MSL
ACADEMIC ENDEAVORS

Cleveland State University

EngagedScholarship@CSU

Business Faculty Publications

Monte Ahuja College of Business

1-31-2017

Determining the multi-scale hedge ratios of stock index futures using the lower partial moments method

Jun Dai

Haigang Zhou
Cleveland State University, h.zhou16@csuohio.edu

Follow this and additional works at: https://engagedscholarship.csuohio.edu/bus_facpub

Part of the Finance and Financial Management Commons

How does access to this work benefit you? Let us know!

Publisher's Statement

Recommended Citation

Determining the multi-scale hedge ratios of stock index futures using the lower partial moments method

Jun Dai   Haigang Zhou   Shaoquan Zhao

A B S T R A C T

This paper considers a multi-scale future hedge strategy that minimizes lower partial moments (LPM). To do this, wavelet analysis is adopted to decompose time series data into different components. Next, different parametric estimation methods with known distributions are applied to calculate the LPM of hedged portfolios, which is the key to determining multi-scale hedge ratios over different time scales. Then these parametric methods are compared with the prevailing nonparametric kernel metric method. Empirical results indicate that in the China Securities Index 300 (CSI 300) index futures and spot markets, hedge ratios and hedge efficiency estimated by the nonparametric kernel metric method are inferior to those estimated by parametric hedging model based on the features of sequence distributions. In addition, if minimum–LPM is selected as a hedge target, the hedging periods, degree of risk aversion, and target returns can affect the multi-scale hedge ratios and hedge efficiency, respectively.

1. Introduction

Rational investors like portfolio managers and some individuals often enter the futures markets with predetermined hedging horizons that vary from seconds to months and beyond. Understanding the relationship between optimal hedge ratios and time horizons is the key to fully exerting the functions of stock index futures. However, previous literature seldom focuses on the effects of different time horizons on optimal hedge ratios. For example, the traditional Ordinary Least Squares (OLS) and recently widely used Bivariate Generalized Autoregressive Conditional Heteroskedasticity (BV-GARCH) models apply both the volatility and the correlation of spot and futures prices in a single period to calculate optimal hedge ratios over different time horizons. Any ignorance of the impact over the hedging horizon on the optimal hedge ratio could be detrimental to decision making and undermine hedging effectiveness.

In and Kim [1] were the first to employ the wavelet analysis model to study the relationship between stock and futures markets and to compute hedge ratios over different hedging horizons. Lien and Shrestha [2] decomposed a time series into different scales using the maximal overlap discrete wavelet transform (MOWDT). Multi-scale optimal hedge ratios were then calculated by OLS regressions of the spot wavelet coefficients on the futures wavelet coefficients at different scales. Following the model introduced by Lien and Shrestha [2], Chen et al. [3] examined the performance of multi-scale hedge ratios on the future markets of the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX). The empirical results indicated that both hedge ratios and hedge effectiveness went up when hedging horizons increased, and hedge ratios continued to climb until they approached the numerical value of one. This conclusion is consistent with findings in Lien and Shrestha [2]. Conlon and Cotter [4] argued that a utility maximization framework, which incorporates the impacts of degree of risk aversion and hedging horizons on optimal hedge ratios, should be used to determine multi-scale hedge ratios.

In addition, with the growth of investors' risk awareness, drawbacks of variance as a risk measure standard have gradually been perceived because variance, as a traditional risk measure instrument, implies that agents view positive and negative deviation from the expected return as equally undesirable. However, agents frequently perceive risk as a failure to achieve a certain level of return. In this context, downside risk measures, assuming that returns below a certain reference level involve risk and returns above a certain reference level represent better investment opportunities, can be highly relevant. To the best of our knowledge, Sun and He [5] were the first scholars to combine multi-scale hedge ratios with the target of the minimum-LPM hedge. However, their preference to estimate the joint density distribution is a nonparametric method that boasts wide suitability but less efficiency.

In this study, we fill the gap in literature by introducing the parametric estimation method with higher moments including coskewness and cokurtosis to compute multi-scale hedge ratios based on LPM. The new model makes full use of the information from the population like the parametric approach does, reflects the characteristics of non-normal distribution of time series, and allows the estimation of hedge ratios for different horizons.

Our primary empirical results show that, in the CSI 300 index futures and spot markets, the nonparametric universal kernel density method is inferior to parametric methods on the distribution features of time series. If LPM is selected as hedge target, the hedging periods, degree of risk aversion, and target returns can affect the multi-scale hedge ratios and hedge efficiency, respectively.

In the following text, Section 2 constructs the hedging model; Section 3 processes the data from the CSI300 index futures and spot markets; Section 4 expounds upon the empirical analysis and Section 5 draws conclusive findings.

2. Methodology

2.1. A short synopsis of wavelet multi-scale analysis

This section starts from a brief synopsis of the wavelet multi-scale analysis adopted from Percival and Walden [6]. Wavelet analysis, a recently developed signal processing technique in time and frequency domains, has been widely used in the financial engineering fields.

