J Econ Finan (2013) 37:100–121
DOI 10.1007/s12197-011-9171-1

Crack spread option pricing with copulas

Hemantha S. B. Herath · Pranesh Kumar · Amin H. Amershi

Published online: 1 February 2011

Abstract A copula-based approach for pricing crack spread options is described. Crack spread options are currently priced assuming joint normal distributions of returns and linear dependence. Statistical evidence indicates that these assumptions are at odds with the empirical data. Furthermore, the unique features of energy commodities, such as mean reversion and seasonality, are ignored in standard models. We develop two copula-based crack spread option models using a simulation approach that address these gaps. Our results indicate that the Gumbel copula and standard models (binomial, and Kirk and Aron (1995)) mis-price a crack spread option and that the Clayton model is more appropriate. We contribute to the energy derivatives literature by illustrating the application of copula models to the pricing of a heating oil–crude oil "crack" spread option.

Keywords Copulas · Non-linear Dependence · Crack Spread Options · Energy Derivatives · Oil and Gas · Risk Management

JEL Classification C: Mathematical and Quantitative Methods

H. S. B. Herath (✉)
Department of Accounting, Faculty of Business, Brock University, Taro Hall 240,
500 Glenridge Avenue, St. Catharines, ON, Canada L2S 3A1
e-mail: hemantha.herath@brocku.ca

P. Kumar
Department of Mathematics, University of Northern British Columbia, 3333 University Way,
Prince George, BC, Canada V2N 4Z9
e-mail: kumarp@unbc.ca

A. H. Amershi
DeGroote School of Business, McMaster University, 120 Main Street, Hamilton, ON,
Canada L8S 4M4
e-mail: amershi@mcmaster.ca

1 Introduction

Spread options are options on the price difference between two or more traded assets. "Crack" is a term used in refining which refers to the "cracking" of crude oil into other products (for example heating oil or gasoline). A crack spread (also known as a refinery spread) is the simultaneous purchase or sale of crude against the purchase or sale of refined petroleum products. Crack spreads were introduced in October of 1994 on the New York Mercantile Exchange (NYMEX) and are used for managing risk in the oil industry. Crack spread options[1] are puts and calls on the one-to-one ratio between the New York Harbor heating oil futures or New York Harbor unleaded gasoline futures and the Exchange's light crude oil futures contract. Our article considers the application of copula models to the problem of pricing heating oil–crude oil crack spread options.

A heating oil–crude oil crack spread call option struck at a strike price K at maturity is given by $(42 \times [\mathrm{HO}] - [\mathrm{CO}] - K)^{+} = \max\{42 \times [\mathrm{HO}] - [\mathrm{CO}] - K, 0\}$. The underlying spread or refining margin and the option price is expressed in dollars and cents per barrel. Therefore, the (42 gallons = 1 barrel) conversion rate for HO is used as heating oil is priced according to \$ per gallon of product. Crack spread options are widely used as a risk management tool by both refiners and marketers of heating oil. Crack spread options allow refiners to hedge their operating margins at a known up-front cost while simultaneously allowing them to participate in any future widening of refining margins. Puts offer refiners an instrument for locking in crude oil costs. In addition, crack spread options allow refiners to generate income by writing options. Calls, on the other hand, offer marketers protection during unstable spread increases.

Crack spread options are traded between 9:00 am to 2:30 pm through open outcry trading at NYMEX. That is, oral "bids" (i.e., requests to buy) and "offers" (i.e., desire to sell) are made in the pits as arms-length transactions between buyers and sellers. In trades, the "High" and "Low" refers to the highest bid price paid and the lowest offer accepted for the crack spread option during the day or life of the contract. The settlement prices, on the other hand, are the prices given by the exchange at the end of the open outcry period. In regard to crack spread options, the settlement prices are derived from available market information including, but not limited to, outright trades, bids, or offers during the close, relevant spread trades, bids, or offers during the close, the settlement price of the underlying future, and the relevant relationships based on option pricing theory using option pricing models employed by the exchange (CBOT Handbook, Rule 813). They are the synchronous prices of the derivative and its underlying asset, and are determined by a settlement committee with input from floor traders.[2] The settlement price, also referred to as the "actual price" (see Laurence and Wang 2009), best reflects the true market valuation at the time of close.

[1] Throughout the article we use the generic term "crack spread option" although the model specifically refers to a heating oil–crude oil crack spread option (NYMEX trading symbol CH). The copula approach is similar for pricing a gasoline oil–crude oil "crack" spread option (NYMEX trading symbol CG).

[2] We would like to acknowledge Bob Biolsi of the CME Group for providing us with insight into how the settlement prices are determined.

Models that have been proposed for pricing spread options have mainly focused on European style-options. Commonly referred to as standard models are the bivariate binomial model (lattice approach) and the approximate analytical formula (Kirk and Aron (1995) or "Kirk" model). These crack spread option models suffer from three major limitations: (1) the assumption of a normal distribution for underlying asset returns; (2) the use of correlation to model dependence, and (3) the non-consideration of the unique features of crude and heating oil prices, such as seasonality and mean reversion. Dependence between prices is critical in many aspects of multi-asset derivative pricing and modeling price behavior in energy markets. Copula models that allow capturing dependencies have become popular in finance over the last decade. Copulas provide a flexible method for capturing the critical features of financial data such as asymmetry, non-linear dependence, and/or heavy tail behavior (Grégoire et al. 2008). Our research is motivated by the need to address these weak assumptions and an attempt to model the unique features of commodity prices that may affect the valuation of crack spread options.

Zła dynamika modeli

The idea behind the concept of a copula is to split a multivariate distribution function into two parts that describe the dependence structure and marginal behavior. More formally, copula models allow the unknown joint distribution of two or more random variables to be modeled as a combination of their marginal distributions and a copula function to capture the dependence, i.e., let X and Y be continuous random variables with distribution functions $F(x) = \Pr(X \leq x)$ and $G(y) = \Pr(Y \leq y)$, respectively. Following Sklar (1959), there exists a unique function C, such that: $\Pr(X \leq x, Y \leq y) = C(F(x), G(y))$ where the copula function $C(u, v) = \Pr(U \leq u, V \leq v)$ is the distribution of the pair $(U, V) = C(F(X), G(Y))$ whose margins are uniform on [0, 1] and that characterizes the dependence in the pair (X, Y). When the joint distribution of (X, Y) is unknown, it can be modeled by assuming specific parametric forms for F, G, and C (Grégoire et al. 2008).

The objective of this article is to examine the choices of copula C appropriate for pricing crack spread options, determine which copula function provides the best model in terms of goodness of fit, and to compare the copula crack spread option models with the standard models for pricing these options. Several articles have proposed copulas to study the relationship between prices for futures on crude oil and natural gas (Grégoire et al. 2008) and the evolution of electricity and natural gas prices for pricing "spark" spread options (Benth and Kettler 2006). We contribute to the literature by using copulas to study the dependence relationship between futures on crude oil and heating oil, and pricing a heating oil–crude oil crack spread option. More specifically, this article develops a copula-based Monte Carlo simulation model to price crack spread options and compare the copula models with settlement prices and the currently used standard models. To the best of our knowledge, there has been as yet no application of copula methodology to pricing of crack spread options. Our approach does not require stringent normality assumptions (any marginal distribution can be used) and would be of interest to both academia and practice.

The paper is organized as follows. Section 2 provides the motivation and a literature review. We demonstrate empirical evidence that is at odds with the assumptions of normal distribution of returns and linear dependence. The literature review discusses limitations of correlation as a dependence measure and articles that have examined copula models for pricing energy derivatives. In Section 3 we present a copula-based Monte Carlo simulation approach for pricing a crack spread option. Numerical results are presented in Section 4. In Section 5, we discuss the managerial use of the proposed models. Section 6 concludes with a discussion on model limitations. Appendix A provides a brief survey of the copula families used in this article, goodness of fit tests for selecting a copula and simulation algorithms. Appendix B summarizes the standard models.

