Performance of statistical arbitrage in petroleum futures markets

Amir H. Alizadeh
Faculty of Finance, Cass Business School, 106 Bunhill Row, London EC1Y 8TZ, UK; email: a.alizadeh@city.ac.uk

Nikos K. Nomikos
Faculty of Finance, Cass Business School, 106 Bunhill Row, London EC1Y 8TZ, UK; email: n.nomikos@city.ac.uk

This paper investigates the intermarket and intercommodity linkages of petroleum and petroleum product futures markets and proposes trading strategies based on the combination of fundamental and technical analyses. These trading strategies use the cointegration between futures prices as fundamental relationships and implement technical trading rules to determine timing of long-short positions. The robustness of the trading strategies is also tested using the stationary bootstrap approach. Our results indicate that expected market prices in the relative form (spreads) incorporate inefficiencies, which can be translated to abnormal profits through appropriate trading strategies even when high levels of transaction costs (bid-ask spreads) are considered.

1 INTRODUCTION

Energy markets function with a unique structure of supply and demand mechanisms, which introduce a degree of complexity along with significant levels of uncertainty. After the two oil price crises in the 1970s, individual investors and energy market participants have always been faced with high levels of volatility, which in turn create the opportunity for large profits from speculation on oil prices. However, at the same time, this volatility can also lead to large losses if investment strategies are implemented at the wrong phase of the cycle. Within this setting, the formulation of sound trading strategies in the petroleum futures market is essential and can make the difference between success and failure of investment decisions.

The main body of literature on physical and derivatives oil markets concentrates on issues such as price discovery, market interrelationships and hedging effectiveness. For example, the issue of price discovery and efficiency has been investigated by Crowder and Hamed (1993), who find that West Texas Intermediate (WTI)

The author would like to thank Mr Panos Pouliasis for meticulous research assistance. The helpful comments of participants at the 2006 Energy Risk Europe Conference in London are also appreciated. The usual disclaimer applies.

futures are unbiased forecasts of the realized spot prices. Other studies investigate the causal relationship between oil spot and futures prices. For example, Silvapulle and Moosa (1999) report that oil spot and futures prices react simultaneously to the arrival of new information to the market. Haigh and Holt (2002) account for volatility spillovers between the crude, unleaded gasoline and heating oil markets, by using a multivariate error correction GARCH model to simultaneously link all three futures and spot markets and further estimate the minimum variance hedge ratio for the crack spread.[1] Making an allowance for transaction costs and carrying out in- and out-of-sample tests, they report substantial risk reduction compared with alternative hedging strategies.

A number of studies also investigate linkages between physical and futures crude oil markets in different geographical locations. For instance, Ewing and Harter (2000) provide evidence that Brent blend and Alaska North Slope crude oil prices move together over time and react similarly to shocks in the world oil market. Milonas and Henker (2001) investigate the relationship between Brent and WTI, modeling the two futures spread as a function of the convenience yields of the two contracts. They use convenience yields as surrogates for supply and demand conditions in the two markets and find that they can explain the variation in the intercrude spread. Alizadeh and Nomikos (2004) find that WTI futures and freight rates are cointegrated, whereas physical Brent and Nigerian Bonny, as well as the spread between futures and physical prices, are not related to freight rates. This is violating the cost of carry relationship, which in turn indicates the existence of arbitrage opportunities. The findings of these studies indicate that oil markets around the world are linked and prices move together over time.

A common feature of all those studies is that they either concentrate on the relationship between spot and futures prices in one market or investigate the linkages between spot and futures prices in different markets. Given the exacerbated volatility of oil prices, the global nature of energy commodities and the frequent supply chain disruptions and demand shocks, the oil market attracts a lot of interest from speculators. Consequently, the issue of whether profitable trading opportunities exist has been investigated extensively in the market.

Girma and Paulson (1999) investigate the profitability of trading opportunities for petroleum futures spreads traded on the New York Mercantile Exchange (NYMEX). They find that crack spread series are stationary and can be used to set up profitable moving average (MA) trading rules. They argue that the existence of strong long-term relationship among the spreads can justify the use of MA rules to identify departures from equilibrium; otherwise, if the spread is non-stationary, it will deviate without boundaries, and its use for either risk management or speculation will involve a high degree of uncertainty. However, although they report profits

[1] A crack spread is the simultaneous purchase (sale) of crude oil futures and sale (purchase) of petroleum product futures. The crack spreads are 3:2:1 crack (ie, three barrels of crude oil against two barrels of gasoline and one barrel of heating oil), 1:1:0 gasoline crack and 1:0:1 heating oil crack, and their magnitude reflects the cost of refining crude oil into petroleum products plus any profit/loss to refineries.

significantly different from zero, overall, the employed trading rules might not be as good as they promise and, as they state, "one cannot be certain that these opportunities still exist" if a different sample is used. Consequently, their analysis is merely a historical evaluation of risk arbitrage opportunities in petroleum futures spreads, because results from technical trading rules are prone to data snooping. The same drawbacks are also evident in another study of the crack spread relationship by Poitras and Teoh (2003). They explore day trading opportunities in the NYMEX market, using opening and closing (settlement) prices. Overall, they report net transaction fees profits, subject to some filter size that regulates the sensitivity of the trade signals to the actual decision of initiating a position. Other empirical studies that concentrate on commodities spread trading include Wahab *et al* (1994) for the gold-silver spread, Johnson *et al* (1991) for the crush spread in the soybean complex, Emery and Liu (2002) for spark spreads constructed from NYMEX contracts and Liu (2005) for spreads among hog, corn and soybean meal futures.

The study by Dunis *et al* (2006a) employs an alternative procedure by modeling and forecasting the spread. The data set used was comprised of daily closing prices of NYMEX WTI and Intercontinental Exchange (ICE) Brent. They use five trading models: fair value cointegration, MA, autoregressive moving average (ARMA), genaralized autoregressive conditional heteroskedasticity (GARCH) and the neural network regression.[2] The sample (1995–2004) is split into two periods, one for in-sample and one for out-sample testing. The best models proved to be the ARMA and the MA. However, the question whether the results would be qualitatively the same, using a different data set is not answered (data snooping).

This paper's objective is to investigate the relationship between the different pairs of petroleum and petroleum products futures spreads and utilize these linkages to establish trading strategies that make use of statistical arbitrage trading opportunities. Till now, literature has been focused on investigating linkages in the WTI – Brent futures or the NYMEX 3:2:1 crack spread, the unleaded gasoline and/or the heating oil crack spread. However, such strategies have been evaluated on the basis of their historical performance. Since technical trading strategies are prone to data snooping, one should confirm the robustness of such strategies, using also out-of-sample tests, before making inferences. In this study, in order to discount the possibility of data snooping bias, we use bootstrap simulations. In addition, another important question that has not been investigated in the oil spreads market literature is whether the profits produced by different strategies are significantly greater than other benchmark strategies. The profitability and risk-return characteristics of the employed trading strategies are compared with a simple

[2] The study by Dunis *et al* (2006b) focuses on artificial neural networks. In this study, three neural network models are used, namely multi-layer perceptron, recurrent and higher order neural network. They examine trading strategies for an equally weighted portfolio of six spreads, containing WTI, Brent, Unleaded Gasoline and Heating oil futures contracts. The best out-of-sample model is the recurrent neural network with a transitive filter.

benchmark strategy, where one has a long position in the petroleum futures market, all the time. This comparison enables us to assess whether the dynamic strategy of frequent rebalancing according to signals specified by the behavior of price differentials among petroleum futures is superior to static trading strategies. In doing so, we also consider the direction of causality and lead or lag relationships between these two markets. Furthermore, we also provide a framework for identifying deviations from long-run relationships among different petroleum futures markets that may lead to profitable trading positions. Hence, the intention of this paper is filling the gap in the oil market literature by providing evidence for trading opportunities. The study is concentrated on the intermarket and intercommodity linkages between the possible combinations of two NYMEX and two ICE futures contracts, namely WTI crude oil and NYMEX heating oil and ICE Brent crude oil and ICE gas oil, respectively. In- and out-of-sample tests of the suggested strategies versus alternative benchmarks validate inferences about the performance of such strategies.

The structure of this paper is as follows: Sections 2 and 3 present the statistical methodology and the empirical model of this study, respectively; the properties of the data are discussed in Section 4; Section 5 offers the empirical results; and Sections 6 and 7 describe the performance of the employed strategies and the conclusions, respectively.

2 STATISTICAL METHODOLOGY

To investigate the relationship between the cross combinations of the petroleum futures contracts, causality and error correction is of paramount importance. The pairs' causal relationship can be examined by using the following vector error correction model (VECM) (Johansen (1988)):

$$\Delta \mathbf{X_t} = \sum_{i=1}^{p-1} \Gamma_\mathbf{i} \Delta \mathbf{X_{t-1}} + \Pi \mathbf{X_{t-1}} + \boldsymbol{\varepsilon}_\mathbf{t}, \quad \boldsymbol{\varepsilon}_\mathbf{t} \sim N(0, \Sigma) \tag{1}$$

where $\mathbf{X_t}$ is a 2×1 vector of futures prices, each being $I(1)$ such that the first differenced series are I(0); Δ denotes the first difference operator; $\boldsymbol{\Gamma}_\mathbf{i}$ and Π are 2×2 coefficient matrices measuring the short- and long-run adjustment of the system to changes in $\mathbf{X_t}$, respectively; and $\boldsymbol{\varepsilon}_\mathbf{t}$ is a 2×1 vector of the vector of Gaussian stationary white noise processes with constant covariance matrix Σ.

The following steps are involved in our analysis. First, the existence of a stationary relationship between different pairs of futures prices is investigated in the VECM of Equation (1) through the λ_{max} and λ_{trace} statistics (Johansen (1988)), which test for the rank (Π). The rank (Π) in turn determines the number of cointegrating relationships. If Π has a full rank, that is 2, then all the variables in $\mathbf{X_t}$ are $I(0)$ and the appropriate modeling strategy is to estimate a vector autoregressive (VAR) model in levels. If rank $(\Pi) = 0$, Π is a 2×2 null matrix and the VECM of Equation (1) is reduced to a VAR model in first differences. On the other hand, if Π has a reduced rank, that is, then there exists one cointegrating vector and the coefficient matrix Π can be decomposed as $\Pi = \boldsymbol{\alpha}\boldsymbol{\beta}'$, where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}'$ are 2×1 vectors. Using this factorization, $\boldsymbol{\beta}'$ represents the vector of cointegrating parameters and $\boldsymbol{\alpha}$ is the vector of error correction coefficients measuring the speed of convergence to the long-run steady state.