To analyze local features, the given time series should be completely decomposed into different time horizons or frequency components by automatically expanding and contracting the movement of two basic wavelet functions: the father wavelet (or scaling function) $\psi_{j,k}(t)$ and the mother wavelet (or wavelet function) $\phi_{J,k}(t)$, which can be scaled and translated to form a basis for the Hilbert space $L^2(\Re)$ of square integrable functions. The father and mother wavelets are formally defined by the following functions:

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \Phi\left(\frac{t - 2^j(k-1)}{2^j}\right) \tag{1a}$$

$$\phi_{J,k}(t) = \frac{1}{\sqrt{2^j}} \Phi\left(\frac{t - 2^J(k-1)}{2^J}\right) \tag{1b}$$

where $j = 1, 2, \ldots, J$ is the scaling parameter in a J-level decomposition, k is a translation parameter that determines the location of analysis, 2^j is the scale used to evaluate $\Phi(t)$, and t represents time. Given a continuous time series $x(t)$, the wavelet coefficients $d_{j,k}$ and scales coefficients $B_{J,k}$ can be specified by:

$$d_{j,k} = \int x(t) \, \psi_{j,k} dt, \quad j = 1, 2, \ldots, J \tag{2a}$$

$$B_{J,k} = \int x(t) \, f_{J,k} dt \tag{2b}$$

where $d_{j,k}$ is the series of detail coefficients obtained from the mother wavelet at all scales from 1 to J, the maximal scale, and $B_{J,k}$ is the series of coefficients from the father wavelet at maximal scale J. For mathematical convenience, the discrete points have the sample size of N, which is assumed to be divisible by 2^J and may be represented by $w = WX$, where W is

an $N \times N$ orthonormal matrix that summarizes the transformations listed above. Then the series of wavelet coefficients w can be defined by the different subvectors of wavelet coefficients as follows:

$$w = \left[w_1^T, w_2^T, \ldots, w_J^T, b_J^T \right] \tag{3}$$

where w_j and b_J are separate $\left(N/2^j\right) \times 1$ and $\left(N/2^J\right) \times 1$ dimensional vectors, w_j is the arbitrary jth series of wavelet coefficients used to reflect the deviation of the detailed signals from the long trend, and b_J is the series of scaling coefficients that reflect the long-term variations. These discrete wavelet and scaling coefficients are fully equivalent to the information contained in the original series in that a time series can be perfectly reconstructed from these coefficients. This process defines what is known as a multi-resolution analysis (MRA). However, maximal overlap discrete wavelet transforms (MODWT) is a major improvement over Discrete Wavelet Transform (DWT) and is more widely used in practice. The MODWT method does not require N to be an integer multiple of 2^J and overcomes the drawback that DWT is sensitive to the selection of a starting point. Thus, it is much easier for MODWT to reveal the features of an original series in the multi-scale analysis. This paper adopts MODWT instead of DWT, and all wavelet analyses are performed by MATLAB software.

2.2. Minimum-LPM hedge ratios

In recent years, considerable interest in risk measures has been based on one-sided moments. Suppose an individual has a portfolio that generates a random return r_p. Let $F\left(\cdot\right)$ denote the probability distribution function of r_p. An outcome larger than target return c is desirable, and the uncertain shortfalls below c are called the downside risk. LPM serves as an intuitive downside risk measure. For further discussion, see Refs. [7,8]. Some studies have applied the LPM methodology to futures hedging [9].

For any positive integer n, the nth order lower partial moment of r_p is defined as:

$$LPM_n\left(c, r_p\right) = E\left[\max\left(0, c - r_p\right)\right]^n = \int_{-\infty}^c \left(c - r_p\right)^n \, dF\left(r_p\right) \tag{4}$$

where n indicates investors' desire to avoid extreme shortfalls from the selected target return. $r_p = r_s - h \times r_f$ denotes portfolio returns. r_s and r_f are the respective spot and futures logarithmic rates of return. h is the hedge ratio. Upon taking the derivative in Eq. (4), we find that the optimal hedge ratio h^* must satisfy the following first-order condition:

$$nE\left\{\left[\max\left(0, c - r_s + hr_f\right)\right]^{n-1} r_f\right\} = 0, \quad n > 1. \tag{5}$$

The second-order condition is obtained by taking the derivative of Eq. (5) with respect to h,

$$n\left(n - 1\right) E\left\{\left[\max\left(0, c - r_s + hr_f\right)\right]^{n-2} r_f^2\right\} > 0. \tag{6}$$

One can see the first-order condition may be satisfied at unary optimal h^*. However, due to the complexity of Eq. (4), there is no explicit analytical expression for the optimal hedge ratio, so numerical methods are needed. The key to determining the numerical methods lies in estimating the unknown distribution of the joint probability density function $F\left(r_p\right)$ of spot and futures returns. Current approaches include the assumption of $F\left(r_p\right)$ following the joint normal distribution [10] or the adoption of nonparametric kernel estimation approaches [11,12]. However, in practice, some drawbacks still exist in the above two approaches.

First, numerous empirical studies have proven that r_s and r_f rarely obey the normal distribution strictly and instead display features like peaks, fat tails and bias, etc. Therefore, the normal distribution assumption often leads to biased estimates of LPM. Second, although the nonparametric estimation methods proposed by Ref. [10] do not need specific function forms, they can reflect the characteristics of abnormal distribution to some degree. When compared to bivariate random distribution $D\left(0, 1, S, K\right)$ coupled with the coskewness and cokurtosis metrics, $D\left(0, 1, S, K\right)$ can obviously better describe abnormal distribution for series r_s and r_f [13]. As stated above, this paper introduces bivariate random distribution $D\left(0, 1, S, K\right)$ to denote probability as follows:

