Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Hedging with Futures: Does Anything Beat the Naïve Hedging Strategy?

Yudong Wang, Chongfeng Wu, Li Yang

MANAGEMENT SCIENCE
Vol. 61, No. 12, December 2015, pp. 2870–2889
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2014.2028

Hedging with Futures: Does Anything Beat the Naïve Hedging Strategy?

Yudong Wang, Chongfeng Wu
Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai 200052, China
{wangyudongnj@126.com, cfwu@sjtu.edu.cn}

Li Yang
School of Banking and Finance, University of New South Wales, Sydney, New South Wales 2052, Australia,
l.yang@unsw.edu.au

This paper investigates out-of-sample performance of the naïve hedging strategy relative to that of the minimum variance hedging strategy, in which the covariance parameters are estimated from 18 econometric models. Hedging performance is compared across 24 futures markets. Our main findings suggest that it is difficult to find a strategy under the minimum variance framework that outperforms the naïve hedging strategy both consistently and significantly. Our findings are robust to different sample periods, estimation windows, and hedging horizons and can be partly explained by the effects of estimation error and model misspecification.

Data, as supplemental material, are available at http://dx.doi.org/10.1287/mnsc.2014.2028.

Keywords: hedging with futures; naïve strategy; minimum variance hedge ratios; estimation error; model misspecification

History: Received February 11, 2013; accepted May 27, 2014, by Itay Goldstein, finance. Published online in *Articles in Advance* January 23, 2015.

1. Introduction

The development of hedging strategies using futures has been a topic of academic and practical interest since the introduction of futures markets. An optimal strategy increases the efficiency of risk management and minimizes the costs of hedging. From a theoretical perspective, hedging with futures can be considered a special case of asset allocation that involves two assets: the underlying asset and its corresponding futures. However, hedgers differ from portfolio investors in the sense that hedgers are more concerned about the risks they face than the returns they can achieve. Thus, the objective of hedging is slightly different from that of asset allocation. Mean-variance portfolio investors seek a trade-off between asset return and risk when they manage their assets, whereas hedgers aim to minimize their risk but leave their returns unmanaged. Given this difference in objectives, the mean-variance framework for asset allocation has been modified to a minimizing-variance framework in the hedging literature to help hedgers determine an optimal hedging strategy.[1] Under this framework, the optimal hedging strategy is obtained as the ratio of the covariance between spot and futures returns to the variance of futures returns, which is often denoted the optimal hedge ratio (OHR) and can be interpreted as the number of short (long) positions of futures contracts a hedger should optimally take for each unit of long (short) positions held in the underlying asset.

The issue is raised empirically about how to estimate the covariance between spot and futures returns and the variance of futures returns for determining an OHR. A number of econometric models have been utilized to model the joint distribution of spot and futures returns and thus estimate the covariance matrix of the two variables. Linear regression models that assume a constant covariance between spot and futures returns over time have frequently been employed to produce time-invariant hedge ratios, such as the ordinary least

[1] In the asset allocation literature, because errors in the estimates of sample means have a larger impact on portfolio weights than do errors in the estimates of the covariances of asset returns, recent research has been focused on constructing minimum variance portfolios, which rely solely on covariance estimates (e.g., Green and Hollifield 2002, Jagannathan and Ma 2003, Ledoit and Wolf 2003, DeMiguel et al. 2009). Although the objective of the minimum variance portfolio is the same as for hedging, the construction of minimum variance portfolios often imposes various constraints. One of the practical constraints imposed is that the portfolio weights must not be negative because short selling is not easy to implement in the market. In contrast, hedging with futures has no such constraints. One can take short positions as easily as long positions in the futures market. Hence, the optimal solutions to these problems with the same objectives but different constraints are different.

squares (OLS) regression used by Ederington (1979), the vector autoregression (VAR) model employed by Myers and Thompson (1989), and the vector error correction (VEC) model and its variations used by Lien and Yang (2006). However, as discussed by Kroner and Sultan (1993), because the distributions of many asset returns are time-varying, and thus the covariance matrix of the asset returns changes over time, it is more realistic to expect that the OHR is time-varying rather than constant. To address this issue, the multivariate generalized autoregressive conditional heteroskedasticity (MGARCH) model of Bollerslev (1990) and its various extensions have been applied to estimate the time-varying nature of the variance and covariance parameters and, consequently, the time-varying OHR. Baillie and Myers (1991) were the first to estimate a time-varying OHR using a bivariate GARCH (BGARCH) model in the commodity market. Haigh and Holt (2002) demonstrate that a BGARCH model that accounts for volatility spillovers between spot and futures markets results in a better hedging strategy compared with a BGARCH model that ignores spillover effects. Alizadeh et al. (2008) find that the OHR based on the Markov regime-switching BGARCH (RS-GARCH) model performs better than the naïve strategy and time-invariant OHRs. Because asymmetric dependence among asset returns has been well documented in more recent literature (e.g., Andrew and Chen 2002, Patton 2006), copula-based GARCH models that can flexibly capture the nonlinear dependence between asset returns have been applied to estimate the covariance and variance parameters in the OHR. For instance, Hsu et al. (2008) show that an OHR based on copula-based GARCH models can result in higher hedging effectiveness than various BGARCH-type models. Lee (2009) proposes a regime-switching Gumbel–Clayton copula GARCH model and finds that the OHR estimated from this model provides better out-of-sample hedging outcomes compared with the OLS and BGARCH-type models. Although constant progress has been made in applying more sophisticated statistical models to flexibly describe the joint distribution of spot and futures returns and thus improve the accuracy of the covariance matrix estimate, there is still no consensus on the best model strategy.

In this paper, we comprehensively investigate the hedging performance of different strategies. In particular, we are interested in the performance of the naïve hedging strategy with a hedge ratio of 1 and its performance relative to that of strategies derived under the minimum variance framework. Given its simplicity and ease of implementation, the naïve hedging strategy has been widely used by companies as a benchmark. The strategy is also cost efficient because there is no need to rebalance hedging positions so that transaction costs are minimal. For example, Mike Corley, a longtime energy and risk management consultant and president of Houston-based Mercatus Energy Advisors, recounts the story of a family-owned, Oklahoma-based oil and gas producer that hedged conservatively: "The company would sell 70% of PDP [proven developed (oil) products] forward for the front year, 40% for the second, and 20% for the third" (Hickey 2011). However, hedging strategies under the minimum variance framework have been commonly considered optimal. These strategies have been studied extensively in the literature, taught in classrooms, and utilized by individual and institutional hedgers. From a theoretical perspective, one would expect that any minimum variance hedging strategy would outperform the naïve strategy. However, several empirical studies have demonstrated that this may not be the case in the stock index futures market (e.g., Alexander and Barbosa 2007).

In this paper, we examine the performance of naïve and minimum variance hedging strategies across 24 different futures markets, whose underlying assets include commodities, currencies, and stock indices. Our sample spans the period from January 1994 to December 2011, and 18 different econometric models are employed to estimate the covariance parameters for minimum variance hedging strategy. Our evaluation criterion is based on the return variance of the hedged portfolio, and a strategy is considered superior if it reduces sample variance more significantly than another strategy. We use the test of Diebold and Mariano (1995) (DM) to evaluate the difference in hedging performance between two strategies. The DM test produces statistical inferences for the evaluation, which does not depend on a particular sample realization and is more reliable than the conclusion drawn based on a simple comparison of the variances between the two strategies in a particular sample. In addition, when the data are not sufficiently informative for a single best strategy to be selected, we use the recently developed model confidence set (MCS) test of Hansen et al. (2011) to select a smaller set of strategies that contains the best strategy with a given confidence level and, moreover, assess whether the naïve strategy performs as well as the set of selected strategies.

Our empirical results suggest that the naïve strategy cannot be significantly and consistently outperformed by any of the 18 strategies across the 24 futures markets examined. These results are robust to the period of the recent financial crisis in 2008–2009, to a change in the length of the estimation window, to various hedging horizons, and to a new set of strategies created by combining the 18 different strategies considered.

Next, we investigate the reasons for our empirical findings. We consider two possible explanations: estimation error and model misspecification. We first examine the effects of estimation error on out-of-sample hedging performance. Because of errors in the estimates of the covariance matrix of spot and futures returns, the out-of-sample performance of any particular strategy can be much worse than its in-sample performance (e.g., Jorion 1992). The asset allocation literature has extensively documented that the estimation error has considerable effects on the out-of-sample effectiveness of a portfolio (e.g., Bawa et al. 1979, Clarkson et al. 1996, Britten-Jones 1999, Siegel and Woodgate 2007, DeMiguel et al. 2009). Because hedging is a special case of asset allocation, one would expect that the estimation error also plays an important role in the out-of-sample performance of a hedging strategy. More sophisticated and complex strategies do not necessarily generate better hedging outcomes than simple ones because the inclusion of more parameters in the strategy produces larger estimation errors in their estimates. To quantify this effect, we compare the hedging performances of the strategies in which the parameters are assumed to be ex ante known to hedgers during the out-of-sample period with the hedging performances of the strategies in which the parameters are assumed to be unknown. We find that the naïve strategy is significantly outperformed by several strategies, such as OLS and regime switching, for several futures markets when the parameters in the OHR are ex ante known to hedgers. By contrast, the naïve strategy performs as well as any of the 18 strategies across the 24 futures markets when the variance and covariance parameters in the OHRs are unknown. We attribute these contrasting results in part to errors in the estimates of the parameters in the OHRs. When the parameters in the OHR are known and do not need to be estimated, one would expect that for all futures markets, the minimum variance hedging strategy would outperform the naïve strategy, because the minimum variance hedging strategy is optimal and should outperform any other strategies. However, our results indicate that the naïve strategy is outperformed by several, rather than all, minimum variance strategies for a few futures markets. The results suggest that the estimation error indeed plays an important but partial role in explaining our main empirical findings.

Second, we consider model misspecification to be a possible explanation for our findings. Although a wide range of model specifications have been applied to describe the joint dynamics of spot and futures returns, any particular model specification may not be sufficient to flexibly capture the dynamic nature of spot and futures returns over time. Therefore, it is possible that a more sophisticated model performs poorly during the out-of-sample period because of its misspecification. To measure the impact of model misspecification on our results, we reexamine the hedging performance using simulated data. Our results demonstrate that the naïve strategy can be significantly outperformed by the true model strategy but performs as well as some other misspecified model strategies. Hence, model misspecification can partially explain our empirical findings.

This paper addresses a long-standing question in the hedging literature: Do more sophisticated strategies significantly improve hedging performance relative to the simple naïve strategy? Our empirical findings suggest that the naïve strategy performs as well as more sophisticated strategies, consistent with several recent asset allocation studies that challenge the existing portfolio diversification theory. For instance, DeMiguel et al. (2009) find that none of the strategies under consideration is consistently better than the $1/N$ rule for a portfolio consisting of many assets. In addition, the authors conclude that when the number of assets in a portfolio and the correlation of these assets are small, the $1/N$ strategy may be outperformed by sophisticated strategies. However, we find that even for a portfolio of only two assets—the underlying asset and its corresponding futures, which are highly correlated—the naïve strategy still cannot be significantly outperformed. Although DeMiguel et al. (2009) attribute the superiority of the $1/N$ strategy to estimation error, we consider another possible reason: model misspecification. We find that not only estimation error but also model misspecification can partly explain our main empirical results. Our paper is also related to the recent work of Tu and Zhou (2011), who show that, although the Markowitz portfolio rule and its various extended rules underperform the naïve $1/N$ rule, the combinations of several sophisticated models can outperform the $1/N$ rule. In contrast, we find that for a special portfolio of two assets, an underlying asset and its futures, the naïve hedging strategy cannot be significantly outperformed by model combination strategies.

The remainder of this paper is organized as follows. Section 2 describes the minimum variance hedging strategy and the models for estimating the parameters of the OHR. Section 3 describes the hedging performance measure and two statistical tests to draw its inferences. Section 4 discusses the data and the preliminary analysis results. Section 5 presents the main empirical results. Section 6 concludes the paper.

2. Minimum Variance Hedging Strategies

Consider an investor who is holding an underlying asset, such as a stock or commodity, and aims to hedge the price risk of the underlying asset by using its corresponding futures contract. At time t, the investor has to determine an optimal futures position to minimize the risk of the combined positions of the underlying assets and futures at time $t+1$. The combined position is denoted by the hedged portfolio. The risk of the hedged portfolio is measured by the variance of the hedged portfolio returns. To determine optimal futures positions, let $r_{s,t}$ and $r_{f,t}$ be the returns of the spot and futures at time t, respectively, and let γ_t be the hedge ratio at time t, defined as the number of futures positions taken for each unit of spot position held at time t. The hedged portfolio return at time $t+1$, $r_{p,t+1}$, is then given by $r_{p,t+1} = r_{s,t+1} - \gamma_t r_{f,t+1}$. Moreover, the variance of the hedged portfolio return at time $t+1$, $\text{var}(r_{p,t+1})$, is given by $\text{var}(r_{p,t+1}) = \text{var}(r_{s,t+1}) + \gamma_t^2 \text{var}(r_{f,t+1}) - 2\gamma_t \text{cov}(r_{s,t+1}, r_{f,t+1})$. Upon minimization of the above variance, the OHR at time t is determined by

$$\gamma_t^* = \frac{\text{cov}(r_{s,t+1}, r_{f,t+1})}{\text{var}(r_{f,t+1})}, \quad (1)$$

where OHR $\equiv \gamma_t^*$. An issue is raised empirically about how to model the joint dynamics of spot and futures returns so that $\text{cov}(r_{s,t+1}, r_{f,t+1})$ and $\text{var}(r_{f,t+1})$ can be estimated accurately.

Traditionally, regression models such as the OLS, VAR, and VEC models have been utilized to describe the dynamics of spot and futures returns under the assumption that $\text{cov}(r_{s,t}, r_{f,t})$ and $\text{var}(r_{f,t})$ are constant over time. Hence, the OHRs estimated from these models are time invariant. In recent years, GARCH-type models have been broadly employed to capture the time-varying variance and covariance of asset returns. When a BGARCH model is applied to describe the joint dynamics of spot and futures returns, it produces time-varying OHRs. Time-varying OHRs can also be obtained from other types of econometric models, such as regime-switching models, which describe different behaviors of spot and futures returns under different states of the economy or financial conditions.

In this paper, we consider 18 different econometric models that are commonly used in the hedging literature to estimate OHRs. Table 1 lists these models, including the regression models, GARCH-type models, regime-switching models, and copula-based models. We briefly discuss each of them in the following section.

