



Volume 36, March 2013 ISSN 0140-9883

# Energy Economics

## CONTENTS

Energy Economics

Functional form and aggregate energy demand elasticities: A nonparametric panel approach for 17 OECD countries <i>A. Karimu and R. Brännlund</i>	19
Asymmetric adjustment of the dynamic relationship between energy intensity and urbanization in China <i>Y. Liu and Y. Xie</i>	43
Oil price shocks and trade imbalances <i>T.-H. Lee and Y. Chang</i>	78
Modeling a clean energy standard for electricity: Policy design implications for emissions, supply, prices, and regions <i>A. Paul, K. Palmer and M. Wierman</i>	108
Marketbidding in power markets: An experimental study <i>S. Arango, J.A. Castañeda and E.R. Larran</i>	125
If you build it, he will come: Anticipative power transmission planning <i>D. Picos, J. Contreras and E. Sauma</i>	135
Strategic behaviour in international metallurgical coal markets <i>J. Trübky</i>	147
The rebound effect in the aviation sector <i>A. Evans and A. Schäfer</i>	158
Further evidence of an Environmental Kuznets Curve in Spain <i>P. Sotillos and J. Munn</i>	177
A bilevel model for electricity retailers' participation in a demand response market environment <i>M. Rogna, J.M. Morales, P. Pinson and H. Madsen</i>	182

Contents continued on the outside back cover

Includes the Journal of  
Energy Finance and Development  
Available online at [www.sciencedirect.com](http://www.sciencedirect.com)  
SciVerse ScienceDirect

**Editors:**  
B.W. ANG  
R.S.J. TOL  
J. WEYANT

**Associate Editors:**  
P. AGNOLUCCI  
N. APERGIS  
L. BACHMEIER  
E.D. BAKER  
S.P. BROWN  
J. CHEVALLIER  
K. FISHER-VANDEN  
R. GERLAGH  
R.J. GREEN  
L.A. GREENING  
S. HAMMOUDEH  
B. HOBBS  
R. HUISMAN  
B.-N. HUANG  
C.-C. LEE  
T.T. LIN  
H. MA  
A. MAEDA  
H. MOHAMMADI  
W. OH  
C. PASURKA  
J.E. PAYNE  
D. POPP  
R.A. RATTI  
K. REHDANZ  
P. SADOWSKY  
A. SERLETIS  
A. SIDDIQUI  
Y. SMEERS  
U. SOYTAS  
R. SMYTH  
B. SU  
S. WANG  
T. ZACHARIADIS  
P. ZHOU

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



# The (de)merits of minimum-variance hedging: Application to the crack spread<sup>☆</sup>

Carol Alexander<sup>a</sup>, Marcel Prokopczuk<sup>b</sup>, Anannit Sumawong<sup>a,\*</sup>

<sup>a</sup> University of Sussex, UK

<sup>b</sup> Zeppelin University, Germany

## ARTICLE INFO

### Article history:

Received 3 May 2012

Received in revised form 12 November 2012

Accepted 17 November 2012

Available online 29 November 2012

### JEL classification:

G10

C52

### Keywords:

Hedging

Crack spread

GARCH

Minimum-variance hedge

## ABSTRACT

We study the empirical performance of the classical minimum-variance hedging strategy, comparing several econometric models for estimating hedge ratios of crude oil, gasoline and heating oil crack spreads. Given the great variability and large jumps in both spot and futures prices, considerable care is required when processing the relevant data and accounting for the costs of maintaining and re-balancing the hedge position. We find that the variance reduction produced by all models is statistically and economically indistinguishable from the one-for-one “naïve” hedge. However, minimum-variance hedging models, especially those based on GARCH, generate much greater margin and transaction costs than the naïve hedge. Therefore we encourage hedgers to use a naïve hedging strategy on the crack spread bundles now offered by the exchange; this strategy is the cheapest and easiest to implement. Our conclusion contradicts the majority of the existing literature, which favours the implementation of GARCH-based hedging strategies.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

There exists a substantial literature on minimum-variance hedging of spot positions using futures contracts in which sophisticated econometric models are applied for estimating the hedge ratios. The majority of these studies conclude that advanced econometric tools improve hedging performance over the naïve hedging strategy of shorting one futures contract per unit of spot exposure. However, most research ignores margin and transaction costs, and/or does not evaluate the improvement in a statistically meaningful way. Even, in some cases, insufficient care is taken to pre-filter the data for use in the analysis. Our contribution is to conduct an extensive out-of-sample study of minimum-variance hedging for a complex underlying position, with meticulous processing of the relevant data. We compare several popular hedging approaches and covariance estimation techniques with the simple naïve hedge, explicitly taking margin and transaction costs into account. In contrast to the majority of

extant literature we find that none of the sophisticated methods are able to outperform the naïve hedge.

Minimum-variance hedging was pioneered by Johnson (1960) and Stein (1961), and further refined by Ederington (1979), Hill and Schneeweis (1982), Figlewski (1984) and Herbst et al. (1989) amongst many others. Since Fama (1965) found that asset covariance structures are time-varying and Bollerslev (1982) introduced the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) method of estimating conditional variance, the application of the GARCH family for estimating hedge ratios has been rapidly growing in popularity. Baillie and Myers (1991) first derived hedge ratios using the bivariate GARCH model. Kroner and Sultan (1993) utilise the CCC GARCH model in the foreign-exchange market and Gagnon et al. (1998) expand the study for multi-asset portfolios using the BEKK GARCH model. Haigh and Holt (2000, 2002) analyse hedging in the freight and crack spread markets using a modified BEKK GARCH model. Further work on GARCH-based hedging includes Lee and Yorder (2007), Lien (2008), Lee (2009, 2010), Chang et al. (2011), and Ji and Fan (2011). All these works conclude that a GARCH-based strategy is superior to other static hedges.

Supporters of GARCH hedge ratios argue in unison, that the implementation of GARCH is necessary in order to capture the time-varying asset covariance structure. This should allow GARCH-based minimum-variance hedging to provide greater variance reduction than naïve hedging. However, due to uncertainty in the GARCH process specification and in its parameter estimates, this may not be the case in practice. Moreover, typically, the hedge ratios derived from GARCH-type models are extremely volatile, suggesting unrealistically frequent re-balancing and hence very large transaction costs for the hedged portfolio.

<sup>☆</sup> Many thanks to two anonymous referees whose insightful comments helped to improve the paper considerably. This paper was submitted to Energy Economics before Prof. Alexander and Anannit Sumawong moved to the University of Sussex. The handling editor, Richard Tol, was not involved in any part of the recruitment process. We would like to thank Platts for providing the spot price data and NYMEX for providing information on margin requirements.

\* Corresponding author at: School of Business, Management and Economics, Sussex University, Falmer, Brighton, BN1 9SL, UK.

E-mail address: [a.sumawong@sussex.ac.uk](mailto:a.sumawong@sussex.ac.uk) (A. Sumawong).

Most previous papers utilise weekly log returns in the analysis, but log returns are not realised and, for assets with prices that can jump, log returns can be highly inaccurate proxies for percentage returns even when measured at the daily frequency. Additionally, since the hedged portfolio can have zero value, even its percentage return may be undefined. Thus, our hedging analysis is based on profit and loss (P&L) rather than on log or percentage returns. Also, in most previous papers the estimation of GARCH and OLS parameters is based on a very large sample size. This choice can bias results towards the GARCH approach because OLS regression attributes equal weight to all observations, including outdated information at the beginning of the sample. Moreover, it will typically result in a relatively small out-of-sample period, which consequently yields test results having relatively large standard errors.

However, the most important difference between our methodology and that employed by many of the papers cited above is the use of constant-maturity versus rollover futures series. Recently, Nguyen et al. (2011) have highlighted the pitfalls of using futures series which simply roll over into the next contract as expiry approaches. This practice creates a saw-tooth pattern in the basis which has the unintended effect of biasing the OLS minimum-variance hedge ratios. To avoid this bias, our analysis is based on constant-maturity futures P&Ls.

We base our study on the problem of hedging crack spread positions. Although one might argue that most of the previous literature has studied the problem of hedging equity or pure commodity positions, we decided to use this underlying as it is a more complicated hedging problem, where prices are highly variable and subject to frequent jumps. As such, more advanced methods have a greater chance to improve the performance. Indeed, as mentioned above, Haigh and Holt (2002) conclude that multivariate GARCH models are superior for hedging the crack spread, although they use a mean-variance rather than a minimum-variance framework.

