FINANCIAL OPTIONS RESEARCH CENTRE

University of Warwick

Modelling Commodity Futures Spreads: An Empirical Study

Leticia Veruete

and

Stewart Hodges

August 1997

Financial Options Research Centre
Warwick Business School
University of Warwick
Coventry CV4 7AL
Phone: (01203) 524118

FORC Preprint: 97/83

Modelling commodity futures spreads: an empirical study*

Leticia Veruete
PhD Candidate
University of Warwick
Business School
Coventry CV4 7AL
UK

and

Stewart Hodges
Financial Options Research Centre
Warwick Business School
Coventry CV4 7AL
UK

August 1997

*The authors thank Jeffrey Williams for insightful comments on an earlier draft of this paper. Helpful suggestions of Stephen Taylor and other participants of the Annual European Research Symposium of the Chicago Board of Trade are also appreciated.

The first author acknowledges the financial support from the Mexican Institution, Conacyt.

Abstract

Understanding the behaviour of commodity futures spreads is important for traders as well as consumers and producers. However, relatively little research has been done in modelling commodity futures spreads.

On the one hand, a state dependent variance function (nonlinear) for commodity spread has been theoretically acknowledged to be important. On the other hand, commodity spreads are bounded by full carrying charges. However, there is no empirical study which considers simultaneously these two features.

In this paper, a new model which incorporates the above features is fitted empirically. It is hypothesised that a model with a quadratic variance function on spread size and with lower and upper bounds in spreads can capture the observed behaviour of commodity futures spreads. It has the notable advantage of being consistent with no-arbitrage under upper and lower bound constraints.

Empirical results, using commodity futures prices from different American exchanges for the period 1985 to 1995, suggest that the proposed empirical model fits well for most commodities when using weekly data. However, the results do not always support the model, especially when daily data are used.

Modelling commodity futures spreads: an empirical study

I. Introduction

Understanding commodity price dynamics and its volatility is of great importance for different economic agents. First, agents want to know prices at different periods ahead to plan efficiently their production, consumption and storage decisions. Second, hedgers and speculators in futures and option markets need to understand the evolution of future commodity prices and the dynamics of the commodity price volatility.

Theoretically, the competitive storage model has shown that the behaviour of commodity prices is not compatible with the random walk hypothesis. Instead, commodity prices respond asymmetrically to supply shocks. This is because inventories can cushion price falls, but can do little to avoid price increases when inventories are exhausted. The presence of stockouts then creates a nonlinear response in commodity prices and affects the variance of the commodity price, (see Deaton and Laroque, 1992 and Williams and Wright, 1991). Two immediate implications follow for the variance of commodity prices when the nonnegativity constraint on stocks is imposed: i) the variance of the commodity price is conditional on current information summarised by availability at every date and ii) this variance has a nonlinear relationship with availability.

Another important feature of commodity markets is that prices of different maturities are intertemporally related by arbitrage, provided inventories are available. In that case, to avoid arbitrage profits, the difference between the futures prices of two different maturities, given by the spread, is bounded by the cost of carrying charges.[1]

On one side, great effort has been devoted to modelling commodity prices with a time varying variance and an asymmetric response in prices, (see Trivedi, 1995). Brunetti and Gilbert (1995) considered a state dependent variance for metal prices. Recently, Pirrong (1997) estimated a bivariate model for oil futures prices of different maturities in which the variance of each futures price depends on the spread size and past variances. Nevertheless, there is no empirical study in the literature which

[1] The cost of carrying charges (cost of carry) includes storage costs and financial costs involved.

considers a state dependent variance function for commodity futures spreads in the way that is proposed in this paper.

On the other side, in a different area of the literature, option pricing models for commodity derivative securities have generally treated the volatility of the commodity price as constant, (some exceptions are Shimko, 1994 and Kang and Brorsen, 1995). Grabbe (1995) proposed an option pricing model for commodity spreads. His contribution is to take into account the full carrying charges bound observed in spreads. However, this model has the disadvantages that it does not allow for a state dependent variance function and that it admits arbitrage.

In this paper a new empirical model, based on a theoretical framework developed by Rady's (1995), is implemented for modelling the evolution of commodity futures spreads. Unlike previous studies, this model considers a dynamic specification which simultaneously takes into account a state dependent variance (nonlinear) function and the bound imposed by full carrying charges in spreads. Furthermore, this dynamic specification has the appealing feature that it has a closed form solution for a call option on the underlying variable, within a no-arbitrage option pricing model, (see Rady, 1995).

This paper is organised as follows. In the next section, a background and motivation for modelling commodity spreads is provided. Section three explains the empirical model, based on Rady's (1995) framework. A description of the data is provided in section four. Section five contains the estimation results and analysis. Finally, concluding remarks are given.

II. Background and motivation

It is known that the cost-of-carry model establishes an intertemporal relationship between the spot price (S_t) at time t and futures price ($F_{t,T}$) with delivery date T, through stockholding of the physical commodity. To avoid arbitrage profits $F_{t,T} = S_t + k_t$ at any date in time t, where k_t is the cost of carry from time t to delivery date T. An immediate implication is that this relationship ensures that futures prices will be bounded by the spot price plus the cost of carry. In addition, under no-arbitrage, there is a risk-neutral measure such that the futures price follows a martingale process.

Consider two futures contracts, F_1 and F_2 on the same commodity, but with different maturities T_1 and T_2 respectively where $T_2 > T_1$. It follows that,

$F_1 = S + k_1$ and

$F_2 = S + k_2$

where k_1 and k_2 represent the cost of carry for each futures contract respectively, and where $k_2 > k_1$ because a longer time to maturity implies a higher cost of carry of the commodity.

The relationship between the prices of contracts with different maturities, given by the spread, can be expressed as follows:

$$F_1 - F_2 = -(k_2 - k_1) \quad (1)$$

where $-(k_2 - k_1)$ is the cost of carry for the time interval between T_1 and T_2. Several points are worth noting. First, it is possible to see in (1) that when the first nearby futures price, F_1, is at discount over the second nearby (or more distant) futures contract price, F_2, the spread is negative reflecting a contango situation. By contrast, a backwardation situation is observed when the first nearby futures price is at premium over the immediate futures contract price, and the spread is positive.

Second, expression (1) reflects the fact that theoretically spreads (or contango) cannot increase above the full carrying charges between T_1 and T_2, otherwise arbitrage profits will exist. In this context, the maximum level for contango allowed reflects the lower bound in spreads, $-(k_2 - k_1)$.

Third, in presence of stockouts commodities cannot be arbitraged, an important point of difference with financial assets. Therefore, backwardation can increase and reach very high levels, nonetheless it cannot continue to increase forever. This is because a substitution effect switches consumption of the exhausted commodity to another, reducing the high demand for the depleted commodity. Under these circumstances there will be an upper bound associated with the highest backwardation level.

Fourth, provided that F_1 and F_2 follow martingale processes, then the spread between futures prices of different maturities, will follow a martingale process as well. Even under the objective probability it is expected that the spread follows a martingale since the drift will be small.

Finally, because nearer the expiration date the cost of carry reduces, spreads should narrow at the delivery date. But during the futures contracts' trading life other factors besides the cost of carry affect the behaviour of spreads and their volatility. For example, the competitive storage model has shown spreads are state dependent on availability, which consists of the amount of supply at date *t* and the inventories carried from previous date (see Deaton and Laroque, 1992 and Williams and Wright, 1991). This means that spreads contain information about the level of availability. That is, when availability is high (either high supply and/or abundant inventories) the likelihood of having a stockout next date is very low. In that situation the first nearby futures price and a more distant futures price will move in line together. As a consequence, spreads narrow and they will display a low volatility. However, when availability is very low (either low supply or very low inventories) the probability of a stockout occurring will be high. The first nearby futures price increases more than more distant futures prices. More distant futures prices are less affected by supply/demand shocks because in the long run inventories can be accumulated or supplies can be adjusted. In these circumstances spreads can widen dramatically, becoming very volatile, (see Ng and Pirrong, 1994). This reflects the fact that the variance of commodity spreads depends in a nonlinear manner on availability, the latter captured by the spread size.