Table 1 Models Under Consideration

No.	Model	Abbreviation
	Naïve	
0	Hedge ratio = 1	Naïve
	Models that produce constant hedge ratios	
1	Ordinary least squares	OLS
2	Vector autoregressions	VAR
3	Vector error corrections	VEC
4	Vector error correction with spot-futures basis	Adj.VEC
5	Fractional integration vector error corrections	FIVEC
	Models that produce dynamic hedge ratios	
	Bivariate GARCH models	
6	Bivariate BEKK-GARCH	BEKK
7	Asymmetric bivariate BEKK-GARCH	ABEKK
8	Constant conditional correlation GARCH	CCC
9	Asymmetric constant conditional correlation GARCH	ACCC
10	Dynamic conditional correlation GARCH	DCC
11	Asymmetric dynamic conditional correlation GARCH	ADCC
	Copula models	
12	T-copula with GARCH margins	GARCH-copula
13	T-copula with GJR margins	GJR-copula
	Regime-switching models	
14	Regime-switching ordinary least squares	RS-OLS
15	Regime-switching vector autoregressions	RS-VAR
16	Regime-switching vector error corrections	RS-VEC
17	Regime-switching scalar BEKK-GARCH	RS-BEKK
18	Regime-switching dynamic conditional correlation GARCH	RS-DCC

Notes. This table lists the models that describe joint distribution of spot and futures returns considered in our investigation. There are two groups of models. The first group assumes that the variance of and the covariance between spot and futures returns are constant, and hence it produces a constant OHR (or strategy). The second group captures the time-varying nature of variance and covariance between the two variables and produces a dynamic OHR (or strategy). The last column of the table gives the abbreviations of the models used to generate the corresponding OHRs.

2.1. The Naïve Hedge Ratio

A simple way to hedge the risk of the underlying asset price is to take one unit of a short position of a futures contract for each unit of a long position held in the underlying asset. This is denoted as the naïve hedging strategy with a hedge ratio of 1. This is a model-free strategy and does not involve any optimization or estimation. From a theoretical perspective, one would expect that any optimal hedging strategy should outperform the naïve strategy. However, several empirical studies have documented that this may not be the case (e.g., Alexander and Barbosa 2007).

2.2. Constant Hedge Ratios

2.2.1. OLS. A conventional method extensively applied in the hedging literature is to estimate the linear regression model $r_{s,t} = \alpha_1 + \beta_1 r_{f,t} + \varepsilon_{s,t}$, where $\varepsilon_{s,t} \sim \text{iid N}(0,1)$. The OLS estimate of β_1 provides an estimate of the OHR.

2.2.2. The VAR Model. Given that a large body of literature has documented that asset returns are serially correlated, the VAR model that accounts for the impact of past spot and futures returns on current returns is adopted to estimate the parameters in the OHR. We consider the VAR(1, 1) model specification, which is given by

$$r_{s,t} = \alpha_s + \beta_{11} r_{s,t-1} + \beta_{12} r_{f,t-1} + \varepsilon_{s,t},$$
$$r_{f,t} = \alpha_f + \beta_{21} r_{s,t-1} + \beta_{22} r_{f,t-1} + \varepsilon_{f,t},$$

where $\varepsilon_{s,t}$ and $\varepsilon_{f,t}$ follow a bivariate normal distribution; that is, $\varepsilon_t = (\varepsilon_{s,t}, \varepsilon_{f,t})' \sim \text{N}(0, H)$, with H being the variance and covariance matrix of ε_t and $[h_s, h_{sf} : h_{sf}, h_f]$ being its elements. Because H is assumed to be constant over time, OHR $\equiv h_{sf}/h_f$ is time invariant.

2.2.3. The VEC Model and Its Variations. The VAR model does not consider the possibility that spot and futures prices can be cointegrated in the long run. To account for this long-run relation, the VEC model is adopted (e.g., Ghosh 1993, Kroner and Sultan 1993, Lypny and Powalla 1998, Brooks et al. 2002, Lien and Yang 2006). We consider the VEC(1,1) model of Engle and Granger (1987), which is given by

$$r_{s,t} = \alpha_s + \beta_{11} r_{s,t-1} + \beta_{12} r_{f,t-1} + \lambda_s z_{t-1} + \varepsilon_{s,t},$$
$$r_{f,t} = \alpha_f + \beta_{21} r_{s,t-1} + \beta_{22} r_{f,t-1} + \lambda_f z_{t-1} + \varepsilon_{f,t},$$

where z_t is a linear combination of spot and futures prices, $z_t = s_t - \psi f_t$, which is denoted the error correction term with the cointegrating vector $(1, -\psi)$, and $\varepsilon_{s,t}$ and $\varepsilon_{f,t}$ follow a bivariate normal distribution; $\varepsilon_t = (\varepsilon_{s,t}, \varepsilon_{f,t})' \sim \mathrm{N}(0, H)$. Because many empirical studies have found that z_t can be well approximated by the basis, $z_t \approx s_t - f_t$, we reestimate the VEC model using this approximation and denote the corresponding strategy Adj.VEC. Another variation of the VEC model is the fractionally integrated VEC (FIVEC) model, in which z_t follows a fractionally integrated autoregressive moving average. We use the FIVEC$(1,1)$ model specification of Johansen and Nielsen (2012) in our investigation.[2]

The main problem associated with the above models is the assumption of constant variance and covariance between spot and futures returns. This assumption makes the models incapable of capturing changes in market volatility and can therefore produce biased estimates of the model parameters and, consequently, the hedge ratios (e.g., Baillie and Myers 1991, Park and Switzer 1995, Choudhry 2004), particularly when the market is in a volatile period.

2.3. Time-Varying Hedging Ratios

The variance and covariance of asset returns are time-varying (e.g., Bollerslev et al. 1988, Engle 2002, Cappiello et al. 2006, among many others). Hence, the OHRs determined based on time-varying variance and covariance of spot and futures returns also vary over time. Various versions of the BGARCH model have been used to capture the dynamics of variance and covariance processes of spot and futures returns for estimating time-varying OHRs. The general consensus is that the use of BGARCH models yields superior hedging performance, as evidenced by a lower return variance of the hedged portfolio, compared with time-invariant hedge strategies. Following the hedging literature, we consider the BEKK-GARCH model of Engle and Kroner (1995),[3] the asymmetric BEKK-GARCH (ABEKK) model of Kroner and Ng (1998), the CCC-GARCH model of Bollerslev (1990), the asymmetric CCC-GARCH (ACCC-GARCH) model of McAleer et al. (2009), the DCC-GARCH model of Engle (2002), and the asymmetric DCC-GARCH model of Cappiello et al. (2006). The BEKK-GARCH and ABEKK-GARCH models are employed by Brooks et al. (2002) to investigate the effect of asymmetries on OHRs. The DCC-GARCH model is used by Lien and Yang (2006) to investigate the improvement of hedging performance in the currency futures market. The ADCC-GARCH model is included to capture the asymmetric effect of past return innovations on both the conditional variance and the correlation (e.g., Cappiello et al. 2006, Engle and Colacito 2006). To more flexibly capture the nonlinear dependence structure between spot and futures returns, we include the recently proposed copula-based GARCH model to estimate the time-varying OHR (e.g., Hsu et al. 2008, Lee 2009).

2.3.1. GARCH-Type Models. *BEKK-GARCH and ABEKK-GARCH models.* A popular GARCH-type model for estimating time-varying OHRs is the bivariate BEKK model. We consider the BEKK$(1,1)$ specification in which the conditional mean, $r_t = (r_{s,t}, r_{f,t})'$, is specified as $r_t = \mu_t + \varepsilon_t$, where $\mu_t = (\mu_{s,t}, \mu_{f,t})'$ denotes the means of spot and futures returns, respectively, and $\varepsilon_t = (\varepsilon_{s,t}, \varepsilon_{f,t})' \sim \mathrm{N}(0, H_t)$. The conditional covariance matrix H_t is specified as $H_t = \Omega'\Omega + A'\varepsilon_{t-1}\varepsilon'_{t-1}A + B'H_{t-1}B$, where Ω, A, and B are 2×2 full matrices. The elements in H_t are $h_{s,t}$, $h_{f,t}$, and $h_{sf,t}$, where $h_{s,t}$ and $h_{f,t}$ are the conditional variances of spot and futures returns at time t and $h_{sf,t}$ is the conditional covariance between spot and futures returns at time t. Thus, $OHR_t = h_{sf,t}/h_{f,t}$. To capture the asymmetric volatility effect, the ABEKK model is adopted. The conditional covariance matrix H_t is now modified into $H_t = \Omega'\Omega + A'\varepsilon_{t-1}\varepsilon'_{t-1}A + B'H_{t-1}B + D'\xi_{t-1}\xi_{t-1}D$, where A, B, and D are scalars and $\xi_t = \min(\varepsilon_t, 0)$.

CCC-GARCH and DCC-GARCH Models. Two other popular GARCH-type models for estimating time-varying OHRs are the bivariate CCC-GARCH model of Bollerslev (1990) and the bivariate DCC-GARCH model of Engle (2002). These two models can flexibly describe the conditional correlations between spot and futures returns separately from their volatilities. The conditional mean for both the CCC-GARCH and DCC-GARCH models is specified as the standard univariate GARCH$(1,1)$ specification; that is, $r_{i,t} = \mu_{i,t} + \varepsilon_{i,t} = \mu_{i,t} + h_{i,t}^{1/2}\eta_{i,t}$, where $\eta_{i,t} \sim \text{iid N}(0,1)$ and $h_{i,t} = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{i,t-1}$ for $i = s, f$. Because the variances of spot and futures returns are estimated from the univariate GARCH model, the CCC-GARCH and DCC-GARCH models provide the same estimates of the variances of spot and futures returns. The only difference between the CCC-GARCH and DCC-GARCH

[2] The results were obtained by using the computer software by Nielsen and Morin (2012).

[3] The acronym BEKK comes from synthesized work on multivariate GARCH models by Baba, Engle, Kraft, and Kroner.

models is the specification of the correlation coefficient between spot and futures returns. For the CCC-GARCH model, the correlation is specified as a constant ρ_{sf} (e.g., Bollerslev 1990), whereas the correlation for the DCC-GARCH model, time-varying $\rho_{sf,t}$, is specified as an ARMA(1, 1) process (e.g., Engle 2002, Tse and Tsui 2002). An estimate of OHR from the CCC-GARCH or DCC-GARCH is given by $\rho_{sf}\sqrt{h_{s,t}/h_{f,t}}0$ or $\rho_{sf,t}\sqrt{h_{s,t}/h_{f,t}}0$. The variations of the CCC-GARCH and DCC-GARCH models we consider are the asymmetric CCC-GARCH (ACCC-GARCH) and DCC-GARCH (ADCC-GARCH) models, which account for asymmetric correlation effects. In the specification of these asymmetric models, we model the individual asset volatility using the GJR-GARCH model of Glosten et al. (1993) to account for the asymmetric volatility effects.

Copula-Based GARCH Model. One of the advantages of copula-based GARCH models over the CCC-GARCH or DCC-GARCH models is that it can flexibly model the dependence structure of the two variables through a copula function. We consider a *t*-copula-based GARCH model, in which the *t*-copula function captures both the linear correlation and tail dependence between spot and futures returns. As in the CCC-GARCH or DCC-GARCH models, the OHR can also be calculated by $\rho_{sf}\sqrt{h_{s,t}/h_{f,t}}$, where ρ_{sf} is the dependence parameter implied by the *t*-copula function and $h_{s,t}$ and $h_{f,t}$ can be obtained from their univariate GARCH(1, 1), respectively.

2.3.2. Regime-Switching Models. Since Hamilton (1989) developed an algorithm to estimate a regime-switching (RS) model efficiently through maximum likelihood estimation, RS models have been widely used in characterizing the dynamic behaviors of financial markets during periods of high and low volatility or bull and bear markets (e.g., Jeanne and Masson 2000, Ang and Bekaert 2002, Cerra and Saxena 2005, Guidolin and Timmermann 2008). More recent hedging literature has applied the RS model to determine the OHR (e.g., Alizadeh and Nomikos 2004, Lee et al. 2006, Lee and Yoder 2007).

The RS-OLS Model. The first model we consider is the regime-switching OLS (RS-OLS) model, which is given by $r_{s,t} = \alpha_{s_t} + \beta_{s_t} r_{f,t} + \sigma_{s_t}\varepsilon_t$, where $\varepsilon_t \sim \text{iid N}(0, 1)$ and $s_t = 1$ or $s_t = 0$ represents one of the two regimes[4] that follow a first-order Markov process with a constant transition probability $\Pr(s_t = j \mid s_{t-1} = i) = p_{ij}$, $i, j \in (1, 0)$.[5] Under the RS framework, we obtain regime-dependent OHRs: $\beta_{s_t=1}$ and $\beta_{s_t=0}$. The final OHR used is the weighted OHR in each of the two regimes.

The RS-VAR Model. A single-regime VAR model is extended to an RS-VAR model as

$$r_{s,t} = \alpha_{s,s_t} + \beta_{11} r_{s,t-1} + \beta_{12} r_{f,t-1} + (h_{s,s_t})^{1/2}\varepsilon_{s,t},$$
$$r_{f,t} = \alpha_{f,s_t} + \beta_{21} r_{s,t-1} + \beta_{22} r_{f,t-1} + (h_{f,s_t})^{1/2}\varepsilon_{f,t},$$

where the intercept ($\alpha_{s,s_t=1}$ and $\alpha_{f,s_t=1}$) and the covariance matrix ($H_{s_t=1}$ and $H_{s_t=0}$) are regime dependent. Consequently, the regime-dependent OHRs are given by $\rho_{s_t}\sqrt{h_{s,s_t}/h_{f,s_t}}$, for $s_t = 1, 0$. Again, the weighted OHRs are used as the final OHR.

The RS-VEC Model. Similarly, a single-regime VEC is extended to an RS-VEC model as

$$r_{s,t} = \alpha_{s,s_t} + \beta_{11} r_{s,t-1} + \beta_{12} r_{f,t-1} + (h_{s,s_t})^{1/2}\varepsilon_{s,t},$$
$$r_{f,t} = \alpha_{f,s_t} + \beta_{21} r_{s,t-1} + \beta_{22} r_{f,t-1} + (h_{f,s_t})^{1/2}\varepsilon_{f,t}.$$

The RS-BEKK and RS-DCC Models. We focus on the BEKK-GARCH and DCC-GARCH models to allow the covariance matrix H_t to switch between regimes in the GARCH process. Following Gray (1996), the univariate RS-GARCH can be generalized to the RS-BGARCH framework.

3. Hedging Performance Measure and Its Inferences

To evaluate out-of-sample hedging performance, we first split the entire sample from January 3, 1994 to December 26, 2011 into two subsamples. The first subsample covers the period from January 3, 1994 to August 4, 2003 with a total of 500 weekly observations, and it is used to estimate the model parameters. We denote this period as the estimation window. The second subsample spans from August 11, 2003 to December 26, 2011 with a total of 438 weekly observations, and it is used to evaluate out-of-sample performance. We denote it as the evaluation window. The estimation window is then rolled forward by sequentially dropping the first observation in the window each time and including one new observation from the evaluation window; this is done so that we can maintain the length of the estimation window fixed at 500 observations. We reestimate the model parameters each time we roll the estimation window forward and calculate one-step-ahead weekly forecasts of the OHR for a given model.

