# Markov Switching GARCH models for Bayesian Hedging on Energy Futures Markets

Monica Billio† Roberto Casarin† Anthony Osuntuyi†

February 2013

**Abstract:** We propose Bayesian Markov Switching Generalized Autoregressive Conditional Heteroscedasticity (MS-GARCH) models for determining time-varying Minimum Variance (MV) hedge ratio in energy futures markets. We apply an efficient simulation based technique for inference and suggest a robust hedging strategy which accounts for model parameter uncertainty. The hedging model is further applied to crude oil and gasoline spot and futures markets.

**Keywords:** Markov Switching; Hedge ratio; Energy futures; GARCH.

## 1. Introduction

Hedging is an investment position taken by investors to mitigate the adverse effect arising from changes in the price of a companion investment. A crucial issue, which has been subject to both theoretical discussions and econometric specifications, is the determination of the optimal hedge ratio, i.e. the number of derivative contracts to buy (or sell) for each unit of the underlying asset on which the investor bears risk. See Chen *et al.* [2003] for a review. In this paper, we focus on the econometric specification and estimation procedure of the optimal hedge ratio proposed by Johnson [1960] and called Minimum Variance (MV) hedge ratio.

The MV hedge ratio is defined as the ratio of the covariance between the underlying spot and futures returns to the variance of the futures return. To apply this optimum hedge ratio in practice, Ederington [1979] suggests regressing the underlying spot returns against the futures returns and to use the estimate of the slope as MV hedge ratio. This approach has been widely criticized on the grounds that some of the well known stylized facts about asset returns are ignored. For example, it is well known that asset returns are usually not strictly stationary. To this end and to improve hedging performance (Myers [1991]), time-varying hedge ratios are proposed in the literature.

Two main approaches have been developed in the literature to estimate time-varying MV hedge ratios. One approach involves the estimation of the conditional second order moment of the underlying and futures returns captured by Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models. See Haigh and Holt [2002], Chang *et al.* [2010] among others for illustration. The later approach treats the hedge ratio as a time-varying regression coefficient and focuses on the estimation of such a parameter (Lee *et al.* [2006], Chang *et al.* [2010] e.t.c.). Note that this hedge ratio works by re-balancing the hedged portfolio on a period by period basis. This may involve huge transaction costs and therefore it may not be worthwhile using this particular instrument for hedging. Also, it has been well documented in the empirical literature that the class of GARCH models exhibit high persistence of conditional variance, i.e. the process is close to being nearly integrated. In view of this, a few authors allow the optimal hedge ratio to be state-dependent. In line with the first approach of estimating time-varying hedge ratios, Alizadeh *et al.* [2008], Lee and Yoder [2007a], Lee and Yoder [2007b] among others propose various regime switching models. These models differ mainly by the

†Address: Department of Economics, University Ca' Foscari of Venice, Fondamenta San Giobbe 873, 30121, Venice, Italy. Corresponding author: Anthony Osuntuyi, aosuntuyi@stud.unive.it. Other contacts: billio@unive.it (Monica Billio); r.casarin@unive.it (Roberto Casarin).

characterization of time-varying covariance matrix. Alizadeh and Nomikos [2004] on the other hand follow the second approach and estimate time-varying hedge ratio by specifying a Markov switching variance model. More specifically, the time-varying variance is defined as an exponential function of the lagged 4-week moving average of the difference between the logarithm of the underlying and the logarithm of the futures. They further test their model with the FTSE 100 and S&P 500 indexes data and find that the Markov switching approach can improve hedging performance in terms of variance reduction and utility maximization.