Any model which aims at capturing the behaviour of commodity futures spreads needs to take the above theoretical regularities into account. This paper argues that Rady's (1995) framework represents an appropriate theoretical model which precisely incorporates the above features identified in commodity futures spreads. Moreover, as far as we know this framework is the first one to consider these features. Furthermore, Rady's specification has the advantages of providing closed form solution to option prices and encompassing other models such as Black and Scholes' (1973) and Merton's (1973).

Specifically, Rady's framework ensures that in presence of a lower and upper bounds and no-arbitrage, the underlying variable can be martingale which respects the bounds: the variance of the underlying variable goes to zero on the boundaries. In other words, because under no-arbitrage the change of measure from the objective to risk neutral has to satisfy some kind of boundedness (Lipschitz) condition, the property that the variance is zero on the boundary is preserved under the objective measure as well. For example, if we consider a stochastic process in which its underlying variable, S, follows $dS = \frac{k}{S}dt + \sigma dZ$, although S cannot fall below zero because of the form of the drift, the model nevertheless admits arbitrage.

In what follows it is explained how the presence of the lower and upper bounds, under a martingale process for spreads, accords with a quadratic variance function on spread size. Intuitively, once the spread has reached the lower bound, it is expected that its immediate future level will remain at the same level given the constraint of cash-and-carry arbitrage, and that the spread will not decrease below this arbitrage bound. Similarly, at the upper bound when the spread is close to its maximum backwardation level substitution effects between commodities ensure that this high level is unlikely to continue to increase. In that case, at the boundaries spreads should be very close to each other, and hence the volatility of spread returns goes to zero. However, when spreads are not binding (in between the bounds), spreads are expected to have complete freedom to vary since availability changes over time accordingly to supply conditions and inventory levels. It is possible to be in contango and switch to backwardation when a stockout occurs randomly.

Figure 1 displays a quadratic relationship between the spread returns volatility and spread size, in which the lower bound (*l*) is associated with arbitrage full carrying charges bound and the upper bound (*u*) is the highest backwardation level.

[Insert Figure 1]

III. A model for commodity futures spreads

Based on Rady's (1995) framework, it is hypothesised that the relative futures spread[2], $X = {}^{F_1}\!/_{F_2}$, solves the following stochastic differential equation;

$$dX_t = \sigma\ (X_t - l)\ (1 - \tfrac{1}{u} X_t)\ dW_t \qquad (2)$$

with initial value X_0, $l < X_0 < u$, $\sigma > 0$ and $0 \le l < u \le +\infty$ are given. l and u stand for the lower and upper bounds respectively.

---

[2] **Altough the spread is given by the difference between two futures prices of different maturities, in this study the relative futures spread is calculated as the ratio of the first nearby futures contract price and the second nearby futures contract price.**

The term $\sigma_{q,t} \equiv \sigma\ (X_t - l)\ (1 - \frac{1}{u} X_t)$ is the instantaneous standard deviation of the process and W_t is a Wiener process.[3] The two terms together reflect the noise term or stochastic component of the futures spread returns.

The above stochastic process represents a martingale without drift and its standard deviation has a quadratic form in the spread size represented in terms of two bounds.

To ensure a nonnegative standard deviation in (2), this diffusion term can be re-expressed as its absolute value in the following way:

$$\sigma_{q,t} = \sigma \left| -l + (1 + \tfrac{l}{u}) X_t - \tfrac{1}{u} X_t^2 \right|$$

and if it is set,

$$\beta_0 \equiv -l$$

$$\beta_1 \equiv (1 + \tfrac{l}{u})$$

$$\beta_2 \equiv -\tfrac{1}{u}$$

then it is possible to write, $\sigma_{q,\ t} = \sigma \left|\beta_0 + \beta_1 X_t + \beta_2 X_t^2\right|$. A key feature of this quadratic specification is that $\sigma_{q,t}$ depends on the spread size (contango or backwardation), in a nonlinear manner, and also depends upon the lower and upper bounds in the futures spread.

There is also a clear interdependence among β_0, β_1 and β_2 in the standard deviation of the process when the lower and upper bounds are considered. To satisfy Rady's specification the parameter β_1 should satisfy the restriction $\beta_1 = 1 + \beta_0 \beta_2$.

Notice moreover that the shape of the quadratic volatility function $\sigma_{q\ t} = \sigma \left|\beta_0 + \beta_1 X_t + \beta_2 X_t^2\right|$ is determined by the sign and size of the parameters β_0, β_1 and β_2 for different levels of spread. As long as β_0 and β_2 have opposite sign to β_1, it is ensured that $l > 0$ and $u > 0$ are satisfied. Additionally, β_2 plays an important role in determining the curvature of the volatility function on the spread size. Provided that $\beta_2 < 0$, the function $\sigma_{q,t}$ will have a maximum. For an appropriate

---

[3] The Wiener process (Wt) by definition is a random variable which is normally distributed with zero mean, variance Δt, and for any two different short interval of time Δt, ΔWt and ΔWt+1 are independent.

magnitude of these parameters such that the restriction $\beta_1 - \beta_0\beta_2 = 1$ is satisfied, and when $\beta_2 < 0$, this quadratic volatility function will resemble an inverted parabola which goes to zero at the boundaries, see Figure 1. In the next section the role of this restriction is examined in more detail in implementing Rady's specification.

It has been documented theoretically and empirically that commodity prices display heteroscedasticity, (see Williams and Wright, 1991 and Kang and Brorsen, 1995). So, spreads are expected to be unconditional heteroscedastic. If commodity spread returns are identically and independently distributed and heteroscedasticity in spread returns volatility is taken into account, it is expected that spread returns will be conditionally normal distributed. In that context, Maximum Likelihood Estimation (MLE) provides an easy way to calculate the transitional probability density function of yesterday's state and today's state, even when discrete observations of commodity spreads are used.

The discrete version for a small interval, *Δt*, of the continuous process of commodity futures spreads of (2) is given by:

$$\Delta X_t = \sigma \ (X_t - l)(1 - \tfrac{1}{u} X_t) \ \Delta W_t \tag{3}$$

with $\Delta W \sim N(0, \Delta t)$, where Δt is the variance. Since the statistical properties of ΔX_t depend crucially on ΔW_t and $\sigma_{q,t}$, it follows that $\Delta X_t \sim N(0, \sigma_{q,t}{}^2)$.

Thus, the approximated conditional density function for each point in time for futures spread returns, $f(\Delta X_t)$, is given by a normal distribution with mean *0* and variance $\sigma_{q,t}{}^2$:

$$f(\Delta X_t) = \frac{1}{\sigma_{q,\ t}\sqrt{2\pi}} \exp^{-\frac{1}{2}\left(\frac{\Delta X_t}{\sigma_{q,t}}\right)^2} \tag{4}$$

The corresponding likelihood function for spread returns, ΔX_t, is given by:

$$f(\Delta X_1, \ldots, \Delta X_T \mid \beta_0, \beta_1, \beta_2) = \prod_{t=1}^{T} \frac{1}{\sigma_{q,\ t}\sqrt{2\pi}} \exp^{-\frac{1}{2}\left(\frac{\Delta X_t}{\sigma_{q,t}}\right)^2} \tag{5}$$

where:

$$\sigma_{q,\ t} = \sigma \left| \beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-1}{}^2 \right| \sqrt{\Delta t}\ .$$

In this model X_t is conditional on X_{t-1} and has a variance which depends on X_{t-1}. The important feature in this specification is that relevant information up to *t-1* about the state variable, contained in the spread, is used to determine the subsequent spread at *t*, but also that lower and upper bounds in the spread are incorporated. Another interesting characteristic is that spread returns , ΔX_t, follow a martingale difference process in which its volatility depends on the degree of backwardation in the market and upon the arbitrage bound.