Because the objective of hedging is to minimize the variance of the hedged portfolio returns, a straightforward measure of hedging performance is to evaluate the sample variance of the hedged portfolio returns for a given strategy. A better strategy is the one that generates smaller variance and is often inferred based on a simple comparison of the out-of-sample variances

[4] Two regimes are commonly defined in studies of hedging performance (e.g., Alizadeh and Nomikos 2004, Lee et al. 2006, Lee and Yoder 2007).

[5] We use a time-invariant transition probability in our analysis for model parsimony considerations, although a time-varying transition probability is feasible in both the maximum likelihood estimation (Diebold et al. 1994, Filardo 1994) and Bayesian frameworks (Filardo and Gordon 1998).

of the hedged portfolio returns from different strategies. However, the fact that the sample variance of one strategy is smaller than that of the other in a particular sample does not mean that the one is necessarily better than the other in the population. Even if in the population the variances of the hedged portfolio returns from the two strategies are the same, in any particular sample one or the other of the two strategies must be better because the two values are most likely not the same. To draw a statistical inference, we test whether the sample variance of one strategy is significantly different from another using the DM test developed by Diebold and Mariano (1995). In addition, when the data are not sufficiently informative to select a single best strategy, we use a procedure proposed by Hansen et al. (2011) to produce an MCS that contains a smaller set of the best strategies with a given level of confidence and assess whether the naïve strategy performs as well as other strategies.

3.1. The DM Test

To illustrate, we focus on assessing the performance of the naïve strategy relative to that of one of the strategies listed in Table 1. The DM and MCS tests both require the specification of a loss function to conduct test statistics. A natural choice for the purposes of our investigation is the out-of-sample variance of the hedged portfolio returns of a strategy.[6] Therefore, let $g_{n,\tau} = (r_{s,\tau} - r_{f,\tau})^2$ be the loss at time τ associated with the naïve strategy in which $\gamma^*_{\tau-1} = 1$, and let $g_{m,\tau} = (r_{s,\tau} - \gamma^*_{\tau-1,m} r_{f,\tau})^2$ be the loss at time τ associated with the strategy in which $\gamma^*_{\tau-1,m}$ is the one-step-ahead forecast of the OHR at time $\tau - 1$ based on model m. The loss differential at time τ is then given by $d_\tau = g_{m,\tau} - g_{n,\tau}$. The null hypothesis H_0 is that the expected losses for the two strategies are equal; that is, $E(d_\tau) = E(g_{m,\tau} - g_{n,\tau}) = 0$. Diebold and Mariano (1995) show that the test statistic $S_{d_\tau} = E(d_\tau)/\sqrt{\widehat{\text{var}}[E(d_\tau)]}$, where $\widehat{\text{var}}[E(d_\tau)]$ is a consistent estimate of $\text{var}[E(d_\tau)]$,[7] has an asymptotic standard normal distribution under the null hypothesis.

3.2. The MCS Test

Hansen et al. (2011) argue that the available data are occasionally not sufficiently informative to yield a single model that significantly dominates others. Hence, one can only obtain a smaller set of models, an MCS, that contains the best model with a given level of confidence. Therefore, the objective of the MCS procedure is to choose a model set M^* that consists of the best model(s) from the initial model set M^0. The model selection is conducted through a sequence of significance tests in which significantly inferior models are eliminated. In each round of the model selection, the equivalence test, denoted as δ_M, is used to test whether any two models in M^0 perform equally well. If δ_M is rejected, the elimination rule e_M is used to eliminate the model with poor sample performance. This procedure is repeated until δ_M is accepted at a given significance level α. The procedure is then stopped and produces an MCS that contains the models that survive the δ_M tests without being eliminated by e_M with a probability that is greater than $1 - \alpha$. Each time the δ_M test is applied, an associated p-value is recorded.[8] The smallest p-value obtained from applying the δ_M test for comparing model i with each of the alternative models in M is reported as the p-value for model i. As mentioned above, if δ_M is rejected, e_M is used to eliminate the model i. Hansen et al. (2011) introduce and report an MCS p-value associated with model i. The MCS p-value is defined as the largest p-value from the application of the δ_M test in the comparison of model i with the other models that have been eliminated. The interpretation of the MCS p-value is that model i is excluded from the MCS with a probability p.

To briefly illustrate how to construct the test statistics, we let the loss differential between the two models i and j be $d_{ij,\tau} = g_{i,\tau} - g_{j,\tau}\ \forall\, i, j \in M^0$, where $g_{i,\tau}$ and $g_{j,\tau}$ are the losses of models i and j, respectively, at time τ. We define the relative sample loss statistic as $\bar{d}_{ij} \equiv T^{-1}\sum_{\tau=1}^{T} d_{ij,\tau}$ and the sample loss of model i relative to the average across (n) models in M statistic as $\bar{d}_i \equiv n^{-1}\sum_{j\in M}\bar{d}_{ij}$. The t-statistics are constructed as $t_{ij} = \bar{d}_{ij}/\sqrt{\widehat{\text{var}}(\bar{d}_{ij})}$ and $t_i = \bar{d}_i/\sqrt{\widehat{\text{var}}(\bar{d}_i)}$ for $i, j \in M$, where $\widehat{\text{var}}(\bar{d}_{ij})$ and $\widehat{\text{var}}(\bar{d}_i)$ are the estimates of $\text{var}(\bar{d}_{ij})$ and $\text{var}(\bar{d}_i)$, respectively. They are associated with the hull hypothesis that $H_{ij}\colon\ \mu_{ij} = 0$ and $H_i\colon\ \mu_i = 0$, where $\mu_{ij} = E(d_{ij,\tau})$ and $\mu_i = E(\bar{d}_i)$. These t-statistics form the basis of the test of the null hypothesis $H_{0,M}\colon\ \mu_{ij} = 0$ and $\mu_i = 0\ \forall\, i, j \in M^0$; more specifically, $T_{\max,M} = \max_{i\in M} t_i$ and $T_{R,M} = \max_{i\in M} |t_{ij}|$. The asymptotic distributions of $T_{\max,M}$ and $T_{R,M}$ are nonstandard, and hence a block bootstrap procedure is employed.[9] If the null hypothesis is rejected at a significance level of α%, the model with the largest $T_{\max,M}$ is removed. This process is repeated until the nonrejection of the null hypothesis occurs. We then construct a $(1-\alpha)$% confidence set for the best models in M^0.[10]

In the application of the MCS test, our initial model set M^0 contains $n + 1$ models with $n = 18$ listed in

[6] The out-of-sample variance of the hedged portfolio returns is approximately equivalent to the mean squared error (MSE) of the hedged portfolio returns, which is a commonly-used loss function.

[7] See Diebold and Mariano (1995) for a detailed discussion.

[8] The p-values are obtained based on the superior predictive ability test of Hansen (2005).

[9] We perform 10,000 bootstrap iterations.

[10] For the detailed discussion, see the work of Hansen et al. (2011). The application of the MCS procedure can be found in several recent studies (e.g., Laurent et al. 2011).

Table 1, plus the naïve hedging strategy. The initial hypothesis is that all 19 models in M^0 have equal hedging performance. As in the DM test, the loss differential associated with the two different models i and j at time τ is given by $d_{ij,\tau} = g_{i,\tau} - g_{j,\tau}$ $\forall\, i, j = 1, \ldots, n+1$, where $g_{i,\tau} = (r_{i,\tau} - \gamma^*_{\tau-1,i} r_{f,\tau})^2$ and $g_{j,\tau} = (r_{j,\tau} - \gamma^*_{\tau-1,j} r_{f,\tau})^2$, and $d_{ij,\tau}$ measures the relative hedging performance between models i and j at time τ.

Overall, to draw a statistical inference, we conduct two tests. The first is the DM test, which assesses whether there is a significant difference between the sample variances of the hedged portfolio returns of two strategies. The second is the MCS test, which assesses a set of models that contain the best model(s) with a given level of confidence when the data are not sufficiently informative to select a single best model. These two tests help us to evaluate more accurately whether the naïve strategy performs as well as the minimum variance strategies listed in Table 1.

4. Data Description and Preliminary Analysis

We use weekly data to investigate the out-of-sample hedging performance across the 24 futures contracts. Our sample spans the period from January 3, 1994 to December 26, 2011. Table 2 lists the futures contracts, their corresponding underlying assets, and data sources. Monday closing prices are used to calculate weekly returns. If Monday is a holiday, the Friday closing price in the previous week is used. The return is calculated as the log-difference between the current and previous week prices multiplied by 100. Table 3 reports a set of preliminary statistics for the weekly spot and futures returns for each market. Panel A of Table 3 provides descriptive statistics, including the mean, maximum, minimum, standard deviation, skewness, and kurtosis. Several observations are of note. First, the mean and volatility of returns differ between two groups of markets: (1) energy and metal markets and (2) grain, softs, currency, and stock markets. The weekly return has a higher mean and is more volatile for the first group than for the second group. Second, the two groups noted above also differ in terms of the volatility of the futures market relative to that of the spot market. For the first group, spot volatility is always higher than futures volatility, with the exception of the silver market. In contrast, for the second group, futures volatility is higher than spot volatility in 8 of the 13 markets. Crain and Lee (1996) argue that the relatively higher spot market volatility can be attributed partly to the illiquidity of the commodity spot market. Third, the skewness and kurtosis statistics suggest that the return distribution is left skewed and fat tailed in most cases.

Panel B of Table 3 reports the Jarque and Bera (1980) statistics, the Q-statistics of Ljung and Box (1978), and ARCH test statistics of Engle (1982). The Jarque and Bera (1980) statistics again confirm the fat-tailed distributions in financial markets by rejecting the null hypothesis of a Gaussian distribution at the 1% significance level for all markets. The Q-statistics for the autocorrelation of returns and squared returns show that the null hypothesis of nonautocorrelation is rejected for most markets. In particular, the autocorrelation of squared returns is significant at the 1% significance level for all 24 markets. When considered with the mean autocorrelation, which is significant in 13 of the 24 markets, the result implies that the volatility autocorrelation is much stronger than the mean autocorrelation in both the spot and futures markets. Furthermore, the F-statistics of the ARCH test consistently indicate an ARCH effect in the spot and futures markets. The presence of a strong ARCH effect explains why GARCH-class models are widely used in financial time-series analysis.

5. Main Empirical Results

5.1. Does Anything Outperform the Naïve Hedging Strategy?

To address our main issue of the relative performances of different hedging strategies, we compute the out-of-sample variance of the hedged portfolio returns and then conduct the DM test to draw statistical inferences on the performance of the naïve strategy relative to that of each of the strategies listed in Table 1. The results are reported in Table 4. The first and second rows for each market report the hedged portfolio return variance of each strategy and the DM statistics for testing the null hypothesis of equal performance between the naïve strategy and each of the strategies listed in Table 1, respectively. The smallest variance is underlined to indicate the best performance of the hedging strategy based on our sample realizations in each market. A positive (negative) DM statistic indicates that the corresponding sample variance is greater (smaller) than that of the naïvely hedged portfolio returns.

Under the minimum variance hedging framework, the better strategy is that which produces a smaller variance of the hedged portfolio returns. A simple comparison of the variances suggests that no single hedging strategy dominates the others across all the 24 markets examined. For example, among the strategies based on the models under the assumption of constant variance and covariance between spot and futures returns—that is, Models 1–5 listed in Table 1—the FIVEC strategy generates the smallest variance for 4 of the 24 markets. The OLS strategy generates the smallest variance for 2 of the 24 markets, and the Adj.VEC strategy generates the smallest variance for only 1 of

Table 2 Spot and Futures Data

Asset	Spot (data source, symbol)	Futures exchange (data source, symbol)	Contract code	Price quotation
Energies				
Crude oil	West Texas Intermediate (Datastream, CRUDOIL)	CME Group/NYMEX (EIA website, contract1)	CL	U.S. dollars/barrel
Heating oil	New York Harbor No. 2 Heating Oil (Datastream, EIAHONY)	CME Group/NYMEX (EIA website, contract1)	HO	U.S. dollars/gallon
Natural gas	Henry Hub Natural Gas (Datastream, NATGHEN)	CME Group/NYMEX (EIA website, contract1)	NG	U.S. dollars/mmBtu
Gasoline	New York Harbor RBOB Gasoline (EIA website, RBOB)	CME Group/NYMEX (EIA website, contract1)	RB	U.S. dollars/gallon
Metals				
Aluminum	High Grade Primary Aluminum (Datastream, LAHCASH)	London Metal Exchange (Datastream, LAH3MTH)	AH	U.S. dollars/tonne
Copper	Grade A Copper (Datastream, LCPCASH)	London Metal Exchange (Datastream, LCP3MTH)	CA	U.S. dollars/tonne
Lead	Lead of 99.97% Purity (Datastream, LEDCASH)	London Metal Exchange (Datastream, LED3MTH)	PB	U.S. dollars/tonne
Tin	Tin of 99.85% Purity (Datastream, LTICASH)	London Metal Exchange (Datastream, LTI3MTH)	SN	U.S. dollars/tonne
Zinc	Zinc of 99.995% Purity (Datastream, LZZCASH)	London Metal Exchange (Datastream, LZZ3MTH)	ZS	U.S. dollars/tonne
Gold	Handy and Harman Base NY (Datastream, GOLDHAR)	CME Group/NYMEX/COMEX (Datastream, NGCCS00)	GC	U.S. dollars/troy ounce
Silver	Handy and Harman Base NY (Datastream, SILVERH)	CME Group/NYMEX/COMEX (Datastream, NSLCS00)	SL	U.S. dollars/troy ounce
Grains				
Corn	No. 2 Yellow Corn (Datastream, CORNUS2)	CME Group/CBOT (Datastream, CC.CS00)	C	U.S. cents/bushel
Soybeans	No. 1 Yellow Soybean (Datastream, SOYBEAN)	CME Group/CBOT (Datastream, CS.CS00)	S	U.S. cents/bushel
Soybean oil	Soya Oil Crude Decatur (Datastream, SOYAOIL)	CME Group/CBOT (Datastream, CB.CS00)	BO	U.S. cents/pound
Wheat	No. 2 Soft Red (Datastream, WHEATSF)	CME Group/CBOT (Datastream, CW.CS00)	W	U.S. cents/bushel
Oats	No. 2 Milling Minneapolis (Datastream, OATSMP2)	CME Group/CBOT (Datastream, CO.CS00)	O	U.S. cents/bushel
Softs				
Cotton	11/16STR Low-Middle Memphis (Datastream, COTTONM)	ICE/NYBOT/NYMEX (Datastream, NCTCS00)	CT	U.S. cents/pound
Sugar	Raw Sugar (Datastream, WSUGDLY)	ICE/NYBOT/NYMEX (Datastream, NSBCS00)	SB	U.S. cents/pound
Currencies				
British pound (GBP)	GBP/USD (Bloomberg, GBP)	CME Group (Bloomberg, BP1)	BP	U.S. dollars/GBP
Canadian dollar (CAD)	CAD/USD (Bloomberg, CAD)	CME Group (Bloomberg, CD1)	CD	U.S. dollars/CAD
Japanese yen (JPY)	JPY/USD (Bloomberg, JPY)	CME Group (Bloomberg, JY1)	JY	U.S. dollars/JPY
Stock indices				
FTSE 100	London Stock Exchange (Bloomberg, UKX)	NYSE Euronext (Bloomberg, Z1)	Z	British pounds
NIKKEI 225	Tokyo Stock Exchange (Bloomberg, NKY)	CME Group (Bloomberg, NK1)	NIY	U.S. dollars
S&P 500	New York Stock Exchange (Bloomberg, SPX)	CME Group (Bloomberg, SP1)	SP	U.S. dollars