The crack spread represents the profits from a simultaneous purchase and sale of crude oil and its refined products, mainly consisting of gasoline and heating oil. An  $a:b:c$  crack spread is defined as buying  $a$  units of crude oil, and selling  $b$  and  $c$  units of gasoline and heating oil, respectively. These ratios are set according to the refineries production technologies. In 1994 the New York Mercantile Exchange (NYMEX), which offers the highest trading volume on oil-related futures amongst all exchanges worldwide, introduced the possibility for refineries to put up a single margin for the 3:2:1 crack spread. Thus, if refineries hedge this position as a whole using futures contracts on crude oil, gasoline and heating oil in this fixed ratio, margin costs are reduced and maintaining the account is simplified for both parties. The popularity of this product led NYMEX to introduce single margins for any  $a:b:c$  crack spread position.

The rest of this paper is structured as follows: Section 2 presents the methodological framework; Section 3 describes the data; Section 4 presents the results; Section 5 concludes.

## 2. Background and Methodology

When hedging spot exposure with futures, the problem of hedge ratio estimation typically arises due to a maturity mismatch. The greater this mismatch, the more opportunity for minimum-variance hedge ratios to perform better than the naïve. If the futures contract matures exactly on the due date of the spot positions, the hedge ratio will be one and the naïve hedge would always be optimal. Like the studies mentioned in Section 1, we consider the typical case where the spot position is due weekly but the futures mature on the monthly cycle. This way, the spot exposure arises before the futures contract matures.<sup>1</sup> This is often

<sup>1</sup> One may argue that, under longer hedging horizons, i.e. every two weeks, transaction costs would be reduced on minimum-variance hedging models. Hence they may perform better when rebalanced less frequently. However, this also reduces the maturity mismatch which is the reason for considering minimum-variance hedging in the first place.

referred to as the *delta hedging problem* and is discussed in seminal works, such as Cecchetti et al. (1988) and Baillie and Myers (1991).<sup>2</sup>

There are two common hedging frameworks in the literature: minimum-variance, where the variance of the hedged portfolio is minimised; and mean-variance, where the investor's utility is maximised based on a mean-variance criterion. The origins of minimum-variance hedging stem from Johnson (1960) where, by minimising the variance of the hedged portfolio under first-order conditions, the hedge ratios can be expressed as a function of the variances and covariances of the spot and futures P&L. The hedged portfolio P&L,  $\Delta\Pi_t = \Pi_t - \Pi_{t-1}$ , can be described by vectors of  $m$  spot P&Ls, futures P&Ls and hedge ratios ( $\Delta S'_t$ ,  $\Delta F_t$  and  $\beta$ ) such that

$$\Delta\Pi_t = \mathbf{1}'_m \Delta S_t + \beta' \Delta F_t,$$

where  $\mathbf{1}_m$  is a  $m \times 1$  vector of ones. The variance,  $V[\Delta\Pi_t]$ , of this portfolio is given by

$$V[\Delta\Pi_t] = \mathbf{1}'_m V[\Delta S_t] \mathbf{1}_m + \beta' V[\Delta F_t] \beta + 2\beta' Cov[\Delta F_t, \mathbf{1}'_m \Delta S_t],$$

where  $V[\Delta S_t]$  and  $V[\Delta F_t]$  are the covariance matrices of  $\Delta S_t$  and  $\Delta F_t$ , respectively, and  $Cov[\Delta F_t, \mathbf{1}'_m \Delta S_t]$  is a vector of covariances between  $\mathbf{1}'_m \Delta S_t$  and the individual elements of  $\Delta F_t$ . Minimising the variance yields the optimal hedge ratio vector

$$\beta = -V[\Delta F_t]^{-1} Cov[\Delta F_t, \mathbf{1}'_m \Delta S_t].$$

On the other hand, for the mean-variance criterion, following Working (1953), the investor's expected utility value of the hedged portfolio is measured by the certainty equivalent income,  $CEI[\Delta\Pi_t]$ . Under certain restrictive conditions this can be expressed analytically as

$$CEI[\Delta\Pi_t] = E[\Delta\Pi_t] - \frac{1}{2\gamma} V[\Delta\Pi_t],$$

where  $\gamma$  is the coefficient of absolute risk tolerance. Maximising the CEI gives the optimal mean-variance hedge ratio vector as

$$\beta^* = V[\Delta F_t]^{-1} (\gamma E[\Delta F_t] - Cov[\Delta F_t, \mathbf{1}'_m \Delta S_t]).$$

Although the latter framework is often employed, we have chosen the former for the following reasons: first, the analytic solution for mean-variance hedging imposes several restrictions (normally distributed P&Ls and an exponential or quadratic utility function) which are unrealistic for our hedging problem; Second, many studies that employ this framework, e.g. Lee and Yorder (2007), make the additional assumption that the speculative (futures) component of the mean-variance hedge follows a martingale process – but then the mean-variance framework ultimately collapses to the minimum-variance framework; Lastly, an out-of-sample analysis of the speculative component depends critically on the expected P&L from the futures series, which typically over-shadows the risk adjustment in the utility.<sup>3</sup> For volatile commodities such as oil, reducing the variance should be the main aim of hedging.

Estimation of hedge ratios for a multi-asset problem is carried out in two steps: First make a decision on the hedging model to be used (the models differ in the number of equations and variables considered); second, estimate the relevant variance and covariance parameters for the chosen model. We consider each decision in turn.

<sup>2</sup> It should be contrasted with the *rollover* hedging problem, where the futures in the hedged portfolio mature before the spot positions.

<sup>3</sup> See Chopra and Ziemba (1993) amongst others.

2.1. Hedging models

Let the  $a:b:c$  crack spread spot and futures prices,  $S_t^z$  and  $F_t^z$ , be given by

$$S_t^z = -aS_t^c + bS_t^g + cS_t^h, \quad F_t^z = -aF_t^c + bF_t^g + cF_t^h,$$

where  $S_t^c, S_t^g, S_t^h, F_t^c, F_t^g$  and  $F_t^h$  denote the spot and futures prices for crude oil, gasoline and heating oil, respectively. The realised hedged portfolio P&L is given by

$$\Delta\Pi_t = \Delta S_t^z + a\beta^c \Delta F_t^c - b\beta^g \Delta F_t^g - c\beta^h \Delta F_t^h, \tag{1}$$

where  $\beta^c, \beta^g, \beta^h$  are the hedge ratios. For the naïve hedge,  $\beta^c = \beta^g = \beta^h = 1$ .

The hedge ratios that minimise the variance of (1) can be obtained by solving the first-order conditions

$$\begin{bmatrix} a\beta^c \\ -b\beta^g \\ -c\beta^h \end{bmatrix} = - \begin{bmatrix} a^2\sigma_{\Delta F_t^c \Delta F_t^c} & -ab\sigma_{\Delta F_t^c \Delta F_t^g} & -ac\sigma_{\Delta F_t^c \Delta F_t^h} \\ -ab\sigma_{\Delta F_t^g \Delta F_t^c} & b^2\sigma_{\Delta F_t^g \Delta F_t^g} & bc\sigma_{\Delta F_t^g \Delta F_t^h} \\ -ac\sigma_{\Delta F_t^h \Delta F_t^c} & bc\sigma_{\Delta F_t^h \Delta F_t^g} & c^2\sigma_{\Delta F_t^h \Delta F_t^h} \end{bmatrix}^{-1} \begin{bmatrix} a\sigma_{\Delta S_t^z \Delta F_t^c} \\ -b\sigma_{\Delta S_t^z \Delta F_t^g} \\ -c\sigma_{\Delta S_t^z \Delta F_t^h} \end{bmatrix}, \tag{2}$$

where  $\sigma_{ij}$  denotes the covariance between  $i$  and  $j$ . This method is analogous to Ordinary Least Squares (OLS) regressions of the spot P&L on the hedging instrument(s). In other words, finding the hedge ratios in (2) is analogous to performing the single-equation, multiple-variable regression

$$\Delta S_t^z = \alpha - a\beta^c \Delta F_t^c + b\beta^g \Delta F_t^g + c\beta^h \Delta F_t^h + \epsilon_t, \tag{3}$$

where  $\epsilon_t$  denotes the regression residuals. We refer to this hedging model as the *single-equation, multiple-variable model*. Since each commodity is exposed to closely related risk factors, it is expected that multicollinearity is present in this setting. Consequently, hedge ratios derived from (3) may have biased standard errors yielding to imprecise hedge ratio estimates.