It can be shown that the sample log likelihood function, for a normal density function with the above quadratic variance function including the bounds, is given by:

$$L(\theta) = -\frac{T}{2}[\log(2\pi) + \log(\sigma^2)] - \frac{1}{2}\sum_{t=1}^{T}\log(\sigma_{q,t})^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\frac{(\Delta X_t)^2}{\sigma_{q,t}^2} \quad (6)$$

where:

$$\sigma_{q,t} = \left|\beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-1}^2\right| = \sigma f(\theta, X)$$

and T is number of observations.

The function $\sigma_{q,t}^2 = \sigma^2 f_t(\theta)$ allows a general form of heteroscedasticity to be considered. If the parameter β_1 is not constrained in the above log likelihood function, then this model represents a general case in which the disturbance term of spread returns has a state dependent variance. In particular, if $f_t(\theta)$ is a quadratic function, this has the advantage that of encompassing different types of heteroscedasticity: i) a quadratic variance function on spread, ii) a linear variance function on spread, and iii) a constant variance independent of spread movements, i.e. homoscedasticity.

The estimation of the diffusion process (3) enables the determination of the values of β_0, β_1 and β_2, which maximise the conditional log likelihood function (6). These parameters can be estimated using gradient algorithms. Once the asymptotic covariance matrix for the maximum likelihood estimates is computed the square roots of the diagonal elements give the estimated asymptotic standard errors.

Bollerslev and Wooldrige (1992) provided a quasi-maximum likelihood estimation technique in which the coefficients of the covariance matrix from a Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model are robust to nonnormality of the conditional variance of the disturbance term. In this paper, robust standard errors for the estimated parameters are reported. These standard errors, calculated in RATS, are asymptotically similar to the ones developed in Bollerslev and Wooldridge (1992).

Similar procedures can be used to implement Rady's specification by imposing the above restriction in the maximisation of the log likelihood function (6).

IV. Data description

Daily settlement futures prices for coffee from the Coffee, Sugar and Cocoa Exchange; for crude oil, gold and heating oil from New York Mercantile Exchange-COMEX; and for soybeans from the Chicago Board of Trade were used in estimating the above model. All futures delivery contracts were obtained from the January 1985 futures contract to the September 1995 futures contract, except for crude oil which was obtained from the January 1987 futures contract to the February 1991 futures contract.

To construct continuous time series for futures prices, all traded contracts for each commodity were used, and when the futures contract expires it was rolled over to the next immediate futures contract a day after the first notice day. Two advantages of doing this are that any noise prior to the delivery month is eliminated and overlapping observations of futures prices at different maturities are avoided.

It is common to define the spread as the difference between the nearby futures price and the distant futures price. In this study, the nearby futures price is referred to the spot month and the distant futures price represents the most immediate futures contract available for each commodity. However, for purposes of implementation of Rady's (1995) framework, futures spreads must be calculated as the ratio of the nearby futures over the distant futures price on the same day, (relative futures spread). Note that this is just a change of scale since the same pattern is preserved.

Constructing the first and second nearby futures prices as above, commodity spreads returns were computed as the first logarithmic difference of the ratio of these two futures prices.

In Table 1 summary statistics are presented for the five different commodity spot (futures) spread returns. A preliminary analysis of spread returns suggests that means of spot spread returns and futures spread returns are close to zero for the five commodities analysed. However, the futures spread returns appeared to be less volatile than spot spread returns. Moreover, gold futures spreads seem to display less erratic behaviour than the agricultural and energy commodities, as expected.

It is evident as well that commodity (spot) futures spreads display asymmetric behaviour. All commodities, except crude oil and coffee, displayed negative skewness in futures spread returns.

Another interesting, but not surprising, feature is the presence of high kurtosis in spot spread returns and futures spread returns. Thus, more extreme changes are observed than would be the case for a normal distribution in spot (futures) spreads, but also more observations concentrated around their mean value. Futures spread returns display higher leptokurtosis than spot spread returns.

There is evidence of negative serial correlation for futures spreads returns for the five commodities. The square of spread returns can be considered as a proxy for the variance of the noise component of spread returns. Their high first order correlation coefficient suggests a serially correlated variance of the noise term of spread returns.

A. Adjusted data

As mentioned before, the noise term of the spread returns is mainly determined by the availability of the commodity, captured by the spread. Nevertheless, there are other factors which can affect this noise term.

For example, one possible factor is the presence of price limits in commodity futures markets. It is known that price limits are present in American futures exchanges. When the market becomes more volatile it is more likely that price limits are hit in commodity futures markets. The inclusion of price limits in the sample could introduce biases in the estimation. In order to deal with this problem, futures prices affected by price limits were removed from the sample.[4]

In addition, it is well-known that a higher variability will be observed in higher frequency data for commodity prices. Thus, weekly data were used as well as daily data. This has a number of advantages, including that a smaller number of price limit

[4] Spread returns were computed using corresponding adjacent spreads for a day before and a day after the price limit was reached. When a two or more subsequent days displayed price limit moves, we used adjacent spreads for two or more days before and after price limits were reached.

moves are observed. So, futures prices on Wednesdays were used to determine weekly spreads and their spread returns.

Another source of noise in the estimations might be that spreads are constructed from a continuous time series of futures prices. This creates jumps, at each maturity month, in spread returns when the spot month is rolled over to the immediate futures contract. For those days in which the spot month contract was rolled over to the immediate futures contract, spread returns were calculated using the spread at the first notice day of the spot month and the spread using the immediate futures contract.

Also, seasonality plays an important role, at least for agricultural and energy commodities, in affecting spread returns volatility. Since in 'old' crop periods, availability is low, spread volatility in these periods reacts more to supply/demand shocks. But around the harvest time, availability is high, and spreads should be more stable. For example, for soybeans, during summer time (old crop period) spreads tend to display high backwardation levels and to be very volatile. However, in November when the harvest takes place, spreads very often display contango situations and less volatile spreads.[5] A dummy variable could be included to distinguish between old crop (high season) and new harvest periods (low season) for agricultural (energy) commodities. Two types of dummy variables were considered; an additive dummy variable and a multiplicative one. The latter has the feature that it accounts for the impact that the degree of backwardation has on the volatility of spread returns.

Another factor which affects the spread returns volatility is the noise which prevails near the period in which agents need to give notice of their intentions for delivering the physical commodity. Therefore, to reduce this noise the spot month was rolled over to the immediate futures contract 15 days prior to the first notice day. In the next section the estimation results, taking into account these adjustments, are discussed in more detail.

V. Estimation results

A. A state dependent variance function model

---

[5] It is evident that soybeans November-January spreads consistently display contango situations and are very stable across years. However, the soybeans August-September spreads have different patterns across years. In some years, August-September spread displays backwardation situations only, but in some other years it is observed both regimes; backwardation and contango situations.