$$\begin{cases} \mu = \begin{pmatrix} \mu_s \\ \mu_f \end{pmatrix} = E\begin{pmatrix} r_s \\ r_f \end{pmatrix}, \qquad H = \begin{pmatrix} h_{ss} & h_{sf} \\ h_{fs} & h_{ff} \end{pmatrix} = \text{cov}\left(r_s, r_f\right) \\[12pt] \eta = \begin{pmatrix} \eta_s \\ \eta_f \end{pmatrix} = H^{-1/2}\begin{pmatrix} r_s - \mu_s \\ r_f - \mu_f \end{pmatrix} \\[12pt] \eta \sim D \begin{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} s_{sss} & s_{ssf} & s_{sfs} & s_{sff} \\ s_{fss} & s_{fsf} & s_{ffs} & s_{fff} \end{pmatrix} \\ \begin{pmatrix} k_{ssss} & k_{sssf} & k_{ssfs} & k_{ssff} & k_{sfss} & k_{sfsf} & k_{sffs} & k_{sfff} \\ k_{fsss} & k_{fssf} & k_{fsfs} & k_{fsff} & k_{ffss} & k_{ffsf} & k_{fffs} & k_{ffff} \end{pmatrix} \end{pmatrix} \end{cases} \tag{7}$$

where μ_i $(i = s, f)$ is the expectation of spot or futures returns, H is the matrix of covariance, η is the standardized rate of returns, and $s_{ijm} = E\left[(r_i - \mu_i)(r_j - \mu_j)(r_m - \mu_m)\right]$ $(i, j, m = s, f)$ is a skewness matrix if $i = j = m$ and becomes a coskewness matrix once i, j and m are not the same. Likewise, $k_{ijml} = E\left[(r_i - \mu_i)(r_j - \mu_j)(r_m - \mu_m)(r_l - \mu_l)\right]$ $(i, j, m, l = s, f)$ denotes a kurtosis matrix if $i = j = m = l$, and becomes a co-kurtosis matrix once i, j, m and l are not the same.

With no explicit analytical expression for $D(0, 1, S, K)$, we have to use the research findings in Ref. [13] as reference points and the cut-off expression of Gram–Charlier's expansion of normal density function at the fourth order as the approximation for $D(0, 1, S, K)$. We then substitute the approximate expression into Eq. (4) and yield the LPM of hedge portfolio as the following:

$$LPM_n\left(c, r_p\right) = \int_{-\infty}^{\frac{c-\mu_p}{\sigma_p}} \left(c - \mu_p - \sigma_p \eta_p\right)^n \frac{1}{\sqrt{2\pi}} e^{-\frac{\eta_p^2}{2}} \bullet \frac{\left[1 + \frac{s_p}{3!}\left(\eta_p^3 - 3\eta_p\right) + \frac{k_p - 3}{4!}\left(\eta_p^4 - 6\eta_p^3 + 3\right)\right]^2}{1 + \frac{s_p}{3!} + \frac{k_p - 3}{4!}} dr_p$$

$$s_p = E\left(r_p - \mu_p\right)^3 / \sigma_p^4 = E\left[(r_s - \mu_s) - h_p\left(r_f - \mu_f\right)\right]^3 / \sigma_p^3 = s_{sss} - h_p s_{ssf} + h_p^2 s_{sff} - h_p^3 s_{fff} / \sigma_p^3$$

$$k_p = E\left(r_p - \mu_p\right)^4 / \sigma_p^4 = \left(k_{ssss} - 4h_p k_{sssf} + 6h_p^2 k_{ssff} - 4h_p^3 k_{sfff} + h_p^4 k_{ffff}\right) / \sigma_p^4$$

(8)

where η_p denotes the normalized returns of the hedged portfolio, s_p and k_p are respectively the matrices of coskewness and cokurtosis. As $LPM_n\left(c, r_p\right)$ is the function of μ_p, σ_p, s_p and k_p of the hedged portfolio, and all of these variables are also functions of hedge ratio h via Eq. (8), the optimal hedge ratio h_* can be determined as the value that minimizes the $LPM_n\left(c, r_p\right)$.

Since the MODWT coefficients at different scales may show various distribution features, the minimum-LPM hedge ratios h_* over different time scales should be calculated differently. The other methods used to calculate h_* in this paper will only be introduced briefly as follows.

In the first scenario, if the test shows that the joint probability density function $F\left(r_p\right)$ follows a normal distribution with $\mu_p = \mu_s - h\mu_f$ and $\sigma_p^2 = \sigma_s^2 - 2h\sigma_{sf} + h^2\sigma_f^2$, then after substituting the normal distribution expression into Eq. (4) and further simplifying the expression, we can draw the implicit formulation of futures optimal hedging ratio h_* with respect to the three degrees of risk aversion $n = 1, 2$, or 3 as follows [14]:

Scenario 1 $(n = 1)$:

$$LPM_1 = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu_z}{\sigma}}^{\infty} \left(\mu_z + \sigma\varepsilon\right) \exp\left(-\frac{\varepsilon^2}{2}\right)$$

$$\frac{\partial LPM_1}{\partial h} = \frac{1}{\sqrt{2\pi}} \left[\mu_f \int_{-\frac{\mu_z}{\sigma}}^{\infty} \exp\left(-\frac{\varepsilon^2}{2}\right) d\varepsilon + \frac{1}{\sigma}\left(h\sigma_f^2 - \sigma_{sf}\right) \exp\left(-\frac{\mu_z^2}{2\sigma^2}\right)\right].$$

(9)

Scenario 2 $(n = 2)$:

$$LPM_2 = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu_z}{\sigma}}^{\infty} \left(\mu_z + \sigma\varepsilon\right)^2 \exp\left(-\frac{\varepsilon^2}{2}\right) d\varepsilon$$

$$\frac{\partial LPM_2}{\partial h} = \frac{2}{\sqrt{2\pi}} \left\{\sigma\mu_f \exp\left(-\frac{\mu_z^2}{2\sigma^2}\right) + \left[\left(\sigma_f^2 + \mu_f^2\right)h - \sigma_{sf} + (c - \mu_s)\mu_f\right] \int_{-\frac{\mu_z}{\sigma}}^{\infty} \exp\left(-\frac{\varepsilon^2}{2}\right) d\varepsilon\right\}.$$