2 Motivation and literature review

Pricing and hedging spread options can be quite challenging since closed-form solutions are rare. Often, analytical approximations, numerical methods, and Monte Carlo simulation are used to price spread options. Carmona and Durrleman (2003) provide a detailed survey of spread options and the theoretical and computational problems associated with pricing and hedging them. In Appendix B, we provide a description of the two commonly used European-type spread option models, the bivariate binomial model and the Kirk and Aron (1995) model, that we compare with the copula models. These standard models suffer from the assumption of symmetric returns, linear dependency, and the non-consideration of unique features such as mean reversion and seasonality in the stochastic behavior of underlying asset prices.

In order to motivate the use of copulas in pricing crack spread options, we use the daily NYMEX commodity market data (519 parallel observations) over the period February 3, 2003, to March 3, 2005. We use the Cushing Oklahoma Crude Oil Futures Contract 1 (dollars per barrel) and the New York Harbor No. 2 Heating Oil Future Contract 1 (cents per gallon) for crude oil (S_1) and heating oil (S_2) respectively. We selected this specific data range because the dependence between prices of heating oil and crude oil measured by the Kendall's Tau over successive periods of three months is found to be roughly constant with time. Grégoire et al. (2008) indicate that this constant feature is an implicit assumption for using copulas.

Wypada problem dynamiki kopuły

Our analysis based on the NYMEX data shows that the above assumptions are at odds with the empirical data. First, statistical evidence based on log-returns for crude and heating oil prices for the data shown in Table 1 indicates that the return distributions exhibit tail behavior. The observed distribution for heating oil is larger in the tails (heavy tailed) than a normal distribution. We confirm this with the descriptive statistics in Table 1 as the kurtosis measure is 4. 669 (>3, the kurtosis value for a normal distribution). The observed distribution for crude oil as seen in Table 1 (kurtosis measure 1.630<3) is less peaked than a normal distribution. We further confirm the above statistical evidence with the return distributions and quantile-quantile plots shown in Fig. 1. Independent from our investigation analyzing commodity futures data, Ross (1999) also provides evidence that log returns of crude oil prices may be non-normal.

Second, the correlation measure used to capture the dependence among the underlying future prices in the standard models has several limitations (Rachev et al. 2005). First, correlation is the appropriate dependence measure only if the marginal

Table 1 Descriptive statistics

	Crude oil	Heating oil
Frequency	519	519
Mean	.000948	.000934
Standard deviation	.02370	.0278
Skewness	−.503	−.668
Standard error of skewness	.107	.107
Kurtosis	1.630	4.669
Standard error of kurtosis	.214	.214
Minimum	−.11540	−.19249
Maximum	.06300	.08671
Range	0.1784	0.2792

Descriptive statistics of log-returns of crude and heating oil prices, February 3, 2003, to March 3, 2005 (Data size=519 parallel observations)

and joint distributions of the random variables are multivariate normal.[3] Second, the widely used Pearson product moment correlation ignores any non-linear dependence. Third, correlation is not invariant under a strictly increasing monotonic transformation of the random variables. From our data analysis, it is clear that the marginal distributions for crude oil and heating oil returns do not fit a normal distribution and thus, correlation (linear dependence) as a dependence measure is not appropriate (Rachev et al. 2005; Alexander 2005). Grégoire et al. (2008) argue that copula models are well suited to accurately modeling the non-linear dependencies in the above instances.

Nieliniowość

In order to capture asymmetry and non-linear dependence, researchers have advocated the use of copulas in an energy price series and derivative modeling. Accurately modeling price behavior in energy markets is crucial because the commodities have become increasingly intertwined (i.e., both natural gas and crude oil are used to generate electricity, and heating oil and natural gas are also used to extract oil from tar sands) (Grégoire et al. 2008). Grégoire et al. (2008) propose the use of copulas since crude oil and natural gas log-returns are found to exhibit non-linear dependence. Various families of copulas are fitted and the best copula is selected to predict the joint distribution for the prices of crude oil and natural gas. Similar to Grégoire et al. (2008), we also investigate the best fit of the Clayton (1978), Frank (1979), and Gumbel (1960) families of copulas and use Monte Carlo simulation for forecasting with a copula.

Benth and Kettler (2006) propose a non-symmetric copula to model the evolution of electricity and natural gas prices, and to investigate the pricing effects on "spark" spread options. They find that joint electricity and gas price discovery is better modeled by normal inverse Gaussian marginal processes using a specialized copula to couple them than by a binormal model. Benth and Kettler (2006) develop a

Relacje nie-multi-normalne

[3] More generally, the marginal distributions and joint distribution of the random variables must be elliptical distributions (such as multivariate normal, multivariate t-distribution, logistic distribution, and Laplace distributions).

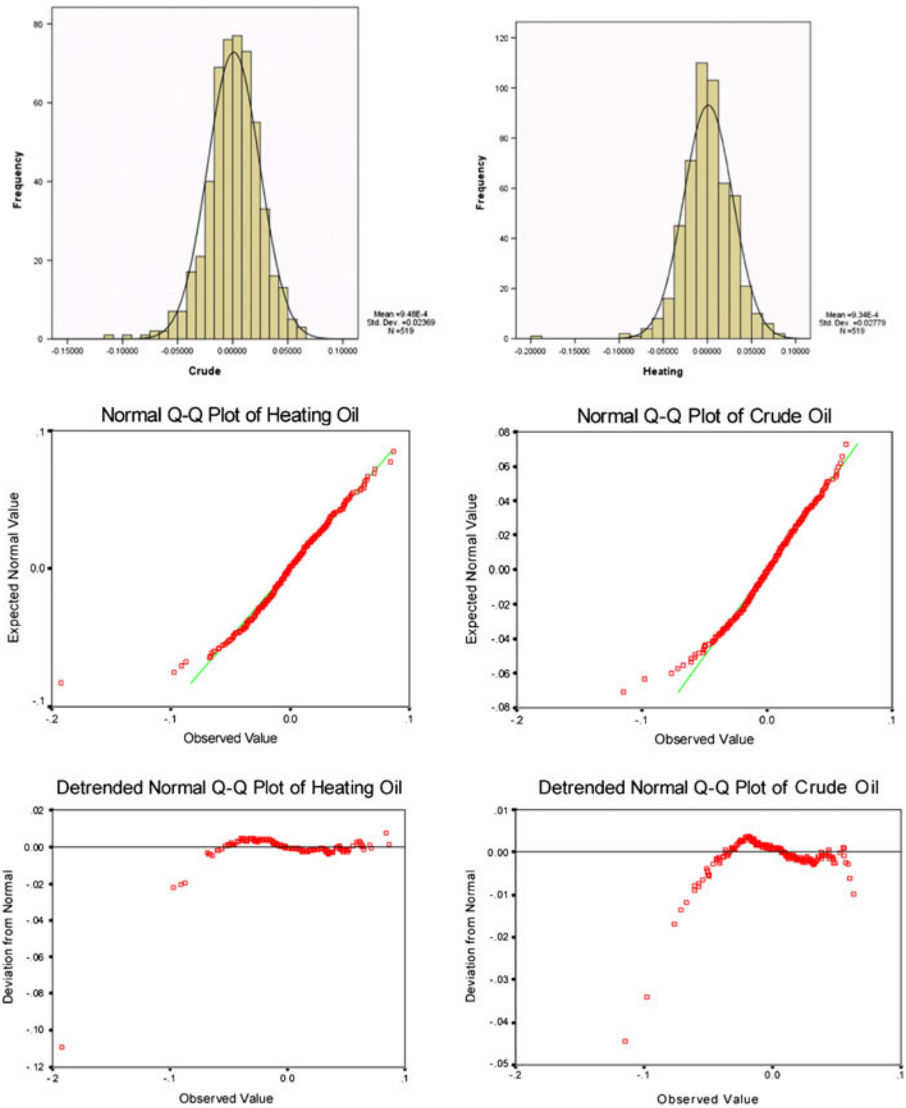


Fig. 1 Log-return distributions and normal Q-Q plots

discrete-time model that allows the computation of the theoretical copula which links electricity and gas prices as a function of a copula difference function and the independent copula. The price paths are simulated by taking random draws from the theoretical copula. Biglova et al. (2008) use a *t*-copula to capture the dependence. They find that in modeling price returns for energy futures such as crude oil, natural gas, and heating oil, a Stable distribution reflects the well documented characteristics of energy prices (for example, mean reversion, seasonality, etc.) much better than a normal distribution. Although, the above articles use copulas to study the price dependences in energy markets, our article is distinct from these as we model the interdependence between crude oil and heating oil to price a heating oil–crude oil crack spread option which, to the best of our knowledge, has not been done before. Thus, our article contributes to the growing body of copula literature in energy markets.