Alternatively, one might employ a *multiple-equation, single-variable model* in which the hedge ratios are estimated via three asset-by-asset regressions, as follows:

$$\begin{bmatrix} a\Delta S_t^c \\ b\Delta S_t^g \\ c\Delta S_t^h \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} a\beta^c \Delta F_t^c \\ b\beta^g \Delta F_t^g \\ c\beta^h \Delta F_t^h \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \epsilon_{3,t} \end{bmatrix}.$$

This model does not account for the covariances between the futures P&Ls at all. In other words, the hedge ratio calculation is the same as in (2), but the off-diagonal elements in the square matrix are assumed to be zero.<sup>4</sup>

The first model requires estimates of several covariances and variances and each are prone to estimation errors. In contrast, the second model assumes all futures covariances to be zero. A third possibility is to simply impose the constraint  $\beta^z := \beta^c = \beta^g = \beta^h$ , whereby one obtains the parsimonious *single-equation, single-variable model* for which the estimation errors might be significantly reduced. This model, which is nested in (3), is given by

$$\Delta S_t^z = \alpha + \beta^z \Delta F_t^z + \epsilon_t, \tag{4}$$

<sup>4</sup> We also estimated hedge ratios using generalised least squares (GLS) in a seemingly unrelated regression equations (SURE) system for this model. As the hedging effectiveness results were indistinguishable; we do not report them.

with optimal hedge ratio given by

$$\beta^z = \frac{\sigma_{\Delta S_t^z \Delta F_t^z}}{\sigma_{\Delta F_t^z}^2}.$$

The price to pay for a parsimonious model is the implicit assumption of constant correlations between the futures P&Ls. This model has not previously been considered in the literature, but when the correlations between the components of a multiple hedge portfolio are high, then so are the estimation errors in the covariances of the futures P&Ls in (3). Hence, one might expect a superior performance from the *single-equation, single-variable model* despite its restrictive assumptions on correlation.

2.2. Estimation Methods

We now turn to the econometric methods used to estimate the variances and covariances in the hedging models. We employ four different popular estimation methods: OLS; exponentially weighted moving averages (EWMA); the standard symmetric GARCH; and an asymmetric GARCH model. To conduct an out-of-sample study we re-estimate all parameters of the OLS and GARCH models using a rolling window of length  $n$ . The parameter of the EWMA model is fixed, a priori.

With OLS, variances and covariances of two assets  $Y_1$  and  $Y_2$  are simply estimated by their sample counterparts

$$\hat{\sigma}_{\Delta Y_{1,t}}^2 = \frac{1}{n-1} \sum_{i=0}^{n-1} (\Delta Y_{1,t-i} - \Delta \bar{Y}_{1,t})^2,$$

and

$$\hat{\sigma}_{\Delta Y_1 \Delta Y_2,t} = \frac{1}{n-1} \sum_{i=0}^{n-1} (\Delta Y_{1,t-i} - \Delta \bar{Y}_{1,t})(\Delta Y_{2,t-i} - \Delta \bar{Y}_{2,t}),$$

respectively. EWMA variances and covariances are estimated via the recursions

$$\hat{\sigma}_{\Delta Y_{1,t}}^2 = (1-\lambda)\Delta Y_{1,t-1}^2 + \lambda\hat{\sigma}_{\Delta Y_{1,t-1}}^2,$$

and

$$\hat{\sigma}_{\Delta Y_1 \Delta Y_2,t} = (1-\lambda)\Delta Y_{1,t-1}\Delta Y_{2,t-1} + \lambda\hat{\sigma}_{\Delta Y_1 \Delta Y_2,t-1},$$

where  $\lambda$  is the EWMA decay coefficient which takes a value between 0 and 1. With a lower  $\lambda$  more emphasis is placed on the most recent observations and the model hence becomes more reactive to changing market conditions.

GARCH variances and covariances are obtained using the BEKK model specification of Engle and Kroner (1995). For a vector of zero mean P&Ls  $\Delta Y_t$ , the multivariate GARCH covariance matrix estimate  $H_t$  is based on the dynamics

$$H_t = A'A + (B'\Delta Y_{t-1})(B'\Delta Y_{t-1})' + C'H_{t-1}C,$$

where  $A, B, C$  are  $m \times m$  matrices of the BEKK parameters for  $m$  assets. The parameter estimates are obtained by maximising the log-likelihood function<sup>5</sup>

$$\ln L(\theta) = -\frac{1}{2} \sum_{t=1}^n (\ln(|H_t|) + \Delta Y_t' H_t^{-1} \Delta Y_t).$$

<sup>5</sup> We use Kevin Sheppard's UCSD GARCH toolbox for the estimation, available at [http://www.kevinshppard.com/wiki/UCSD\\_GARCH](http://www.kevinshppard.com/wiki/UCSD_GARCH).

As it is well known that the symmetric GARCH specification can be improved by allowing for an asymmetric variance response to shocks we also employ the asymmetric GARCH BEKK specification (AGARCH) of Grier et al. (2004). Here the variances dynamics are specified as

$$\hat{H}_t = \mathbf{A}\mathbf{A}' + (\mathbf{B}'\Delta\mathbf{Y}_{t-1})(\mathbf{B}'\Delta\mathbf{Y}_{t-1})' + \mathbf{C}'\hat{H}_{t-1}\mathbf{C} + (\mathbf{D}'\Delta\mathbf{Y}_{t-1}^*)(\mathbf{D}'\Delta\mathbf{Y}_{t-1}^*)'$$

where  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ ,  $\mathbf{D}$  are  $m \times m$  matrices of the asymmetric BEKK parameters for  $m$  assets and  $\mathbf{Y}_t^*$  is a vector of  $\max\{Y_t, 0\}$  for a positively skewed sample or  $\min\{Y_t, 0\}$  for a negatively skewed sample.

For ease of presentation, we abbreviate the hedging models and estimation techniques by  $model_{ij}$  where  $model$  denotes the estimation method, i.e.  $model = \{\text{OLS}, \text{EWMA}, \text{GARCH}, \text{AGARCH}\}$ , and  $i = 1, 3$  and  $j = 1, 3$  denote the number of equations and variables in the regression system respectively. For example,  $\text{EWMA}_{13}$  refers to the single-equation, multiple-variable model as specified in (3) where the variances and the covariances are estimated using the EWMA method, etc.

In total, seven hedging models are analysed: naïve,  $\text{OLS}_{31}$ ,  $\text{OLS}_{13}$ ,  $\text{OLS}_{11}$ ,  $\text{EWMA}_{11}$ ,  $\text{GARCH}_{11}$  and  $\text{AGARCH}_{11}$ . For the EWMA, GARCH and AGARCH estimation methods we omit results for multiple-equation or multiple-variable models because preliminary results, based only on the OLS models, show that the three regression configurations are more or less equally effective. Moreover, the proliferation of parameters when GARCH and AGARCH models are applied to multiple-equation or multiple-variable models exacerbates the problem of parameter estimate instability, which is discussed later on with reference to Fig. 3.

### 2.3. Margin and transaction costs

When trading futures on the NYMEX, transaction costs arise from the round trip commission charged by the exchange and from the bid-ask spread. Since the early 2000s, the NYMEX has reduced the round trip commission costs from \$15.00 to \$1.45 per futures contract bought and sold. Although the NYMEX is an open-outcry market (which allows limit orders) the hedger is assumed to place market orders, to prioritise the variance reduction over possible gains from trading. The bid-ask spreads of the three considered commodity futures,  $x$ ,  $y$  and  $z$  are defined as the average spread between the bid and the ask price divided by the average mid price of each commodity futures. The dollar value of the bid-ask spreads,  $TC_t$ , arises from re-balancing the portfolio and is given by:

$$TC_t = aF_t^c |\Delta\beta_t^c| x + bF_t^g |\Delta\beta_t^g| y + cF_t^h |\Delta\beta_t^h| z.$$

The modulus signs are placed to indicate how the hedger loses the spread regardless of the direction of trade. We follow Dunis et al. (2008) and set  $x$ ,  $y$  and  $z$  to be 1 bps, 10 bps and 12 bps, respectively. Although these are bid-ask spreads of first-to-mature rollover series, we assume that these are constant through the first two months of the term structure and hence use the same for our constant-maturity futures.

