Hedging Performance of Multiscale Hedge Ratios

Jahangir Sultan
Professor of Finance
Bentley University
Waltham, Massachusetts, USA

Antonios K. Alexandridis*
Senior Lecturer in Finance
Kent Business School, University of Kent
Canterbury, Kent, UK

Mohammad Hasan
Senior Lecturer in Finance
Kent Business School, University of Kent
Canterbury, Kent, UK

Xuxi Guo
Doctoral Student in Finance
Georgia State University
Atlanta, Ga, USA

ABSTRACT

In this study, the wavelet multiscale model is applied to selected assets to hedge time-dependent exposure of an agent with a preference for a certain hedging horizon. Based on the in-sample and out-of-sample portfolio variances, the wavelet-based GARCH model produces the lowest variances. From a utility standpoint, wavelet networks combined with GARCH have the highest utility. Finally, the wavelet GARCH model has the lowest minimum capital risk requirements (MCRR). Overall, the wavelet GARCH and wavelet networks offer improvements over traditional hedging models.

Keywords: wavelet analysis, multiscale hedge ratio, hedging effectiveness, GARCH model, wavelet networks.

JEL Classification: G1; G13; G15.

The authors would like to thank the editor and the anonymous referee for the constructive comments that helped to improve substantially the final version of this paper. We are responsible for all remaining errors.
*Corresponding author: Antonios Alexandridis, University of Kent, Kent Business School, Room 336, Sibson Building, Canterbury, Kent, CT2 7FS, UK. Tel: 0044 1227 827702. Email: A.Alexandridis@kent.ac.uk.

1 INTRODUCTION

We study the impact of hedging-horizon on the multi-scale and multi-period hedge ratio using wavelet decomposed returns from three representative classes of assets: commodities, currency, and stock index. Multi-scale and multi-period hedging decisions stem from the hedger's preference for a certain hedge horizon, hedging instruments, and risk tolerance. A wavelet is a small wave (signal) that grows over time but decays within a finite period. It has both the time and frequency domains that characterize its evolution. A wavelet-transform allows researchers to decompose time-series data into orthogonal components with different frequencies (scales) to accommodate structural changes, discontinuity, and regime shifts (Conlon and Cotter, 2012). The wavelet analysis accommodates multi-period decision-making models for heterogeneous economic agents weighing identical assets differently (see Kamara et al. (2016) for more on 'clientele effects')[1]. Overall, for effective risk management, it is important to measure risk at multiple scales of time.

Surprisingly, there is limited research to assess the hedging performance of the wavelet-based hedge ratios from scale-dependent data. For instance, previous studies have utilized nonparametric wavelets (In and Kim, 2006), ordinary least squares (Lien and Shrestha, 2007) and moving window ordinary least squares (Conlon and Cotter, 2012) methods to compute multi-scale hedge ratios and evaluate hedging effectiveness. None of the previous research accounted for the time-variation of the hedge ratios when return distributions are not normal. As Conlon and Cotter (2012) noted, by smoothing time series data, traditional approaches to determine static multiscale hedge ratios underestimate the information content of large dynamic changes. Furthermore, an analysis of the behavior of dynamic hedge ratios using alternative variants of econometric models, such as nonparametric wavelet, ordinary least squares, and GARCH models and a comparative assessment of hedging performances of the optimal hedge ratios across those models is lacking in the literature. Also, it is unclear whether hedgers derive higher utility from multi-period and multiscale hedging when portfolio returns have negative skewness and excess kurtosis. Finally, within a wavelet-based time-varying hedging framework, the use of VaR and minimum capital risk requirement (MCRR) as indicators of hedge effectiveness is limited.

In this study, we use wavelet-decomposed returns from Brent crude oil, FTSE100 Index, Gold, and the U.S. dollar (USD) index for the period January 3, 2005 to December 14, 2018 to evaluate five hedging models. We compare the performance of these five models to evaluate their incremental contributions to portfolio variance reduction, utility maximization, and reduction in the regulatory capital requirement. The in-sample hedging models are wavelet-unhedged (WU), wavelet-full hedge (WFH), wavelet-OLS (WOLS), wavelet-GARCH (WG), and wavelet-hedge (WH). The same models are used for out-of-sample evaluation with the exception that the WH strategy is replaced with the wavelet neural networks hedging model (WN) that combines wavelet transformations and artificial neural networks. The WN model is justified (to be discussed later) since the wavelet networks can be used for forecasting time-varying out-of-sample hedge ratios which the standard wavelets by themselves cannot do.

Based on the in-sample portfolio variance of the assets considered, the WG model performs best, followed by the WOLS strategy. For out-of-sample hedging, WG again is the best performing model, followed by WN. From a utility standpoint using in-sample wavelet-decomposed returns, WH is the best strategy overall, followed by WG and WOLS, respectively. In terms of out-of-sample performance based on wavelet-decomposed returns, WG is the overall winner, followed by WOLS. Finally, based on the MCRR, the WG model outperforms alternative models. The next best hedging model is WOLS. Overall, WG offers improvements over traditional hedging models.

A key result in this study is that for all assets across all horizons and hedging strategies, the portfolio variance based on the original returns exceeds the wavelet-based portfolio variance. Furthermore, the standard GARCH model performs worse than the wavelet-GARCH model in terms of hedged portfolio variance. Overall, wavelet-based multiscale hedging performs far better than conventional and dynamic hedging.

The study makes several unique contributions to the literature on hedging. First, it applies the GARCH method to combine time-varying hedging, multiscale hedging horizon, and heterogeneous investors in a synthetic wavelet-GARCH framework. Unlike conventional approaches to estimating static multiscale hedge ratios, the synthetic wavelet-GARCH framework captures dynamic information content to produce time-varying multiscale hedge ratios when asset returns are not normal. Second, this study applies a new class of artificial neural networks, namely the Wavelet Networks (WNs) to examine out-of-sample hedging effectiveness of multiscale hedge ratios. A combination of wavelet analysis and neural networks improves significantly the forecast accuracy of out-of-sample hedge ratios even for longer horizons. Third, in addition to variance reduction, hedging effectiveness is judged based on mean-variance (MEV) and exponential utility functions, certainty equivalent wealth (CE), and risk-adjusted information ratio (AIR). The utility analysis tests whether or not there is an increase in utility from multi-period and multiscale hedging especially when portfolio returns have negative skewness and excess kurtosis. Finally, the MCRR is calculated to confirm the practical usefulness of wavelet-based hedging models for keeping risk capital requirement low. In other words, since the hedge ratios of various portfolios are predictable, to achieve maximum risk reduction, a hedger would prefer a portfolio with the lowest MCRR.

The study proceeds as follows. Section 2 provides a brief survey of the hedging models considered in this paper. In Section 3, empirical results are reported. The final section offers a summary of the key findings.

2 METHODOLOGY

Several recent studies have suggested that a wavelet based multi-horizon hedging is a preferred strategy over conventional methods (see Lien and Shrestha (2007) and Conlon et al. (2017)). The rationale is that a hedger makes decisions in a multi-period setting in the real world, taking into account hedging preference, hedging horizon length, and hedge effectiveness. In short, the hedger's exposure to the financial market depends upon magnitude, variability, and location of the shock. Consequently, there is a unique hedge ratio for each hedging horizon (Geppert, 1995). Most often, a single period hedging model is preferred due to its computational simplicity though it may not adequately minimize risk when the hedger faces time-dependent multi-horizon exposure (Lien and Luo (1993).

Multi-horizon hedging also accommodates selective hedging. According to Conlon *et al.* (2016), selective hedging is a form of speculation when hedgers and speculators prefer a policy of no-hedge, partial-hedge, and horizon-specific hedge, as opposed to complete hedging.[2] For example, the three biggest air carriers in China - the Air China, the China Eastern Airlines, and the China Southern Airlines - did not hedge fuel purchases for some time after the 2008-2009 financial crisis.[3] In contrast, it is common among Asian airlines to reduce hedging cost by undertaking a partial-hedge or at least hedging for a short horizon.[4] Consequently, wavelet analysis provides an appropriate framework to approximate the multiscale nature of the hedge ratio in these circumstances. Furthermore, as explained later, the wavelet approach overcomes data reduction problem for low-frequency data and captures information associated with all available data. In other words, the hedger picks a model that is robust to different time scales or horizons, without having to run out of data.[5]

Surprisingly, the application of wavelet-based multiscale hedging has been rather limited (see, for example, In and Kim (2006), Fernandez (2008), Lien and Shrestha (2007), and Conlon et al. (2012)). For example, Geppert (1995) employed a permanent/transitory decomposition model to investigate the behavior of the multi-period hedge ratio for five assets. Chen et al. (2004) studied the effects of the length of the hedging-horizon[6] on the optimal hedge ratio and hedge effectiveness. Their results are similar to those reported in Geppert (1995), the only exception is that the authors found that hedging effectiveness increases with the length of hedging-horizon. Conlon and Cotter (2012) applied a moving-window OLS method to wavelet-decomposed data for crude oil, currency, and stock indices to estimate minimum variance hedge ratios for horizon-hedging. They evaluated hedging performance using two criteria, i.e., variance reduction and scale-dependent value-at-risk (VaR). Both in-sample and out-of-sample results indicate that the hedge ratio is increasing in scale, and long-horizon hedging has lower transaction costs and higher utility. Furthermore, portfolio VaR shows that unhedged tail risk exists in all scales due to excess kurtosis.

Wavelet-based hedging models are also associated with a higher utility for a hedger. The papers by In and Kim (2006) and Conlon et al. (2016) examined the utility-based hedge effectiveness of wavelet models. Their simulation results indicate that scale-dependent hedging effectiveness hinges on the risk aversion of the hedger. Specifically, a hedger with extremely high-risk aversion derives higher benefits from long-term hedging; a hedger with extremely low-risk aversion attains most hedging benefits at a short-term scale, and a hedger with moderate risk aversion achieves maximum utility at the intermediate scale. Some of the other notable findings in the literature include: a unique hedge ratio corresponds to each hedging horizon length, the long-run hedge ratio converges to one, in-sample hedging effectiveness[7] converges to one as investment horizon increases, and out-of-sample hedging effectiveness tends to decrease as hedging horizon increases.