In a recent article, Laurence and Wang (2009) derive closed form distribution free lower bounds and optimal sub-replicating strategies for spread options in a one-period static arbitrage setting. They provide lower bounds for spread options on two assets subject to the constraints that forward prices of each asset are known and call prices corresponding to one strike per asset are also known. The authors provide a partial solution in the two strikes per asset case by reducing it to a linear programming problem. They test the models using NYMEX heating oil–crude oil crack spread option data from October 2006 to December 2006. Their findings indicate that the one strike per asset discrete lower bound can be far from as well as very close to the traded price. In our article, we use the optimal lower bounds for the one strike case developed in Laurence and Wang (2009), to evaluate the copula models and the standard models for pricing crack spread options.[4]

Although not specific to energy markets, several articles have used copulas to model dependence relationships between underlying assets for valuing financial options, including spread options. van den Goorbergh et al. (2005) propose dynamic copula models to price bivariate options such as outperform and underperform options. They use the daily returns on the S&P 500 and Nasdaq but do not empirically test the valuation models as price data is not available. Cherubini and Luciano (2002) use copulas as pricing devices to price bivariate options such as binary digital options, options on the minimum of two assets, and exchange options. Copulas are suggested as a strategy to address joint issues of non-normality of returns and dependence in pricing bivariate contingent claims.

The standard models for pricing crack spread options assume that the underlying prices for crude oil and heating oil follow a geometric Brownian motion (GBM). A GBM process for crude oil is also assumed in Paddock et al. (1988). Furthermore, Dias (2004) and Paddock et al. (1988) indicate that there are instances where the GBM is a good approximation for modeling the short run evolution of crude oil prices. Specialists, however, have argued that a realistic model for stochastic behavior of oil prices should include mean-reversion features (Dias 2004; Pindyck 1999; Schwartz 1997). Several researchers have also modeled other unique features of energy commodity prices such as strong seasonality[5] in heating oil and gasoline, time to maturity effect, etc. (Girma and Paulson 1998, 1999; Haigh and Holt 2002; Dunis et al. 2006). More recently, Biglova et al. (2008) have argued that characteristics such as mean reversion should be incorporated in modeling energy prices. As there is mixed evidence of mean reversion for crude oil (Biglova et al. 2008; Dias 2004; Paddock et al. 1988), for generality, we incorporate this feature in the proposed copula model.[6]

GBM nie ok

---

[4] We wish to sincerely thank the anonymous referees for suggesting the Laurence and Wang (2009) article which has helped improve an earlier version of this article considerably.

[5] In the daily NYMEX crude oil and heating oil futures price data (519 parallel observations) over the period February 3, 2003, to March 3, 2005, we find no evidence of seasonality for both price series. The chi-square test statistic values are 6.84 (p-value=0.81) for crude oil and 14.52 (p-value=0.21) for heating oil respectively.

[6] Dornier and Queruel (2000) include an additional term in order to consider seasonality in a mean-reverting model.

3 Copula-based Monte Carlo simulation model for pricing a crack spread option

"Copulas" are functions that join or couple multivariate distributions to their one-dimensional marginal distribution functions. Nelsen (1999) provides a good foundation of copula techniques. The copula-based Monte Carlo Simulation approach to price a crack spread option is quite versatile. It allows the modeling of unique features such as mean reversion, which are associated with energy commodities, and the capturing of dependence using copula families. The approach is as follows: First, we construct the price series for crude oil and heating oil using a copula-based Monte Carlo simulation approach. The copula functions capture the dependencies. We consider three Archimedean copula families discussed in Appendix A and use the Genest and Rivest (1993) fit test procedure (summarized in Appendix A.2) to identify the best-fit copula. Next, we use the simulated asset prices to compute the crack spread option price. The mean value of the option price computed over a large number of simulations provides an estimate of the crack spread option premium.

In order to demonstrate that the proposed copula-based approach is appropriate for capturing the unique features of commodity prices, we develop two models: (A) a Geometric Brownian motion crack spread option model (GBM-CSOM) and (B) a mean reversion crack spread option model (MR-CSOM)

(A) **GBM-CSOM:** The two commodities, crude oil and heating oil, are traded assets, and thus, both $S_1(t)$ and $S_2(t)$ in this basic model are assumed to follow a geometric Brownian motion (GBM). In the simulation approach, we discretize the option's life into n steps of length $\Delta t = T/n$ to construct the price series[7] for crude oil and heating oil. The dependence is modeled by an Archimedean copula. The primary modeling feature of the copula-based simulation is that the random variate X_i is unique to the copula. To generate the random paths, we use the following recursion for each asset $i\epsilon[1,2]$:

$(X_1;X_2)$ będą zależne poprzez kopułę

$$S_i(t) = S_i(t-1)\exp\left\{\left(r - 0.5\sigma_i^2\right)\Delta t + \sigma_i X_i \Delta t\right\} \quad (1)$$

(B) **MR-CSOM:** In this model, we assume that the two commodities, crude oil ($S_1(t)$) and heating oil ($S_2(t)$) follow an arithmetic Ornstein-Uhlenbeck (pure mean reversion) process where the prices revert to a long run equilibrium level. There are several types of mean reversion models that can be used but we adopt the risk neutral version of the pure mean reversion model (see Dias (2004) for a nice exposition). The Monte Carlo simulation of a stationary first order autoregressive process provides the correct discretization of a continuous mean-reversion stochastic process (Dixit and Pindyck 1994). Define a AR(1)

[7] While we only need the payoff at maturity to price the European-type crack spread option discussed in this paper, we generate the full random paths for the assets as an illustration since the copula simulation technique can be used for pricing other path dependent options (e.g., American type) in the energy sector.

stochastic variable $S_i(t) = \ln S_i(t)$ which is reverting towards an equilibrium level (mean) $\bar{s}_i = \ln \bar{S}_i$, where η_i, σ_i^* and μ_i are positive parameters measuring the speed of reversion, the volatility of the process, and the risk-adjusted discount rate for the price of commodity S_i. The risk-neutral mean reversion process that allows for valuing derivatives can be simulated directly using the following sample path simulation equation:

tu to samo

$$S_i(t) = \exp\left\{ \left[\ln[s_i(t-1)]e^{-\eta_i \Delta t}\right] + \left[\ln(\bar{s}_i) - \left[\frac{\mu_i - r}{\eta_i}\right]\left(1 - e^{-\eta_i \Delta t}\right)\right] - \left[\left(1 - e^{-2\eta_i t}\right)\frac{\sigma_i^{*2}}{4\eta_i}\right] + \sigma_i^* \sqrt{\frac{1 - e^{-2\eta_i \Delta t}}{2\eta_i}} \underline{\underline{X_i}} \right\} \quad (2)$$

The parameter estimation for the simulation Eq. 2 can be done easily using the following estimation procedure for an arithmetic Ornstein-Uhlenbeck process. For each commodity we run the regression equation $s_i(t) - s_i(t-1) = a + bs_i(t-1) + \varepsilon(t)$. From the regression results we can estimate: long run equilibrium mean $\widehat{\bar{s}}_i = -a/b$, speed of reversion $\eta_i = -\ln(1+b)$, and volatility $\sigma_i^* = \sigma_{i\varepsilon}\sqrt{\frac{2\eta_i}{(1-e^{-2\eta_i})}}$ where $\sigma_{i\varepsilon}$ is the standard error of the regression. It is important to notice that in the above estimation equations, the values of a, b should be statistically significant and that $b<0$. If $b>0$ and/or is not statistically significant, then there is no mean reversion observed in the process.