Margin costs arise from raising the initial margin and from marking-to-market the maintenance margins. In the past decade there have been several changes to the NYMEX margin requirement rules. When trading an  $a:b:c$  crack spread, NYMEX calculates the initial margins based on the portfolio Value-at-Risk at the 5% or 1% level for commercial (hedgers) and non-commercial (speculators) traders, respectively. For a hedger, who shorts a crack spread expiring in 1 month, the initial margin is approximately \$11, \$18 and \$7 per 3:2:1, 5:3:2 and 2:1:1 crack spread bundles respectively (as opposed to \$15, \$25 and \$10 for a speculator).<sup>6</sup> We shall focus on the costs

incurred by refineries, which are generally treated as hedgers by the clearing house. The total cost  $m_t^i$  from raising the initial margin is

$$m_t^i = |\beta_t^z| N (r_t^d - r_t^f),$$

where  $r_t^d$  is the cost of raising the initial margin,  $r_t^f$  is the risk-free rate of return gained from depositing in the margin account,  $N$  is the initial margin required per crack spread bundle and  $\beta_t^z$  is the number of crack spread bundles purchased. In the cases where the hedge ratios do not allow for exact transaction of the bundles (i.e.  $\beta^c \neq \beta^g \neq \beta^h$ ), the approximation  $\beta^z \approx \frac{a\beta^c + b\beta^g + c\beta^h}{a+b+c}$  is taken instead. The refinery is assumed to raise debt for the initial margin,  $r_t^d$  is set as the average cost of debt in the industry. The top ten US refineries are currently, on average, rated AA by Moody's. Hence Moody's AA bond index is chosen as a proxy for the cost of debt  $r_t^d$ . The initial margin is set at \$11, \$18 and \$7 for the 3:2:1, 5:3:2 and 2:1:1 bundles, respectively. These were the values quoted by NYMEX on 06/06/2011. Three-months US T-bill rates are used as a proxy for  $r_t^f$ .

The gains and losses from the maintenance margin arise from the movement in the futures prices every day. These are marked-to-market daily but as we work with weekly data. We employ a linear approximation of the daily changes in the margin account. The weekly interest on the margin account  $m_t^m$  is therefore approximated as

$$m_t^m = \frac{1}{2} (-a\beta_t^c \Delta F_t^c + b\beta_t^g \Delta F_t^g + c\beta_t^h \Delta F_t^h) r_t^f.$$

The total hedged portfolio P&L including margin and transaction costs  $\Delta\Pi_t^*$  may now be expressed as

$$\Delta\Pi_t^* = \Delta\Pi_t + m_t^m - m_t^i - TC_t. \tag{5}$$

### 2.4. Performance measurement

Hedging effectiveness is measured by the Ederington Effectiveness (EE) calculated as

$$EE = \frac{\sigma_u^2 - \sigma_h^2}{\sigma_u^2},$$

where  $\sigma_u^2$  and  $\sigma_h^2$  are the variances of the unhedged and the hedged portfolios, respectively. We compute the EE for each model in two ways: (i) using unconditional variances over the whole sample period and (ii) using a rolling window of EWMA variances with  $\lambda = 0.97$ , i.e. we use the conditional EE measures employed by Alexander and Barbosa (2007). The EWMA method is preferred to a rolling window of unconditional variances because the latter produces ghost features where the variances are augmented as long as a spike in the P&L remains inside the window. This is also to avoid any possible bias the unconditional EE may have over the conditional EE, as highlighted by Lien (2005, 2009). The conditional EE also allows us to examine how the models' performance changes, as price series move through volatile and non-volatile time periods. To test whether the variance reduction from each model is significantly different from the variance reduction obtained using the naïve hedge, we apply the standard F-test for equality of variances.

## 3. Data

### 3.1. Spot prices

Wednesday spot prices from 30/12/1992 to 23/02/2011 of Cushing WTI light-sweet crude oil, New York Harbour heating oil No. 2, unleaded gasoline and RBOB gasoline barges are taken from Platts. In the rare cases where Wednesday is not a trading day, the price

<sup>6</sup> For an  $a:b:c$  crack spread, a "bundle" indicates simultaneously going long  $a$  barrels of crude oil, short  $b$  barrels of gasoline and short  $c$  barrels of heating oil.

on Tuesday is taken instead.<sup>7</sup> The delivery location of the spot prices is the same as their corresponding NYMEX futures.

We use Platts prices as these are collected from a window of physical commodity buyers which truly reflect the spot of the physical commodity trades. Platts prices are determined at 4:30 pm GMT as opposed to the NYMEX futures prices which are determined at 5:00 pm GMT-5/6 (depending on summer/winter time zones), posing a non-synchronicity problem between the two sources. This may invoke a downward bias on the daily correlation between the spot and futures prices, but our analysis is on weekly data with weekly hedging horizons. As such, this relatively minor time difference will have negligible effect on the empirical results.

### 3.2. Futures prices

Wednesday NYMEX futures prices of crude oil, heating oil, and gasoline from 30/12/1992 to 23/02/2011 are based on the NYMEX closing price. Among these three commodities, gasoline production has undergone some changes over time and therefore, since 2006, the NYMEX has no longer offered the original unleaded gasoline futures, replacing them by Reformulated Blendstock for Oxygen Blending (RBOB) gasoline futures. Due to data availability and low liquidity in the early years of the RBOB futures market, we switch from unleaded to RBOB gasoline in different years in the spot (2003) and futures markets (2006). This problem is of limited importance as both types of gasoline face the same demand and supply trends so that the prices are extremely highly correlated.

There are two ways to create a continuous series of futures P&Ls: the rollover method and the constant-maturity method. A standard rollover series is constructed by taking a futures price series up to a rollover date, the price series then jumps to the prompt futures series which is taken up to the next rollover date and so on. Often, the rollover dates are roughly a week before maturity to avoid thin market trading but for the commodities we study there is no need for this adjustment since trading continues in high volumes right up to the maturity date.

However, there are two problems associated with the rollover futures series. First, as explained by Nguyen et al. (2011) where unlike constant-maturity series, any regression relating spot data to futures data will be contaminated by the “saw-tooth” pattern in the basis. Second, gasoline and heating oil futures contain breaks in the term structure due to strong seasonality effects: prices at year-end periods are expected to be lower/higher than the rest of the year, and start of the year prices are expected to be higher/lower than the rest of the year. Hence, futures contracts which mature at these periods are often priced at a significant step above or below the rest of the year. When rolling from a mid-year contract to a year-end contract (say, October to November) an artificial jump is created through this seasonality effect. In fact, we obtain up to four jumps in the rollover series per year, which produces outliers in the P&L data that are merely an artefact of the rollover methodology and are difficult to deal with statistically.

Such artefactual jumps are not present in the constant-maturity futures P&L series that are constructed by linearly interpolating between two futures with maturities above and below the constant-maturity target.<sup>8</sup> Galai (1979) compares two methods for creating constant-maturity futures which he terms the value index (interpolation between prices) and the return index (interpolation between returns). Galai shows that the return index method is the only one

that provides realisable investments. As we require realisable investments to implement the optimal hedge ratios in practice, but our analysis must also be based on P&L rather than returns, we adapt Galai's return index method to the P&L as follows:

$$\Delta F_{t,T} = \eta_t \Delta F_{t,T_1} + (1 - \eta_t) \Delta F_{t,T_2}, \quad 0 \leq \eta_t \leq 1,$$

where  $\Delta F_{t,T}$  is the constant-maturity futures P&L expiring in  $T$  days,  $\Delta F_{t,T_1}$  and  $\Delta F_{t,T_2}$  are the futures P&Ls expiring at  $T_1$  and  $T_2$  respectively, and

$$\eta_t = \frac{T_2 - (t + T)}{T_2 - T_1}, \quad T_1 < T < T_2.$$

A reasonable choice for  $T$  is 44 calendar days, i.e. approximately 1.5 months. With this choice there will always be two maturities straddling the constant-maturity. Of course, to maintain a constant-maturity series of  $\Delta F_t^2$  for the regression (4) the constant-maturity must be the same for all futures.

### 3.3. Summary statistics

Tables 1 and 2 report summary statistics and correlations of the weekly spot and constant-maturity futures P&L distributions based on the entire sample period. Crude oil spot and futures are less volatile than gasoline and heating oil spot and futures, and in each case the spot is more volatile than the futures. Each P&L except spot heating oil is slightly negatively skewed and all series are highly leptokurtic.

Fig. 1 displays the P&L time series for all six variables. We observe that all series show rising volatility from the year 2000 onwards. Surges in prices produced by unexpected supply shortages result in frequent jumps in all the series. In many cases a decoupling of spot and futures prices results in a jump in the basis which is difficult to hedge effectively with the one-for-one ratio, and possibly also with a minimum-variance hedge ratio. Only one, very extreme spike in the data was removed. This was during the week of Hurricane Katrina, during which we assume no trades were made.