Second, if the pairs of futures prices are cointegrated, then causality must exist in at least one direction (Granger (1986)). A time series, say WTI_t, is said to Granger cause another time series, say Brent_t, if the present values of Brent_t can be predicted more accurately by using past values of WTI_t than by not doing so, considering also other relevant information including past values of Brent_t (Granger (1969)). If both WTI_t and Brent_t Granger cause each other, then there is a two-way feedback relationship between the two markets. The VECM of Equation (1) provides a framework for valid inference in the presence of $I(1)$ variables. Moreover, modeling the series using the Johansen (1988) procedure results in more efficient estimates of the cointegrating relationship than the Engle and Granger (1987) estimator (see Gonzalo (1994)). In addition, Johansen (1988) tests are shown to be fairly robust to the presence of non-normality (Cheung and Lai (1993)) and heteroskedastic disturbances (Lee and Tse (1996)).

If $\mathrm{F}_{1,t}$ and $\mathrm{F}_{2,t}$ are the logprices of the two legs of the constructed spread at time t, the comovement and linkage between petroleum futures contracts can be examined using the VECM of Equation (1). The short-run price dynamics are expressed by the lagged cross-market terms, whereas the long-run price processes are reflected in the cointegration vector. The important element of the established cointegrating relationship is the error correction term (ECT), which can be regarded as the spread between logfutures prices $(\beta_1 F_{1,t} - \beta_2 F_{2,t} - \beta_0)$. In particular, the intercept term in the ECT, β_0, represents the long-run equilibrium relationship or the average spread.[3]

3 TRADING STRATEGIES

The aim of the cointegration analysis is to investigate the relationship between the different pairs of petroleum futures and then to develop a trading strategy that utilizes this relationship to identify investment timing opportunities. Therefore, we make use of the historical correlation and cointegration of the prices as indicators of market movements and, consequently, as signals for buying and/or selling decisions.

In practice, one can devise limitless trading rules and strategies, as there are multiple combinations of relationships between variables that can produce a trading signal and multiple parameterizations for a given family of rules. For instance, there are different combinations of MA rules reflecting different time spans in the estimation of MA prices; similarly, there are numerous parameterizations of filter rules, depending on how many standard deviations one allows before reversing a position. As it is beyond scope of this study to evaluate an exhaustive set of

[3] If rank $(\Pi) = 0$, spreading is not justified because zero rank means that the two-legged positions will tend to drift apart over time.

trading rules, we focus our efforts on three simple cases of MA rules, based on different petroleum futures spreads, to identify departures of the differential from the long-run equilibrium relationship.

MA trading strategies are based on the comparison of one fast (short) and one slow (long) MA of the spread of futures prices. For example, one such strategy is to compare a three-month MA of the spread with a one-week MA of the same series. In this setting, in any given month, a positive difference between the three-month MA and the one-week MA of the spread indicates a buy decision. This is because when the three-month MA is greater than the one-week MA, this can be interpreted as the series being below the long-run average (which is represented by the three-month MA). Consequently, this implies that the spread is lower than its long-run average or, alternatively, that the futures prices of the one commodity are undervalued relative to the other.

4 DESCRIPTION OF THE DATA AND PRELIMINARY ANALYSIS

The data set for this study comprises weekly futures prices for four energy commodities: NYMEX WTI crude oil, NYMEX heating oil, ICE Brent crude oil and ICE gas oil, covering the period August 9, 1989, to September 20, 2006, resulting 894 weekly observations. Futures prices are Wednesday prices; when a holiday occurs on Wednesday, Tuesday's observation is used. Data is collected from Datastream.

Nowadays, there are two major exchanges providing oil-derivative contracts: NYMEX and ICE in London. Other exchanges that trade oil-related contracts are the Tokyo Commodity Exchange (TOCOM) since 1999 (crude oil, gas oil, gasoline and kerosene futures) and the Dubai Mercantile Exchange (DME) since 2006, which constitutes the first energy futures exchange in the Middle East. This study concentrates on NYMEX and ICE futures, where the data set is sufficient for a weekly based analysis. NYMEX WTI contracts are traded for all deliveries within the next 30 consecutive months as well as for specific long-dated deliveries such as 36, 48, 72 and 84 months from delivery. Each contract is traded until the close of business on the third business day prior to the 25th calendar day of the month preceding the delivery month. NYMEX heating oil contracts are traded for all deliveries within the next 18 months and each contract of NYMEX heating oil is terminated on the last business day of the month preceding the delivery month. ICE Brent crude oil contracts are traded for all deliveries within the next 30 consecutive months and then half-yearly out to a maximum of seven years. Each contract is traded until the close of business on the business day immediately preceding the 15th day prior to the first day of the delivery month. ICE gas oil contracts are traded for all deliveries for 12 consecutive months forward, then quarterly out to 24 months and then half-yearly out to 36 months. Contracts expire two business days prior to the 14th calendar day of the delivery month.

Since the contracts' expiration dates are not matching, it is assumed that the investor will roll over to the front month pair of contracts, the first day of the last trading month. For instance, the February 2001 ICE Brent crude oil contract (for delivery in March 2001) expired on Tuesday, January 16, and the February 2001 NYMEX heating oil contract (for delivery in March 2001) expired on Wednesday, January 30. Consequently, the spread is constructed as February 2001 ICE crude oil – NYMEX heating oil. Switching to the front month pair of contracts occurs simultaneously on the first day of the expiration month for those contracts; that is on Wednesday, January 3, 2001, when the spread becomes March 2001 ICE crude oil – NYMEX heating oil, we call this spread the one-month spread. This way we ensure that the spreads are measured at the same point in time and we also avoid problems associated with thin trading and expiration effects, as these spreads are always liquid. In the same way, for the two-month spread, rolling over occurs on the first day of the month preceding the expiration month and so on. Having constructed the four continuous time series for the futures contracts, futures prices are then converted to the same unit of measure; that is US dollars per barrel (US$/bbl). Prices are then transformed to natural logarithms, and then the spreads are constructed, as discussed above.

Descriptive statistics of the one-, two-, three- and four-month spreads indicate that the spread series are serially correlated, heteroskedastic and non-normal. In addition, unit root tests reveal that all crude and petroleum futures prices are difference stationary. The spread series, on the other hand, are stationary, indicating that the VECM specification is the appropriate tool to uncover the relationships and mean-reverting properties of the spread series.[4]

5 EMPIRICAL RESULTS

Cointegration techniques are employed next to investigate the existence of a long-run relationship between the time series. The lag length of the VECM of Equation (1) is chosen on the basis of the Schwarz Bayesian Information Criterion (SBIC) (Schwarz (1978)). Johansen (1988) cointegration tests, presented in Table 1, indicate that all oil futures prices stand in a long-run relationship with each other. Since this condition is met, the pairs of futures prices evolve in proximity to one another, and any deviation from this relationship signals a trading opportunity, as cointegration implies that this departure will be restored. The normalized coefficient estimates of the cointegrating vector $\boldsymbol{\beta}' = (\beta_1 \beta_2 \beta_0)$ represent this long-run relationship between the series. Since the asymptotic distributions of the cointegration test statistics are dependent upon the presence of deterministic terms in the VECM, it is important to validate the inclusion or not of constant and/or linear trends in the system. Likelihood ratio tests[5] indicate that an intercept term should be included in the long-run relationship. The inclusion of an intercept term is also justified on the basis that it may capture the impact of constant parameters such

[4] To save space, descriptive statistics are not presented but are available from the authors.
[5] These tests follow Johansen (1991). The results are not presented here and are available from the authors.

TABLE 1 Johansen cointegration tests for petroleum futures spreads.

			Statistic		Error correction coefficients		CV $(1\ \beta_2\ \beta_0)$	Causality test	
	Lags	H_0:	λ_{max} test	λ_{trace} test	a_1	a_2	Normalized	$F_2 \rightarrow F_1$	$F_1 \rightarrow F_2$
Panel A: One-month futures									
$CB_t - GO_t$	2	$r = 0$	33.31	34.76	0.016	0.091	(1 −1.012 0.237)	1.510	16.84
		$r = 1$	1.458	1.458	(0.025)	(0.025)***		{0.470}	{0.000}***
$CL_t - HO_t$	1	$r = 0$	40.08	41.89	−0.004	0.077	(1 −0.978 0.078)	0.024	10.14
		$r = 1$	1.810	1.810	(0.027)	(0.024)***		{0.878}	{0.001}***
$CB_t - CL_t$	3	$r = 0$	28.99	30.65	−0.096	−0.004	(1 −1.044 0.205)	1.683	8.537
		$r = 1$	1.653	1.653	(0.079)	(0.076)		{0.641}	{0.036}**
$CB_t - HO_t$	1	$r = 0$	42.63	44.24	0.006	0.090	(1 −1.022 0.291)	0.058	16.02
		$r = 1$	1.610	1.610	(0.025)	(0.022)***		{0.810}	{0.000}***
$CL_t - GO_t$	1	$r = 0$	55.48	57.35	0.004	0.106	(1 −0.971 0.036)	0.026	14.11
		$r = 1$	1.868	1.868	(0.025)	(0.028)***		{0.872}	{0.000}***
$GO_t - HO_t$	2	$r = 0$	55.33	56.86	−0.025	0.142	(1 −1.006 0.040)	14.48	10.95
		$r = 1$	1.529	1.529	(0.061)	(0.052)***		{0.001}***	{0.004}***
Panel B: Two-month futures									
$CB_t - GO_t$	2	$r = 0$	25.27	26.75	0.013	0.075	(1 −1.016 0.253)	1.107	17.44
		$r = 1$	1.485	1.485	(0.026)	(0.024)***		{0.575}	{0.000}***
$CL_t - HO_t$	1	$r = 0$	32.12	33.80	0.021	0.083	(1 −0.979 0.085)	0.745	10.79
		$r = 1$	1.679	1.679	(0.025)	(0.025)***		{0.388}	{0.001}***
$CB_t - CL_t$	3	$r = 0$	33.17	34.74	−0.220	−0.127	(1 −1.045 0.207)	10.81	4.045
		$r = 1$	1.576	1.576	(0.085)	(0.082)		{0.013}**	{0.256}
$CB_t - HO_t$	1	$r = 0$	33.07	34.60	0.005	0.071	(1 −1.025 0.301)	0.030	8.122
		$r = 1$	1.532	1.532	(0.026)	(0.025)***		{0.863}	{0.004}***

TABLE 1 Continued.