Notes. The 18 commodity futures, 3 currency futures, and 3 stock index futures used in our analysis, for which the prices of futures and their corresponding spots are collected from Bloomberg, Datastream, and the EIA website, are listed in the table. The futures contracts are traded on the Chicago Board of Trade (CBOT), the Chicago Mercantile Exchange (CME), the Intercontinental Exchange (ICE), and the New York Mercantile Exchange (NYMEX). Note that CME recently took over NYMEX, COMEX, and CBOT and was renamed the CME Group.

the 24 markets. Among the strategies based on time-varying variance and covariance—that is, Models 6–18—the regime-switching strategies perform better than the GARCH-type strategies. However, for some markets, the regime-switching regression strategies RS-OLS, RS-VAR, and RS-VEC generate higher variance than the linear regression strategies OLS, VAR, and VEC. Among the GARCH-type strategies, the BEKK performs the best but produces the smallest variance for only 2 of the 24 markets. More interestingly, the

Table 3 Summary Statistics of Spot and Futures Returns

		Panel A: Descriptive statistics						Panel B: Test statistics			
		Mean	Maximum	Minimum	Std. dev.	Skewness	Kurtosis	JB test	LB $Q(10)$	LB $Q^2(10)$	ARCH test
Crude oil	Spot	0.205	27.33	−36.08	5.545	−0.573	8.025	1,038.0***	49.46***	317.2***	21.92***
	Futures	0.205	23.38	−26.92	5.305	−0.481	6.093	410.07***	36.58***	147.1***	10.44***
Heating oil	Spot	0.194	36.17	−29.14	4.590	0.078	10.190	2,021.8***	70.37***	230.5***	21.91***
	Futures	0.194	20.61	−18.87	4.085	−0.128	4.248	63.41***	47.79***	99.73***	6.93***
Natural gas	Spot	0.039	43.95	−35.52	8.496	0.328	5.292	222.12***	3.80	205.9***	13.29***
	Futures	0.051	44.57	−36.86	7.727	0.271	5.188	198.59***	11.09	31.9***	2.67***
Gasoline	Spot	0.198	37.36	−22.49	5.108	0.017	6.576	499.7***	49.12***	81.06***	6.70***
	Futures	0.195	21.08	−23.27	4.642	−0.363	5.198	209.38***	43.03***	80.28***	5.26***
Aluminum	Spot	0.063	10.00	−13.66	2.890	−0.170	3.989	42.75***	12.51	61.39***	4.04***
	Futures	0.062	9.445	−13.20	2.764	−0.210	4.163	59.77***	12.26	81.75***	5.25***
Copper	Spot	0.156	12.05	−19.81	3.663	−0.574	5.160	233.87***	23.31*	200.0***	11.80***
	Futures	0.155	12.01	−16.05	3.491	−0.471	4.819	164.01***	28.68***	296.2***	15.60***
Lead	Spot	0.153	19.97	−20.13	4.650	−0.097	5.126	178.11***	11.91	212.2***	12.49***
	Futures	0.151	19.78	−20.37	4.349	−0.197	5.932	342.16***	11.48	276.6***	15.32***
Tin	Spot	0.147	18.65	−22.25	3.584	−0.492	8.036	1,028.8***	18.85**	138.4***	8.25***
	Futures	0.146	18.96	−22.20	3.509	−0.501	8.531	1,234.8***	22.60**	139.4***	8.44***
Zinc	Spot	0.064	14.89	−21.98	4.045	−0.461	5.948	372.93***	7.39	109.8***	6.23***
	Futures	0.063	14.58	−20.56	3.850	−0.384	5.439	255.43***	7.92	145.1***	7.48***
Gold	Spot	0.151	13.72	−11.57	2.370	0.255	6.674	537.74***	18.99**	245.3***	15.02***
	Futures	0.150	14.30	−10.93	2.316	0.292	6.961	626.46***	18.39**	170.2***	9.94***
Silver	Spot	0.184	18.92	−28.13	4.161	−0.695	7.625	911.34***	21.24**	111.1***	7.10***
	Futures	0.183	19.01	−26.78	4.193	−0.688	7.167	752.73***	20.47**	92.19***	5.71***
Corn	Spot	0.079	16.00	−21.83	4.213	−0.405	5.240	221.79***	11.05	93.04***	6.13***
	Futures	0.075	16.44	−19.05	4.005	−0.031	4.595	99.56***	13.58	61.71***	5.40***
Soybeans	Spot	0.053	13.14	−21.49	3.669	−0.544	5.656	321.96***	15.40	118.4***	6.88***
	Futures	0.053	12.12	−17.65	3.532	−0.407	5.114	200.48***	11.11	128.7***	7.88***
Soybean oil	Spot	0.056	13.12	−18.57	3.638	−0.068	4.297	66.47***	8.11	198.6***	10.69***
	Futures	0.058	14.58	−17.03	3.513	−0.062	4.229	59.68***	5.69	154.8***	8.48***
Wheat	Spot	0.048	17.23	−16.09	4.086	0.234	4.512	97.91***	9.46	59.27***	4.57***
	Futures	0.055	19.85	−15.76	4.213	0.377	4.051	65.47***	8.66	27.13***	2.30**
Oats	Spot	0.079	24.15	−24.95	4.774	0.005	5.953	340.75***	26.76***	34.4***	3.56***
	Futures	0.088	24.15	−19.72	4.822	−0.093	4.719	116.86***	21.97**	39.03***	3.14***
Cotton	Spot	0.036	14.14	−17.39	4.126	−0.057	3.999	39.21***	8.87	162.6***	9.49***
	Futures	0.027	15.76	−36.92	4.487	−0.935	11.110	2,685.0***	17.02*	30.26***	2.75***
Sugar	Spot	0.083	18.30	−19.51	4.491	−0.217	4.580	104.8***	17.56*	29.6***	2.38***
	Futures	0.044	14.93	−23.72	4.774	−0.410	4.782	150.6***	15.22	41.41***	3.01***
GBP	Spot	0.006	4.349	−9.793	1.250	−0.605	7.099	713.9***	7.24	163.1***	10.11***
	Futures	0.006	4.389	−9.005	1.248	−0.515	6.159	431.3***	7.32	171.9***	11.23***
CAD	Spot	0.027	8.276	−7.813	1.162	−0.109	8.942	1,382.0***	16.62*	432.4***	32.25***
	Futures	0.027	8.159	−7.685	1.162	−0.095	8.568	1,213.0***	15.82	426.4***	30.73***
JPY	Spot	0.039	13.28	−5.886	1.642	0.918	8.749	1,424.0***	10.57	61.49***	5.01***
	Futures	0.039	14.35	−5.864	1.649	1.027	10.290	2,240.0***	11.28	50.67***	4.29***
FTSE 100	Spot	0.051	14.26	−14.14	2.598	−0.419	6.517	510.8***	25.94***	167.9***	8.86***
	Futures	0.050	14.46	−14.19	2.628	−0.354	6.391	469.0***	24.23***	184.2***	9.70***
NIKKEI 225	Spot	−0.076	18.02	−23.54	3.356	−0.515	8.517	1,231.0***	16.76*	341.9***	29.52***
	Futures	−0.077	16.57	−26.47	3.426	−0.706	9.704	1,834.0***	19.02**	407.1***	37.25***
S&P 500	Spot	0.107	12.95	−14.91	2.688	−0.370	6.922	622.5***	31.32***	376.8***	20.16***
	Futures	0.106	14.97	−17.10	2.745	−0.410	7.769	915.0***	37.84***	305.5***	20.07***

Notes. This table reports summary statistics of weekly spot and futures returns (in percent). The JB test refers to the Jacque and Bera (1980) normality test. The null hypothesis under the JB test is a joint hypothesis of both skewness and excess kurtosis being zero. LB $Q(10)$ and LB $Q^2(10)$ refer to the Ljung and Box (1978) Q-test for serial correlation of returns and squared returns, respectively. The null hypothesis is that there is no autocorrelation of series up to the 10th orders. The ARCH test refers to the ARCH test of Engle (1982) on squared return series. The F-statistics of the ARCH test are for the null hypothesis of no ARCH effects up to the 10th orders. The sample period covers January 1994 to December 2011.

*, **, and *** denote rejections of the null hypothesis at the 10%, 5%, and 1% significance levels, respectively.

naïve strategy generates a smaller variance than any of the other strategies considered for 5 of the 24 markets, particularly for financial markets.

In addition to the relative performance of different strategies, several interesting patterns can be observed cross-sectionally. First, ARCH effects are significant in each of the 24 markets, as shown in Table 3, but, for nearly all markets, the GARCH-type strategies produce higher variance than the linear, regime-switching regression, and naïve strategies do. This result seems

Table 4 Out-of-Sample Hedging Performance of the Naïve Strategy Relative to That of One of Minimum Variance Strategies Based on the DM Test