## 4. Empirical results

We study the hedging performance of seven different models: naïve, OLS<sub>31</sub>, OLS<sub>13</sub>, OLS<sub>11</sub>, EWMA<sub>11</sub>, GARCH<sub>11</sub> and AGARCH<sub>11</sub> both, in-sample and out-of-sample. For the in-sample analysis, parameters are estimated using the entire data set, i.e. 939 weekly observations. Hedge ratios are then calculated based on these parameters and held constant for computing the hedge performance. But clearly, the in-sample analysis is just a data-fitting exercise – it is the out-of-sample analysis that matters for practical purposes. Here, the parameters are estimated using a rolling window of 260 weeks.<sup>9</sup> The hedge ratios estimated at time  $t$  are then applied to the one step ahead P&L. The hedger is assumed to re-estimate the parameters every week.<sup>10</sup> Since the EWMA parameter  $\lambda$  is always constant, EWMA results are the same both in-sample and out-of-sample.

All empirical results presented are for the 3:2:1 crack spread bundle, as many refineries have this approximate crack spread and the original NYMEX margin bundles were also based on this spread.

<sup>9</sup> For the OLS methods we have also employed windows of length 104, 156, and 208 to ensure that this choice is not the driver of our results. No significant differences were found. For the GARCH models, shorter windows were not feasible due to the number of parameters to be estimated. In some few instances, the optimisation of the GARCH parameters failed to converge. We then used the estimates from the previous week.

<sup>10</sup> A separate analysis for annual recalibration of parameters was also carried out, attempting to tame the volatility of the GARCH parameter estimates. Hedging performance was found to be inferior to the alternative weekly re-estimation method. Parameters also jumped on a year-to-year basis. Hence the results are not presented here but are available on request.

<sup>7</sup> And in the case where Tuesday is not a trading day as well as Wednesday, Monday's price is taken instead. In the circumstance where none of those days are trading days, the week is omitted entirely.

<sup>8</sup> With constant-maturity futures the hedger is required to re-balance at the frequency of the data (in our case every week) but the transaction costs incurred are negligible.

**Table 1**

Summary statistics for weekly constant-maturity futures and spot P&Ls for the sample period 30/12/1992 to 23/02/2011. The total number of observations is 939 for each series.  $\mu$ ,  $\sigma$ ,  $\tau$  and  $\kappa$  denote the mean, standard deviation, skewness and excess kurtosis, respectively.

	$\Delta F^c$	$\Delta F^g$	$\Delta F^h$	$\Delta S^c$	$\Delta S^g$	$\Delta S^h$
$\mu$	0.0324	0.1100	0.0569	0.0820	0.0964	0.1045
$\sigma$	2.3609	2.7426	2.7518	2.4586	3.1717	2.9239
$\tau$	-0.2804	-0.3470	-0.0239	-0.1122	-0.3049	0.0431
$\kappa$	6.3410	3.3010	7.0075	6.1150	3.2928	5.8725

**Table 2**

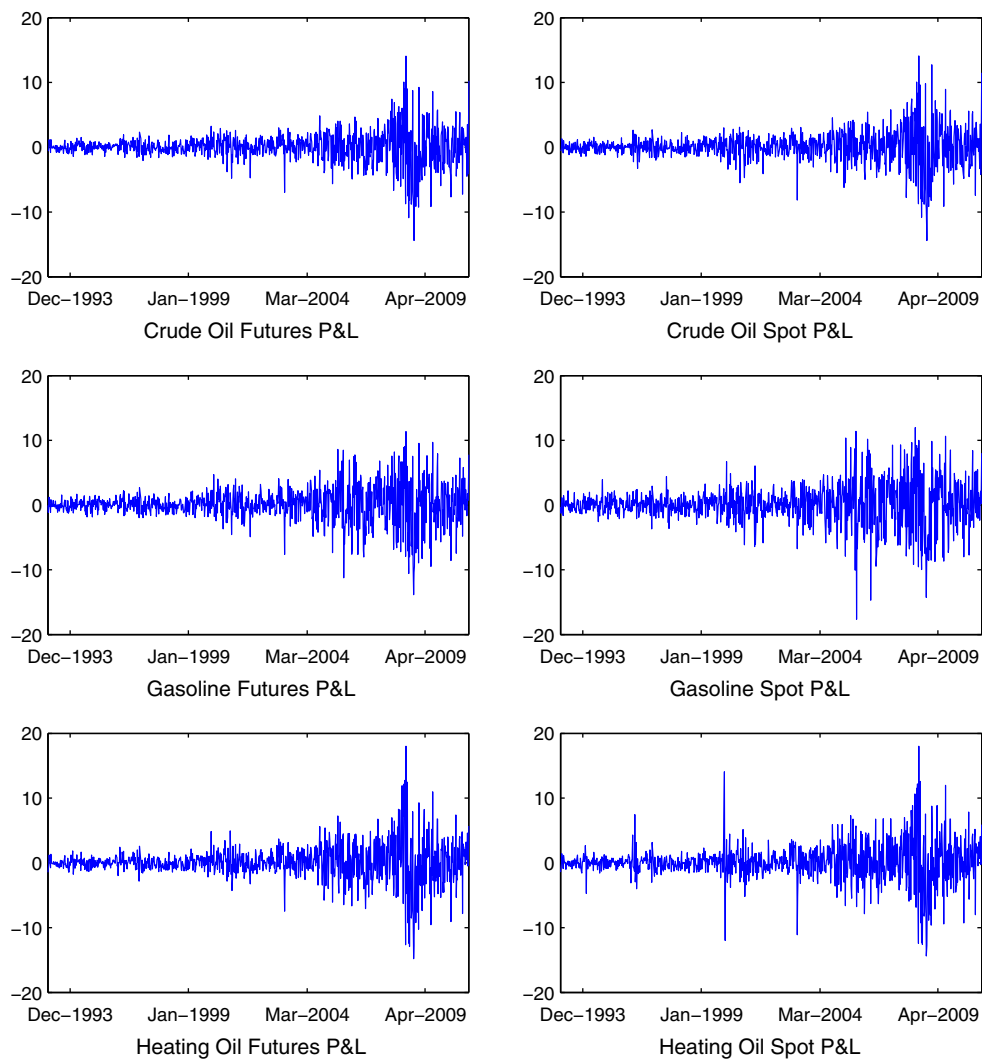
Correlation matrix between spot and futures P&Ls for the sample period 30/12/1992 to 23/02/2011. The total number of observations is 939 for each series.

	$\Delta F^c$	$\Delta F^g$	$\Delta F^h$	$\Delta S^c$	$\Delta S^g$	$\Delta S^h$
$\Delta F^c$	1	-	-	-	-	-
$\Delta F^g$	0.8539	1	-	-	-	-
$\Delta F^h$	0.9006	0.8395	1	-	-	-
$\Delta S^c$	0.9718	0.8268	0.8683	1	-	-
$\Delta S^g$	0.7357	0.9334	0.7423	0.7106	1	-
$\Delta S^h$	0.8385	0.7859	0.9507	0.8128	0.6981	1

**Table 3**

Average in-sample and out-of-sample hedge ratios with standard deviations in parentheses. In-sample ratios are estimated using the entire sample period: 30/12/1992–23/02/2011. Out-of-sample hedge ratios are estimated using a moving window of 260 weeks.

	in-sample	OLS <sub>13</sub>	OLS <sub>31</sub>	OLS <sub>11</sub>	EWMA <sub>11</sub>	GARCH <sub>11</sub>	AGARCH <sub>11</sub>
$\beta^c$	1.285	1.012	1.277	1.381	1.363	1.381	
	-	-	-	(0.215)	(0.262)	(0.205)	
$\beta^g$	1.296	1.079	1.277	1.381	1.363	1.381	
	-	-	-	(0.215)	(0.262)	(0.205)	
$\beta^h$	1.202	1.010	1.277	1.381	1.363	1.381	
	-	-	-	(0.215)	(0.262)	(0.205)	
<i>out-of-sample</i>							
$\beta^c$	1.417	1.035	1.412	1.381	1.392	1.368	
	(0.162)	(0.021)	(0.135)	(0.215)	(0.343)	(0.326)	
$\beta^g$	1.432	1.147	1.412	1.381	1.392	1.368	
	(0.109)	(0.039)	(0.135)	(0.215)	(0.343)	(0.326)	
$\beta^h$	1.349	1.084	1.412	1.381	1.392	1.368	
	(0.363)	(0.061)	(0.135)	(0.215)	(0.343)	(0.326)	



**Fig. 1.** Spot and constant-maturity futures P&L series for each commodity. Period: 30/12/1992–23/02/2011. Prices for the week of 28/08/2005–02/09/2005 have been removed due to abnormal market conditions caused by hurricane Katrina. The investor is assumed to make no trades on this week.

