Trading Futures Spreads: An Application of Correlation

by

Ben Evans*
Christian L. Dunis**
Jason Laws***

CIBEF+ and Liverpool John Moores University

October 2004

Abstract

Original motivation for this paper is the investigation of a correlation filter to improve the risk/return performance of the trading models. Further motivation is to extend the trading of futures spreads past the "Fair Value" type of model used by Butterworth and Holmes (2003).

The trading models tested are the following; the cointegration "fair value" approach, MACD, traditional regression techniques and Neural Network Regression. Also shown is the effectiveness of the two types of filter, a standard filter and a correlation filter on the trading rule returns.

Our results show that the best model for trading the WTI-Brent spread is an ARMA model, which proved to be profitable, both in- and out-of-sample. This is shown by out-of-sample annualised returns of 34.94% for the standard and correlation filters alike (inclusive of transactions costs).

Keywords

Efficient market hypothesis (EMH), Cointegration, Trading Filters, Neural Networks, Futures Spreads.

***Ben Evans** is currently working on his PhD thesis at Liverpool John Moores University (Email: B.Evans@livjm.ac.uk).
****Christian Dunis** is professor of Banking and Finance at Liverpool John Moores University and director of CIBEF (Email: cdunis@totalise.co.uk).
*****Jason Laws** is Lecturer and course leader of MSc International Banking and Finance at Liverpool John Moores University (Email: J.Laws@livjm.ac.uk).
+**CIBEF** – Centre for International Banking, Economics and Finance, JMU, John Foster Building, 98 Mount Pleasant, Liverpool, L35UZ

1. Introduction

The prospect of trading spreads should be of interest to technical traders who are limited in the amount they can invest. As stated in Tucker (2000) "spread margins can be as much as 80% lower or more…in some commodities margin requirements are prohibitive. Spreads, however, offer an affordable alternative approach." In addition to this it is important to note that spreads are less likely to suffer from information shocks, as the movement of the two legs will offset each other. A further point of note is that spreads are less likely to be subject to speculative bubbles. Sweeney (1986) notes that speculative bubbles are a big source of market inefficiency. This effect is less likely to happen in spread markets because any bubble effect will be replicated in the opposing leg of the spread (assuming the two legs are sufficiently correlated), the effect of the bubble therefore being largely offset.

This paper will focus on the most liquid of futures markets, that of oil. In particular the WTI-Brent spread. This spread is the difference between two types of crude oil, West Texas Intermediate (WTI) on the long side and Brent Crude (Brent) on the short side; this spread will from here on in be referred to as WTI-Brent.

The two oils concerned differ only in the ability of WTI to produce slightly more Gasoline in the cracking ratio (this is due to lower levels of sulphur and lower gravity, these concepts are further explained in section 3). Gasoline being traditionally more valuable than the other derivatives of the cracking process is the cause of WTI's slight pricing margin over Brent. It is not surprising given the above two points that this spread shows significant signs of a predictable pricing structure.

This paper extends work by Butterworth and Holmes (2003) by expanding on the fair value approach. In this paper we include a fair value model, but compare this to traditional time series models and Neural Network Regressions (henceforth NNR). We also pick up on the point made by Butterworth and Holmes (2003) that "the overall profitability of the strategy is seriously impaired by the difficulty, which traders face, in liquidating their positions" by using both a standard and a correlation filter to further refine the performance of the trading models. The results prove positive with the ARMA time series model significantly out-performing the Fair Value model. Further, in most cases the application of a filter improves the results of the model, in terms of the out-of-sample Sharpe ratio.

The remainder of this paper will be structured as follows; Section 2 will present some of the relevant literature, section 3 gives details of the dataset used, section 4 shows the methodology and gives details of the transaction costs. Section 5 presents the trading rules used and explains how the two filters are applied. Sections 6 and 7 will give the results and conclusions respectively.

2. Literature Review

Spread trading was first introduced formally into the finance literature by Working (1949), who investigated the effects of the cost of storage on pricing relationships. It was demonstrated that futures traders could profit from the existence of abnormalities in the pricing relationships between futures contracts of different expiry.

Meland (1981) gives further justification for interest in spread trading stating that "although spread trading has been used to speculate on the cost of carry between different futures contracts, spread trading also serves the functions of arbitrage and hedging, together with providing a vital source of market liquidity". It is therefore surprising that whilst there has been interest in cash-futures arbitrage[1], inter-commodity spread trading has been largely ignored among the academic fraternity[2].

Spread trading is also of benefit since it increases the amount of investment opportunities, Peterson (1977) and Francis and Wolf (1991) explain this further.

Studies such as Sweeney (1986), Pruitt and White (1988) and Dunis (1989) directly support the use of technical trading rules as a means of trading financial markets. Trading rules such as moving averages, filters and patterns seemed to generate returns above the conventional buy and hold strategy. Lukac and Brorsen (1990) carried out a comprehensive test of futures market trading. It was found that all but one of the trading rules tested generated significantly abnormal returns. Sullivan, Timmerman and White (1998), investigated the performance of technical trading rules over a 100-year period of the Dow Jones Industrial Average, they conclude "there is no evidence that any trading rule outperforms [the benchmark buy and hold strategy] over the sample period."