In order to price the crack spread option, we simulate a large number of bivariate data $\left(S_1^k(t), S_2^k(t)\right)$ using either the GBM-CSOM or the MR-CSOM model by repeating the algorithms (Appendix A.3.1 and A.3.2) M number of times. Then the price of a crack spread option can be estimated as:

$$\widehat{C} = \underbrace{e^{-\int_1^T r dt}}_{e^{-rT}} \cdot \underbrace{\frac{1}{M}\sum_{k=1}^{M} \max\left[S_1^k - S_2^k - K, 0\right]}_{\mathbb{E}[\text{Payoff}]} \quad (3)$$

4 Numerical results

The copula pricing of a crack spread option is illustrated using the NYMEX daily futures prices of crude oil (S_1) and heating oil (S_2) over the period February 3, 2003, to March 3, 2005 (519 parallel observations). In the copula approach, the Kendall's Tau, a nonparametric measure of dependence, is often used to measure the dependence between the two assets (crude oil and heating oil) and derive the copula dependence measure (θ). Using the statistical package SPSS, we compute the Kendall's Tau for the above data as 0.687. Next, we measure the degree of dependence by the copula parameter (θ) for each copula function—Frank, Clayton, and Gumbel—using the equation (d) in Columns A, B, and C of Appendix Table 4. The estimated values of parameter (θ) for the Frank, Clayton, and Gumbel copulas are 0, 4.3898, and 3.19219, respectively.

In order to identify the appropriate copula, we follow the fit-test procedure outlined in Appendix A.2. The empirical distribution $K_E(z)$ and the $K(z)$ values for the Frank, Clayton, and Gumbel copulas based on the Eqs. A1, A2, and A3 in Appendix A.2 are shown in Fig. 2. Based on a visual fit, it is evident from Fig. 2 that Clayton and Gumbel provide the best fit (both are relatively close). However, minimizing the mean square error would be a more robust method to select the appropriate copula. Model selection

We first determine which of the two models, GBM-CSOM or MR-CSOM, would be applicable for our data. We run the regression equation $s_i(t) - s_i(t-1) = a + bs_i(t-1) + \varepsilon(t)$ on the 519 parallel observations of the daily futures price for crude oil and heating oil. The values of a, b and the P-values (in brackets) for crude oil are $a = 0.01546917(P = 0.424 > 0.05)$ and $b = -0.00403481(P = 0.453 > 0.05)$, which indicates that there is no mean reversion observed since the parameters are not statistically significant at 95% confidence level. Similarly, for heating oil we obtain $a = 0.01830239(P = 0.376 > 0.05)$ and $b = -0.00466399(P = 0.400 > 0.05)$, which again does not support the hypothesis regarding mean reversion. One of the factors affecting this lack of support may be the short time span of the data collection period.

≈ 0

$r_t = a + b r_{t-1}$

Based on the above analysis, we choose the GBM-CSOM as the applicable copula-based heating oil–crude oil crack spread option model. In this section, we first perform an out-of-the-sample analysis using a sample of heating oil–crude oil crack spread option trade prices[8] obtained directly from the NYMEX during the period February 3, 2003, to March 3, 2005, and the computed Laurence and Wang (2009) lower bounds for spread options. Using the lower bounds, we are able to partially narrow down the choices of pricing models from the set of Clayton, Gumbel, bivariate binomial, and Kirk and Aron (1995) models. A more robust approach is to compare the models with the settlement prices since these synchronous prices best reflect the true market valuation at time of close. Thus, we evaluate the proposed copula model and the standard models using the following two measures: the mean absolute pricing error and the median absolute pricing error. Second, in an in-sample analysis, we compare the Clayton copula-based option price with the Gumbel, bivariate binomial, and Kirk and Aron (1995) models for particular cases to assess the pricing bias. The copula-based crack spread option prices (Clayton and Gumbel) are computed from 5000 simulation runs.

Error metrics

Laurence and Wang (2009) derive the optimal lower bounds for the one strike per asset case for spread options that can be directly applied to partially evaluate the copula and standard models. To be thorough, we present the results from Section 2.2 of Laurence and Wang (2009). Let c_i be the call prices pertaining to asset x_i with a strike K_i and spot price m_i, $D = K_1 - K_2 - K$ and $F_i = \frac{m_i - c_i}{K_i}$ for (i=1,2). Then the optimal lower bounds are given as follows:

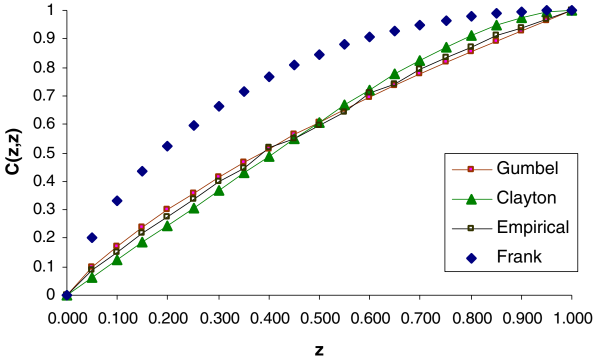
(i) if $D \geq 0$, $\max\{c_1 - c_2, m_1 - m_2 - K, m_1 - m_2 - K - (c_1 - c_2) - DF_2, 0\}$;

(ii) if $D \leq 0$, $K_1 < K$, $\max\{(c_1 - c_2)^+ + (K_1 - K)F_1 - K_2F_2,\ (c_1 - c_2)^+ + DF_1, 0\}$;

(iii) if $D \leq 0$, $K_1 < K$, $\max\{(c_1 - c_2)^+ + DF_1, (c_1 - c_2)^+ + DF_2, (c_1 - c_2)^+ + K_1F_1 - K_2F_2 - K,\ 0\}$. We use the Black and Scholes (1973) option pricing model to determine the values for c_1 and c_2 over a large range of asset prices. We choose nine strike prices [−10, −7.5, −5, −2.5, 0, 2.5, 5, 7.5, 10], four below and four above the at-the-money strike for each underlying asset x_i, to

[8] The options data include the following items: "symbol," "trading date," "contract month," "open interest," "Call (c)/Put (p)," "strike price," "High," "Low," "Last," "Settlement Price (Actual Price)," and "Total Volume."

Fig. 2 Comparison of copulas



compute an average lower bound for each crack spread option. Using the average lower bound is sufficiently robust since spread options data are difficult to obtain and it is also difficult to track each options contract pertaining to the Black-Scholes price (Laurence and Wang 2009).

The option pricing parameters estimated from the NYMEX commodity data over the period February 3, 2003, to March 3, 2005, are as follows. For crude oil and heating oil, the initial price and volatility parameters are S_1=53.57, σ_1=0.376, S_2= 62.61, and σ_2=0.441. The non-linear dependence parameters obtained from the data for the Clayton and Gumbel copulas are θ_C=4.3898, θ_G=3.19219, and the linear correlation is ρ=0.799. Throughout the analysis, we assumed a risk-free rate of r= 0.04, which was approximately the Treasury Bill rate in 2005. In Fig. 3, we compare the copula-based crack spread option prices (Clayton and Gumbel) and standard (binomial and Kirk and Aron 1995) models with the Laurence and Wang (2009) lower bounds and settlement prices for a series of 10-month options with exercise prices of 8, 8.5, and 9, and a series of 11-month options with exercise prices 9.5, 11, and 15.

Implausible results

The Laurence and Wang (2009) bounds provide a theoretically optimal lower bound for traded prices. Based on Fig. 3, we observe that for both the 10-month and 11-month crack spread option series, the settlement prices are consistently above the lower bounds, while the Kirk and Aron (1995) model prices are consistently below the lower bounds (except the 11-month option with K=15). The binomial model prices are found to be below the lower bounds for only the 10-month option series. For both the 10-month and 11-month option series, the copula models are found to be at or above the lower bounds. Thus, by using the Laurence and Wang (2009) lower bound for prices, we are able to provide partial evidence favoring the copula models. We observe that the settlement price range is smaller for the 10-month options compared to the 11-month crack spread options. We also observe that the Clayton model prices (10-month option with exercise prices 8.5 and 9) are within the settlement price range.