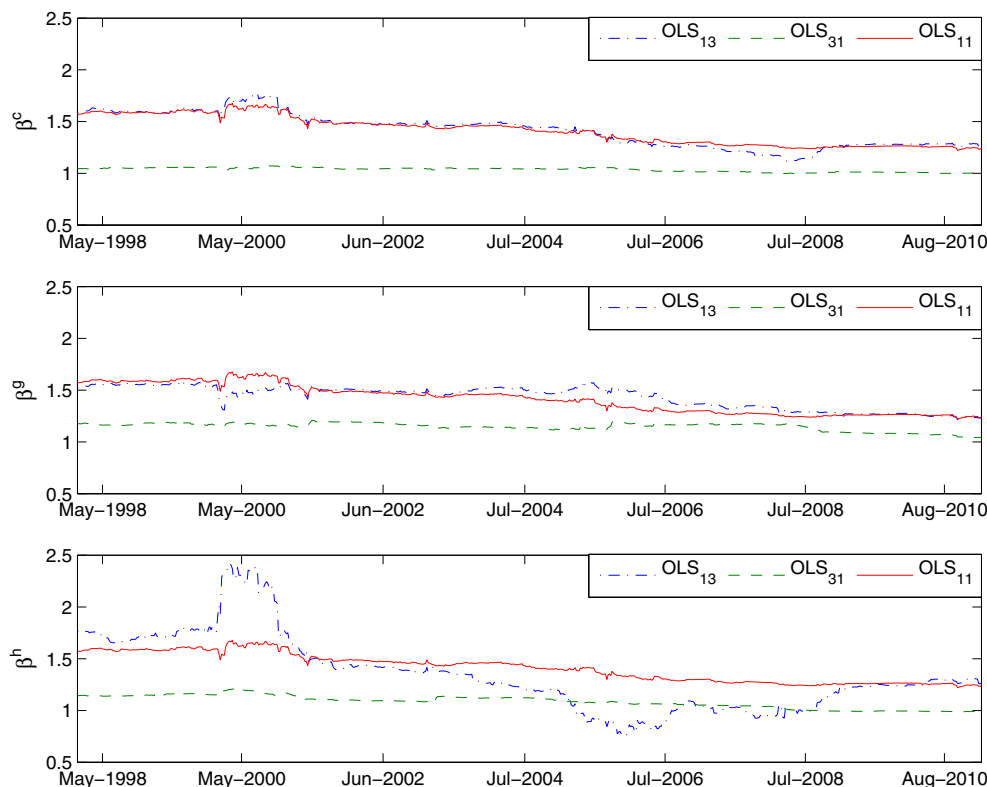


Fig. 2. OLS hedge ratios calculated using a moving window of 260 weeks. Period: 14/01/1998–23/02/2011.

Robustness checks were carried out by repeating our analysis for two other crack spread ratios, namely 5:3:2 and 2:1:1. Since our qualitative conclusions remained the same for these other crack spread bundles detailed results are not reported here. However, they are available from the authors on request.

#### 4.1. Hedge ratios

Table 3 reports the average hedge ratios for each model and their standard deviations. In-sample hedge ratios are reported for completeness, we focus the following discussion on the out-of-sample hedge ratios. The multiple-equation model  $OLS_{31}$  yields hedge ratios closer to 1, yet the single-equation models produce hedge ratios nearer to 1.3. It is tempting to conclude that the  $OLS_{13}$  model produced these higher hedge ratios because of multicollinearity. However, both  $OLS_{13}$  and  $OLS_{11}$  produce hedge ratios of roughly the same magnitude, which brings into question any such conclusion.

The  $OLS_{31}$  model produces smaller hedge ratios, closer to 1, because all cross-market correlations are assumed to be zero. As they are certainly not (see Table 2) this produces a substantial bias. On the other hand, the  $OLS_{13}$  assumes an equal cross-market correlation across all commodities – an assumption that seems reasonable in light of Table 2.

Fig. 2 displays the evolution of the OLS models' out-of-sample hedge ratios over time. One can observe that  $OLS_{31}$  and  $OLS_{11}$  are relatively stable. In contrast, the hedge ratio for the heating oil contract of  $OLS_{13}$  in particular exhibits some substantial transitions over time. This is also reflected by the relatively high standard deviation in Table 3. The  $AGARCH_{11}$  estimation method produced the most volatile out-of-sample hedge ratios. Although this characteristic is expected given that GARCH parameters are generally more sensitive with respect to innovations in the data, the volatility of the hedge ratios should

roughly be of the same magnitude as the  $EWMA_{11}$  hedge ratios.<sup>11</sup> According to Table 3 however, the out-of-sample  $GARCH_{11}$  hedge ratios are roughly 33% more volatile than the  $EWMA_{11}$  hedge ratios.

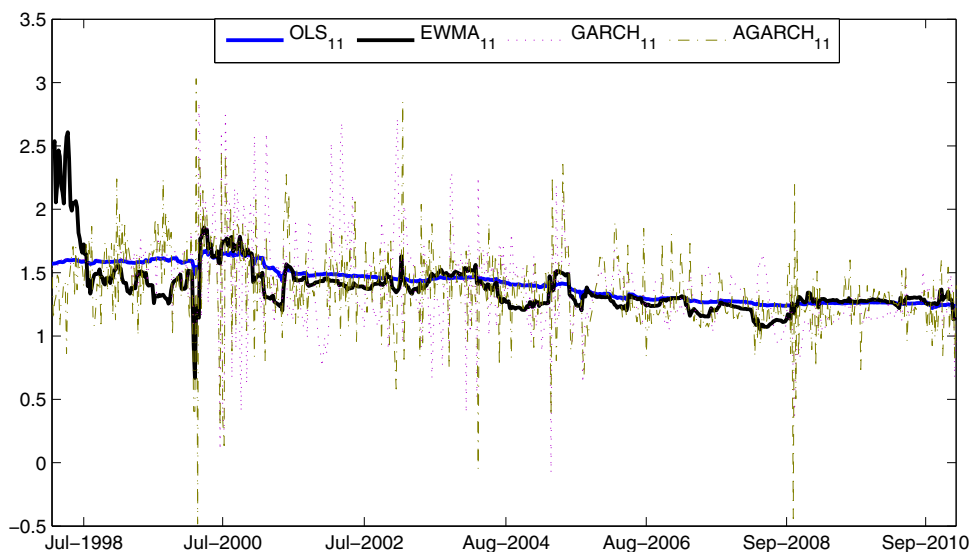
This is also shown in Fig. 3, which compares the behaviour of the hedge ratios derived for the single-equation, single-variable models over time. Note how volatile the GARCH hedge ratios are over time. Would a serious risk manager implement a hedging strategy that involved re-balancing more than 100% of the hedging portfolio from week to week? This casts serious doubts on the merits of GARCH-based hedge ratios.

The GARCH model parameter estimates are highly volatile over time. Table 4 displays the means of the estimated GARCH parameters, and their standard deviations measured over the entire out-of-sample period. Clearly, the estimates are far from being stable. Extending the length of the estimation windows up to 8 years did not produce substantially more stable estimates. Hence, the problem is not one of convergence to local optima instead of a global optimum, but rather an intrinsic problem with applying GARCH models for hedging when there are frequent jumps in a highly volatile basis. In this situation, large changes in the conditional variance parameter estimates are only to be expected. Indeed, the finding of highly unstable GARCH hedge ratios is nothing peculiar for our data set. Previous studies, e.g. Lee and Yorder (2007) and Lee (2010), have found similar results.

Another problem concerns transaction costs. Re-balancing a hedge with such extreme swings will amount to much higher transaction costs in comparison to the other methods having more stable hedge ratios. Table 5 presents the average transaction costs (including

<sup>11</sup> For some periods, the GARCH hedge ratios are unmanageably large for the investor (e.g. up to  $\pm 5$  times the spot investment). To control these, when the absolute value of the GARCH hedge ratios exceed twice the absolute value of the EWMA hedge ratio, the GARCH hedge ratio from the previous time step is used instead.





**Fig. 3.** Comparisons between out-of-sample hedge ratio estimates of the OLS<sub>11</sub>, EWMA<sub>11</sub>, GARCH<sub>11</sub> and AGARCH<sub>11</sub> models. EWMA<sub>11</sub> hedge ratios estimated with  $\lambda = 0.97$ . Period: 04/01/1998–23/02/2011.

margin costs) of the seven hedging strategies. One can see that the GARCH<sub>11</sub> and AGARCH<sub>11</sub> models produce average transaction costs of \$0.06 per bundle. A refinery that purchases 50,000 3:2:1 crack spread bundles per week for example, would be paying \$156,000 per year only to implement their hedging strategy. This is very large in comparison to the other models, especially the naïve strategy, where hedging does not require re-balancing and the associated margin and transaction costs are much smaller.