With the increasing processing power of computers, rule induced trading has become far easier to implement and test. Kaastra and Boyd (1995) investigated the use of Neural Networks for forecasting financial and economic time series. They concluded that the large amount of data needed to develop working forecasting models involved too much trial and error. On the contrary Chen *et al*. (1996), study the 30-year US Treasury bond using a neural network approach. The results prove to be good with an average buy prediction accuracy of 67% and an average annualised return on investment of 17.3%.

In recent years there has been an expansion in the use of computer trading techniques, which has once again called into doubt the efficiency of even very liquid financial markets. Lindemann *et al.* (2004) suggest that it is possible to achieve abnormal returns on the Morgan Stanley High Technology 35 index using a Gaussian mixture neural network trading model. Lindemann *et al.*

[1] See for example Mackinlay and Ramaswamy (1988), Yadav and Pope (1990), and Chung (1991), among others.
[2] Notable exceptions include Billingsley and Chance (1988), Board and Sutcliffe (1996) and Butterworth and Holmes (2003)

(2003) justified the use of the same model to successfully trade the EUR/USD exchange rate, an exchange rate noted for its liquidity.

The paper from which we take our lead, Butterworth and Holmes (2003) states "an analysis of spread trading is important since it contributes to the economics of arbitrage and serves as an alternative to cash-futures arbitrage for testing for futures market efficiency". They test the fair value cointegration model on the FTSE250 – FTSEMID100 spread and conclude "while there are many deviations from fair value, these are generally quite small in actual magnitude, indicating that both contracts tend to be efficiently priced". Their paper forms the platform of this research, however we have extended the number of trading models used, into a full technical trading analysis. We have also included an investigation of the correlation filter, which potentially provides a new methodology for filtering trades on spread markets.

3. Data

The dataset used in this study is daily closing prices from 1995 until 2004 of the WTI, Light[3], Sweet[4] crude oil futures contract and the Brent Crude oil futures contract. Prices have been taken for this period in order to maintain closing time synchronicity. Since November 1994 the International Petroleum exchange (IPE) and New York Mercantile Exchange (NYMEX) have closed at identical times; therefore lending more tradability to the results. These two time series are combined to form the WTI-Brent spread simply by subtracting the price of the Brent contract from the price of the WTI contract of the same expiry. All data was taken from Datastream®.

4. Methodology

The spread returns series is calculated in the following way:

$$R_s = \left[\frac{(WTI_t - WTI_{t-1})}{WTI_{t-1}}\right] - \left[\frac{(Brent_t - Brent_{t-1})}{Brent_{t-1}}\right]$$

Where:
WTI_t = Price of WTI at time t
WTI_{t-1} = Price of WTI at time t-1
$Brent_t$ = Price of Brent at time t
$Brent_{t-1}$ = Price of Brent at time t-1
This convention allows for the calculation of annualised returns and annualised standard deviation to be done in the usual way.

[3] Light refers to the level of "gravity" of the oil. Light oil will have not less than 37° API gravity or more than 42° API gravity. Gravity is an arbitrary scale representing the viscocity and density of the oil (see; www.emis.platts.com/thezone/ guides/platts/**oil**/glossary). API is the "American Petroleum Institute".
[4] Sweet refers to the Sulphur level of the oil. A sweet oil will have not more than 0.42% Sulphur by weight.

For all trading rules apart from the Neural Network Regression the data has been split into two subsections; they are:

Subset	**Purpose**	**Period**
In-Sample	Optimise Model	03/01/1995 – 02/07/2002
Out-of-Sample	Test Model	03/07/2002 – 01/01/2004

The first subset (the in-sample subset) will be used to test all the models and find the optimum of each model type. The second subset (the out-of-sample subset) will be used as a simulation of future prices to trade the optimised models. It is very important that the out-of-sample dataset does not affect our trading decision. For this reason it has been kept as a separate file and only used when the trading model has been decided upon.

In the case of the NNR model, and in order to avoid overfitting, the data will be split into three subsets; they are as follows:

Subset	Purpose	Period
1 (Training)	Optimise model	03/01/1995 – 29/12/1999
2 (Test)	Stop model optimisation	01/01/2000 – 02/07/2002
3 (Validation)	Test model	03/07/2002 – 01/01/2004

The NNR model is trained slightly differently to the other models. The training dataset is used to train the network, the minimisation of the error function being the criteria optimised. The training of the network is stopped when the profit on the test dataset is at a maximum. This model will then be traded on the validation subset, which for comparison purposes is identical to the out-of-sample dataset used for the other models. This technique restricts the amount of noise that the model will fit, whilst also ensuring that the structure inherent in the Training and Test subsets is modelled. Detailed explanation of this is contained in section 5.4.

4.1 Forming the Continuous Series

Trading on futures markets is slightly more complex than trading on cash markets, because futures contracts have limited lifetimes. If a trader takes a position on a futures contract, which subsequently expires, he can take the same position on the next available contract. This is called “rolling forward”. The problem of rolling forward is that two contracts of different expiry may not (and invariably do not) have the same price. When the roll forward technique is applied to the futures time series it will cause the time series to exhibit periodic blips in the price of the futures contract. Whilst the cost of carry (which actually causes the pricing differential) can be mathematically taken out of each contract, this does not leave us with an exactly tradable futures series.

In this study, since we are dealing with futures spreads, we have rolled forward both contracts on the same day of each month (irrespective of exact expiry dates). Since both contracts are on similar underlying the short leg roll forward will cancel out the long leg roll forward[5]. We are left with a tradable time series with no periodic roll forward price blips.