	Un-hedged	Naïve	OLS	VAR	VEC	Adj. VEC	FIVEC	BEKK	ABEKK	CCC	ACCC	DCC	ADCC	GARCH-copula	GJR-copula	MRS-OLS	MRS-VAR	MRS-VEC	MRS-BEKK	MRS-DCC
Crude oil	34.04	3.629	3.630	3.625	3.622	3.622	*3.598*	4.160*	4.035	5.503*	5.262	4.633*	7.761*	5.550	5.321	3.618	3.618	3.624	4.503	4.196
			(0.053)	(−0.207)	(−0.590)	(−0.636)	(−1.224)	(1.738)	(1.350)	(1.665)	(1.392)	(1.768)	(1.795)	(1.587)	(1.522)	(−0.392)	(−0.358)	(−0.687)	(1.260)	(1.195)
Heating oil	17.06	0.467	0.495**	0.476	0.469	0.481	0.468	0.524***	0.520***	0.525***	0.618***	0.535***	1.004***	0.543***	0.599***	0.477**	0.470	*0.467*	0.601***	0.532***
			(2.185)	(1.269)	(1.178)	(1.606)	(0.241)	(3.660)	(2.662)	(2.410)	(3.429)	(3.261)	(6.974)	(3.279)	(3.852)	(2.373)	(0.670)	(−0.124)	(4.213)	(2.885)
Natural gas	81.00	48.77	*47.05*	47.47	47.36	47.36	47.61	51.10	53.48	56.24	53.91	55.29	54.78	55.25	54.96	47.98	47.25	47.49	52.75	52.98
			(−0.629)	(−0.728)	(−0.819)	(−0.815)	(−0.972)	(0.671)	(1.030)	(1.143)	(1.137)	(1.157)	(1.181)	(1.154)	(1.182)	(−0.391)	(−0.797)	(−0.723)	(1.018)	(1.031)
Gasoline	27.81	4.277	4.300	4.301	4.295	4.289	4.183	4.191	4.540	4.684	4.555	4.865	4.779*	4.796	4.792	4.232	4.205	4.198	4.286	*4.175*
			(0.646)	(0.630)	(0.490)	(0.323)	(−1.619)	(−0.496)	(1.237)	(1.101)	(0.837)	(1.306)	(1.692)	(1.324)	(1.515)	(−0.422)	(−0.583)	(−0.670)	(0.038)	(−0.478)
Aluminum	11.05	0.150	0.150	0.150	0.149	*0.149*	0.151	0.152	0.163	0.169***	0.177***	0.170**	0.259***	0.216***	0.173**	0.149	0.155	0.156	0.165	0.159
			(−0.261)	(−0.263)	(−0.300)	(−0.311)	(0.171)	(0.460)	(1.381)	(2.607)	(3.418)	(2.534)	(3.279)	(3.572)	(2.436)	(−0.177)	(0.726)	(0.975)	(1.084)	(0.838)
Copper	19.16	0.208	*0.199*	0.202	0.203	0.203	0.201	0.211	0.209	0.230	0.213	0.236	0.214	0.228	0.221	0.201	0.201	0.201	0.216	0.206
			(−0.702)	(−0.654)	(−0.596)	(−0.570)	(−1.011)	(0.226)	(0.061)	(1.330)	(0.359)	(1.400)	(0.548)	(1.075)	(0.700)	(−0.667)	(−0.766)	(−0.727)	(0.529)	(−0.183)
Lead	35.30	0.744	0.776	0.782	0.796	0.786	0.785	*0.731*	0.795	0.916*	0.805**	0.928*	0.953**	0.777	0.802	0.792	0.781	0.805	0.766	0.767
			(0.776)	(0.892)	(1.143)	(0.974)	(0.953)	(−0.417)	(1.087)	(1.810)	(2.017)	(1.841)	(2.367)	(1.240)	(1.630)	(1.070)	(0.867)	(1.293)	(0.711)	(0.568)
Tin	21.00	0.369	0.375**	0.374**	0.374**	0.375**	0.371	0.385	0.378	0.472**	0.454**	0.399	0.446**	0.441	0.469**	*0.359*	0.382	0.372	0.416	0.420
			(2.097)	(2.248)	(2.262)	(2.012)	(1.147)	(0.508)	(0.443)	(2.246)	(2.098)	(0.851)	(2.201)	(1.377)	(2.155)	(−0.309)	(1.275)	(0.538)	(1.138)	(1.234)
Zinc	26.28	0.139	0.153	0.154	0.148	0.149	0.149	*0.132*	0.153	0.174***	0.152*	0.146	0.182***	0.145	0.150	0.141	0.157	0.157	0.150	0.141
			(1.414)	(1.489)	(0.924)	(1.031)	(1.027)	(−1.117)	(1.186)	(4.126)	(1.766)	(1.272)	(5.277)	(1.163)	(1.215)	(0.186)	(1.521)	(1.526)	(0.930)	(0.202)
Gold	8.100	0.971	0.968	0.970	0.968	0.968	*0.965*	1.012*	1.033*	1.049***	1.055**	1.042***	1.069***	1.040***	1.054***	1.033	0.970	0.966	1.013*	1.017**
			(−0.188)	(−0.036)	(−0.145)	(−0.148)	(−0.319)	(1.884)	(1.742)	(3.289)	(2.295)	(3.001)	(2.826)	(3.105)	(3.159)	(1.010)	(−0.010)	(−0.275)	(1.783)	(2.007)
Silver	27.22	2.138	2.153	2.136	2.126	2.123	*2.109*	2.256	2.458*	2.353**	2.244	2.219	2.271**	2.226	2.282**	2.149	2.129	2.126	2.343	2.288
			(0.321)	(−0.024)	(−0.263)	(−0.335)	(−0.638)	(1.299)	(1.712)	(2.233)	(1.480)	(1.045)	(2.281)	(1.212)	(2.265)	(0.217)	(−0.175)	(−0.261)	(1.432)	(1.320)
Corn	23.53	2.317	2.392	2.379	2.372	2.383	2.346	2.357	2.449	2.386	2.398	2.446	2.559	2.413	2.585	2.353	2.360	2.391	*2.295*	2.388
			(1.081)	(0.961)	(0.915)	(0.968)	(0.471)	(0.472)	(1.408)	(0.889)	(0.979)	(1.434)	(1.125)	(1.455)	(1.304)	(0.595)	(0.877)	(1.645)	(−0.239)	(0.839)
Soybeans	19.23	*2.727*	2.761	2.762	2.769	2.771	2.731	4.240**	3.662	3.100	3.052	3.926**	3.598**	3.170*	3.177	2.936	2.852	2.861	3.394	3.366
			(0.438)	(0.450)	(0.537)	(0.539)	(0.053)	(2.550)	(1.277)	(1.504)	(1.486)	(2.574)	(2.324)	(1.732)	(2.152)	(1.296)	(1.196)	(1.206)	(1.249)	(1.262)
Soybean oil	17.50	0.789	0.789	0.789	0.790	0.790	*0.783*	0.784	0.792	0.809	0.823	0.805	0.815	0.805	0.814	0.786	0.791	0.789	0.788	0.801
			(0.118)	(0.015)	(0.161)	(0.239)	(−1.127)	(−0.226)	(0.161)	(0.901)	(1.442)	(0.649)	(1.052)	(0.697)	(1.099)	(−0.396)	(0.412)	(−0.016)	(−0.028)	(0.554)
Wheat	21.76	6.975	6.977	6.971	6.966	6.968	6.916**	7.081*	7.341**	7.349**	7.124*	7.154*	7.279*	7.153*	7.284**	6.999	6.927*	*6.908***	7.225*	7.339**
			(1.051)	(−1.094)	(−1.116)	(−1.104)	(−2.082)	(1.793)	(2.022)	(2.177)	(1.805)	(1.806)	(1.945)	(1.748)	(1.980)	(1.462)	(−1.873)	(−2.290)	(1.778)	(1.969)
Oats	24.13	4.444	5.252**	5.209**	5.019*	5.003*	4.929	4.661	4.414	5.184**	5.178*	4.520	4.640	*4.407*	4.508	4.589	4.508	4.480	4.621	4.471
			(2.360)	(2.283)	(1.828)	(1.770)	(1.592)	(0.619)	(−0.114)	(1.928)	(1.858)	(0.223)	(0.230)	(−0.141)	(0.244)	(0.542)	(0.245)	(0.141)	(0.647)	(0.111)
Cotton	23.09	6.085	5.655	5.656	5.625	5.632	5.367	5.221	5.158	5.676	5.690	5.251	5.263	5.775	5.803	*4.821*	5.369	4.965	5.335	5.097
			(−0.774)	(−0.776)	(−0.974)	(−0.932)	(−1.059)	(−0.742)	(−0.753)	(−0.494)	(−0.498)	(−0.630)	(−0.643)	(−0.468)	(−0.445)	(−0.967)	(−1.300)	(−1.675)*	(−0.701)	(−0.763)
Sugar	19.40	10.04	6.799**	6.812**	6.781**	6.771**	6.427**	5.144***	5.864***	5.452***	5.568**	5.365***	5.411***	5.471***	5.503***	5.065***	4.831***	*4.812****	5.333***	5.492***
			(−2.325)	(−2.340)	(2.381)	(2.378)	(−2.205)	(−2.729)	(−2.938)	(−2.621)	(−2.433)	(−2.907)	(−2.867)	(−3.007)	(−3.067)	(−3.579)	(−3.909)	(−3.815)	(−3.037)	(−2.921)
GBP	2.092	*0.074*	0.076	0.076	0.076	0.076	0.075	0.085	0.083	0.096	0.097	0.095	0.109	0.096	0.097	0.076	0.077	0.075	0.084	0.085
			(0.952)	(0.882)	(0.899)	(0.917)	(0.573)	(1.075)	(1.568)	(1.180)	(1.098)	(1.152)	(1.647)	(1.160)	(1.080)	(0.877)	(0.885)	(0.262)	(0.702)	(1.079)
CAD	2.159	*0.006*	0.006	0.006	0.006	0.006	0.006	0.007	0.007**	0.007	0.007*	0.007	0.008	0.007	0.007*	0.007	0.008	0.007*	0.008	0.007
			(0.948)	(1.208)	(1.226)	(0.829)	(0.061)	(1.449)	(2.497)	(1.363)	(1.822)	(1.621)	(1.615)	(1.460)	(1.844)	(1.349)	(1.388)	(1.661)	(1.538)	(1.423)
JPY	2.292	*0.009*	0.011***	0.010***	0.010**	0.010***	0.010**	0.009	0.009	0.012*	0.011	0.012	0.011*	0.017***	0.016***	0.011***	0.010***	0.010***	0.009	0.009
			(3.536)	(2.649)	(2.452)	(2.622)	(2.503)	(0.730)	(1.178)	(1.725)	(1.626)	(1.661)	(1.738)	(3.294)	(3.432)	(3.667)	(3.148)	(3.315)	(0.280)	(0.056)
FTSE	7.043	*0.071*	0.072*	0.072	0.072	0.072	0.072	0.076	0.078	0.076	0.075	0.076	0.080***	0.076	0.075	0.073*	0.072	0.072	0.073	0.073
			(1.768)	(1.589)	(1.545)	(1.864)	(0.612)	(1.485)	(1.622)	(1.493)	(1.030)	(1.508)	(2.961)	(1.491)	(1.025)	(1.742)	(1.227)	(1.241)	(0.888)	(0.604)
NIKKEI	12.731	0.177	*0.163*	0.164	0.168	0.169	0.163*	0.206	0.188	0.194	0.187	0.194	0.216	0.194	0.212	0.182	0.171	0.170	0.208	0.175
			(−1.242)	(−1.595)	(−1.589)	(−1.569)	(−1.753)	(0.644)	(0.495)	(0.435)	(0.262)	(0.445)	(1.065)	(0.432)	(0.845)	(0.214)	(−0.656)	(−0.885)	(0.639)	(−0.080)
S&P 500	7.985	0.100	0.097	0.096	0.096	0.097	0.094	0.103	0.116	0.109	0.107	0.109	0.107	0.100	0.107	*0.092*	0.096	0.095	0.105	0.100
			(−0.740)	(−0.856)	(−0.992)	(−0.925)	(−1.428)	(0.299)	(1.530)	(0.859)	(0.676)	(0.847)	(0.657)	(0.916)	(0.675)	(−1.063)	(−0.665)	(−0.808)	(0.721)	(0.045)

Notes. This table reports the variances of the hedged portfolio returns of the strategies and the DM test statistics. The italicized value in a row is the smallest variance across that row. The values reported in parentheses are the DM statistics to test for the null hypothesis of equal hedging performance between the naïve strategy and each of 18 minimum variance strategies.

*, **, and *** denote rejections of the null hypothesis at the 10%, 5%, and 1% significance levels, respectively.

puzzling, given that GARCH-type models are developed for capturing the ARCH effect in the data. One of the potential explanations could be that the hedging effectiveness of the GARCH-type strategies is significantly affected by estimation error when the sample size is small. The GARCH-type models have more parameters to estimate and hence higher estimation errors than other models. Model misspecification could also be a potential factor that makes the GARCH-type models perform worse than other models. However, it is unlikely that the GARCH-type models deviate more than others from the true model for all 24 markets. Therefore, this observed pattern is more likely due to estimation errors.

The second observation is that a relative small change in model specification can cause a good model to stop working well. For example, in the oats market, the GARCH-copula model generates smaller variance compared with the naïve strategy. However, a change from the GARCH-copula to the GJR-copula produces a higher variance. In the gasoline market, the model specifications of the RS-DCC and the RS-BEKK are very similar, but these models generate lower and higher variances, respectively, compared with the naïve hedging strategy. In the zinc market, a change from the BEKK to the ABEKK also results in different hedging outcomes relative to the naïve hedging strategy. In the sugar market, the VAR model performs better than the naïve strategy, whereas the VEC performs worse than the naïve strategy. In each of these markets, the two models have relatively similar specifications, but one performs better and the other does worse than the naïve strategy. This pattern suggests that the hedging outcomes of more sophisticated models are sensitive to model misspecification.

The third observation is that time-series properties of a market cause some model strategies to work better than others. For example, supply and demand shocks often cause the returns in the oil and its related markets to switch from one regime to the other. Our results show that the regime-switching regression models generate smaller variance than the nonswitching ones, including the naïve strategy, in the crude oil, heating oil, and gasoline markets, suggesting that the strategy works better when the model is closer to the true data-generating process (DGP). Hence, both the estimation error and model misspecification could play significant roles in the relative hedging performances of the strategies.

As discussed in §3, to draw statistical inferences, we apply the DM test, for which the null hypothesis is that the naïve strategy performs as well as each of the 18 strategies considered. The results show that the null hypothesis cannot be rejected for all 18 strategies in the S&P 500 index futures market and for 17 of the 18 strategies in the NIKKEI futures market. In the FTSE futures market, the naïve strategy can significantly beat 3 of the 18 strategies and performs as well as the other 15 strategies. Overall, the results in the stock index market suggest that the naïve strategy performs as well as almost all other strategies. This finding is consistent with the findings of DeMiguel et al. (2009), who suggest that model-driven portfolio strategies are not consistently better than the $1/N$ portfolio strategy. In addition to the stock market, consistent results are found in the currency market for British pounds and Canadian dollars and in the energy market for natural gas and gasoline. In the agricultural commodity market, except for wheat, cotton, and sugar, the results again suggest that the naïve hedging strategy performs statistically as well as other strategies. In summary, based on a simple comparison of the sample variance and the DM test, we find little evidence that any of the minimum variance hedging strategies considered can consistently and significantly outperform the naïve hedging strategy.

The second test that we use to assess hedging performance of the naïve strategy relative to that of other alternative strategies is the MCS test, as discussed in §3.2. Recall that the test procedure involves finding an MCS that contains the best model(s). Table 5 reports the MCS p-value of the test obtained from 10,000 bootstrap iterations with a block length of 2. The out-of-sample variances of the hedged portfolio returns are the same as those reported in Table 4, and we thus do not report them in this table. The MCS p-value determines whether the model is included in the MCS with a given significance level. For instance, in the crude oil market, the MCS p-value reported under the naïve strategy is 0.664, which suggests that the naïve strategy performs as well as other strategies in the MCS and hence should be included in the MCS. Note that we use one, two, and three asterisks in the table to indicate that the strategies are excluded from the MCS at the significance levels of 10%, 5%, and 1%, respectively.

Given a significance level of 10%, we find that the out-of-sample hedging performance results of different strategies vary sharply among the markets. For example, for the commodities such as crude oil, natural gas, tin, gasoline, corn, copper, and soybean oil, and five of the six financial markets, the MCS contains almost all of the 18 models. However, for the commodities such as lead and zinc, the MCS excludes most of the models. Furthermore, the models that generate dynamic hedging ratios are excluded from the MCS for gold, silver, and JPY, whereas the models that generate constant hedging ratios are excluded for oats. Interestingly, we find that the naïve strategy is included in the MCS for 23 of the 24 markets. Overall, the MCS test results again indicate that it is difficult to find a strategy under the minimum variance framework that consistently outperforms the naïve hedging strategy.

Table 5 Out-of-Sample Hedging Performance of the Naïve Strategy Relative to That of One of Minimum Variance Strategies Based on the MCS Test

	Naïve	OLS	VAR	VEC	Adj. VEC	FIVEC	BEKK	ABEKK	CCC	ACCC	DCC	ADCC	GARCH-copula	GJR-copula	MRS-OLS	MRS-VAR	MRS-VEC	MRS-BEKK	MRS-DCC
Crude oil	0.664	0.664	0.664	0.664	0.664	1	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.728	0.728	0.728	0.26	0.26
Heating oil	0.953	0.039**	0.546	0.896	0.142	0.953	0.002***	0.002***	0.006***	0.001***	0.002***	0***	0.002***	0***	0.357	0.896	1	0***	0.002***
Natural gas	0.324	1	0.324	0.324	0.324	0.324	0.324	0.324	0.324	0.324	0.324	0.324	0.324	0.324	0.324	0.487	0.324	0.324	0.324
Gasoline	0.954	0.58	0.447	0.738	0.878	0.998	0.998	0.375	0.429	0.438	0.369	0.132	0.261	0.187	0.997	0.998	0.998	0.985	1
Aluminum	0.751	0.751	0.751	0.751	1	0.751	0.751	0.026**	0.007***	0.008***	0.002***	0***	0***	0.001***	0.751	0.165	0.15	0.049**	0.199
Copper	0.121	1	0.469	0.278	0.194	0.681	0.121	0.121	0.054*	0.121	0.063*	0.121	0.088*	0.121	0.771	0.771	0.771	0.121	0.278
Lead	0.681	0.071*	0.011**	0.004***	0.008***	0.008***	1	0.008***	0.001***	0.003***	0.001***	0***	0.214	0.002***	0.008***	0.021**	0.002***	0.391	0.404
Tin	0.931	0.848	0.922	0.932	0.932	0.932	0.785	0.932	0.241	0.146	0.684	0.209	0.413	0.296	1	0.739	0.932	0.561	0.512
Zinc	0.23	0***	0***	0***	0***	0***	1	0***	0***	0***	0***	0***	0.001***	0***	0.127	0***	0***	0***	0.008***
Gold	0.775	0.775	0.683	0.775	0.775	1	0.059*	0.091*	0.005***	0.038**	0.011**	0.016**	0.022**	0.08*	0.214	0.653	0.775	0.149	0.045**
Silver	0.128	0.025**	0.054*	0.128	0.128	1	0.025**	0.022**	0.014**	0.025**	0.025**	0.022**	0.025**	0.022**	0.025**	0.128	0.128	0.025**	0.025**
Corn	0.805	0.381	0.498	0.503	0.404	0.782	0.654	0.157	0.503	0.404	0.109	0.172	0.319	0.087*	0.654	0.654	0.381	1	0.381
Soybean	1	0.834	0.616	0.263	0.389	0.956	0.098*	0.154	0.155	0.155	0.066*	0.085*	0.154	0.154	0.155	0.155	0.155	0.155	0.155
Soybean oil	0.846	0.734	0.821	0.639	0.575	1	0.952	0.821	0.573	0.523	0.575	0.567	0.575	0.558	0.952	0.575	0.887	0.952	0.575
Wheat	0.11	0.11	0.24	0.145	0.717	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.448	0.483	1	0.071	0.11
Oats	0.992	0.013**	0.022**	0.098*	0.359	0.562	0.774	0.992	0.039**	0.151	0.924	0.953	1	0.978	0.89	0.968	0.992	0.842	0.992
Cotton	0.158	0.051*	0.102	0.436	0.43	0.436	0.436	0.558	0.026**	0.026**	0.436	0.436	0.027**	0.033**	1	0.436	0.721	0.436	0.713
Sugar	0.012**	0.168	0.251	0.003***	0***	0.381	0.361	0.522	0.059*	0.007***	0.147	0.076*	0.115	0.039**	1	0.931	0.676	0.837	0.931
GBP	1	0.345	0.345	0.345	0.345	0.527	0.345	0.345	0.345	0.345	0.345	0.323	0.345	0.345	0.345	0.345	0.345	0.345	0.345
CAD	1	0.387	0.199	0.267	0.683	0.94	0.199	0.199	0.199	0.199	0.199	0.199	0.199	0.199	0.199	0.199	0.199	0.17	0.199
JPY	1	0.013**	0.049**	0.54	0.154	0.348	0.959	0.801	0.018**	0.075*	0.018*	0.045**	0.002***	0.003***	0.007***	0.032**	0.045**	0.905	0.959
FTSE 100	1	0.399	0.517	0.492	0.42	0.617	0.35	0.294	0.289	0.377	0.282	0.112	0.291	0.377	0.377	0.552	0.517	0.399	0.617
NIKKEI 225	0.4	1	0.963	0.56	0.4	0.975	0.4	0.4	0.4	0.4	0.4	0.321	0.4	0.366	0.4	0.4	0.4	0.4	0.4
S&P 500	0.131	0.272	0.272	0.632	0.272	0.704	0.131	0.117	0.117	0.131	0.117	0.131	0.131	0.131	1	0.272	0.632	0.117	0.272

