness index guarantees that there exists no other portfolio which second-order stochastic dominates the optimal portfolio. In another word, it is also true that, for all risk averse investors, there exists no other portfolio whose expected utility is larger than that of the optimal one.

There are several other advantages for the riskiness index proposed by A&S over the traditional risk measures. For example, variance, as a measure of dispersion, takes little account of the direction of deviations from the mean. In comparison, the A&S riskiness index gives less credit for gains but penalizes more for losses. Another commonly used risk measure is the Sharpe ratio. The implicit assumption for building the Sharpe ratio is that the mean and standard deviation of an asset return completely characterize the risk of the asset. However, it is well known that returns in the financial markets are often skewed with excess kurtosis. The A&S riskiness index, on the other hand, is derived by a moment generating function restriction that has an expected utility interpretation. Finally, VaR is a widely used risk measure, but it ignores the gain side of an asset. Even on the loss side, VaR focuses only on the losses that are below a pre-specified threshold. The A&S riskiness index takes into account not only the loss side but also the gain side.

With the above motivations, we contribute to the literature by calculating the optimal hedge ratio[3] on the basis of the A&S riskiness index and by providing the associated estimation and testing methods. Specifically, we first derive the optimal hedge ratio by minimizing the riskiness index of the hedged portfolio. We provide the analytical solution of this riskiness-minimizing (*R*-min) hedge ratio under the normality assumption, and propose a general rule to calculate the *R*-min hedge ratio when the returns do not follow normal distributions. We also compare the *R*-min hedge ratio with the variance-minimizing (*V*-min) hedge ratio, and show the differences in their empirical implications. By using the method-of-moments, we can easily estimate these optimal hedge ratios and test whether the *R*-min hedge ratio is statistically different from the *V*-min hedge ratio (or any other fixed hedge ratios).

Finally, we further demonstrate an empirical application of our method using the daily spot and futures prices of a set of US stock indices. We find that the *R*-min hedged portfolio can effectively reduce the riskiness of the spot measured by the A&S riskiness index, and has higher mean and variance and much lower riskiness than the *V*-min hedged portfolio. These empirical findings are all consistent with the theoretical results that will be explored later in the paper. We also find that the *R*-min hedge ratio is significantly different from zero and one, implying that the decision makers should neither choose the naked portfolio (that is, the spot) nor the fully-hedged portfolio for minimizing the riskiness. Moreover, we find that the *R*-min hedged ratio is significantly smaller than the *V*-min hedged ratio. This suggests that the empirical differences between the *R*-min hedge portfolio and the *V*-min hedge portfolio are statistically meaningful.

The remainder of this paper is organized as follows. Section 2 presents a description of the A&S riskiness index, and explores how this riskiness index behaves under non-normality. Section 3 discusses the *R*-min method, and compares it with the *V*-min method. Section 4 offers the estimation and testing methods and provides an analysis of the empirical example. Finally, in Section 5, we present our conclusions. Appendix A collects some mathematical derivations.

2. Riskiness index

Let x_t be the return on an asset at time t with $\mathrm{P}(x_t < 0) > 0$ and $\mathbb{E}[x_t] > 0$, and w be the initial wealth of a decision maker in regard to this asset.[4] This individual has the utility $U(w)$ at time $t+1$ if she does not invest in this asset at time t, and has the expected utility $\mathbb{E}[U(w+x_t)]$ at time $t+1$ if she makes the investment at time t. Given the constant absolute risk aversion (CARA hereafter) utility function: $U(w) = -\gamma^{-1}\exp(-\gamma w)$ with the risk aversion coefficient $\gamma > 0$, the individual would be indifferent between making and not making the investment if $\gamma = \gamma^*$, where γ^* is given by the expected utility restriction:

$$-\frac{1}{\gamma^*}\exp(-\gamma^* w) = \mathbb{E}\left[-\frac{1}{\gamma^*}\exp(-\gamma^*(w+x_t))\right].$$

By denoting that $R_x := 1/\gamma^* > 0$, we can write this restriction as:

$$\mathbb{E}\left[\exp\left(-\frac{x_t}{R_x}\right)\right] = 1. \tag{1}$$

A&S proposed using R_x as an economic index for the riskiness of investing in this asset, and proved that this index has the appealing properties mentioned in Section 1.

Riskiness index of A&S does not require that the underlying utility function is CARA. They apply the CARA utility function as a bridge to derive a measurement which satisfies monotonicity with respect to stochastic dominance. In another word, for all risk averse individuals, if lottery *A* is preferred to lottery *B*, then the riskiness of lottery *A* is smaller than that of lottery *B*. Thus, in theory, we can only conclude that there does not exist a portfolio preferred by all risk averse investors to the portfolio under the optimal hedge ratio minimizing A&S riskiness index. However, for a specific utility function, the portfolio minimizing A&S riskiness index may not be always preferred. To illustrate the above argument, later in Section 3.3, we further provide quadratic utility function to compare the expected utility of the portfolio minimizing A&S riskiness index and the portfolio minimizing variance.

To explore the statistical features of R_x, note that $\mathbb{E}[\exp(\tau x_t)]$ with $\tau \in \mathbb{R}$ is the moment generating function of x_t when it exists. Thus, according to (1), we may derive R_x from a moment generating function restriction for some particular distributions. For example, x_t has the moment generating function:

$$\mathbb{E}[\exp(\tau x_t)] = \exp\left(\mu_x\tau + \frac{1}{2}\sigma_x^2\tau^2\right),$$

with $\mu_x := \mathbb{E}[x_t] > 0$ and $\sigma_x^2 := \mathrm{var}[x_t]$, when $x_t \sim N(\mu_x, \sigma_x^2)$. In this normality case, we can write (1) as

$$\exp\left(-\frac{\mu_x}{R_x} + \frac{\sigma_x^2}{2R_x^2}\right) = 1 \tag{2}$$

[3] We note that our choice of optimal hedge ratio is not based on a dynamic setting.

[4] To be precise, x_t is the return from t to $t+1$.

Table 1
R_x under non-normal distributions.

$t(v)$				$L^-(\omega)$				$L^+(\omega)$			
v	$\mu_x = 0.1$	0.25	0.5	ω	$\mu_x = 0.1$	0.25	0.5	ω	$\mu_x = 0.1$	0.25	0.5
4	5.249	2.249	1.249	0.05	5.035	2.046	1.049	0.05	4.955	1.951	0.951
6	5.082	2.146	1.249	0.10	5.084	2.098	1.102	0.10	4.906	1.901	0.903
8	5.062	2.070	1.153	0.15	5.135	2.151	1.158	0.15	4.856	1.852	0.856
10	5.055	2.049	1.093	0.20	5.187	2.207	1.218	0.20	4.805	1.802	0.809
12	5.051	2.038	1.067	0.25	5.242	2.266	1.283	0.25	4.753	1.752	0.763
14	5.049	2.032	1.053	0.30	5.300	2.330	1.355	0.30	4.699	1.701	0.717
16	5.048	2.028	1.044	0.35	5.361	2.400	1.436	0.35	4.643	1.649	0.671
18	5.046	2.025	1.038	0.40	5.429	2.477	1.529	0.40	4.585	1.595	0.625
20	5.046	2.022	1.033	0.45	5.502	2.564	1.639	0.45	4.523	1.539	0.579
22	5.045	2.021	1.030	0.50	5.584	2.665	1.770	0.50	4.457	1.481	0.533

Note: The entries are the (simulated) riskiness of x_t under non-normal distributions.

and obtain a particular R_x:

$$R_x^n = \frac{\sigma_x^2}{2\mu_x}; \tag{3}$$

see A&S (p. 820) for a different derivation of (3). This simple formula demonstrates that, under the normality assumption, the riskiness of x_t not only increases with σ_x^2 but also decreases with μ_x.

In general, it is difficult to obtain an analytical form of R_x in the presence of non-normality.[5] Nonetheless, R_x^n might be regarded as a sensible approximation to R_x when μ_x is sufficiently small. The rationale is that, by fixing σ_x^2, the riskiness index R_x^n increases as the expected return μ_x decreases. This implies that $\exp(-x_t/R_x^n)$ with an arbitrarily fixed x_t should converge to one (and hence $\mathbb{E}[\exp(-x_t/R_x^n)]$ should converge to $\mathbb{E}[\exp(-x_t/R_x)]$) when μ_x approaches zero. Observing this relationship is useful because daily stock returns are typically non-normal and have rather small means.

To explore how R_x behaves under non-normality by means of a simulation, we consider the standardized $t(v)$ distribution with the probability density function (PDF):

$$g_t(u, v) = \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)\pi}}\left(1 + \frac{u^2}{v-2}\right)^{-\frac{v+1}{2}}, \tag{4}$$

where $u \in \mathbb{R}$ and $\Gamma(\cdot)$ denotes the gamma function, and the standardized log–normal distribution $L^+(\omega)$ with the PDF:

$$g_L^+(u, \omega) = \frac{\sigma_\omega}{\sqrt{2\pi}\omega\left(\mu_\omega^{1/2} + \sigma_\omega u\right)} \times \exp\left(-\frac{1}{2}\left(\frac{1}{\omega}\ln\left(\mu_\omega^{1/2} + \sigma_\omega u\right)\right)^2\right), \tag{5}$$

where $u > -\sigma_\omega^{-1}\mu_\omega$ (otherwise, zero), $\mu_\omega := \exp(\omega^2)$, and $\sigma_\omega^2 := \exp(\omega^2)(\exp(\omega^2) - 1)$. Like $N(0,1)$, these two distributions are both of zero mean and unit variance. Unlike $N(0,1), t(v)$ is heavily-tailed when v is small and $L^+(\omega)$ is right-skewed when $\omega > 0$. The former converges to $N(0,1)$ as $v \to \infty$, and the latter degenerates to $N(0,1)$ as $\omega = 0$. We also consider the "left-skewed log–normal distribution" $L^-(\omega)$ which is defined to have the PDF $g_L^-(u,\omega) := g_L^+(-u,\omega)$.