4.2 Transactions Costs

Transaction costs for futures markets are a fraction of those for equity markets. A commission fee of $25 per round trip trade (ie. opening and closing a position on a single contract) has been taken from www.Sucden.co.uk. This is the equivalent of 0.03% of the price of Brent and would allow a non-member to trade the WTI-Brent spread.

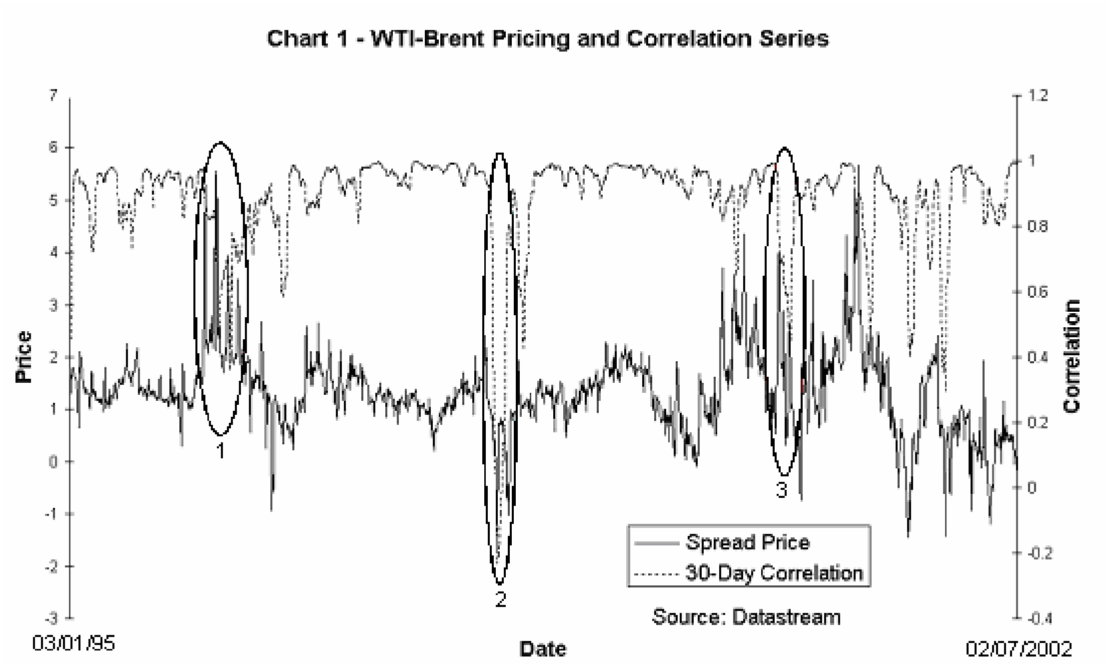
A much bigger consideration when trading spreads is the bid-ask spread, which in the case of spreads has to be covered twice in order to generate profits. The bid-ask spread has been taken as a single percentage of the investment, as in Butterworth and Holmes (2003). A bid-ask spread of 0.17% of the price of the futures contract underlying has been deemed reasonable. Therefore a figure of 0.2%[6] per trade (ie. opening or closing a position on the spread) has been used for the total trading costs.

5. Trading Rules and Time Dependency

The in-sample structure of the spread is mean reverting and this should dictate which trading rules are deemed most appropriate.

Chart 1 below shows the daily closing prices of the WTI-Brent spread for the in-sample period, it is evident that while the spread shows large deviations from the long-term equilibrium, the general pattern is one of mean reversion. Also shown is the rolling 30-day correlation of the spread.

[5] The cost of carry is the difference between the cash and futures price. This is determined by the cost of buying the underlying in the cash market now and holding until futures expiry. Since the cost of storage of both underlying is identical, they will exactly offset each other.
[6] This figure consists of a $0.03 bid ask spread per barrel and a commission fee of $25 per lot (1,000 barrels) both have been taken from www.sucden.co.uk.



It is evident in the above chart that large deviations away from the long term average value are accompanied by large drops in correlation between the two legs of the spread, as shown in the highlighted circles 1, 2 and 3 above. Bearing this consistent feature in mind a correlation filter has been proposed. The idea of which is to filter out any stable pricing periods of the spread. This filter will be optimised in-sample to investigate the impact on annualised percentage returns, annualised standard deviation, maximum drawdown and Sharpe ratio. Various levels of the correlation filter will then be used to see the impact of this on annualised returns. A complete explanation of the correlation filter is included in section 5.5.

The trading rules that are to be tested range from the cointegration "fair value" approach, onto traditional regression analysis, moving averages and finally neural networks. The trading rules and filters used are formally described below.

5.1 Cointegration "Fair Value" Approach

The cointegration "fair value" approach relies on the assumption that there is some underlying value for the spread, which can be considered the long-term equilibrium. Any movement away from this value will constitute a non-equilibrium value and will therefore be short lived. Firstly a Johansen cointegration test is done to find the cointegrating vector of the in-sample dataset. The Johansen cointegration test for the WTI-Brent spread is shown in appendix 1.