Notes. This table reports the MCS p-values obtained from 10,000 bootstraps with a block length of 2. Note that the variances of the hedged portfolio returns of the strategies are the same as those reported in Table 4 and, hence, are not included in this table.

*, **, and *** indicate that the model (or strategy) is excluded from the MCS at the 10%, 5%, and 1% significance levels, respectively.

5.2. The 2008–2009 Financial Crisis

During the recent financial crisis in 2008–2009, asset prices, particularly in the commodity and stock markets, experienced large fluctuations because of shocks to global economic and financial activities. The crisis provides an opportunity to examine the hedging performance of different strategies during the period when markets are extremely volatile. We conduct the MCS test based on the sample period from January 1, 2008 to December 31, 2009. The model parameters are estimated by using 500 weekly observations immediately before January 1, 2008. It is interesting to note that a comparison of the sample variances of the hedged portfolio return with that of the unhedged portfolio reveals that the former in any of the markets is much smaller than the latter, which suggests that the use of futures contracts could have greatly reduced the uncertainty of the underlying asset's price movement during the recent financial crisis. Consistent with the MCS test results over the whole sample period, the results for the recent financial crisis period show that the naïve strategy is included in the MCS for each of the 24 markets.[11] However, the more sophisticated model strategies are excluded for several markets. For example, for the cotton, lead, and wheat markets, the regression models that produce constant hedge ratios are excluded from the MCS. For the heating oil, aluminum, copper, lead, wheat, sugar, and JPY markets, the GARCH-type models are excluded from the MCS. The regime-switching models are outperformed by other models for the lead, tin, and wheat markets. The strategies based on copula models are not good choices for the aluminum, copper, and JPY markets. The above evidence suggests that none of the minimum variance hedging strategies could consistently outperform the naïve hedging strategy during the recent financial crisis, and it confirms the robustness of the naïve hedging strategy.

5.3. The Effect of the Length of the Estimation Window

We also examine the out-of-sample hedging performance using the rolling window method with a fixed size of 500 observations. The choice of the estimation window size has always been a concern for practitioners because the use of different window sizes can lead to different empirical results. For a robustness check, we adjust the estimation window size from 500 to 750 and reexamine the out-of-sample hedging performance. The results reveal that the naïve strategy performs best for 6 of the 24 markets and is included in the MCS for each of the 24 markets, which is consistent with the results obtained by using an estimation window size of

[11] To save space, the results are not included in the paper but are available upon request.

500. Hence, the adjustment of the estimation window size does not change our main empirical finding; that is, it is difficult to find a strategy that consistently outperforms the naïve strategy.

5.4. The Effect of the Hedging Horizon

It has been found that the out-of-sample OHRs depend on the hedging horizon, that is, the time interval for measuring price changes (e.g., Chen et al. 2004). Several studies have shown that the in-sample hedge effectiveness tends to increase as the hedging horizon increases (e.g., Benet 1992, Geppert 1995). Therefore, the out-of-sample hedging performance could also be sensitive to the choice of hedging horizon. For a robustness check, we use the biweekly nonoverlapped returns to repeat our comparison analysis of the relative hedging performance of the naïve strategy and each of the alternative strategies. We find that, although the variance of the hedged portfolio returns of a particular strategy is different from that of the strategy based on a horizon of one week, the overall results are consistent. Specifically, the naïve strategy produces the smallest variances for 5 of the 24 markets and is included in the MCS for 23 of the 24 markets, suggesting that the adjustment of the hedging horizon does not change our main empirical findings on the performance of the naïve strategy relative to that of the minimum variance hedging strategies.

5.5. Model Combination Strategies

It has been shown in the forecasting literature that the performance of a single model may be unstable and a set of combined models may perform better than a single model over time (e.g., Stock and Watson 2003). For a robustness check of our main findings, we construct a new set of strategies based on the 18 models listed in Table 1 using the method of model combination (e.g., Stock and Watson 2004). We consider four combination strategies that are commonly utilized. The first combination strategy is the mean combination, which is the equally weighted average of the OHRs based on the 18 different models. The second combination strategy is the trimmed mean combination—that is, excluding the OHR with the worst past performance before calculating the average of the OHRs. The third combination strategy involves the use of the most recent best performance of the OHRs. The fourth combination strategy is the weighted average of the individual strategies, in which the weights depend inversely on the historical performance of each individual strategy. The MCS results indicate that the naïve strategy is still included in the MCS for all 24 markets, suggesting that the model combination strategies do not substantially improve the hedging performance of the minimum variance hedging strategies relative to that of the naïve strategy.[12]

5.6. Why Is It Difficult to Beat the Naïve Hedging Strategy?

Thus far, we have demonstrated that the naïve strategy cannot be outperformed consistently and significantly by any of the 18 minimum variance hedging strategies across the 24 markets considered. In this section, we investigate two possible explanations: estimation error and model misspecification. Estimation error is known to affect the out-of-sample effectiveness of asset allocation (e.g., Bawa et al. 1979, Jorion 1992, Clarkson et al. 1996, Britten-Jones 1999, Siegel and Woodgate 2007, DeMiguel et al. 2009). Because hedging with futures is a special case of asset allocation, one would expect that estimation error might also affect the out-of-sample hedging performance.

In addition to estimation error, our empirical findings discussed in the previous section suggest that model misspecification could play a crucial role in the hedging performance of the minimum variance strategies relative to that of the naïve strategy. Recall that we find that a small change in the model specification results in a large difference in the hedging outcomes. The naïve hedging strategy is robust to changes in market conditions, as shown during the recent financial crisis.

5.6.1. The Role of Estimation Error. To investigate the effect of estimation error on hedging performance, we compare the results when the parameters in the model are assumed to be ex ante known to hedgers during the out-of-sample period with those when the parameters in the model are assumed to be unknown. The difference between the two sets of the results indicates the effect of estimation error.[13] The investigation of relative hedging performance when the parameters are assumed to be ex ante known is equivalent to an in-sample analysis, whereas, when the parameters are assumed to be unknown, the investigation is equivalent to an out-of-sample analysis. The key difference between an out-of-sample analysis and an in-sample analysis is that in the former, the hedger uses all of the information available at time t to estimate the model parameters and makes one-step forecasts of the variance and covariance based on the estimates of the model parameters. Each time a new observation becomes available, the hedger incorporates it and reestimates the model to update the estimates of the model

[12] To save space, the results are not included in the paper but are available upon request.

[13] Note that our comparison between in-sample and out-of-sample hedging performance is based on the same model. If the model is misspecified, the misspecification errors should equally affect both in-sample and out-of-sample performance. Hence, our comparison results are robust to model misspecification.

parameters. In an in-sample analysis, for the same model, the hedger estimates the model parameters only once using the whole out-of-sample observations and treats the estimates as the true values of these parameters.

The results from the out-of-sample analysis are reported in Table 4. We now conduct the in-sample analysis. We first estimate the model parameters and then treat these estimates as their true values. With these values, particularly the estimates of the variance and covariance, we calculate the OHRs and, moreover, the hedged portfolio returns. We then repeat our earlier comparison analysis, using the MCS test on the performance of the naïve strategy relative to that of the minimum variance hedging strategies conducted in §5.1. The results are reported in Table 6. The variances of the hedged portfolios for most strategies become smaller when the model parameters are assumed to be known, particularly for the regression model strategies and the regime-switching model strategies. In particular, for each of the 24 markets, the variance of the hedged portfolio returns of the OLS strategy is smaller than that of the naïve strategy when the model parameters are assumed to be known, in sharp contrast to the results obtained when the parameters are unknown, in which the variance of the OLS strategy is higher than that of the naïve strategy. Moreover, the MCS results show that the naïve strategy is included in the MCS for 17 of the 24 markets when the parameters are known but is included in the MCS for 23 of the 24 markets when the parameters are unknown. The results suggest that if the model parameters are ex ante known, the hedging performance of the minimum variance strategies that depend on the estimation of the model parameters can be greatly improved and outperform the naïve strategy in 7 of the 24 markets. Hence, our main finding that it is difficult to find a strategy that significantly beats the model-free naïve strategy can be partly explained by errors in the estimates of the model parameters.

In practice, the true model is unknown. For a given model, the parameters in the model need to be estimated. Estimation errors, in the form of bias and variability, can be checked by plotting the parameter estimates in the hedging exercise using recursive windows.[14] To focus only on estimation error and control for model misspecification, we use simulated data. The advantage of using simulated data is that we know the true model specification and true values of its parameters. Thus, we can assess the role of estimation error independently from model misspecification. We select five markets, one from each asset group: crude oil (energy), gold (metal), corn (agricultural commodity), the British pound (currency), and the S&P 500 index (stock). To generate the data, we first estimate the model parameters using the entire sample observations assuming that the true data-generating process is described by either the BEKK or RS-OLS model for each of the five pairs of spot and futures markets we have chosen. Given the values of these parameters in the model, we use Monte Carlo sampling to generate weekly returns of $T = 5{,}000$ observations and then estimate the model parameters using recursive windows. We start with the first 300 observations and then expand the windows by adding 10 more observations each time and reestimating the model. The estimation error of a parameter estimate is measured by the MSE scaled by the squared true value of the parameter. Figure 1 displays the estimation errors of all the parameters in each of the two models and for each of the five selected markets.

For a smaller estimation window, the parameter estimates are more volatile and deviate more greatly from their true values, suggesting a larger variation of the parameter estimates and implying that the data are too noisy to identify the true value of the parameters. With the expansion of the estimation window, the deviations become smaller, implying that the parameter estimates are closer to their true values. However, it should be noted that the parameter estimates approach their true values only for very long estimation windows. For example, for crude oil, the estimates of the BEKK and the RS-OLS model parameters converge to their true value for a window length longer than 4,000 weeks (i.e., 80 years). The window length of convergence for the BEKK model parameter estimates varies from 2,500 weeks (S&P 500) to more than 4,000 weeks (corn). The window length of convergence for the RS-OLS model parameter estimates across all five selected markets is greater than 3,000 weeks. It is impossible to obtain such long sample data in the real world. The more complicated models require more data to estimate the model parameters correctly. Therefore, there is always a small sample issue for estimating model parameters in practice. As shown in §5.3, when the estimation window increases from 500 to 750 observations, the hedging performance improves, and the GARCH-type model strategies improve the most. However, the naïve strategy still performs as well as the other strategies, suggesting that 750 observations may still not be long enough to obtain a more accurate parameter estimate, which is consistent with our finding here based on the simulation study. Hence, our main finding that it is difficult to find a strategy that significantly beats the naïve strategy can be partly explained by estimation error in the form of bias in the parameter estimates caused by small sample size.

5.6.2. Model Misspecification. Another potential explanation for the unbeatable hedging performance of the naïve strategy may result from model

[14] We thank the referee for the constructive suggestion.