In Table 1, we show the simulated R_x's of $x_t := \mu_x + \sigma_x u_t$, with $\sigma_x = 1$, under different combinations of (μ_x, u_t)'s, including $\mu_x = 0.1, 0.25, 0.5$, $u_t \sim t(v)$ with $v = 4, 6, \ldots, 22$, and $L^-(\omega)$ and $L^+(\omega)$ with $\omega = 0.05, 0.1, \ldots, 0.5$. We compute these R_x's by solving the equation: $T^{-1}\sum_{t=1}^T \exp(-x_t/R_x) = 1$, where $T = 10^6$ and $\{x_t\}_{t=1}^T$ is a random sample drawn from the distribution being considered. From this table, we can observe that, by fixing μ_x, R_x tends to increase when $t(v)$ becomes more heavily-tailed (v decreases), $L^-(\omega)$ becomes more left-skewed (ω increases), or $L^+(\omega)$ becomes less right-skewed (ω increases). Meanwhile, R_x becomes closer to $R_x^n = (2\mu_x)^{-1}$ as v increases or ω decreases. Moreover, as expected, the discrepancy between R_x and R_x^n decreases (that is, the ratio R_x/R_x^n is closer to one) when μ_x becomes smaller.

We also consider the Gram–Charlier type-A expansion as another example in exploring the behavior of R_x under non-normality.[6] Let S_x and K_x be, respectively, the skewness coefficient and the kurtosis coefficient of x_t. This expansion for the PDF of x_t is of the form:

$$g(x) = \frac{1}{\sigma_x}\phi\left(\frac{x-\mu_x}{\sigma_x}\right)\left(1 + \frac{\kappa_3}{6\sigma_x^3}H_3\left(\frac{x-\mu_x}{\sigma_x}\right) + \frac{\kappa_4}{24\sigma_x^4}H_4\left(\frac{x-\mu_x}{\sigma_x}\right)\right), \tag{6}$$

in which $H_3(z) := z^3 - 3z$ and $H_4(z) := z^4 - 6z^2 + 3$ are, respectively, the third and fourth Hermite polynomials, $\phi(z) := (2\pi)^{-1/2}\exp(-z^2/2)$, with $z \in \mathbb{R}$, is the PDF of $N(0,1)$, and $\kappa_3 := \sigma_x^3 S_x$ and $\kappa_4 := \sigma_x^4(K_x - 3)$ are, respectively, the third and fourth cumulants of x_t; see, e.g., Gallant and Tauchen (1989) and Jondeau and Rockinger (2001) for its financial applications. In Appendix A, we show that if the PDF of x_t follows (6), then R_x is the solution of the nonlinear equation:

$$1 - \frac{S_x}{6}\left(\frac{\sigma_x}{R_x}\right)^3 + \frac{(K_x-3)}{24}\left(\frac{\sigma_x}{R_x}\right)^4 = \exp\left(-\frac{\mu_x}{\sigma_x}\left(\frac{\sigma_x}{R_x}\right) + \frac{1}{2}\left(\frac{\sigma_x}{R_x}\right)^2\right). \tag{7}$$

In this example, R_x is determined by the first four moments of x_t, and includes R_x^n as a special case where $S_x = 0$ and $K_x = 3$ (that is, the skewness and kurtosis coefficient restrictions under normality.)

In Table 2, we present the R_x numerically solved from (6) under the same first four moment combinations considered by Table 1. Specifically, we set $\mu_x = 0.1$, 0.25, and 0.5, $\sigma_x^2 = 1$, and $(S_x, K_x) = (\lambda_S(\omega), 3)$ and $(0, \lambda_K(v))$, with $\lambda_S(\omega) := (\exp(\omega^2) + 2)\sqrt{\exp(\omega^2) - 2}$ denoting the skewness coefficient of $L^+(\omega)$ and $\lambda_K(v) := 3 + 6/(v-4)$ denoting the kurtosis coefficient of $t(v)$, for the (v, ω)'s shown in Table 1. Because $\lambda_K(v)$ is only defined when $v > 4$, the case where $v = 4$ is infeasible in Table 2. In Fig. 1, we plot $\lambda_S(\omega)$ and $\lambda_K(v)$ and compare the PDFs of $N(\mu, \sigma^2), t(v)$, and $L^-(\omega)$ with the Gram–Charlier expansion $g(\cdot)$ with $\mu_x = 0.25, \sigma_x^2 = 1$, $v = 6$ or 8, and $\omega = 0.3$ or 0.5. This figure shows that the distribution shape of $g(\cdot)$ is quite different from that of $t(v)$ or $L^-(\omega)$ even though they have the same μ_x, σ_x^2, S_x, and K_x. Like Tables 1 and 2 shows that R_x converges to R_x^n when (S_x, K_x) approaches to (0,3), and the discrepancy between R_x and R_x^n decreases when μ_x becomes smaller. However, unlike Tables 1 and 2 indicates that

[5] Schulze (2010) derives numerical and closed-form solutions for the riskiness of certain particular distributions, including exponential, Poisson, Gamma, and variance-Gamma distributions.

[6] We thank an anonymous referee for this comment.

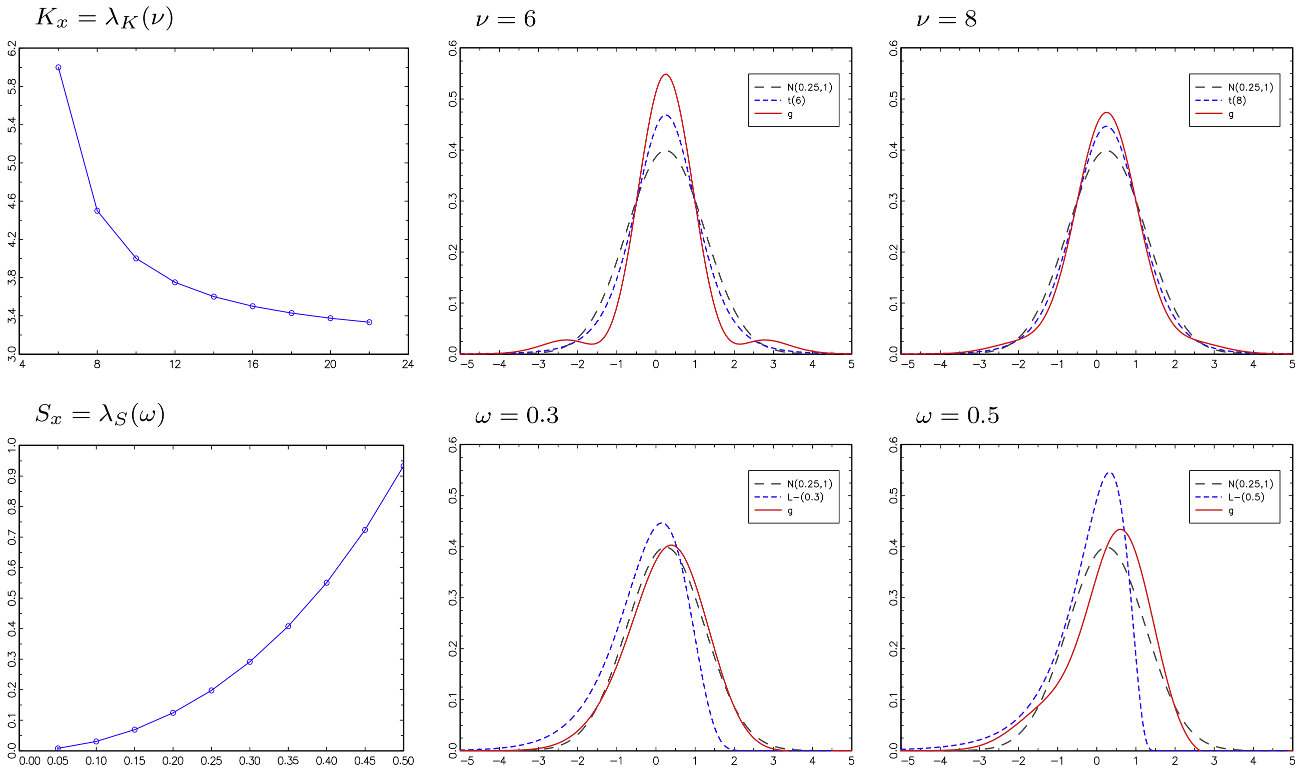

Fig. 1. The PDFs of $N(\mu_x, \sigma_x^2), t(\nu)$, and $L^-(\omega)$ and the Gram–Charlier expansion $g(\cdot)$.

Table 2
R_x under the Gram–Charlier expansion.

$(S_x, K_x) = (0, \lambda_K(\nu))$				$(S_x, K_x) = (-\lambda_S(\omega), 3)$				$(S_x, K_x) = (\lambda_S(\omega), 3)$			
ν	$\mu_x = 0.1$	0.25	0.5	ω	$\mu_x = 0.1$	0.25	0.5	ω	$\mu_x = 0.1$	0.25	0.5
4	–	–	–	0.05	4.998	1.998	0.998	0.05	5.003	2.002	1.003
6	4.948	1.856	0.431	0.10	4.990	1.990	0.990	0.10	5.010	2.010	1.010
8	4.975	1.933	0.829	0.15	4.978	1.977	0.977	0.15	5.023	2.023	1.022
10	4.983	1.957	0.900	0.20	4.958	1.958	0.957	0.20	5.040	2.040	1.040
12	4.988	1.968	0.929	0.25	4.933	1.932	0.931	0.25	5.066	2.064	1.063
14	4.990	1.974	0.945	0.30	4.902	1.898	0.895	0.30	5.094	2.093	1.091
16	4.993	1.979	0.955	0.35	4.859	1.854	0.848	0.35	5.133	2.128	1.124
18	4.993	1.982	0.962	0.40	4.810	1.797	0.786	0.40	5.176	2.170	1.163
20	4.993	1.984	0.967	0.45	4.746	1.723	0.706	0.45	5.230	2.219	1.207
22	4.995	1.986	0.971	0.50	4.666	1.624	0.612	0.50	5.294	2.275	1.257

Note: The entries are the riskiness of x_t numerically solved from (7).