The in-sample cointegrating vectors are:

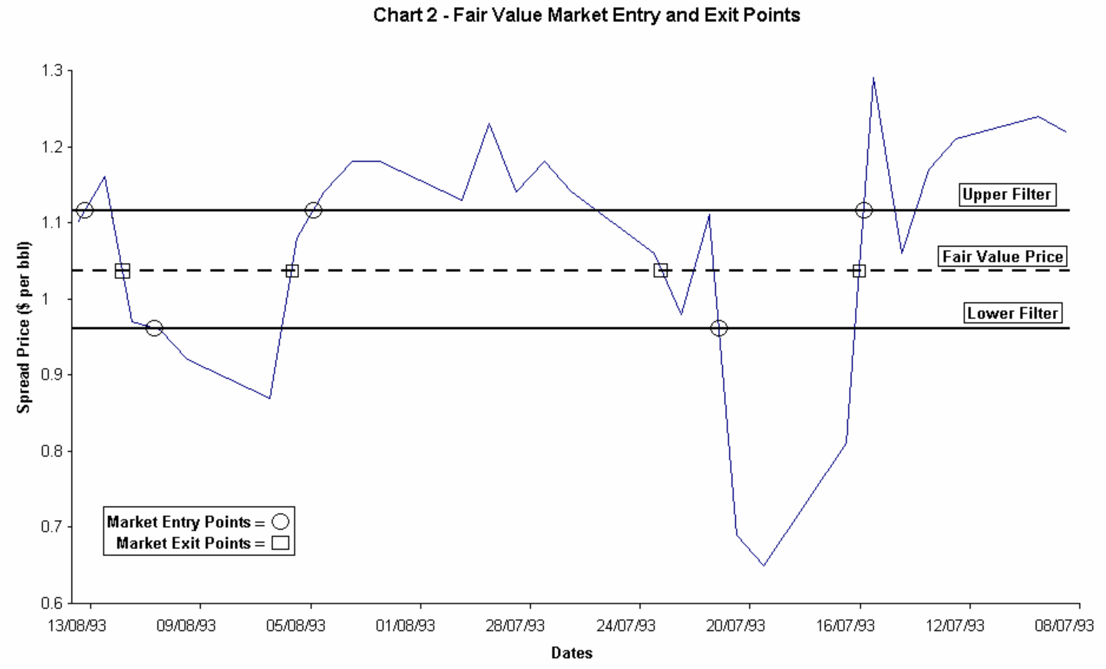
$$(1.00000 \times WTI) - (1.060353 \times Brent)$$

Using these vectors it is fairly simple to calculate the fair value of the spread relative to the price of either leg. This is done in the following way:

$$WTI - Brent = (1.060353 \times Brent) - Brent \text{ , therefore:}$$
$$WTI - Brent = Brent \times (1.060353 - 1)$$

Formally the fair value trading signals can be written as follows:
If WTI<(1.060353)Brent, then go long the spread, until fair value is regained.
If WTI>(1.060353)Brent, then go short the spread, until fair value is regained.

Since we are only interested in spread mispricings that are far enough away from fair value for their return to fair value to net a trading profit, a filter should be employed. This is done in two ways. Firstly a standard filter is used. This dictates that we enter the market when the price of the spread is above (*below*) the fair value plus (*minus*) X, X being the optimised[7] level of the filter. We then exit the market when the spread returns to the fair value price. This is illustrated in chart 2 below:



The second method of filtering is the correlation filter, using this we limit our trading to those periods for which the correlation between the two legs of the spread falls below a certain level. This is further explained in section 5.5.

5.2 Moving Averages

The main problem with the fair value cointegration approach is that the fair value is stationary. This is a problem because any fundamental change in the underlying relationship could cause a massive drawdown resulting in the trader being priced out of the market. A logical extension of this model would be to re-estimate the fair value every day based on the most recent data. Whilst this would be a large undertaking for the cointegration fair value model,

[7] The optimizing parameter for both filters was the net Sharpe Ratio. That is the earnings after costs, divided by the Standard Deviation.

a simpler method is to use an n-day moving average as a proxy for the fair value price.

The “reverse moving average” in which the traditional rule positions are reversed therefore provides the trader with a dynamic model for exploring the situation when markets are not trending but mean-reverting, this rule should also help limit the potential problem of large drawdowns affecting the results.

The formalism for the traditional moving average is given below:

$$MA_t = \sum_{t-n}^{t} \frac{p_{t-n}}{N}$$

The trader should go long if $p_t > MA_t$, and the trader should go short if $p_t < MA_t$
Where:
p_{t-n} price at time $t-n$
$n = (1,2,\ldots,N)$
p_t price at time t
N = number of days of moving average.

The reverse moving average rule will therefore be:
Trader should go long if $p_t < MA_t$, trader should go short if $p_t > MA_t$, with MA_t calculated in the same way.
The standard filter can be applied to this model in the following way:
The trader should go long if $X + p_t > MA_t$, until moving average is regained.
trader should go short if $X - p_t < MA_t$. until moving average is regained.
Where; X = filter level (optimised in-sample).
The correlation filter described in section 5.5 will also be applied to this trading rule and optimised.

5.3 Time Series Analysis

Regression analysis has for a long time been the mainstay of econometric forecasting techniques. With any regression analysis it is normal to start with some tests of normality, ARCH effects and Stationarity. The results of which will impact on the type of model we need to use. The ADF test on the spread levels showed it to be a stationary series and an ARMA(1245678,12367)[8] model was arrived at, from an initial ARMA(10,10), to model the spread. An ARCH test showed some evidence of ARCH effects so a GARCH(2,2) model has been similarly arrived at to model the spread. For complete mathematical explanations of the GARCH and ARMA models see Engle and Bollerslev (1986) and Box, Jenkins and Reinsel (1994) respectively.