Our contribution to the literature on time-varying hedge ratio is two-folds. First, we propose a MS-GARCH approach. A direct translation of Bollerslev [1986] GARCH model to regime switching setting is known to be affected by path dependence which makes parameter estimation a challenging exercise. Based on this, Gray [1996] propose a generalized version of the MS-GARCH models which avoids the path dependence structure inherent in the former. To this end, a commonly used model in the hedging literature is the multivariate extension of Gray [1996] model. While Gray's model is attractive, its analytical intractability is a drawback. We differ from this approach by taking the direct translation of Bollerslev [1986] GARCH model to a regime switching setting as given. We address the inherent path dependence problem by following a Bayesian approach based on Markov Chains Monte Carlo (MCMC). The estimation exercise is efficiently carried out by following Billio *et al.* [2012] technique for efficiently sampling the state variable trajectory. The second contribution is the use of robust hedging that accounts for parameter uncertainty. Most studies on hedging are empirically implemented by using the 'plug-in' principle i.e. once the optimal decision hedge ratio is derived, sample estimates of the model's parameters are substituted directly. This approach totally ignores parameter estimation risk. We follow the simple Bayesian model averaging procedure as illustrated in Lence and Hayes [1994a] and Lence and Hayes [1994b] for incorporating model uncertainty into the hedging strategies. In the empirical applications, we also efficiently estimate time-varying MV hedge ratios for crude oil and gasoline spot and futures prices used in Chang *et al.* [2010] and compare the result to conventional OLS method proposed by Ederington [1979].

The structure of this paper is as follows. In the next section, we present the MV hedge ratio methodology and illustrate the empirical model that we used in this paper followed by the estimation technique. In Section 3, we illustrate with real data how the proposed model is implemented. Section 4 concludes the paper and provides possible extensions.

## 2. Measuring the Optimal Hedge Ratio

Johnson [1960] derives the optimal hedge ratio by minimizing the variance of the hedged portfolio. Let $RS_t$ and $RF_t$ represent the underlying spot and futures returns at a given time, $t$, respectively. Then, the optimal hedge ratio, $h$, is given as

$$h = \frac{Cov(RS_t, RF_t)}{Var(RF_t)}. \tag{1}$$

The assumption of constant variance and covariance implies a time-invariant hedge ratio which suggests the use of OLS for its estimation,

$$RS_t = \mu + \nu RF_t + \epsilon_t \qquad \epsilon_t \overset{iid}{\sim} (0, \sigma^2), \tag{2}$$

where $\mu$, $\nu$ and $\sigma$ are the regression parameters. The coefficient of $RF_t$, $\nu$, estimated by ordinary least squares (OLS) is the MV hedging ratio.

Equation (2) represents an oversimplification of assets returns dynamics because it ignores some of the well known stylized facts, such as conditional heteroscedasticity and volatility clustering, commonly observed in financial data. In view of this and to allow for changes in the market conditions to affect the hedge ratios, Equation (2) is extended to an $M$ state Markov switching model with a time-varying volatility process also characterized by regime switching. Let $s_t$ be a discrete, unobserved, state variable which could be interpreted as the state of the world at time $t$. Then,

$$
\begin{aligned}
RS_t &= \mu(s_t) + \nu(s_t)RF_t + \sigma_t\eta_t \qquad \eta_t \overset{iid}{\sim} \mathcal{N}(0,1), \\
\sigma_t^2 &= \gamma(s_t) + \alpha(s_t)\epsilon_{t-1}^2 + \beta(s_t)\sigma_{t-1}^2, \\
RF_t &= a(s_t) + \tau_t\zeta_t \qquad \zeta_t \overset{iid}{\sim} \mathcal{N}(0,1), \\
\tau_t^2 &= \kappa(s_t) + \omega(s_t)\xi_{t-1}^2 + \psi(s_t)\tau_{t-1}^2,
\end{aligned} \tag{3}
$$

where, $\epsilon_t = \sigma_t\eta_t$, $\xi_t = \tau_t\zeta_t$, $\mu(s_t)$, $\nu(s_t)$, $\gamma(s_t) > 0$, $\alpha(s_t) \geq 0$, $\beta(s_t) \geq 0$, $a(s_t)$, $\kappa(s_t) > 0$, $\omega(s_t) \geq 0$, $\psi(s_t) \geq 0$, and $s_t \in \{1, \ldots, M\}$, $t = 1, \ldots, T$, is assumed to follow a $M$-state first order Markov chain with transition probabilities $\{\pi_{ij}\}_{i,j=1,2,\ldots,M}$:

$$
\pi_{ij} = p(s_t = i | s_{t-1} = j), \quad \sum_{i=1}^{M} \pi_{ij} = 1 \quad \forall\, j = 1, 2, \ldots, M. \tag{4}
$$

The parameter shift functions $\mu(s_t)$, $\nu(s_t)$, $a(s_t)$, $\gamma(s_t)$, $\alpha(s_t)$, $\beta(s_t)$, $\kappa(s_t)$, $\omega(s_t)$ and $\psi(s_t)$ describe the dependence of parameters on the realized regime $s_t$ e.g.