			Statistic		Error correction coefficients		CV $(1\ \beta_2\ \beta_0)$	Causality test	
	Lags	H_0:	$\lambda_{\max}$ **test**	λ_{trace} **test**	a_1	a_2	**Normalized**	$F_2 \to F_1$	$F_1 \to F_2$
$CL_t - GO_t$	1	$r = 0$	50.97	52.63	0.021	0.113	(1 −0.977 0.057)	0.765	16.13
		$r = 1$	1.668	1.668	(0.024)	(0.028)***		{0.382}	{0.000}***
$GO_t - HO_t$	2	$r = 0$	58.46	59.99	−0.049	0.129	(1 −1.003 0.032)	13.55	7.901
		$r = 1$	1.523	1.523	(0.059)	(0.054)**		{0.001}**	{0.019}**
Panel C: Three-month futures									
$CB_t - GO_t$	2	$r = 0$	22.03	23.66	0.005	0.064	(1 −1.020 0.269)	1.100	18.58
		$r = 1$	1.633	1.633	(0.027)	(0.024)***		{0.578}	{0.000}***
$CL_t - HO_t$	1	$r = 0$	26.60	28.17	0.017	0.071	(1 −0.983 0.101)	0.390	7.263
		$r = 1$	1.570	1.570	(0.028)	(0.027)***		{0.533}	{0.007}***
$CB_t - CL_t$	3	$r = 0$	37.32	38.71	−0.291	−0.189	(1 −1.044 0.202)	14.04	4.570
		$r = 1$	1.393	1.393	(0.096)***	(0.099)*		{0.003}***	{0.206}
$CB_t - HO_t$	1	$r = 0$	29.88	31.41	−0.016	0.047	(1 −1.029 0.317)	0.295	2.735
		$r = 1$	1.530	1.530	(0.029)	(0.029)*		{0.587}	{0.098}*
$CL_t - GO_t$	1	$r = 0$	50.61	52.15	0.028	0.120	(1 −0.983 0.079)	1.229	16.72
		$r = 1$	1.537	1.537	(0.025)	(0.029)***		{0.268}	{0.000}***
$GO_t - HO_t$	2	$r = 0$	64.31	65.85	−0.123	0.065	(1 −1.004 0.032)	15.10	3.618
		$r = 1$	1.539	1.539	(0.063)*	(0.062)		{0.001}***	{0.164}
Panel D: Four-month futures									
$CB_t - GO_t$	2	$r = 0$	21.10	22.96	−0.014	0.047	(1 −1.026 0.291)	0.374	12.58
		$r = 1$	1.857	1.857	(0.028)	(0.024)**		{0.829}	{0.002}***
$CL_t - HO_t$	1	$r = 0$	25.28	26.88	−0.013	0.035	(1 −0.990 0.122)	0.294	1.867
		$r = 1$	1.595	1.595	(0.024)	(0.026)		{0.588}	{0.171}

TABLE 1 Continued.

	Lags	H_0:	Statistic λ_{max} test	λ_{trace} test	Error correction coefficients a_1	a_2	CV $(1\ \beta_2\ \beta_0)$ Normalized	Causality test $F_2 \to F_1$	$F_1 \to F_2$
$CB_t - CL_t$	3	$r = 0$	43.33	44.64	−0.348	−0.232	(1 −1.044 0.202)	16.39	6.427
		$r = 1$	1.323	1.323	(0.104)***	(0.098)**		{0.001}***	{0.093}*
$CB_t - HO_t$	1	$r = 0$	31.05	32.53	−0.055	0.010	(1 −1.038 0.345)	2.458	0.087
		$r = 1$	1.482	1.482	(0.035)	(0.032)		{0.117}	{0.768}
$CL_t - GO_t$	1	$r = 0$	51.75	53.27	0.023	0.118	(1 −0.987 0.097)	0.867	17.58
		$r = 1$	1.516	1.516	(0.025)	(0.028)***		{0.352}	{0.000}***
$GO_t - HO_t$	2	$r = 0$	65.80	67.34	−0.166	0.004	(1 −1.005 0.035)	19.97	1.511
		$r = 1$	1.539	1.539	(0.057)***	(0.059)		{0.000}***	{0.470}

Lags is the lag length of the unrestricted VAR model in levels. A VAR with p lags of the dependent variable can be reparameterized in a VECM with p–1 lags of first differences of the dependent variable plus the error correction term. The lag length is chosen on the basis of Schwarz Bayesian Information Criterion (Schwarz (1978)).
λ_{max} tests the null hypothesis of r cointegrating vectors against the alternative of $r + 1$. The 5% critical values for H_0: $r = 0$ and H_0: $r = 1$ are 15.67 and 9.24, respectively. Critical values obtained from Osterwald-Lenum (1992).
λ_{trace} tests the null hypothesis that there are at most r cointegrating vectors against the alternative that the number of cointegrating vectors is greater than r. The 5% critical values for H_0: $r = 0$ and H_0: $r = 1$ are 19.96 and 9.24, respectively. Critical values obtained from Osterwald-Lenum (1992).
The coefficients of the error correction term are estimated using the VECM of Equation (1). Figures in parentheses are standard errors, which are calculated using a Newey–West (1987) correction for serial correlation and heteroskedasticity. *, ** and *** indicate significance at 10%, 5% and 1%, respectively.
$\beta' = (1\beta_1, \beta_2)$ are the coefficient estimates of the cointegrating vector, where the coefficient of $F_{1,t-1}$ is normalized to be unity, β_1 is the intercept term and β_2 is the coefficient of $F_{2,t-1}$.
The statistic of the causality test is distributed as χ^2 with degrees of freedom equal to the number of the restrictions. Figures in braces are the corresponding p-values.

as premia and discounts, insurance charges, and quality and location differentials for the different types of the oil commodities.

Along with the normalized coefficients of the unrestricted cointegrating vectors, Table 1 reports the estimated error correction coefficients from the VECM. The standard errors are corrected for heteroskedasticity and serial correlation using the Newey–West (1987) method, for all the regressions. The speed of adjustment of futures prices to their long-run relationship, measured by the $\alpha 1$ and $\alpha 2$ estimated coefficients, is expected to be negative in the first equation and positive in the second equation. This implies that in response to a positive deviation from their long-run relationship at period $t-1$, ie, $F_{1,t-1} - \beta_2 F_{2,t-1} - \beta_0 > 0$, the futures price of the first (second) leg of the spread will decrease (increase) in value, in order to restore the long-run equilibrium. As it can be seen in Table 1, all the significant error correction coefficients have the correct sign, with the exception of the three- and four-month intercrude spread.

More rigorous investigation of the interactions between the variables can be obtained by performing Granger causality tests. According to the Granger (1986) representation theorem, if two prices are cointegrated, causality must exist in at least one direction. Several observations merit attention regarding the joint dynamics of the price processes. The complex structure of the oil market implies that great caution should be taken when making inferences about causality, because its direction is not known *a priori*. The assumption that crude oil prices are expected to Granger cause petroleum product prices can be based on the fact that, first, crude oil prices are determined by the worldwide supply and demand as opposed to refined products where regional supply and demand dynamics are important.[6] Of course, refined products are linked to the international market through crude oil prices, which represent a significant input production cost. Second, demand for, say WTI crude oil, is not likely to be driven by the demand for heating oil alone, since a substantial amount of the refined crude oil is transformed to other products such as gasoline, naphtha and kerosene. Test for the joint significance of the lagged cross-market returns and ECT confirm the above setting;[7] that is the existence of one-way relationship, with crude oil leading the information discovery process – in the equations for either NYMEX or ICE crude oil futures paired with ICE gas oil and NYMEX heating oil across all maturities. The exceptions are the four-month ICE Brent (NYMEX heating oil and NYMEX WTI) heating oil pairs where there is no significant lead–lag relationship. Finally, the estimates of the error correction coefficients, overall, in terms of magnitude and significance,

[6] Crude oil dominates the world trade because its transportation is carried in large vessels and economies of scale are achieved. On the other hand, transportation of higher-value refined products is carried in smaller vessels, is more expensive and usually is a restricted service for shorter distance routes. In general, refineries are located next to demand sources to avoid the need for transportation, government import–export barriers, etc.

[7] Nevertheless, it is not unlikely for the refined product prices to pull crude oil prices, since demand for crude oil is derived from petroleum products' demand, which in turn is generated from transportation, industrial and residential needs.

indicate that, basically, heating oil and gas oil prices move to adjust the long-run equilibrium, whereas crude oil prices are not responsive to departures from the long-run mean of the differential.

In the intercrude market, the first nearby spread indicates that Brent has explanatory power on WTI futures only in the short run, but the two are not responsive to the differential. When the horizon increases to two-, three- and four-month futures, the picture is reversed and now NYMEX WTI Granger causes ICE Brent futures, at 5% significance level. This is expected since the United States reflects by far the largest oil consumer and importer of crude oil, and this dependency introduces a high degree of sensitivity of the international market to the US oil prices, which perhaps makes the WTI market dominant in terms of information discovery (see, for instance, Lin and Tamvakis (2001)). However, this is not reflected in the one-month futures prices. This can be explained by the fact that the volatility of futures prices increase as time to maturity approaches (Samuelson (1965)), because futures prices tend to converge to the actual spot prices and the contracts become more sensitive to information flows, resulting an interruption of the core lead–lag relationship. Error correction coefficients in the Brent futures equation have the correct negative sign, and they are all significant except the one-month case. This indicates that in response to a positive shock, the price will decrease to restore the long-run mean (spread is constructed as log-Brent minus log-WTI). In the WTI futures equation, error correction coefficients for the three- and four-month spreads are also negative and at 5% significance level. However, the magnitude is lower compared with Brent, indicating that the degree of responsiveness of WTI to the spread is inferior and, actually, adjustment to restore the long-run equilibrium is mainly due to ICE Brent crude futures price movements.