(10)

Scenario 3 $(n = 3)$:

$$LPM_2 = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu_z}{\sigma}}^{\infty} \left(\mu_z + \sigma\varepsilon\right)^3 \exp\left(-\frac{\varepsilon^2}{2}\right) d\varepsilon$$

$$\frac{\partial LPM_2}{\partial h} = \frac{3}{\sqrt{2\pi}} \left\{\sigma\left(\mu_f\mu_z + 2h\sigma_f^2 - 2\sigma_{sf}\right) \exp\left(-\frac{\mu_z^2}{2\sigma^2}\right) + \left[\mu_f\left(\sigma^2 + \mu_z^2\right) + 2\mu_z\left(h\sigma_f^2 - \sigma_{sf}\right)\right]\right.$$

$$\left. \times \int_{-\frac{\mu_z}{\sigma}}^{\infty} \exp\left(-\frac{\varepsilon^2}{2}\right) d\varepsilon\right\}$$

(11)

where $\varepsilon = (z - \mu_z)/\sigma$ represents a standardized variable, and $z \sim N\left(\mu_z, \sigma_z^2\right)$ is a random variable given by $z = c - r_p = c - r_s + h \cdot r_f$. Here c is the target return. From Eqs. (9)–(11), it can be seen that the implicit expression of $\partial LPM_n/\partial h$ is the unary function of h. Thus, the numerical method is needed to find h^*, which corresponds to the point where $\partial LPM_n/\partial h$ is closest to zero.

In the second scenario, assuming that $F\left(r_p\right)$ does not follow normal distribution, we can employ the nonparametric kernel density method to determine the optimal futures hedge ratio. For any given hedge ratio, the hedge portfolio returns

series $m_1, m_2, \ldots, m_N$, where N is the sample size, can be constructed from data series r_s and r_f. The probability density function of the portfolio return that is less than m_i at given point y, denoted by $f(y)$, can be estimated by:

$$f(y) = (1/N\theta) \sum_{i=1}^{N} g\left((y - m_i)/\theta\right) \tag{12}$$

where $g(y, m_i, \theta)$ is an alternative kernel function. Here, we choose $g(y, m_i, \theta)$ to be the normal kernel. θ is called the bandwidth, which controls the degrees of smoothing. According to the rule of thumb proposed by Silverman [15], based on the ruler of Mean Integrated Squared Error (MISE), the bandwidth can be expressed in a more simplistic manner. In particular, for the standard second-order normal kernel, the bandwidth can be expressed as $\theta \approx 1.06\sigma n^{-1/5}$, where σ is the sample standard deviation and n is the coefficient of risk aversion. Consequently, by substituting Eq. (12) into Eq. (4), LPM can be estimated by:

$$LPM_n\left(c, r_p\right) = \int_{-\infty}^{c} (c - r)^n \cdot \frac{1}{N\theta(h)} \sum_{t=1}^{N} g\left(\frac{r - r_t^p}{\theta(h)}\right) \mathrm{d}r \tag{13}$$

where $g(\mu) = \left(1/\sqrt{2\pi}\right) \exp\left(-\mu^2/2\right)$ is the density function of a standard normal random variable. In Eq. (13), obviously $LPM_n\left(c, r_p\right)$ is the function of h. Then we can employ a grid search method to obtain the minimum-LPM hedge ratio. However, in reality, calculation of the nonparametric kernel estimation method is often complicated. Therefore, Homaifar and Graddy [16] suggested that the kernel estimation method would be similar to the empirical distribution function method when the bandwidth was small enough. With the employment of the empirical distribution function method, the LPM can be calculated as follows:

$$LPM_n\left(c, r_p\right) = \sum_{m_i < c} (1/N)\, (c - m_i)^n. \tag{14}$$

2.3. Hedging performances

In order to evaluate the effectiveness of the minimum-LPM hedge ratios estimated by the parametric and nonparametric methods, we introduce the percentage reduction in LPM ρ and the reward-to-semivariance (R/SV) ratios (Nawrocki, 1983) [17] as hedging performance measures:

$$\rho = \frac{LPM_n\left(c, r_p\left(h^*\right)\right) - LPM_n\left(c, r_s\right)}{LPM_n\left(c, r_s\right)} \tag{15}$$

where $LPM_n\left(c, r_s\right)$ denotes the LPM when there is no futures hedging for the given coefficient of risk aversion n and target return c; $LPM_n\left(c, r_p\left(h^*\right)\right)$ is the minimum-LPM at the optimal hedge ratio h^* with the same above n and c. In general, ρ lies between zero and one with a bigger decimal indicating a better hedging performance.

Ratio R/SV measures the excess return per unit of deviation in an investment asset:

$$R/SV = \left(E\left(r_p\right) - r_f\right)/LPM\left(c, n, h\right) \tag{16}$$

where $E\left(r_p\right)$ is the expected return of the hedged portfolio, r_f is the risk-free rate of interest. The higher the ratio, the higher the portfolio return is per unit of risk.

3. Data description

The data employed in this paper pertain to the CSI 300 index futures and spot markets. The CSI 300 index, which was launched on April 8, 2005, is a capitalization-weighted stock market that is designed to replicate the performance of 300 stocks traded in Shanghai and Shenzhen stock exchanges, accounting for approximately 70% of market capitalization of both stock exchanges. To provide investors with a tool to hedge risk in the stock market, the CSI 300 index futures started trading in China Financial Futures Exchange (CFFEX) on April 16, 2010. Our datasets consist of daily observations of spot index and futures prices from April 16, 2010, to October 30, 2015. There are four CSI 300 index futures contracts (current calendar month, the next calendar month and the following two calendar quarterly months) being traded simultaneously. To calculate the returns of the futures contracts, we use the closing prices of the nearest and usually the most liquid contract. To avoid thin markets and expiration effects, we roll over to the next nearest contract one week prior to the expiration of the current contract. After matching spot index prices with index futures prices, there are only 1343 prices observations for each time series, which are log differenced as $\log(P_t) - \log(P_{t-1})$ for convenience of analysis.