This section focuses on determining which type of form of heteroscedasticity represents better the behaviour of commodity futures spreads returns. For each commodity, three different specifications for the spread returns volatility were considered; i) quadratic variance function, ii) linear variance function and iii) constant variance case. This has the advantage that it allows a comparison to be made between a model with a state dependent variance function and a constant variance model.

Daily and weekly data were used when price limits were not binding. The approximated log likelihood function (6) without the restriction on β_1 is maximised based on the BFGS method. When using both daily and weekly frequency data, for several different starting values for the parameters considered, convergence was not achieved for any of the three variance functions specifications. To overcome the problem of nonconvergence the relative futures spread and relative spread returns for the five commodities were scaled (standardized) to ease the iteration procedure. So, the deviation of the relative spread about its mean value was computed and divided by its standard deviation. The spread returns were divided only by its standard deviation, since their mean is very close to zero.

Notice that the scaling in spreads preserves the sign of the parameter β_2, but not those of the parameters β_0 and β_1, since for standardized spreads the volatility of the spread returns is re-expressed as $\sigma_{q,t}* = \sigma\left|\beta_o * + \beta_1 * X_{t-1} * + \beta_2 * X_{t-1} *^2\right|$ where $X* = \dfrac{X - \mu_X}{\sigma_X}$. It can be shown that the equivalence between nonstandardized and standaridized spreads is given by $\beta_0 \equiv \beta_0 * - \dfrac{\beta_1 \mu_X}{\sigma_X} + \dfrac{\beta_2 * \mu_X}{\sigma_X^2}$, $\beta_1 \equiv \dfrac{\beta_1 *}{\sigma_X} - \dfrac{2\beta_2 * \mu_X}{\sigma_X^2}$ and $\beta_2 \equiv \dfrac{\beta_2 *}{\sigma_X^2}$.

Once standardized series are used to maximise the log likelihood function, convergence is achieved. The results for daily and weekly data of the estimated parameters along with their robust t-statistics are displayed in Panels A and B of Table 2 respectively. In addition, the statistical properties of the standardized residuals, $\dfrac{\varepsilon_t}{\hat{\sigma}_{q,t}}$, for each model are provided.

Based on daily data, for the three variance function specifications, all coefficient parameters are significantly different from zero at the 5 % significance level. For the quadratic variance case, for all commodities the estimated parameters appear all positive, but for gold the parameter β_2 is negative, see Panel A. There is evidence however that the standardized residuals for the different models are skewed and highly leptokurtic.

For coffee, crude oil, heating oil and gold, the parameter β_2 remained negative and significantly different from zero, when weekly data were used instead of daily data. A clear distinction though is that the standardized residuals of all models display much lower skewness and leptokurtosis than the ones for daily data.

Given the estimated parameters, the quadratic volatility function ($\sigma_{q,t}$) can be calculated for different levels of the spread, and at the bounds. These calculations are reported in Panels A and B of Table 3 for daily and weekly data respectively.

Although the evidence indicates that there is a quadratic relationship for coffee, crude oil, heating oil and soybeans, contrary to expectations, a *U*-shaped between spread returns volatility and spread size is observed. A similar pattern, using the crude oil implied volatility, was obtained by Simons (1996). Unlike the above commodities, the volatility function for gold reflected a different pattern in which volatility is higher at lower level of spreads than at bigger level of spreads.

It is apparent that this *U*-shaped relationship (for both daily and weekly data) reflects a failure to bound spread returns volatility at the lower and upper bounds, i.e volatility does not go to zero at the boundaries.

To test whether or not the noise component of the commodity futures spreads returns display a state dependent volatility, two nested hypotheses for the five commodities were performed, using the Likelihood ratio test (LR). [6] A comparison of the log likelihood function of the quadratic variance case was made with the one of a linear variance case, LR_2, and similarly for a linear variance function against the constant case, LR_1. The LR statistic tests for both hypotheses showed that a quadratic

---

[6] **The likelihood ratio (LR) test compares two different log likelihood functions under a restricted case, $\log(\theta\, r)$, and another unrestricted case, $\log(\theta\, ur)$. The LR test-statistic is constructed by the ratio of these different log likelihood functions and, it is distributed as a χ_k^2, with *k* degrees of freedom i.e. $LR = -2[\log(\theta\, r) - \log(\theta\, ur)] \sim \chi_k^2$ with k being the number of restrictions.**

diffusion term was accepted in favour over the linear diffusion term for the five commodities, except for heating oil; and the linear one was accepted in favour over a constant variance case, at the 5% significance level. For the latter hypothesis, crude oil was the only exception.

Also, the log likelihood function for a quadratic variance function was compared with the constant variance case, LR_0. For all commodities, the quadratic variance function was accepted in favour over the constant variance case at the 5 % significance level.

Two main features are worth mentioning from the above results. First, the standard deviation of weekly spread returns process displays a quadratic form of heteroscedasticity. That is, a quadratic relationship is observed between spread returns volatility and relative spread size. Second, the diffusion term of the weekly spread returns is better represented by a state dependent volatility, taking into account the arbitrage bound.

A.1 Importance of the arbitrage bound

As mentioned before, weekly data support a state dependent variance for spread returns and subject to the arbitrage bound. Notice though that the standardized residuals of the different models remain slightly skewed and leptokurtic.

An arbitrage argument however can be used instead of a goodness of fit criteria to exhibit the importance of the arbitrage bound when modelling commodity futures spreads. That is, a model which recognises the arbitrage bound will tend to predict less futures spread levels which lie outside this lower bound.

In order to show this point more clearly, the number of violations of the arbitrage bound was calculated in a model which assumes a constant variance of the process. When daily spreads are close to the full carry at time *t*, this model with constant variance will tend to predict more cases (compared to a model which recognises the arbitrage bound) in which spreads in a subsequent time $t+1$ will be above the full carrying charges. Levels L_1 and L_0 in futures spreads were chosen, for each commodity, in such a way that after an interval H, the distance between L_1 and L_0 is a one standard deviation event, i.e. $L_1 - L_0 = \sigma\sqrt{H}$, where H is the time horizon (three weeks). Taking all cases in which the futures spread was at L_1 the probability that after 15 days later the spread is below L_0 is calculated. These calculations indicate that for coffee, there was a 6% of probability of being below L_0 after 15 days; for gold 9%; for heating oil 8%; and for soybeans 17.4%.[7] Thus, a high number of violations of the no-arbitrage bounds are observed in a naive model which ignores the arbitrage bound.

B. **Empirical implementation of Rady's model**

The implementation of Rady's specification involves the maximisation of the log likelihood function (6), but imposing the restriction $\beta_1 = 1 + \beta_0\beta_2$. In addition, this log likelihood function was maximised to determine β_0, β_1 and β_2 making the following adjustments (see above): i) removing data affected by price limits and rolling the spot month over to the immediate futures contract 15 days prior to the first notice day, ii) as i) but also including a seasonal dummy variable and iii) using weekly data and removing data affected by price limits. The first two cases were estimated using only daily data.

For each of the above adjustments, Table 4 provides a summary of the results for the quadratic variance function for the five different commodities, along with the statistical properties of their standardized residuals.

Estimated parameters for the quadratic variance function specification, with the restriction on β_1, were used to calculate the standard deviation of the spread returns process -volatility, $\sigma_{q,\ t} = \sigma\ \left|\beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-1}^{\ 2}\right|$. Table 5 display volatility functions of spread returns for different levels of spread for the five commodities with the above three adjustments.

