Original Article

Trading and filtering futures spread portfolios: Further applications of threshold and correlation filters

Received (in revised form): 17th November 2008

Christian L. Dunis

is Professor of Banking and Finance at Liverpool Business School, Liverpool John Moores University, where he also heads the Centre for International Banking, Economics and Finance. He is also a consultant to asset management firms, specialising in the application of non-linear approaches to financial management problems. He has published widely in the field of financial markets analysis and forecasting, and has organised the annual Forecasting Financial Markets Conference since 1994.

Jason Laws

is Reader in Finance and Course Leader of the MSc in International Banking and Finance at Liverpool Business School, Liverpool John Moores University. His research interests are primarily in the area of risk management and how forecasting methods, allied with suitable derivative securities, can generate an optimal hedging strategy.

Ben Evans

is currently Risk Manager Trading for WINGAS Natural Gas Trading in Kassel, Germany. He is also an associate researcher for the Centre for International Banking Economics and Finance. Ben has been published in a wide range of academic journals, including the *European Journal of Finance*, *Derivatives Use Trading*, and *Regulation and Applied Financial Economics*. He continues to work extensively in the use, pricing and trading of commodity derivatives.

Correspondence: Ben Evans GBT – Trading, Friedrich-Ebert-Straße 160, 34119 Kassel, Germany

ABSTRACT This study researches the topic of trading futures spreads, that is, trading the pricing differential between two futures contracts. We trade an equally weighted portfolio of three oil spreads using four trading models: the fair value cointegration, Generalised AutoRegressive Conditional Heteroskedastic, Moving Average Convergence Divergence and Neural Network Regression (NNR) models. The motivation of this research is twofold. First, the profitability of spread markets has been tested further than the traditional fair value model. Second, the correlation filter has been extended by investigating the effect of combining inputs from both a threshold filter and a correlation filter. The results indicate that the best model type for trading spreads is the NNR, with an out-of-sample annualised percentage return of 10.76 per cent and drawdown of –1.52 per cent, resulting in a leverage factor of 7.0964 in the case of the hybrid filter. Further, the results show the hybrid filter to be a sound development, proving to be the best out-of-sample filter on the occasion it is selected.

Journal of Derivatives & Hedge Funds (2010) 15, 274–287. doi:10.1057/jdhf.2009.24

Keywords: cointegration; fair value models; futures spreads trading; neural networks; spreads portfolios; trading filters

This study derives its motivation from those by Butterworth and Holmes¹ and Dunis *et al.*² Butterworth and Holmes¹ investigate the use of a theoretical no-arbitrage model to test for weak form efficiency in the spread between FTSE100 and FTSEMID250. They conclude that 'While the inter-market spread is found to trade within its transactions costs limits on the majority of occasions, large and sustained deviations from fair value exist in both directions, resulting in the triggering of spread arbitrage transactions'. This would indicate that with some effective filtering a profitable arbitrage model could be built.

Dunis *et al.*² investigate the trading of the West Texas Intermediate (WTI)-Brent spread with a similar fair value model; this time the model is based on the cointegration vector of the two underlying series. This fair value model produces an out-of-sample return of 17.46 per cent; however, this is not the only model investigated. Dunis *et al.*² also investigate the Autoregressive Moving Average (ARMA), General Autoregressive Conditional Heteroskedastic (GARCH), Moving Average Convergence Divergence (MACD) and Neural Network Regression (NNR) models. Every model tested produced positive out-of-sample returns inclusive of transactions costs.

This study extends the work of Dunis *et al.*² in two ways. First, we test the same trading rule models on a portfolio of three spreads, thus testing the efficiency of more markets. By looking at the drawdown and standard deviation of these portfolios, we will also be able to draw some conclusions as to the diversifying effect of spread portfolios.

The second way we have extended previous work is by the development of a new filtering technique. Dunis *et al.*² look at threshold and correlation filters, and in this study this is

developed further with the hybrid filter. This filter uses a combination of inputs from both a threshold and a correlation filter to further refine the trading rules that are used. The correlation filter is explained in the section entitled 'The correlation filter', and the hybrid filter is explained in the section entitled 'The hybrid filter'.

The case for spread trading has been made by many academics, most of whom call on reduced margin requirements³⁻⁶ or consistently tradable patterns⁷⁻¹⁰ to encourage interest in spread trading. However, these same studies fail to explicitly explain that with reduced margin comes reduced potential reward. The reason for the reduced margin when trading a spread is the reduced chance of large moves. This is because the two legs of the spread are highly correlated over the long term and will therefore move in generally the same direction.