The models were optimised in-sample and then forecasted to include the out-of-sample subset. Once the results were obtained they were transferred to Excel where the standard and correlation filters were applied.
The standard filter was applied in the following way:
If $p^e{}_t - X > p^a{}_{t-1}$, then go long the spread.
If $p^e{}_t + X < p^a{}_{t-1}$, then go short the spread.

[8] This is simply an ARMA(8,8) with the 3rd auto regressive term and the 4th 5th and 8th moving average terms removed.

Where; p^e_t = Estimated price for time t (estimated at t-1)
p^a_{t-1} = actual price at time t-1
X = size of filter.
An optimised correlation filter was also applied to this trading model, as described in section 5.5.

5.4 Neural Network Regression

The most basic type of model, which is used in this paper, is the MultiLayer Perceptron (MLP). The network has three layers; they are the input layer (explanatory variables), the output layer (the model estimation of the time series) and the hidden layer. The number of nodes in the hidden layer defines the amount of complexity that the model can fit. The input and hidden layers also include a bias node (similar to the intercept for standard regression), which has a fixed value of 1.
The network processes information as shown below:

1. The input nodes contain the values of the explanatory variables (in this case 10 lagged values of the spread).
2. These values are transmitted to the hidden layer as the weighted sum of it's inputs.
3. The hidden layer passes the information through a nonlinear activation function and, if the calculated value is above the threshold, onto the output layer.

The connections between neurons for a single neuron in the net are shown in chart 3 below:

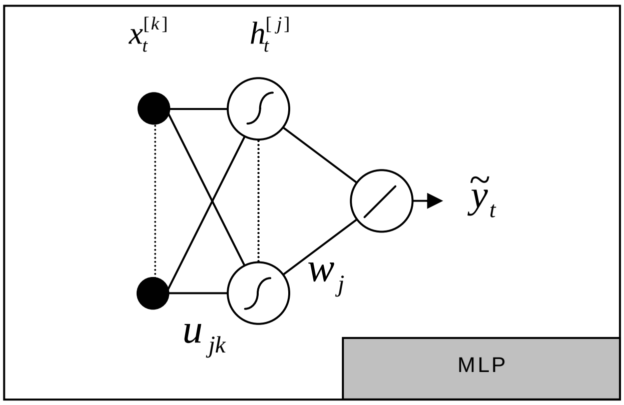


Chart 3: A single output, fully connected MLP model

where:

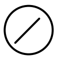
$x_t^{[n]}$ $(n = 1,2,\cdots,k+1)$ are the model inputs (including the input bias node) at time t

$h_t^{[m]}$ $(m = 1,2,...,m+1)$ are the hidden nodes outputs (including the hidden bias node)

$\tilde{y}_t$ is the MLP model output

u_{jk} and w_j are the network weights

 is the transfer sigmoid function: $S(x) = \frac{1}{1+e^{-x}}$,

 is a linear function: $F(x) = \sum_i x_i$

The error function to be minimised is:

$E(u_{jk}, w_j) = \frac{1}{T}\sum (y_t - \tilde{y}_t(u_{jk}, w_j))^2$ with y_t being the target value.

The training of the neural network is of utmost importance, since it is possible for the network to learn the training data subset exactly (commonly referred to as overfitting). For this reason the network training must be stopped early. This is achieved by dividing the dataset into 3 different components (as shown in section 3). Firstly a training subset is used to optimise the model, the "back propagation of errors" algorithm is used to establish optimal weights from the initial random weights.

Secondly a test subset is used to stop the training subset from being overfitted. Optimisation of the training subset is stopped when the test subset is at maximum positive return. These two subsets are the equivalent of the in-sample subset for all other models. This technique will prevent the model from overfitting the data whilst also ensuring that any structure inherent in the spread is captured.

Finally a validation subset is used to simulate future values of the time series, which for comparison is the same as the out-of-sample subset of the other models.

Since the starting point for each network is a set of random weights, we have used a committee of ten networks to arrive at a trading decision (the average change estimate decides on the trading position taken). This helps to overcome the problem of local minima effecting the training procedure. The trading model predicts the change in the spread from one closing price to the next, therefore the average result of all trading models was used as the forecast of the change in the spread.

The standard filter applied to this model was to stay out of the market if the predicted change in the spread was smaller in magnitude than X, X being the optimised filter level. A correlation filter as described in section 5.5 was also applied to this model and optimised.

5.5 The Correlation Filter

As well as the application of a standard filter, we also filtered the spreads in terms of correlation. This is presented as a new methodology to filter trading rule returns on spread markets. The idea is to enable the trader to filter out periods of static spread movement (when the correlation between the underlying legs is increasing) and retain periods of dynamic spread movement (when the correlation of the underlying legs of the spread is decreasing). This was done in the following way.

A rolling Z-day correlation is produced from the two legs of the spread. The Y-day change of this series is then calculated. From this a binary output of either 0 if the change in the correlation is above X, or 1 if the change in the correlation is below X. X being the filter level. This is then multiplied by the returns series of the trading model.

By using this filter it should also be possible to filter out initial moves away from fair value which are generally harder to predict than moves back to fair value. Chart 4 below shows the entry and exit points of the filter.