2.1 Conventional and Time-Varying Hedging Models

In this section, we review the conventional hedging models and make necessary modifications to introduce the wavelet analysis. The hedger in our model holds one unit of the spot asset and wishes to hedge by shorting x units of futures contracts. Various theoretical approaches, such as minimum variance, mean-variance, expected utility, mean-expected Gini coefficient as well as semivariance, have been considered for identification and estimation of the optimal hedge ratio. Chen et al. (2003) contend that these approaches generate similar hedge ratios under the Martingale assumption and joint-normality.

We begin by considering Johnson's (1960) risk-minimizing hedge ratio h^*, defined as:

$$h^* = -\frac{\sigma_{s,f}}{\sigma_f^2} = -\frac{\text{cov}(\Delta S_t, \Delta F_t)}{\text{var}(\Delta F_t)} \quad (1)$$

where S and F denote log of spot and futures prices, respectively, and Δ is the first difference operator. The OLS hedge ratio is computed as the slope coefficient of the following regression:

$$\Delta S_t = \alpha + \beta \Delta F_t + \varepsilon_t \quad (2)$$

where ε_t is an iid error term[8]. In equation (2), a β=1 yields a fully hedged (FH) position. When asset returns are not normally distributed, the variance/covariance terms are changing over time, and the optimal hedge ratio changes over time. So, equation (1) changes to:

$$h_t^* = -\frac{\text{cov}_t(\Delta S_t, \Delta F_t \mid \Omega_t)}{\text{var}_t(\Delta F_t \mid \Omega_t)}. \quad (3)$$

where conditional moments change when the information set Ω_t, is updated. Therefore, the optimal hedge ratio h_t^* changes through time.

Equation (2) may not be specified correctly if the spot and futures prices are cointegrated. Theory of cointegration suggests that when the basis becomes large, arbitrageurs exploit this temporary disequilibrium (Brenner and Kroner, 1995) to restore a long-run equilibrium. In other words, a stationary basis reinforces the cost-of-carry assumption though the assumption may not hold in some cases (Chen et al., 2004). The following bivariate GARCH (p, q) model, which has become standard in the literature, incorporates both non-normal asset returns and an error correction (EC) term to produce time-varying hedge ratios:

$$\begin{aligned} \Delta S_t &= \alpha_0 + \alpha_1 u_{t-1} + \varepsilon_s \\ \Delta F_t &= \beta_0 + \beta_1 u_{t-1} + \varepsilon_f \end{aligned} \quad (4)$$

$$\begin{bmatrix} \varepsilon_s \\ \varepsilon_t \end{bmatrix} \mid \Omega_{t-1} \sim N(0, H_t) \quad (5)$$

$$vech(H_t) = C + \sum_{i=1}^{p} A_i vech(\varepsilon_{t-i})^2 + \sum_{j=1}^{q} B_j vech(H_{t-j}) \quad (6)$$

where u_{t-1} in the conditional mean equations are the lagged error-correction terms (lagged basis)[9]. The terms ε_s and ε_f are the residuals from the mean equations (4). Equation (5) describes their joint density function which is time-varying, given the information set Ω_{t-1}. In equation (6), H_t is a (2x2) conditional covariance matrix, C is (3x1) parameter vector of constants (unconditional variance and covariance), A_i and B_j are (3x3) ARCH and GARCH parameter matrices, respectively, and *vech* is the column stacking operator that stacks the lower triangular portion of a symmetric matrix. In equation (6), both lagged squared residuals, as well as past volatility, are assumed to be key determinants of the current volatility. The GARCH hedge ratios are defined as:

$$h_t^* = \hat{H}_{sf,t} / \hat{H}_{ff,t} \quad (7)$$

where $\hat{H}_{sf,t}$ is the estimated conditional covariance between the spot and futures returns, and $\hat{H}_{ff,t}$ is the estimated conditional variance of futures returns. The GARCH model is estimated by maximizing a likelihood function using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm in RATS with unconditional variance and covariance used as starting values.

2.2 Wavelet-Hedge (WH)

The wavelet decomposition of a time series combines both spatial analysis and Fourier transformation to detect the properties of quick variation of values (Alexandridis and Zapranis, (2013 and 2014). There are several distinct benefits from using the wavelet decomposition. A spatial analysis of a time series reveals the value of a function at a particular location but does not offer any information on the magnitude of the variability. A Fourier transformation documents the magnitude of the variability but does not say anything about where the variability is located (see Lien and Shrestha, 2007, and Lindsay et al., 1996). The wavelet-transform combines both the magnitude and location of the variability. Using wavelets, a time series can be decomposed into various scales to capture maximum information from the data. The low (high) scale represents the high (low) frequency. The wavelet-transform is localized in both time and frequency and it also overcomes the fixed time-frequency partitioning. This means that the wavelet transform has good frequency resolution for low-frequency events and good time resolution for high-frequency events. Further, the wavelet analysis captures the structures of the original time-series such as trends, jumps or periodicities. Such structures are common in daily returns.

Another principal benefit of the wavelet transform is that it does not suffer from the sample reduction problem which has been identified in the literature. For example, to estimate multi-period hedge ratios using the conventional method, a hedger matches the frequency of the data or the differencing interval to the hedging horizon. One needs to use weekly, monthly and annual data to obtain hedge ratios consistent with weekly, monthly and annual investment horizons, respectively (see, for example, Chen et al., 2004), Geppert, 1995), and Lien and Shrestha, 2007). In short, for a *k*-period hedging horizon, one needs to use a *k*-period differenced data, which results in a reduction of the sample, especially when the long-horizon period (e.g., annual) is considered (because it would require sampling data at annual intervals for differencing)[10]. Overall, the wavelet analysis is appropriate for estimating scale-dependent dynamic hedge ratios (see Lien and Shrestha (2007)).

In this study, the Maximal Overlap Discrete Wavelet Transformation (MODWT) is applied. It has many desirable properties compared to the classic DWT (see Percival and Walden (2000)). First, a father wavelet function, $\varphi_{j,k}(t)$, representing the smooth component and a mother wavelet function, $\psi_{j,k}(t)$, representing the deviations are selected. Next, wavelet coefficients are estimated through the convulsion of the mother wavelet function with the time-series $f(t)$:

$$d_{j,k} = \int_{-\infty}^{+\infty} f(t)\psi_{j,k}(t)\,dt \tag{8}$$

$$s_{j,k} = \int_{-\infty}^{+\infty} f(t)\varphi_{j,k}(t)\,dt \tag{9}$$

where $j = 1,...,J$ is the number of scales and k indicates the k^{th} coefficient. In this study the LA5 (Least Asymmetric of length 5) wavelet transform filter is used. The analysis is performed at 5 levels of the decomposition and the reflection method was used for the boundary conditions. This follows the previous findings in the literature that as much as 90% of the return variance in many commodities comes from shorter time scales (Lien and Shrestha (2007)). The original time-series can be reconstructed by

$$f(t) = \sum_{k\in Z} s_{J,k}\varphi_{J,k}(t) + \sum_{j\in Z}\sum_{k\in Z} d_{j,k}\psi_{j,k}(t). \tag{10}$$

By setting $S_{J,t} = \sum_{k} s_{J,k}(t)\varphi_{J,k}(t)$ and $D_{j,t} = \sum_{k} d_{j,k}(t)\psi_{j,k}(t)$, equation (10) can be written as:

$$f(t) = S_{J,t} + D_{J,t} + D_{J-1,t} + \cdots + D_{1,t} \tag{11}$$

which is known as the multi-resolution analysis (MRA). The original time-series is denoted as the approximation S_0. At each level j of the MODWT, the approximation S_{j-1} is split into two parts, the new approximation S_j and a detail signal D_j that captures short-term deviations in the time-series. The variance of the original time-series (X) can be decomposed by scale, identifying the contribution of each scale to the variance (Percival and Walden (2000). An unbiased estimator of the wavelet variance is given by:

$$\hat{\upsilon}^2_{X,j} = \frac{1}{N_j}\sum_{t=L_j-1}^{n-1}\left(d^{(X)}_{j,t}\right)^2 \tag{12}$$

where $d_{j,t}$ is the MODWT wavelet coefficients at scale j, n is the sample size, L_j is the length of the scale j wavelet filter where $L_j = (2^j - 1)(L-1)+1$, and N_j is the number of the MODWT coefficients unaffected by the boundary where $N_j = n - L_j + 1$. Similarly, the wavelet covariance can be computed by decomposing the sample covariance at different time scales. Given two time-series X and Y, the unbiased estimator of the wavelet covariance can be computed by:

$$\hat{\upsilon}^2_{XY,j} = \frac{1}{N_j}\sum_{t=L_j-1}^{N-1} d^{(X)}_{j,t} d^{(Y)}_{j,t} \ . \tag{13}$$

Hence, the wavelet hedge ratios for asset i at scale j can be computed as follows:

$$h_{i,j} = \frac{\hat{\upsilon}^2_{\Delta S \Delta F,j}}{\hat{\upsilon}^2_{\Delta F,j}} \tag{14}$$

where $\hat{\upsilon}^2_{\Delta S \Delta F,j}$ is the wavelet covariance of the wavelet-decomposed spot and futures returns at scale j, and $\hat{\upsilon}^2_{\Delta F,j}$ is the wavelet variance of the wavelet-decomposed futures returns at scale j. This model is referred to as the wavelet-hedge (WH). Both the spot and futures returns can be decomposed into different time scales (see Lien and Shrestha (2007)):

$$\Delta S_t = S^S_{J,t} + D^S_{J,t} + D^S_{J-1,t} + \cdots + D^S_{1,t} \tag{15}$$

$$\Delta F_t = S^F_{J,t} + D^F_{J,t} + D^F_{J-1,t} + \cdots + D^F_{1,t} \tag{16}$$

In this study, a scale-dependent version of equation (2) is estimated, i.e., estimates of J regressions using the j^{th} scale decomposition. The model is referred to as the wavelet-OLS (WOLS):

$$D^s_{j.t} = \alpha_{j,0} + \beta_{j,1} D^f_{j,t} + \varepsilon_{j,t} \tag{17}$$

The OLS hedge ratio associated with the j^{th} scale is denoted by $\beta_{j,1}$. When $\beta_{j,1}$=1, the wavelet fully hedged (WFH) position is obtained. The GARCH model described earlier (equations 4-6) can also be applied to the equation (17) at each level j using wavelet-decomposed returns:

$$h^*_{t,j} = \frac{\hat{H}_{sf,t,j}}{\hat{H}_{ff,t,j}} \tag{18}$$

where $h^*_{t,j}$ is the optimal time-varying GARCH hedge ratio estimated at scale *j*. At scale *j*, $\hat{H}_{sf,t,j}$ is the estimated conditional covariance between the in-sample wavelet-decomposed spot and the futures returns and $\hat{H}_{ff,t,j}$ is the conditional variance of in-sample wavelet-decomposed futures returns. This is referred to as the wavelet GARCH (WG)[11] model.