Figures A1 and A2 show the PDF of the WTI and the Brent series differences, respectively. Figure A3 shows the PDF of the WTI-Brent spread differences. It is evident¹¹ that the average change of the spread is smaller than the average change for either underlying by a significant amount. This validates the decision to offer a lower margin for spreads. In contrast, the kurtosis of the spread PDF is extremely high(11.48), indicating that large moves of the spread are consistent features. Further, the maximum move of the spread is similar to the maximum move of each of the underlying series. This seems to indicate that if spread trading rules were selective enough to pick out these large movements, the potential profitability of the rule would not be that much smaller than that of a rule on a single market, but could still retain the ability of the spread to hedge the position should a large, unfavourable, information-driven move occur on either leg.

BACKGROUND

Spread trading was first formally introduced into the finance literature by Working,¹² 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 among futures contracts of different expiries.

Meland¹³ gives further justification for research into 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 although there has been interest in cash-futures arbitrage,^{14–17} inter- and intra-commodity spread trading has been largely ignored among the academic fraternity.^{1,18–22}

Studies such as those by Sweeney,²³ Pruitt and White,²⁴ and Dunis²⁵ directly support the use of technical trading rules as a means of trading financial markets. Trading rules such as moving averages, filters and patterns seem to generate returns above the conventional buy and hold strategy. Nevertheless, Lukac and Brorsen²⁶ 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 *et al*²⁷ investigated the performance of technical trading rules over a 100-year period of the Dow Jones Industrial Average, they conclude that ‘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 Boyd²⁸

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. In contrast, Chen *et al*²⁹ 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 per cent and an average annualised return on investment of 17.3 per cent.

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. Kaastra and Boyd²⁸ 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*³⁰ justified the use of the same model to successfully trade the EUR/USD exchange rate, an exchange rate noted for its liquidity.

Butterworth and Holmes²¹ state that ‘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 a fair value model on the FTSE250 – FTSEMID100 spread, and conclude that ‘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’. This statement does not indicate whether these deviations from fair value can be captured by either a more sophisticated trading rule or an appropriate filter.

Recently, Dunis *et al*² looked at the trading of the WTI-Brent spread using a variety of techniques, and concluded that it is possible, using only past prices, to generate abnormal

out-of-sample profits. They also investigate the use of filters in improving the Sharpe ratios of the trading rules; the two filters used are a threshold filter based on the size of the predicted market move and a correlation filter based on a rolling correlation. This study further extends this work in two ways: first by testing the profitability of this trading rule on spread portfolios, and second by introducing a further logical development of a correlation filter, namely, a hybrid filter, the exact description of which is given in the section entitled ‘The hybrid filter’.

DATA AND METHODOLOGY

The spread returns series is calculated in the following way:

$$R_s = \left[\frac{(Leg_{1,t} - Leg_{1,t-1})}{Leg_{1,t-1}} \right] - \left[\frac{(Leg_{2,t} - Leg_{2,t-1})}{Leg_{2,t-1}} \right]$$

where $Leg_{1,t}$ is price of Leg_1 at time t ; $Leg_{1,t-1}$ is price of Leg_1 at time $t-1$; $Leg_{2,t}$ is price of Leg_2 at time t ; and $Leg_{2,t-1}$ is price of Leg_2 at time $t-1$.

This convention allows for the calculation of annualised returns and annualised standard

deviation to be carried out in the usual manner.

Data set

For all trading rules, except the NNR rule, the data have been split into two subsections; these are as shown in Table 1.

The first subset (the in-sample subset) is used to test and optimise the models. The second subset (the out-of-sample subset) is used as an unseen data set, to test our optimised models.

In the case of the NNR model, and in order to avoid overfitting, the data will be split into three subsets, as is standard in the literature (for example Lindemann *et al*³⁰ and Kaastra and Boyd²⁸); these are as shown in Table 2.

The NNR model is trained slightly differently from the other models. The training data set is used to train the network, the minimisation of the error function being the criterion optimised. The training of the network is stopped when the profit on the test data set is at a maximum. This model is then traded on the validation subset, which for comparison purposes is identical to

Table 1 In-Sample and Out-of-Sample trading periods

<i>Subset</i>	<i>Purpose</i>	<i>Period</i>	<i>Data points</i>
In-sample	Optimise model	03/01/1995 to 25/07/2002	2175
Out-of-sample	Test model	26/07/2002 to 01/01/2005	435