In addition, Granger causality tests in Table 1 indicate that ICE gas oil is not only Granger caused by NYMEX heating oil, but also Granger causes NYMEX heating oil for the one- and two-month spreads. This two-way feedback relationship holds at 1% and 5% significance level for the one- and two-month futures, respectively. The error correction estimates have the correct sign, negative for the ICE petroleum product and positive for the NYMEX petroleum product, to ensure convergence in the long run. The long-run equilibrium relationship is restored after adjustment of heating oil prices in the one- and two-month futures prices, whereas in the three- and four-month case heating oil prices are not responsive to the differential and any possible adjustment originates from the ICE gas oil market. One would normally expect response to the differential mainly from the gas oil market because it is smaller than the US heating oil market. Another reason is that, since the demand for crude oil is driven by the demand for refined products and since the United States is the biggest importer of crude oil, increased demand in the United States is more likely to put pressure in oil prices. Hence, ICE gas oil market is more likely to be driven by the US market. However, the change in pattern from the shorter- to the longer-term maturities may be due to the fact that uncertainty of future expectations regarding prices, demand, supply, inventories and unknown weather conditions is relatively higher in the longer term but the responsiveness of prices to new information is slower, and slower is the price transmission mechanism as well.

6 PERFORMANCE OF MOVING AVERAGE TRADING RULES

The trading strategy employed in this paper, which combines the fundamental relationship between variables with technical trading rules, is based on the deviation of the spread from its long-run mean. In order to determine the timing of buy or sell, we devise four MA series using the differential between log futures prices: one fast [MA(1)] and three slow [MA(4), MA(8) and MA(12)]. The difference between the two constructed MA series is then used as an indicator that signals whether to buy or sell in the petroleum futures spread. The signals are based on the sign of the difference between the slow and the fast MA in such a way that a positive difference is a buy signal, while a negative difference is a sell signal. For instance, regarding the ICE – NYMEX intercrude spread (constructed as log-Brent minus log-WTI), if MA(8) > MA(1), then a long position on the spread will be initiated by purchasing one ICE Brent crude oil contract and selling one NYMEX WTI crude oil contract. The position will be held until the relationship between the two MA series is inverted, ie, MA(8) < MA(1). Then, simultaneously, the long position will be closed and a short position on the ICE – NYMEX intercrude spread will be initiated.[8] For comparison purposes, we also consider the performance of a benchmark buy and hold strategy. The performance of this strategy reflects the income of an investor who maintains a long outright position on the petroleum futures market (across the whole sample period).

One important element when evaluating dynamic trading strategies is the incurred transaction cost that these strategies involve, arising from the frequent rebalancing of the portfolio of interest. For the purposes of this study, a transaction cost of 0.2% for every round trip of initiating and reversing trade is deemed reasonable. Our assumption is comparable with other studies in the literature (Dunis *et al* (2006b); Poitras and Teoh (2003); and Girma and Paulson (1999)). Commission charges (ie, any fixed fees such as brokerage and other transaction fees), on the other hand, are very low and, usually, are negligible.[9]

[8] The exercise was repeated for the historical and bootstrap simulations, using different filter rules. We used rolling windows for standard deviations, in order to filter the signal for entry and exit points in the market. For instance, in the ICE – NYMEX intercrude spread, a long position on the spread will be initiated if MA(8) > MA(1) + $X\sigma$, where X is the number of standard deviation units. The position will be held until the relationship between the two MA series becomes MA(8) > MA(1) – $X\sigma$, where a short position will be initiated simultaneously. Then, the long position will be closed and a short position on the ICE – NYMEX intercrude spread will be initiated. Application of such filters did not change the results qualitatively, but in general, as the filter size increased, annualized returns decreased. The filters used were $\pm 0.25\sigma$, $\pm 0.5\sigma$, $\pm\sigma$ and $\pm 1.5\sigma$, and results are available from the authors upon request.

[9] For instance, in the NYMEX division, the half-turn trading fee is approximately between 0.04 and 0.18 cents per barrel subject to whether the trade is undertaken by a member or non-member of the exchange (see www.nymex.com).

TABLE 2 Historical and bootstrap simulation of three-, two- and one-month MA trading strategies.

	Three-month strategy (MA12 vs MA1)				Two-month strategy (MA8 vs MA1)				One-month strategy (MA4 vs MA1)			
	Mean Ret	**SD**	**Sharpe ratio**	**Sharpe R improve-ment**	**Mean Ret**	**SD**	**Sharpe ratio**	**Sharpe R improve-ment**	**Mean Ret**	**SD**	**Sharpe ratio**	**Sharpe R improve-ment**
Panel A: One month												
$CB_t - GO_t$	18.19	18.08	1.006	0.762***	20.60	18.03	1.142	0.882 ***	25.60	(17.84)	1.435	1.212 ***
$CL_t - HO_t$	4.573	15.75	0.290	0.015	5.859	15.74	0.372	0.094	8.797	15.69	0.561	0.264
$CB_t - CL_t$	3.659	9.469	0.386	0.160	6.632	9.431	0.703	0.374 *	11.02	9.308	1.184	0.856 ***
$CB_t - HO_t$	6.009	16.66	0.361	0.086	4.535	16.68	0.272	−0.030	7.093	16.64	0.426	0.164
$CL_t - GO_t$	19.93	19.41	1.027	0.862***	24.42	19.28	1.267	1.093 ***	30.03	19.11	1.571	1.349 ***
$GO_t - HO_t$	23.01	15.20	1.513	1.296***	27.12	15.03	1.804	1.582 ***	29.17	14.96	1.950	1.737 ***
Panel B: Two month												
$CB_t - GO_t$	18.88	15.71	1.202	0.920 ***	20.85	15.65	1.332	1.090 ***	27.71	15.40	1.800	1.577 ***
$CL_t - HO_t$	−1.461	12.46	−0.117	−0.380	0.293	12.45	0.024	−0.216	6.927	12.37	0.560	0.295*
$CB_t - CL_t$	3.108	7.840	0.397	0.075	3.738	7.817	0.478	0.144	4.495	7.851	0.572	0.246
$CB_t - HO_t$	7.228	13.17	0.549	0.224	8.025	13.16	0.610	0.299	7.202	13.18	0.546	0.211
$CL_t - GO_t$	18.61	16.66	1.117	0.934 ***	20.52	16.61	1.235	1.046 ***	26.12	16.43	1.590	1.352 ***
$GO_t - HO_t$	25.35	13.44	1.886	1.554 ***	26.87	13.36	2.011	1.719 ***	24.82	13.49	1.840	1.580 ***
Panel C: Three month												
$CB_t - GO_t$	20.23	14.77	1.370	1.126 ***	24.68	14.61	1.689	1.423 ***	29.06	(14.42)	2.015	1.730 ***
$CL_t - HO_t$	−4.285	10.70	−0.400	−0.570	−3.026	10.70	−0.283	−0.497	5.556	10.63	0.523	0.223
$CB_t - CL_t$	3.468	7.455	0.465	0.127	3.266	7.456	0.438	0.093	4.380	7.437	0.589	0.172
$CB_t - HO_t$	1.914	11.64	0.164	−0.007	3.695	11.65	0.317	0.067	10.21	11.54	0.885	0.464 *
$CL_t - GO_t$	19.09	15.33	1.246	1.053 ***	21.73	15.24	1.426	1.227 ***	26.76	15.05	1.779	1.552 ***
$GO_t - HO_t$	26.83	12.74	2.106	1.807 ***	28.42	12.68	2.242	1.998 ***	27.50	12.73	2.160	1.873***

TABLE 2 Continued.

	Three-month strategy (MA12 vs MA1)				Two-month strategy (MA8 vs MA1)				One-month strategy (MA4 vs MA1)			
	Mean Ret	SD	Sharpe ratio	Sharpe R improvement	Mean Ret	SD	Sharpe ratio	Sharpe R improvement	Mean Ret	SD	Sharpe ratio	Sharpe R improvement
Panel D: Four month												
$CB_t - GO_t$	21.50	14.25	1.509	1.222***	24.67	14.12	1.747	1.470***	27.73	13.98	1.983	1.675***
$CL_t - HO_t$	−6.059	10.04	−0.604	−0.782	−3.594	10.03	−0.358	−0.654	2.305	10.03	0.230	−0.107
$CB_t - CL_t$	3.833	7.656	0.501	0.141	4.878	7.615	0.641	0.278	5.986	7.585	0.789	0.338
$CB_t - HO_t$	0.903	11.05	0.082	−0.185	1.733	11.06	0.157	−0.213	6.791	11.00	0.617	0.210
$CL_t - GO_t$	19.67	14.52	1.355	1.114***	21.86	14.45	1.513	1.252***	25.33	14.31	1.770	1.546***
$GO_t - HO_t$	24.75	12.32	2.009	1.696***	25.67	12.28	2.090	1.762***	26.16	12.29	2.129	1.859***

Panel E: Buy and hold strategies

	One month			Two month			Three month			Four month		
	Mean Ret	SD	Sharpe ratio	Mean Ret	SD	Sharpe ratio	Mean Ret	SD	Sharpe ratio	Mean Ret	SD	Sharpe ratio
CB_t	6.719	31.00	0.217	6.907	29.10	0.237	7.150	27.68	0.258	7.400	25.89	0.286
GO_t	6.980	32.05	0.218	7.198	29.90	0.241	7.352	28.13	0.261	7.470	26.89	0.281
CL_t	6.276	31.51	0.199	6.469	29.37	0.220	6.724	27.74	0.242	7.072	26.22	0.270
HO_t	6.662	32.28	0.206	6.787	29.27	0.232	6.903	27.08	0.255	7.005	25.24	0.278

Mean Ret are the % annualized returns and SD are the % annualized standard deviations.
Sharpe ratios are calculated using the formula R / STD.
Mean Ret, SD and Sharpe ratios are those from the historical simulation of the different strategies.
Improvement in Sharpe ratio is the excess Sharpe ratio of the spread MA-based trading compared with the buy and hold strategy of petroleum futures, across 1,000 simulations.
The 1,000 realizations of the trading strategies are based on the stationary bootstrap of Politis and Romano (1994).
***,** and * measure the significance level for which we can reject a one-tail test on the null that Sharpe ratios are not different between the MA and BH strategies at 1%, 5% and 10% significance level, respectively.
Historical and bootstrap simulation is performed assuming transaction costs of 0.2%.