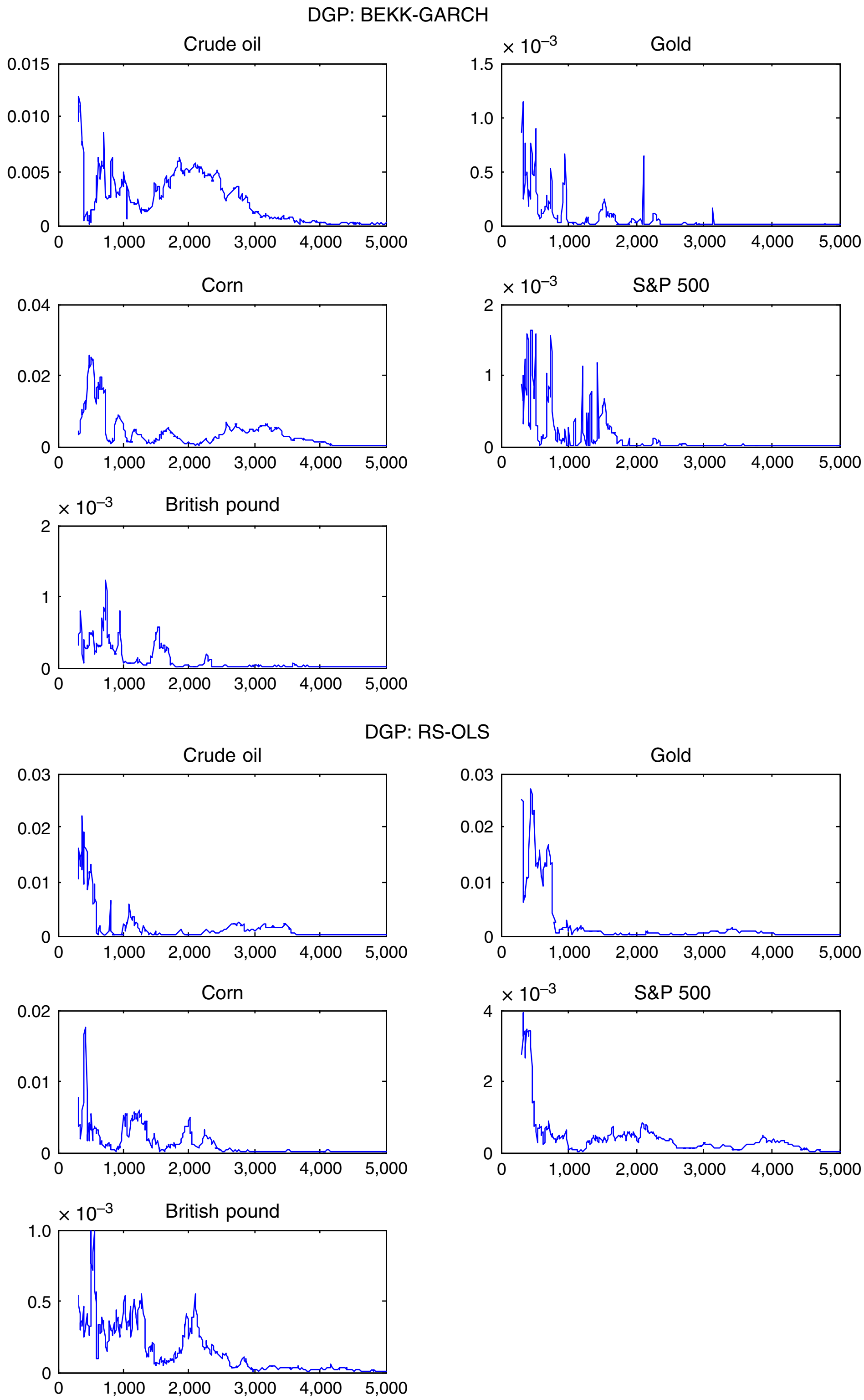
Table 6 **MCS Test Results Under the Assumption That the Parameters in the Model Are Ex Ante Known to Hedgers**

	Un-hedged	Naïve	OLS	VAR	VEC	Adj. VEC	FIVEC	BEKK	ABEKK	CCC	ACCC	DCC	ADCC	GARCH-copula	GJR-copula	MRS-OLS	MRS-VAR	MRS-VEC	MRS-BEKK	MRS-DCC
Crude	34.04	3.629	3.602	3.602	3.605	3.607	3.610	3.944	4.974	4.722	4.931	4.160	4.315	5.287	3.646	3.606	3.619	*3.541*	4.310	4.199
oil		(0.859)	(0.859)	(0.859)	(0.859)	(0.859)	(0.859)	(0.337)	(0.256)	(0.256)	(0.256)	(0.256)	(0.256)	(0.256)	(0.803)	(0.859)	(0.859)	(1)	(0.337)	(0.256)
Heating	17.06	0.467	0.467	0.467	0.467	0.467	0.467	0.492**	0.496**	0.494**	0.486*	0.492**	0.484	0.495**	0.469	*0.462*	0.467	0.467	0.500**	0.499**
oil		(0.324)	(0.324)	(0.324)	(0.324)	(0.318)	(0.312)	(0.026)	(0.027)	(0.029)	(0.080)	(0.046)	(0.122)	(0.027)	(0.324)	(1)	(0.324)	(0.324)	(0.025)	(0.026)
Natural	81.00	48.77	46.69	46.69	46.69	46.70	46.70	45.61	53.48	53.39	53.07	53.16	52.83	54.66	46.73	46.59	46.95	46.75	*45.37*	53.71
gas		(0.443)	(0.443)	(0.443)	(0.443)	(0.443)	(0.443)	(0.911)	(0.112)	(0.112)	(0.112)	(0.112)	(0.112)	(0.112)	(0.443)	(0.443)	(0.443)	(0.443)	(1)	(0.112)
Gasoline	27.81	4.277	4.236	4.236	4.240	4.240	4.241	4.802	4.043	4.255	4.195	4.349	4.235	4.336	4.251	*3.236*	4.209	4.132	4.089	4.155
		(0.308)	(0.308)	(0.308)	(0.308)	(0.308)	(0.308)	(0.293)	(0.312)	(0.308)	(0.312)	(0.301)	(0.308)	(0.302)	(0.308)	(1)	(0.312)	(0.329)	(0.312)	(0.312)
Aluminum	11.05	0.150**	0.143	*0.142*	0.143	0.143	0.143	0.150**	0.157**	0.151**	0.145**	0.153**	0.147**	0.151**	0.150**	0.143	0.146**	0.143	0.156**	0.157**
		(0.023)	(0.992)	(1)	(0.992)	(0.946)	(0.992)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.946)	(0.023)	(0.992)	(0.023)	(0.023)
Copper	19.16	0.208	0.203	0.203	0.203	0.203	0.203	*0.196*	0.202	0.197	0.202	0.197	0.203	0.197	0.213	0.197	0.203	0.205	0.202	0.202
		(0.446)	(0.481)	(0.449)	(0.632)	(0.451)	(0.480)	(1)	(0.480)	(0.931)	(0.451)	(0.931)	(0.449)	(0.931)	(0.416)	(0.931)	(0.769)	(0.449)	(0.682)	(0.532)
Lead	35.30	0.744	0.720	0.720	0.720	0.720	0.720	0.717	0.718	0.726	0.727	0.720	0.718	0.725	0.763	*0.697*	0.726	0.712	0.727	0.723
		(0.418)	(0.646)	(0.646)	(0.646)	(0.645)	(0.645)	(0.646)	(0.646)	(0.620)	(0.628)	(0.646)	(0.646)	(0.628)	(0.199)	(1)	(0.618)	(0.646)	(0.605)	(0.628)
Tin	21.00	0.369	0.369	0.369	0.369	0.369	0.369	0.377	0.383	0.390	0.392	0.373	0.371	0.389	0.371	0.370	*0.369*	0.473	0.388	0.382
		(0.996)	(0.996)	(0.996)	(0.986)	(0.996)	(0.981)	(0.679)	(0.826)	(0.468)	(0.341)	(0.934)	(0.981)	(0.598)	(0.934)	(0.956)	(1)	(0.188)	(0.695)	(0.678)
Zinc	26.28	0.139	0.126	0.126	0.126	0.126	0.126	0.129	0.131	0.129	0.127	0.129	0.127	0.129	0.142*	0.127	0.126	*0.126*	0.132	0.132
		(0.275)	(0.952)	(0.952)	(0.952)	(0.952)	(0.952)	(0.747)	(0.678)	(0.762)	(0.952)	(0.870)	(0.952)	(0.806)	(0.085)	(0.952)	(0.952)	(1)	(0.561)	(0.607)
Gold	8.100	0.971	0.964	*0.964*	0.969	0.969	0.969	1.031**	1.022**	1.022**	1.023**	1.030**	1.027**	1.022**	0.973	0.965	0.978	0.964	1.029**	1.025**
		(0.186)	(0.995)	(1)	(0.106)	(0.106)	(0.874)	(0.011)	(0.016)	(0.011)	(0.013)	(0.013)	(0.016)	(0.011)	(0.106)	(0.974)	(0.106)	(0.982)	(0.016)	(0.016)
Silver	27.22	2.138	2.109	2.109	2.110	2.110	2.110	2.146	2.193	2.198	2.233	2.154	2.183	2.198	2.110	*2.108*	2.117	2.110	2.110	2.193
		(0.856)	(0.987)	(0.991)	(0.980)	(0.980)	(0.982)	(0.946)	(0.686)	(0.482)	(0.388)	(0.807)	(0.639)	(0.565)	(0.980)	(1)	(0.976)	(0.979)	(0.991)	(0.732)
Corn	23.53	2.317*	2.316*	2.316*	2.317*	2.316*	2.317*	2.288*	2.284*	2.363*	2.600**	2.373**	2.345*	2.366*	2.351**	2.274*	*2.123*	2.322*	2.313*	2.318*
		(0.067)	(0.065)	(0.064)	(0.067)	(0.066)	(0.067)	(0.067)	(0.067)	(0.067)	(0.030)	(0.049)	(0.056)	(0.059)	(0.044)	(0.067)	(1)	(0.067)	(0.067)	(0.067)
Soybeans	19.23	2.727	2.721	2.721	2.723	2.723	2.721	3.147	3.221	2.964	2.958	3.074	3.191	3.007	2.724	*2.715*	2.795	2.716	3.219	3.228
		(0.925)	(0.988)	(0.967)	(0.915)	(0.925)	(0.948)	(0.202)	(0.202)	(0.202)	(0.188)	(0.202)	(0.130)	(0.152)	(0.913)	(1)	(0.202)	(0.991)	(0.152)	(0.202)
Soybean	17.50	0.789	0.786	0.786	0.786	0.786	0.786	0.784	0.790	0.784	0.782	0.781	*0.780*	0.783	0.803	0.795	0.786	0.786	0.788	0.792
oil		(0.882)	(0.901)	(0.901)	(0.901)	(0.901)	(0.901)	(0.901)	(0.857)	(0.901)	(0.901)	(0.901)	(1)	(0.901)	(0.818)	(0.842)	(0.901)	(0.901)	(0.882)	(0.842)
Wheat	21.76	6.975**	6.862**	6.862**	6.862**	6.862**	6.863**	6.663**	6.949**	6.908**	6.808**	6.871**	6.761**	6.911**	6.874**	*5.792*	6.940**	6.774**	6.966**	6.976**
		(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(1)	(0.023)	(0.023)	(0.023)	(0.023)
Oats	24.13	4.444***	4.216***	4.216***	4.217***	4.217***	4.216***	4.210***	4.263***	4.422***	4.336***	4.392***	4.262***	4.492***	4.248***	*3.917*	4.148***	3.940	4.298***	4.308***
		(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(1)	(0.006)	(0.350)	(0.006)	(0.006)
Cotton	23.09	6.085***	5.331***	5.332***	5.361**	5.349***	5.364**	5.279***	5.101**	5.411***	5.199***	4.963**	4.680*	5.552***	5.696**	*3.670*	4.011	4.349	4.978**	4.931**
		(0.003)	(0.002)	(0.002)	(0.018)	(0.002)	(0.022)	(0.002)	(0.018)	(0.002)	(0.002)	(0.022)	(0.063)	(0.002)	(0.034)	(1)	(0.238)	(0.238)	(0.022)	(0.022)
Sugar	19.40	10.04**	6.449**	6.449**	6.453**	6.452***	6.454**	5.392**	5.867**	5.505**	5.325**	5.407**	5.345**	5.479**	5.375**	*4.501*	5.963**	5.463**	5.370**	6.036**
		(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.026)	(0.017)	(0.017)	(0.017)	(0.017)	(1)	(0.017)	(0.017)	(0.017)	(0.017)
GBP	2.092	0.074	0.074	0.074	0.074	0.074	0.074	0.085	0.086	0.086	0.086	0.083	0.082	0.086	0.075	*0.057*	0.076	0.076	0.835	0.085
		(0.156)	(0.156)	(0.156)	(0.156)	(0.156)	(0.156)	(0.155)	(0.155)	(0.155)	(0.156)	(0.155)	(0.156)	(0.156)	(0.156)	(1)	(0.156)	(0.156)	(0.156)	(0.155)
CAD	2.159	0.0063	*0.0062*	0.0062	0.0062	0.0062	0.0062	0.0068	0.0071	0.0068	0.0068	0.0069	0.0069	0.0068	0.0063	0.0066	0.0065	0.0064	0.0068	0.0068
		(0.481)	(1)	(0.957)	(0.481)	(0.936)	(0.936)	(0.158)	(0.158)	(0.158)	(0.158)	(0.158)	(0.158)	(0.158)	(0.472)	(0.158)	(0.158)	(0.481)	(0.158)	(0.158)
JPY	2.292	0.0091	0.0091	0.0091	0.0091	0.0091	0.0091	0.0091	*0.0089*	0.0092	0.0091	0.0092	0.0091	0.0092	0.0092	0.0091	0.0091	0.0091	0.0091	0.0091
		(0.958)	(0.959)	(0.959)	(0.959)	(0.959)	(0.959)	(0.959)	(1)	(0.958)	(0.959)	(0.958)	(0.959)	(0.958)	(0.958)	(0.959)	(0.959)	(0.959)	(0.959)	(0.959)
FTSE 100	7.043	0.071***	0.071***	0.071***	0.071***	0.071***	0.071***	0.074***	0.071***	0.074***	0.073***	0.074***	0.073***	0.074***	0.071***	*0.054*	0.070***	0.071***	0.072***	0.072***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(1)	(0.002)	(0.002)	(0.002)	(0.002)
NIKKEI	12.731	0.177	0.158	0.158	0.160	0.160	0.159	0.176	0.158	0.166	0.146	0.166	0.146	0.165	0.172	*0.141*	0.147	0.144	0.171	0.175
225	(0.260)	(0.461)	(0.462)	(0.462)	(0.462)	(0.462)	(0.228)	(0.324)	(0.261)	(0.462)	(0.260)	(0.462)	(0.295)	(0.324)	(1)	(0.462)	(0.462)	(0.324)	(0.206)	
S&P	7.985	0.100	0.094	0.094	0.094	0.094	0.095	0.105	0.100	0.107	0.105	0.107	0.105	0.107	0.097	*0.092*	0.092	0.096	0.103	0.102
500		(0.417)	(0.598)	(0.598)	(0.598)	(0.598)	(0.598)	(0.260)	(0.417)	(0.198)	(0.303)	(0.217)	(0.237)	(0.184)	(0.591)	(1)	(0.955)	(0.598)	(0.319)	(0.319)

Notes. This table reports the variance of the hedged portfolio returns of the strategies and the MCS p-values. The italicized value in a row is the smallest variance across that row. The values reported in parentheses are the MCS p-values obtained from 10,000 bootstraps with a block length of 2.

*, **, and *** indicate that the model (or strategy) is excluded from the MCS at the significant levels of 10%, 5%, and 1%, respectively.

Figure 1 (Color online) The Mean Squared Error Scaled by the Squared True Value of the Parameter Using Simulated Data Based on Recursive Windows



Note. The prespecified parameters of groups 1–5 are from the estimates of actual spot and futures data of crude oil, gold, corn, the S&P 500 index, and the British pound, respectively.

misspecification. To investigate this possibility, we use the simulated data generated in the previous section and repeat our earlier comparison analysis of the naïve strategy relative to that of the minimum variance hedging strategies based on the in-sample hedging performance. For the in-sample case, we use all of the observations to estimate the parameters once to minimize the estimation errors.