2.3 Wavelet Networks (WNs)

Using wavelet analysis, a static WH can be estimated which in turn can be used as a naïve estimate of the hedge ratio for the next period[12]. However, it is important to recognize that wavelets by themselves cannot forecast time-varying hedge ratios. A solution is to fit a GARCH model to wavelet-decomposed returns and then use the fitted values to make out-of-sample forecasts. Alternatively, one can forecast the decomposed signals $D_1,...,D_5$ and then compute the hedge ratios. Based on the preceding, the WNs are chosen that combine the classic neural networks architecture with the wavelet analysis. The WNs are different from classical neural networks because the activation function in each neuron (or hidden unit) is a wavelet function instead of the classic sigmoid one. Figure 1 presents the WN architecture. It also shows the transformation in each layer and the node. A WN consists of three layers: the input, hidden units, and output. The input variables, $\mathbf{x} = \{x_1,..,x_m\}$, are inserted into the model through the input layer. The hidden layer consists of the hidden units,

$$\Psi_j(\mathbf{x}) = \prod_{i=1}^{m} \psi\left(\frac{x_i - w^{[1]}_{(\xi)ij}}{w^{[1]}_{(\zeta)ij}}\right). \tag{19}$$

In the hidden layer, the inputs are transformed into dilated and translated versions of the mother wavelet. Finally, in the output layer, the output of each neuron is linearly combined to produce the network's output. The approximation of the target values, $\hat{y}(\mathbf{x})$, is estimated as:

$$g_\lambda(\mathbf{x};\mathbf{w}) = \hat{y}(\mathbf{x}) = w^{[2]}_{\lambda+1} + \sum_{j=1}^{\lambda} w^{[2]}_j \cdot \Psi_j(\mathbf{x}) + \sum_{i=1}^{m} w^{[0]}_i \cdot x_i \tag{20}$$

where $\boldsymbol{x}$ is the input vector, m is the number of network inputs λ is the number of hidden units and w stands for a network weight. Finally, $\Psi_j(\mathbf{x})$ is a multidimensional wavelet, which is the product of m scalar wavelets.

[Insert Figure 1 About Here]

In this study the second derivative of the Gaussian, the so-called "Mexican Hat" wavelet is used. The complete vector of the network parameters comprises: $w = \left(w_i^{[0]}, w_j^{[2]}, w_{\lambda+1}^{[2]}, w_{(\xi)ij}^{[1]}, w_{(\zeta)ij}^{[1]} \right)$. These parameters are adjusted during the training phase, which is described next.

To forecast the out-of-sample hedge ratios, the WNs and $g_\lambda(\mathbf{x};\mathbf{w})$ need to be trained first. For each asset, a different WN was trained for each detail D_j and smoothed component S_j using the model identification algorithm presented in Alexandridis and Zapranis (2014). The algorithm also provides an efficient determination of the lag series and the network topology. The backward elimination algorithm was used to initialize the WN that has been proven to be more efficient than alternative initialization procedures (see Alexandridis and Zapranis (2013)). A rolling window was applied to produce one-step ahead out-of-sample forecasts. More precisely, in order to train the WNs on a particular detail D_j, the vector $\mathbf{x} = \left\{ D_{j,t-m}, D_{j,t-m+1}, \ldots, D_{j,t-1} \right\}$ is used as input values and the target values are given by $\mathbf{y} = D_{j,t}$. A similar approached was followed for the smoothed components. Finally, the hedge ratio at each time-step is estimated by applying equation (14) to the forecasted time-series.

3 EMPIRICAL ANALYSIS

Daily spot and futures prices for Brent crude oil, FTSE100, Gold and the USD are collected from the Bloomberg terminal. Brent crude oil spot price per barrel is the 'Europe price'. Brent futures contracts trade on the Intercontinental Exchange (ICE). FTSE100 is the Financial Times Stock Index from the London Stock Exchange. The FTSE futures contracts trade on the ICE. Gold price (COMEX) is the spot price per troy ounce. Gold futures contracts trade on the Chicago Mercantile Exchange (CME). The spot U.S. dollar Index (USD) is based on geometric averages of the six component currencies: euro, Japanese yen, British pound, Canadian dollar, Swedish krona, and Swiss franc. The USD futures contracts trade on the ICE. All futures prices in this study are based on active contracts with rollover 15 days before expiration. The in-sample[13] period is January 3, 2005 to March 2, 2016, while the out-of-sample period is March 3, 2016 to December 14, 2018. As noted earlier, the study uses two types of data: spot and futures return (log-returns), referred to as 'original returns' and wavelet-decomposed returns (Scales 1-5). Days (in parenthesis) represented by these horizons are scale 1 (1-2), scale 2 (2-4), scale 3 (4-8), scale 4 (8-16), and scale 5 (16-32). The in-sample empirical work proceeds as follows. First, the conventional hedging models (unhedged (U), full hedge (FH), OLS, and GARCH) are estimated using original returns. Next, wavelet-based hedging models (WU, WFH, WOLS, WG, and WH) are applied to the wavelet decomposed returns (scales 1-5). The process is also repeated for the out-of-sample with the exception that WH is replaced with WN.

As an example, Figure 2 shows the wavelet decomposition at 5 levels for both spot (top panel) and futures (bottom panel) returns of the FTSE 100 Index. The wavelet decomposition splits the original time series data into high and low-frequency parts. A closer inspection of the top panel reveals that wavelet analysis brings out periods of high and low variability (e.g. during the financial crisis an increase in the volatility is evident while the volatility is significantly lower during the period 2012-2015) in the decomposed wavelet spot returns. Finally, the wavelet-decomposed FTSE100 futures returns exhibit similar patterns (bottom panel). The wavelet decomposition for the Brent, USD, and gold exhibit similar patterns (not reported to save space).

[Insert Figure 2a About Here]
[Insert Figure 2b About Here]

In Panels A-D, Table 1, summary statistics for each asset are presented. The mean return is zero for all horizons and all assets. Also, there is negative skewness, lower excess kurtosis, and lower standard deviation at a higher scale. For Brent, the skewness is negative for lower scales and it becomes positive at higher scales. Kurtosis (spot and futures) is positive for all scales. For FTSE100, both excess kurtosis and skewness decline at higher scales while at lower scales, the skewness is positive. Skewness is mostly negative for returns on Gold but there is no systematic pattern as one moves from low to high scales. Excess kurtosis, on the other hand, tapers off with higher scales. Original returns and wavelet-decomposed returns for the U.S. dollar have mostly negative skewness for all scales. Excess kurtosis declines as higher scales are considered. The Jarque-Berra statistic confirms non-normality for both original returns and wavelet-decomposed returns of all assets.

Finally, Table 1 also presents Engle's TR^2 statistic which is the Lagrange Multiplier test to confirm the presence of autoregressive conditional heteroscedasticity (ARCH) errors. Both the original returns and the wavelet-decomposed returns have ARCH properties. The estimated chi-square is significantly higher than the critical value of 3.841 at 5% with 1 degree of freedom. The use of the bivariate GARCH model (a more generalized version of ARCH) is appropriate given that these returns have ARCH properties.

In Table 2, hedge ratios across all horizons from both WOLS and wavelet hedge (WH) are shown. Consistent with the previous findings in the literature, the hedge ratios tend to increase with higher horizons.

3.1 In-Sample Portfolio Variance

In this section, we examine the performance of the hedging models by evaluating the variance of in-sample (January 3, 2005, to March 2, 2016) hedged portfolios ($\Delta S_t - h_t^* \Delta F_t$). For calculating the wavelet-based hedged portfolio variance, ΔS_t and ΔF_t are replaced with wavelet-decomposed spot and futures returns, respectively. Since the main focus of this study is on the wavelet hedging models, unless otherwise noted, discussion of the results will concentrate on wavelet-based portfolio variances estimated from wavelet-decomposed returns.

There are two criteria for judging the hedging effectiveness. First, using the unhedged strategy (WU) as the benchmark, one can examine the portfolio variance reduction from an alternative strategy. For example, to compare between the WG and unhedged (WU) portfolios, calculate the percentage reduction in the variance as $(V_{WU} - V_{WG})/V_{WU}$. Second, for each hedging horizon, rank the portfolio variances of all models using a 1-5 measurement chart. The strategy that has the least portfolio variance is ranked 1. Since both of these criteria produces similar conclusions, the results using the second method are presented in this study. To facilitate comparison, the aggregate rank is found by simply adding the individual rank for each hedging horizon across all assets.

In-sample portfolio variances of these models are presented in Panel A, Table 3. For Brent, the WG model has the lowest variance (ranked 1) compared to the remaining hedging strategies across all horizons. The strategy is ranked 1 across all 5 horizons, producing an aggregate rank of 5. The remaining strategies have the following aggregate ranks (in parenthesis): WOLS (12), WH (13), WFH (20), and WU (25). The WG model also performs the best in the case FTSE100 and the rank of the hedging models is similar as in the case of Brent. Interestingly, WH and WOLS have similar performance as both are ranked 13. The best performing hedging model for Gold is WG (8), followed by WH (11), WOLS (12), WFH (19), and WU (25), respectively. Finally, the WG (5) model turns out to be the best performer in the case of the USD.

