Co-integration in Crude Oil Components and the Pricing of Crack Spread Options

Jin-Chuan Duan* and Annie Theriault†

(December 6, 2007)

Abstract

The crack spread options traded at the New York Mercantile Exchange are American-style futures spread options on the one-to-one by volume difference between the futures price of a refined petroleum product and that of light sweet crude oil. We investigate the importance of co-integration and maturity effects in pricing the two most common of these options, namely Heating Oil/Crude and Gasoline/Crude spread options. We compare the performance of five models: a bivariate constant-volatility model, a bivariate GARCH model with and without maturity effects, and the same two GARCH models being augmented with co-integration. The model parameters are estimated using futures prices, and the theoretical option prices are computed using a primal simulation technique. The evidence for co-integration, stochastic volatility and maturity effects is strong in the futures prices. The option prices also show support for these data features.

1 Introduction

Spread options are options on the price difference between two or more assets. In this study, we focus on two types of standardized exchange traded spread options – crack spread options on the one-to-one by volume price difference between the futures prices of New York Harbor no. 2 heating oil and light sweet crude oil (Heating Oil/Crude) and between the futures prices of New York Harbor unleaded gasoline and light sweet crude oil (Gasoline/Crude), respectively.[1] Both are American-style options that lead to the delivery of two offsetting

*Duan is with Risk Management Institute and Department of Finance, National University of Singapore, and Rotman School of Management, University of Toronto. E-mail: bizdjc@nus.edu.sg.

†Theriault is with TD Securities. E-mail: Annie.Theriault@tdsecurities.com.

[1] Note that several other types of spread options are available at the NYMEX. They are across grades, times to maturity or other criteria and can be American- or European-style options.

futures positions if exercised. When a call option is exercised, the buyer receives a long product futures position and a short crude oil futures position. Conversely, when a put option is exercised, the buyer of the option receives a short product futures position and a long crude oil futures position. Note that there are no negative strikes needed, since a negative strike on a call is equivalent to a put with the same strike.

Several pricing models have been proposed for spread options, with most focusing on European-style options. Carmona (2003) provided an extensive review of these techniques and we refer the interested readers to that article for more details. In one of the earliest attempts at developing a model for pricing spread options, Margrabe (1978) focused on options with a strike price of zero, also referred to as "outperformance" or "exchange" options. For this special case, Margrabe (1978) showed that an analytical formula, akin to the Black-Scholes (1973) formula, can be developed under the assumption of correlated lognormal diffusions. However, numerical procedures had to be considered for the more general non-zero strike price case since exact analytical formulas cannot be derived under the bivariate normality assumption due to the fact that a linear combination of correlated lognormals is not lognormal. One of the commonly referenced numerical pricing techniques was proposed in Boyle (1988). In that study, he suggested pricing options with payoffs depending on two state variables under correlated lognormal diffusions with a five-point lattice. This approach has the advantage of being applicable to both European- and American-style options, but is computationally demanding. To reduce computational time, Kirk (1995) derived an analytical approximation for European options. This approximation was shown to yield prices with similar properties as the 5-point lattice method and is commonly cited in research aimed at practitioners.

Other underlying asset processes have been used to price spread options. In fact, some researchers have developed models based on dynamics of the spread directly. This approach is no longer common in the academic literature since it ignores the distribution of each asset and the correlation between them. Readers are referred to Carmona (2003) for a discussion of these models. More recent research in the area of spread option pricing has focussed on developing models with more realistic assumptions for the underlying assets' bivariate distributions and estimation procedures to solve for prices under these assumptions. Dempster and Hong (2000) developed a numerical technique using Fast Fourier transforms. They show that their procedure is flexible and can be used for models with stochastic volatility, jump diffusion and variance-gamma assumptions. Alexander and Scourse (2004) developed a model with changing volatility starting with the assumption that the marginal distribution of each asset is a mixture of lognormal distributions. They performed a theoretical study of their model for European-style options using an analytic approximation to solve for prices and showed that changing volatility impacts option prices and that their model is consistent with a correlation frown.

Lastly, Duan and Pliska (2004) developed the co-integration option pricing model that we use here to study crack spread options. Their approach is based on a discrete-time model where the assets are co-integrated process with multivariate GARCH volatility. They showed that the co-integration variable enters the pricing model only when volatility is time-varying. They performed a numerical study for European-style spread options, comparing the option prices obtained with their model to those obtained with a constant volatility model and with those obtained with a multivariate GARCH model without co-integration. Their results indicate that the time-varying volatility and the presence of co-integration both impact the prices obtained for spread options. Our work builds on this theoretical research by providing the first empirical study on the performance of this pricing approach on spread options.

Crack spread options present an ideal setting for investigating the impact of bivariate GARCH volatility and co-integration because their underlying assets intuitively exhibit these properties. First, as shown in Theriault (2007), the univariate series of crude oil, heating oil and gasoline futures returns exhibit volatility clustering, which justifies the use of a GARCH volatility formulation for pricing these options. Intuitively, crude oil, heating oil and gasoline futures prices should be unit root processes individually and that taken together, they should be co-integrated. Serletis (1992,1994) showed that crude oil, heating oil and gasoline futures are indeed unit root processes and are co-integrated, respectively. Another important feature to be considered is the presence of maturity effects in the volatility of commodity futures. The Samuelson effect and of an increase in volatility that coincides with a period of heavy contract switching related to hedging rollover have been documented in Theriault (2007) for the crude oil complex. These maturity effects are also found by Theriault (2007) to be important for pricing American futures options on individual commodities of the crude oil complex. Naturally, we will need to incorporate these maturity effects into the model of Duan and Pliska (2004) in order to adequately price the crack spread options in our sample.

The complexity of co-integration model with stochastic volatility leads us to the employment of Monte Carlo simulations in pricing options. Since crack spread options are of American style, one needs to deal with the challenge of dealing early exercise of an option using Monte Carlo simulations. For this, we adopt the primal Monte Carlo simulation technique pioneered by Carriere (1996) and made popular by Longstaff and Schwartz (2001). Particularly relevant to our paper is the application of the primal simulation to the GARCH option pricing model by Stentoft (2005). We adapt Stentoft's (2005) procedure to our spread option pricing model by adding four regressors, namely the price and variance of the second asset and their squares. The model parameters are estimated using the maximum likelihood method on the futures prices of crude oil, heating oil and gasoline.

This paper is organized as follows. Section 2 documents the joint time-series properties of the underlying futures prices and describes our options sample. Section 3 discusses our option pricing model and describes the estimation technique. Section 4 provides the empirical results, and Section 6 concludes.

2 Joint Time-Series Process

The univariate time-series properties of the returns series of crude oil, heating oil and gasoline futures have been documented in Theriault (2007) using settlement prices of futures contracts expiring between January 1995 and December 2005 (obtained from Reuters BridgeStation). The volatilities of these time-series of returns were shown to be time-varying with autoregressive properties that are consistent with a GARCH process. They were also shown to change systematically as futures contracts approach maturity. The Samuelson effect is a gradual increase in volatility as the time-to-maturity of a futures contract decreases and it is evident in the data. The contract switching effect related to hedging rollover was found to cause an increase in the volatility of futures returns when investors switch into futures positions with one additional month to maturity after closing expiring positions. We therefore incorporate these properties into our time-series process and pricing model.