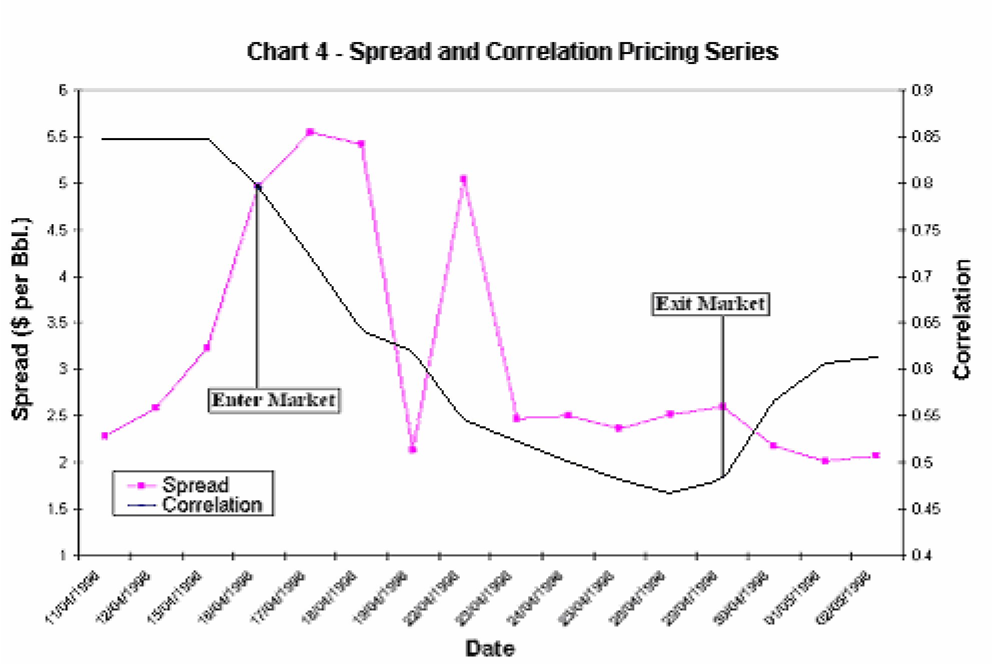


Chart 4 shows that we enter the market the day after the correlation falls and exit the market the day after the correlation rises. Doing this we can not only filter out periods when the spread is stagnant, but also the initial move away from fair value, which is less predictable than the move back to fair value.

There are several optimising parameters, which have been used for this type of filter, namely the length of correlation lag, period of correlation drawdown and amount of correlation drawdown (labelled Z, Y and X respectively).

6. Results

The results for all trading models used are shown in tables 1 to 5 below. The first two rows of the table illustrate the effectiveness of the trading rule without any trading filters in place. The middle two rows of the table show the trading rule performance with the application of a standard filter as described in the trading rule subsections. The final two rows indicate the performance of the trading rule with the application of the correlation filter, as described in section 5.5.

The three right hand columns show the parameters used in optimising the filter. For the trading rule with no filter there are obviously no optimising parameters. For the standard filter there is one optimising parameter, which is the level of the filter. For the correlation filter we have three optimising parameters, which are; the filter level (or amount of correlation change), the period over which the correlation change takes place and the initial correlation series length (labelled X, Y and Z).

Table 1 below shows the results of the “Fair Value” model similar to that used by Butterworth and Holmes (2003). It is evident that the standard filter improves both the in-sample and out-of-sample performance of the trading rule in terms of the Sharpe ratio and the Drawdown.

The correlation filter improves the model but is slightly out performed by the standard filter in terms of out of sample Sharpe ratio. We can however note here a significant fall in the drawdown in the case of the correlation filter.

Table 1 - Fair Value Cointegration Model

							Parameters		
Period	**Filter**	**Returns**	**Drawdown**	**Stdev**	**Sharpe**	**Ann. Trades**	**Level(X)**	**Period(Y)**	**Cor. Length(Z)**
In-Sample	none	31.47%	-12.45%	22.58%	1.39	22.69	-	-	-
Out-of-Sample	none	28.51%	-9.12%	17.53%	1.63	25.93			
In-Sample	Standard	31.42%	-11.97%	22.11%	1.42	21.41	0.12	-	-
Out-of-Sample	Standard	28.66%	-9.12%	17.29%	1.66	28.59			
In-Sample	Correlation	29.17%	-9.69%	17.83%	1.64	22.69	0.0001	2-Days	14-Days
Out-of-Sample	Correlation	20.85%	-8.61%	12.63%	1.65	25.93			

Table 2 below shows the trading statistics of the MACD model. The MACD model finally arrived at was the single 20-day reverse moving average model. This model shows a distinct improvement over the Fair value model, having higher Sharpe ratio and lower out-of-sample drawdown.

With the application of the standard filter the MACD model performance does improve. There are slight improvements in the returns and the standard deviation. Conversely the application of a correlation filter shows no improvement over the un-filtered model indicated by a lower out-of-sample Sharpe ratio and a bigger Drawdown.