We evaluate the proposed copula and the standard models using the following two measures: mean absolute percentage pricing error ($\mathrm{MAPE} = 100\frac{1}{N}\sum_{k=1}^{N}\left(\frac{abs\left(P^{\mathrm{Model}} - P^{\mathrm{Settlement}}\right)}{P^{\mathrm{Settlement}}}\right)$) and the median absolute percentage pricing error

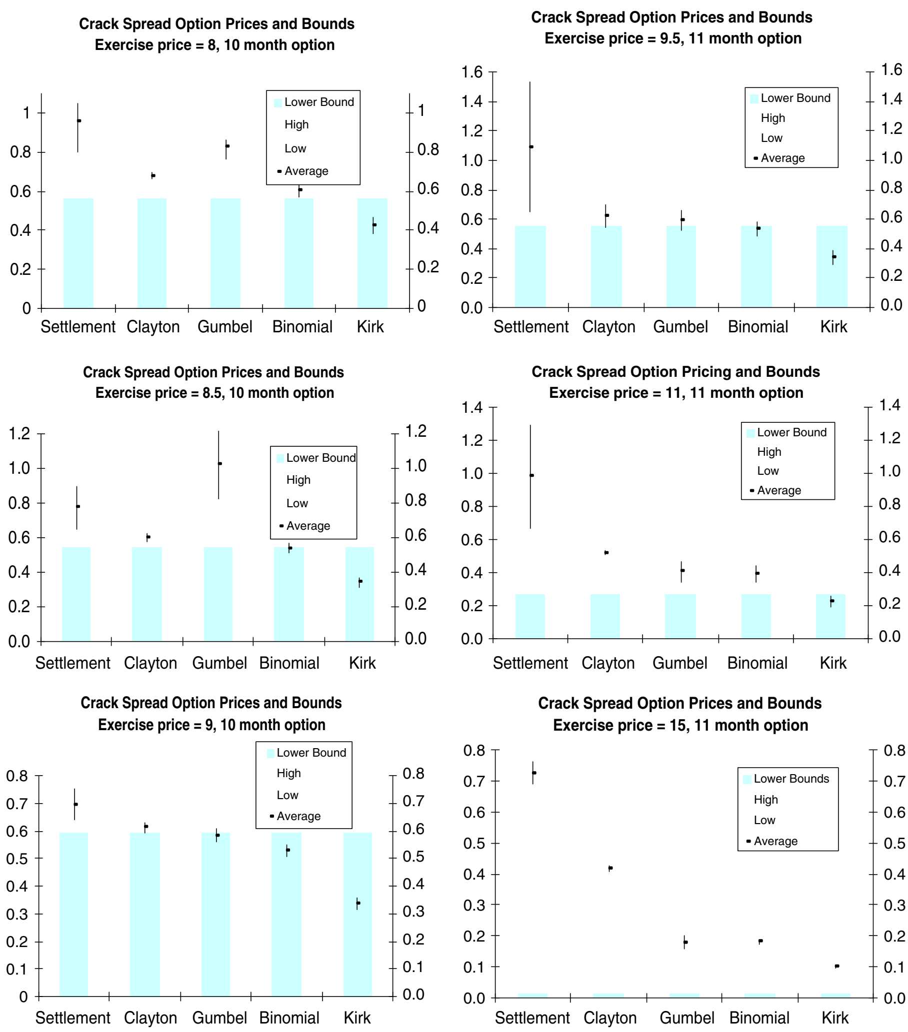


Fig. 3 Model comparison based on lower bounds and actual (settlement) prices in US$

$\left(\text{MdAPE} = 100\text{median}\left[\frac{abs\left(P^{\text{Model}} - P^{\text{Settlement}}\right)}{P^{\text{Settlement}}}\right]\right)$. Table 2 summarizes (as a percentage) the pricing errors pertaining to high, low, and average values for the heating oil–crude oil crack spread option obtained using the Clayton, Gumbel, binomial, and Kirk and Aron (1995) models. Our results show that the Kirk and Aron (1995) and binomial models are outperformed by the copula models. In fact, the mean (MAPE) and median (MdAPE) pricing errors are lower for the copula models. In our sample, the mean and the median absolute pricing errors for the Clayton model are 39.5% and 38.8% for high, 19.6% and 16.8% for low, and 38.6% and 36.2% for the average price, whereas for the Gumble model, they are 44.3% and 46.1% for high, 31.7% and 23.1% for low, and 46.1% and 38.7% for the average price.

Table 2 Pricing errors

	High	Low	Average
Mean Absolute Pricing Error (MAPE)			
Clayton (C)	39.513	19.568	38.557
Gumbel (G)	44.342	31.682	46.091
Binomial (B)	50.901	36.876	52.087
Kirk (K)	68.038	61.143	71.721
Median Absolute Pricing Error (MdAPE)			
Clayton (C)	38.793	16.802	36.226
Gumbel (G)	46.132	23.082	38.744
Binomial (B)	50.717	27.266	43.997
Kirk (K)	67.071	53.522	62.442

The mean absolute pricing error ($\text{MAPE} = 100\frac{1}{N}\sum_{k=1}^{N}\left(\frac{abs\left(p^{\text{Model}} - p^{\text{Settlement}}\right)}{p^{\text{Settlement}}}\right)$) and the median absolute pricing error ($\text{MdAPE} = 100\text{median}\left[\frac{abs\left(p^{\text{Model}} - p^{\text{Settlement}}\right)}{p^{\text{Settlement}}}\right]$) expressed as a percentage

Next, we compare the numerical results obtained from the Clayton (C) model with the Gumbel (G), bivariate binomial (B), and Kirk and Aron (1995) (K) models for a one year (T=1) crack spread option with an exercise price K=3. When S_1= 53.57 and S_2=62.61, the crack spread option values based on the bivariate binomial and the Kirk and Aron (1995) model are found to be \$1.86 and \$1.81. The Clayton and Gumbel model prices are \$1.99 and \$1.92. In Table 3, we compare the crack spread option prices obtained using the Clayton copula with the Gumbel, bivariate binomial, and Kirk and Aron (1995) models for particular cases.

In Fig. 4, we show the price relationship between the Clayton model and the Gumbel, bivariate binomial, and Kirk and Aron (1995) models. Our results indicate that in general, the Gumbel and bivariate binomial models tend to misprice the crack spread options when the asset spread is low. When the spread between the commodities is high, the Gumbel, bivariate binomial, and Kirk and Aron (1995) model prices are closer to the Clayton crack spread option prices. In Fig. 5, we show the effect of changing the exercise price while keeping the spread constant. Our findings indicate that the Gumbel and bivariate binomial crack spread prices are always higher (negative bias) than the Clayton model prices for in-the-money crack spread options. On the other hand, the Kirk and Aron (1995) model shows both a negative and positive bias in pricing.

5 Potential use of the model

Petroleum refiners and marketers face enormous risks when the price of crude oil rises while the price of heating oil remain static or even decline. Crack spread options provide added flexibility to both refiners and marketers of heating oil for managing risks. We provide a stylized example to illustrate the practical use of the copula-based crack spread

Table 3 Comparison of copula models with bivariate binomial and Kirk and Aron (1995) models

		Clayton (C)	C-G	Gumbel (G)	C-B	Binomial (B)	C-K	Kirk (K)
S1	S2	Price		Price		Price		Price
44	60	0.645	0.167	0.478	0.031	0.614	0.162	0.483
47	62	0.713	0.041	0.672	−0.065	0.778	0.033	0.680
43	56	0.644	0.011	0.633	−0.088	0.732	−0.008	0.652
46	59	0.845	0.045	0.800	0.000	0.845	0.043	0.803
47	60	0.810	−0.009	0.820	−0.077	0.887	−0.046	0.856
50	61	1.369	0.082	1.287	0.125	1.244	0.107	1.262
52	63	1.331	0.092	1.240	−0.041	1.372	−0.062	1.393
48	57	1.479	−0.108	1.587	0.061	1.418	0.070	1.409
52	61	1.651	0.005	1.646	−0.083	1.734	−0.047	1.698
51	59	1.789	−0.166	1.955	−0.081	1.870	−0.006	1.795
48	53	2.140	−0.230	2.370	−0.168	2.308	−0.007	2.147
51	56	2.296	−0.669	2.965	−0.287	2.584	−0.110	2.406
52	56	2.462	−0.875	3.337	−0.466	2.928	−0.276	2.738
53	56	3.127	−0.809	3.936	−0.146	3.273	0.031	3.096
52	52	3.895	−1.432	5.327	−0.047	3.941	−0.038	3.933

Parameters are volatility of crude oil, volatility of heating oil, time to maturity, exercise price, risk free rate, correlation, dependence parameter for the Clayton and Gumbel copulas. $\sigma_1 = 0.376$, $\sigma_2 = 0.441$, $T = 1\text{year}$, $K = 3$, $r = 0.04$, $\rho = 0.799$, $\theta_C = 4.3898$, $\theta_G = 3.19219$

option pricing model developed in this article. Consider the following example wherein the futures prices of crude oil and heating oil are given as S_1=42.12 and S_2=50.07. The value of the crack spread can be computed as \$7.95 per barrel (bbl). For simplicity, assume that this crack spread is currently trading at \$8.00 per bbl. The corresponding heating oil–crude oil crack spread option price (or premium) computed using the Clayton and Kirk and Aron (1995) models with an exercise price of \$9.00 per bbl. are found to be \$0.68 and \$0.36, respectively.