Table 2 In-Sample and Out-of-Sample trading periods for NNR model

<i>Subset</i>	<i>Purpose</i>	<i>Period</i>	<i>Data points</i>
1 (Training)	Optimise model	03/01/1995 to 29/12/2000	1740
2 (Test)	Stop model optimisation	01/01/2001 to 25/07/2002	435
3 (Validation)	Test model	26/07/2002 to 01/01/2005	435

Table 3 Time Series

<i>Time series</i>	<i>Currency</i>	<i>Fixing times (GMT)</i>	<i>Exchange</i>	<i>Bid-ask spread(%)</i> ³⁸
WTI Cude	US\$	17.30	Nymex	0.0940
Brent Crude	US\$	17.30	IPE	0.0289
Unleaded Gasoline	US\$	17.30	Nymex	0.2008
Heating Oil	US\$	17.30	Nymex	0.2482

the out-of-sample data set used for the other models. This technique restricts the amount of noise that the model will fit, while also ensuring that the structure inherent in the training and test subsets is modelled. Further explanation of this is presented in the section entitled 'Neural Network Regression'.

Table 3 shows details of the time series used to form the portfolio.

The series shown in Table 3 are then combined to form the following spreads:

1. WTI crude versus Brent Crude;
2. WTI crude versus Unleaded Gasoline; and
3. WTI Crude versus Heating Oil.

It is evident from Table 3 that these time series can be traded as spreads, as they are denoted in the same currency and each leg of the spread has an identical fixing time to the opposing leg. The portfolio that has been traded is an equally weighted portfolio of the three spreads shown above, and therefore each spread return is given a one-third weighting for its effect on the portfolio.

Constructing a continuous spread 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 that 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. Although the cost of carry (which actually causes the pricing differential) can be mathematically taken out of each contract, this does not leave us with exactly tradable futures series.

In this study, as we are dealing with futures spreads, we have rolled forward both contracts on the first trading day of the maturity month of the earliest maturing contract. As both contracts are on largely similar underlying, the short leg roll forward will cancel out the long leg roll forward.³¹ We are therefore left with a tradable time series with no periodic roll-forward price blips.

Transactions costs

The bid-ask spread is an average of four intra-day bid-ask spreads, and is presented here as a percentage of the underlying price. The bid-ask spread of each contract, as a percentage of the contract price, can be seen in Table 3. The cost of trading the spread is the total cost of trading

one leg in addition to the total cost of trading the other leg. For example, the cost of trading the WTI-Brent spread is (0.0289 per cent + 0.0940 per cent) = 0.1229 per cent.³²

TRADING DECISION MODELS

The trading decision models have been arranged such that each generic set of trading rules is used to form a portfolio of trading models. There are therefore four trading model portfolios; these are described below.

Fair value cointegration trading rule

The fair value model that has been used in this study is the cointegration fair value model as used by Dunis *et al.*² The model itself is based on the cointegration test,³³ which shows the long-run relationship between multiple assets. This can be extended to a trading model as shown below.

Taking the in-sample cointegrating vectors as $\alpha_{12} + \beta_{12}$ = Cointegrating Vector of Leg_1 with respect to Leg_2 .

Using this vector we can find the residuals of the cointegrating equation. This is achieved in the following way:

$$(\alpha_{12} + (\beta_{12} \times Leg_1)) - Leg_2 = \mu_t$$

where Leg_1 is the long leg of the spread and Leg_2 is the short leg of the spread.

Any deviation of Leg_2 from the theoretical price imposed on it by Leg_1 could be seen as a deviation from fair value. We can therefore present a trading rule as follows:

If $\mu_t < 0$ then go long the spread, until $\mu_t = 0$ is regained.

If $\mu_t > 0$ then go short the spread, until $\mu_t = 0$ is regained.

MACD trading rule

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 regularly re-estimate the fair value based on the most recent data. However, although this would be a computationally intensive undertaking for the fair value cointegration model, a faster 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 wherein 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 as follows:

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

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} is price at time $t-n$; $n = (1, 2, \dots, N)$; p_t is price at time t ; and N is number of days of moving average.

The reverse moving average rule will therefore be that the trader should go long if $p_t < MA_t$ and should go short if $p_t > MA_t$, with MA_t calculated in the same way.

Traditional regression analysis

The third trading decision models to be investigated are traditional regression analysis

models, that is, ARMA and GARCH models. An ARMA(10, 10) model was used to estimate the percentage change in the spread, and a restricted model was estimated using the Akaike information criterion as the optimising parameter (over the in-sample period). Autocorrelation was then tested for and removed with the addition of lags of the percentage change in the spread. If heteroskedasticity was present, an alternative GARCH(1, 1) model was similarly estimated. The final models are those free from heteroskedasticity and autocorrelation, with an optimised Akaike information criterion. These models were then used to estimate the out-of-sample using a 'one size fits all' estimation.