$$
\mu(s_t) = \sum_{m=1}^{M} \mu_m \mathbb{I}_{s_t=m}, \qquad \text{with,} \quad \mathbb{I}_{s_t=m} = \begin{cases} 1, & \text{if } s_t = m \\ 0, & \text{otherwise.} \end{cases}
$$

Given information up to time $t-1$, the conditional hedge ratio at time $t$, within this setting, is given by

$$
h_t | \Omega_{t-1} = \frac{Cov(RS_t, RF_t | \Omega_{t-1})}{Var(RF_t | \Omega_{t-1})}, \tag{5}
$$

where $\Omega_{t-1}$ denotes the information set available up to time $t$,

$$
\begin{aligned}
Cov(RS_t, RF_t|\Omega_{t-1}) &= Cov(\mu(s_t), RF_t|\Omega_{t-1}) + Cov(\nu(s_t)RF_t, RF_t|\Omega_{t-1}) \\
&= Cov(\mu(s_t), a(s_t)|\Omega_{t-1}) + E[\nu(s_t)(\tau_t^2 + a(s_t)^2)|\Omega_{t-1}] \\
&- E[\nu(s_t)a(s_t)|\Omega_{t-1}]E[a(s_t)|\Omega_{t-1}], \\
\text{and} \quad V(RF_t|\Omega_{t-1}) &= V(a(s_t)|\Omega_{t-1}) + E[\tau_t^2|\Omega_{t-1}].
\end{aligned} \tag{6}
$$

The model parameters in Equation (5) are often not known with certainty. Based on this, we derive the optimal hedge ratio under a simple Bayesian model averaging (see Lence and Hayes [1994a]) to account for the parameter uncertainty to obtain

$$
h_t | \Omega_{t-1} = \frac{E_\theta[Cov(RS_t, RF_t | \Omega_{t-1})]}{E_\theta[Var(RF_t | \Omega_{t-1})]}, \tag{7}
$$

where $E_\theta[\cdot]$ is the expectation operator with respect to the posterior distribution of the parameter $\theta$ of the model. Observe that the optimal hedge ratio given in Equation (7) does not only depend on the coefficient of $RF$, $\nu(s_t)$, as it is the case in the conventional OLS approach, a precise Econometric specification of the return process on futures is required.

## 2.1 Estimating time-varying MV hedge ratio

Following Billio *et al.* [2012], we describe an efficient simulation based technique for Bayesian inference on the proposed hedging model. The Bayesian approach is based on MCMC Gibbs algorithm which allows us to circumvent the path dependence problem and efficiently sample the state trajectory. The purpose of this algorithm is to generate samples from the posterior distribution which are then used for its characterization. We follow a data augmentation framework by treating the state variables as parameters of the model and construct the completed likelihood function assuming the states are known. Let $s_{s:t} = (s_s, \ldots, s_t)$, $RS_{s:t} = (RS_s, \ldots, RS_t)$, $RF_{s:t} = (RF_s, \ldots, RF_t)$ whenever $s < t$, $\theta_\pi = (\pi_{11}, \ldots, \pi_{M1}, \ldots, \pi_{1M}, \ldots, \pi_{MM})$, $\theta_u^{RS} = (\mu_1, \ldots, \mu_M, \nu_1, \ldots, \mu_{2,M})$, $\theta_a^{RF} = (a_1, \ldots, a_M)$, $\theta_\sigma = (\gamma_1, \ldots, \gamma_M, \alpha_1, \ldots, \alpha_M, \beta_1, \ldots, \beta_M)$, $\theta_\tau = (\kappa_1, \ldots, \kappa_M, \omega_1, \ldots, \omega_M$