Tables 4 and 5 reveal several interesting features. First, in contrast to the general case, in which all commodities except gold display a positive and significant value for the parameter β_2, it is evident that once the restriction is imposed on β_1 for coffee, crude oil and heating oil the parameter β_2 becomes negative and statistically significant at the 5% significance level. Although there is evidence that the standardized residuals for the Rady's specification are not normally distributed for all

---

[7] **Results are available from the authors upon request.**

commodities, these residuals appear to have less skewness and leptorkurtosis than the ones when the restriction is not imposed on β_1 at least for coffee, crude oil and soybeans.

Unlike the general case, when the restriction is taken into account, the volatility functions of spread returns for all commodities, except for soybeans, reflect the expected parabola shape, in which at the bounds the volatility goes to zero, see Table 5. For instance, for the case of heating oil, it is observed that when price limits are not present the volatility function reflects a pattern in accordance with Rady's specification, see Figure 2.

Second, the results indicate the statistical significance of the coefficient associated with a multiplicative seasonal dummy variable, (for all commodities except for gold in which a dummy variable was not included in the estimation). Also larger values of the log likelihood function are observed once a seasonal dummy variable is introduced. Further, the standardized residuals for this model in which seasonal volatility adjustments are considered, show lower skewness and leptokurtosis than the ones obtained without the dummy variable, except of crude oil case. As with the earlier pattern of the spread returns volatility function based only on adjustments for price limits, the results suggest that for all commodities, except soybeans, the volatility function increases more rapidly outside the bounds, but the volatility goes to zero at the boundaries.

Third, for soybeans the parameter β_2 remained positive despite the adjustments for seasonal volatility and price limits. Empirical evidence shows that the pattern of the volatility function ($\sigma_{q,t}$) for soybeans increases more with the degree of backwardation, see Figure 3.

Fourth, for weekly data, we observe a better fit of the Rady's specification for most commodities. In particular, for crude oil and heating oil their standardized residuals for the quadratic variance function imposing the restriction on β_1 are much less skewed and leptokurtic. Moreover, the standardized residuals for the models for crude oil and heating oil do not display serial correlation at the 5% significance level and 1% significance level respectively. For all commodities expect soybeans their volatility functions of spread returns display a maximum as predicted in Rady's specification in which the volatility decreases at the boundaries.

Finally, part of the nonnormality of the standardized residuals in a quadratic variance function model is explained by the presence of price limits. However, there is evidence that higher serial correlation and higher frequency variability in daily data are observed for most of the commodities.

VI. Empirical Findings

Several empirical findings emerge from the above results. First, a list of the empirical findings consistent with the theory is given:

1. The noise component of commodity spreads returns appear to be better represented by a quadratic variance function rather than a linear one to account for heteroscedasticity.

2. Generally speaking, Rady's specification fits well for most commodities for weekly data. Nevertheless, the higher frequency variability and serial correlation displayed in daily data makes it difficult to fit the above process for all commodities considered.

Second, other empirical findings that were not predicted by the theory are listed below:

3. In the majority of the cases, the spread returns volatility appears to increase more at the extreme contango and backwardation situations when the restriction that $\beta_1 = 1 + \beta_0 \beta_2$ is not imposed.

4. Although the volatility of soybeans spreads are determined by fundamental factors in physical commodity markets; presence of price limits, changing the maturity month did not explain the failure to fit Rady's model.

5. It appears for soybeans that there were several days in which backwardation was erratic in the sense that it widened and narrowed quickly in the period analysed. In that case spreads at backwardation were not necessarily all close to the upper bound.

VII. Implications for options on futures

The model considered in this paper assumed that the commodity spread follows the following stochastic process:

$$d(F_1 - F_2) = \sigma f(F_1 - F_2) dW \qquad (7)$$

How does this model restrict the volatility of the first nearby futures contract price, σ_1, and the volatility of a more distant futures contract price, σ_2? The stochastic processes of the two futures prices with different maturities, F_1 and F_2 respectively are given as follows:

$$dF_1 = \sigma_1(.)dW_1 \tag{8}$$

$$dF_2 = \sigma_2(.)dW_2 \tag{9}$$

with a correlation coefficient ρ between F_1 and F_2.

Assuming that dW_1 and dW_2 are independent (8) can be expressed as follows;

$dF_2 = \sigma_2(.)[\rho(.)dW_1 + \sqrt{1-\rho(.)^2}dW_2]$. From (8) and (9) the process followed by the difference of these two futures prices is $dF_1 - dF_2 = [\sigma_1(.) - \rho(.)\sigma_2(.)]dW_1 + \sqrt{1-\rho(.)^2}\sigma_2(.)dW_2$, where its volatility is $\sqrt{(\sigma_1(.) - \rho(.)\sigma_2(.))^2 + (1-\rho(.)^2)\sigma_2{}^2(.)}$. Notice however that (7) restricts the volatility of (8) and (9) depending on whether the market is in contango or backwardation. This is because under contango, when inventories are available, F_1 and F_2 will be close with $\rho = 1$ and $\sigma_1{}^2 \approx \sigma_2{}^2$ to avoid arbitrage, and volatility of the changes in $dF_1 - dF_2$ is given by $\sqrt{(\sigma_1(.) - \sigma_2(.))^2}$ which becomes close to zero. By contrast, under backwardation F_1 and F_2 will tend to move apart, and $\sigma_1{}^2 > \sigma_2{}^2$, (see Ng and Pirrong, 1994).

The state dependent volatility of spreads, along with the lower and upper bounds have implications on the option value. Under the stochastic process of the spread returns followed by (2) the option pricing formula derived by Rady's (1995) is given by:

$$C_t = \frac{1}{1-u^{-1}l}\left\{(1-u^{-1})(F_1 - lF_2)\phi(e_t{}^+) - (1-l)(F_2 - u^{-1}F_1)\phi(e_t{}^-)\right\} \tag{10}$$

where:

$$e^{\pm}{}_t = \frac{1}{(1-u^{-1}l)\sigma\sqrt{T-t}}[\log\frac{F_1 - lF_2}{F_2 - u^{-1}F_1} - \log\frac{1-l}{1-u^{-1}} \pm \frac{1}{2}(\sigma - u^{-1}l\sigma)^2(T-t)]$$

ϕ denotes the standard normal distribution function, and (*T-t*) is the time to maturity.

Thus, the call value will be higher in backwardation than in contango.

VIII. Conclusions

This paper provides an empirical implementation of a theoretical model, which respects two key features of commodity futures spreads; i) a state dependent variance function (nonlinear) on spreads and ii) the existence of arbitrage bound(s) in commodity spot (futures) spreads.

We have argued that the above features can be captured by a stochastic diffusion process with a quadratic term with lower and upper bounds in spreads, similar to the process proposed by Rady's (1995) framework.

An empirical estimation of this stochastic process with a quadratic term was undertaken for five different commodity spreads. Empirical evidence shows that unconstrained Black-Scholes models do not capture the behaviour of commodity spreads. Furthermore, data support a nonlinear variance function for spread returns subject to the arbitrage bound.

For crude oil and heating oil cases the Rady's specification fits well for weekly data. However, more reservations about accommodating this process arise for daily frequency for all commodities considered. This is because the standardized residuals for this model remain skewed and highly leptokurtic.

For soybeans despite the seasonal volatility and price limit adjustments, the stochastic process with a quadratic term is rejected. Part of the failure to fit a quadratic diffusion term for spread returns is attributed to the higher frequency variability observed in daily data. Nevertheless, it maybe that volatility increases with the degree of backwardation. It is also possible that other factors not considered in this paper may have played an important role in determining the volatility of the spread returns of soybeans. But this conjecture deserves further investigation.

To provide more evidence of the importance of the arbitrage bound a different model is needed in which the arbitrage bound is imposed rather than estimated. The likelihood function for the variance should be such that there is no probability of passing below that bound. This however awaits for further research.