$R_x > R_x^n$ in the excess-kurtosis case where $(S_x, K_x) = (0, \lambda_K(\nu))$ or in the negative-skewness case where $(S_x, K_x) = (-\lambda_S(\omega), 3)$, and $R_x < R_x^n$ in the positive-skewness case where $(S_x, K_x) = (\lambda_S(\omega), 3)$. This may be related to the fact that, as shown by Fig. 1, the distribution shape of $g(\cdot)$ is quite different from those of $t(\nu)$ and $L^-(\omega)$ under non-normality. The function $g(\cdot)$ has a higher peakedness than the PDFs of $t(\nu)$ and $N(\mu_x, \sigma_x^2)$ in the excess-kurtosis case, and has a larger (lower) mode than $L^-(\omega)$ and $N(\mu_x, \sigma_x^2)$ in the negative-skewness (positive-skewness) case.

Generally speaking, the results in Tables 1 and 2 suggest that (μ_x, σ_x^2) may play a more essential role than the distribution symmetry and tails, or (S_x, K_x), in determining R_x under non-normality; moreover, various non-normal distributions could have various R_x's even when these distributions have the same $(\mu_x, \sigma_x^2, S_x, K_x)$. This means that it is important to estimate R_x based on the general moment restriction in (1), rather than its particular formula, in the presence of non-normality.

3. The riskiness-minimizing method

Let s_t and f_t be, respectively, the spot return and the futures return at time t. We denote $\mu_s := \mathbb{E}[s_t], \sigma_s^2 := \text{var}[s_t], \mu_f := \mathbb{E}[f_t], \sigma_f^2 := \text{var}[f_t], \sigma_{sf} := \text{cov}(s_t, f_t)$, and $\rho_{sf} := \text{corr}(s_t, f_t)$, and assume that $\mu_s > 0, \mu_f > 0$, and $|\rho_{sf}| < 1$. The purpose of hedging is to mitigate certain undesirable features, such as the risk measures mentioned in Section 1, of a portfolio which comprises a long spot position and a short futures position. Let $p_t(\alpha) := s_t - \alpha f_t$ be the return on a hedged portfolio, with $\alpha \in \mathbb{A}$ standing for the hedge ratio and with $\mathbb{A}$ denoting the parameter space of α. Also, denote

$$\mu_p(\alpha) := \mathbb{E}[p_t(\alpha)] = \mu_s - \alpha\mu_f \tag{8}$$

and

$$\sigma_p^2(\alpha) := \text{var}[p_t(\alpha)] = \sigma_s^2 - 2\alpha\sigma_{sf} + \alpha^2\sigma_f^2. \tag{9}$$

The optimal choice of α is determined by the risk measure to be minimized.

Conventionally, researchers consider the variance $\sigma_p^2(\alpha)$ as the risk measure to be minimized. This V-min method generates an optimal hedge ratio:

$$\alpha_V := \arg\min_{\alpha\in\mathbb{A}} \operatorname{var}[p_t(\alpha)],$$

which can be written as:

$$\alpha_V = \frac{\sigma_{sf}}{\sigma_f^2} = \rho_{sf}\left(\frac{\sigma_s}{\sigma_f}\right). \tag{10}$$

By applying (1) to $x_t = p_t(\alpha)$, we can define the riskiness of $p_t(\alpha)$ as $R_p(\alpha)$, which is given by the moment restriction:

$$\mathbb{E}\left[\exp\left(-\frac{p_t(\alpha)}{R_p(\alpha)}\right)\right] = 1, \tag{11}$$

provided that $\mu_p(\alpha) > 0$ (that is, $\alpha < \mu_s/\mu_f$). Unlike the V-min method, the R-min method considers $R_p(\alpha)$ as the risk measure to be minimized. Correspondingly, the R-min hedge ratio is defined as:

$$\alpha_R := \arg\min_{\alpha\in\mathbb{A}'} R_p(\alpha), \tag{12}$$

where $\mathbb{A}' := \{\alpha < \mu_s/\mu_f\}$. The result in (10) indicates that the V-min hedge ratio α_V is fully determined by the second moments of s_t and f_t. In comparison, because $R_p(\alpha)$ is more complicated than $\operatorname{var}[p_t(\alpha)]$, the R-min hedge ratio α_R is determined by not only the second moments but also other features of the distribution of (s_t, f_t).

3.1. Normality

In the case where (s_t, f_t) follows a bivariate normal distribution, $p_t(\alpha)$ is of the normal distribution $N\left(\mu_p(\alpha), \sigma_p^2(\alpha)\right)$. Since we are in the normality case, $R_p(\alpha)$ is equal to $R_p^n(\alpha)$ defined in the particular form:

$$R_p^n(\alpha) = \frac{\sigma_s^2 - 2\alpha\sigma_{sf} + \alpha^2\sigma_f^2}{2(\mu_s - \alpha\mu_f)}, \tag{13}$$

which is obtained by plugging $x_t = p_t(\alpha)$ into (3). In Appendix A, we show that the R-min hedge ratio can be expressed as:

$$\alpha_R^n = \left(\frac{\mu_s}{\mu_f}\right) - A^{1/2}, \tag{14}$$

where

$$A = \left(\frac{\mu_s}{\mu_f} - \frac{\sigma_s}{\sigma_f}\right)^2 + 2(1-\rho_{sf})\left(\frac{\mu_s}{\mu_f}\right)\left(\frac{\sigma_s}{\sigma_f}\right),$$

in this case. It is easy to see that $A > 0$, because by assuming $\mu_s/\mu_f > 0$. This means that the restriction: $\alpha_R^n < \left(\frac{\mu_s}{\mu_f}\right)$ is satisfied and the riskiness index $R_p^n(\alpha)$ is well-defined when it is evaluated at $\alpha = \alpha_R^n$.

3.2. General distributions

In the case where (s_t, f_t) does not have a normal distribution, we can still first identify $R_p(\alpha)$ from (11) for each $\alpha \in \mathbb{A}'$ and then solve α_R as the minimizer of $R_p(\cdot)$ as defined in (12). The R-min hedge ratio α_R can be understood as the solution of the first-order condition:

$$R_p'(\alpha_R) = 0, \tag{15}$$

provided that the second-order condition:

$$R_p''(\alpha_R) > 0 \tag{16}$$

is satisfied.

To be specific on conditions (15) and (16), we denote

$$B(\alpha) := -\mathbb{E}\left[\exp\left(-\frac{p_t(\alpha)}{R_p(\alpha)}\right)f_t\right]$$

and

$$C(\alpha) := \mathbb{E}\left[\exp\left(-\frac{p_t(\alpha)}{R_p(\alpha)}\right)p_t(\alpha)\right].$$

Note that $C(\alpha)$ may tend to be negative because of the functional form of $\exp(-x/R)x$. Specifically, by fixing an arbitrary $R > 0$, $\exp(-x/R)x$ is negative (positive) when $x < 0$ (when $x > 0$), and diverges to negative infinity (converges to zero) as x decreases (increases) from zero. Thus, $C(\alpha)$ ought to be negative provided that the distribution of $p_t(\alpha)$ is not extremely concentrated on the gain side. Indeed, as proved in Appendix A, we have the result:

$$C(\alpha) = -\mu_p(\alpha) \tag{17}$$

under normality; by construction, $C(\alpha)$ is negative when $\mu_p(\alpha) > 0$. Thus, we assume that $C(\alpha) < 0$ in the following discussions.

By differentiating (11) with respect to α, we have the result:

$$\mathbb{E}\left[\exp\left(-\frac{p_t(\alpha)}{R_p(\alpha)}\right)\left(\frac{f_t}{R_p(\alpha)} + \frac{p_t(\alpha)R_p'(\alpha)}{R_p(\alpha)^2}\right)\right] = 0.$$

Since $R_p(\alpha)$ and $R_p'(\alpha)$ do not depend on s_t and f_t, the above equation can be rewritten as

$$-R_p(\alpha)\mathbb{E}\left[\exp\left(-\frac{p_t(\alpha)}{R_p(\alpha)}\right)f_t\right] = \mathbb{E}\left[\exp\left(-\frac{p_t(\alpha)}{R_p(\alpha)}\right)p_t(\alpha)\right]R_p'(\alpha),$$

which implies that

$$R_p'(\alpha) = \left(\frac{B(\alpha)}{C(\alpha)}\right)R_p(\alpha). \tag{18}$$

Consequently, we can reexpress the first-order condition in (15) as $B(\alpha_R) = 0$; that is,

$$\mathbb{E}\left[\exp\left(-\frac{p_t(\alpha_R)}{R_p(\alpha_R)}\right)f_t\right] = 0. \tag{19}$$

By differentiating (18) with respect to α, we can further show that

$$R_p''(\alpha) = \left(\frac{B(\alpha)}{C(\alpha)}\right)R_p'(\alpha) + \left(\frac{B'(\alpha)}{C(\alpha)} - \frac{B(\alpha)C'(\alpha)}{C(\alpha)^2}\right)R_p(\alpha),$$

where

$$B'(\alpha) := -\mathbb{E}\left[\exp\left(-\frac{p_t(\alpha)}{R_p(\alpha)}\right)f_t\left(\frac{f_t}{R_p(\alpha)} + \frac{p_t(\alpha)R_p'(\alpha)}{R_p(\alpha)^2}\right)\right].$$

Using the first-order condition: $R_p'(\alpha_R) = 0$ or $B(\alpha_R) = 0$, we can obtain that

$$R_p''(\alpha_R) = \frac{B'(\alpha_R)}{C(\alpha_R)}R_p(\alpha_R), \tag{20}$$

where

$$B'(\alpha_R) = -\mathbb{E}\left[\exp\left(-\frac{p_t(\alpha_R)}{R_p(\alpha_R)}\right)\left(\frac{f_t^2}{R_p(\alpha_R)}\right)\right] < 0.$$

Thus, the second-order condition: $R_p''(\alpha_R) > 0$ holds when $C(\alpha_R) < 0$.

As a consequence, we can interpret the explicit form of the first-order condition in (19) as the moment condition of the R-min hedge ratio α_R. Introducing this moment condition is important for the method-of-moments estimation for α_R that will be discussed in Section 4.