As outlined in Section 2.1, the spot and futures logarithmic returns are decomposed by MODWT. For this study, we selected the least asymmetric (LA) wavelet, which exhibits near symmetry at about the filter midpoint and has the property of aligning wavelet coefficients with the unfiltered time series. LA filters are designed in even widths, and the optimal filter
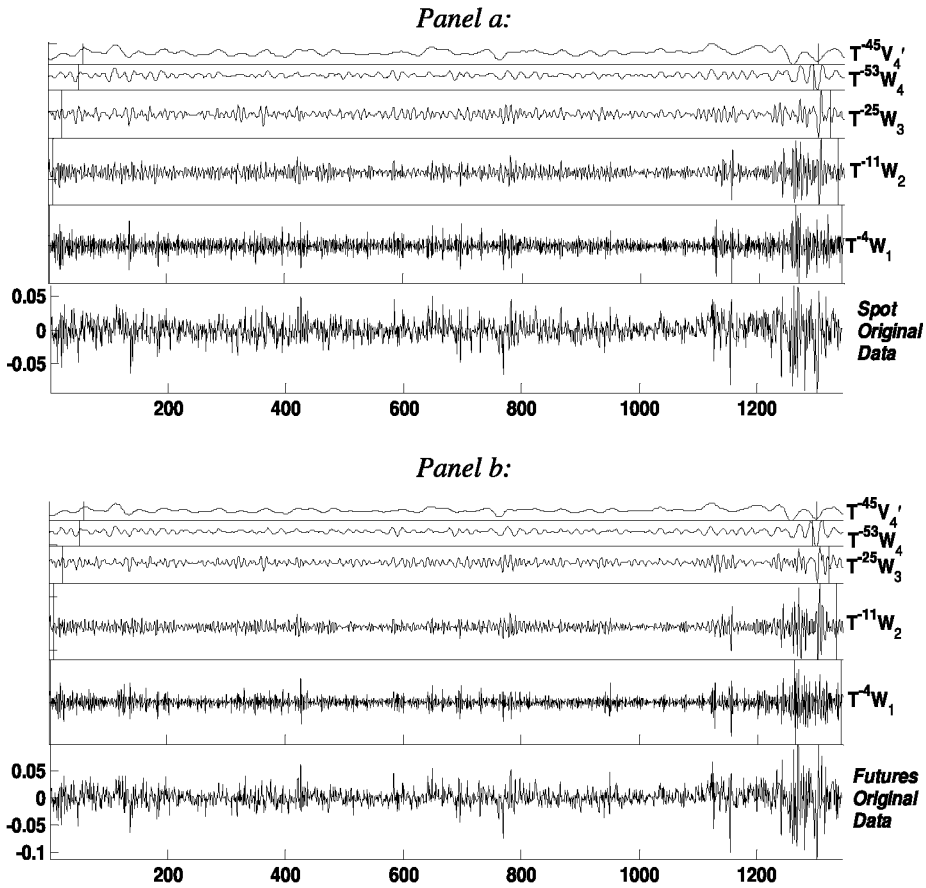
Fig. 1. MODWT MRA of the spot returns (Panel a) and futures returns (Panel b) for the CSI300 futures and spot markets. These plots are obtained from a wavelet Daubechie filter up to four time scales. $T^{-v}W_j$ and $T^{-v}V_J$ are wavelet coefficients and scaling coefficients respectively. T^{-v} is the shift vector that ensures the wavelet filter is a zero-phase filter.

width is dependent upon characteristics of the signal and the length of the data series. Here, the filter width is 8, which is the optimal value for financial series [2]. Considering the length of the sample is only 1343 days, the upper bound of scales is 10 ($2^{10} < 1343 < 2^{11}$). In practice, we do not need to consider all possible scales in the decomposition because large scales often result in smooth detail series, thus failing to provide much useful information about the original time series. Since wavelet coefficients at scales larger than 4 are too smooth, it is difficult to define a specific optimal hedge ratio in most situations. Consequently, we use 4-scale decomposition for each dataset and produce four components, which correspond to horizons including scale 1 associated with 2–4 day dynamics, scale 2 with 4–8 day dynamics, scale 3 with 8–16 day dynamics and scale 4 with 16–32 day dynamics.

Fig. 1 shows the multi-resolution approximation (MRA) of the logarithmic returns of spot (Panel a) and future (Panel b) using various time scales (four different wavelet details). MRA could be viewed as a filter version of the series that retains the most important parts of the series, but denoises the series to a greater or lesser degree. Seeing the set of series in Fig. 1, we may observe that as the time scale increases from the lower time scales (W_1) to the highest time scale (W_4), wavelet coefficients become gradually smoother. This means short-term noise vanishes as the scale increases, and the high-scale wavelet coefficients can remain the long-term relationships between spot and futures markets very well.