Table 2 - MACD Model

							Parameters		
Period	**Filter**	**Returns**	**Drawdown**	**Stdev**	**Sharpe**	**Ann. Trades**	**Level(X)**	**Period(Y)**	**Cor. Length(Z)**
In-Sample	None	40.71%	-13.45%	21.65%	1.88	48.18	-	-	-
Out-of-Sample	None	33.24%	-7.08%	18.05%	1.84	41.75			
In-Sample	Standard	40.77%	-13.45%	21.45%	1.90	51.40	0.02	-	-
Out-of-Sample	Standard	33.79%	-7.12%	18.01%	1.88	42.75			
In-Sample	Correlation	40.76%	-13.45%	21.48%	1.90	52.84	0.049	1-Day	30-Days
Out-of-Sample	Correlation	29.74%	-8.35%	17.91%	1.66	45.77			

Table 3 below shows the summary statistics for the ARMA model, the model in question is an ARMA(1245678,12367), which was arrived at from an ARMA(10,10) model. The use of this model is justified giving returns well in excess of the Fair Value model. The application of the standard filter shows no improvement in any out-of-sample statistics. The correlation filter shows a slight improvement in in-sample trading performance over the un-filtered model, but the effect of the filter is too small to change the out-of-sample performance.

Table 3 - ARMA Model Parameters

Period	Filter	Return	DrawDown	Stdev	Sharpe	Ann. Trades	Level(X)	Period(Y)	Cor. Length(Z)
In-Sample	None	41.63%	-16.17%	21.24%	1.96	88.89	-	-	-
Out-of-Sample	None	34.94%	-9.08%	17.32%	2.02	58.15			
In-Sample	Standard	41.63%	-16.17%	21.24%	1.96	88.89	0	-	-
Out-of-Sample	Standard	34.94%	-9.08%	17.32%	2.02	58.15			
In-Sample	Correlation	41.70%	-13.45%	21.19%	1.97	90.93	0.07	1-Day	30-Days
Out-of-Sample	Correlation	34.94%	-9.08%	17.32%	2.02	58.15			

Table 4 below shows the trading statistics for the GARCH trading model, the model in question being a GARCH(2,2). The use of this model cannot be justified over a Fair Value model since this model exhibits a far worse out-of-sample Sharpe ratio than the Fair Value model. Further the application of the standard filter provides us with no improvement in results. The application of the correlation filter does improve the results slightly but this is still a very poor out-of-sample performance.

Table 4 - GARCH Model **Parameters**

Period	Filter	Return	DrawDown	Stdev	Sharpe	Ann. Trades	Level(X)	Period(Y)	Cor. Length(Z)
In-Sample	None	33.30%	-16.49%	21.30%	1.56	94.09	-	-	-
Out-of-Sample	None	10.56%	-11.33%	17.42%	0.61	82.71	-		
In-Sample	Standard	25.18%	-14.71%	18.18%	1.39	105.10	0.068	-	-
Out-of-Sample	Standard	8.50%	-12.43%	15.25%	0.56	124.06			
In-Sample	Correlation	34.10%	-18.53%	21.27%	1.60	97.25	0.06	1-Day	30-Days
Out-of-Sample	Correlation	10.87%	-12.36%	17.41%	0.62	84.00			

Table 5 below shows the trading statistics of the Neural Network Regression. Surprisingly the Fair Value model outperforms this model, and its use cannot be justified. The plus point of this model is that it does respond well to filters, showing marked improvements in the out-of-sample Sharpe ratio in the case of both filters.

Comparing both filters we can see that the correlation filter shows a significant improvement in out-of-sample performance over the standard filter, with higher returns and Sharpe ratio and lower drawdown and standard deviation. The poor level of return is obviously badly affected by the high number of trades, and therefore trading costs, when compared to any other model.

Table 5 - Neural Network Regression Parameters

Period	Filter	Return	Drawdown	Stdev	Sharpe	Ann. Trades	Level(X)	Period(Y)	Cor. Length(Z)
In-Sample	none	21.32%	-32.58%	20.90%	1.02	154.89	-	-	-
Out-of-Sample	none	22.35%	-29.33%	17.86%	1.25	151.81			
In-Sample	Standard	26.86%	-21.22%	19.43%	1.38	122.53	0.022	-	-
Out-of-Sample	Standard	23.57%	-23.18%	16.70%	1.41	120.83			
In-Sample	Correlation	28.56%	-20.64%	20.13%	1.42	105.47	0.01	1-Day	30-Days
Out-of-Sample	Correlation	26.64%	-26.94%	17.40%	1.53	127.55			

7. Concluding Remarks

It is evident from the above tables that the Fair Value model whilst popular with academics does not capture the structure of the WTI-Brent spread as well as an ARMA model. Saying this, the Fair Value model did perform surprisingly well and was the second best model type tested. A further point of note is that an ARMA model out performed the ANN model, which is regarded by many academics as “state of the art”. The reason for this becomes obvious when the trading costs are considered. The ARMA model has annualised trades of between 58.15 and 90.93. The ANN model on the other hand has annualised trades of between 105.47 and 154.89. This is a significant increase and can be blamed, in part, for the ANN model’s poor showing. This is particularly evident in spreads where two sets of trading costs have to be covered.

In the case of 4 out of 5 of the models tested, the act of filtering improved the results, with the correlation filter being the best filter type in 2 of these 4, vindicating the idea of the correlation filter as an alternative to the standard filter for some trading models.

It is therefore concluded that for this spread, the correlation filter can, in some cases, improve the performance of trading rules over and above the performance of both the un-filtered model and the standard filtered model. Further, it is concluded that the use of the Fair Value cointegration model is vindicated, but that a better performing model is the ARMA model illustrated here.

Finally we conclude that the correlation filter can provide traders with a useful way of improving the performance of spread trading models and subsequently should be studied further, with further research concentrating on the application of a combination, or hybrid, filter.