3.3. Comparison with the $\boldsymbol{V}$-min method

To compare the R-min method with the V-min method in an easier and clearer way, we focus on the normality case where α_R has the analytical solution in (14). To facilitate this comparison, we first reexpress the component A in (14) as:

$$A = \left[\left(\frac{\mu_s}{\mu_f}\right) - \alpha_V\right]^2 + \left(1 - \rho_{sf}^2\right)\left(\frac{\sigma_s}{\sigma_f}\right)^2.$$

Accordingly, we can use the restriction: $|\rho_{sf}| < 1$ to show that

$$A^{1/2} > \left(\frac{\mu_s}{\mu_f}\right) - \alpha_V. \tag{21}$$

Combining (14) with (21), we can obtain the result: $\alpha_R^n < \alpha_V$. This indicates that the R-min hedge ratio α_R^n is smaller than the V-min hedge ratio α_V. Since $R^n(\alpha_R^n) < R^n(\alpha_V)$ holds by construction, this also means that the V-min method is over-hedged if the riskiness of $p_t(\alpha)$ to be hedged is measured by $R_p(\alpha)$, rather than by $\sigma_p^2(\alpha)$. Moreover, unlike α_V which is independent of μ_s and μ_f, α_R^n is dependent on the mean ratio μ_s/μ_f as in (14). In particular, the R-min hedge ratio α_R^n would increase with the ratio μ_s/μ_f because

$$\frac{\partial\alpha_R^n}{\partial(\mu_s/\mu_f)} = 1 - A^{-1/2}\left(\frac{\mu_s}{\mu_f} - \alpha_V\right)$$

is positive as implied by (21).

Following (8) and (9), and the aforementioned result: $\alpha_V - \alpha_R^n > 0$, we can further see that

$$\mu_p(\alpha_R^n) - \mu_p(\alpha_V) = (\alpha_V - \alpha_R^n)\mu_f > 0$$

and

$$\begin{aligned}\sigma_p^2(\alpha_R^n) - \sigma_p^2(\alpha_V) &= 2(\alpha_V - \alpha_R^n)\sigma_{sf} - \left(\alpha_V^2 - (\alpha_R^n)^2\right)\sigma_f^2\\ &= 2(\alpha_V - \alpha_R^n)\alpha_V\sigma_f^2 - \left(\alpha_V^2 - (\alpha_R^n)^2\right)\sigma_f^2\\ &= \left(2(\alpha_V - \alpha_R^n)\alpha_V - \left(\alpha_V^2 - (\alpha_R^n)^2\right)\right)\sigma_f^2\\ &= (\alpha_V - \alpha_R^n)^2\sigma_f^2 > 0,\end{aligned}$$

in which the second equality is due to (10). Put differently, the R-min portfolio return $p_t(\alpha_R^n)$ has a larger mean and variance than the V-min portfolio return $p_t(\alpha_V)$. Thus, the choice between these two methods is determined by the utility function of investors.

Recall that $U(\cdot)$ denotes a utility function. Using the fact that $p_t'(\alpha) = -f_t$ and $\partial^k p_t(\alpha)/\partial\alpha^k = 0$ when $k > 1$, we may use Taylor's expansion to show that

$$U(p_t(\alpha_R^n)) = U(p_t(\alpha_V)) + (\alpha_V - \alpha_R^n)U'(p_t(\alpha_V))f_t;$$

the second-order and higher-order terms of this expansion are zeros because

$$\frac{\partial^k U(p_t(\alpha))}{\partial\alpha^k} = \frac{\partial^k U(p_t(\alpha))}{\partial p_t(\alpha)^k} \times \frac{\partial^k p_t(\alpha)}{\partial\alpha^k} = 0$$

when $k > 1$. Accordingly, the expected utilities generated by $p_t(\alpha_R^n)$ and $p_t(\alpha_V)$ are of the discrepancy:

$$\mathbb{E}[U(p_t(\alpha_R^n))] - \mathbb{E}[U(p_t(\alpha_V))] = (\alpha_V - \alpha_R^n)\mathbb{E}[U'(p_t(\alpha_V))f_t], \tag{22}$$

which is positive (negative) if $\mathbb{E}[U'(p_t(\alpha_V))f_t]$ is positive (negative).

In (22), we do not assume that $U'(\cdot)$ is a constant which is the case of risk neutrality. Following (22), we further discuss two cases, one is assumed risk neutral preference and the other exhibits mean–variance preference. If $U(\cdot)$ is the risk-neutral utility function such that $U(p_t(\alpha)) = p_t(\alpha)$, then $\mathbb{E}[U'(p_t(\alpha_V))f_t] = \mu_f$ and hence $\mathbb{E}[U(p_t(\alpha_R^n))] > \mathbb{E}[U(p_t(\alpha_V))]$. In this example, the investor would certainly prefer the R-min portfolio to the V-min portfolio. If $U(\cdot)$ is the quadratic utility function such that

$$U(p_t(\alpha)) = a + bp_t(\alpha) - cp_t(\alpha)^2,$$

with $b > 0$ and $c > 0$, then

$$\mathbb{E}[U'(p_t(\alpha_V))f_t] = (b/(2c) - (\mu_s - \alpha_V\mu_f))2c\mu_f.$$

In this example, $\mathbb{E}[U(p_t(\alpha_R^n))] \geqslant \mathbb{E}[U(p_t(\alpha_V))]$ if $b/(2c) \geqslant (\mu_s - \alpha_V\mu_f)$. By fixing $p_t(\alpha)$, since $b/(2c)$ decreases with the Arrow–Pratt absolute risk-aversion measure:

$$-\frac{U''(p_t(\alpha))}{U'(p_t(\alpha))} = \frac{1}{b/(2c) - p_t(\alpha)},$$

the condition: $b/(2c) \geqslant (\mu_s - \alpha_V\mu_f)$ is more likely to be valid when the investor is less risk-averse. The underlying intuition is that the R-min portfolio considers both mean and variance, whereas the V-min portfolio only takes into account the variance. Since less risk-averse individuals attach more weight to the mean than the variance, they would favor the R-min method.

4. Empirical example

In this empirical example, we show the applicability of the proposed R-min portfolio to real data by using the method-of-moments. Let S_t and F_t be, respectively, the spot and futures prices of one of the three US stock indices: Standard & Poor's 400 (S&P 400), S&P 500, and NASDAQ 100 on date t. Among these indices, the S&P 400 index includes 400 medium-sized companies, the S&P 500 index is a value-weighted index consisting of the largest 500 companies in the US stock markets, and the NASDAQ 100 index consists of the largest 100 firms listed on the NASDAQ. The price data are retrieved from Datastream. The sampling period is from June 1996 to May 2011 with the sample size $T = 3912$. Let r_t be the risk-free rate, which is represented by the 1-month Treasury bill rate. The daily spot and futures returns are, respectively, computed as $s_t = 100 \times (S_{t+1} - S_t)/S_t$ and $f_t = 100 \times (F_{t+1} - e^{-r_t}F_t)/(e^{-r_t}F_t)$.[7]

Denote the sample means $\hat{\mu}_x := T^{-1}\sum_{t=1}^T x_t$, $\hat{\mu}_s := T^{-1}\sum_{t=1}^T s_t$, and $\hat{\mu}_f := T^{-1}\sum_{t=1}^T f_t$, the sample variances $\hat{\sigma}_x^2 := T^{-1}\sum_{t=1}^T (x_t - \hat{\mu}_x)^2$, $\hat{\sigma}_s^2 := T^{-1}\sum_{t=1}^T (s_t - \hat{\mu}_s)^2$, and $\hat{\sigma}_f^2 := T^{-1}\sum_{t=1}^T (f_t - \hat{\mu}_f)^2$, the sample covariance $\hat{\sigma}_{sf} := T^{-1}\sum_{t=1}^T (s_t - \hat{\mu}_s)(f_t - \hat{\mu}_f)$, and the sample correlation $\hat{\rho}_{sf} := \hat{\sigma}_{sf}/(\hat{\sigma}_s\hat{\sigma}_f)$. Recall that the parameters $R_x^n, \alpha_V, R_p^n(\alpha)$, and α_R^n have analytical solutions. Thus, we can easily and consistently estimate the parameters: $R_x^n, \alpha_V, R_p^n(\alpha)$, and α_R^n using their sample analogs:

$$\widehat{R}_x^n = \frac{\hat{\sigma}_x^2}{2\hat{\mu}_x}, \tag{23}$$

$$\hat{\alpha}_V = \frac{\hat{\sigma}_{sf}}{\hat{\sigma}_f^2}, \tag{24}$$

$$\widehat{R}_p^n(\alpha) = \frac{\hat{\sigma}_s^2 - 2\alpha\hat{\sigma}_{sf} + \alpha^2\hat{\sigma}_f^2}{2(\hat{\mu}_s - \alpha\hat{\mu}_f)}, \tag{25}$$

and

$$\hat{\alpha}_R^n = \left(\frac{\hat{\mu}_s}{\hat{\mu}_f}\right) - \widehat{A}^{1/2}, \tag{26}$$

with

[7] This formula implicitly assumes that an investor buys a futures contract, invests $e^{-r_t}F_t$ dollars in Treasury bills for the settlement of his futures position, and posts Treasury bills as margin (see Bodie and Rosansky, 1980).

Table 3
Summary statistics for the spot and futures returns.