$, \psi_1, \ldots, \psi_M)$ and $\theta = (\theta_\pi, \theta_u^{RS}, \theta_a^{RF}, \theta_\sigma, \theta_\tau)$. We assume fairly informative prior for $\theta_\pi$ and independent uniform prior for $\theta_u^{RS}$, $\theta_a^{RF}$, $\theta_\sigma$ and $\theta_\tau$ and denote with $f(\theta)$ the joint prior density. To avoid label switching we assume that $\gamma_1 < \gamma_2 < \cdots < \gamma_M$ i.e. identifiability restriction. The posterior density of the augmented parameter vector given by

$$f(\theta, s_{1:T}|RS_{1:T}, RF_{1:T}) \propto f(RS_{1:T}|s_{1:T}, \theta, RF_{1:T}) f(RF_{1:T}|s_{1:T}, \theta) p(s_{1:T}|\theta) f(\theta) \tag{8}$$

cannot be identified with any standard distribution, hence we cannot sample directly from it. Our Gibbs sampler generate samples from posterior distribution, by iteratively sampling from the following full conditional distributions:

- $p(s_{1:T}|\theta, RS_{1:T}, RF_{1:T})$,
- $f(\theta_\pi|\theta_u^{RF}, \theta_a^{RS}, \theta_\sigma, \theta_\tau, s_{1:T}, RS_{1:T}, RF_{1:T}) = f(\theta_\pi|s_{1:T})$, and
- $f(\theta_u^{RS}, \theta_a^{RF}, \theta_\sigma, \theta_\tau|\theta_\pi, s_{1:T}, RS_{1:T}, RF_{1:T}) = f(\theta_u^{RS}, \theta_a^{RF}, \theta_\sigma, \theta_\tau|s_{1:T}, RS_{1:T}, RF_{1:T})$.

The full joint distribution of the state variables, $s_{1:T}$, given the parameter values and return series

$$p(s_{1:T}|\theta, RS_{1:T}, RF_{1:T}) \propto f(RS_{1:T}|RF_{1:T}, \theta, s_{1:T}) f(RF_{1:T}|\theta, s_{1:T}) \tag{9}$$

is a non-standard distribution. We consider a Metropolis Hastings (MH) strategy for generating proposals for the state variables. We construct the proposal distribution by first considering an analytical approximation of the regime switching GARCH model and then derive the joint distribution of the state variables. See Billio *et al.* [2012] for alternative approximations. For expository purpose, we apply the Basic approximation given in Billio *et al.* [2012]. Samples of the state trajectory are then drawn by Forward Filter Backward sampling scheme.

While the full conditional of $\theta_\pi$ is Dirichlet under Dirichlet prior distribution assumption, the posterior density of $(\theta_u^{RS}, \theta_a^{RF}, \theta_\sigma, \theta_\tau)$

$$\begin{aligned} &f(\theta_u^{RS}, \theta_a^{RF}, \theta_\sigma, \theta_\tau|s_{1:T}, RS_{1:T}, RF_{1:T}) \\ &\propto \prod_{t=1}^{T} \frac{1}{\sigma_t} \exp\left(\frac{(RS_t - \mu(s_t) - \nu(s_t) RF_t)^2}{\sigma_t^2}\right) \prod_{t=1}^{T} \frac{1}{\tau_t} \exp\left(\frac{(RF_t - a(s_t))^2}{\tau_t^2}\right) \end{aligned} \tag{10}$$

is non-standard. Hence, we apply adaptive Metropolis-Hastings (MH) sampling technique for this step of the Gibbs algorithm.

## 3. Empirical Applications

To illustrate the proposed method, we use daily closing energy prices for West Texas Intermediate (WTI) crude oil futures and gasoline futures for the period January 1, 1996 to December 31, 2005 (2500 observations). Both spot and futures daily settlement prices are available from the database of the Commodity Research Bureau. Futures contracts with a 1 week rolling period prior to expiration of the current contracts to the next contract are used to circumvent the effects of thin markets and expiration. The daily returns are computed using the first difference of the natural logarithm of the daily settlements. Figure (1) displays the sample path of the crude oil and gasoline squared returns on spot and futures. We observe volatility clustering, which calls for the use of MS-GARCH models. The estimated model considered is a two regime ($M = 2$) restricted version ($a(s_t) = a$ and $\tau_t = \tau$) of the MS-GARCH model. In this case, the hedge ratio reduces to

$$h_t|\Omega_{t-1} = E_\theta E[\nu(s_t)|\Omega_{t-1}] = E_\theta \left[\sum_{m=1}^{M} \nu_m p(s_t = m|\Omega_{t-1})\right], \tag{11}$$

and estimated using,

$$h_t|\Omega_{t-1} = \frac{1}{G} \sum_{i=1}^{G} \sum_{m=1}^{M} \nu_m^{(i)} p^{(i)}(s_t = m|\Omega_{t-1}), \tag{12}$$