Table 1 represents the summary statistics, which includes the mean, standard deviation, skewness, kurtosis and Jarque–Bera normality test for the two assets over each scale. From Table 1, we observe that the original returns and wavelet coefficients, at scales from 1 to 4, display obviously different distribution features. Specifically, for the original returns and the first-level wavelet coefficients series of the both CSI 300 spot and futures log returns, it is easy to find abnormal distribution characteristics, such as peaks, fat tails, bias, etc. With the growth of scales, the wavelet coefficients gradually converge to the point of normal distribution. As to the wavelet coefficients at scale 2, although the hypothesis of normality is rejected by the Jarque–Bera test, their distributions are very close to normality. For wavelet coefficients at scales 3 or 4, reflecting the difference of long-term weighted averages, normality cannot be rejected by the Jarque–Bera tests. Thus, in the CSI 300 index futures and spot markets, the calculation of multi-scale hedge ratios with the LPM approach should employ different estimation methods according to the distribution features of wavelet coefficients at each scale. Employment of this method can help achieve better estimations of the optimal hedge ratio.

Table 1
Descriptive statistics for log-returns of futures and spot series at different time-scales for the CSI 300 Equity Index.

	Series	Mean(10^{-4})	Standard deviation%	Skewness	Kurtosis	Jarque–Bera
Spot	Original data	−5.061	1.495	−0.354	6.102	333.655*
	Scale 1	0.00000	1.105	−0.071	7.462	656.87*
	Scale 2	0.00000	0.719	0.072	3.438	6.983**
	Scale 3	0.00000	0.494	−0.037	3.203	1.539
	Scale 4	0.00000	0.326	−0.057	3.093	0.717
Futures	Original data	−5.382	1.475	−0.087	5.680	237.81*
	Scale 1	0.00000	1.065	−0.005	4.511	75.212*
	Scale 2	0.00000	0.719	0.099	3.938	30.30*
	Scale 3	0.00000	0.501	−0.105	3.264	3.745
	Scale 4	0.00000	0.336	0.021	3.237	1.917

This table reports the descriptive statistics and Jarque–Bera tests for the continuous returns of the spot and futures prices of the CSI 300 indices. The sample period is from April 16, 2010, to October 30, 2015, with a total of 1343 observations. The mean and deviation of each series are given in ten-thousandth and percentage terms respectively, while a skewness of zero indicates no skewness and a kurtosis of 3 indicates no excess kurtosis. The Jarque–Bera statistic tests the null hypothesis that the distribution is normal, Scale i ($i = 1, 2, 3$ and 4) is the wavelet coefficients of spot or futures returns at scales from 1 to 4.
* Denotes statistical significance at 1%.
** Denotes statistical significance at 5%.

Further, Table 1 also summarizes some statistical properties of the data series. For both the original data and the estimated wavelet coefficients at each scale, parameters (such as mean, standard deviation, skewness and kurtosis) vary across different hedge horizons. Since all of the above parameters are determining factors of LPM, as defined in Eq. (8), generalizing the hedge ratio calculated from a single hedge horizon to different hedge horizons will inevitably lead to major errors and bad investment decisions.

4. Empirical results

Depending on whether the original data and the wavelet coefficients at different scales follow a normal distribution, two approaches are employed to estimate LPM. For original data and wavelet coefficients at scales 1 or 2 that do not follow a normal distribution, LPM can be estimated by Eq. (8), which includes the matrices of coskewness and cokurtosis. Because the wavelet coefficients at scales 3 or 4 follow a normal distribution, LPM can be directly calculated via Eqs. (9)–(11). Then we use numerical methods to calculate multi-scale minimum-LPM hedge ratios and evaluate the performance of the above hedge ratios.

To illustrate the advantage of parametric estimations, we also calculate optimal hedge ratios using the nonparametric method and compare them with those estimated using the parametric method. As a nonparametric way to estimate the probability density function of a random variable, Kernel Density Estimation (KDE) works better because it is not dependent upon the population fitting any parameterized distribution. Therefore, it can be applied to the estimation of LPM for any time series. However, once the understanding of the population is sufficient to determine the distribution of time series, the nonparametric estimation method becomes inferior to the parametric estimation method because the former usually fails to take full advantage of the information available in the sample. Therefore, sometimes estimation in nonparametric ways will display lower efficiency and be considered much more conservative.

To derive the optimal hedge ratios, we set a list of possible target returns, $c = -1\%$, 0% or 1% and a list of risk-aversion coefficients, $n = 1, 2$ or 3. In addition, the percentage reduction in LPM ρ and reward-to-semivariability R/SV are adopted to compare the hedge performance of minimum-LPM hedge ratios estimated respectively by the parametric and nonparametric methods. While estimating the R/SV, 3.47% is taken as the riskless return, which is the average benchmark six-month lending rate during the years 2010–2015 in the China Interbank Offered Rate. All estimation results are shown in Table 2.

First, consider changes in hedge ratios. As indicated in Lien and Shrestha [2], with increasing wavelet time scales, the wavelet coefficients will be smooth, implying that short-term noise in the market is canceled out as the wavelet time scale increases. Consequently, the 'true' underlying economic relationship between stock and futures prices will prevail in the long run, which suggests hedge ratios will increase when wavelet time scales go up. However, Table 2 shows that, only under the conditions of Targeted Return $c = -1\%$, 0% and Parametric Based on the Feature of Series, the multi-scale optimal hedge ratios (h^*) will increase monotonically, converging toward the long-horizon hedge ratio of one as hedging horizons (scales) increase. For the other target return $c = 1\%$ and Nonparametric Kernel Density Method, h^* does not obey the above-mentioned rule.

In addition, h^* shows close connections with degrees of risk aversion n and target returns c only in the original return series and the wavelet coefficients at short horizon scales 1 and 2. Specifically, h^* will go up as n increases and will go down as c increases, and the above law would become more significant as the time scales get smaller. So for a hedger who is willing to undertake small losses but is otherwise extremely cautious about large losses, he should choose the optimal multi-hedge ratios more strictly according to his degree of risk aversion and target returns, especially in short horizons. Furthermore, as can be seen in Table 2, under the same n, overall smaller hedge ratios indicating larger ventures can bring higher hedge portfolio returns.