Finally, when commodity spread option data become available, models of stochastic volatility with a quadratic variance function on spread can be tested. Note, that the empirical implementation of Rady's (1995) framework can be tested for different financial instruments and is not specific to commodities.

References

Black, F. and M. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy*, 81, 637-654.

Bollerslev, T. and J.M. Wooldridge, 1992, Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, Econometric Reviews, 11, 143-179.

Brunetti, C. and C. Gilbert, 1995, Metals price volatile, 1972-95, *Resources Policy*, 21, 237-254.

Deaton, A. and G. Laroque, 1992, On the behaviour of commodity prices, *Review of Economics Studies,* 59, 1-23.

Grabbe, O. , 1995, Copper-bottom pricing, *Risk Magazine* vol 8, no. 5, 63-66.

Greene, W., 1993, Econometric Analysis, New York Macmillan International.

Kang, T. and W. Brorsen, 1995, Conditional heteroskedasticity, asymmetry, and option pricing, *The Journal of Futures Markets,* vol 15, no. 8, 901-928.

Kolb, R., 1991, "Understanding futures markets". Third Edition. Simon & Schuster.

Merton, R.C., 1973, Theory of rational option pricing, *Bell Journal of Economics and Management*, 4, 141-183.

Ng, V. K. and S. C. Pirrong, 1994, Fundamentals and volatility: storage, spreads, and the dynamics of metals prices, *Journal of Business* vol. 67, no 2, 203-230.

Pirrong, S., 1997, Metallgesellschaft: a prudent hedger ruined, or a wildcatter on NYMEX?, *The Journal of Futures Markets,* vol 17, no. 5, 543-578

Rady, S.,1995, Option pricing with a quadratic diffusion term, Discussion paper no. 226, *LSE Financial Markets Group*.

Shimko, D.C., 1994, Options on futures spreads: hedging, speculation, and valuation, *The Journal of Futures Markets,* vol. 14, 183-213.

Simons, H., 1996, Measuring market tension, *Futures Cedar*, February.

Trivedi, P.K., 1995, Tests of some hypotheses about the time series behaviour of commodity prices in "Advances in Econometrics and Quantitative Economics", Maddala, G.S. et al. Basil Blackwell, 382-412.

Wright, B. and J. Williams, 1991, "Storage and commodity markets", Cambridge University Press.

Table 1. Summary statistics of commodity spot (futures) spreads

	Coffee (cents/pound)	Crude oil ($/barrels)	Gold ($/oz)	Heating oil (cents/gallon)	Soybeans (cents/bushel)
Mean:					
(ΔB)	-1.66×10^{-5}	-	-0.81×10^{-6}	-7.40×10^{-6}	-1.11×10^{-6}
(ΔX)	-2.50×10^{-6}	-3.90×10^{-5}	-1.13×10^{-6}	-1.00×10^{-6}	-4.79×10^{-6}
Standard Deviation:					
(ΔB)	0.01834	-	0.009617	0.01484	0.005677
(ΔX)	0.00929	0.00945	0.000331	0.008669	0.002950
Skewness					
(ΔB)	0.8574	-	0.1222	2.0546	-0.8917
(ΔSP)	3.8151	1.1366	-0.5432	-0.3491	-3.4424
Kurtosis [a]					
(ΔB)	63.1819	-	7.4031	40.7520	37.1901
(ΔX)	126.6012	40.3049	23.9815	42.4613	70.0913
ρ_1 [b]					
(ΔB)	-0.23	-	-0.51	-0.15	-0.21
$(\Delta B)^2$	0.46	-	0.44	0.22	0.34
(ΔX)	-0.07	-0.24	-0.05	-0.13	-0.11
$(\Delta X)^2$	0.06	0.37	0.16	0.32	0.16
nobs					
(ΔX)	2889	940	2498	2629	2672

ΔB measures basis (spot spread) returns and ΔX represents futures spread returns.
nobs: number of observations

For crude oil the spot price is not a reliable, so basis (spot spread) was not calculated.

[a] **The fourth moment reported here is measured in the following way** $\gamma_2 = \frac{\mu_4}{\mu_2^2}$

b First order autocorrelation coefficient.

Table 2. Different variance functions specifications for commodity futures spreads

Panel A, Adjusted daily data

Quadratic variance function	Coffee	Crude oil	Gold	Heating oil	Soybeans
β_0	1.4809	0.9788	-1.7577	-4.9199	1.2167
	(16.575)	(46.510)	(-77.692)	(-5.0856)	(103.059)
β_1	0.0383	0.9095	0.7545	0.3514	0.4780
	(0.3750)	(11.890)	(30.797)	(0.51193)	(19.865)
β_2	0.1462	0.6337	-0.2362	-1.5279	0.2404
	(2.1146)	(8.810)	(-9.916)	(-4.2310)	(9.313)
σ	0.6037	0.6203	0.4681	0.1615	0.6064
	(16.465)	(73.322)	(147.372)	(4.8282)	(91.576)
Log L	-1280.788	-208.999	-1057.928	-1277.561	-696.690
nobs	2714	858	2498	2633	2612
Properties of standardized residuals					
Skewness	-2.254	-1.811	2.219	0.1676	-0.592
Kurtosis	65.635	52.295	46.788	54.487	204.772

Linear variance function	Coffee	Crude oil	Gold	Heating oil	Soybeans
β_0	1.7057	1.4923	1.4363	1.0155	1.1026
	(27.310)	(14.210)	(144.768)	(35.541)	(121.733)
β_1	0.0329	0.2975	-0.4362	0.0509	0.4121
	(0.4916)	(3.796)	(-34.877)	(0.8451)	(30.877)
σ	0.5858	0.6133	0.6419	0.9815	0.7813
	(28.373)	(15.457)	(146.377)	(17.614)	(133.743)
Log L	-1353.788	-334.353	-1112.783	-1304.250	-766.978
nobs	2714	858	2498	2633	2612
Properties of standardized residuals					
Skewness	-1.846	-2.826	3.971	-0.915	0.764
Kurtosis	54.012	53.649	89.136	64.943	25.680

Continued Panel A.

Constant case	**Coffee**	**Crude oil**	**Gold**	**Heating oil**	**Soybeans**
β_0	0.9998 (13.914)	0.9995 (5.601)	1.0002 (6.311)	0.9998 (13.011)	0.9998 (15.425)
Log L	-1356.000	-428.125	-1503.646	-1315.622	-1305.039
nobs	2714	858	2498	2633	2612
Properties of standardized residuals					
Skewness	-1.903	-5.664	10.399	1.113	-3.028
Kurtosis	54.226	107.816	300.422	60.368	41.966

The asymptotic robust t-statistics are in parentheses, based on the BFGS numerical algorithm method. The critical value at 5% significance level is 1.96.