An investigation of the joint dynamics of the two pairs of commodities underlying our options shows that they are correlated and co-integrated. In particular, Figure 1 shows two plots of daily closing prices, for crude oil with heating oil and for crude oil with gasoline. The data plotted are rolling time-series of alternating March and September delivery futures contract prices for the two pairs of series of interest. The prices of heating oil and gasoline were multiplied by 42 for these plots so that they represent the same amount of product by volume since crude oil is priced per barrel of product, while heating oil and gasoline are priced per gallon of product (42 gallons = 1 barrel). The plots show that the spreads are positive, with average spreads of \$4.44 per barrel for the heating oil/crude pair and \$5.89 per barrel for the gasoline/crude pair, which is consistent with refiners of crude oil making a return on their operations. Both plots exhibit high correlation since shocks to one series are accompanied by shocks to the paired series. The plots also show that the price series follow each other through time and the distance between the price series appear to be mean reverting. These properties are characteristic of co-integrated data sets.

Table 1 formalizes this observation by providing the correlations of prices and returns of the series as well as the Phillips-Ouliaris statistics for each pair of commodity futures prices using ten years of rolling March and September futures prices for the period beginning in June 1995 and ending in June 2005. The price correlations for the heating oil/crude and gasoline/crude pairs were 0.986 and 0.988, respectively, while the corresponding measures for returns were 0.873 an 0.874, respectively. These significant correlations indicate that futures prices are related by a positive long-run relationship as well as react to shocks in the same direction on a short-run basis. In the well known that treating such series as two correlated processes are inappropriate and can lead to seriously erroneous conclusions. Using a co-integration regression (allowing for a constant and a time trend) between two logarithmic price series, the Phillips-Ouliaris statistics (12 lags) are respectively −73.780 and −79.666 for the heating oil/crude and gasoline/crude pairs, which are statistically significant at the 5% level of confidence. Not surprisingly, these price serious are co-integrated which is a fact reflected in our pricing model in the next section.

3 Methodology

3.1 Pricing Model

Our pricing model is an application of the Duan and Pliska (2004) model to commodity futures. As discussed in the previous section, the empirical joint time-series process of crude oil and heating oil futures as well as the joint time-series process of crude oil and gasoline futures exhibit the GARCH-type volatility, maturity effects and are co-integrated. We therefore derive our pricing model using the following process for futures prices under the physical probability measure P:[2]

$$
\begin{aligned}
\ln\left(\frac{F_{1,t,T_1}}{F_{1,t-1,T_1}}\right) &= \lambda_1\sqrt{h_{1,t}} - \frac{1}{2}h_{1,t} + \delta_1 Z_{t-1} + \sqrt{h_{1,t}}\epsilon_{1,t}, & (1)\\
\ln\left(\frac{F_{2,t,T_2}}{F_{2,t-1,T_2}}\right) &= \lambda_2\sqrt{h_{2,t}} - \frac{1}{2}h_{2,t} + \delta_2 Z_{t-1} + \sqrt{h_{2,t}}\epsilon_{2,t}, & (2)\\
h_{1,t} &= \beta_{1,1}h_{1,t-1} + \beta_{1,2}h_{1,t-1}\left(\epsilon_{1,t-1} - \theta_1\right)^2 & \\
&\quad + \beta_{1,0}(T_1 - t)^{\gamma_1} + \eta_1 SW_{1,t}, & (3)\\
h_{2,t} &= \beta_{2,1}h_{2,t-1} + \beta_{2,2}h_{2,t-1}\left(\epsilon_{2,t-1} - \theta_1\right)^2 & \\
&\quad + \beta_{2,0}(T_2 - t)^{\gamma_2} + \eta_2 SW_{2,t}, & (4)\\
Z_t &= a + bt + c\ln\left(F_{2,t,T_2}\right) + \ln\left(F_{1,t,T_1}\right), & (5)
\end{aligned}
$$

where F_{i,t,T_i} is the futures price of commodity i at time t expiring at time T_i, λ_i is the constant unit risk premium for the i-th commodity price risk, $h_{i,t}$ is the time-varying conditional variance of the futures returns of commodity i, Z_{t-1} is the stationary error-correction variable and $\epsilon_{i,t}$ is the time-t innovation to returns with the bivarate normal distribution with mean $\mathbf{0}$ and covariance matrix $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$.

[2]Since $\ln\left(\frac{F_{i,t,T_i}}{F_{i,t-1,T_i}}\right) = \ln\left(\frac{S_{i,t}}{S_{i,t-1}}\right) - (r - y_i)$, where $S_{i,t}$ is the spot price of commodity i at time t, r is the risk free rate of interest and y_i is the net convenience yield of commodity i. This expression is consistent with constant interest rate and constant convenience yield.

We impose the restrictions $\beta_{i,0} > 0$, $\beta_{i,1} \geq 0$ and $\beta_{i,2} \geq 0$ to ensure that the conditional variances are positive. The restriction $\beta_{i,1} + \beta_{i,2}(1 + \theta_i^2) < 1$ may be imposed to ensure that the process is variance stationary (after removing the cyclical maturity effects). In addition, the restriction $|\delta_1 + \delta_2 * c| \geq 1$ is needed for the existence of co-integration.[3] The parameter θ_i is used to model the asymmetric reaction of the variance process to negative and positive innovations. Positive values for these parameters indicate that the conditional variance reacts more strongly to negative innovations than to positive ones, whereas negative ones imply a stronger response to positive innovations.

The term $(T_i - t)^{\gamma_i}$ is used to model the Samuelson effect, where $(T_i - t)$ is the time-to-maturity (in days) of the futures contract and γ_i is a parameter. A negative value for γ_i is consistent with the Samuelson effect. The dummy variable, $SW_{i,t}$, which is set to 1 during the switching period of each commodity futures and to 0 otherwise, allows us to model the increase in volatility that occurs during the contract switching period, as defined and modeled in Theriault (2007). Switching effects in volatility are observed between 24 and 32 trading days to maturity for crude oil, between 25 and 39 for heating oil and between 25 and 38 for gasoline. They correspond to periods where investors switch into futures positions while closing shorter maturity ones. As explained in Theriault (2007), switching effects near expiration occur together with Samuelson effect and cannot be separated from the Samuelson effect.

As in Duan and Pliska (2004), the local risk-neutralization principle of Duan (1995) can be applied to the physical joint time-series process to give the following process under the risk-neutral pricing measure Q, which can be used to price any derivative claim on two assets such as spread options:

$$
\begin{aligned}
\ln\left(\frac{F_{1,t,T_1}}{F_{1,t-1,T_1}}\right) &= -\frac{1}{2}h_{1,t} + \sqrt{h_{1,t}}\xi_{1,t}, && (6)\\
\ln\left(\frac{F_{2,t,T_2}}{F_{2,t-1,T_2}}\right) &= -\frac{1}{2}h_{2,t} + \sqrt{h_{2,t}}\xi_{2,t}, && (7)\\
h_{1,t} &= \beta_{1,1}h_{1,t-1} + \beta_{1,2}h_{1,t-1}\left(\xi_{1,t-1} - \lambda_1 - \theta_1 - \delta_1\frac{Z_{t-2}}{h_{1,t-1}}\right)^2 \\
&\quad +\beta_{1,0}(T_1 - t)^{\gamma_1} + \eta_1 SW_{1,t}, && (8)\\
h_{2,t} &= \beta_{2,1}h_{2,t-1} + \beta_{2,2}h_{2,t-1}\left(\xi_{2,t-1} - \lambda_2 - \theta_2 - \delta_2\frac{Z_{t-2}}{h_{2,t-1}}\right)^2 \\
&\quad +\beta_{2,0}(T_2 - t)^{\gamma_2} + \eta_2 SW_{2,t}, && (9)\\
Z_t &= a + bt + c\ln\left(F_{2,t,T_2}\right) + \ln\left(F_{1,t,T_1}\right), && (10)
\end{aligned}
$$

[3] These properties are discussed in Duan (1997) and Duan and Pliska (2004).

where $\xi_{i,t} = \epsilon_{i,t} + \lambda_i$ and has a bivariate standard normal distribution with respect to probability measure Q with mean of $\mathbf{0}$ and covariance matrix $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$.