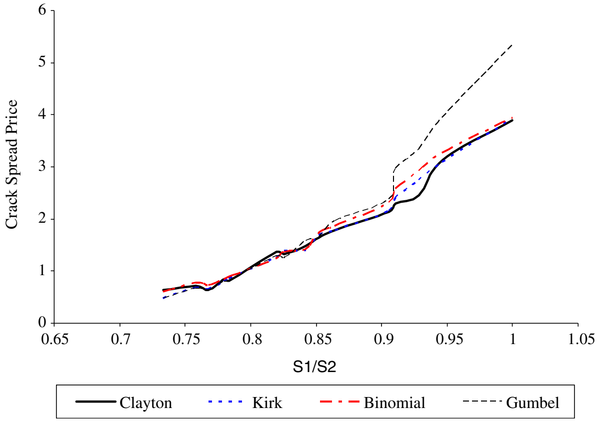


Fig. 4 Pricing bias. Parameters are volatility of crude oil, volatility of heating oil, time to maturity, exercise price, risk free rate, correlation, dependence parameter for Clayton and Gumbel copula. $\sigma_1 = 0.376$, $\sigma_2 = 0.441$, $T = 1\text{year}$, $K = 3$, $r = 0.04$, $\rho = 0.799$, $\theta_C = 4.3898$, $\theta_G = 3.19219$

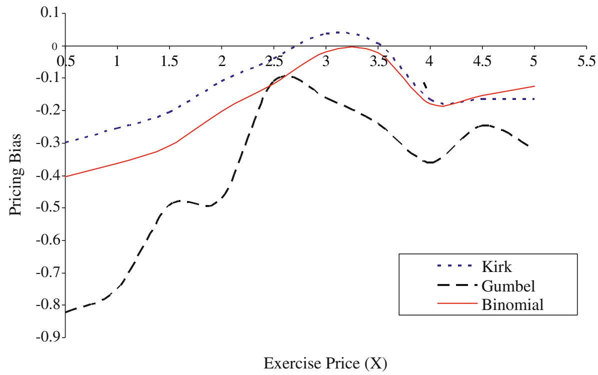


Fig. 5 Effect of varying the exercise price K. Parameters are futures price of crude oil at t=0, futures price of heating oil t=0, volatility of crude oil, volatility of heating oil, time to maturity, risk free rate, correlation, dependence parameter for the Clayton and Gumbel copula. $S_1 = 53.57,\ S_2 = 62.61,\ \sigma_1 = 0.376,\ \sigma_2 = 0.441,\ T = 1\text{year},\ r = 0.04, \rho = 0.799, \theta_C = 4.3898,$ $\theta_G = 3.19219$

Refiners are natural sellers (writers) of crack spread call options as they continuously buy crude oil and sell heating oil. In the above example, selling (or writing) a call option with an exercise price of \$9.00 allows the refiner to participate in favorable increase in the spread from \$8.00 per bbl. up to \$9.00 per bbl. If the crack spread widens to more than \$9.00 per bbl, the refiner will give up any profits above \$9.00 for the portion of its market exposure but it ensures the refiner has a continuous supply of crude oil. If the refinery plans to process an additional 50,000 barrels of crude oil, the refiner should sell (write) 50 crack spread call options as each call option contract represents 1,000 barrels of crude oil and 42,000 gallons (or 1,000 barrels) of heating oil. The buyer of the call (marketer) purchases the right to sell the crude oil and buy heating oil. If the marketer (buyer) exercises the call, the refiner is obligated to "sell the spread," that is, buy crude oil and sell heating oil.

If the spread widens to more than \$9.00 per bbl. (say \$10.50 per bbl.), then the call option sold by the refiner will be exercised by the marketer. The refiner will give up profits of \$ (11.50–9.00)=\$2.50 per bbl., totaling \$2.50×50×1,000=\$125,000. Selling the crack spread call option, however, enhances the margin since the premium income can be used to offset this loss. Suppose the Kirk and Aron (1995) model was used to price the crack spread option; then the premium income would be only \$0.36×50× 1,000=\$18,000. However, if the Clayton copula model was used, the premium income would be much higher: \$0.68×50×1,000=\$34,000. In a similar vein, refiners can use a "collar" strategy (sell a call and buy a put) to reduce the cost of buying a put option to protect the spread from falling below a specific amount. The cost of the hedge can be lowered by using the proceeds from the sale (writing) of the crack spread call to reduce the cost of purchasing an opposite options (put) position.

6 Conclusion

Recently there has been a growing body of articles examining price dependencies in energy markets that advocate the use of copula models for pricing energy derivatives. We contribute to this literature by illustrating the application of copula models for pricing a heating oil–crude oil crack spread option. In general, our results indicate that copula models are more appropriate to price crack spread options than the standard binomial and Kirk and Aron (1995) models. The copula models are relatively easy to implement and the copula-based simulation approach can readily incorporate mean reversion and seasonality in instances where such features are observed in the data. To the best of our knowledge, this is the first article in the literature to value a heating oil–crude oil crack spread option using copula functions. The application of copula functions, which do not require the normality assumption to value energy derivatives, will be of interest to both academia and practitioners.

Limitations

There are, however, several limitations to the copula approach. Although the implementation of a copula model is moderately easy, often closed form formulas are preferred by practitioners because they are easy and quick to implement. The closed form solution will always give the same answer, whereas the copula approach, which is based on Monte Carlo simulation, contains some randomness (Carmona and Durrleman 2003). Therefore, in practice, one would have to weigh the costs versus the benefits of using analytical approximations or more accurate copula models. From a theoretical viewpoint, however, as shown in this article, the copula approach is more methodologically robust in capturing non-linear dependencies and the unique features of energy commodities. Other limitations in the current approach that we plan to address in future research include estimating the marginal distributions from the price data and then inverting the option price for the copula parameter. The calibrated model can be used to price other options on the same asset. This would allow one to see the fit of different copula families for different options on the same asset. Alternatively, the appropriate copula and marginal distributions could be exploited to develop a copula-based analytical model to price a crack spread option that would be preferred by practitioners.

Acknowledgements We also wish to acknowledge valuable comments from participants at the Institute of Industrial Engineers Research Conference, Nashville, Tennessee, May, 2007 and Institute of Industrial Engineers Research Conference, Orlando, Florida, May, 2006. Dr. Herath acknowledges research funding from the Social Sciences and Humanities Research Council (SSHRC) of Canada.

Appendix A

Survey of Frank, Clayton, and Gumbel copulas

Three one parameter bivariate Archimedean copulas are given in Table 4. Nelsen (1999: 94–97) tabulates 22 one-parameter Archimedean copula families. The parameter θ in each case measures the degree of dependence and controls the dependence between the two variables. For instance, when $\theta \to 0$ there is no dependence and if $\theta \to \infty$ there is perfect dependence. The dependence parameter θ which characterizes each family of Archimedean copulas can be related to Kendall's Tau. This property is used to empirically determine the applicable copula form.