	S&P 400		S&P 500		NASDAQ 100	
	$x_t = s_t$	$x_t = f_t$	$x_t = s_t$	$x_t = f_t$	$x_t = s_t$	$x_t = f_t$
$\hat{\mu}_x$	0.04593	0.04779	0.02611	0.02880	0.05205	0.04939
$\hat{\sigma}_x^2$	1.88073	1.99779	1.65066	1.71781	4.12794	4.02786
$\widehat{R}_x^n$	20.47506	20.90397	31.60898	29.82506	39.65054	40.77923
$\widehat{R}_x$	20.60285	21.06038	31.64085	29.79565	39.46765	40.66929
$\hat{\rho}_{sf}$	0.96927		0.97712		0.97272	
$\widehat{S}_x$	−0.17428	−0.19145	0.00521	0.17729	0.33457	0.23055
$\widehat{K}_x$	9.34883	11.48109	10.97327	13.35146	8.30095	8.53895
$\hat{n}_x$	14.24446**	10.93818**	13.24167**	10.37304**	11.06737**	20.33798**

Note: The returns are in percentage unit. The notations: $\widehat{S}_x := T^{-1}\sum_{t=1}^{T}((x_t - \hat{\mu}_x)/\hat{\sigma}_x)^3$, $\widehat{K}_x := T^{-1}\sum_{t=1}^{T}((x_t - \hat{\mu}_x)/\hat{\sigma}_x)^4$, and $\hat{n}_x$ are, respectively, the sample skewness coefficient, the sample kurtosis coefficient, and the skewness-kurtosis-based normality test statistic of Bai and Ng (2005, Theorem 4). Denote $\Phi := \lim_{T\to\infty} \mathrm{var}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} Z_t\right]$ with $Z_t := \left((x_t - \mu_x),\ (x_t - \mu_x)^2 - \sigma_x^2,\ (x_t - \mu_x)^3,\ (x_t - \mu_x)^4 - 3\sigma_x^4\right)^\top$. The test statistic $\hat{n}_x$ is defined as:

$$\hat{n}_x := \widehat{Y}^\top\left[\hat{\gamma}\hat{\Phi}\hat{\gamma}^\top\right]^{-1}\widehat{Y}, \quad \text{where} \quad \widehat{Y} := \begin{bmatrix} \hat{\sigma}_x^3\sqrt{T}\widehat{S}_x \\ \hat{\sigma}_x^4\sqrt{T}(\widehat{K}_x - 3) \end{bmatrix}, \quad \hat{\gamma} := \begin{bmatrix} -3\hat{\sigma}_x^2 & 0 & 1 & 0 \\ 0 & -6\hat{\sigma}_x^2 & 0 & 1 \end{bmatrix},$$

and $\hat{\Phi}$ is an HAC (heteroskedasticity–autocorrelation-consistent) estimator for Φ, and has the asymptotic distribution: $\hat{n}_x \xrightarrow{d} \chi^2(2)$ under the null of normality; see Appendix A for how the HAC estimator is computed in this paper. The 99% critical value of $\chi^2(2)$ is 9.2103.
** Indicates that the normality test statistic is significant at the 1% level.

$$\widehat{A} := \left(\frac{\hat{\mu}_s}{\hat{\mu}_f}\right)^2 - 2\hat{\rho}_{sf}\left(\frac{\hat{\mu}_s}{\hat{\mu}_f}\right)\left(\frac{\hat{\sigma}_s}{\hat{\sigma}_f}\right) + \left(\frac{\hat{\sigma}_s}{\hat{\sigma}_f}\right)^2,$$

respectively. In general, the parameters: $R_x, R_p(\alpha)$, and α_R may not have analytical solutions. Nonetheless, we can still consistently estimate R_x and $R_p(\alpha)$ by $\widehat{R}_x$ and $\widehat{R}_p(\alpha)$ that are, respectively, solved from the estimating equations:

$$\frac{1}{T}\sum_{t=1}^{T}\exp\left(-\frac{x_t}{\widehat{R}_x}\right) = 1 \tag{27}$$

and

$$\frac{1}{T}\sum_{t=1}^{T}\exp\left(-\frac{p_t(\alpha)}{\widehat{R}_p(\alpha)}\right) = 1 \tag{28}$$

that are, respectively, the sample analogs of the moment conditions: (1) and (11). Using the fact that α_R is the minimizer of $R_p(\cdot)$, we can also estimate α_R by the minimizer of $\widehat{R}_p(\cdot)$ which is denoted as $\hat{\alpha}_R$. Corresponding to the moment condition of α_R in (19), the estimator $\hat{\alpha}_R$ is of the estimating equation:

$$\frac{1}{T}\sum_{t=1}^{T}\exp\left(-\frac{p_t(\hat{\alpha}_R)}{\widehat{R}_p(\hat{\alpha}_R)}\right)f_t = 0; \tag{29}$$

(19) and (29) are, respectively, the first-order conditions of minimizing $R_p(\cdot)$ and minimizing $\widehat{R}_p(\cdot)$. Following a standard asymptotic method discussed in Newey and McFadden (1994) and many others, it is not difficult to derive the asymptotic distributions of these estimators and the associated significance tests. We provide the derivations in a supplementary appendix which is not reported here for the sake of brevity but is available upon requested.

In Table 3, we show a set of summary statistics of the return sequences: $\{s_t\}$ and $\{f_t\}$. From this table, we can see that the sample mean $\hat{\mu}_x$ is positive for all the spot and futures returns. This permits us to estimate the riskiness of these assets using $\widehat{R}_x^n$ or $\widehat{R}_x$. It is straightforward to compute $\widehat{R}_x^n$ according to (23). The $\widehat{R}_x^n$'s of $\{s_t\}$ and $\{f_t\}$ are about 20.475 and 20.904 for the S&P 400, 31.609 and 29.825 for the S&P 500, and 39.651 and 40.779 for the NASDAQ 100. On the other hand, we compute $\widehat{R}_x$ as the minimizer of the objective function:

$$\widehat{Q}(R) := \left(\frac{1}{T}\sum_{t=1}^{T}\exp(-x_t/R) - 1\right)^2$$

because the associated first-order condition is the same as the estimating Eq. (27). We implement the numerical optimization using the OPTMUM of *GAUSS*™ with the Newton–Raphson method. Since we already have the sample of x_t, the procedure of optimization starts with assigning an initial point of R by $\widehat{R}_x^n$. Then, the change in R is determined by the sign of $\widehat{Q}'(R)$. The searching process stops when $\widehat{Q}'(R)$ is sufficiently close to zero.

By this method, the $\widehat{R}_x$'s for $\{s_t\}$ and $\{f_t\}$ are about 20.603 and 21.060 for the S&P 400, 31.641 and 29.796 for the S&P 500, and 39.468 and 40.669 for the NASDAQ 100; see Table 3. In Fig. 2, we plot the sequences of the $\{\widehat{Q}(R)^{0.25}\}$'s with a range of R's around the $\widehat{R}_x$'s. This figure verifies that these $\widehat{R}_x$'s are the minimizers of the associated $\widehat{Q}(R)$'s. Clearly, $\widehat{R}_x^n$ is very close to $\widehat{R}_x$ for all the (s_t, f_t)'s considered. This is consistent with our theoretical finding in Section 2 that R_x^n ought to be close to R_x when μ_x is sufficiently small.

Theoretically, the choice between these two riskiness estimates is dependent on whether the hypothesis of normality is satisfied. To check this hypothesis, we conduct the skewness-kurtosis-based test of Bai and Ng (2005), which also allows the return sequences to be serially dependent. The test statistics are also shown in Table 3; see the footnote of this table for more details about this test. Not surprisingly, the normality hypothesis is strongly rejected by this test for all the returns considered. This suggests that we should choose $\widehat{R}_x$ as the riskiness estimate and explore the empirical performance of the R-min method without using the normality assumption. Thus, the normality-based estimates: $\widehat{R}_p^n(\alpha)$ and $\hat{\alpha}_R^n$ will not be considered in the following discussions.

To compare the R-min method with the V-min method in terms of their empirical performance, we show the optimal hedge ratio estimates: $\hat{\alpha}_R$ and $\hat{\alpha}_V$ and a set of summary statistics of the hedged portfolio returns: $\{p_t(\hat{\alpha}_R)\}$ and $\{p_t(\hat{\alpha}_V)\}$ in Table 4. In this table, we report the riskiness estimates: $\widehat{R}_p(\hat{\alpha}_R)$ and $\widehat{R}_p(\hat{\alpha}_V)$ for all but the case of the S&P 500 because $\hat{\mu}_p(\hat{\alpha}_V)$ is negative and hence $\widehat{R}_p(\hat{\alpha}_V)$ is undefined (or infinite) in this case. As we detail in this section, $\hat{\alpha}_V$ is estimated from (24), while $\hat{\alpha}_R$ is estimated from the

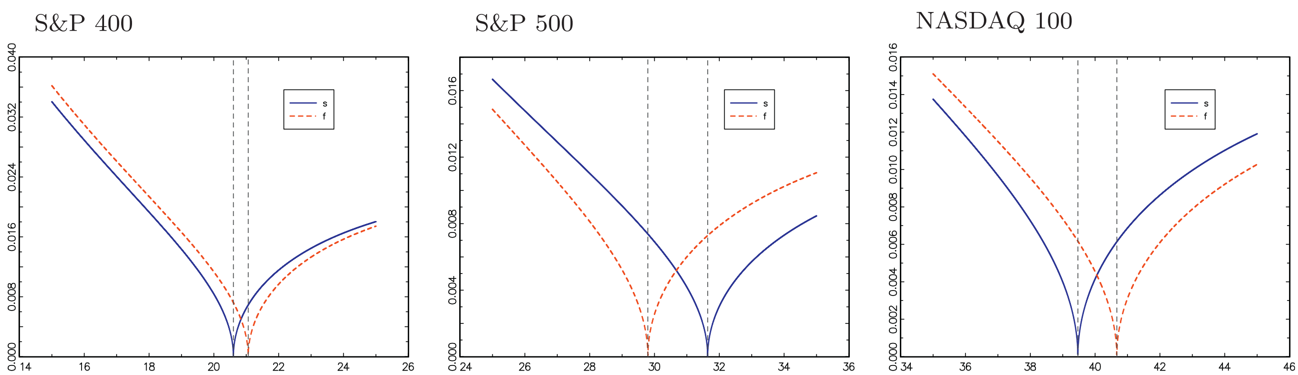

Fig. 2. The sequence $\{\widehat{Q}(R)^{0.25}\}$.

Table 4
The *R*-min and *V*-min hedged portfolios.

	S&P 400		S&P 500		NASDAQ 100	
	$\hat{\alpha} = \hat{\alpha}_R$	$\hat{\alpha} = \hat{\alpha}_V$	$\hat{\alpha} = \hat{\alpha}_R$	$\hat{\alpha} = \hat{\alpha}_V$	$\hat{\alpha} = \hat{\alpha}_R$	$\hat{\alpha} = \hat{\alpha}_V$
$\hat{\alpha}$	0.72339	0.94044	0.69345	0.95783	0.80815	0.98473
$\hat{\mu}_p(\hat{\alpha})$	0.01136	0.00099	0.00614	−0.00147	0.01214	0.00342
$\hat{\sigma}_p^2(\hat{\alpha})$	0.20724	0.11292	0.19415	0.07386	0.34609	0.22014
$\widehat{S}_p(\hat{\alpha})$	0.03183	0.07903	−0.20657	−0.26857	0.12362	−0.50218
$\widehat{K}_p(\hat{\alpha})$	8.31906	11.44760	8.28384	9.76390	10.41737	17.71967
$\widehat{R}_p(\hat{\alpha})$	9.12671	57.12148	15.84411	–	14.24178	32.16472

Note: The notations: $\widehat{S}_p(\hat{\alpha})$ and $\widehat{K}_p(\hat{\alpha})$ are, respectively, the sample skewness and kurtosis coefficients of the hedged portfolio returns $\{\hat{p}_t(\hat{\alpha})\}_{t=1}^T$.

second-step estimation in (29), where the first-step estimation of $\widehat{R}_p(\alpha)$ is based on (28).