3.2 Numerical Pricing Procedure

Since our data are a sample of American-style options, the pricing problem becomes solving an optimal stopping problem in a complex system with co-integration and stochastic volatility. Specifically, the time-t price of a spread option, P_t, depends only on four state variables: $F_{1,t,T_1}, F_{2,t,T_2}, h_{1,t+1}$ and $h_{2,t+1}$. Conceptually, its value can be obtained by solving the following expression:

$$\begin{aligned} & P_t(F_{1,t,T_1}, F_{2,t,T_2}, h_{1,t+1}, h_{2,t+1}) \\ = & \sup_{t \leq \tau \leq T^{op}} E^Q \left\{ e^{-r\tau} \max \left[\omega \left(42 * F_{2,\tau,T_2} - F_{1,\tau,T_1} - K \right), 0 \right] | \mathcal{F}_t \right\}, \end{aligned} \tag{11}$$

where F_{1,τ,T_1} is the futures price per barrel of crude oil at any stopping time, τ, F_{2,τ,T_2} is the futures price per gallon of heating oil or gasoline at any stopping time, τ, depending on the case of interest, T^{op} is the time-to-maturity of the spread option such that $T^{op} \leq \min\{T_1, T_2\}$, K is the exercise spread of the option and ω is a constant equal to 1 for a call option and -1 for a put option, and the time t information set is $\mathcal{F}_t$. The risk-free rate (r), although does not appear in the futures price system, will be needed for option pricing. We use in the empirical study the U.S. Treasury 3-month, constant-maturity middle rate (series FRTCM3M) obtained from DataStream.

Numerically, we adapt the primal simulation approach of Stentoft (2005) developed in the context of the GARCH option pricing model to spread options. In order to evaluate this expression by the least squares Monte Carlo simulation method, we follow the procedure described in Stentoft (2005) but add monomials that are functions of the second commodity's price and variance to account for the two additional state variables. More specifically, we use a total of four regressors for the constant volatility model; they are the price and price squared of each of two commodities. For models with stochastic volatiliy, we use eight regressors with the added four corresponding to the variance and variance squared for each of two commodities. We use 5,000 simulation paths and retain common random numbers for the all models in order to ensure that any difference in performance is not caused by using different sets of random numbers. We make the empirical martingale adjustment technique of Duan and Simonato (1998) and the antithetic simulation technique to improve simulation efficiency.

3.3 Parameter Estimates

As was done in Stentoft (2005), we obtain via maximum likelihood estimation the parameters needed for option pricing. We update the parameter estimates on a semi-annual basis using ten years of rolling series of alternating March and September contracts for all estimation. We control for the statistical impact of changing futures contract semi-annually to build a continuous time-series of returns. This is accomplished by using an actual return on the first day after we have changed contracts so that successive prices of the same series of futures are used in computing that return. We used the optimized model parameters along with the price time series to obtain starting variances and the initial co-integration variable where applicable. The parameters used for pricing were obtained with data available to a trader at the time of pricing. Option prices are not used in parameter estimation. They are only used as a means to examine the performance of different pricing spread pricing models.

A sample of parameters used for this pricing study is given in Table 2. This table gives the parameters obtained for the sample period ending in June 2005 for each pair of commodities for the five models of interest which are all nested in the physical time-series process used to develop our pricing model (for $T - t \geq 1$). The constant volatility model (CV) restricts all parameters except for $\beta_{i,0}$, λ_i and ρ to zero. We use this model as our benchmark because it is the commonly used model in the literature on spread options. The simple bivariate GARCH model (G11) restricts the maturity and error-correction parameters, γ_i, η_i and δ_i, to zero, and the co-integrated bivariate GARCH model (G11-C) restricts only the parameters related to maturity effects, γ_i and η_i, to zero. These models are similar to those given in Duan and Pliska (2004). The bivariate GARCH model with maturity effects (GM) restricts only the parameters related to co-integration, δ_i, to zero, and the co-integrated bivariate GARCH model with maturity effects (GM-C) is the most general model discussed above.

In addition to parameter values, we also include the maximum log-likelihood values and values for $Stat_i = \beta_{i,1} + \beta_{i,2}(1 + \theta_i^2)$ and $CS = |\delta_1 + \delta_2 * c|$ that help determine whether these parameters are consistent with a stationary process (apart from the maturity effects) and co-integration. The log-likelihood values provided in Table 2 can be used to perform likelihood ratio tests to show that the CV model is outperformed by all other models. Likelihood ratios are significant at the 5% level for all models tested against the CV model when compared with the appropriate critical values of χ^2. For instance, likelihood ratios comparing the GM-C model to the CV model are greater than the 5% critical χ^2_{12} value of 21.03. The GM model outperforms the G11 model since the log-likelihood ratios that are greater than the 5% critical χ^2_4 value of 9.48 for both pairs of commodities. The GM-C model outperforms the G11-C model with log-likelihood ratio that are also greater than the χ^2_4 value.

Additionally, the values for $\beta_{1,1} + \beta_{1,2}(1 + \theta_1^2)$ and $\beta_{2,1} + \beta_{2,2}(1 + \theta_2^2)$ are both less than one for all models, showing that our parameters are consistent with stationary volatility and $|1 + \delta_1 + c\delta_2|$ is less than one for all co-integration models thereby meeting a necessary conditions for co-integration.

4 Empirical Analysis of Crack Spread Options

4.1 Data

Our sample of spread option prices was obtained directly from the NYMEX and corresponds to the daily settlement prices of options traded between January 2004 and December 2005.[4] Each option contract relates to the futures contracts that expire in the same month as the option contract and trades until the last business day prior to the expiration of the crude oil futures contract, which in turn expires three business days before the 25^{th} of each month. There are up to 18 consecutive months of contracts available for Heating Oil/Crude spread options and 12 consecutive months of contracts available for Gasoline/Crude spread options, meaning that Heating Oil/Crude spread options have at most 378 trading days to maturity, while Gasoline/Crude spread options have at most 252 trading days to maturity.

Our sample has delivery dates between February 2004 and April 2006 as there are few options with positive open interest at the longest available times to maturity. Given the limited number of days with positive trading volumes for each option, we use all options that were within the 0.9 to 1.1 moneyness range, where moneyness is measured by $e^{-r(T^{op}-t)}(42 * F_{2,t,T_2} - F_{1,t,T_1})/K$. Our sample consists of $3,002$ Heating Oil/Crude and $3,430$ Gasoline/Crude spread options prices. Their distributions by option type, moneyness and time-to-maturity are given in Table 3. We define out-of-the money call options and in-the-money put options as those with a moneyness measurement smaller than 0.98, at-the-money options as those with a moneyness measurement between 0.98 and 1.02 and in-the-money call options and out-of-the-money put options as those with a moneyness measurement greater than 1.02. Short-term options are those with 21 trading days to maturity or less, medium-term options are those with between 22 and 60 trading days to maturity and long-term options are those with 61 trading days to maturity or more. Table 3 shows that call options outnumber put options for both samples, with calls accounting for 54% and 57% of the options sample for Heating Oil/Crude and Gasoline/Crude spread options, respectively. It also shows that our sample contains more out-of-the-money options than at- or in-of-the-money options, with at-the-money options having the smallest sample sizes in both cases. Our sample has more medium- and long-term option prices than short-term ones.