4.2. Hedging effectiveness

We now consider the hedging effectiveness of each model. The main question is whether the effort to implement more advanced models and the associated transaction costs pay off in a superior hedging performance? Table 6 shows the overall hedging performance measured by the unconditional *EE* of each model both, in- and out-of-sample. In the more relevant out-of-sample test, all models produce variance reductions in the range of 64–71% with the OLS<sub>11</sub> as the most effective model. The AGARCH<sub>11</sub> model performs worst, with an *EE* of 64.81%.

**Table 4**  
Out-of-sample mean and standard deviation of GARCH<sub>11</sub> and AGARCH<sub>11</sub> parameter estimates using a 260 weeks rolling window. Period: 14/01/1998–23/02/2011.

Parameter	GARCH <sub>11</sub>		AGARCH <sub>11</sub>	
	Mean	St. Dev.	Mean	St. Dev.
A <sub>11</sub>	-0.423	(0.791)	-0.292	(0.910)
A <sub>21</sub>	-0.785	(1.790)	-0.887	(2.054)
A <sub>22</sub>	0.311	(0.381)	0.163	(0.203)
B <sub>11</sub>	-0.271	(0.257)	0.058	(0.393)
B <sub>12</sub>	-1.039	(0.548)	-0.060	(0.808)
B <sub>21</sub>	0.022	(0.131)	0.031	(0.186)
B <sub>22</sub>	0.540	(0.194)	-0.466	(0.685)
C <sub>11</sub>	0.818	(0.274)	0.865	(0.354)
C <sub>12</sub>	0.158	(0.191)	1.599	(0.668)
C <sub>21</sub>	0.034	(0.148)	-0.023	(0.182)
C <sub>22</sub>	0.728	(0.136)	-0.244	(0.423)
D <sub>11</sub>	-	-	0.139	(0.224)
D <sub>12</sub>	-	-	0.271	(0.669)
D <sub>21</sub>	-	-	0.017	(0.209)
D <sub>22</sub>	-	-	-0.021	(0.421)
Log-likelihood	-1170.116	(240.992)	-1161.063	(236.468)

Fig. 4 displays the out-of-sample conditional *EE* for each model over time. It is variable throughout the sample period and occasionally reacts to the jumps in the basis. For instance, during the first quarter of 2000 the hedging effectiveness of all models drops below 0% but then rises to about 40% after about 3 months. This is due to the surge in heating oil prices (note the spike at this time in the bottom, right-hand graph in Fig. 1). We have not excluded data from this event because the price shift occurred over a period of two months, and hedging would have been necessary over such a long period. Although the GARCH models are expected to perform better under these conditions since they are more capable to react to changing market conditions, here they produce roughly the same hedging effectiveness as all the other models.

From Fig. 4 we can conclude that all models have similar effectiveness throughout the entire sample period. To test this more formally, we perform a standard F-test, for equality of variances: between the variance of P&L resulting from the naïve hedge and the P&L variance from each of the models. We use the out-of-sample P&L and evaluate the F-statistic using a rolling window to calculate the individual variances.<sup>12</sup> Fig. 5 depicts these F-statistics together with lines showing the critical values at the 90% and 95% confidence level. We fail to reject the null hypothesis that the hedge portfolio variance produced by more advanced models is significantly smaller than the naïve strategy in every instance. No model is able to improve upon the naïve hedge, utilising the 3:2:1 bundle offered by NYMEX. The same conclusion is reached for all the a:b:c crack spreads considered, although those results have not been reported for brevity.

A further robustness check was carried out to assess the dependence between the *EE* and different volatility regimes. We re-estimated the average hedging effectiveness of each model: during high volatility and low volatility time periods. As a test, the first half of the sample (sample A) was taken as the low volatility regime (30/12/1992–11/08/2004, average underlying variance 10.12 \$<sup>2</sup>), and the second half of the sample (sample B) was the high volatility regime (11/08/2004–23/02/2011, average underlying variance 58.24 \$<sup>2</sup>). Empirical results confirm that all models again produce roughly the same *EE*, as for the full-sample results. We also find that the models perform better in sample B than in sample A: in sample B the *EE* ranges from 67% to 74%.

<sup>12</sup> We have used a one year window (52 weeks) in order to obtain a long out-of-sample period. Results for longer windows (260 weeks) yield identical conclusions.

**Table 5**  
Average margin and transactions costs in \$ per spot bundle, numbers in parentheses represent standard deviations. Period 14/01/1998–23/02/2011.

	naïve	OLS <sub>13</sub>	OLS <sub>31</sub>	OLS <sub>11</sub>	EWMA <sub>11</sub>	GARCH <sub>11</sub>	AGARCH <sub>11</sub>
In-sample	0.008 (0.004)	0.010 (0.005)	0.008 (0.005)	0.009 (0.005)	0.014 (0.011)	0.024 (0.027)	0.050 (0.050)
Out-of-sample	0.008 (0.004)	0.013 (0.005)	0.009 (0.006)	0.011 (0.007)	0.014 (0.011)	0.040 (0.047)	0.059 (0.070)

**Table 6**  
Whole-sample *EE* of each model in percentage points. Hedge ratios calculated using a 260-week rolling window. Whole-sample data points: 679. Period 14/01/1998–23/02/2011. Whole-sample unhedged portfolio variance: 44.34 \$<sup>2</sup>. Hedged portfolio includes margin and transaction costs.

	naïve	OLS <sup>13</sup>	OLS <sub>31</sub>	OLS <sub>11</sub>	EWMA <sub>11</sub>	GARCH <sub>11</sub>	AGARCH <sub>11</sub>
In-sample	67.28%	69.98%	68.01%	70.47%	69.71%	67.33%	68.58%
Out-of-sample	67.28%	69.70%	67.78%	69.98%	69.71%	67.40%	64.81%

compared with 53–58% in sample A. One might suppose that this was because the correlation between the spot and futures crack spread increases during volatile times. However, upon further investigation we find that this is not the case. The GARCH correlation levels remain roughly 0.80 throughout both samples, with the exception of the period 26/01/2000–09/02/2000, where the futures-spot correlation drops to -0.12 (attributed to the heating oil surge in January–March 2000). The downward jump in the *EE* that is evident in Fig. 4 is due to this momentary shock alone. It so happens that the shock occurred during sample A, and this is the reason why models perform better during sample B. These results are available upon request.

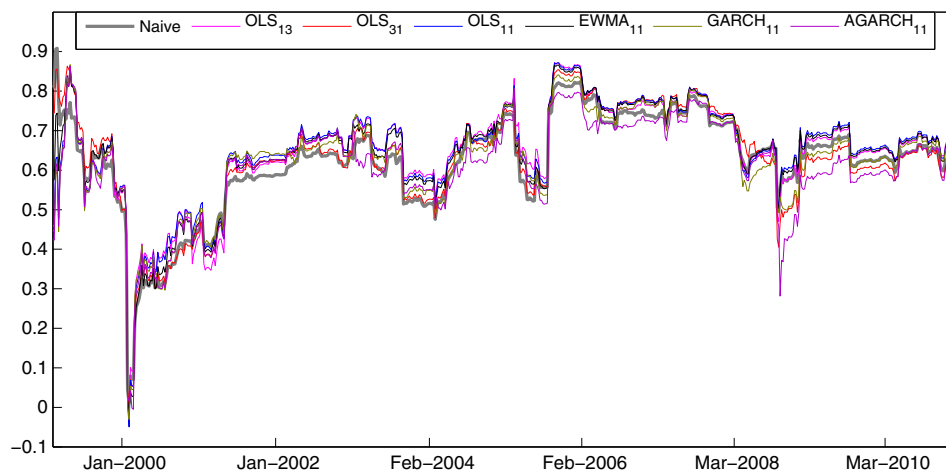
Lastly, to ensure that the hedging effectiveness is not driven by our assumption regarding the size of margin and transaction costs, we repeat the analysis, this time ignoring those costs. The change in *EE* resulting from excluding margin and transaction costs is found to be very small and mostly positive, the largest being 0.14% for the AGARCH model. One could also account for the correlation between the margin costs, or transactions costs, and the spot and futures P&L when minimising the portfolio variance, i.e. minimising the variance of  $\Delta\Pi_t^h$  in (17) as opposed to the variance of  $\Delta\Pi_t$  in (1). However, the correlations between the hedged portfolio P & L and the margin costs, or transaction costs, are very small (in the region of -0.09 to 0.03). Thus, we do not account for these correlations; it would have

only minimal effect on the empirical results, whilst augmenting the complexity of the model.