The performance of different strategies is presented in Table 2. As indicated by the annualized returns across all cases, it can be noted that for every spread, there exists an MA strategy to produce higher returns than the benchmark strategies. When the MA strategy is used, the profit potential is, overall, increased and the corresponding average annualized returns lie between −6.06% and 29.17%, whereas the interval of returns for the buy and hold strategies is in the range of 6.28−7.47%. Apart from higher returns (overall), the results of MA trading rules also indicate reduction in the standard deviations. Therefore, Sharpe ratios (the ratio of average return and standard deviation) also indicate that most MA strategies outperform the buy and hold strategy. All MA(4,1), MA(8,1) and MA(12,1) strategies outperform the buy and hold strategies, with some exceptions in the NYMEX crude – heating oil and ICE Brent – NYMEX heating oil spreads. The interval of the benchmarks' Sharpe ratio is 0.199–0.286, whereas under the MA trading rule applied for the spread, the range becomes −0.604 to 2.242.

Figures 1 and 2 plot the cumulative returns of the three MA strategies [MA(12,1), MA(8,1) and MA(4,1)] against the static trading strategies for the spreads under consideration. The initial investment for the calculation of the cumulative returns is set to US$1. We plot these cumulative returns for the first nearby contracts of the spreads CB_t – CL_t and GO_t – HO_t under transaction costs of 0.2%. These graphs clearly illustrate the benefits of using trading signals to identify relative mispricing in the petroleum futures markets, as the cumulative returns based on MA rules reached levels of more than 100 US dollars [MA(4,1)] in the three spreads of ICE Brent\NYMEX WTI\NYMEX heating oil against ICE gas oil.

FIGURE 1 Historical cumulative return of MA strategies against buy and hold in the one-month ICE Brent and NYMEX WTI futures markets (CB_t – CL_t).

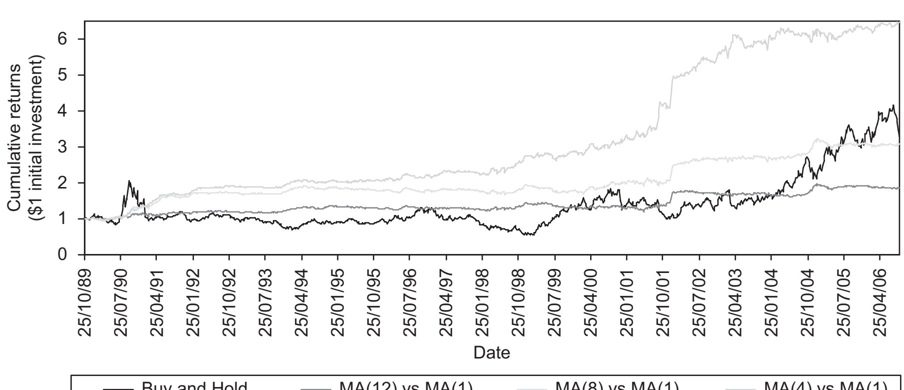
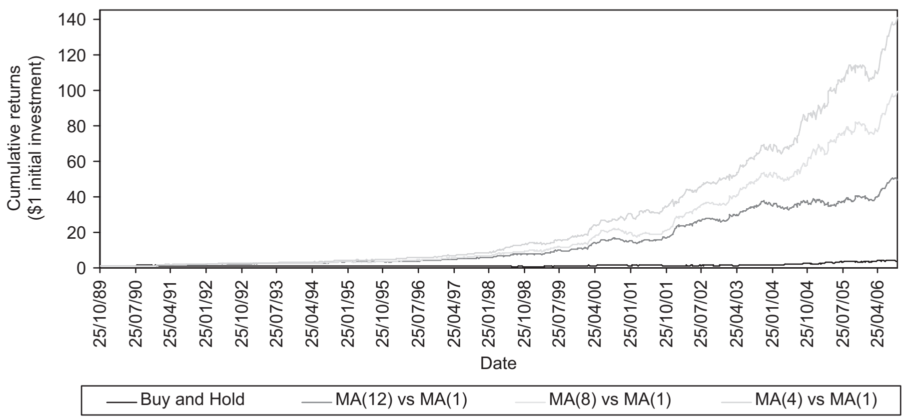


FIGURE 2 Historical cumulative return of MA strategies against buy and hold in the one-month ICE Gas oil and NYMEX heating oil futures markets ($GO_t - HO_t$).



7 DATA SNOOPING AND THE STATIONARY BOOTSTRAP

The results in Section 6 are encouraging regarding the performance of our proposed trading strategies. However, an important issue that arises when evaluating trading rules is that of data snooping. According to Sullivan *et al* (1999) and White (2000), data snooping occurs when a data set is used more than once for data selection and inference purposes. In other words, using the same data set frequently for testing trading strategies may increase the probability of having satisfactory results purely due to chance or due to the use of posterior information rather than the superior ability of the trading strategies.

The method most commonly used in the literature to assess the performance of trading strategies and test for data snooping is bootstrap. The bootstrap, introduced by Efron (1979), is a resampling method that uses the empirical distribution of the statistic of interest, rather than the theoretical distribution implied by the statistical theory, to conduct statistical inference. The main advantage of bootstrap is that it can approximate the properties of the sampling distribution of the underlying statistic even when such a distribution is not parametrically defined, or the underlined statistic is complex and not easy to obtain. Bootstrap techniques have also been used by Brock *et al* (1992), who test whether trading results from some trading rules can be explained by time-series models, and Sullivan *et al* (1999), who use bootstrap to test the joint performance of several technical rules.

However, when ordinary bootstrap techniques are applied to serially dependent observations, as in the case of petroleum and petroleum product futures prices, the resampled series will not retain the statistical properties of the original data set and yield inconsistent results and statistical inference (see Ruiz and Pascual (2002)).

In view of that, we employ the stationary bootstrap method of Politis and Romano (1994). This procedure is based on resampling blocks of random length, where the length of each block follows a geometric distribution. This procedure generates random samples that preserve the serial dependence property of the original series and are also stationary. This is important since our proposed trading strategy relies on the premise that the differentials of the futures prices under study are stationary; see Appendix for technical details.

Therefore, in order to assess the performance of our trading strategies, we use the stationary bootstrap technique to regenerate random paths that futures prices may have possibly followed over the sample period, while maintaining the distributional properties of the original series. We then implement the proposed trading strategies using the simulated prices' series which, in turn, generate a distribution of trading statistics under the different trading rules. Therefore, our approach in using bootstrap is different from the previous literature in the sense that we bootstrap to generate paths of the spread series and subsequently assess the profitability of MA-based trading rules. This approach follows Alizadeh and Nomikos (2006).

We start by bootstrapping the logdifference series. Then these bootstrapped series are transformed back into levels to construct the spreads that are used to trigger buy and sell decisions based on the MA trading strategies. In implementing these strategies, we consider 0.2% transaction costs. As benchmark models, we also consider the buy and hold trading strategy in which one is always long in either leg of the spread and, hence, benefits from a possible income irrespective of the level and fluctuation of the spread. Both the MA and the two buy and hold strategies (two legs of the spread) are implemented for each one of the 1,000 bootstrapped series, thus generating a series of empirical distributions of mean returns and Sharpe ratios. Under the null hypothesis that a dynamic strategy is no better than a buy and hold strategy, or, equivalently, that there is no information or signals in the original spread series, the profit from the MA strategies should be no better than the profit from the buy and hold strategies.

The results of the bootstrap simulations are reported in Table 2 in terms of improvement in Sharpe ratios. The mean annualized returns (obtained as the mean return from the trading strategies implemented on the 1,000 bootstrapped series) and average Sharpe ratio across the bootstrapped series are not presented here since they are very similar to those observed in the empirical series under the same trading rule. Furthermore, the comparative performance is also similar; for instance, the higher returns come from the three spreads that include ICE gas oil in the one leg, ie, ICE Brent\NYMEX WTI\NYMEX heating oil versus ICE gas oil, whereas the worst performance is achieved by the four-month NYMEX crude – heating oil spread for the MA(12,1) strategy. Overall, the MA-based trading rules seem to outperform the static investment tactics both in terms of increasing average returns and in terms of the Sharpe ratios. Buy and hold oil futures strategies' Sharpe ratios vary from 0.220 to 0.307. MA(12,1)-based Sharpe ratios lie between −0.782 and 1.807, MA(8,1) from −0.351 to

2.281 and the MA(4,1) spread strategy has a better downside limit with Sharpe ratios in the range of 0.195–2.156. In all MA rules and maturities, the highest Sharpe ratio is achieved by the ICE gas oil – NYMEX heating oil spread. With the exception of the MA(12,1)- and MA(8,1)-based strategies of the NYMEX crude – heating oil spread and ICE Brent – NYMEX heating oil spread, all the differences in the Sharpe ratios have a positive sign, denoting improvement. For example, the Sharpe ratios of the ICE gas oil – NYMEX heating oil spread reveal a more than eightfold increase compared with the buy and hold strategies.

More formal statistical tests are conducted by testing whether the excess performance of the Sharpe ratio, based on the bootstrap simulations, is significantly different from zero. More specifically, for each simulated series, we estimate the excess Sharpe ratio of the MA trading strategy relative to the buy and hold strategy. We then construct the p-values for the tests in Table 2. These are simply calculated as the ratio of frequency of occurrence of negative excess Sharpe ratios over the total number of simulations (1,000 replications) and reflect significance level, for which the null hypothesis that there is no significant difference between the Sharpe ratios can be rejected, using a one-tail test. Overall, these results indicate that the MA strategies can provide significant increases in Sharpe ratios compared with the ordinary buy and hold strategies. More specifically, four out of six, one-month spreads achieve significantly higher Sharpe ratios compared with the buy and hold strategy. The same is true for the two- and three-month case, whereas in the case of four-month spreads, the figure is reduced to three out of six. Overall, p-values provide additional support for the robustness of the superiority of the MA trading strategies compared with buy and hold benchmarks. Significantly higher Sharpe ratios at 1% significance level are achieved in the one-, two-, three- and four-month spreads of ICE Brent, NYMEX WTI and heating oil against ICE gas oil for all MA strategies as well as the one-month MA(4,1) strategy for the ICE – NYMEX intercrude spread. Significance is also achieved at 10% significance level for the latter spread [MA(8,1) strategy] as well as for the one- and two-month MA(4,1) strategy of NYMEX WTI – heating oil and finally for the three-month MA(4,1) strategy of ICE Brent – NYMEX heating oil spread.