[4]Since settlement prices are given by the Exchange at the end of the open outcry period and since the futures and options all trade until 2:30 p.m., this sample contains synchronous prices of the derivative and its underlying.

Next, Tables 4 and 5 show the average open interest and trading volumes of the options analyzed, with the trading volumes data being limited to days where the trading volume was positive for averaging purposes. The average open interest of Heating Oil/Crude spread options ranges from 233 to 529 contracts, whereas the same measure for Gasoline/Crude spread options ranges from 152 to 511 contracts. The average trading volume for long- and medium-term options range from 67 to 286 for Heating Oil/Crude spread options and from 50 to 231 for Gasoline/Crude spread options. The average trading volumes when options have less than 21 trading days to maturity are lower, especially for out- and at-the-money options, not exceeding 56 contracts for Heating Oil/Crude and not exceeding 68 contracts for Gasoline/Crude, with smaller numbers for call options than put options, showing that the trades in these options near maturity are less active.

Lastly, Table 6 shows that the mean prices of these spread options are consistent with intrinsic value trends, with in-the-money options being priced higher than out-of-the-money options. They are also mostly consistent with the usual maturity trends since the mean price of longer-term American options of similar moneyness are greater than the mean price of shorter maturity options. There are some exceptions where the average prices of medium-term options are greater for both samples. This is not inconsistent with theory because the maturity trend is neither an intrinsic feature of commodity options nor that of American-style options. The mean prices of short-term Heating Oil/Crude spread options range from 0.84 to 1.44, whereas those of Gasoline/Crude spread options range from 0.76 to 1.32. The mean prices of long-term Heating Oil/Crude spread options range from 1.73 to 2.39, whereas those of Gasoline/Crude spread options range from 1.58 to 1.85.

4.2 Results without moneyness-maturity breakdowns

Table 7 gives the pricing errors obtained with the constant volatility model (CV), the bivariate GARCH model (G11), the co-integrated bivariate GARCH model (G11-C), the bivariate GARCH model with maturity effects (GM) and the co-integrated bivariate GARCH model with maturity effects (GM-C). Pricing errors are studied using the following five measures: the mean percentage pricing error ($MPE = 100\frac{1}{N}\sum_{i=1}^{N}\frac{P^{Mod}-P}{P}$), the median percentage pricing error ($MdPE = 100 median\left(\frac{P^{Mod}-P}{P}\right)$), the mean absolute percentage pricing error ($MAPE = 100\frac{1}{N}\sum_{i=1}^{N}\frac{abs(P^{Mod}-P)}{P}$), the median absolute percentage pricing error ($MdAPE = 100 median\left(\frac{abs(P^{Mod}-P)}{P}\right)$) and the relative root mean square error ($RRMSE = 100\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{P^{Mod}-P}{P}\right)^2}$).

Our results show that the CV model is outperformed by more sophisticated models and that maturity effects are important for pricing purposes. We find the simple bivariate

GARCH and co-integrated bivariate GARCH models without maturity effects, that is the G11 and G11-C models, do not outperform the CV model. For both of these models, most of our pricing error measures are greater for these two models than for the CV model. When the maturity effects are included, we note significant improvements in pricing accuracy. The last two rows of Panels A and B show that both the bivariate GARCH and co-integrated bivariate GARCH models with maturity effects (GM and GM-C) outperform the CV model in terms of most pricing error measures. In fact, the median percentage error, mean and median absolute percentage error and relative root mean square error are lower with the maturity models for both types of crack spread options studied. For Heating Oil/Crude spread options the GM-C model also gives a lower mean percentage error than the CV model. For the Heating Oil/Crude spread options sample, we find that the median percentage pricing error, mean and median absolute percentage pricing error and relative root mean square measures decrease from −7.35%, 29.99%, 21.72% and 45.14% with the CV model to −1.79%, 24.57%, 17.92% and 39.70% with the GM model. We also find that these pricing error measures decrease to −5.87%, 25.24%, 19.78% and 39.10% with the GM-C model. In addition, with the GM-C model, the mean percentage pricing error of 1.69% obtained with the CV model decreases to 0.58%.

Our findings are similar for Gasoline/Crude spread options. We find that the median percentage pricing error, mean and median absolute percentage pricing error and relative root mean square measures decrease from −5.85%, 26.45%, 24.87% and 30.67% with the CV model to −2.55%, 21.26%, 19.26% and 25.46% with the GM model and to 2.83%, 22.12%, 18.28% and 28.14% with the GM-C model with maturity effects. However, neither the GM model nor the GM-C model yields lower mean percentage pricing error than the CV model.

Unfortunately, the results obtained for both samples, do not allow us to decide where the model with the co-integration feature is preferred. Between the two models with maturity effects, some measures of pricing errors are lowest for the GM model, while others are lowest for the GM-C option pricing model. Thus, the overall result is mixed. Next we will examine the pricing results by moneyness and maturity category and with a regression analysis of the absolute percentage pricing errors to gain further insight into these models.

4.3 Results by Moneyness and Maturity Categories

Tables 8 and 9 show the pricing errors obtained with each pricing model for our option sample classified into different moneyness and maturity categories.

Table 8 provides the pricing errors obtained for Heating Oil/Crude spread options. These results are consistent with those in Table 7 discussed earlier. We find that the G11 and G11-C models do outperform the CV model in some moneyness categories for short- and medium-term options. They are outperformed by the CV model by most measures for long-term options. The GM and GM-C models outperform all other models by most measures for short- and long- term options and the results are mixed for medium-term options. We find that the GM-C model, tends to outperform in the most moneyness categories for long-term options.

The results given in Table 9 for Gasoline/Crude spread options are more mixed. Like for Heating Oil/Crude spread options, the G11 and G11-C models only outperform the CV model in some moneyness categories for short- and medium-term options. We find that the GM model outperforms the CV model by most measures for short-term options and that the GM-C model does have lower MPE and MAPE than the GM model in all moneyness categories for these options. For medium- and long-term options, we find that the two maturity models have lower error measures than the CV model in all moneyness categories. Both the GM and GM-C models outperform the CV model by most measures for medium- and long-term options. The GM-C outperforms the GM model by most measures for medium-term options, wile the GM model outperforms by all error measures for long-term options.

In summary, these results from different moneyness-maturity combinations are consistent with those observed in Table 7 in that they show that the G11 and G11-C models generally outperform the CV model, and that the two models with maturity effects (GM and GM-C) outperform the other models by most measures. Since there is no clear rankings between the GM and GM-C models, the evidence is not in support of co-integration as far as the crack spread options are concerned.

4.4 Absolute Percentage Pricing Error Regression

We regressed the absolute percentage pricing errors against some potential sources of systematic pricing errors in order to gain a better understanding of our results. More specifically, we run the following regression:

$$\begin{aligned} APE &= a_1 + a_2(T-t) + a_3(T-t)^2 + a_4 M_t + a_5 M_t^2 + a_6\sigma_{1,t} + a_7\sigma_{1,t}^2 + a_8\rho_t \\ &\quad + a_9(T-t)I^P + a_{10}(T-t)^2 I^P + a_{11}M_t I^P + a_{12}M_t^2 I^P + a_{13}\sigma_{1,t} I^P + a_{14}\rho_t I^P + \nu_t, \end{aligned}$$

where $(T-t)$ is the time to maturity of the options contract in years, M_t is the moneyness of the option, $\sigma_{1,t}$ is the 20-day standard deviation of crude oil futures, ρ_t is the 20-day correlation between crude oil futures and heating oil (or gasoline) futures, and I^P is an indicator equal to 1 for put options and 0 otherwise and ν_t is an error term. This regression allows us to determine whether the more sophisticated models are able to remove some of the systematic pricing errors associated with the benchmark CV model.