Log L: Log likelihood function
nobs: number of observations

Table 2, Panel B, Adjusted weekly data

Quadratic variance function	Coffee	Crude oil	Gold	Heating oil	Soybeans
β_0	1.2415 (139.510)	0.6965 (9.817)	-1.0840 (-27.212)	1.1114 (47.503)	0.8205 (29.131)
β_1	0.1797 (13.200)	0.1826 (1.906)	0.2061 (5.300)	0.2287 (5.236)	0.5826 (12.145)
β_2	-0.0379 (-5.992)	0.4878 (5.577)	-0.0703 (-1.753)	-0.0206 (-1.181)	0.3657 (6.222)
σ	0.8220 (133.674)	0.7985 (19.995)	0.8522 (35.842)	0.9041 (50.733)	0.7624 (43.919)
Log L	-267.524	-43.511	-242.905	-244.344	-77.414
nobs	566	147	518	518	2672
LR_2	11.398	57.979	9.100	1.080	72.248
Properties of standardized residuals					
Skewness	-0.916	1.019	0.184	0.732	0.040
Kurtosis	29.783	3.389	4.165	7.595	17.848

Linear variance function	Coffee	Crude oil	Gold	Heating oil	Soybeans
β_0	1.2499 (16.792)	1.2409 (18.8826)	1.145 (36.372)	1.4000 (30.094)	1.1954 (23.682)
β_1	0.2153 (1.174)	0.0107 (0.1686)	-0.1678 (-1.933)	0.2618 (3.098)	0.4921 (12.297)
σ	0.7988 (16.325)	0.8034 (15.9523)	0.6970 (29.729)	0.7058 (29.527)	0.6718 (21.868)
Log L	-273.223	-72.551	-247.455	-244.884	-113.538
nobs	566	147	518	518	2672
LR_1	17.646	1.960	21.124	26.744	311.942
Properties of standardized residuals					
Skewness	-1.004	1.193	0.254	0.744	-0.012
Kurtosis	32.403	6.296	4.848	7.723	14.414

Continued Panel B.

Constant case	**Coffee**	**Crude oil**	**Gold**	**Heating oil**	**Soybeans**
β_0	0.9991 (8.444)	0.9972 (8.549)	0.9999 (16.627)	0.9995 (14.777)	0.9991 (9.120)
Log L	-282.046	-72.591	-258.017	-258.256	-269.509
nobs	566	147	518	518	541
LR_0	29.044	58.160	30.224	27.824	384.190
Properties of standardized residuals					
Skewness	-1.117	-1.132	0.651	0.158	-1.813
Kurtosis	29.985	6.099	5.652	7.729	24.834

The asymptotic robust t-statistics are in parentheses, based on the BFGS numerical algorithm method. The critical value at 5% significance level is 1.96.

Log L: Log likelihood function
nobs: number of observations

The Likelihood ratio (LR) tests were performed for the following nested hypotheses: a) H_0: β2=0, LR_2, b) H_0: β1=0, LR_1, and c) H_0: β1=0 and β2=0, LR_0, in expression (6) respectively. The critical values at the 5% significance level are X_1^2=3.84 and X_2^2=5.99 respectively.

Table 3. Quadratic variance function, adjusted data

Panel A. Daily data

	Coffee				Crude oil				Gold				Heating oil				Soybeans			
	ll	X>ll	X<ul	ul	ll	X>ll	X<ul	ul	ll	X>ll	X<ul	ul	ll	X>ll	X<ul	ul	ll	X>ll	X<ul	ul
relative spread (X)	0.879	0.959	1.039	1.089	0.957	0.978	1.020	1.087	0.978	0.990	0.994	0.998	0.925	0.965	1.135	1.225	0.973	0.993	1.033	1.083
standardized X	-4.811	-1.193	2.424	4.684	-2.732	-1.632	1.119	4.419	-2.686	0.282	1.024	2.013	-3.147	-1.560	5.186	8.757	1.472	0.169	3.452	7.555
volatility $(\sigma q,t)$	2.825	0.992	1.468	2.939	2.000	0.733	1.730	10.78	3.399	0.645	0.260	0.063	3.417	1.483	7.137	19.22	1.559	3.294	8.691	19.05

Panel B. Weekly data

	Coffee				Crude oil				Gold				Heating oil				Soybeans			
	ll	X>ll	X<ul	ul	ll	X>ll	X<ul	ul	ll	X>ll	X<ul	ul	ll	X>ll	X<ul	ul	ll	X>ll	X<ul	ul
relative spread (X)	0.899	0.939	1.030	1.089	0.938	0.978	1.048	1.098	0.985	0.989	0.994	0.999	0.964	0.984	1.074	1.174	0.965	0.985	1.045	1.095
standardized X	-3.734	-1.978	1.971	4.604	-2.817	-1.262	1.460	3.404	-1.895	-0.565	1.097	3.092	-1.377	-0.713	2.276	5.597	-1.985	-0.49	3.973	7.697
volatility $(\sigma q,t)$	0.035	0.606	1.191	1.040	3.236	0.992	1.599	5.565	1.472	1.042	1.097	0.953	0.685	0.848	1.379	1.579	0.842	0.474	6.791	20.56

Estimated parameters β_0, β_1 and β_2 from the Maximum Likelihood Estimation of the log likelihood function (6), displayed in Table 2 were used to calculate the standard deviation (volatility) of spread returns.

Notice that ll and ul are the lowest and highest level of spreads respectively for the sample analysed for each commodity.

Table 4. Quadratic variance function specification, with different adjustments

Commodity	15 days prior to the first notice day and price limits	seasonal dummy variable *	weekly basis and removing data affected by price limits
Coffee			
β_0	4.6032 (69.796)	3.8389 (12.091)	3.271 (173.693)
β_1	0.2229	0.00457	0.5054
β_2	-0.1688 (-5.772)	-0.2593 (-30.645)**	-0.1512 (-3.938)
σ	0.2430 (14.062)	0.2640 (12.039)	0.3191 (9.824)
Log L	-1550.992	-1332.904	-269.630
nobs	2714	2714	566
Properties of standardized residuals			
Skewness	-1.625	-1.838	-0.801
Kurtosis	53.448	51.458	28.073
LM test	244.853	244.854	127.161
Crude oil			
β_0	0.8226 (4873.857)	1.4509 (25.682)	0.777 (722.208)
β_1	0.7519	0.1319	0.6211
β_2	-0.3016 (-126.656)	-0.5983 (-49.120)**	-0.4876 (-28.816)
σ	1.969 (149.389)	0.8025 (6.561)	1.7488 (33.777)
Log L	-721.431	-484.648	-89.993
nobs	858	858	147
Properties of standardized residuals			
Skewness	-0.501	-4.227	0.989
Kurtosis	27.983	63.608	5.853
LM test	3.032	3.032	2.139

Table 4. (continued)

Commodity	15 days prior to the first notice day and price limits	seasonal dummy variable *	weekly basis and removing data affected by price limits
Heating oil			
β_0	3.0476 (91.436)	1.7154 (5.798)	3.656 (6.636)
β_1	0.6595	0.6909	0..7525
β_2	-0.1117 (-6.949)	-0.1802 (-6.170)**	-0.0677 (-2.555)
σ	0.3612 (11.098)	0.1116 (5.107)	0.2748 (6.217)
Log L	-1427.937	-1314.344	-244.344
nobs	2633	2633	518
Properties of standardized residuals			
Skewness	-1.243	-1.106	0.732
Kurtosis	91.048	57.419	7.595
LM test	134.981	134.981	5.404
Soybeans			
β_0	5.0885 (6.129)	3.0443 (5.604)	3.284 (10.102)
β_1	1.9999	1.9985	1.8443
β_2	0.1965 (9.067)	0.3280 (10.897) **	0.2571 (8.276)
σ	0.1633 (6.070)	0.2455 (5.194)	0.1973 (6.095)
Log L	-736.329	-695.222	-96.207
nobs	2612	2612	541
Properties of standardized residuals			
Skewness	0.957	0.420	0.504
Kurtosis	25.284	13.684	15.660
LM test	6.279	6.279	20.013

Continued

Commodity	15 days prior to the first notice day and price limits	seasonal dummy variable *	weekly basis and removing data affected by price limits
Gold			
β_0	37.6107 (4.045)		32.5524 (1.995)
β_1	-10.7157		-3.5898
β_2	-0.3115 (-42.778)	-	-0.1410 (-6.557)
σ	-0.0244 (-3.699)	-	0.0305 (1.736)
Log L	-1124.459	-	-248.209
nobs	3007		518
Properties of standardized residuals			
Skewness	-4.349	-	0.272
Kurtosis	99.309	-	4.937
LM test	0.0069	-	35.018

The asymptotic t-statistics are in parentheses, based on the BFGS numerical gradient algorithm. The critical value at 5% significance level is 1.96.