Figures 3 and 4 plot the distributions of returns of three MA strategies [MA(12,1), MA(8,1) and MA(4,1)] against the static trading strategies using the bootstrap technique for the spreads under consideration. We plot these distributions for the first nearby contracts of the spreads $CB_t - CL_t$ and $GO_t - HO_t$, under transaction costs of 0.2%. These graphs clearly illustrate the benefits of using trading signals to identify relative mispricing in the petroleum futures markets, as the distribution of simulated returns based on MA rules show relatively lower dispersion, in all cases, and significant shifts to the right, in three out of six cases (the three spreads that include ICE gas oil).

Provided that transaction costs are a significant trading cost factored in dynamic trading strategies and, also, given that these costs are subject to variations,

FIGURE 3 Bootstrapped distribution of returns of MA strategies against buy and hold in the one-month ICE Brent and NYMEX WTI futures markets ($CB_t - CL_t$).

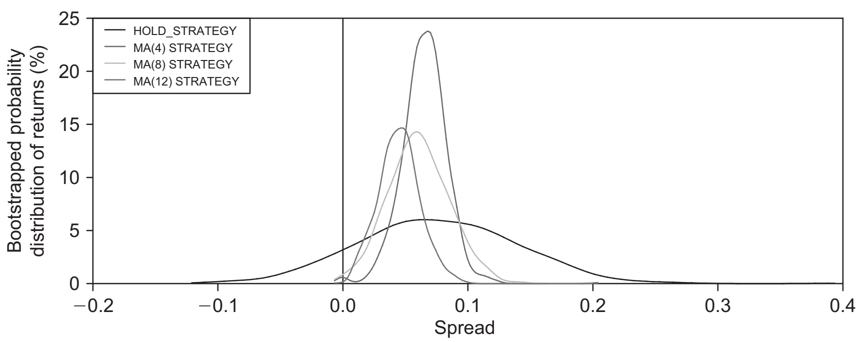
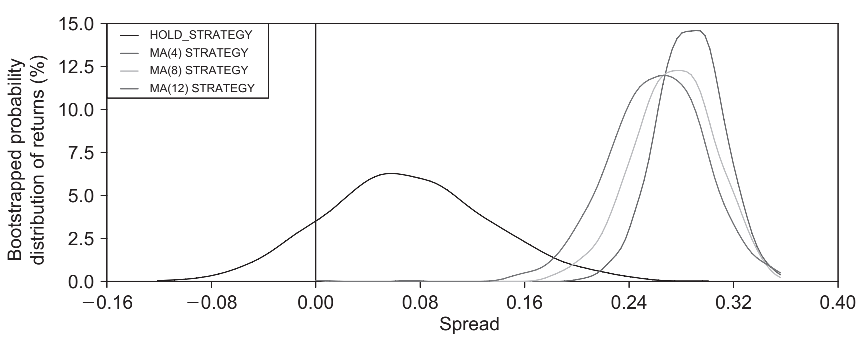


FIGURE 4 Bootstrapped distribution of return of MA strategies against buy and hold in the one-month ICE gas oil and NYMEX heating oil futures markets ($GO_t - HO_t$).



depending on the type of trader (eg, member or non-member) and the market (ie, United States, United Kingdom or both) involved, the performance of the MA trading strategies is tested, applying different sets of transaction costs in the bootstrap simulation, through sensitivity analysis.[10] Table 3 reports the results

[10] For economy of space, sensitivity analysis results present the outcome of the simulation of the one-month-based MA strategy that proved to be (overall) the best, under the 0.2% transaction costs' case. Another reason is that the specific strategy is expected to have the higher degree of sensitivity to transaction costs, as the fastest MA [MA(4,1)] requires more frequent rebalancing of the portfolio. Sometimes, the number of trades increases to double compared with MA(12,1) or MA(8,1) strategy. Results, regarding the number of trades in both historical and bootstrap simulations as well as sensitivity analysis, are available from the authors upon request.

TABLE 3 Sensitivity of MA trading strategies to different levels of transaction costs – one-month MA model.

	Transaction costs 0.3%				Transaction costs 0.4%				Transaction costs 0.5%			
	Mean return	**Sharpe ratio**	**Improvement in Sharpe ratio**		**Mean return**	**Sharpe ratio**	**Improvement in Sharpe ratio**		**Mean return**	**Sharpe ratio**	**Improvement in Sharpe ratio**	
Panel A: One month												
$CB_t - GO_t$	23.06	1.344	1.101	{0.000}***	20.61	1.206	0.985	{0.000}***	18.88	1.097	0.862	{0.000}***
$CL_t - HO_t$	5.547	0.367	0.143	{0.300}	3.776	0.250	0.029	{0.406}	2.415	0.158	−0.044	{0.499}
$CB_t - CL_t$	8.135	0.869	0.628	{0.024}**	6.126	0.641	0.407	{0.099}*	4.107	0.420	0.181	{0.274}
$CB_t - HO_t$	4.970	0.308	0.074	{0.347}	3.239	0.203	−0.029	{0.513}	1.470	0.093	−0.155	{0.654}
$CL_t - GO_t$	27.07	1.465	1.250	{0.000}***	24.93	1.345	1.131	{0.000}***	23.04	1.241	1.019	{0.000}***
$GO_t - HO_t$	26.45	1.823	1.580	{0.000}***	24.07	1.645	1.395	{0.000}***	21.88	1.488	1.249	{0.000}***
Panel A: Two month												
$CB_t - GO_t$	24.79	1.663	1.409	{0.001}***	22.79	1.518	1.259	{0.000}***	20.70	1.371	1.105	{0.000}***
$CL_t - HO_t$	4.545	0.351	0.094	{0.264}	2.763	0.228	−0.017	{0.449}	1.009	0.085	−0.166	{0.686}
$CB_t - CL_t$	1.949	0.248	−0.013	{0.466}	0.144	0.013	−0.239	{0.720}	−1.560	−0.217	−0.477	{0.851}
$CB_t - HO_t$	3.937	0.306	0.038	{0.406}	2.124	0.167	−0.093	{0.554}	0.340	0.028	−0.213	{0.696}
$CL_t - GO_t$	23.16	1.447	1.203	{0.000}***	21.29	1.331	1.094	{0.000}***	19.41	1.207	0.961	{0.000}***
$GO_t - HO_t$	21.96	1.672	1.405	{0.001}***	19.07	1.492	1.228	{0.001}***	17.55	1.321	1.048	{0.000}***

TABLE 3 Continued.

	Transaction costs 0.3%				Transaction costs 0.4%				Transaction costs 0.5%			
	Mean return	**Sharpe ratio**	**Improvement in Sharpe ratio**		**Mean return**	**Sharpe ratio**	**Improvement in Sharpe ratio**		**Mean return**	**Sharpe ratio**	**Improvement in Sharpe ratio**	
Panel C: Three month												
$CB_t - GO_t$	25.89	1.834	1.550	{0.000}***	23.57	1.676	1.383	{0.000}***	21.79	1.533	1.241	{0.002}***
$CL_t - HO_t$	3.108	0.301	0.016	{0.409}	1.243	0.133	−0.147	{0.643}	−0.050	−0.054	−0.324	{0.806}
$CB_t - CL_t$	1.510	0.192	−0.084	{0.545}	−0.433	−0.084	−0.374	{0.756}	−2.519	−0.369	−0.653	{0.860}
$CB_t - HO_t$	6.607	0.586	0.295	{0.155}	4.754	0.424	0.154	{0.302}	3.011	0.269	−0.007	0.458
$CL_t - GO_t$	24.43	1.674	1.417	{0.000}***	22.34	1.525	1.263	{0.000}***	20.12	1.373	1.100	{0.000}***
$GO_t - HO_t$	24.35	1.974	1.684	{0.000}***	21.87	1.767	1.486	{0.000}***	19.78	1.582	1.297	{0.000}***
Panel D: Four month												
$CB_t - GO_t$	24.34	1.802	1.497	{0.000}***	22.06	1.639	1.330	{0.000}***	19.92	1.471	1.167	{0.000}***
$CL_t - HO_t$	0.001	0.002	−0.280	{0.737}	−1.671	−0.177	−0.464	{0.834}	−3.436	−0.360	−0.650	{0.875}
$CB_t - CL_t$	2.435	0.324	0.015	{0.451}	0.452	0.034	−0.262	{0.680}	−1.532	−0.249	−0.557	{0.825}
$CB_t - HO_t$	3.631	0.342	0.040	{0.407}	1.938	0.185	−0.109	{0.569}	0.270	0.026	−0.272	{0.705}
$CL_t - GO_t$	23.48	1.704	1.409	{0.000}***	21.21	1.525	1.234	{0.001}***	19.10	1.372	1.088	{0.001}***
$GO_t - HO_t$	23.41	1.972	1.658	{0.000}***	21.27	1.783	1.504	{0.001}***	18.78	1.567	1.247	{0.000}***

See notes in Table 2.

for transaction costs in the range of 0.3%, 0.4% and 0.5%, respectively, for all the maturities of each spread trading strategies, based on the fastest MA.

One first important observation is that the three out of the six spreads, namely $CB_t - GO_t$, $CL_t - GO_t$ and $GO_t - HO_t$, produce significantly higher (at 1% significance level) Sharpe ratios than the buy and hold strategies, even in the high zone of 0.5% transaction costs. The annualized returns are in the range of 18.9–23.0%, which means more than twofold to fourfold increase compared with the buy and hold strategies. This implies that in the specific three spread markets, the mean-reverting tendencies, as implied by the fair value of the spread (estimated using Johansen (1988) cointegration procedure), are not fully impounded in the expectations of the market and abnormal excess returns become possible. Regarding the other three spreads, $CL_t - HO_t$, $CB_t - CL_t$ and $CB_t - HO_t$, relative mispricing is also discernible, subject to transaction costs. When transaction costs are 0.3% and 0.4% significantly higher Sharpe ratios are also achieved by the one-month intercrude $CB_t - CL_t$ spread.