The parameters and explanatory power of the regressions, as measured R^2, are given in Table 10. Consistent with our earlier findings that GARCH models without maturity effects do not outperform the CV model, we find that the R^2 of the regressions for the G11 and G11-C models are either greater than or near that of the CV model for both types of crack spread options. More specifically, for both options samples, the R^2 of these regressions are all greater than 30%. On the other hand, we find that the explanatory power of the regressions are significantly lower for models that include maturity effects. The R^2 decrease from 0.301 with the CV model to 0.121 with the GM model and 0.111 with the GM-C for Heating Oil/Crude spread options, and from 0.363 with the CV model to 0.135 with the GM model and 0.070 with the GM-C model for Gasoline/Crude spread options.

In summary, our regression analysis shows that the prices obtained with the GM-C model are less prone to systematic errors than than those obtained with all other models. This suggests that both maturity effects and co-integration are important for pricing these options.

5 Conclusion

Our work finds support for the use of a model that incorporates both co-integration and maturity effects. The GM-C model yields pricing errors that are the lowest for both spread option types and our regression analysis shows that this model's pricing errors are less related to systematic factors than the other models tested.

However, we note that the pricing errors of the GM-C model remains relatively large. We believe that the GM-C model has the potential to perform better than our simple empirical study was able to show. We remind the reader that many implementation choices are needed and they undoubtedly will impact the performance of the models. More specifically, we had to select specific futures contracts and sample periods for time-series analysis as well as to decide on a frequency for rolling futures returns data to construct the time-series. Even though these choices do not theoretically matter, in practice, they are likely to impact our results. In pre-tests, we found that although the ordering of the models did not usually change when different time-series are constructed for price estimation, these choices can significantly change parameter estimates and the size of the pricing errors.

Two avenues for future research are suggested. One way is through further study of the underlying price series. Unlike studies on stock options, where all options share the same underlying asset, commodity futures options or crack spread options of different times to maturity are written on different underlying futures contracts. In our implementation, we used the same time-series to obtain parameters for all options and rolled futures contracts semi-annually for both studies. A user of the model may wish to try different combinations of futures contracts and rolling frequencies. Furthermore, although there are no theoretical reasons for using different time-series models for modeling the futures contracts expiring in different months, it is possible that there are additional market effects not modeled herein (such as seasonality) that could be incorporated in the model for different maturity months.

A second avenue is to use recent option prices to calibrate parameter values to take advantage of the information embedded in option data. We note that such an approach has been implemented in some study of the GARCH option pricing model on European stock options. This procedure is generally computer intensive. The fact that crack spread options are American-style will make such an approach even more intensive. The increasing availability of grid computing networks may soon make such an approach practical.

References

1. Alexander, C and A. Scourse, 2004, "Bivariate normal mixture spread option valuation," *Quantitative Finance* 4, 637–648.

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

3. Boyle, P., 1988, "A lattice framework for option pricing with two state variables," *Journal of Financial and Quantitative Analysis* 23, 1–12.

4. Carmona, R., 2003, "Pricing and hedging spread options," *Society for Industrial and Applied Mathematics* 45, 627–685.

5. Carriere, J., 1996, "Valuation of the early-exercise price for options using simulations and nonparametric regression," *Insurance: Mathematics and Economics* 19, 19–30.

6. Dempster, M. and G. Hong, 2000, "Spread option valuation and the fast fourier transform," Research Papers in Management Studies Working Paper.

7. Duan, J.-C., 1995, "The GARCH option pricing model," *Mathematical Finance* 5, 13–32.

8. Duan, J.-C., 1997, "Augmented GARCH(p,q) Process and its Diffusion Limit," *Journal of Econometrics* 79, 97–127.

9. Duan, J.-C. and S. Pliska, 2004, "Option valuation with co-integrated asset prices," *Journal of Economic Dynamics and Control* 28, 727–754.

10. Duan, J.-C. and J.-G. Simonato, 1998, "Empirical martingale simulation for asset prices," *Management Science* 44, 1218–1233.

11. Hsieh, K. and P. Ritchken, 2005, "An empirical comparison and GARCH option pricing models," *Review of Derivatives Research* 8, 129–150.

12. Kirk, E., 1995, "Correlation in the energy markets: managing energy price risk," *Risk*, 71–78.

13. Margrabe, W., 1978, "The value of an option to exchange one asset for another," *Journal of Finance* 33, 177-186.

14. Longstaff, F. and E. Schwartz, 2001, "Valuing American options by simulation: a simple least-squares approach," *Review of Financial Studies* 14, 113–147.

15. Serletis, A., 1992, “Unit root behaviour in energy futures prices,” *Energy Journal* 13, 119–128.

16. Serletis, A., 1994, “A cointegration analysis of petroleum futures prices,” *Enery Economics* 16, 93–97.

17. Stentoft, L., 2005, “Pricing American options when the underlying asset follows GARCH processes,” *Journal of Empirical Finance* 12, 576–611.

18. Theriault, A., 2007, “An empirical investigation of futures and options in the oil and gas market,” PhD thesis, University of Toronto.

Figure 1: Co-integration

Rolling series of alternating March and September delivery futures prices for crude oil together with heating oil (top chart) and together with gasoline (bottom chart) to illustrate the co-integration property of these data sets. These plots show that the futures returns of the two pairs of commodities underlying the crack spread options are highly correlated and likely co-integrated.

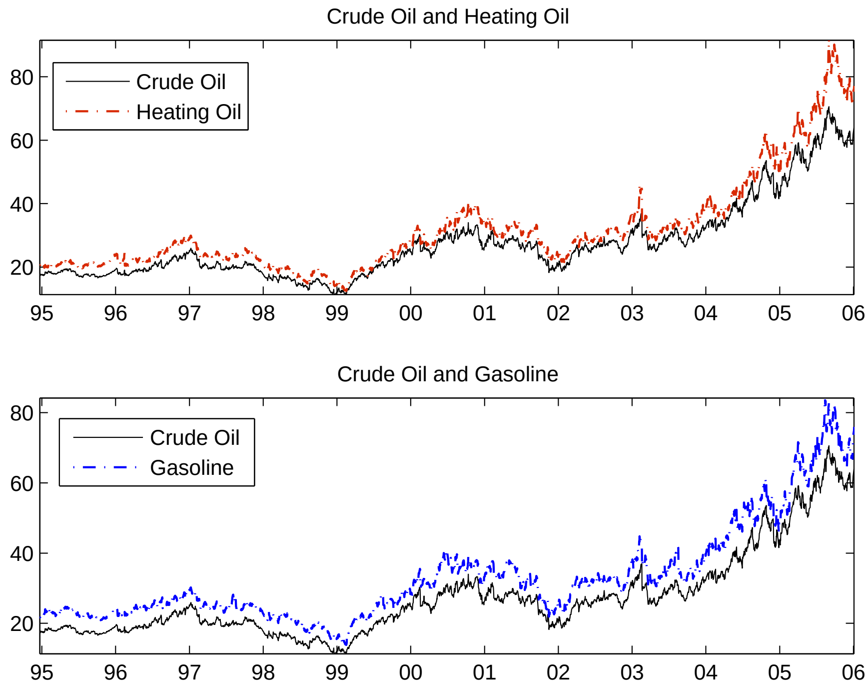

Table 1: Bivariate time-series properties

These basic statistics were obtained for time-series of alternating March and September delivery contracts for a ten-year sample period beginning in June 1995 and ending in June 2005. These statistics show that the futures returns of the two pairs of commodities underlying the crack spread options are highly correlated and co-integrated.