Calibration

Table 4 Frank, Clayton, and Gumbel Families

	Column A: Frank Copula (1979)	Column B: Clayton Copula (1978)	Column C: Gumbel Copula (1960)
(a) Generator	$\varphi_\theta(t) = -\ln\frac{e^{-\theta t}-1}{e^{-\theta}-1}$	$\varphi_\theta(t) = (t^{-\theta} - 1)$	$\varphi_\theta(t) = (-\ln(t))^\theta$
(b) Bivariate Copula Function	$C_\theta(u,v) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-ut}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right)$	$C_\theta(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$	$C_\theta(u,v) = \exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right\}$
(c) Laplace Transform:	$\varphi(t) = \varphi_\theta^{-1}(t) = \theta^{-1}\ln[1 + e^t(e^\theta - 1)]$	$\varphi(t) = \varphi_\theta^{-1}(t) = (1-t)^{-\frac{1}{\theta}}$	$\varphi(t) = \varphi_\theta^{-1}(t) = \exp\left(-t^{\frac{1}{\theta}}\right)$
(d) Kendall's Tau	$\tau_\theta = 1 - \frac{4}{\theta}[1 - D_1(\theta)]$ where $D_k(x)$ is the Debye function for any positive integer k, given by $D_k(x) = \int_0^x \frac{t^k}{e^t-1}dt$.	$\tau_\theta = \frac{\theta}{\theta+2}$	$\tau_\theta = \frac{\theta-1}{\theta}$

Empirically identifying a copula form based on distance $\| C_{param} - C_{emp} \|$

The first step in modeling and simulation is identifying the appropriate copula form. Genest and Rivest (1993) provide the following procedure (fit test) to identify an Archimedean copula. The method assumes that a random sample of bivariate data (X_i, Y_i) for i=1,2.....,n is available. Assume that the joint distribution function H has an associated Archimedean copula C_θ, and then the fit allows us to select the appropriate generator φ. The procedure involves verifying how closely different copulas fit the data by comparing the closeness of the copula (parametric version) with the empirical (non-parametric) version. The steps are:

Step 1 Estimate from the data the Kendall's correlation using the non-parametric or distribution free measure: !

$$\tau_E = \binom{n}{2}^{-1} \sum_{ij} Sign\left[\left(X_i - X_j\right)\left(Y_i - Y\right)_i\right]$$

Step 2 Identify an intermediate variable $Z_i = F\,(X_i, Y_i)$ with a distribution function $K(z) = \Pr(Z_i \le z)$. Construct an empirical (non parametric) estimate of this distribution as follows:

$$Z_i = \frac{\text{number}\left\{\left(X_i, Y_j\right)\text{such that } X_j < X_i \text{ and } Y_j < Y_i\right\}}{n-1}$$

The empirical version of the distribution function $K(z)$ is $K_E(z)$ = proportion of $Z_i \le z$

Step 3 Construct the parametric estimate of $K\,(z)$. The relationship between this distribution function and the generator is given by $K(z) = z - \frac{\varphi(z)}{\varphi'(z)}$, where $\varphi'(z)$ is the derivative of the generator and $0 \le z \le 1$. We provide below the specific form of $K(z)$ for the three Archimedean copulas surveyed in this paper.

minimize $\| K_p(\Theta) - K_e \|$

(i) Frank Copula
$$K(z) = \frac{\theta z - [1 - \exp(\theta z)] \ln\left[\frac{\exp(-\theta z) - 1}{\exp(-\theta)}\right]}{\theta} \tag{A1}$$

(ii) Clayton Copula
$$K(z) = \frac{z(1 + \theta - z^\theta)}{\theta} \tag{A2}$$

(iii) Gumbel Copula
$$K(z) = \frac{z(\theta - \ln z)}{\theta} \tag{A3}$$

Repeat Step 3 for several different families of copulas, i.e., several choices of generator functions, $\varphi(.)$. By visually examining the graph of $K(z)$ versus $K_E(z)$ or using statistical measures such as minimum square error analysis, one can choose the *best* copula. This copula can be used in modeling dependence and simulation.

Copula simulation procedures for the Clayton and Gumbel families

Use R/Python packages

In order to simulate outcomes from the bivariate distribution of asset prices, we use the procedures developed by Genest (1987) and Lee (1993) for the Clayton copula:

Algorithm A.3.1 Generating bivariate outcomes from the Clayton family

Step 1 Generate independent (0, 1) uniform random numbers u_1 and u_2
Step 2 Set $X_1 = F^{-1}(u_1)$
Step 3 Compute $u_2^* = \left[1 + u_1^{-\theta}\left(u_2^{-\theta/1+\theta} - 1\right)\right]^{-1/\theta}$
Step 4 $X_2 = G^{-1}\left(u_2^*\right)$

The above algorithm is computationally expensive for simulating bivariate data from the Gumbel distribution (see discussion in Frees and Valdez 1998). An alternate algorithm suggested by Marshall and Olkin (1988) for the construction of compound copulas can be used instead. The algorithm is as follows:

For the Gumbel copula, the generator and Laplace transform are given in Equation (c) Column C, Appendix Table 4. The inverse generator is equal to the Laplace transform of a positive stable variate $\gamma \sim \mathrm{St}\left(\widehat{\alpha}, 1, \Theta, 0\right)$ where $\Theta = \left(\cos\left(\frac{\prod}{2\theta}\right)\right)^{\theta}$ and $\theta>0$.

Algorithm A.3.2 Generating bivariate outcomes from the Gumbel family

Step 1 Simulate a positive Stable variate $\gamma \sim \mathrm{St}(\widehat{\alpha}, 1, \Theta, 0)$
Step 2 Simulate two independent uniform [0,1] random numbers u_1 and u_2
Step 3 Set $X_1 = F^{-1}\left(u_1^*\right)$ and $X_2 = G^{-1}\left(u_2^*\right)$ where $u_i^* = \varphi\left(\frac{1}{\gamma}\ln u_i\right)$ and $\varphi(t) = \exp\left(-t^{\frac{1}{\theta}}\right)$ for $i \in [1, 2]$

Cherubini et al. (2004) suggest the following procedure to simulate a positive random variable $\gamma \sim \mathrm{St}(\widehat{\alpha}, 1, \Theta, \delta)$:

Step 1 (a) Simulate a uniform random variable $\upsilon = U\left(\frac{-\prod}{2}, \frac{\prod}{2}\right)$

Step 1 (b) Independently draw an exponential random variable (ε) with mean=1

Step 1 (c) $\theta_0 = \arctan\left(\tan\left(\frac{\prod \widehat{\alpha}}{2}\right)\right)/\widehat{\alpha}$ and compute

$$z = \frac{\sin \widehat{\alpha}(\theta_0 + \upsilon)}{\left(\cos \widehat{\alpha}\theta_0 \cos \upsilon\right)^{\frac{1}{\widehat{\alpha}}}} \left[\frac{\cos\left(\widehat{\alpha}\theta_0 + \left(\widehat{\alpha} - 1\right)\upsilon\right)}{\varepsilon}\right]^{\frac{\left(1-\widehat{\alpha}\right)}{\widehat{\alpha}}}$$

Step 1 (d) $\gamma = \Theta z + \delta$

Appendix B

Current approaches to pricing crack spread options

Spread options are options whose payoffs depend on the difference between two or even three assets. The payoff of a two asset spread option at maturity for a call is

$$C(S_T) = \max(S_1 - S_2 - K, 0) \tag{B1}$$

where K is the strike price of the option, S_i is the price of the underlying asset i. When K=0, the spread option is equivalent to an exchange option (see Margrabe 1978). There is no closed-form solution for spread options with correlated asset prices and a non-zero strike price (see Boyle et al. (2004) and Carmona and Durrleman (2003) for discussions of pricing spread options with non-zero strike prices).