Finally, one vital question that arises is why some spreads perform better than others and how we can identify these trading opportunities. Although petroleum futures prices move together in the long run, as indicated by λ_{trace} and λ_{max} statistics of Johansen (1988) cointegration tests (Table 1), they tend to show a different behavior in the short run. Long-run comovements are due to the fact that these prices are driven by the same underlying factor, which are the conditions prevailing in the world oil markets. Highly correlated assets are possible to generate trading opportunities by simultaneously taking opposite positions in these assets and profit from the divergence or convergence in prices. Divergence of futures prices occurs mainly in the short run and can be attributed to seasonal factors, changes in regional supply/demand and flow of new information. This pattern is also shown in Figures 5 and 6, where the 52-week rolling correlation is plotted. The strength of these short-run comovements is determined by the correlation coefficient of changes in futures prices. High correlation between the legs of the pair in the long run, together with the relatively lower correlation in the short run, are two of the essential indicators one should examine before initiating a spread trading strategy, since these measures provide information on the frequency of divergence from the aforementioned long-run relationship. Another important indicator is the level of volatility of both prices and returns, which determines the degree of divergence from the aforementioned long-run relationship. Normally, we would expect the higher the correlation in the long run and the lower the correlation in the short run to produce the best possible outcome. However, the volatility effect cannot be neglected because the individual relationship between correlation and standard deviation is market specific and indicates that it would be more appropriate to adjust these coefficients in order to get a relative measure, subject to the levels of volatility both in the longer and in the shorter term. This is based on the fact that in spite of the theoretical negative relationship between correlation and volatility (that is the higher the correlation, the lower the volatility), the strength of this relationship, the “steepness” of the slope between these two measures and the existence or not of non-linearities are

FIGURE 5 Time-varying correlation in levels and first differences between the one-month ICE Brent and NYMEX WTI futures markets ($CB_t - CL_t$).

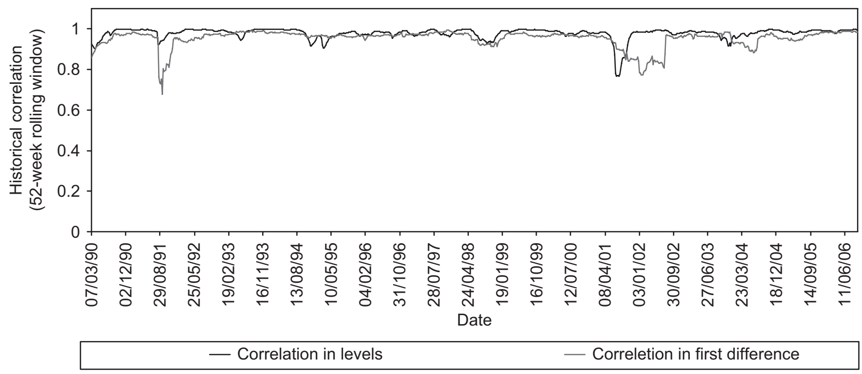
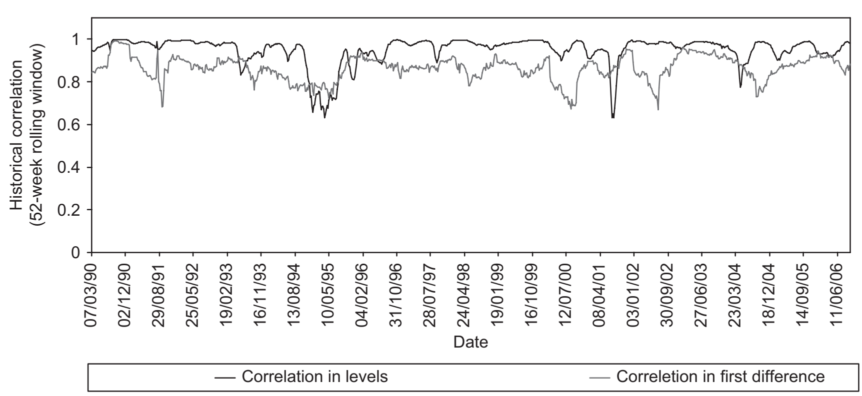


FIGURE 6 Time-varying correlation in levels and first differences between the one-month ICE gas oil and NYMEX heating oil futures markets ($GO_t - HO_t$).



of paramount importance. Hence, long- and short-run correlation justifies pairs trading and long- and short-run volatility includes information on the potential profitability of the spread under study, on the grounds that as the spread deviates from the long-run equilibrium, the probability of convergence is increased along with the probability of achieving higher profits. Table 4 reports the following correlation volatility ratio:

$$CVR = \frac{\rho_{F1,F2} / \sigma_{(F1-F2)}}{\rho_{dF1,dF2} / \sigma_{(dF1-dF2)}} \tag{2}$$

TABLE 4 Long- and short-term unconditional correlations between petroleum futures.

	Three-month strategy (MA12 vs MA1)		Two-month strategy (MA4 vs MA1)		One-month strategy (MA12 vs MA1)		Correlation		Standard Deviation		CVR ratio	CVR rank
	Sharpe ratio	Rank	Sharpe ratio	Rank	Sharpe ratio	Rank	Levels	Returns	Spread	Changes in spread		
Panel A: One month												
$CB_t - GO_t$	1.006	3	1.142	3	1.435	3	0.9907	0.8316	0.417	0.184	0.528	3
$CL_t - HO_t$	0.290	6	0.372	5	0.561	5	0.9911	0.8774	0.398	0.158	0.449	6
$CB_t - CL_t$	0.386	4	0.703	4	1.184	4	0.9983	0.9556	0.210	0.097	0.485	4
$CB_t - HO_t$	0.361	5	0.272	6	0.426	6	0.9911	0.8606	0.409	0.169	0.475	5
$CL_t - GO_t$	1.027	2	1.267	2	1.571	2	0.9890	0.8064	0.444	0.198	0.549	2
$GO_t - HO_t$	1.513	1	1.804	1	1.950	1	0.9972	0.8741	0.224	0.158	0.801	1
Panel B: Two month												
$CB_t - GO_t$	1.202	2	1.332	2	1.800	2	0.9922	0.8529	0.383	0.161	0.488	3
$CL_t - HO_t$	−0.117	6	0.024	6	0.560	5	0.9926	0.9090	0.360	0.125	0.379	6
$CB_t - CL_t$	0.397	5	0.478	5	0.572	4	0.9988	0.9644	0.185	0.081	0.450	4
$CB_t - HO_t$	0.549	4	0.610	4	0.546	6	0.9929	0.8991	0.367	0.134	0.402	5
$CL_t - GO_t$	1.117	3	1.235	3	1.590	3	0.9909	0.8300	0.400	0.171	0.510	2
$GO_t - HO_t$	1.886	1	2.011	1	1.840	1	0.9979	0.8840	0.193	0.141	0.825	1

TABLE 4 Continued.

	Three-month strategy (MA12 vs MA1)		Two-month strategy (MA4 vs MA1)		One-month strategy (MA12 vs MA1)		Correlation		Standard Deviation		CVR ratio	CVR rank
	Sharpe ratio	Rank	Sharpe ratio	Rank	Sharpe ratio	Rank	Levels	Returns	Spread	Changes in spread		
Panel C: Three month												
$CB_t - GO_t$	1.370	2	1.689	2	2.015	2	0.9920	0.8516	0.361	0.149	0.482	3
$CL_t - HO_t$	−0.400	6	−0.283	6	0.523	6	0.9937	0.9230	0.329	0.108	0.352	6
$CB_t - CL_t$	0.465	4	0.438	4	0.589	5	0.9991	0.9635	0.173	0.077	0.460	4
$CB_t - HO_t$	0.164	5	0.317	5	0.885	4	0.9940	0.9112	0.341	0.118	0.378	5
$CL_t - GO_t$	1.246	3	1.426	3	1.779	3	0.9913	0.8344	0.365	0.153	0.499	2
$GO_t - HO_t$	2.106	1	2.242	1	2.160	1	0.9964	0.8749	0.182	0.136	0.851	1
Panel D: Four month												
$CB_t - GO_t$	1.509	2	1.747	2	1.983	2	0.9939	0.8432	0.345	0.148	0.506	3
$CL_t - HO_t$	−0.604	6	−0.358	6	0.230	6	0.9945	0.9232	0.307	0.101	0.356	6
$CB_t - CL_t$	0.501	4	0.641	4	0.789	4	0.9991	0.9565	0.170	0.079	0.486	4
$CB_t - HO_t$	0.082	5	0.157	5	0.617	5	0.9945	0.9102	0.329	0.113	0.375	5
$CL_t - GO_t$	1.355	3	1.513	3	1.770	3	0.9933	0.8294	0.338	0.150	0.531	2
$GO_t - HO_t$	2.009	1	2.090	1	2.129	1	0.9980	0.8750	0.182	0.131	0.795	1

Ranking is carried out by ranking the performance of the spreads under study, according to their historical Sharpe ratios.
The ratio is calculated as in Equation (2).

where $\rho_{F1,F2}$ and $\rho_{dF1,dF2}$ are the loglevel (long run) and logchange (short run) correlations between the legs of the spread, ie, $F1$ and $F2$, respectively. $\sigma_{(F1-F2)}$ is the variance (long run) of a portfolio consisting of F1 and F2 and $\sigma_{(dF1-dF2)}$ is the variance (short run) of changes in the spread series. As it can be seen in Table 4, this ratio is able to distinguish the most profitable petroleum spreads from the unprofitable ones. However, this ratio is only a relative measure that can provide a broad indicator since a spread trading strategy is not considered successful unless transaction costs are covered and the benchmark strategy is outperformed. Consequently, this ratio can be used only in combination with fundamental and technical analyses. For instance, the two spreads of $CB_t - HO_t$ and $GO_t - HO_t$ display similar correlation dynamics both in the long (0.991 vs 0.997) and in the short run (0.861 vs 0.874). However, for the first spread, the bootstrap exercise resulted in annualized returns not more than 7% and 3% under 0.3% and 0.5% transaction costs, respectively. On the other hand, the $GO_t - HO_t$, spread realized returns in excess of 17.5%, in the conservative case of 0.5% transaction costs. This is an example of the effect of volatility on the aforementioned ratio. $CB_t - HO_t$ has significantly greater volatility than $GO_t - HO_t$ in the long run, whereas for the changes in the spread (short run), the volatility ratio is reduced and is not significantly different (at 5% significance level)[11] for the one- and two-month spread, resulting a stronger correlation in the long run for the $GO_t - HO_t$ spread. Time-varying correlation based on a 52-week rolling window indicated that in the case of one-month $CB_t - HO_t$, the minimum value was 0.41, whereas for the $GO_t - HO_t$ spread the same figure is 0.88. This is also evident in Table 1 where cointegration is stronger in the $GO_t - HO_t$ pair, according to λ_{trace} and $\lambda_{\max}$ statistics. On the other hand, in the case of the $CB_t - HO_t$, speed of adjustment coefficients are of less magnitude, indicating lower responsiveness to the differential, and in the three- and four-month case, there is no causality at 5% significance level.