(*) This quadratic variance function specification model was estimated removing data affected by price limit moves, rolling over the first nearby futures contract that matures 15 days prior to the first notice day and including a seasonal dummy variable. For coffee, the dummy variable is defined as 1 during March, April, May 0 in the rest of the year; for crude oil and heating oil, the dummy variable takes the value 1 during April, May, June, July, August and September and 0 in the other months of the year; for soybeans, the dummy variable is 1 during June, July, August and September and 0 otherwise. Notice that for the case of gold, a dummy variable was not introduced.

**** Indicating that a multiplicative dummy variable fitted better than an additive dummy variable.**

Table 5. Quadratic volatility function of the spread returns on spread size, with restriction on β_1.

Removing data affected by price limits , rolling over 15 days prior to the first notice day

	Coffee				Crude oil				Gold				Heating oil				Soybeans	
	l	X>l	X<u	u	l	X>l	X<u	u	l	X>l	X<u	u	l	X>l	X<u	u	l	u
relative spread (X)		0.949	1.009			0.997	1.047			0.9857	0.995			0.935	1.125			
standardized X	-4.603	-1.646	1.067	5.924	-0.822	-0.532	2.219	3.316	-37.611	-0.946	1.518	3.210	-3.047	-2.750	4.789	8.953	-5.0885	<u>-5.0891</u>
volatility $(\sigma q,t)$	8.9 X 10^{-5}	0.918	1.130	1.2 X 10^{-4}	6.2 X 10^{-6}	0.665	1.981	2.5 X 10^{-5}	1.4 X 10^{-15}	1.160	0.503	2.2 X 10^{-17}	0	0.140	1.316	0	0	1.5 X 10^{-16}

Removing data affected by price limits , rolled over 15 days prior to the first notice day and adjusting for seasonal dummy

	Coffee				Crude oil				Heating oil				Soybeans	
	l	X>l	X<u	u	l			u	l	X>u	X<l	u	l	u
relative spread (*X*)		0.949	1.009			0.977	1.057			0.965	1.125			
standardized spread	-3.839	-1.646	1.067	3.856	-1.451	-1.632	2.769	1.671	-2.000	-1.560	4.789	3.559	-5.553	<u>-5.554</u>
volatility $\sigma\left\|\beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-1}^2\right\|$	1.2 X 10^{-16}	0.826	0.937	0	0	0.287	2.225	0	4.9 X 10^{-5}	0.355	1.315	8.7 X 10^{-5}	0.5054	0.5055

Note: For the case of gold, only one presence of price limit move was found for the period analysed. Therefore, no adjustments were considered to estimate the log likelihood function (6). Also, it was not considered a seasonal dummy for gold.

Table 5. (continued)

Weekly data, removing data affected by price limits and rolling over a day after first notice day

	Coffee				Crude oil				Gold				Heating oil				Soybeans			
	l	X>l	X<u	u	l	X>l	X<u	u	l	X>l	X<u	u	l	X>l	X<u	u	l	X>l	X<u	u
relative spread (X)		0.969	1.039			0.997	1.048			0.987	0.994			0.974	1.154					
standardized X	-3.271	0.662	2.410	6.614	-0.777	-0.484	1.460	2.052	-32.55	-0.897	1.430	7.092	-3.656	-1.377	4.933	14.765	-3.284			-3.889
volatility ($\sigma q,t$)	2.6 X 10^{-5}	0.856	1.152	5.2 X 10^{-5}	0	0.633	1.129	0	1.9 X 10^{-13}	1.088	0.825	1.6 X 10^{-14}	5.9 X 10^{-6}	0.768	1.572	2.2 X 10^{-6}	1.1 X 10^{-5}			1.3 X 10^{-5}

Note: l and u are the estimated lower and upper bounds from the maximisation of the log likelihood function (6) considering the restriction on β_1 with weekly data.

Figure 1

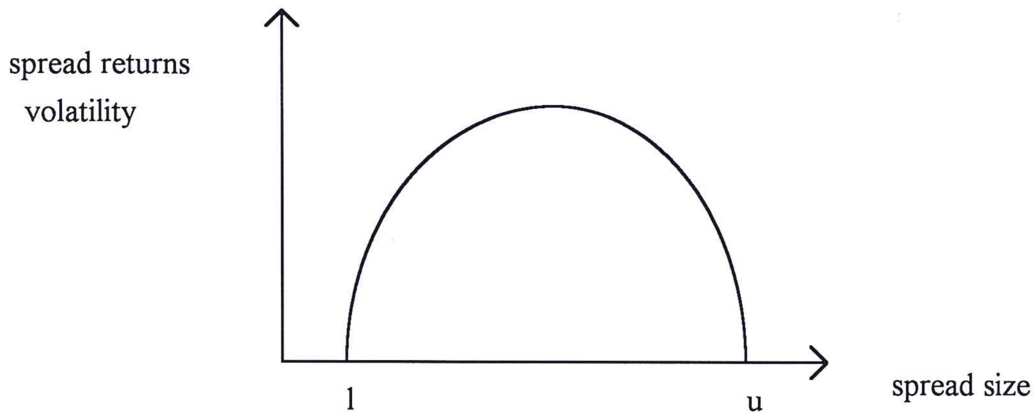

Figure 2

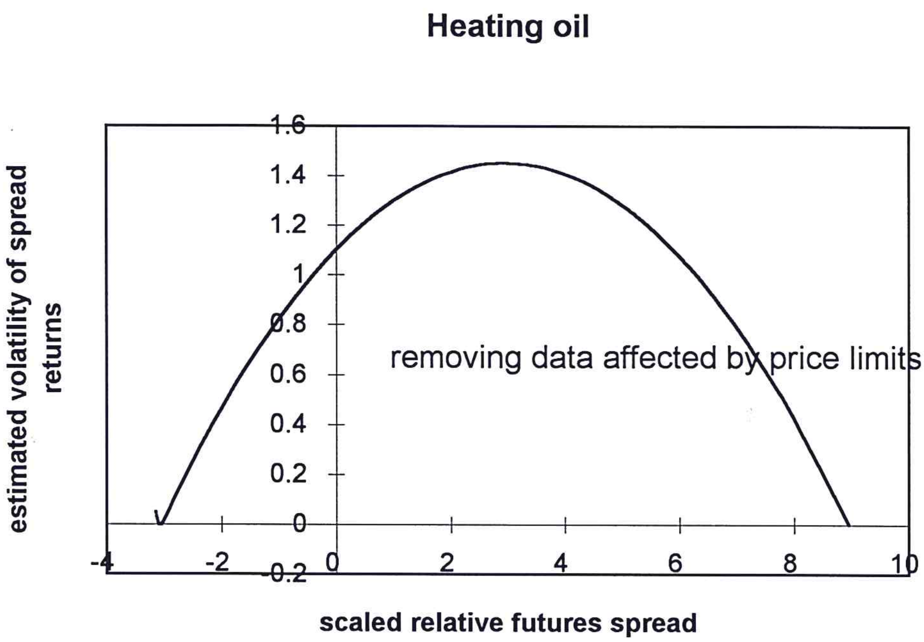

Figure 3