Overall, the results confirm the relative superiority of the wavelet-based hedging models. Aggregate ranking of hedging models for all assets and all hedging horizons shows WG having a rank[14] of 23 while the wavelet-unhedged model has an overall rank of 100, which makes it the worst hedging strategy. The next best performer (after WG) is WH with a rank of 47. These results suggest that the wavelet-GARCH hedging model has superior potentials for minimizing risk across all hedging horizons.

3.2 Out-of-Sample Portfolio Variance

Out-of-sample (March 3, 2016 to December 14, 2018) performance of the hedging models is presented next. As discussed earlier, this part of the exercise requires out-of-sample forecasting and training of the WNs to obtain time-varying hedge ratios. Out-of-sample portfolio returns ($\Delta S_{t+1} - h^*_{t+1}\Delta F_{t+1}$) are based on forecasted hedge ratios using the WN model[15]. Using a recursive updating procedure, the in-sample values of D_1 to D_5 and S (the final approximation) are utilized to train the WNs. Few lags were used to forecast the next values of D_1 to D_5 and S (similar to a nonlinear autoregressive model). Subsequently, the forecasted return at time 1, denoted as $\sum D_{1..5} + S$ is generated. Next, another sample observation is added and the next out-of-sample observation is forecasted. The wavelet analysis is applied to this new sample and the new wavelet coefficients W and V are derived, from which the wavelet variance/covariance and subsequently the hedge ratio at each scale are obtained. Using a recursive window, this process is repeated for all out-of-sample observations across all the five hedging models described earlier.

In Panel B, Table 3, out-of-sample portfolio variances are reported for all hedging strategies. Similar to Panel A, the models are estimated using both original returns and wavelet-decomposed returns. Overall, across all assets, the WG model turns out to be the best performer, except for Gold, where both WG and WN are equally effective in reducing the portfolio variance. The second best hedging model overall is WN, followed by WOLS, WFH, and WU. A no-hedge policy (WU) has a rank of 4, which makes it the worst strategy[16]. We also notice, consistent with the evidence reported in the literature, that the hedge ratio and hedging effectiveness tend to increase with the length of hedging-horizon. These results support the notion that wavelet multiscale hedging offers significantly better risk minimization capability. Finally, in Panel C of Table 3, in-sample and out-of-sample average ranking of each hedging strategy across all scales and assets is shown. WG ranks first both in-sample and out-of-sample, followed by WN. WOLS ranks third, followed by the full hedge (WFH).

An important finding in this study is that for all assets across all horizons and hedging strategies and for both in-sample and out-of-sample exercises, portfolio variance estimated using the original returns exceeds the wavelet-based portfolio variance. In particular, the wavelet-GARCH model using wavelet decomposed returns produces a lower variance than the standard GARCH model applied to the original returns for all assets. This further reconfirms the notion that wavelet-based multiscale hedging is far superior to conventional and dynamic hedging.

In Table 4, pairwise F-tests are conducted at each scale and for all assets to confirm whether the portfolio variances (reported in Table 3) are statistically different across the hedging models. In-sample results are reported in Panel A while out-of-sample results are reported in Panel B. A closer inspection of Table 4 reveals that the results from WG are statistically significant compared to the remaining methods (F-tests rejected equality of the variances). There is no difference between the WH and WOLS (F-tests accepted the hypothesis of equality of variances). The out-of-sample results are similar. In the majority of the cases, F-tests reject the hypothesis that WG portfolio variances are equal to portfolio variances from WOLS and WN strategies. By combining the results of Table 3 and Table 4, we can state that WG generates portfolios with lower variances and the variance reduction is statistically different from the other hedging models. Based on the F-tests, WN ranks second overall, followed by WOLS.

Finally, the effectiveness of the WN model for forecasting the decomposed series is examined by comparing the results against two benchmarks (results are not reported to save space). First, the estimated wavelet hedge ratios are applied to the next hedging period to obtain a one-step forecast (referred to as Wavelet-Bench)[17]. This method is often used in practice. The second benchmark is a simple random walk (referred to as Wavelet-RW) model. Here, a simple Wavelet-RW is used to forecast D_j and then estimate the hedge ratios. At each step t, the $D_{j,t+1}$ is forecasted as the average value of D_j up to time t. As presented in Welch and Goyal (2007), the simple RW is very difficult to beat and it is the usual benchmark used in the literature (see Rapach et al., 2010)). The forecasting power of the WN, Wavelet-RW and the Wavelet-Bench is compared by performing the Clark and West (2007) test. The WN model delivers statistical significantly better results for all assets for both the spot and futures returns. It is ranked first, followed by the Wavelet-Bench and Wavelet-RW. The Wavelet-RW performs worse than the naïve method while the Wavelet-Bench performs similar to the WOLS model.[18]

3.3 Utility of hedging models

From a practical standpoint, small size reductions in portfolio risk do not imply that the economic viability of the proposed strategy is insignificant (Kroner and Sultan (1993)). In other words, a multiscale WG model should be selected if it also increases the investor's utility net of transaction costs. To demonstrate this, the economic significance of the time-varying hedge ratio is analyzed using three criteria: the mean-variance utility function (MEV), the certainty equivalent exponential utility function (CE), and the adjusted information ratio (AIR) (Alexander and Barbosa, 2008). The mean-variance utility function is defined as

$$EU\left(\Delta S_t - h_t^* \Delta F_t\right) = E(\Delta S_t - h_t^* \Delta F_t) - Q - \lambda \sigma^2 \left(\Delta S_t - h_t^* \Delta F_t\right) \quad (21)$$

where Q is the transaction cost. The expected return to the hedged portfolio is assumed to be zero (Kroner and Sultan (1993)), and the coefficient of risk tolerance (λ) is 4. The average utility from hedging on a given trading day is: $-Q - 4\sigma^2(\Delta S_t - h_t^* \Delta F_t)$. In this study, a round-trip cost of 0.005% is assumed.[19]

Traditionally, the mean-variance rule and conventional hedging models are both designed to select portfolios which are expected to generate the lowest risk for a given expected return. Therefore, the choice of the optimal hedge ratio focuses on the first two moments of the return distribution and hedging effectiveness is measured by the proportional reduction in the variance of portfolio return. An alternative measure of hedging performance in recent research underscores the role of skewness and kurtosis of portfolio returns (Alexander and Barbosa (2008)). As the authors noted, performance evaluation based on the proportional variance reduction does not incorporate the effect of variance reduction on skewness and kurtosis. In other words, a conventionally hedged portfolio may have a very low return volatility but a high kurtosis indicates that the hedge can backfire some days. With a negative skewness, the hedged position would be losing rather than making money. Therefore, the second measure of hedging effectiveness which accounts for both skewness and kurtosis is derived from the following exponential utility function:

$$U(p) = -\lambda \exp\left(-p/\lambda\right) \quad (22)$$

where p signifies wealth. The exponential function has the property, $U\left(p\right) = E\left[U\left(p\right)\right]$. Using Taylor expansion of U(p) around the mean and taking the expectation up to the fourth term, the certainty equivalent utility function (CE) may be approximated as:

$$CE = \mu - \frac{\sigma^2}{2\lambda} + \frac{\phi}{6\lambda^2} - \frac{\kappa}{24\lambda^3} \quad (23)$$

where the third and fourth moments $\phi = E\left[\left(p-\mu\right)^3\right]$ and $\kappa = E\left[\left(p-\mu\right)^4\right]$ signify skewness and kurtosis, respectively. When $\lambda > 0$, risk aversion increases with increasing variance, negative skewness, and higher kurtosis.

Finally, hedging performance is also evaluated on the basis of AIR (Alexander and Barbosa, 2008):

$$AIR = IR + \frac{\hat{\phi}}{6} IR^2 + \frac{\hat{\kappa}}{24} IR^3 \qquad (24)$$

where *IR* refers to the information ratio defined as the ratio of mean return to the volatility of return, $\hat{\phi}$ denotes the estimated sample skewness and $\hat{\kappa}$ signifies excess kurtosis.

First, using in-sample original returns for all four assets, utility functions and the information ratio are estimated (to save space, these results are not reported). Hedging performance of all four hedging strategies (U, FH, OLS, and GARCH) is analyzed using their relative ranks for each asset on three utility-based criteria (MEV, CE, and AIR). The ranking is done on a 1-3 measurement chart (1=highest MEV, for example). For Brent, both FH and OLS models produce the highest values in terms of MEV, CE, and AIR. In the case of FTSE100, the OLS is the best model. Both unhedged and GARCH models are the next best. There is evidence that the OLS is the best model for hedging Gold as it is ranked best in terms of all 3 criteria. Finally, for hedging USD, the OLS model is the best performer. The next best performer is the GARCH model. Across all four assets, the OLS is the best strategy.

In contrast, for the wavelet-decomposed returns, there are four evaluation criteria: MEV, ΔMEV (Δ refers to the first difference), CE, and AIR. Across all horizons and assets, WH is the best hedging model, followed closely by WG and WOLS, respectively. Both WFH and WU models have the worst performance. Individually, WG has the best performance for hedging Brent crude oil, followed by WH. Both of these models surpass the competing models. For hedging the FTSE100, the winning model is WH, followed by WG. The WH model is again the winner for hedging Gold, followed closely by WOLS. The performance of the remaining models (WU, WFH, and WG) is similar. Finally, for hedging USD, the WH is the best model, followed by WG, WOLS, WFH, and WU, respectively.

A similar analysis is performed using the out-of-sample original and wavelet-decomposed returns. Recall that for out-of-sample utility evaluations, WN replaces WH. First, for the original returns, the GARCH model is the best hedging model across all four assets. Next, the best wavelet-hedging model is WOLS. The remaining models are ranked in the following order: WG, WFH, WU, and WN. Individually, for Brent, WG and WFH are the best models. For the FTSE100, WG has the best ranking. For the remaining assets (Gold and USD), WOLS and WFH have better performance than competing models. WN is the worst hedging model. Overall, utility-based analysis reveals that hedging effectiveness increases as one considers higher moments in higher scales[20].