Kirk and Aron (1995) model

The Kirk and Aron (1995) crack spread option model is based on the Black (1976) futures contract model. It is a European-style spread option on futures contracts modified by equating the futures price to a general asset price S_i, with strike price K, volatility of the respective assets σ_i, the correlation (linear dependence) between the two assets ρ and the risk-free rate r. The Kirk and Aron (1995) formula to price the crack spread is as follows:

$$C = (S_2 + K)\left[e^{-rT}(S_A N(d_1) - N(d_2))\right] \tag{B2}$$

where

$$S_A = \frac{S_1}{S_2 + K}$$

$$\sigma = \sqrt{\sigma_1^2 + \left[\sigma_2 \frac{S_2}{S_1 + K}\right]^2 - 2\rho\sigma_1\sigma_2 \frac{S_2}{S_1 + K}}$$

$d_1 = \frac{\ln(S_A) + \left(r + \sigma^2/2\right)(T-t)}{\sigma\sqrt{T-t}}$ and $d_2 = d_1 - \sigma\sqrt{T-t}$

The Kirk and Aron (1995) formula is an *approximate* model since there is no closed-form solution to a spread option on correlated assets with a non-zero exercise price.

Binomial lattice approach

To price a crack spread option, we can use the binomial lattice approach for two correlated assets (see Kamrad and Ritchken 1991; Boyle 1988; Pickles and Smith 1993). Define the asset pair $\{S_1(t), S_2(t)\}$over time t as having a bivariate lognormal density. Let the drift rate be μ_i where σ_i is the instantaneous standard deviation of the

i^{th} asset and r is the risk-free rate. Define the instantaneous linear correlation between two assets as ρ. In the two assets binomial tree, each node has four branches and the risk-neutral probabilities for each branch is given by:

$$p_1 = \frac{1}{4}\left[1 + \sqrt{\Delta t}\left(\frac{\mu_1}{\sigma_1} + \frac{\mu_2}{\sigma_2}\right) + \rho\right] \tag{B3.a}$$

$$p_2 = \frac{1}{4}\left[1 + \sqrt{\Delta t}\left(\frac{\mu_1}{\sigma_1} - \frac{\mu_2}{\sigma_2}\right) - \rho\right] \tag{B3.b}$$

$$p_3 = \frac{1}{4}\left[1 - \sqrt{\Delta t}\left(\frac{\mu_1}{\sigma_1} + \frac{\mu_2}{\sigma_2}\right) + \rho\right] \tag{B3.c}$$

$$p_4 = \frac{1}{4}\left[1 - \sqrt{\Delta t}\left(\frac{\mu_1}{\sigma_1} - \frac{\mu_2}{\sigma_2}\right) - \rho\right] \tag{B3.d}$$

The value for the up and down movements for each asset i=1,2 is given by $\widehat{u}_i = e^{\sigma_i \Delta t}$ and $\widehat{d}_i = e^{-\sigma_i \Delta t}$.

Non-copula Monte Carlo simulation approach

Monte Carlo simulation can be used to price spread options including crack spread options. If the volatility of the assets and their correlation are assumed to be deterministic, then one could estimate the joint distribution of the underlying assets which can be used directly in the simulation. For example, if the marginal distribution of the log-returns of each asset is assumed to be normal, then in the simulation, one can sample directly from a joint normal distribution (or joint lognormal if returns instead of log-returns are used). In this case, payoff values can be simulated from the terminal distribution. For a European-type spread option, it is not necessary to simulate the entire price path. The non-copula Monte Carlo simulation approach, however, is also based on the same two weak assumptions (i.e., normal distribution of log-returns and linear dependence) as the Kirk and Aron (1995) model and the binomial approach.

References

Alexander C (2005) Correlation in crude oil and natural gas markets. In Managing energy price risk: the new challenges and solutions. Risk Books, London, pp 573–604

Benth FE, Kettler PC (2006) Dynamic copula models for the spark spread. Pure Math 14:1–31

Biglova A, Kanamura T, Rachev ST, Stoyanov SV (2008) Modeling, risk assessment and portfolio optimization of energy futures. Invest Manag Financ Innovat 5(1):17–31

Black F (1976) The pricing of commodity contracts. J Financ Econ 3:167–179

Black F, Scholes M (1973) The pricing of options and corporate liabilities. J Polit Econ 81(3):637–654

Boyle PP (1988) A lattice framework for pricing options with two state variables. J Financ Quant Anal 23:1–12

Boyle PP, Lai Y, Tan KS (2004) Pricing options using lattice rules. NA Actuar J 9(3):50–76

Carmona R, Durrleman V (2003) Pricing and hedging spread option. SIAM Rev 45(4):627–685

Cherubini U, Luciano E (2002) Bivariate option pricing with copulas. Appl Math Financ 9:69–85

Cherubini U, Luciano E, Vecchiato W (2004) Copula methods in finance. Wiley, West Sussex

Clayton DG (1978) A model for association in bivariate life tables and its applications in epidemiological studies of familial tendency in chronic disease incidence. Biometrika 65:141–151

Dias MAG (2004) Valuation of exploration and production assets: an overview of real options models. J Pet Sci Eng 44:93–114

Dixit AK, Pindyck RS (1994) Investment under uncertainty. Princeton University Press, Princeton

Dornier F, Queruel M (2000) Caution to the wind. Energ Power Risk Manag Aug 30–32

Dunis CL, Laws J, Evans B (2006) Modelling and trading the gasoline crack spread: a non-linear story. Deriv Use Trading Regul 12(1–2):126–145

Frank MJ (1979) On the simultaneous associativity of F(x, y) and x + y –F(x, y). Aequationes Math 19:194–226

Frees EW, Valdez E (1998) Understanding relationships using copulas. NA Actuar J 2(1):1–25

Genest C (1987) Frank's family of bivariate distributions. Biometrika 74:549–555

Genest C, Rivest L (1993) Statistical inference procedures for bivariate archimedean copulas. J Am Stat Assoc 88:1034–1043

Girma PB, Paulson AS (1998) Seasonality in petroleum futures spreads. J Futures Mark 18(5):581–598

Girma PB, Paulson AS (1999) Risk arbitrage opportunities in petroleum futures spreads. J Futures Mark 19(8):931–955

Grégoire V, Genest C, Gendron M (2008) Using copulas to model price dependence in energy markets. Energy Risk March 58–64

Gumbel EJ (1960) Distributions des valeurs extremes en plusiers dimensions. Publ Inst Stat Univ Paris 9:171–173

Haigh MS, Holt MT (2002) Crack spread hedging: accounting for time-varying volatility spillovers in the energy futures markets. J Appl Econ 17(3):269–289

Kamrad B, Ritchken P (1991) Multinomial approximating model for options with k-state variables. Manag Sci 37(12):1640–1652

Kirk E, Aron J (1995) Correlation in energy markets. In Managing energy price risk. Risk Books, London

Laurence P, Wang T (2009) Sharp distribution free lower bounds for spread options and the corresponding optimal subreplicating portfolios. Insur Math Econ 44:35–47

Lee AJ (1993) Generating random binary deviates having fixed marginal distributions and specified degrees of association. Am Stat 47:209–215

Margrabe W (1978) The value of an option to exchange one asset for another. J Financ 33:177–186

Marshall AW, Olkin I (1988) Families of multivariate distributions. J Am Stat Assoc 83:834–841

Nelsen RB (1999) An introduction to copulas. Springer, New York

Paddock JL, Siegel DR, Smith JL (1988) Option valuation of claims on real assets: the case of offshore petroleum leases. Quar J Econ 479–508

Pickles E, Smith JL (1993) Petroleum property valuation: a binomial lattice implementation of option pricing theory. Energ J 14(2):1–26

Pindyck RS (1999) The long-run evolution of energy prices. Energ J 20(2):1–27

Rachev ST, Menn C, Fabozzi FJ (2005) Fat-tailed and skewed asset return distributions-implications for risk management, portfolio selection and option pricing. Wiley, New York

Ross SM (1999) An introduction to mathematical finance. Cambridge University Press

Schwartz ES (1997) The stochastic behavior of commodity prices: implications for valuation and hedging. J Financ 52(3):923–973

Sklar A (1959) Functions de repartition a n dimensions et leurs merges. Publ Inst Stat Univ Paris 8:229–231

van den Goorbergh WWJ, Genest C, Weker BJM (2005) Bivariate option pricing using dynamic copula models. Insur Math Econ 37:101–114