In Fig. 3, we plot the adaptive kernel density estimates of $\{s_t\}, \{f_t\}, \{p_t(\hat{\alpha}_R)\}$, and $\{p_t(\hat{\alpha}_V)\}$; see Breiman et al. (1977). This figure shows that the density estimate of s_t is essentially indistinguishable from that of f_t for all the stock indices considered. Moreover, the density estimates of s_t and f_t have much thicker tails than those of $p_t(\hat{\alpha}_R)$ and $p_t(\hat{\alpha}_V)$. This suggests that the *R*-min method and the *V*-min method are both useful for reducing the dispersion of s_t (or f_t). However, from this figure, we can also observe that $p_t(\hat{\alpha}_R)$ and $p_t(\hat{\alpha}_V)$ have different distributions. In the following, we summarize the main differences between these two hedged portfolios according to the results in Table 4, and explain that these differences are consistent with our theoretical findings.

First, the *R*-min hedge ratio $\hat{\alpha}_R$ is systematically smaller than the *V*-min hedge ratio $\hat{\alpha}_V$ for all the stock indices considered. The value of $\hat{\alpha}_R$ ($\hat{\alpha}_V$) is about 0.723, 0.693, and 0.808 (0.940, 0.958, and 0.985) for the S&P 400, S&P 500, and NASDAQ 100, respectively. The $\hat{\alpha}_V$'s are close to one for all the indices considered. This reflects the fact that the ratios of the standard deviations and the sample correlation coefficients between $\{s_t\}$ and $\{f_t\}$, shown in Table 3, are all close to one. In comparison, the $\hat{\alpha}_R$'s are of a greater variation, and their variation is consistent with that of the sample mean ratio $\hat{\mu}_s/\hat{\mu}_f$ which is about 0.961, 0.906, and 1.054 for the S&P 400, S&P 500, and NASDAQ 100, respectively. This empirical finding reflects the theoretical results: $\alpha_R^n \leqslant \alpha_V^n$ and $\partial\alpha_R^n/\partial(\mu_s/\mu_f) > 0$ shown in Section 3.3, and suggests that the *V*-min method tends to generate an over-hedged portfolio if the undesirable feature to be minimized is the riskiness, rather than the variance, of a portfolio.

Second, the sample mean $\hat{\mu}_p(\hat{\alpha}_R)$ is systematically greater than $\hat{\mu}_p(\hat{\alpha}_V)$ for all the stock indices considered. Put differently, the *R*-min method is useful for generating a higher averaged return than the *V*-min method. This is consistent with the theoretical result: $\mu_p(\alpha_R) > \mu_p(\alpha_V)$ mentioned in Section 3.3. Nonetheless, the sample variance $\hat{\sigma}_p^2(\hat{\alpha}_R)$ is also systematically greater than $\hat{\sigma}_p^2(\hat{\alpha}_V)$ by construction. Thus, as mentioned before, the *R*-min method does not stochastically dominate the *V*-min method and vice versa. This empirical finding suggests that the *R*-min method may be preferred to the *V*-min method by investors who are more concerned about the mean than the variance of returns.

Third, by construction, the *R*-min method is much more effective for remedying the riskiness of spot (and futures) than the *V*-min method. By fixing $x_t = s_t$, the *R*-method substantially reduces the riskiness from $\widehat{R}_x = 20.603$ to $\widehat{R}_p(\hat{\alpha}_R) = 9.127$ for the S&P 400, from 31.609 to 15.844 for the S&P 500, and from 39.468 to 14.242 for the NASDAQ 100. By contrast, the *V*-min method only slightly reduces the riskiness of NASDAQ 100 from 39.468 to 32.165. Importantly, the *V*-min method does not reduce but rather increases the riskiness in the cases of the S&P 400 and S&P 500. For the S&P 400 (S&P 500), the riskiness increases from 20.603 to 57.122 (from 31.609 to infinity) by this method.

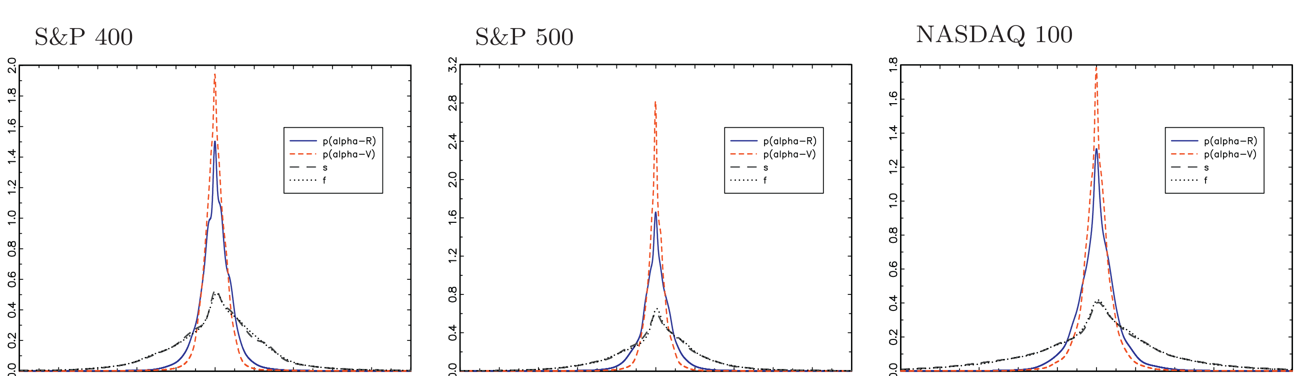

Fig. 3. Adaptive kernel density estimates of $\{s_t\}, \{f_t\}, \{p_t(\hat{\alpha}_R)\}$, and $\{p_t(\hat{\alpha}_V)\}$.

Table 5
Significance test statistics.

Hypothesis	S&P 400	S&P 500	NASDAQ 100
$H_o : \alpha_R = \alpha_V$	−3.31305**	−2.04201*	−2.36062*
$H_o : \alpha_R = 0$	10.98460**	5.35700**	10.74883**
$H_o : \alpha_R = 1$	−4.20029**	−2.36814*	−2.55171*

Note: See Appendix A for the formulae and the asymptotic null distribution of the significance test statistics.
* Indicates that the test statistic is significant at the 5% level.
** Indicates that the test statistic is significant at the 1% level.

Fourth, Table 4 also shows that the sample kurtosis coefficients of $\{p_t(\hat{\alpha}_R)\}$ are systematically smaller than those of $\{p_t(\hat{\alpha}_V)\}$ for all the stock indices considered. This is consistent with Fig. 3 which shows that the distribution of $p_t(\hat{\alpha}_R)$ is of a lower peakedness than that of $p_t(\hat{\alpha}_V)$ for these stock indices. This suggests that the R-min portfolio returns are "more regular" than the V-min portfolio returns.

Since the R-min hedged portfolio return $p_t(\alpha_R)$, the spot return $s_t = p_t(0)$, the fully-hedged portfolio return $p_t(1)$, and the V-min hedged portfolio return $p_t(\alpha_V)$ are only different in terms of their hedge ratios: $\alpha = \alpha_R, \alpha = 0, \alpha = 1$, and $\alpha = \alpha_V$, we can further check whether $p_t(\alpha_R)$ is significantly different from $p_t(0), p_t(1)$, and $p_t(\alpha_V)$ by testing the parameter restrictions: $\alpha_R = \alpha_V, \alpha_R = 0$, and $\alpha_R = 1$, respectively. In Table 5, we show the significance test statistics for these parameter restrictions. The formulae and the asymptotic null distribution of the significance test statistics are provided in Appendix A. As shown by this table, the test statistics are all significant at the 5% level for all the cases considered. Thus, the R-min portfolio is significantly different from the spot, the fully-hedged portfolio, and the V-min portfolio in this empirical example.

5. Conclusions

In this paper, we propose calculating the optimal hedge ratio by minimizing the A&S riskiness index of the hedged portfolio returns. Compared to the alternative risk measures used in the literature on hedged portfolios, this riskiness index has better economic interpretations and theoretical properties. We derive the analytical solution of the R-min hedge ratio under normality, and provide a general solution to this optimal hedge ratio in the presence of non-normality. In addition, we also demonstrate a number of differences between the empirical implications of the R-min portfolio and the V-min portfolio. In addition, we further facilitate the practical applications of the R-min portfolio by using a set of method-of-moments-based estimators and significance tests. The empirical example shows that the R-min portfolio is effective for reducing the riskiness of the spot. Moreover, it has quite different empirical performance from the V-min portfolio, and the empirical differences between these two hedged portfolios are consistent with our theoretical findings. We also apply the significance tests to show that the R-min portfolio is statistically different from the V-min portfolio (and the naked and fully-hedged portfolios).