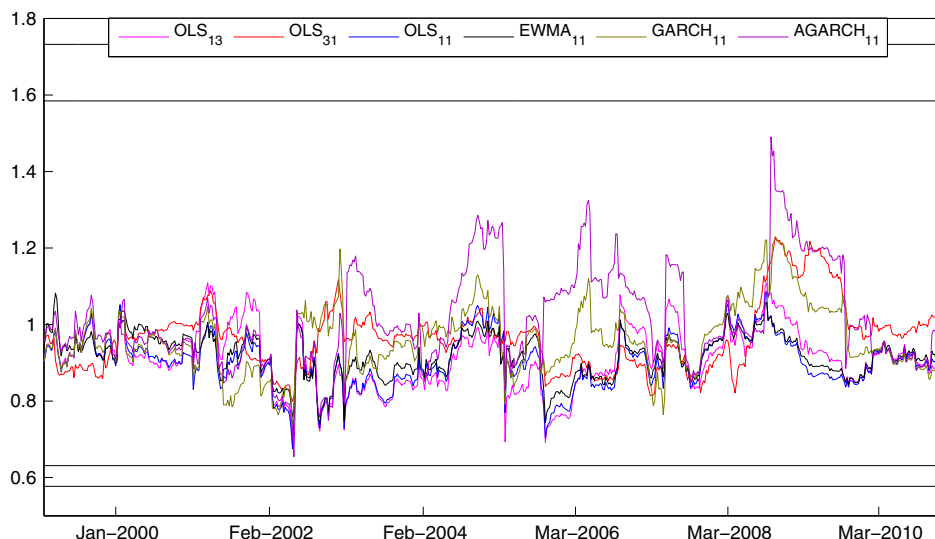
**5. Conclusions**

We have compared seven different models for estimating hedge ratios for crack spread delta hedging. Although all models are found to produce a healthy amount of variance reduction (roughly 68% on average). The most complex models (i.e. GARCH models) deliver the worst hedging results. The hedging strategies derived from GARCH models are not only more complicated to implement, they also generate the highest transaction costs. Moreover, instability in parameter estimates is another problem which can lead to unrealistically high hedge ratios associated with the most complex models.

Our findings contradict a fair body of existing literature which concludes that model-based minimum-variance hedging is superior, particularly when GARCH models are employed. In contrast, we find that the hedging effectiveness is statistically indistinguishable between all the models considered. This finding is based on a very long out-of-sample period, but we would have reached the same conclusion had we used much shorter sub-periods or, indeed, had we based conclusions on in-sample analysis alone.



**Fig. 4.** Out-of-sample analysis: EWMA *EE* of each model. Period: 13/01/1999–23/02/2011. Hedge ratios calculated using a 260-week rolling window.



**Fig. 5.** Rolling-moving  $F$ -statistic for testing the equality of variances between each hedging model and the naïve hedged portfolio. Rolling-moving variances are calculated using a 52-week rolling window. Hedge ratios calculated using a 260-week rolling window. Period: 13/01/1999–23/02/2011. Horizontal lines indicate two-sided critical values at 5% and 10% significance levels, respectively.

We have taken much more care with the data than the previous studies that have analysed the hedging of the crack spread. We use the best (Platts) spot prices and we replace the rollover log return series, which are typically used in studies of this type, by constant-maturity P&L series which are not affected by the saw-tooth pattern in the basis that biases the OLS hedge ratios. Moreover, we take meticulous care to account for all the costs involved in hedging. The margin and transaction costs of minimum-variance hedging have a very small effect on the hedging effectiveness, even for the excessively variable hedge ratios prescribed by GARCH models. However, these costs are important to analyse, because they reveal that GARCH hedging models would be too expensive to implement in practice, even if they did provide statistically significant superior performance (which they do not).

The main point for end-users to take away from our study is that, even for complex underlyings such as spreads on oil-related commodities which produce a basis that is extremely variable and jumpy, the maturity mismatch justification for minimum-variance hedging is simply not viable. The naïve hedge ratio performs as well as any other model, and it requires the least re-balancing of all. It may be that minimum-variance hedging can improve on the so-called naïve hedge when a proxy futures contract must be used – but even this remains an open question waiting for a thorough empirical analysis.

## References

- Alexander, C., Barbosa, A., 2007. Effectiveness of minimum variance hedging. *J. Portf. Manag.* 33 (2), 46–59.
- Baillie, R.T., Myers, R.J., 1991. Bivariate GARCH estimation of the optimal commodity futures hedge. *J. Appl. Econ.* 6 (2), 109–124.
- Bollerslev, T., 1982. Generalized autoregressive conditional heteroskedasticity. *J. Econ.* 31, 301–327.
- Cecchetti, S., Cumby, R., Figlewski, S., 1988. Estimation of the optimal futures hedge. *Rev. Econ. Stat.* 70, 623–630.
- Chang, C.-L., McAleer, M., Tansuchat, R., 2011. Crude oil hedging strategies using dynamic multivariate GARCH. *Energy Econ.* 33, 912–923.
- Chopra, V.K., Ziemba, W.T., 1993. The effect of errors in means variances, and covariances on optimal portfolio choice. *J. Portf. Manag.* 19 (2),

- Dunis, C., Laws, J., Evans, B., 2008. Trading futures spread portfolios: applications of higher order recurrent networks. *Eur. J. Finance* 14, 503–521.
- Ederington, L.H., 1979. The hedging performance of the new futures markets. *J. Finance* 34 (1).
- Engle, R.F., Kroner, K.K., 1995. Multivariate simultaneous generalized ARCH. *Econ. Theory* 11 (1), 122–150.
- Fama, E.F., 1965. The behaviour of stock market prices. *J. Bus.* 38, 34–105.
- Figlewski, S., 1984. Hedging performance and basis risk in stock index futures. *J. Finance* 39 (3), 657–669.
- Gagnon, L., Lupny, G.L., McCurdy, T.H., 1998. Hedging foreign currency portfolios. *J. Empir. Finance* 5, 197–220.
- Galai, N., 1979. A proposal for indexes for traded call options. *J. Finance* 34, 1157–1172.
- Grier, K.B., Henry, O.T., Olekalns, N., Shields, K., 2004. The asymmetric effects of uncertainty on inflation and output growth. *J. Appl. Econ.* 19, 551–565.
- Haigh, M.S., Holt, M.T., 2000. Hedging multiple price uncertainty in international grain trade. *Am. J. Agric. Econ.* 82 (4), 881–889.
- Haigh, M.S., Holt, M.T., 2002. Crack spread hedging: accounting for time varying volatility spillovers in the energy futures markets. *J. Appl. Econ.* 17, 269–289.
- Herbst, A.F., Kare, D.D., Caples, S.C., 1989. Hedging effectiveness and minimum risk hedge ratios in the presence of autocorrelation: foreign currency futures. *J. Finance Eng.* 9 (3), 185–197.
- Hill, J., Schneeweis, T., 1982. The hedging effectiveness of foreign currency futures. *J. Finance Res.* 5, 95–104.
- Ji, Q., Fan, Y., 2011. A dynamic hedging approach for refineries in multiproduct oil markets. *Energy* 36, 881–887.
- Johnson, L.L., 1960. The theory of hedging and speculation in commodity futures. *Rev. Econ. Stud.* 27 (3), 139–151.
- Kroner, K.F., Sultan, J., 1993. Time-varying distributions and dynamic hedging with foreign currency futures. *J. Financ. Quant. Anal.* 28 (4), 535–551.
- Lee, H.-T., 2009. Optimal futures hedging under jump switching dynamics. *J. Empir. Finance* 16, 446–456.
- Lee, H.-T., 2010. Regime switching correlation hedging. *J. Empir. Finance* 34, 2728–2741.
- Lee, H.-T., Yorder, J.K., 2007. A bivariate markov regime switching GARCH approach to estimate time varying minimum variance hedge ratios. *Appl. Econ.* 39, 1253–1265.
- Lien, D., 2005. The use and abuse of the hedging effectiveness measure. *Int. Rev. Financ. Anal.* 14 (2), 277–282.
- Lien, D., 2008. Asymmetric effect of basis on dynamic futures hedging: empirical evidence from commodity markets. *J. Bank. Finance* 32, 187–198.
- Lien, D., 2009. A note on the effectiveness of GARCH models. *Int. Rev. Econ. Finance* 18, 110–112.
- Nguyen, V.T.T., Carbonez, K.A.E., Sercu, P., 2011. Hedging with two futures contracts: simplicity pays. *Eur. Financ. Manag.* 17, 806–834.
- Stein, J.L., 1961. The simultaneous determination of spot and futures prices. *Am. Econ. Rev.* 51, 1012–1025.
- Working, H., 1953. Futures trading and hedging. *Am. Econ. Rev.* 43 (3), 314–343.