	Correlation of Prices	Correlation of Returns	Phillips-Ouliaris Statistics
Heating Oil/Crude	0.986	0.873	-73.780
Gasoline/Crude	0.988	0.874	-79.666

Table 2: Parameter estimates

This table shows the parameter estimates and corresponding t-ratios for the model in equations (1)-(5). The constant volatility model (CV) restricts all parameters except for $\beta_{i,0}$, λ_i and ρ to zero, the simple bivariate GARCH model (G11) restricts the maturity and error-correction parameters, γ_i, η_i and δ_i, to zero, the co-integrated bivariate GARCH model (G11-C) restricts γ_i and η_i, to zero the bivariate GARCH model with maturity effects (GM) restricts δ_i to zero and the co-integrated bivariate GARCH model with maturity effects (GM-C) is as shown above. These parameters were obtained for a time-series of alternating March and September contracts for ten-year sample period ending in June 2005. $Stat_i$ and CS are the test statistics as described in the text. t-statistics are provided for each parameter estimate in brackets. Likelihood ratio tests performed on these data show that our proposed model outperforms the restricted models for the data samples analyzed.

	Heating Oil/Crude					Gasoline/Crude				
	CV	**G11**	**G11-C**	**GM**	**GM-C**	**CV**	**G11**	**G11-C**	**GM**	**GM-C**
Mean Parameters										
λ_1	0.0288	0.0338	0.0380	0.0250	0.0095	0.0328	0.0423	0.0447	-0.0462	0.1031
	(1.36)	(1.55)	(1.76)	(1.18)	(7.60)	(1.57)	(2.06)	(2.10)	(-2.32)	(4.92)
λ_2	0.0296	0.0361	0.0396	0.0474	0.0134	0.0308	0.0432	0.0405	-0.0612	0.1024
	(1.43)	(1.68)	(1.83)	(2.23)	(8.94)	(1.49)	(2.04)	(1.89)	(-3.01)	(4.86)
δ_1			-0.0084		-0.0034			-0.0183		-0.0130
			(-1.03)		(-0.39)			(-3.08)		(-1.96)
δ_2			0.0238		0.0031			0.0166		0.0036
			(2.79)		(0.35)			(2.81)		(0.54)
Volatility Parameters										
$\beta_{1,0}$	4.03E-04	3.73E-05	5.10E-05	3.67E-05	3.09E-06	4.16E-04	1.72E-05	4.24E-05	5.30E-05	2.28E-05
	(70.72)	(6.17)	(4.34)	(8.03)	(4.96)	(72.07)	(2.87)	(8.06)	(11.19)	(5.04)
$\beta_{1,1}$		0.8694	0.7995	0.9459	0.9753		0.9469	0.8196	0.9476	0.9410
		(50.43)	(17.51)	(158.57)	(357.52)		(53.79)	(50.44)	(180.57)	(162.15)
$\beta_{1,2}$		0.0329	0.0054	0.0154	0.0095		0.0080	0.0431	0.0123	0.0114
		(8.06)	(2.78)	(7.01)	(7.60)		(3.66)	(10.13)	(7.49)	(8.34)
θ_1		-0.4212	-3.5608	0.3278	0.7232		-0.4618	-0.9584	0.3262	-1.0667
		(-4.82)	(-2.74)	(1.93)	(5.03)		(-3.60)	(-13.67)	(2.40)	(-8.01)
γ_1				-0.2775	-0.1747				-0.3675	-0.1825
				(-9.72)	(-2.91)				(-12.21)	(-5.23)
η_1				1.75E-05	4.70E-05				1.29E-05	6.23E-05
				(4.09)	(18.23)				(2.85)	(7.57)
$\beta_{2,0}$	4.21E-04	4.57E-05	4.73E-05	8.19E-06	1.02E-05	3.80E-04	4.70E-07	2.39E-05	1.42E-05	2.17E-05
	(104.03)	(9.58)	(8.92)	(3.69)	(4.08)	(67.52)	(3.11)	(9.55)	(2.93)	(3.71)
$\beta_{2,1}$		0.8156	0.8121	0.9600	0.9677		0.9969	0.8669	0.9665	0.9215
		(55.60)	(44.82)	(259.65)	(335.63)		(2600.60)	(77.45)	(250.20)	(130.38)
$\beta_{2,2}$		0.0668	0.0258	0.0140	0.0134		0.0012	0.0566	0.0099	0.0134
		(11.37)	(6.87)	(8.43)	(8.94)		(6.53)	(11.92)	(6.69)	(6.92)
θ_2		-0.4094	-1.3624	0.1383	0.2674		0.7237	-0.6378	0.4340	-1.4212
		(-6.99)	(-7.52)	(1.57)	(3.05)		(2.42)	(-11.59)	(3.48)	(-7.63)
γ_2				-0.0704	-0.2628				-0.2373	-0.1613
				(-1.21)	(-5.35)				(-3.15)	(-2.86)
η_2				3.96E-05	3.55E-05				3.06E-05	7.44E-05
				(12.97)	(20.75)				(8.24)	(7.79)
Correlation and Co-integration Parameters										
ρ	0.8790	0.8747	0.8790	0.8767	0.8800	0.8735	0.8735	0.8831	0.8776	0.8922
	(489.30)	(332.99)	(352.32)	(336.00)	(347.86)	(311.11)	(311.11)	(341.99)	(292.84)	(296.67)
a			-3.5421808		-3.542181			-3.5419684		-3.54197
			(-1088)		(-1088)			(-1028)		(-1028)
b			-2.70E-05		-2.70E-05			-6.25E-06		-6.25E-06
			(-18.94)		(-18.94)			(-3.61)		(-3.61)
c			-0.9558		-0.9558			-1.0548		-1.0548
			(-271.29)		(-271.29)			(-265.98)		(-265.98)
Log-Lik	18975	18989	19034	19125	19128	18950	18972	19043	19136	19186
$Stat_1$		0.908	0.873	0.963	0.990		0.957	0.902	0.961	0.965
$Stat_2$		0.894	0.886	0.974	0.982		0.999	0.947	0.978	0.962
CS			0.969		0.994			0.964		0.983

Table 3: Distribution by type, moneyness and maturity

Moneyness is measured by $e^{-r(T^{op}-t)}(42 * F_{2,t,T} - F_{1,t,T})/K$. Out-of-the money call options and in-the-money put options are those with moneyness measurement less than 0.98, at-the-money call and put options are those with moneyness measurement between 0.98 and 1.02 and in-the-money options are those with moneyness measurement greater than 1.02. Short-term options are those with 21 trading days to maturity or less, medium-term options are those with between 22 and 60 trading days to maturity and long-term options are those with 61 trading days to maturity or more.

	All				**Calls**				**Puts**			
Moneyness:	**out**	**at**	**in**	**Total**	**out**	**at**	**in**	**Total**	**out**	**at**	**in**	**Total**
A. Heating Oil/Crude Futures Spread Options												
Short-term	383	170	279	**832**	218	89	142	**449**	165	81	137	**383**
Medium-term	515	198	315	**1028**	278	117	180	**575**	237	81	135	**453**
Long-term	506	262	374	**1142**	221	148	228	**597**	285	114	146	**545**
Total	**1404**	**630**	**968**	**3002**	**717**	**354**	**550**	**1621**	**687**	**276**	**418**	**1381**
B. Gasoline/Crude Futures Spread Options												
Short-term	414	179	317	**910**	234	102	161	**497**	180	77	156	**413**
Medium-term	608	308	526	**1442**	346	186	309	**841**	262	122	217	**601**
Long-term	422	252	404	**1078**	206	127	286	**619**	216	125	118	**459**
Total	**1444**	**739**	**1247**	**3430**	**786**	**415**	**756**	**1957**	**658**	**324**	**491**	**1473**

Table 4: Average open interest by type, moneyness and maturity

The average open interest is lower for short-term options than for options with longer times to maturity.