Appendix A

A.1. Derivation of (7)

If the PDF of x_t follows (6), then the expectation operator $\mathbb{E}[\cdot]$ is taken with respect to (6). Accordingly, we can rewrite the moment condition in (1) as:

$$\int \eta(x,\mu_x,\sigma_x,R_x)\left(1+\frac{S_x}{6}H_3\left(\frac{x-\mu_x}{\sigma_x}\right)+\frac{(K_x-3)}{24}H_4\left(\frac{x-\mu_x}{\sigma_x}\right)\right)\mathrm{d}x=1, \tag{A1}$$

where

$$\eta(x,\mu_x,\sigma_x,R_x) := \frac{1}{\sigma_x}\phi\left(\frac{x-\mu_x}{\sigma_x}\right)\exp\left(-\frac{x}{R_x}\right). \tag{A2}$$

Let $\psi(x,\mu_x,\sigma_x^2,R_x)$ be the PDF of $N(\mu_x-\sigma_x^2/R_x,\sigma_x^2)$. Note that

$$\eta(x,\mu_x,\sigma_x,R_x) = \frac{1}{\sqrt{2\pi}\sigma_x}\exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right)\exp\left(-\frac{x}{R_x}\right)$$
$$= \frac{1}{\sqrt{2\pi}\sigma_x}\exp\left(-\frac{R_x(x-\mu_x)^2+2\sigma_x^2 x}{2\sigma_x^2 R_x}\right),$$

which can be further expressed as:

$$\eta(x,\mu_x,\sigma_x,R_x)=\frac{1}{\sqrt{2\pi}\sigma_x}\exp\left(-\frac{x^2-2(\mu_x-\sigma_x^2/R_x)x+(\mu_x-\sigma_x^2/R_x)^2}{2\sigma_x^2}\right)$$
$$\times\exp\left(\frac{1}{R_x}\left(\mu_x-\frac{\sigma_x^2}{2R_x}\right)\right):=\psi(x,\mu_x,\sigma_x,R_x)\exp\left(\frac{1}{R_x}\left(\mu_x-\frac{\sigma_x^2}{2R_x}\right)\right).$$

Accordingly, we can rewrite (A1) as

$$\int\left(1+\frac{S_x}{6}H_3\left(\frac{x-\mu_x}{\sigma_x}\right)+\frac{(K_x-3)}{24}H_4\left(\frac{x-\mu_x}{\sigma_x}\right)\right)\psi(x,\mu_x,\sigma_x^2,R_x)\mathrm{d}x$$
$$=\exp\left(-\frac{1}{R_x}\left(\mu_x-\frac{\sigma_x^2}{2R_x}\right)\right). \tag{A3}$$

Denote the standardized variable: $x_* := \sigma_x^{-1}(x-(\mu_x-\sigma_x^2/R_x))$. Using the fact that $\sigma_x^{-1}(x-\mu_x) = x_* - (\sigma_x/R_x)$ and the definition of $H_3(z)$ and $H_4(z)$, we can show that

$$H_3\left(\frac{x-\mu_x}{\sigma_x}\right) = x_*^3-3x_*^2\left(\frac{\sigma_x}{R_x}\right)+3x_*\left(\frac{\sigma_x}{R_x}\right)^2-\left(\frac{\sigma_x}{R_x}\right)^3-3x_*+3\left(\frac{\sigma_x}{R_x}\right) \tag{A4}$$

and

$$H_4\left(\frac{x-\mu_x}{\sigma_x}\right) = x_*^4-4x_*^3\left(\frac{\sigma_x}{R_x}\right)+6x_*^2\left(\frac{\sigma_x}{R_x}\right)^2-4x_*\left(\frac{\sigma_x}{R_x}\right)^3$$
$$+\left(\frac{\sigma_x}{R}\right)^4-6x_*^2+12\left(\frac{\sigma_x}{R_x}\right)x_*-6\left(\frac{\sigma_x}{R_x}\right)^2+3. \tag{A5}$$

Furthermore, using the fact that $\psi(x,\mu_x,\sigma_x^2,R_x)$ is the PDF of $N(\mu_x-\sigma_x^2/R_x,\sigma_x^2)$, we also have the restriction:

$$\int x_*^k\psi(x,\mu_x,\sigma_x^2,R_x)\mathrm{d}x = \begin{cases}1, & k=0,\\ 0, & k=1,\\ 1, & k=2,\\ 0, & k=3,\\ 3, & k=4.\end{cases} \tag{A6}$$

According to A4, A5 and A6, we obtain that

$$\int\psi(x,\mu_x,\sigma_x^2,R_x)\mathrm{d}x=1, \tag{A7}$$

$$\int H_3\left(\frac{x-\mu_x}{\sigma_x}\right)\psi(x,\mu_x,\sigma_x^2,R_x)\mathrm{d}x=-\left(\frac{\sigma_x}{R_x}\right)^3, \tag{A8}$$

and

$$\int H_4\left(\frac{x-\mu_x}{\sigma_x}\right)\psi(x,\mu_x,\sigma_x^2,R_x)\mathrm{d}x=\left(\frac{\sigma_x}{R_x}\right)^4. \tag{A9}$$

The result in (7) is obtained by introducing A7, A8 and A9 into (A3). □

A.2. Derivation of (14)

To minimize $R_p^n(\alpha)$, we take the derivative of (13) with respect to α and set it equal to zero. This yields the first-order condition:

$$\left(-2\sigma_{sf}+2\alpha\sigma_f^2\right)(\mu_s-\alpha\mu_f)+\left(\sigma_s^2-2\alpha\sigma_{sf}+\alpha^2\sigma_f^2\right)\mu_f=0.$$

Dividing both sides of the above equation by $\sigma_f^2\mu_f$, we can rewrite this condition as:

$$-2\left(\frac{\sigma_{sf}}{\sigma_f^2}-\alpha\right)\left(\frac{\mu_s}{\mu_f}-\alpha\right)+\left(\frac{\sigma_s^2}{\sigma_f^2}-2\alpha\frac{\sigma_{sf}}{\sigma_f^2}+\alpha^2\right)=0.$$

Using the fact that $\sigma_{sf}=\rho_{sf}\sigma_s\sigma_f$, we can further write this condition as:

$$-2\left(\rho_{sf}\left(\frac{\sigma_s}{\sigma_f}\right)-\alpha\right)\left(\frac{\mu_s}{\mu_f}-\alpha\right)+\left(\left(\frac{\sigma_s}{\sigma_f}\right)^2-2\alpha\rho_{sf}\left(\frac{\sigma_s}{\sigma_f}\right)+\alpha^2\right)=0;$$

that is,

$$-\alpha^2+2\alpha\left(\frac{\mu_s}{\mu_f}\right)-2\rho_{sf}\left(\frac{\sigma_s}{\sigma_f}\right)\left(\frac{\mu_s}{\mu_f}\right)+\left(\frac{\sigma_s}{\sigma_f}\right)^2=0.$$

Accordingly, we have the equation:

$$\left(\alpha-\frac{\mu_s}{\mu_f}\right)^2=\left(\frac{\mu_s}{\mu_f}\right)^2-2\rho_{sf}\left(\frac{\sigma_s}{\sigma_f}\right)\left(\frac{\mu_s}{\mu_f}\right)+\left(\frac{\sigma_s}{\sigma_f}\right)^2$$

that has the solutions: $\alpha=\left(\frac{\mu_s}{\mu_f}\right)\pm A^{1/2}$. It can be checked that $\alpha=\left(\frac{\mu_s}{\mu_f}\right)-A^{1/2}$ makes $\frac{\partial^2R_p^n(\alpha)}{\partial\alpha^2}>0$, whereas $\alpha=\left(\frac{\mu_s}{\mu_f}\right)+A^{1/2}$ makes $\frac{\partial^2R_p^n(\alpha)}{\partial\alpha^2}<0$. Thus, the optimal hedge ratio is $\alpha_R^n=\left(\frac{\mu_s}{\mu_f}\right)-A^{1/2}$.

A.3. Derivation of (17)

Under the normality assumption: $p_t(\alpha)\sim N\left(\mu_p(\alpha),\sigma_p^2(\alpha)\right)$, we can write that

$$\begin{aligned}\mathbb{E}\left[\exp\left(-\frac{p_t(\alpha)}{R_p(\alpha)}\right)p_t(\alpha)\right]&=\int\exp\left(-\frac{p_t(\alpha)}{R_p(\alpha)}\right)p_t(\alpha)\frac{1}{\sqrt{2\pi}\sigma_p(\alpha)}\\&\quad\times\exp\left(-\frac{(p_t(\alpha)-\mu_p(\alpha))^2}{2\sigma_p^2(\alpha)}\right)\mathrm{d}p_t(\alpha)\\&=\int p_t(\alpha)\frac{1}{\sqrt{2\pi}\sigma_p(\alpha)}\exp\left(-\left(\frac{(p_t(\alpha)-\mu_p(\alpha))^2}{2\sigma_p^2(\alpha)}+\frac{p_t(\alpha)}{R_p(\alpha)}\right)\right)\mathrm{d}p_t(\alpha).\end{aligned}\tag{A10}$$

Denote $\mu_p^*(\alpha):=\mu_p(\alpha)-\sigma_p^2(\alpha)R_p(\alpha)^{-1}$. By combining (A10) with the result:

$$\begin{aligned}&\frac{(p_t(\alpha)-\mu_p(\alpha))^2}{2\sigma_p^2(\alpha)}+\frac{p_t(\alpha)}{R_p(\alpha)}\\&=\frac{p_t(\alpha)^2-2\mu_p(\alpha)p_t(\alpha)+\mu_p(\alpha)^2+2\sigma_p^2(\alpha)R_p(\alpha)^{-1}p_t(\alpha)}{2\sigma_p^2(\alpha)}\\&=\frac{p_t(\alpha)^2-2\mu_p^*(\alpha)p_t(\alpha)+\mu_p(\alpha)^2}{2\sigma_p^2(\alpha)}\\&=\frac{p_t(\alpha)^2-2\mu_p^*(\alpha)p_t(\alpha)+\mu_p^*(\alpha)^2}{2\sigma_p^2(\alpha)}+\frac{\mu_p(\alpha)^2-\mu_p^*(\alpha)^2}{2\sigma_p^2(\alpha)}\\&=\frac{\left(p_t(\alpha)-\mu_p^*(\alpha)\right)^2}{2\sigma_p^2(\alpha)}+\frac{\mu_p(\alpha)^2-\mu_p^*(\alpha)^2}{2\sigma_p^2(\alpha)},\end{aligned}$$

we can further write that

$$\begin{aligned}\mathbb{E}\left[\exp\left(-\frac{p_t(\alpha)}{R_p(\alpha)}\right)p_t(\alpha)\right]&=\exp\left(-\frac{\mu_p(\alpha)^2-\mu_p^*(\alpha)^2}{2\sigma_p^2(\alpha)}\right)\\&\quad\times\int p_t(\alpha)\frac{1}{\sqrt{2\pi}\sigma_p(\alpha)}\\&\quad\times\exp\left(-\left(\frac{\left(p_t(\alpha)-\mu_p^*(\alpha)\right)^2}{2\sigma_p^2(\alpha)}\right)\right)\mathrm{d}p_t(\alpha)\\&=\exp\left(-\frac{\mu_p(\alpha)^2-\mu_p^*(\alpha)^2}{2\sigma_p^2(\alpha)}\right)\\&\quad\times\left[\int\left(p_t(\alpha)-\mu_p^*(\alpha)\right)\frac{1}{\sqrt{2\pi}\sigma_p(\alpha)}\right.\\&\quad\times\exp\left(-\left(\frac{\left(p_t(\alpha)-\mu_p^*(\alpha)\right)^2}{2\sigma_p^2(\alpha)}\right)\right)\\&\quad\left.\times\mathrm{d}p_t(\alpha)+\mu_p^*(\alpha)\right]\\&=\exp\left(-\frac{\mu_p(\alpha)^2-\mu_p^*(\alpha)^2}{2\sigma_p^2(\alpha)}\right)\mu_p^*(\alpha).\end{aligned}\tag{A11}$$

Recall that $R_p(\alpha)=\sigma_p^2(\alpha)/(2\mu_p(\alpha))$ and $\mu_p^*(\alpha)=\mu_p(\alpha)-2\mu_p(\alpha)=-\mu_p(\alpha)$ hold under the normality assumption. We obtain (17) by plugging this result into (A11).