8 CONCLUSIONS

In this study, we tested the performance of MA trading strategies in the petroleum futures markets. In particular, we utilized linkages between different pairs of petroleum futures and devised strategies to identify timing for statistical arbitrage. Error correction models were employed to discover the fundamental drivers of the market and causality direction. Using the cointegration relationship, we further developed a trading strategy that is based on the deviation of oil futures spreads from their long-run mean, combining both the NYMEX and the ICE markets. The turning points are identified using MA trading rules, which are used as indicators for buy and sell opportunities. Simple MA trading strategies proved to be able to

[11] F-tests of equality of variance for the a) spread and b) changes in spread resulted in the following statistics (p-values): a) One-month: 3.327 (0.000), two-month: 3.603 (0.000), three-month: 3.543 (0.000) and four-month: 3.081 (0.000) b) One-month: 1.144 (0.045), two-month: 1.119 (0.093), three-month: 1.315 (0.000) and four-month: 1.357 (0.000).

improve Sharpe ratios and provided evidence that statistical arbitrage opportunities are present in all the combinations of petroleum futures spreads under study, subject to transaction costs. Allowing for variation in transaction costs, maturities of futures prices and MA windows, three out of six spreads proved to be very profitable having realized annualized returns of 18.2–30.0%. Namely, ICE Brent versus gas oil, NYMEX WTI versus ICE gas oil and NYMEX heating oil versus ICE gas oil achieved annualized returns of 17–29%, after accounting for data snooping; this corresponds to Sharpe ratios from 1.00 to 2.28 (3–10 times improvement compared with the buy and hold strategies). Less profitable, yet significant improvement from the employment of technical trading rules, proved to be the one-month intercrude differential.

By and large, our results revealed that the relationship between petroleum futures prices can be used for investment timing, and investors can benefit from technical trading rules, had these rules been based on sound fundamental arguments. The high volatility and high long-run correlation of petroleum futures renders the use of spread trading strategies, where relatively lower-margin requirements and risk exposure along with netting effects can offset the high bid–ask spreads. The rationale behind spread trading is based on the mean-reverting properties of the differential, which are described by the time aggregation of both correlation (which justifies pairs trading) and volatility (which includes information on the potential profitability of the spread). Hence, the long- and short-run correlation and volatility dynamics can provide a good indicator whether statistical arbitrage opportunities are open. Overall, the results indicate that simple MA indicators are able to exploit statistical arbitrage opportunities among futures prices, and by using such indicators, market agents may be able to obtain superior gains, measured in terms of annualized returns and increase in Sharpe ratios.

APPENDIX

Here, we present the algorithm that is used to implement the stationary bootstrap resampling technique of Politis and Romano (1994). The description of the algorithm here follows from Appendix C of Sullivan *et al* (1999).

The stationary bootstrap is calculated as follows: given the original sample of T observations, $X(t), t = \{1, \dots, T\}$, we start by selecting a "smoothing parameter", $q = qT, 0 < qT \leq 1, Tq_T \to \infty$ as $T \to \infty$, and from the bootstrapped series, $X(t)^*$, as follows:

1) At $t = 1$, select $X(1)^*$ at random, independently and uniformly from $\{X(1), \dots, X(t)\}$. Say for instance that X(1)* is selected to be the Jth observation in the original series, $X(1)^* = X(J)$, where $1 \leq J \leq T$.
2) Increment t by 1. If $t > T$, then stop. Otherwise draw a standard uniform random variable U independently of all other random variables.

(a) If $U < q$, then select $X(2)^*$ at random, independently from $\{X(1), \ldots, X(T)\}$.

(b) If $U > q$, then expand the block by setting $X(2)^* = X(J+1)$, so that $X(2)^*$ is the next observation in the original series following $X(J)$. If $J+1 > T$, then reset $J+1$ to 1, so that the block continues from the final observation in the sample.

3) Repeat Step 2 until we reach $X(T)^*$.

4) Repeat Steps 1–3 1,000 times.

Therefore, the stationary bootstrap resamples blocks of varying length from the original data, where the block length follows a geometric distribution, with mean block length 1/q. In general, given that $X(t)^*$ is determined by the Jth observation $X(J)$, in the original series, then $X(t+1)^*$ will be equal to the next observation in the block $X(J+1)$ with probability 1–q and picked at random from the original observations with probability q. Regarding the choice of q, a large value of q is appropriate for data with little dependence, and a smaller value of q is appropriate for data that exhibit serial dependence. The value of q chosen in our experiments is 0.1, corresponding to a mean block length of 10. This follows other studies in the literature, most notably Sullivan *et al* (1999). Furthermore, we also perform sensitivity tests with different values of q and find that the results presented in this section are not sensitive to the choice of q.

REFERENCES

Alizadeh, A. H., and Nomikos, N. (2004). The cost of carry, causality and arbitrage between oil futures and tanker freight markets. *Transportation Research, Part E* **40**, 297–316.

Alizadeh, A. H., and Nomikos, N. (2006). Investment timing and trading strategies in the sale and purchase market for ships. *Transportation Research, Part* B **41**(1), 126–143.

Brock, W., Lakonishok, J., and LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *Journal of Finance* **47**(5), 1731–1764.

Cheung, Y., and Lai, K. (1993). A cointegration test for oil futures market efficiency. *The Journal of Futures Markets* **13**, 933–941.

Crowder, W., and Hamed, A. (1993). A cointegration test for oil futures market efficiency. *The Journal of Futures Markets* **13**, 933–941.

Dunis, C. L., Laws, J., and Evans, B. (2006a). Trading futures spreads: an application of correlation and threshold filters. *Applied Financial Economics* **16**, 903–914.

Dunis, C. L., Laws, J., and Evans, B. (2006b). Trading futures spreads portfolios: application of higher order and recurrent networks. Working Paper, Centre for International Banking, Economics and Finance (CIBEF) and Liverpool John Moores University (JMU), UK.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics* **7**, 1–26.

Emery, G. W., and Liu, Q. (2002). An analysis of the relationship between electricity and natural-gas futures prices. *The Journal of Futures Markets* **22**, 95–122.

Engle, R. F., and Granger, C. W. (1987). Cointegration and error correction: representation, estimation, and testing. *Econometrica* **55**, 251–276.

Ewing, B., and Harter, C. (2000). Co-movements of Alaska North Slope and UK Brent crude oil prices. *Applied Economic Letters* **7**, 555–558.

Girma, P. B., and Paulson, A. S. (1999). Risk arbitrage opportunities in petroleum futures spreads. *The Journal of Futures Markets* **19**, 931–955.

Gonzalo, J. (1994). Five alternative methods of estimating long-run equilibrium relationships. *Econometrica* **60**, 203–233.

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**, 424–438.

Granger, C. (1986). Developments in the study of cointegrated variables. *Oxford Bulletin of Economics and Statistics* **48**, 213–227.

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

Johansen, S. (1988). Statistical analysis of cointegrating vectors. *Journal of Economic Dynamics and Control* **12**, 231–254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models. *Econometrica* **59**, 1551–1558.

Johansen, S., and Juselius, K. (1991). Maximum likelihood estimation and inference on cointegration – with applications to the demand for money. *Oxford Bulletin of Economics and Statistics* **52**, 169–211.

Johnson, R. L., Zulauf, C. R., Irwin, S. H., and Gerlow, M. E. (1991). The soybean complex spread: an examination of market efficiency from the viewpoint of production process. *The Journal of Futures Markets* **1**, 565–596.

Lee, T., and Tse, Y. (1996). Tests with conditional heteroscedasticity. *Journal of Econometrics* **73**, 401–410.

Lin, X. S., and Tamvakis, M. (2001). Spillover effects in energy futures markets. *Energy Economics* **23**, 43–56.

Liu, Q. (2005). Price relations among hog, corn, and soybean meal futures. *The Journal of Futures Markets* **25**, 491–514.

Milonas, N., and Henker, T. (2001). Price spread and convenience yield behaviour in the international crude oil market. *Applied Financial Economics* **11**, 23–26.

Newey, W., and West, K. (1987). A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* **55**, 703–708.

Osterwald-Lenum, M. (1992). A note with the quantiles of the asymptotic distribution of the ML cointegration rank test statistics. *Oxford Bulletin of Economics and Statistics* **54**, 461–472.

Poitras, G., and Teoh, A. (2003). The crack spread: day trading the oil futures complex. *Derivatives Use, Trading & Regulation* **9**, 102–116.

Politis, D. N., and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association* **89**(428), 1303–1314.

Ruiz, E., and Pascual, L. (2002). Bootstrapping financial time series. *Journal of Economic Surveys* **16**, 271–300.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review* **6**, 41–63.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* **6**, 461–464.

Silvapulle, P., and Moosa, I. (1999). The relationship between spot and futures prices: evidence from the crude oil market. *The Journal of Futures Markets* **19**, 175–193.

Sullivan, R., Timmermann, A., and White, H. (1999). Data-snooping, technical trading rule performance, and the bootstrap. *Journal of Finance* **54**(5), 1647–1691.

Wahab M., Cohn, R., and Lashgari, M. (1994). The gold–silver spread: integration, cointegration, predictability, and ex-ante arbitrage. *The Journal of Futures Markets* **14**, 709–756.

White, H. (2000). A reality check for data snooping. *Econometrica* **68**, 1097–1126.