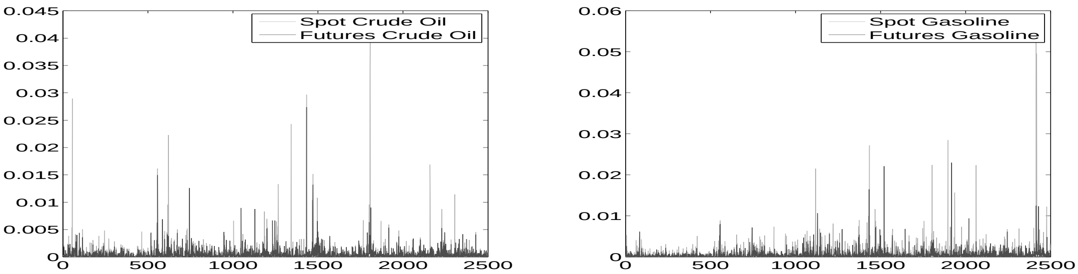


**Figure 1:**: Graphs for daily squared returns on WTI crude oil and gasoline spot and futures from 01/01/1996 to 31/12/2005

where $G$ is the number of Gibbs samples. In this paper, the estimate of the hedge ratio is based on $G = 15000$ Gibbs samples. Figure (2) displays the dynamics of estimated hedge ratio for both the MS-GARCH (blue line) and OLS (red line) models. We find that 95% of the time-varying hedge ratio lies within the credible interval of $(0.31, 0.94)$ for crude oil and (0.86, 1.10) for gasoline. The estimates of the hedge ratios, $\nu(s_t)$, of being state 1 is 0.9981 (1.1530) with standard deviation 0.0116 (0.0075) and 0.2006 (0.5266) with standard deviation 0.0508 (0.0423) for being in state two for crude oil (gasoline). The OLS hedge ratio of Equation (2) lies between these two hedge ratios for both commodities. The probability of the hedge ratio for crude oil for staying lower volatility regime is estimated to be $0.94$ with standard deviation $0.0077$ and $0.87$ with standard deviation $0.0787$ for staying in the higher volatility regime. In the case of gasoline, the probability of staying in the lower volatility regime is $0.92$ with standard deviation $0.0082$ and $0.49$ with standard deviation $0.0285$ for remaining in the higher volatility regime. In both cases, the probability of staying in regime one is very high suggesting a low transaction cost because the investor only needs to reblance his portfolio ocassionaly. Following the estimation of the MS-GARCH and subsequently the hedge ratio using Equation (12), we formally assess the performance of these hedges by first constructing the portfolio implied by the computed hedge ratios daily. Then we calculate the variance of the returns of these portfolios over the sample. In mathematical forms, we evaluate

$$Var(RS_t - h_t^* RF_t) \tag{13}$$

where $h_t^*$ are the estimated hedge ratios. The incremental variance improvement of the MS-GARCH model against the OLS model, using Equation (2), is calculated as follows

$$\frac{Var(\text{OLS}) - Var(\text{MS-GARCH})}{Var(\text{OLS})}, \tag{14}$$

where $Var(\text{OLS})$ and $Var(\text{MS-GARCH})$ are respectively the variance of the returns on the hedged portfolio (Equation (13)) estimated using hedge ratios obtained from the OLS and MS-GARCH models. For crude oil, the MS-GARCH model provides a better variance reduction ($7\%$) when compared to the OLS model. Whereas, the OLS model performs slightly better ($0.84\%$) than MS-GARCH model in terms of risk reduction for gasoline.