Appendix 1 – Johansen Cointegration Test

Date: 07/05/04 Time: 15:24
Sample(adjusted): 8 1959
Included observations: 1952 after adjusting endpoints
Trend assumption: No deterministic trend
Series: BRENT WTI
Lags interval (in first differences): 1 to 4

Unrestricted Cointegration Rank Test

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	5 Percent Critical Value	1 Percent Critical Value
None **	0.030592	60.65109	12.53	16.31
At most 1	1.80E-06	0.003509	3.84	6.51

*(**) denotes rejection of the hypothesis at the 5%(1%) level
Trace test indicates 1 cointegrating equation(s) at both 5% and 1% levels

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	5 Percent Critical Value	1 Percent Critical Value
None **	0.030592	60.64758	11.44	15.69
At most 1	1.80E-06	0.003509	3.84	6.51

*(**) denotes rejection of the hypothesis at the 5%(1%) level
Max-eigenvalue test indicates 1 cointegrating equation(s) at both 5% and 1% levels

Unrestricted Cointegrating Coefficients (normalized by b'*S11*b=I):

BRENT	WTI
-1.309067	1.234557
0.034583	0.012068

Unrestricted Adjustment Coefficients (alpha):

D(BRENT)	0.016666	0.000644
D(WTI)	-0.057688	7.88E-05

1 Cointegrating Equation(s): Log likelihood -1981.793

Normalized cointegrating coefficients (std.err. in parentheses)

BRENT	WTI
1.000000	-0.943082
	(0.00436)

Adjustment coefficients (std.err. in parentheses)

D(BRENT)	-0.021816
	(0.01454)
D(WTI)	0.075517
	(0.00980)

References

Billingsley, R. and Chance, D. (1988), the Pricing and Performance of Stock Index Futures Spreads, *Journal of Futures Markets*, 8, 303-318.

Board, J. and Sutcliffe, C. (1996), The Dual Listing of Stock Index Futures: Arbitrage, Spread Arbitrage and Currency Risk, *Journal of Futures Markets*, 16, 29-54.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994), Time Series Analysis: Forecasting and Control, Third Edition, *Englewood Cliffs*, NJ: Prentice Hall, 197-199.

Butterworth, D. and Holmes, P. (2003), Mispricings in Stock Index Futures Contracts: Evidence For FTSE100 and FTSEmid250 Contracts, *Working Papers Series in Economics and Finance,* University of Durham.

Chen, W., Wagner, L. and Lin, C. H. (1996), Forecasting the 30-year U.S. Treasury Bond With a System of Neural Networks, *NeuroVe$t Journal*, Jan/Feb 96.

Chung, Y. P. (1991), A Transactions Data Test of Stock Index Futures Market Efficiency and Index Arbitrage Profitability, *Journal of Finance*, 46, 1791-1809.

Dunis, C. (1989), Computerised Technical Systems and Exchange Rate Movements, Edited by M. Feeny, Exchange Rate Forecasting, *Probus Publishing Company, Cambridge, UK*, 165-205.

Engle, R. and Bollerslev, T. (1986), Modeling Persistence of Conditional Variances, *Economic Review*, 5-50.

Francis, J. C. and Wolf, A. (1991), On The Theory Of Spreading, Edited by A Chen, A Research Annual, Greenwich, Conn. And London: JAI Press, 9, 171-85.

Kaastra, I. and Boyd, M. (1995), Designing a Neural Network for Forecasting Financial and Economic Time Series, *Neurocomputing,* 10, 215-236.

Lindemann, A. Dunis, C. and Lisboa, P. (2004), Probability Distributions and Leveraged Strategies: An Application of Gaussian Mixture Model to the Morgan Stanley High Technology 35 Index, *CIBEF Working Papers*, www.cibef.com.

Lindemann, A. Dunis, C. and Lisboa, P. (2003), Level Estimation, Classification and Probability Distribution Architectures for Trading the EUR/USD Exchange rate, *Neural Computing and Applications,* Forthcoming.

Lukac, L. and Brorsen, B. (1990), A Comprehensive Test of Futures Market Disequilibrium, *The Financial Review*, 25, 4, 593-633.

Mackinlay, C. and Ramaswamy, K. (1988), Index Futures Arbitrage and The Behaviour of Stock Index Futures Prices, *Review of Financial Studies*, 1, 137-158.

Meland, L. (1981), Futures Market Liquidity and The Technique of Spreading, *Journal of Futures Markets,* 1, 405-411.

Peterson, R. (1977), Investor Preferences for Futures Straddles, *Journal of Financial and Quantitative Analysis,* 12, 105-120.

Pruitt, S. W. and White, R. E. (1988), The CRISMA Trading System: Who Says Technical Analysis Can't Beat the Market?, *Journal of Portfolio Management,* 14, 55-58.

Sullivan, R. Timmerman, A. and White, H. (1998), Datasnooping, Technical Trading Rule Performance and the Bootstrap, *Journal of Finance*, 55, 5, 1647-1691

Sweeney, R. J. (1988), Some New Filter Rule Tests: Methods and Results, *Journal of Financial and Quantitative Analysis,* 23, 285-300.

Tucker, S. (2000), Spreading the Wealth, *www.Futuresmag.com,* January 2000.

Working, H. (1949), The Theory of Price of Storage, *American Economic Review,* 39, 1254-1262.

Yadav, P. K. and Pope, P. F. (1990), Stock Index Futures Arbitrage: International Evidence, *Journal of Future Markets,* 10, 573-604.