Table 2
Estimated optimal hedge ratios and hedging performance for given (c, n).

		Risk aversion					
		$n = 1$		$n = 2$		$n = 3$	
		Parametric based on the feature of series	Nonparametric Kernel density method	Parametric based on the feature of series	Nonparametric Kernel density method	Parametric based on the feature of series	Nonparametric Kernel density method
A. Original return series (daily returns)							
c Targeted return	−1.0%	0.942[0.97]	0.972[0.94]	0.943[0.99]	1.038[0.88]	0.944[0.99]	1.233[0.83]
	0%	0.935[0.93]	0.958[0.70]	0.939[0.97]	0.992[0.87]	0.941[0.98]	1.083[0.85]
	1.0%	0.932[0.91]	0.927[0.19]	0.936[0.95]	0.968[0.57]	0.938[0.97]	1.027[0.77]
B. The wavelet coefficients (2 days) at scale 1							
c Targeted return	−1.0%	0.876[0.98] (0.825)	0.933[0.90] (0.617)	0.883[0.99] (98.179)	0.994[0.72] (29.607)	0.887[0.99] (839.450)	1.184[0.51] (−203.548)
	0%	0.844[0.94] (0.047)	0.830[0.64] (0.035)	0.857[0.97] (6.748)	0.947[0.79] (4.742)	0.866[0.98] (645.235)	0.959[0.70] (374.795)
	1.0%	0.833[0.92] (0.098)	0.827[0.08] (0.010)	0.844[0.94] (0.749)	0.923[0.41] (0.545)	0.852[0.96] (50.543)	0.935[0.61] (35.968)
C. The wavelet coefficients (4 days) at scale 2							
c Targeted return (%)	−1.0%	0.926[0.99] (1.173)	0.875[0.82] (0.879)	0.931[0.99] (97.654)	0.945[0.83] (81.090)	0.940[0.99] (780.493)	0.946[0.89] (706.541)
	0%	0.905[0.94] (0.067)	0.852[0.72] (0.059)	0.914[0.98] (7.805)	0.940[0.88] (6.693)	0.927[0.99] (670.101)	0.957[0.87] (519.524)
	1.0%	0.762[0.90] (0.032)	0.418[0.02] (0.019)	0.770[0.93] (1.341)	0.510[0.29] (1.319)	0.776[0.96] (76.291)	0.614[0.53] (60.943)
D. The wavelet coefficients (8 days) at scale 3							
c Targeted return (%)	−1.0%	0.954[0.99]	–	0.954[0.99]	–	0.954[0.99]	–
	0%	0.954[0.81] (0.108)	0.930[0.81] (0.098)	0.954[0.96] (14.226)	0.950[0.83] (13.933)	0.954[0.99] (144.685)	0.973[0.84] (127.305)
	0.1%	0.661[0.00] (0.453)	0.704[0.00] (0.427)	0.954[0.09] (1.554)	0.954[0.00] (1.143)	0.954[0.23] (107.265)	0.954[0.00] (77.239)
E. The wavelet coefficients (16 days) at scale 4							
c Targeted return (%)	−1.0%	0.964[0.99]	–	0.964[0.99]	–	0.964[0.99]	–
	0%	0.964[0.79] (0.071)	0.977[0.84] (0.069)	0.964[0.95] (8.364)	0.968[0.95] (8.232)	0.964[0.99] (712.359)	0.971[0.98] (669.421)
	1.0%	0.857[0.00]	–	0.964[0.18] (1.182)	0.964[0.18] (0.893)	0.964[0.40] (74.741)	0.965[0.40] (62.578)

This table reports multi-scale hedge ratios based on LPM from the nonparametric kernel density method and the parametric method based on the features of series. In each blank, the first number is the hedge ratio, the second number within the square brackets [] is the percentage reduction in LPM ρ, and the third number within the round brackets () is the semivariability (R/SV) ratio. c is the target return. n indicates the investor's willingness to avoid extreme shortfalls from the selected target return. '–' denotes incapability of determining a unique optimal hedge ratio. R/SV is not estimated when '–' appears, since it has no sense for comparison.

Secondly, we compare the performance of hedge ratios at different scales. Lien and Shrestha [2] took the minimum-variance as the target and found that hedging performance would improve itself with longer hedging horizons. However, as shown in Table 2, for the target of minimum-LPM, no obvious correlation would exist between hedging performance and time scales. In addition, only under the Parametric Based on the Feature of Series column, the hedging performance of percentage reduction in LPM ρ displays obviously significant correlation with target return c and degree of risk aversion n. Specifically, under each hedging horizon and any given n, when target return c increases, the hedging performance measure ρ strictly decreases. Similarly, when target returns c are the same, the higher the degree of risk aversion, the more efficient the hedge performance becomes.

In particular, wavelet coefficients at scales 3 or 4 can be regarded as longer hedging horizons, when c is set to −1%, ρ reaches 0.99. That means the downside risk of hedged portfolios has been completely eliminated. From April 16, 2010, to October 30, 2015, the CSI 300 index fell from 3356.3 points to 2249.1 points in the first four years and displayed a major price reversal in the most recent year. During such an overall declining market, it is obviously unrealistic for hedgers to reap overall positive returns on hedged portfolios, but if hedgers only seek to minimize risk, things will be materially different. In addition, the positive correlation between hedging performance and degrees of risk aversion indicates that stock index futures can perform better in hedging LPM risk with a greater weight imposed on any loss.