A.4. Significance tests

The derivation of the following significance tests is presented in a supplementary appendix which is available upon requested. Denote

$$J(\alpha_V):=\left(0,0,0,1/\sigma_f^2,-\sigma_{sf}/\sigma_f^4\right),$$

$$J(\alpha_R):=-R_p(\alpha_R)\mathbb{E}\left[\exp\left(-\frac{p_t(\alpha_R)}{R_p(\alpha_R)}\right)f_t^2\right]^{-1},$$

$$J(R_p(\alpha)):=-R_p(\alpha)^2\mathbb{E}\left[\exp\left(-\frac{p_t(\alpha)}{R_p(\alpha)}\right)p_t(\alpha)\right]^{-1},$$

$$\begin{aligned}\xi_t(\alpha_V):=&\left(s_t-\mu_s,f_t-\mu_f,(s_t-\mu_s)^2-\sigma_s^2,(s_t-\mu_s)(f_t-\mu_f)-\sigma_{sf},\right.\\&\left.(f_t-\mu_f)^2-\sigma_f^2\right)^\top,\end{aligned}$$

$$\begin{aligned}\xi_t(\alpha_R):=&\exp\left(-\frac{p_t(\alpha_R)}{R_p(\alpha_R)}\right)f_t+\frac{1}{R_p(\alpha_R)^2}\mathbb{E}\left[\exp\left(-\frac{p_t(\alpha_R)}{R_p(\alpha_R)}\right)p_t(\alpha_R)f_t\right]\\&\times J(R_p(\alpha_R))\xi_t(R_p(\alpha_R)),\end{aligned}$$

and

$$\xi_t(R_p(\alpha)):=\exp\left(-\frac{p_t(\alpha)}{R_p(\alpha)}\right)-1.$$

For the null hypothesis: $\alpha_R=\alpha_V$, the significance test statistic is of the form:

$$S(\hat{\alpha}_R,\hat{\alpha}_V):=\sqrt{T}(\hat{\alpha}_R-\hat{\alpha}_V)/\hat{\omega}(\hat{\alpha}_R,\hat{\alpha}_V)^{1/2},$$

where $\hat{\omega}(\hat{\alpha}_R,\hat{\alpha}_V)$ is an HAC estimator for $\omega(\alpha_R,\alpha_V):=\lim_{T\to\infty}\mathrm{var}\left[T^{-1/2}\sum_{t=1}^T\zeta_t(\alpha_R,\alpha_V)\right]$ with

$$\zeta_t(\alpha_R,\alpha_V):=J(\alpha_R)\xi_t(\alpha_R)-J(\alpha_V)\xi_t(\alpha_V),$$

and has the asymptotic null distribution: $S(\hat{\alpha}_R,\hat{\alpha}_V)\xrightarrow{d}N(0,1)$. For the null hypothesis: $H_o:\alpha_R=\alpha_o$, with a known α_o (such as $\alpha_o=0$ or $\alpha_o=1$), the significance test statistic is of the form:

$$S(\hat{\alpha}_R,\alpha_o):=\sqrt{T}(\hat{\alpha}_R-\alpha_o)/\hat{\Omega}(\hat{\alpha}_R)^{1/2},$$

where $\hat{\Omega}(\hat{\alpha}_R) := \widehat{J}(\hat{\alpha}_R)^2\hat{\Sigma}(\hat{\alpha}_R)$ with $\widehat{J}(\hat{\alpha}_R) := -\widehat{R}_p(\hat{\alpha}_R)\left[T^{-1}\sum_{t=1}^{T}\exp\left(-\frac{p_t(\hat{\alpha}_R)}{\widehat{R}_p(\hat{\alpha}_R)}\right)f_t^2\right]^{-1}$ and $\hat{\Sigma}(\hat{\alpha}_R)$ denotes an HAC estimator for $\Sigma(\alpha_R) := \lim_{T\to\infty}\mathrm{var}\left[T^{-1/2}\sum_{t=1}^{T}\xi_t(\alpha_R)\right]$, and has the asymptotic null distribution: $S(\hat{\alpha}_R, \alpha_o)\xrightarrow{d}N(0,1)$. Let $\hat{\xi}_t(\hat{\alpha}_R)$ be the sample analog of $\xi_t(\alpha_R)$. In computing $S(\hat{\alpha}_R, \alpha_o)$, we make use of the HAC estimator $\hat{\Sigma}(\hat{\alpha}_R) := \sum_{\ell=1-T}^{T-1}\kappa\left(\frac{\ell}{\tau}\right)\hat{\gamma}(\ell|\hat{\theta})$, with the lag-$\ell$ sample autocovariance:

$$\hat{\gamma}(\ell|\hat{\alpha}_R) := \frac{1}{T}\sum_{t=|\ell|+1}^{T}\left(\hat{\xi}_t(\hat{\alpha}_R) - \frac{1}{T}\sum_{t=1}^{T}\hat{\xi}_t(\hat{\alpha}_R)\right)\left(\hat{\xi}_{t-|\ell|}(\hat{\alpha}_R) - \frac{1}{T}\sum_{t=1}^{T}\hat{\xi}_t(\hat{\alpha}_R)\right),$$

the kernel function $\kappa(\cdot)$, and the smoothing parameter τ; see, e.g., Newey and West (1987) and Andrews (1991). In this paper, we set $\kappa(\cdot)$ as the Bartlett kernel and set $\tau = 3$, as in the empirical study of Hong et al. (2007). The HAC estimator $\hat{\omega}(\hat{\alpha}_R, \hat{\alpha}_V)$ is also computed in a similar way. The 95% and 99% critical values of the test statistics are, respectively, about ±1.96 and ±2.576.

References

Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817–858.

Aumann, R.J., Serrano, R., 2008. An economic index of riskiness. Journal of Political Economy 116, 810–836.

Bai, J., Ng, S., 2005. Tests for skewness, kurtosis, and normality for time series data. Journal of Business & Economic Statistics 23, 49–60.

Bali, T.G., Cakici, N., Chabi-Yo, F., 2011. A generalized measure of riskiness. Management Science 57, 1406–1423.

Bodie, Z., Rosansky, V.I., 1980. Risk and return in commodity futures. Financial Analysts Journal 36, 27–39.

Breiman, L., Meisel, W., Purcell, E., 1977. Variable kernel estimates of multivariate densities. Technometrics 19, 135–144.

Brooks, C., Henry, O.T., Persand, G., 2002. The effect of asymmetries on optimal hedge ratio. Journal of Business 75, 333–352.

Chen, S.S., Lee, C.F., Shrestha, K., 2001. On the mean-generalized semi-variance approach to determining the hedge ratio. Journal of Futures Markets 21, 581–598.

Chen, S.S., Lee, C.F., Shrestha, K., 2003. Futures hedge ratios: a review. Quarterly Review of Economics and Finance 43, 433–465.

Cheung, C.S., Kwan, C.C.Y., Yip, C.Y., 1990. The hedging effectiveness of options and futures: a mean-Gini approach. Journal of Futures Markets 21, 581–598.

De Jong, A., De Roon, F., Veld, C., 1997. Out-of-sample hedging effectiveness of currency futures for alternative models and hedging strategies. Journal of Futures Markets 17, 817–837.

Ederington, L.H., 1979. The hedging performance of the new futures markets. Journal of Finance 34, 157–170.

Foster, D.P., Hart, S., 2009. An operational measure of riskiness. Journal of Political Economy 117, 785–814.

Gallant, R., Tauchen, G., 1989. Seminonparametric estimation of conditionally constrained hetergeneous processes: asset pricing applications. Econometrica 57, 1091–1120.

Homm, U., Pigorsch, C., 2012. Beyond the Sharpe ratio: an application of the Aumann–Serrano index to performance measurement. Journal of Banking and Finance 36, 2274–2284.

Hong, Y., Tu, J., Zhou, G., 2007. Asymmetries in stock returns: statistical tests and economic evaluation. Review of Financial Studies 5, 1547–1581.

Howard, C.T., D'Antonio, L.J., 1984. A risk-return measure of hedging effectiveness. Journal of Financial and Quantitative Analysis 19, 101–112.

Jondeau, E., Rockinger, M., 2001. Gram-Charlier densities. Journal of Economic Dynamics and Control 25, 1457–1483.

Kolb, R.W., Okunev, J., 1993. Utility maximizing hedge ratios in the extended mean Gini framework. Journal of Futures Markets 13, 597–609.

Lien, D., Luo, X., 1993. Estimating multiperiod hedge ratios in cointegrated markets. Journal of Futures Markets 13, 909–920.

Lien, D., Tse, Y.K., 1998. Hedging time-varying downside risk. Journal of Futures Markets 18, 705–722.

Newey, W.K., McFadden, D., 1994. Large sample estimation and hypothesis testing. In: Engle, R., McFadden, D. (Eds.), The Handbook of Econometrics, vol. 4. North-Holland, Amsterdam.

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Schreiber, A., 2012. An Economic Index of Relative Riskiness. Working Paper, Hebrew University of Jerusalem.

Schulze, K., 2010. Existence and Computation of the Aumann–Serrano Index of Riskiness. Working Paper, McMaster University.