3.4 Hedging Effectiveness: Minimum Capital Risk Requirement (MCRR)

In this section, hedging effectiveness is evaluated by comparing the MCRR for portfolios from the hedging models considered in this study. Given that the hedge ratios of portfolios obtained from the hedging models are predictable, the hedger prefers a portfolio with the lowest MCRR. The estimation of the MCRR takes into account the VaR of the portfolio, which is a statistical measure of the expected maximum loss on the portfolio, given some level of confidence. VaR can be derived from the probability distribution of the future portfolio as the worst possible realization R^* (R denotes the value of the portfolio) such that the probability of a value lower than R^* is:

$$P(R \leq R^*) = \int_{-\infty}^{R^*} f(R)\, dR = 1 - c \tag{25}$$

and (1-c) represents the probability of a lower-tail event. The MCRR for a 1-day investment horizon is calculated by simulating densities of portfolio returns using Efron's (1982) bootstrapping methodology (which is based on a multivariate GARCH (1,1) model). The Monte Carlo simulation procedure used 10,000 simulated paths of portfolio returns based on a GARCH (1,1) model to generate an empirical distribution of the maximum loss.

Table 5 presents the estimated in-sample MCRR for each hedging model. Note that the MCRRs for portfolios from conventional hedging models (unhedged, full hedge, OLS, and GARCH) are based upon original returns. The MCRRs are also reported for the following wavelet hedging models: WU, WFH, WOLS, WG, and WH. Panel A presents MCRRs for Brent. The results show that the WG model outperforms competing models uniformly across all scales. For the original returns, relative to an unhedged position, there is a substantial reduction of MCRR for the wavelet-full hedge model. Panel B presents estimated MCRRs for the FTSE100 stock index. The WG model outperforms competing models mostly at lower and intermediate scales (scales 1 and 4). The WOLS has the lowest MCRR at intermediate scales (scales 2 and 3). The WFH strategy performs the best for scale 5.

Panel C presents the estimated MCRR for gold. The WG model dominates competing models at most scales, i.e., scales 1, 2, 4 and 5. The WOLS has the lowest MCRR at an intermediate scale (scale 3). Panel D presents MCRR for USD. The result shows that the WG model outperforms competing hedging models at scales 1, 2, 3 and 4. The WFH model performs best at scale 5. When the results are compared across Panels A-D, it is evident that position in Brent requires more capital than for the remaining assets. The position in USD requires the least amount of MCRR compared to the positions in other assets. These comparative results also suggest that positions in currency and stock markets are less risky than positions in the commodities markets.

4 CONCLUSIONS

In this study, the wavelet decomposition of spot and futures data from three representative assets including currency, commodity and a stock index is derived to investigate the relative effectiveness of various hedging models by taking into account the hedger's exposure to time and frequency domain issues. The wavelet model allows heterogeneous hedgers with a preference for multiscale hedging which is not possible in a conventional setting. The hedging models are applied to both original returns and wavelet-decomposed returns to estimate the hedge ratios. Finally, hedging effectiveness at different time-scales is compared based on in-sample and out-of-sample portfolio variance, utility functions, minimum capital risk requirement, and VaR. Both in-sample and out-of-sample results show that the wavelet-GARCH (WG) model has the best performance overall. The WG model also dominates other hedging models in most cases when hedging performance is judged based on the mean-variance utility function. When performance is based upon exponential utility function, the wavelet-networks (WN) models outperformed other models (with the exception that the WG model exhibits superior results in some cases). Since WG based portfolio returns exhibit larger excess kurtosis at most scales, it attenuates the utility level produced by the WG model when the CE utility function is considered. Overall, the wavelet-GARCH model has the lowest MCRR.

The results of this study also confirm that wavelet-based hedging is more effective in managing risk. For both in-sample and out-of-sample, non-wavelet based hedging models have larger portfolio variances, compared to the wavelet-based hedging models.[21] Finally, the wavelet-GARCH model produces much lower portfolio variance than the standard GARCH model when applied to the original returns for all assets.

Overall, the wavelet-based hedging strategies offer superior risk reduction in most cases for the assets included in the study. Furthermore, the MCRRs from wavelet-based hedging models confirm the practical usefulness of the models by keeping regulatory capital requirements low.

References

Alexander, C. and Barbosa, A. (2008). Hedging index exchange traded fund. Journal of Banking and Finance, 32(2), 326-337.

Alexandridis, A. K. and Zapranis, A. D. (2013). Wavelet neural networks: A practical guide. Neural Networks, 42, 1-27.

Alexandridis, A. and Zapranis, A. D. (2014). *Wavelet networks: methodologies and applications in financial engineering, classification and chaos*. New Jersey, USA: Wiley.

Brenner, R., & Kroner, K. (1995). Arbitrage, Cointegration, and Testing the Unbiasedness Hypothesis in Financial Markets. The Journal of Financial and Quantitative Analysis, 30(1), 23-42. doi:10.2307/2331251

Brooks, C., Henry O. T. and Persand, G. (2002). The effects of asymmetries in optimal hedge ratios. Journal of Business, 75, 333-352.

Brown, G. W., Crabb, P. R. and Haushalter, G. D. (2006). Are firms successful at selective hedging? Journal of Business, 79, 2925-2949.

Chen, S. S., Lee, C. F. and Shrestha, K. (2004). An empirical analysis of the relationship between the hedge ratio and hedging horizon: a simultaneous estimation of the short-and long-run hedge ratios. Journal of Futures Markets, 24(4), 359-386.

Clark ,T.E. and West K.D. (2007). Approximately Normal Tests for Equal Predictive Accuracy in Nested Models. *Journal of Econometrics* 138, 291

Conlon, T., Cotter, J. and Gencay, R. (2016). Commodity futures hedging, risk aversion and the hedging horizon. European Journal of Finance, 22 (15), 1534-1560.

Conlon, J. and Cotter, J. (2012). An empirical analysis of dynamic multiscale hedging using wavelet decomposition. Journal of Futures Markets, 32(3), 272-299.

Conlon, J., Lucy, Brian, and Uddin, Gazi (2017). Is Gold a hedge against inflation? A wavelet time-scale perspective. *Review of Quantitative Finance and Accounting*, 51 (2), 317-345.

Efron, B. (1982). *Jacknife, the bootstrap, and other resampling plans*. Philadelphia: Society for Industrial and Applied Mathematics.

Fernandez, V. (2008). Multi-period hedge ratios for a multi-asset portfolio when accounting for returns co-movement. Journal of Futures Markets, 28(2), 182-207.

Geppert, J. (1995). A statistical model for the relationship between futures contract hedging effectiveness and investment horizon length. Journal of Futures Markets, 15(5), 507-536.

Harris, R. D. F. and Shen, J. (2003). Hedging and value at risk. Journal of Futures Markets, 26(4), 369-390.

Haushalter, G. D. (2000). Financing policy, basis risk, and corporate hedging: evidence from oil and gas producers. Journal of Finance, 55, 107-152.

In, F. and Kim, S. (2006). The hedge ratio and the empirical relationship between stock and futures markets: a new approach using wavelet analysis. Journal of Business, 79(2), 799:820.

Johnson, L. (1960). The theory of hedging and speculation in commodity Futures. Review of Economic Studies,27,139-151.

Kamara, A., Korajczyk, R. A., Lou, X. and Sadka, R. (2016). Horizon pricing. Journal of Financial and Quantitative Analysis, 51(6), 1769-1793.

Kroner, K. F. and Sultan, J. (1993). Time-varying distributions and dynamic hedging with foreign currency futures. Journal of Financial and Quantitative Analysis,28(4), 535-51.

Lien, D. and Luo, X. (1993). Estimating multiperiod hedge ratios in cointegrated markets. Journal of Futures Markets, 13 (8), 909-920.

Lien, D. and Shrestha, K. (2007). An empirical analysis of the relationship between hedge ratio and hedging horizon using wavelet analysis. Journal of Futures Markets, 27(2), 127-150.

Lindsay, R. Percival, D. and Rothrock, D (1996). The discrete wavelet transform and the scale analysis of the surface properties of sea ice. IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 34 (3), 771-787.

Maharaj E. A., Galagedera D. U. A. and Dark J. (2011). A comparison of developed and emerging equity market return volatility at different time scales. Managerial Finance, 2011, 37(10), 940-952.

Percival, D. and Walden, A. (2000) *Wavelets analysis for time series analysis*. Cambridge, UK: Cambridge University Press.

Rapach, D. E., Strauss, J. K., and Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies*, 23(2), 821-862

Welch, I., and Goyal, A. (2007). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4), 1455-1508.
Yang, M. J. and Lai, Y. C. (2009). An out-of-sample comparative analysis of hedging performance of stock index futures: dynamic versus static hedging. *Applied Financial Economics*,19, 1059-1072.

Data citation: The data that support the findings of this study are available from Bloomberg terminal. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the authors with the permission of Bloomberg.