Table 7 In-Sample Hedging Performance of the Naïve Strategy Relative to One of Four Selected Strategies Using Simulated Data

	Naïve	BEKK	ABEKK	RS-OLS	RS-VAR
			DGP: BEKK-GARCH		
Crude oil	2.058***	*1.819*	1.896***	2.023***	1.882***
	(0)	(1)	(0)	(0)	(0)
Gold	0.103**	*0.097*	0.097	0.101*	0.098
	(0.024)	(1)	(0.786)	(0.059)	(0.117)
Corn	0.163***	*0.124*	0.135***	0.128*	0.127*
	(0)	(1)	(0.002)	(0.077)	(0.077)
GBP	0.114***	*0.058*	0.060	0.096***	0.085***
	(0)	(1)	(0.181)	(0)	(0)
S&P 500	0.062***	*0.054*	0.057**	0.056*	0.055
	(0)	(1)	(0.045)	(0.088)	(0.367)
			DGP: RS-OLS		
Crude oil	3.145	3.150	3.179***	*3.145*	3.145
	(0.852)	(0.626)	(0)	(1)	(0.913)
Gold	0.520***	0.520***	0.547***	*0.513*	0.513
	(0)	(0)	(0)	(1)	(0.783)
Corn	1.821***	1.815***	1.855***	*1.770*	1.771
	(0.004)	(0.004)	(0.004)	(1)	(0.041)
GBP	0.1047***	0.1047***	0.1054***	*0.1034*	0.1046***
	(0)	(0)	(0)	(1)	(0)
S&P 500	0.0540***	0.0541***	0.0549***	*0.054*	0.054
	(0.004)	(0.004)	(0.004)	(1)	(0.482)

Notes. This table reports the variance of the hedged portfolio returns of the strategies using the data simulated based on either the BEKK-GARCH or RS-OLS models. The italicized value in a row is the smallest variance across that row. For each model, the prespecified parameters are from the estimates of actual spot and futures data of crude oil, gold, corn, the British pound, and the S&P 500 index, respectively. The numbers in parentheses are the MCS *p*-values.

*, **, and *** indicate that the model (or strategy) is excluded from the MCS at the significant levels of 10%, 5%, and 1%, respectively.

Table 7 reports the variances of the hedged portfolio returns and the MCS test results for the selected model strategies. The hedged portfolio variances of the true model strategies are always smallest. In addition, when the true DGP is the BEKK model, the variance of the ABEKK model is smaller than that of the other misspecified models, possibly because the ABEKK model is closest to the true model specification. The naïve strategy and regime-switching model strategies are both misspecified, and the former generates higher variance. This may be because the true DGP of the BEKK model assumes time-varying variance and covariance between spot and futures returns and, hence, a time-varying hedge ratio, whereas the hedge ratio from the naïve strategy is constant. Although the RS-OLS model is also misspecified, it allows the hedge ratio to vary from one regime to the other, which is more flexible than the constant hedge ratio of one given by the naïve strategy. The MCS test results indicate that the true model strategy performs the best and that nearly all other model strategies are excluded from the MCS.

When the true DGP is the RS-OLS model, the RS-VAR model strategy produces smaller variance than other misspecified models. Again, this may be because it is the closest to the true model specification. The naïve strategy, BEKK, and ABEKK are all misspecified. But the variances of the hedged portfolio returns suggest that the naïve strategy performs no worse than the BEKK or ABEKK model strategies. This may result from the fact that the hedge ratios from the BEKK and ABEKK models are fluctuated more relative to the one from the RS-OLS model. The constant hedge ratio from the naïve strategy is closer to the hedge ratio from the true model. Hence, when the naïve strategy and more complicated model strategies are both far from the true DGP, the naïve strategy performs at least as well as other misspecified model strategies.

We also conduct the above analysis using actual data. We take the BEKK model strategy as a typical case and compare its hedging performance with the naïve strategy. Similar to our simulation analysis, we choose one market from each of the five asset groups: crude oil (energy), gold (metal), corn (agricultural commodity), the S&P 500 (stock index), and the British pound (currency). We find that the OHRs of the BEKK model strategy for these five markets are quite unstable. Table 4 indicates that the naïve strategy generates a smaller variance of the hedged portfolio returns than the BEKK model does for these five markets, confirming that the naïve strategy works better for the markets where the hedge ratios are unstable. Overall, the results based on both simulated and actual data suggest that model misspecification can partly explain why it is very difficult to significantly beat the naïve strategy in practice.

6. Conclusion

Hedging with futures is an important issue in financial risk management. Although there have been a considerable number of studies on hedging spot price risk by using futures contracts, there is still no consensus on the optimal hedging strategy. This paper revisits this issue by investigating the performance of a simple naïve hedging strategy and 18 minimum variance hedging strategies. Our findings, based on the two statistical tests, indicate that it is difficult to find a minimum variance strategy that performs consistently better than the simple naïve strategy before the transaction costs. If the costs of transactions are considered, the strategies with time-varying hedge ratios, which require frequently rebalancing hedging positions, may be even worse. As a robustness check, we also investigate the out-of-sample hedging performance of different strategies during the recent financial crisis and find the consistent performance of the naïve strategy. We vary the length of the estimation window and the hedging horizon, and we find that the results still support our main finding—that is, the naïve strategy performs as well as the other strategies we evaluate. To explain why it is difficult to find a strategy that consistently outperforms the naïve hedging strategy, we consider two possible explanations, estimation error and model misspecification, and observe that our main findings can be partially explained by both estimation error and model misspecification.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2014.2028.

Acknowledgments

The authors thank Itay Goldstein (the department editor), an anonymous associate editor, and two anonymous referees for their constructive and insightful comments and suggestions. The authors thank Andrew Patton, Kevin Sheppard, Morten Nielsen, Peter Hansen, and Asger Lunde for kindly sharing their computing codes. Financial support from the National Science Foundation of China [71320107002] and Australian Research Council [DP0988579] is gratefully acknowledged. All errors are the responsibility of the authors.

References

Alexander C, Barbosa A (2007) Effectiveness of minimum variance hedging. *J. Portfolio Management* 33:46–59.
Alizadeh AH, Nomikos N (2004) A Markov regime switching approach for hedging stock indices. *J. Futures Markets* 24:649–674.
Alizadeh AH, Nomikos NK, Pouliasis PK (2008) A Markov regime switching approach for hedging energy commodities. *J. Banking Finance* 32:1970–1983.
Andrew A, Chen J (2002) Asymmetric correlation of equity portfolios. *J. Financial Econom.* 63:443–494.
Ang A, Bekaert G (2002) International asset allocation with regime shifts. *Rev. Financial Stud.* 15:1137–1187.
Baillie RT, Myers R (1991) Bivariate garch estimation of the optimal commodity futures hedge. *J. Appl. Econometrics* 6:109–124.
Bawa VS, Brown SJ, Klein RW (1979) *Estimation Risk and Optimal Portfolio Choice* (North-Holland, Amsterdam).
Benet BA (1992) Hedge period length and ex-ante futures hedging effectiveness: The case of foreign-exchange risk across hedges. *J. Futures Markets* 12:163–175.
Bollerslev T (1990) Modelling the coherence in short-run nominal exchange rate: A multivariate generalized arch approach. *Rev. Econom. Statist.* 72:498–505.
Bollerslev T, Engle RF, Wooldridge JM (1988) A capital asset model with time-varying covariances. *J. Political Econom.* 96:116–131.
Britten-Jones M (1999) The sampling error in estimates of mean-variance efficient portfolio weights. *J. Finance* 54:655–671.
Brooks C, Henry O, Persand G (2002) The effect of asymmetries on optimal hedge ratios. *J. Bus.* 75:333–352.
Cappiello L, Engle RF, Sheppard K (2006) Asymmetric dynamics in the correlations of global equity and bond returns. *J. Financial Econom.* 4:537–572.
Cerra V, Saxena SC (2005) Did output recover from the Asian crisis? *IMF Staff Papers* 52:1–23.
Chen SS, Lee CF, Shrestha K (2004) An empirical analysis of the relationship between the hedge ratio and hedge horizon: A simultaneous estimation of the short- and long-run hedge ratios. *J. Futures Markets* 24:359–386.
Choudhry M (2004) *Structured Credit Products: Credit Derivatives and Synthetic Securitisation* (John Wiley & Sons, Singapore).
Clarkson P, Guedes J, Thompson R (1996) On the diversification, observability and measurement of estimation risk. *J. Financial Quant. Anal.* 31:69–84.
Crain SJ, Lee JH (1996) Volatility in wheat spot and futures markets, 1950–1993: Government farm programs, seasonality, and causality. *J. Finance* 51:325–343.
DeMiguel V, Garlappi L, Uppal R (2009) Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *Rev. Financial Stud.* 22:1915–1953.
Diebold FX, Mariano RS (1995) Comparing predictive accuracy. *J. Bus. Econom. Statist.* 13:253–263.
Diebold FX, Lee JH, Weinbach G (1994) Regime-switching with time-varying transition probabilities. Hargreaves CP, ed. *Non-Stationary Time Series Analysis and Cointegration* (Oxford University Press, Oxford, UK), 283–302.
Ederington LH (1979) The hedging performance of the new futures markets. *J. Finance* 34:157–170.
Engle RF (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50:987–1007.
Engle RF (2002) Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *J. Bus. Econom. Statist.* 20:339–350.
Engle RF, Colacito R (2006) Testing and valuing dynamic correlations for asset allocation. *J. Bus. Econom. Statist.* 24:238–253.
Engle RF, Granger CWJ (1987) Co-integration and error correction: Representation, estimation, and testing. *Econometrica* 55:251–276.
Engle RF, Kroner KF (1995) Multivariate simultaneous generalized arch. *Econometric Theory* 11:122–150.
Filardo AJ (1994) Business cycle phases and their transitional dynamics. *J. Bus. Econom. Statist.* 12:299–308.
Filardo AJ, Gordon SF (1998) Business cycle durations. *J. Econometrics* 85:99–123.
Geppert JM (1995) A statistical model for the relationship between futures contract hedging effectiveness and investment horizon length. *J. Futures Markets* 15:507–536.
Ghosh A (1993) Hedging with stock index futures: Estimation and forecasting with error correction model. *J. Futures Markets* 13:743–752.
Glosten LR, Jagannathan R, Runkle DE (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. *J. Finance* 48:1779–1801.
Gray SF (1996) Modeling the conditional distribution of interest rates as a regime-switching process. *J. Financial Econom.* 42:27–62.
Green RC, Hollifield B (2002) When will mean-variance efficient portfolios be well diversified? *J. Finance* 47:1785–1809.
Guidolin M, Timmermann A (2008) International asset allocation under regime switching, skew, and kurtosis preferences. *Rev. Financial Stud.* 21:889–935.
Haigh MS, Holt MT (2002) Crack spread hedging: Accounting for time-varying volatility spillovers in the energy futures markets. *J. Appl. Econometrics* 17:269–289.
Hamilton JD (1989) A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57:357–384.
Hansen PR (2005) A test for superior predictive ability. *J. Bus. Econom. Statist.* 23:365–380.
Hansen PR, Lunde A, Nason JM (2011) The model confidence set. *Econometrica* 79:453–497.
Hickey L (2011) To hedge or not to hedge. *Oil Gas Financial J.* (June 1), http://www.ogfj.com/articles/print/volume-8/issue-6/features/to-hedge-or-not-to-hedge.html.
Hsu CC, Tseng CP, Wang YH (2008) Dynamic hedging with futures: A copula-based garch model. *J. Futures Markets* 28:1095–1116.
Jagannathan R, Ma T (2003) Risk reduction in large portfolios: Why imposing the wrong constraints helps. *J. Finance* 58:1651–1684.
Jarque CM, Bera AK (1980) Efficient test of normality, homoscedasticity and serial independence of regression residuals. *Econom. Lett.* 6:255–259.
Jeanne O, Masson P (2000) Currency crises, sunspots and Markov-switching regimes. *J. Internat. Econom.* 50:327–350.
Johansen S, Nielsen MØ (2012) Likelihood inference for a fractionally cointegrated vector autoregressive model. *Econometrica* 80:2667–2732.
Jorion P (1992) Portfolio optimization in practice. *Financial Analysts J.* 48:68–74.

Kroner KF, Ng VK (1998) Modeling asymmetric comovements of asset returns. *Rev. Financial Stud.* 11:817–844.

Kroner K, Sultan J (1993) Time-varying distributions and dynamic hedging with foreign currency futures. *J. Financial Quant. Anal.* 28:535–551.

Laurent S, Rombouts JVK, Violante F (2011) On the forecasting accuracy of multivariate GARCH models. *J. Appl. Econometrics* 27:934–955.

Ledoit O, Wolf M (2003) Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empirical Finance* 10:603–621.

Lee HT (2009) A copula-based regime-switching garch model for optimal futures hedging. *J. Futures Markets* 29:946–972.

Lee HT, Yoder JK (2007) Optimal hedging with a regime-switching time-varying correlation GARCH model. *J. Futures Markets* 27:495–516.

Lee HT, Yoder JK, Mittelhammer RC, McCluskey JJ (2006) A random coefficient autoregressive Markov regime switching mode for dynamic futures hedging. *J. Futures Markets* 26:103–129.

Lien D, Yang L (2006) Spot-futures spread, time-varying correlation, and hedging with currency futures. *J. Futures Markets* 26:1019–1038.

Ljung GM, Box GEP (1978) On a measure of lack of fit in time series models. *Biometrika* 65:297–303.

Lypny G, Powalla M (1998) The hedging effectiveness of DAX futures. *Eur. J. Finance* 4:345–355.

McAleer M, Hoti S, Chan F (2009) Structure and asymptotic theory for multivariate asymmetric conditional volatility. *Econometric Rev.* 28:422–440.

Myers RJ, Thompson SR (1989) Generalized optimal hedge ratio estimation. *Amer. J. Agricultural Econom.* 71:858–868.

Nielsen MØ, Morin L (2012) FCVARmodel.m: A Matlab software package for estimation and testing in the fractionally cointegrated VAR model. Working Paper 1273, Queen's University, Kingston, ON, Canada. http://ideas.repec.org/p/qed/wpaper/1273.html.

Park TH, Switzer LN (1995) Bivariate GARCH estimation of the optimal hedge ratios for stock index futures: A note. *J. Futures Markets* 15:61–67.

Patton A (2006) Modelling asymmetric exchange rate dependence. *Internat. Econom. Rev.* 47:527–556.

Siegel AF, Woodgate A (2007) Performance of portfolios optimized with estimation error. *Management Sci.* 53:1005–1015.

Stock JH, Watson MW (2003) Forecasting output and inflation: The role of asset prices. *J. Econom. Perspect.* 41:788–829.

Stock JH, Watson MW (2004) Combination forecasts of output growth in a seven-country data set. *J. Forecasting* 23:405–430.

Tse YK, Tsui AKC (2002) A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *J. Bus. Econom. Statist.* 20:351–362.

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