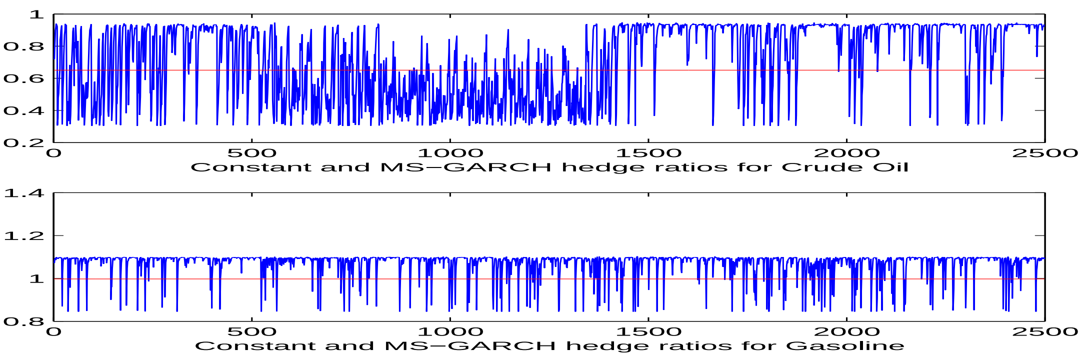


**Figure 2:**: Graphs of estimated hedge ratios from 01/01/1996 to 31/12/2005

## 4. Conclusion

In this paper, we present and examine the performance of a robust Bayesian MS-GARCH model for determining the time-varying hedge ratios in energy futures markets. The use of this type of model is suggested by the fact that the dynamics of both spot and futures returns may be affected by changes in the state of the world, conditional heteroscedasticity and volatility clustering. These features suggest that allowing the hedge ratios to depend on the state of the market may produce more efficient and perhaps superior hedging performance when compared to the conventional method. Hedging strategies in practice are mostly implemented by using the plug-in principle. This approach entirely neglects estimation risk. We account for the parameter uncertainty by considering a Bayesian model averaging approach. We have considered a very restricted version of the MS-GARCH model in our empirical applications for which we obtained mixed results. It is our opinion that if we allow the dynamics of the futures returns to follow an MS-GARCH model a better result in terms of efficiency and performance may be obtained. The in-sample performance of the hedging strategies consider in this paper gives an indication of historical performance. However, investors are more interested in how well they can forecast using different hedging strategies. In this respect, it is our plan to address this issue in further research.

## References


Alizadeh A.H. and Nomikos N.K. (2004) A Markov regime switching approach for hedging stock indices, *Journal of Futures Markets*, 24, 649 – 674.

Alizadeh A.H., Nomikos N.K. and Pouliasis P.K. (2008) A Markov regime switching approach for hedging energy commodities, *Journal of Banking & Finance*, 32, 1970 – 1983.

Billio M., Casarin R. and Osuntuyi A. (2012) Efficient Gibbs sampling for Markov switching GARCH models.

Bollerslev T. (1986) Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307–327.

Chang C.Y., Lai J.Y. and Chuang I.Y. (2010) Futures hedging effectiveness under the segmentation of bear/bull energy markets, *Energy Economics*, 32, 442 – 449.

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

Ederington L. (1979) The hedging performance of the new futures markets, *Journal of Finance*, 34(1), 157 – 170.

Gray S.F. (1996) Modeling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics*, 42, 27–62.

Haigh M.S. and Holt M.T. (2002) Crack spread hedging: accounting for time-varying volatility spillovers in the energy futures markets., *Journal of Applied Econometrics*, 17, 269 – 289.

Johnson L.L. (1960) The theory of hedging and speculation in commodity futures., *Review of Economic Studies*, 27(3), 139 – 151.

Lee H. and Yoder J.K. (2007a) A bivariate Markov regime switching GARCH approach to estimate time varying minimum variance hedge ratio, *Applied Economics*, 39(10), 1253 – 1265.

Lee H. and Yoder J.K. (2007b) Optimal hedging with a regime switching time varying correlation GARCH model, *Journal of Futures Markets*, 27(5), 495 – 516.

Lee H.T., Yoder J.K., Mittelhammer R.C. and McCluskey J.J. (2006) A random coefficient autoregressive Markov regime switching model for dynamic futures hedging, *The Journal of Futures Markets*, 26(2), 103 – 129.

Lence S.H. and Hayes D.J. (1994a) The empirical minimum-variance hedge, *American Journal of Agricultural Economics*, 76(1), 94 – 104.

Lence S.H. and Hayes D.J. (1994b) Parameter-based decision marketing under estimation risk: An application to futures trading, *Journal of Finance*, 49(1), 345 – 357.

Myers R. (1991) Estimating time-varying optimal hedge ratios on futures markets, *Journal of Futures Markets*, 11, 39 – 53.