Figure 1: A Feedforward wavelet networks

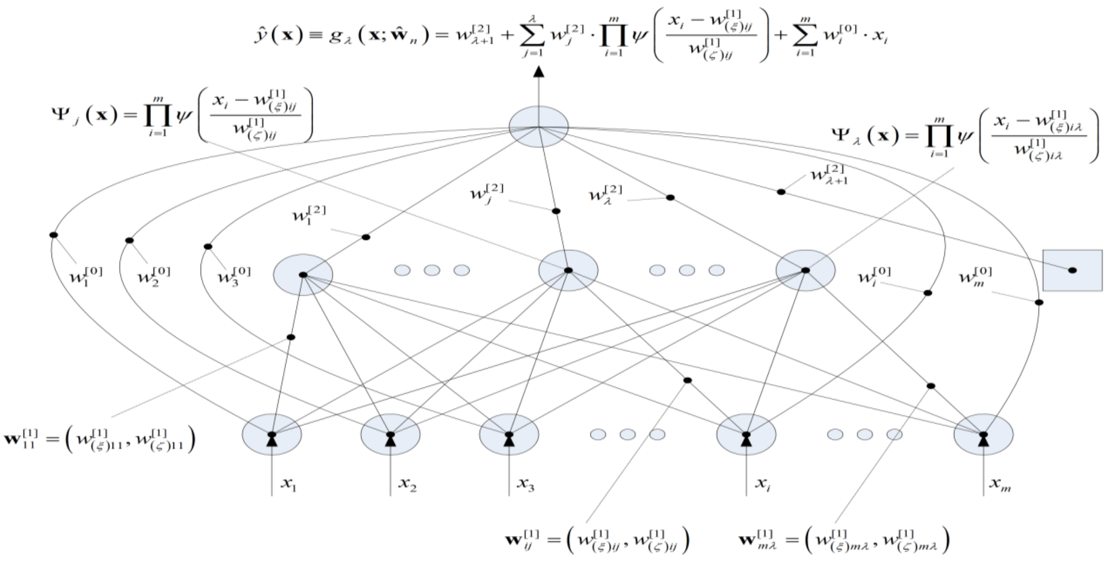

Figure 2. Wavelet decomposition at 5 levels of FTSE100 spot (top panel) and FTSE100 futures (bottom panel) log-returns from 3/1/2004 –14/12/2018.

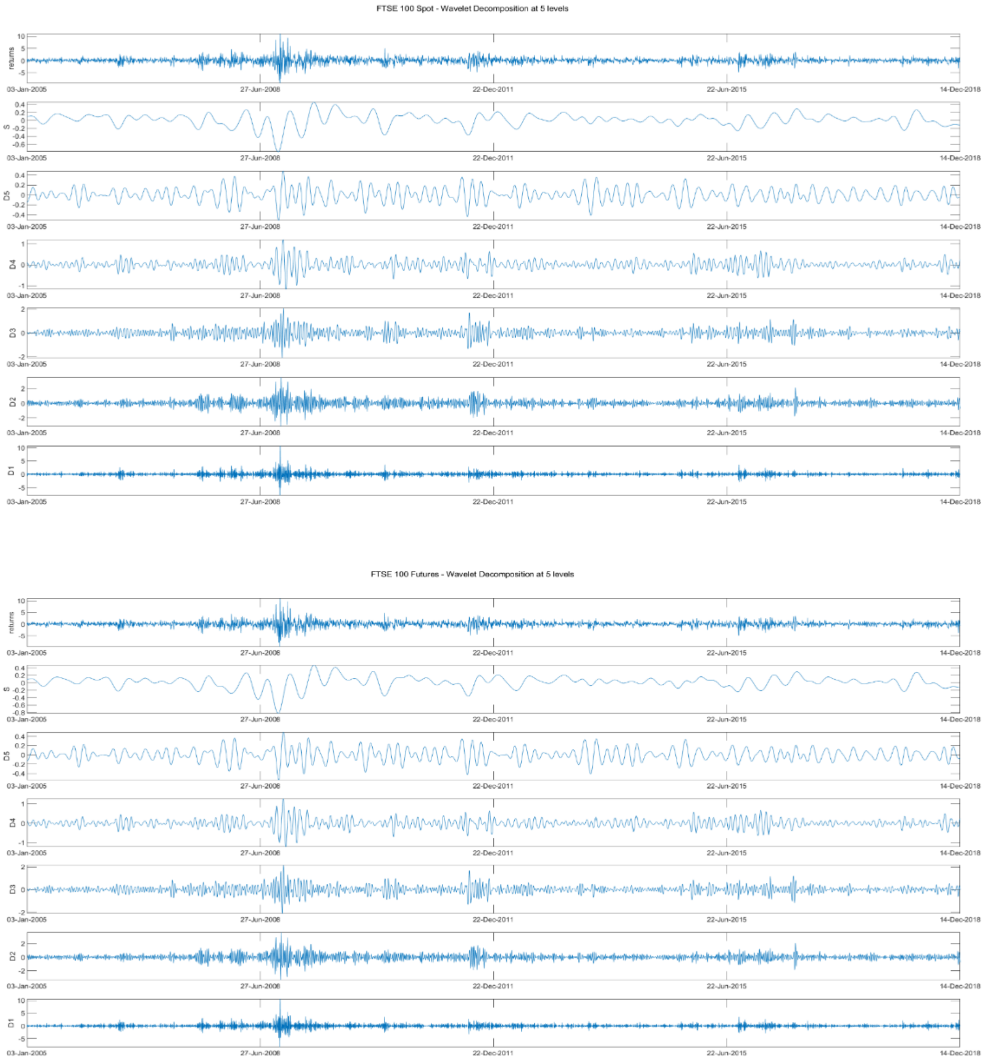

Table 1. Descriptive statistics for the spot and futures returns (original and wavelet decomposed returns)

Panel A											
Brent Spot	Mean	St.Dev	Max	Median	Min	Skewness	Kurtosis	JB	p-value	ARCH Test	p-value
Original	0.0110	2.1371	13.4668	0.0259	-12.4020	-0.1365	6.2815	1577.21	0.0000	181.88	0.0000
D1	0.0000	1.4637	8.8702	-0.0034	-9.0478	-0.1151	7.0007	2335.84	0.0000	1156.81	0.0000
D2	0.0000	0.8134	4.0978	-0.0063	-4.0640	0.0278	4.7527	447.29	0.0000	228.71	0.0000
D3	0.0000	0.5745	4.8965	0.0021	-3.6470	0.1697	9.5657	6287.17	0.0000	2118.58	0.0000
D4	0.0000	0.3985	2.0906	0.0010	-2.0621	0.0682	5.4190	853.89	0.0000	2977.99	0.0000
D5	0.0000	0.2821	1.2125	0.0069	-1.0350	0.0024	3.9240	124.19	0.0000	3355.55	0.0000
S	0.0110	0.3901	0.7466	0.0507	-1.7950	-1.3588	6.1279	2497.44	0.0000	3480.12	0.0000
Brent Futures	Mean	St.Dev	Max	Median	Min	Skewness	Kurtosis	JB	p-value	ARCH Test	p-value
Original	-0.0150	2.0642	12.8817	0.0549	-11.4110	-0.1758	6.5701	1871.88	0.0000	88.57	0.0000
D1	0.0000	1.4337	8.9239	-0.0144	-8.7923	-0.0619	6.5405	1825.61	0.0000	1201.83	0.0000
D2	0.0000	0.7764	3.5406	0.0081	-3.4916	-0.0013	4.8551	500.59	0.0000	241.85	0.0000
D3	0.0000	0.5486	4.1218	0.0009	-3.5781	0.1715	8.4619	4356.48	0.0000	2083.31	0.0000
D4	0.0000	0.3772	1.9908	0.0109	-1.9089	0.0577	5.0364	605.17	0.0000	2968.86	0.0000
D5	0.0000	0.2624	1.0187	-0.0041	-0.9171	0.0129	3.6891	69.17	0.0000	3348.91	0.0000
S	-0.0150	0.3809	0.6110	0.0403	-1.8014	-1.4709	6.4381	2978.13	0.0000	3481.39	0.0000
Panel B											
FTSE 100 Spot	Mean	St.Dev	Max	Median	Min	Skewness	Kurtosis	JB	p-value	ARCH Test	p-value
Original	0.0101	1.1426	11.1124	0.0254	-9.2656	0.0291	13.0994	14836.86	0.0000	452.14	0.0000
D1	0.0000	0.7793	10.6474	-0.0160	-7.9935	0.5330	20.2547	43471.76	0.0000	1253.52	0.0000
D2	0.0000	0.4416	3.5561	-0.0023	-3.1893	0.0888	9.1537	5512.76	0.0000	341.89	0.0000
D3	0.0000	0.3126	2.0766	0.0030	-2.1027	0.0520	7.4047	2823.66	0.0000	1866.84	0.0000
D4	0.0000	0.2132	1.2002	-0.0002	-1.1766	-0.0428	6.5310	1814.66	0.0000	3041.06	0.0000
D5	0.0000	0.1353	0.4796	-0.0029	-0.5128	-0.0910	3.5382	46.94	0.0000	3327.64	0.0000
S	0.0101	0.1529	0.4589	0.0274	-0.7896	-1.1148	6.8456	2874.28	0.0000	3471.38	0.0000
FTSE 100 Futures	Mean	St.Dev	Max	Median	Min	Skewness	Kurtosis	JB	p-value	ARCH Test	p-value
Original	0.0168	1.1439	10.9818	0.0351	-9.7002	0.0047	13.4851	15991.34	0.0000	438.29	0.0000
D1	0.0000	0.7822	10.5837	-0.0100	-8.3093	0.4820	20.4771	44565.14	0.0000	1246.11	0.0000
D2	0.0000	0.4391	3.6919	-0.0020	-3.3991	0.0814	9.6437	6424.29	0.0000	326.83	0.0000
D3	0.0000	0.3123	2.1328	0.0040	-2.1082	0.0524	7.5402	3000.01	0.0000	1866.38	0.0000
D4	0.0000	0.2164	1.2657	-0.0008	-1.2147	-0.0295	6.8703	2179.37	0.0000	3049.96	0.0000
D5	0.0000	0.1355	0.4933	-0.0003	-0.5404	-0.0941	3.6107	59.40	0.0000	3326.66	0.0000
S	0.0168	0.1557	0.4769	0.0347	-0.8197	-1.1667	7.2904	3469.40	0.0000	3471.92	0.0000

St.Dev: Standard deviation

JB: Jarque – Berra statistic

ARCH Test: Engle's Lagrange Multiplier (TR^2) statistic.

D1-D5 are the five wavelet scales.