Moreover, as shown in Table 2, no matter what degree of risk aversion and target returns exists, the hedging performance of the minimum-LPM hedge ratio at each scale will be superior to the transactions with no hedge, which means ρ is always greater than zero.

Table 2 also provides reward-to-semivariability (R/SV) ratios that measure the return-to-risk performance of hedged portfolios. From Table 2, it is easy to perceive that under the same degree of risk aversion and target returns, the R/SV ratios derived from the parametric methods based on features of series are larger than those derived from the nonparametric kernel density method. This means that the parametric method can generate better returns for the same risk than the nonparametric method. Therefore, as to the determination of multi-scale hedge ratios of stock index futures under the target of minimum-LPM, the parametric method on the distribution features of time series is superior to the universal nonparametric kernel density method.

Finally, from the empirical results ρ and R/SV of two different hedging strategies, as shown in the Parametric Based on the Feature of Series and Nonparametric Kernel Density Method columns in Table 2, the cross-comparison clearly shows that regardless of whether the time scale is long or short, whether the degree of risk aversion is high or low, and whether the target return increases or decreases, the hedge effectiveness of adopting different methods based on the distribution features of time series is always superior to the nonparametric kernel density method. In addition, since wavelet coefficients at high scales are too smooth sometimes, the optimal hedge ratio often has no unique value when they are determined by the nonparametric kernel density method. Thus, as theory indicates, though generally suitable, the nonparametric kernel density method is less precise compared to the parametric estimation with known distribution. Therefore, when estimating multi-scale hedge ratios based on LPM, one needs to take sequence distribution features into full consideration and adopt different methods. The nonparametric kernel density should be adopted only when it is impossible to figure out the specific distribution of wavelet coefficients at different scales.

5. Concluding remarks

The hedging model based on LPM can better reflect an investor's concern and show more flexibility. In this paper, we first decomposed the log return series of spot and futures returns into four components using the MODWT and found that parametric estimation methods, based on the distribution features of the original series and wavelet coefficients at different levels, are superior to the traditional nonparametric method in the hedge performance.

We also conducted empirical studies for the validity of the above models in the CSI 300 index futures and spot markets. Analysis shows that upon the multi-scale hedging model based on LPM, hedgers should choose appropriate multi-scale hedge ratios according to their hedging horizons. In addition, the shorter the hedging horizons are, the more attention hedgers should pay to the impact that the degree of risk aversion n and target return c have on optimal hedge ratio h^*. Specifically, in the short hedge horizon, both the parametric estimated h^* and the percentage reduction in LPM ρ display the rules of increasing with the rise of n and decreasing with the increment of c.

Furthermore, given the background of long-term decline in the China stock market, if target returns are adjusted to some negative level, good control of downside risk can be achieved. Finally, no matter whether the time scale is long or short, whether the degree risk aversion is high or low, and whether the target return is big or small, the hedge effectiveness of adopting the specific parametric method based on the features of time series is always better than those estimated by the nonparametric kernel density methods.

Acknowledgments

References

[1] F. In, S. Kim, The hedge ratio and the empirical relationship between the stock and futures markets: A new approach using wavelet analysis, J. Bus. 79 (2006) 799–820.
[2] D. Lien, K. Shrestha, An empirical analysis of the relationship between hedge ratio and hedging horizon using wavelet analysis, J. Futures Mark. 27 (2007) 127–150.
[3] D. Chen, W. Chen, C. Chuang, Multiscale hedge ratio between Taiwan stock and futures index: an application of wavelet analysis, Invest. Manag. Financ. Innov. 6 (2009) 217–230.
[4] T. Conlon, J. Cotter, An empirical analysis of dynamic multiscale hedging using wavelet decomposition, J. Futures Mark. 32 (2012) 272–299.
[5] Y. Sun, J.M. He, The multi-scale hedge ratios based on LPM, Syst. Eng. (Chinese) 32 (2012) 53–58.
[6] D.B. Percival, A.T. Walden, Wavelet Methods for Time Series Analysis, Cambridge University Press, England, 2006.
[7] J.C.T. Mao, Models of capital budgeting, E-V VS E-S, J. Finan. Quant. Anal. 4 (1970) 657–675.
[8] W.Y. Lee, R.K.S. Rao, Mean lower partial moment valuation and lognormally distributed returns, Manage. Sci. 34 (1988) 446–453.
[9] D. Lien, Y.K. Tse, Some recent developments in futures hedging, J. Econ. Surv. 16 (2002) 357–396.
[10] D. Lien, Y.K. Tse, Hedging time-varying downside risk, J. Futures Mark. 18 (1998) 705–722.
[11] W.V. Harlow, Asset allocation in a downside-risk framework, Financ. Anal. J. 47 (1991) 28–40.
[12] D. Lien, Y.K. Tse, A note on the length effect of futures hedging, Adv. Inves. Anal. Portf. Manag. 7 (2000) 131–143.
[13] Á León, G. Rubio, G. Serna, Autoregresive conditional volatility, skewness and kurtosis, Quart. Rev. Econ. Finance 45 (2005) 599–618.
[14] R. Demirer, D. Lien, Downside risk for short and long hedgers, Int. Rev. Econ. Finance 12 (2003) 25–44.
[15] B.W. Silverman, Density Estimation for Statistics and Data Analysis, Taylor & Francis, London, 1986.
[16] G. Homaifar, D.B. Graddy, Variance and lower partial moment betas as alternative risk measures in cost of capital estimation: a defense of the CAPM beta, J. Bus. Financ. Account. 17 (1990) 677–688.
[17] D. Nawrocki, A comparison of risk measures when used in a simple portfolio selection heuritic, J. Bus. Financ. Account. 10 (1983) 183–194.