	All			Calls			Puts		
Moneyness:	**out**	**at**	**in**	**out**	**at**	**in**	**out**	**at**	**in**
A. Heating Oil/Crude Futures Spread Options									
Short-term	241	233	281	199	189	200	298	281	365
Medium-term	423	381	391	363	396	423	493	359	349
Long-term	529	507	450	428	463	454	608	563	445
B. Gasoline/Crude Futures Spread Options									
Short-term	152	190	205	114	140	156	202	255	256
Medium-term	312	319	285	232	256	202	417	414	404
Long-term	511	449	380	470	401	393	550	499	347

Table 5: Average trading volume by type, moneyness and maturity

The average trading volume is lower for short-term options than for options with longer times to maturity.

	All			Calls			Puts		
Moneyness:	**out**	**at**	**in**	**out**	**at**	**in**	**out**	**at**	**in**
A. Heating Oil/Crude Futures Spread Options									
Short-term	61	64	56	31	53	31	122	97	97
Medium-term	94	127	67	62	126	62	129	127	95
Long-term	248	286	116	263	258	94	236	300	149
B. Gasoline/Crude Futures Spread Options									
Short-term	20	68	37	14	62	29	29	78	46
Medium-term	50	80	121	45	103	25	61	48	299
Long-term	231	89	199	199	109	263	291	5	40

Table 6: Average price by type, moneyness and maturity

The distribution of these prices follow the expected moneyness and maturity trend with prices being greater at longer times to maturity and greater moneyness.

	All			Calls			Puts		
Moneyness:	**out**	**at**	**in**	**out**	**at**	**in**	**out**	**at**	**in**
A. Heating Oil/Crude Futures Spread Options									
Short-term	0.84	1.12	1.44	0.87	1.15	1.49	0.81	1.08	1.38
Medium-term	1.39	1.88	2.26	1.40	1.90	2.25	1.37	1.85	2.27
Long-term	1.73	1.95	2.39	1.63	1.87	2.13	1.80	2.05	2.81
B. Gasoline/Crude Futures Spread Options									
Short-term	0.76	1.00	1.32	0.78	1.05	1.41	0.74	0.95	1.23
Medium-term	1.15	1.37	1.74	1.16	1.44	1.89	1.12	1.28	1.52
Long-term	1.61	1.58	1.85	1.60	1.45	1.84	1.62	1.72	1.85

Table 7: Pricing errors

This table shows the mean percentage pricing error ($MPE = 100\frac{1}{N}\sum_{i=1}^{N}\frac{P^{Mod}-P}{P}$), the median percentage pricing error ($MdPE = 100 median\left(\frac{P^{Mod}-P}{P}\right)$), the mean absolute percentage pricing error ($MAPE = 100\frac{1}{N}\sum_{i=1}^{N}\frac{abs(P^{Mod}-P)}{P}$), the median absolute percentage pricing error ($MdAPE = 100 median\left(\frac{abs(P^{Mod}-P)}{P}\right)$) and the relative root mean square error ($RRMSE = 100\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{P^{Mod}-P}{P}\right)^2}$) for the two data samples for the four test models and the benchmark CV model. These results favor the maturity models, with different price measures favoring the GM and GM-C models for both samples of options.

	MPE	MdPE	MAPE	MdAPE	RRMSE
A. Heating Oil/Crude Futures Spread Options					
CV	1.69	-7.35	29.99	21.72	45.14
G11	6.12	-4.36	30.39	20.27	47.90
G11-C	5.61	-4.88	30.35	20.02	48.02
GM	4.08	-1.79	24.57	17.92	39.70
GM-C	0.58	-5.87	25.24	19.78	39.10
B. Gasoline/Crude Futures Spread Options					
CV	-1.31	-5.85	26.45	24.87	30.67
G11	6.51	6.12	28.24	26.39	32.93
G11-C	9.91	7.20	27.25	24.58	32.95
GM	-1.43	-2.55	21.26	19.26	25.46
GM-C	4.12	2.83	22.12	18.28	28.14

Table 8: Heating Oil/Crude pricing errors by moneyness and maturity

This table shows the mean percentage pricing errors, $MPE = 100\frac{1}{N}\sum_{i=1}^{N}\frac{P^{Mod}-P}{P}$, the mean absolute percentage pricing errors, $MAPE = 100\frac{1}{N}\sum_{i=1}^{N}\frac{abs(P^{Mod}-P)}{P}$, and the relative root mean square error, $RRMSE = 100\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{P^{Mod}-P}{P}\right)^2}$, for our heating oil crack spread options sample for by moneyness and maturity categories. These results show that the co-integration model mainly outperforms for long-term options.

i. Short-term									
	Out-of-the-money			**At-the-money**			**In-the-money**		
Model	**MPE**	**MAPE**	**RRMSE**	**MPE**	**MAPE**	**RRMSE**	**MPE**	**MAPE**	**RRMSE**
CV	-34.23	34.58	36.58	-27.38	27.38	26.41	-19.55	19.55	21.07
G11	-26.16	27.50	33.57	-23.20	23.47	24.26	-15.71	16.05	18.95
G11-C	-26.60	27.53	34.45	-22.43	22.50	24.79	-15.72	16.18	19.43
GM	-18.62	24.31	33.85	-16.65	18.42	22.34	-11.52	14.24	19.14
GM-C	-25.20	30.12	35.68	-21.14	23.75	24.48	-14.77	17.97	21.75
ii. Medium-term									
	Out-of-the-money			**At-the-money**			**In-the-money**		
Model	**MPE**	**MAPE**	**RRMSE**	**MPE**	**MAPE**	**RRMSE**	**MPE**	**MAPE**	**RRMSE**
CV	-16.18	19.55	46.44	-17.92	21.20	45.03	-17.21	18.46	22.42
G11	-12.52	19.46	51.10	-14.37	16.90	44.88	-13.23	14.89	22.00
G11-C	-13.76	19.79	51.28	-15.24	17.48	44.67	-14.39	16.08	21.91
GM	1.76	23.50	55.94	-5.23	20.99	49.62	-11.06	17.68	22.07
GM-C	-4.43	24.53	54.71	-9.77	23.50	48.31	-14.18	19.44	24.56
iii. Long-term									
	Out-of-the-money			**At-the-money**			**In-the-money**		
Model	**MPE**	**MAPE**	**RRMSE**	**MPE**	**MAPE**	**RRMSE**	**PME**	**MAPE**	**RRMSE**
CV	16.07	19.48	58.53	17.23	20.67	62.68	12.26	17.00	49.80
G11	19.08	22.88	64.13	19.68	23.87	67.51	14.60	18.94	53.85
G11-C	17.34	22.75	63.74	18.40	21.65	67.33	13.35	18.38	54.58
GM	2.76	13.33	40.95	2.81	12.45	42.49	0.91	8.24	37.72
GM-C	-1.76	13.79	38.63	-0.51	10.91	39.76	-2.12	7.98	36.77