Table 1 (contd.) Descriptive statistics for the spot and futures returns (original and wavelet decomposed returns)

Panel C											
Gold Spot	Mean	St.Dev	Max	Median	Min	Skewness	Kurtosis	JB	p-value	ARCH Test	p-value
Original	0.0297	1.1556	10.2451	0.0467	-9.5121	-0.3540	9.1061	5496.20	0.0000	37.71	0.0000
D1	0.0000	0.7686	6.1289	-0.0098	-4.2799	-0.0146	6.6919	1982.79	0.0000	1111.82	0.0000
D2	0.0000	0.4481	2.3754	0.0025	-2.6212	-0.0872	5.7566	1109.71	0.0000	303.91	0.0000
D3	0.0000	0.3185	1.7715	0.0002	-1.8733	0.0394	6.1528	1446.81	0.0000	1937.80	0.0000
D4	0.0000	0.2274	0.9847	0.0026	-1.1787	-0.1307	4.7448	452.78	0.0000	2955.51	0.0000
D5	0.0000	0.1758	0.7754	-0.0018	-0.8550	-0.1159	4.8850	524.66	0.0000	3370.80	0.0000
S	0.0297	0.1697	0.5745	0.0244	-0.3839	0.2994	2.9011	53.59	0.0000	3464.86	0.0000
Gold Futures	Mean	St.Dev	Max	Median	Min	Skewness	Kurtosis	JB	p-value	ARCH Test	p-value
Original	0.0220	1.1710	8.6243	0.0308	-9.8204	-0.3570	8.5903	4619.88	0.0000	69.07	0.0000
D1	0.0000	0.7726	4.5610	-0.0039	-4.9597	-0.0378	6.1692	1461.76	0.0000	980.85	0.0000
D2	0.0000	0.4573	2.3824	0.0030	-2.4427	-0.0953	5.6969	1063.23	0.0000	293.11	0.0000
D3	0.0000	0.3267	1.9322	-0.0003	-1.7279	0.0262	5.7359	1089.16	0.0000	1931.92	0.0000
D4	0.0000	0.2316	1.0403	0.0029	-1.2053	-0.1172	4.6490	403.54	0.0000	2948.13	0.0000
D5	0.0000	0.1774	0.7745	-0.0025	-0.8504	-0.1096	4.7610	458.06	0.0000	3370.23	0.0000
S	0.0220	0.1687	0.5523	0.0159	-0.3955	0.2805	2.8968	47.32	0.0000	3464.53	0.0000
Panel D											
USD Spot	Mean	St.Dev	Max	Median	Min	Skewness	Kurtosis	JB	p-value	ARCH Test	p-value
Original	0.0054	0.5023	2.5199	0.0000	-2.7263	-0.0291	5.0550	614.77	0.0000	41.67	0.0000
D1	0.0000	0.3300	1.5914	-0.0019	-1.9528	-0.0234	4.2337	221.71	0.0000	970.42	0.0000
D2	0.0000	0.2013	1.0977	0.0018	-1.0325	-0.0215	4.5270	339.45	0.0000	227.75	0.0000
D3	0.0000	0.1394	0.6068	0.0006	-0.6355	0.0247	3.8141	96.75	0.0000	1724.58	0.0000
D4	0.0000	0.0973	0.3784	-0.0005	-0.4590	-0.1345	4.0549	172.39	0.0000	2949.05	0.0000
D5	0.0000	0.0714	0.2421	-0.0003	-0.2815	-0.0779	3.6024	56.32	0.0000	3343.41	0.0000
S	0.0054	0.0832	0.2880	0.0018	-0.1953	0.3038	3.0206	53.77	0.0000	3469.43	0.0000
USD Futures	Mean	St.Dev	Max	Median	Min	Skewness	Kurtosis	JB	p-value	ARCH Test	p-value
Original	0.0031	0.5124	2.3645	0.0050	-2.7455	-0.0206	4.9608	559.47	0.0000	30.44	0.0000
D1	0.0000	0.3379	1.6411	0.0008	-2.0166	-0.0276	4.2686	234.52	0.0000	963.34	0.0000
D2	0.0000	0.2044	1.0952	0.0016	-1.0697	-0.0229	4.5091	331.57	0.0000	220.89	0.0000
D3	0.0000	0.1423	0.5866	-0.0001	-0.5586	0.0289	3.6825	68.23	0.0000	1711.68	0.0000
D4	0.0000	0.0995	0.3981	0.0001	-0.4861	-0.1191	4.0859	179.76	0.0000	2949.86	0.0000
D5	0.0000	0.0728	0.2354	-0.0001	-0.2708	-0.0597	3.5902	52.74	0.0000	3344.95	0.0000
S	0.0031	0.0829	0.2895	-0.0012	-0.1999	0.2750	3.0115	44.01	0.0000	3468.86	0.0000

St.Dev: Standard deviation
JB: Jarque – Berra statistic
ARCH Test: Engle's Lagrange Multiplier (TR^2) statistic.
D1-D5 are the five wavelet scales.

Table 2: In-sample hedge ratios using wavelet and conventional hedging methods

	Brent	FTSE 100	Gold	USD
Wavelet hedge ratio				
D1	0.9348	0.9883	0.8227	0.9694
D2	0.9849	0.9984	0.9106	0.9784
D3	1.0168	1.0011	0.9605	0.9762
D4	1.0405	0.9854	0.9783	0.9748
D5	1.0627	0.9944	0.9914	0.9781
Wavelet-OLS hedge ratio				
D1	0.9273	0.9873	0.8152	0.9683
D2	0.9840	0.9982	0.9178	0.9795
D3	1.0221	1.0011	0.9595	0.9760
D4	1.0424	0.9836	0.9789	0.9748
D5	1.0741	0.9990	0.9901	0.9771

Table 3: Portfolio variance comparison

Panel A: In-sample portfolio variance					
Brent	WU	WFH	WOLS	WG	WH
Original	4.83322319	0.68701807	0.68176754	0.69872385	-
D1	2.28253603	0.44967171	0.43834714	0.40213483	0.43838746
D2	0.68705841	0.07935517	0.07919416	0.07099235	0.07918056
D3	0.35061475	0.01834967	0.0181938	0.01379608	0.01819954
D4	0.17144397	0.00867430	0.00840447	0.00492461	0.00840354
D5	0.08022033	0.00449425	0.00413194	0.00150110	0.00413979
Rank (1=best)	Benchmark	4	2	1	3
FTSE 100	WU	WFH	WOLS	WG	WH
Original	1.49179453	0.03148753	0.03139709	0.03171182	-
D1	0.70050274	0.02139405	0.02128086	0.01946207	0.02127788
D2	0.21968523	0.00394713	0.00394641	0.00348674	0.00394572
D3	0.10986708	0.00059306	0.00059292	0.00044234	0.00059281
D4	0.05182694	0.00016933	0.00015494	0.00006434	0.00015509
D5	0.02037019	0.00006689	0.00006687	0.00001835	0.00006727
Rank (1=best)	Benchmark	4	2	1	3
Gold	WU	WFH	WOLS	WG	WH
Original	1.53112856	0.32981063	0.30812137	0.32486496	-
D1	0.67452187	0.24481083	0.22154316	0.24775794	0.22154171
D2	0.23102572	0.0311357	0.02951757	0.02799138	0.02952441
D3	0.11773517	0.00413423	0.00393169	0.00278599	0.00393110
D4	0.05877433	0.00055319	0.00052604	0.00027318	0.00052596
D5	0.03523801	0.00007065	0.00006714	0.00002855	0.00006719
Rank (1=best)	Benchmark	4	1	3	2
USD	WU	WFH	WOLS	WG	WH
Original	0.27389183	0.00271043	0.00251763	0.00256404	-
D1	0.11842648	0.00158603	0.00146099	0.00113947	0.00146087
D2	0.04400747	0.00036432	0.00034516	0.00029095	0.00034515
D3	0.02092369	0.00012964	0.00011709	0.0000864	0.00011706
D4	0.00985246	0.00003571	0.00002915	0.00001022	0.00002915
D5	0.00584421	0.00002804	0.00002484	0.00000561	0.00002484
Rank (1=best)	Benchmark	4	3	1	2

Table 3 (contd.): Out-of-sample portfolio variance comparison

Panel B: Out-of-sample Portfolio Variance					
Brent	WU	WFH	WOLS	WG	WN
Original	3.50663900	0.42177641	0.41227107	0.41193496	-
D1	1.58302460	0.2660753	0.24738338	0.22666979	0.24718824
D2	0.55863167	0.04917766	0.04974800	0.04365603	0.05013706
D3	0.25035338	0.01958862	0.01970209	0.01404018	0.01957319
D4	0.10770567	0.00521918	0.00578596	0.00449586	0.00556955
D5	0.07412687	0.00125765	0.00185765	0.00040513	0.00155147
Rank (1=best)	Benchmark	4	3	1	2
FTSE 100	WU	WFH	WOLS	WG	WN
Original	0.56205363	0.02693962	0.02620831	0.02457513	-
D1	0.23580221	0.01826307	0.01741400	0.01299250	0.01751350
D2	0.09622728	0.00283116	0.00281487	0.00263592	0.00281354
D3	0.04838609	0.00070847	0.00070641	0.00051819	0.00071037
D4	0.01941908	0.00016596	0.00015794	0.00008511	0.00015830
D5	0.01031176	0.00010486	0.00010414	0.00007557	0.00010275
Rank (1=best)	Benchmark	4	3	1	2
Gold	WU	WFH	WOLS	WG	WN
Original	0.55347337	0.16933982	0.15454291	0.16249467	-
D1	0.25604907	0.12534196	0.11026239	0.11899600	0.11032717
D2	0.08043437	0.01633943	0.01484290	0.01424456	0.01475241
D3	0.03717171	0.00184306	0.00172106	0.00146895	0.00172062
D4	0.02361882	0.00034556	0.00031301	0.00015771	0.00031453
D5	0.01303040	0.00003292	0.00003149	0.00004654	0.00003130
Rank (1=best)	Benchmark	4	3	1	2
USD	WU	WFH	WOLS	WG	WN
Original	0.16613931	0.00097859	0.00099156	0.00096843	-
D1	0.07085864	0.00058671	0.00059914	0.00041252	0.00059542
D2	0.02655027	0.00013194	0.00013364	0.00010218	0.00013290
D3	0.01342547	0.00004329	0.00003995	0.00002057	0.00003945
D4	0.00737932	0.00001269	0.00001352	0.00000551	0.00001251
D5	0.00253773	0.00000430	0.00000606	0.00002239	0.00000566
Rank (1=best)	Benchmark	4	3	1	2