Table 9: Gasoline/Crude pricing errors by moneyness and maturity

This table shows the mean percentage pricing errors, $MPE = 100\frac{1}{N}\sum_{i=1}^{N}\frac{P^{Mod}-P}{P}$, the mean absolute percentage pricing errors, $MAPE = 100\frac{1}{N}\sum_{i=1}^{N}\frac{abs(P^{Mod}-P)}{P}$, and the relative root mean square error, $RRMSE = 100\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{P^{Mod}-P}{P}\right)^2}$, for our gasoline crack spread options sample for by moneyness and maturity categories. These results show that the co-integration model mainly outperforms for short- and medium-term options.

i. Short-term									
	Out-of-the-money			**At-the-money**			**In-the-money**		
Model	**MPE**	**MAPE**	**RRMSE**	**MPE**	**MAPE**	**RRMSE**	**MPE**	**MAPE**	**RRMSE**
CV	-38.61	38.77	39.21	-29.58	29.91	29.63	-19.32	20.31	20.77
G11	-23.30	30.01	34.71	-15.99	23.05	28.25	-11.13	16.64	20.71
G11-C	-21.40	30.34	33.20	-15.44	23.75	26.65	-9.09	15.97	18.97
GM	-15.39	28.79	34.88	-9.98	20.31	29.66	-4.40	15.44	19.84
GM-C	3.05	21.76	40.60	3.06	17.16	36.77	2.61	12.30	24.67
ii. Medium-term									
	Out-of-the-money			**At-the-money**			**In-the-money**		
Model	**MPE**	**MAPE**	**RRMSE**	**MPE**	**MAPE**	**RRMSE**	**MPE**	**MAPE**	**RRMSE**
CV	-16.47	26.46	29.00	-9.21	20.84	22.86	-11.75	18.16	19.79
G11	-4.89	28.66	31.86	2.39	22.84	26.26	-6.70	19.51	22.42
G11-C	-2.27	24.46	29.07	1.72	21.41	24.70	-3.43	16.09	19.96
GM	-11.60	22.25	28.15	-7.94	18.74	22.90	-10.72	17.79	20.36
GM-C	-11.52	21.33	28.57	-8.12	17.24	22.60	-9.30	15.29	19.40
iii. Long-term									
	Out-of-the-money			**At-the-money**			**In-the-money**		
Model	**MPE**	**MAPE**	**RRMSE**	**MPE**	**MAPE**	**RRMSE**	**MPE**	**MAPE**	**RRMSE**
CV	27.46	28.70	37.35	27.87	27.87	38.36	22.21	23.99	33.68
G11	34.71	38.76	43.09	36.86	36.91	43.69	29.45	31.15	38.01
G11-C	38.48	38.48	45.87	40.67	40.67	46.59	31.85	31.85	41.02
GM	6.83	17.50	24.06	10.00	14.36	22.20	8.62	18.52	22.68
GM-C	9.31	19.41	26.61	11.28	15.89	25.50	10.99	20.90	26.55

Table 10: Absolute percentage pricing error regression

This table gives the parameter estimates for the following regression equation:

$$APE = a_1 + a_2(T-t) + a_3(T-t)^2 + a_4 M_t + a_5 M_t^2 + a_6\sigma_{1,t} + a_7\sigma_{1,t}^2 + a_8\rho_t + a_9(T-t)I^P + a_{10}(T-t)^2 I^P + a_{11}M_t I^P + a_{12}M_t^2 I^P + a_{13}\sigma_{1,t}I^P + a_{14}\rho_t I^P + \nu_t,$$

where $(T-t)$ is the time to maturity of the options contract in years, M_t is the moneyness of the option, $\sigma_{1,t}$ is the 20-day standard deviation of crude oil futures, ρ_t is the 20-day correlation between crude oil futures and heating oil (or gasoline) futures, I^P is an indicator equal to 1 for put options and 0 otherwise and ν_t is an error term. These results show that the GM-C models yields the least systematic pricing errors.

A. Heating Oil/Crude Futures Spread Options										
	CV		**G11**		**G11-C**		**GM**		**GM-C**	
a_1	0.97	0.54	0.80	0.41	0.85	0.44	2.40	1.30	2.35	1.32
a_2	-0.39	-2.54	-0.19	-1.11	-0.25	-1.46	-0.09	-0.57	-0.23	-1.50
a_3	1.24	5.24	1.08	4.17	1.13	4.38	0.31	1.28	0.50	2.09
a_4	-1.79	-0.50	-2.47	-0.63	-2.64	-0.68	-5.61	-1.51	-5.75	-1.61
a_5	0.53	0.29	0.86	0.44	0.95	0.49	2.34	1.26	2.39	1.34
a_6	-3.82	-7.25	-2.89	-5.02	-3.49	-6.08	-3.72	-6.81	-1.91	-3.62
a_7	5.30	7.02	3.44	4.18	4.41	5.36	5.14	6.57	2.67	3.55
a_8	1.36	8.81	1.78	10.56	1.92	11.41	1.93	12.05	1.77	11.48
a_9	-0.42	-2.35	-0.55	-2.80	-0.59	-3.00	-0.32	-1.70	-0.24	-1.33
a_{10}	0.43	1.67	0.60	2.16	0.65	2.35	0.24	0.89	-0.02	-0.07
a_{11}	1.64	3.58	1.85	3.70	1.96	3.93	1.58	3.34	1.89	4.14
a_{12}	-0.12	-0.43	-0.20	-0.66	-0.27	-0.88	-0.11	-0.37	-0.26	-0.92
a_{13}	0.67	4.17	0.69	3.94	0.74	4.19	0.33	1.98	0.24	1.50
a_{14}	-1.94	-7.83	-2.08	-7.69	-2.14	-7.90	-1.77	-6.88	-1.92	-7.74
R^2	0.301		0.308		0.319		0.121		0.111	

B. Gasoline/Crude Futures Spread Options										
	CV		**G11**		**G11-C**		**GM**		**GM-C**	
a_1	3.70	5.13	2.83	3.48	2.04	2.33	2.92	3.85	2.10	2.15
a_2	-0.99	-13.85	-0.19	-2.32	-0.32	-3.70	-0.69	-9.16	-0.53	-5.50
a_3	2.50	20.70	1.55	11.39	1.87	12.73	1.37	10.80	1.18	7.26
a_4	-3.55	-2.45	-2.26	-1.39	-2.35	-1.34	-4.09	-2.69	-3.43	-1.75
a_5	1.41	1.95	0.74	0.90	0.81	0.91	1.78	2.33	1.48	1.50
a_6	-0.77	-3.63	-0.27	-1.11	-0.67	-2.58	-1.00	-4.45	-0.41	-1.43
a_7	1.69	5.12	1.14	3.05	1.50	3.72	1.93	5.55	0.88	1.96
a_8	-1.27	-10.62	-1.19	-8.83	-0.24	-1.61	-0.24	-1.93	0.16	0.99
a_9	0.98	10.62	0.36	3.49	0.57	5.04	0.40	4.16	0.31	2.47
a_{10}	-2.04	-14.12	-1.25	-7.67	-1.55	-8.79	-1.18	-7.80	-1.08	-5.54
a_{11}	-0.61	-2.01	-2.82	-8.21	-0.70	-1.88	-1.28	-3.99	-0.68	-1.65
a_{12}	1.00	5.94	2.09	11.07	1.00	4.90	1.19	6.75	0.88	3.86
a_{13}	-0.29	-3.99	-0.66	-8.23	-0.57	-6.52	-0.11	-1.51	-0.01	-0.08
a_{14}	-0.41	-2.49	0.99	5.32	-0.16	-0.81	0.10	0.55	-0.23	-1.00
R^2	0.363		0.322		0.338		0.135		0.070	