Panel C: In-sample and out-of-sample average rank of hedging models across all scales and assets

	WU	WFH	WOLS	WG	WH/WN
Average Rank In Sample	5.00	3.90	2.55	1.15	2.40
Average Rank Out-of-Sample	5.00	3.15	2.95	1.40	2.50

Table 4 F-test of equality of portfolio variance

Panel A: In sample F-Test of equality of variance										
Asset	Horizon	WFH	WOLS	WG	WH	WOLS	WG	WN	WG	WH
Brent	D1		accept	Reject	accept		reject	accept		reject
	D2		accept	Reject	accept		reject	accept		reject
	D3		accept	Reject	accept		reject	accept		reject
	D4		accept	Reject	accept		reject	accept		reject
	D5		reject	Reject	reject		reject	accept		reject
FTSE 100	D1		accept	Reject	accept		reject	accept		reject
	D2		accept	Reject	accept		reject	accept		reject
	D3		accept	Reject	accept		reject	accept		reject
	D4		reject	Reject	reject		reject	accept		reject
	D5		accept	Reject	accept		reject	accept		reject
Gold	D1		reject	Accept	reject		reject	accept		reject
	D2		accept	Reject	accept		accept	accept		accept
	D3		accept	Reject	accept		reject	accept		reject
	D4		accept	Reject	accept		reject	accept		reject
	D5		accept	Reject	accept		reject	accept		reject
USD	D1		reject	Reject	reject		reject	accept		reject
	D2		accept	Reject	accept		reject	accept		reject
	D3		reject	Reject	reject		reject	accept		reject
	D4		reject	Reject	reject		reject	accept		reject
	D5		reject	Reject	reject		reject	accept		reject

Panel B: Out-of-sample F-Test of equality of variance										
Asset	Horizon	WFH	WOLS	WG	WN	WOLS	WG	WN	WG	WN
Brent	D1		accept	reject	accept		accept	accept		accept
	D2		accept	accept	accept		accept	accept		accept
	D3		accept	reject	accept		reject	accept		reject
	D4		accept	reject	accept		reject	accept		reject
	D5		reject	reject	reject		reject	reject		reject
FTSE 100	D1		accept	reject	accept		reject	accept		reject
	D2		accept	accept	accept		accept	accept		accept
	D3		accept	reject	accept		reject	accept		reject
	D4		accept	reject	accept		reject	accept		reject
	D5		accept	reject	accept		reject	accept		reject
Gold	D1		accept	accept	accept		accept	accept		accept
	D2		accept	accept	accept		accept	accept		accept
	D3		accept	reject	accept		reject	accept		reject
	D4		accept	reject	accept		reject	accept		reject
	D5		accept	reject	accept		reject	accept		reject
USD	D1		accept	reject	accept		reject	accept		reject
	D2		accept	reject	accept		reject	accept		reject
	D3		accept	reject	accept		reject	accept		reject
	D4		accept	reject	accept		reject	accept		reject
	D5		reject	reject	reject		reject	accept		reject

Table 5: MCRR estimates

Hedge Horizon	WU	WFH	WOLS	WG	WH
Panel A: Brent					
Original data	6.356960	1.836884	1.881271	2.086830	
D1	1.356155	0.633175	0.623811	0.630667	0.630667
D2	0.678162	0.390851	0.398539	0.360413	0.392816
D3	0.288627	0.063544	0.061831	0.050072	0.064361
D4	0.734238	0.871018	0.854197	0.825759	0.872607
D5	0.045541	0.161417	0.160866	0.099804	0.160682
Panel B: FTSE 100					
Original data	2.640111	0.377684	0.368986	0.362332	
D1	0.920615	0.218726	0.201823	0.136983	0.199673
D2	0.726635	0.134191	0.132165	0.140435	0.133940
D3	0.311913	0.024092	0.023897	0.026718	0.024394
D4	0.315177	0.032373	0.025533	0.010510	0.025239
D5	0.282599	0.000870	0.002743	0.003059	0.002292
Panel C: Gold					
Original data	2.154113	0.562533	0.726309	0.693793	
D1	1.054171	0.292985	0.329068	0.292718	0.322496
D2	0.391429	0.406279	0.344628	0.226207	0.332342
D3	0.125930	0.117914	0.112527	0.118316	0.114712
D4	0.221394	0.024424	0.019411	0.019404	0.019893
D5	0.308875	0.014894	0.011627	0.001814	0.012540
Panel D: USD					
Original data	0.874643	0.040738	0.033507	0.032705	
D1	0.325977	0.021700	0.017967	0.016332	0.018883
D2	0.275179	0.007110	0.008015	0.006809	0.008246
D3	0.093960	0.008395	0.006730	0.006564	0.006731
D4	0.138753	0.004196	0.001331	0.001128	0.001268
D5	0.227967	0.000403	0.005412	0.001565	0.005399

Endnotes

[1]Kamara et al. (2016) noted that in a well-segmented market with horizon clienteles, different assets are priced with different pricing kernels. They found that value (liquidity) risk is priced over intermediate (short) horizons; long-horizon investors focus on investing in less liquid but high-return assets. For instance, highly leveraged hedge funds may prefer short-run horizons and liquid stocks. In comparison, pension funds, mutual funds, and long-term investors prefer to invest in high-yield but less liquid assets.

[2]Haushalter (2000) noted that oil and gas producers on average hedge nearly 30% of one-year production. Brown et al. (2006) also noted that gold producers prefer selective hedging.

[3] www.bloomberg.com/news/articles/2015-08-04/chinese-airlines-benefit-as-oil.

[4] China Eastern Airlines incurred fuel hedging losses totalling $690 million in November, 2008, attributed to a massive price reduction of crude oil for February 2009 delivery, from $147.27 to $40.50 per barrel on the NYMEX. The carrier used most of the contracts with maturities of two to three years to stabilize jet fuel costs. Cathay Pacific Airways also suffered losses in 2008 (see www.chinadaily.com.cn/bizchina/2009-01/13/content_7390689.htm).

[5] For the same time-period, the number of observations obtained is different for different time-frequencies.

[6]Chen et al. (2004) noted that the long-run refers to an investment horizon which is longer than 8 weeks.

[7]The hedging effectiveness is typically measured by the percentage reduction in the variance of the spot returns (unhedged portfolio) relative to the variance of the hedged portfolio. In and Kim (2006) defined this degree of hedging effectiveness as equal to the square of the correlation between the spot and futures price changes.

[8]A $\beta = 0$ implies unhedged position; $\beta = 1$ signifies a fully hedged position; and $\beta < 1$ implies a partial hedge.

[9]The inclusion of the EC is consistent with the finding of cointegration for all spot and futures prices for the assets in this study. As Brenner and Kroner (1995) noted, if markets are cointegrated, the basis is stationary. Cointegration test results are available upon request.

[10]Regarding the consequence of differencing matched with the long-investment horizon, Geppert (1995) noted that the regression-based method results in two complications: (1) the use of non-overlapping data tends to reduce the sample size; and (2) the use of overlapping differences introduces spurious statistical properties into the series. The wavelet method applied to daily data in this study (similar to Lien and Shrestha (2007)) alleviates both problems.

[11]The wavelet GARCH model is essentially the same as in equations 4-6 with the exception that the dependent variables are wavelet-decomposed returns. Also, the error correction term (EC) (equations 4-6) is specific to the original returns. Since the original returns are decomposed into horizon-specific returns, the inclusion of the EC in the mean equations for wavelet models is debatable. The results from alternative models without the EC term are qualitatively similar. The inclusion of the EC term is based on the notion that a long-run no-arbitrage relationship exists even though the investor may have a preference for a short run hedge horizon.

[12] Thanks to an anonymous referee for pointing this.

[13]While the updated sample is arbitrary, it reflects the fact that trading in the DX (USD futures on the Intercontinental Exchange) has increased since 2005.

[14]The aggregate rank for WG (23) is based on its asset-specific ranks: Brent (5), FTSE200 (5), Gold (8), and USD (5).

[15]In particular, in the case of the GARCH and WG, an iterative updating process is followed to forecast the out-of-sample hedge ratios. Simply, when the in-sample maximum likelihood optimization converges, it produces the hedge ratio for the next day (t+1). Next, one out-of-sample observation is added to estimate the hedge ratio for t+2. This iterative process is repeated until all out-of-sample observations are exhausted. At each step, there is a risk that the model may not converge. In case of non-convergence, the earliest observation of the in-sample is dropped for re-optimizing until the model converges. When the model converges, the hedge ratio is retrieved for calculating the portfolio. The purpose is to allow the optimization algorithm to bypass a difficult area of the likelihood surface. Once convergence occurs, the algorithm returns to the normal updating of the hedge ratio using only in-sample data. Non-convergence issues occurred mostly at higher scales.

[16]The rankings of the hedging strategies based on portfolio variance are as follows: WG (28), WN (50), WOLS (59), WFH (63), and WU (100).

[17]Thanks to an anonymous referee for suggesting this test.

[18]Analytical results, including all statistical tests, are available upon request from the authors.

[19]Yang and Lai (2009) noted that transaction cost ranges between 0.005% and 0.01% at major exchanges.

[20]Papers by Conlon and Cotter (2012), Harris, and Shen (2006) reported that the kurtosis of a hedged portfolio is greater than that for an unhedged portfolio across scales. Further, the authors claimed that when hedging effectiveness is measured using variance reduction and the value at risk (VaR), excess kurtosis reduces the effectiveness of a hedged portfolio compared to the VaR minimization metric. This result suggests that researchers should consider higher moments to account for the effects of VaR minimization in reducing tail risk at longer time-horizons.

[21] Thanks to an anonymous referee for pointing this out.