Neural Network Regression

The most basic type of NNR model, which is used in this study, is the MultiLayer Perceptron (MLP). As explained in the study by Lindemann *et al*,³⁴ the network has three layers: 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,³⁵ which has a fixed value of 1.

The network processes information as follows:

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 its inputs.
3. The hidden layer passes the information through a non-linear activation function and, if the calculated value is above a threshold value, onto the output layer.

Figure 1: A single-output, fully connected MLP model.

The connections between neurons for a single output neuron in the net are shown in Figure 1 :

$x_t^{[n]}$ ($n = 1, 2, \dots, k + 1$)	are the model inputs (including the input bias node) at time t ;
$h_t^{[m]}$ ($m = 1, 2, \dots, m + 1$)	are the hidden nodes outputs (including the hidden bias node);
$\tilde{y}_t$	is the MLP model output (the estimate of the change in the spread);
u_{jk} and w_j	are the network weights;
$\int$	is the transfer sigmoid function: $S(x) = 1/(1 + e^{-x})$; and
$\odot$	is a linear function: $F(x) = \sum_i x_i$.

The error function to be minimised is $E(u_{jk}, w_j) = \frac{1}{T} \sum (\gamma_t - \tilde{y}_t(u_{jk}, w_j))^2$, with γ_t (the change in the spread) being the target value.

The training of the neural network is of utmost importance, as 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 data set into three different components (as described at the start of this section). First, 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.

Second, 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, while 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.

As the starting point for each network is a set of random weights, we have used a committee of 10 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 affecting the training procedure. The trading model predicts the change in the spread from one closing price to the next, and therefore the average result of all trading models was used as the forecast of the change in the spread.

FILTER METHODOLOGIES

The filter methodologies employed in this study are as follows.

The threshold filter

The threshold filter is constructed as follows:

If $(\Delta y_t)^2 < x^t$ then stay out of the market; otherwise take the decision of the trading rule,

where Δy_t is the estimate of the percentage change in the spread, and x^t is the size of the threshold filter, which has been optimised in-sample.

In the case of the MACD and fair value models, the position is held until the moving average or fair value is regained.

The correlation filter

In addition to the application of a threshold filter, the trading rules have been filtered in terms of their correlation. This methodology is explained in the study by Dunis *et al*² and is reproduced below.

A rolling 30-day correlation is produced from the two legs of the spread. The change of this series is then calculated. From this there is a binary output of either 0 if the change in the correlation is above X^c or 1 if the change in the correlation is below X^c . X^c being the correlation filter level, which is optimised in-sample. 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. Figure 2 presents an example of the entry and exit points of the filter when $X^c = 0$.

Figure 2 shows that a market entry is triggered the day after the drop in correlation. The market exit is triggered the day after the correlation starts to rise.

The hybrid filter

The hybrid filter has been included here because the threshold and correlation filters seem to filter trades under different circumstances. The hybrid filter is simply a combination of signals from

Figure 2: Operation of the correlation filter.

both filters. The hybrid filter has not been optimised itself, but uses the parameters of the threshold and correlation filters. This filter can be represented as follows.

If either filter shows out of the market signals then stay out of the market; otherwise take the decision of the trading rule.

RESULTS

The results for all four trading techniques can be seen below. Filters are optimised in terms of the in-sample leverage factor.³⁶ This has been used because when trading futures markets, a trader will only need to invest a margin (around 6 per cent of the contract price for the period under review); therefore the underlying investment should be based on a measure of drawdown.

Table 4 shows the trading statistics for the fair value cointegration portfolio. The table covers the percentage of annualised return, the percentage of annualised standard deviation, the percentage of maximum drawdown, the Sharpe ratio and a leverage factor.³⁷

Cointegration fair value portfolio

It is evident from Table 4 that the out-of-sample portfolio does produce returns sufficient to cover transactions costs even without the use of a filter. It is also evident that the use of filtering methodologies is vindicated. From the in-sample results the threshold filter would have been chosen (decision based on highest leverage factor): the out-of-sample results prove this to be the best performer.

MACD portfolio

It is evident from Table 5 that the out-of-sample portfolio does produce returns sufficient to cover transactions costs even without the use of a filter. It is also evident that the use of filtering methodologies is not vindicated. From the in-sample results the unfiltered model would have been chosen (decision based on highest leverage factor); the out-of-sample results also show the threshold filter to subsequently be the best-performing model.

Table 4 Fair-value cointegration portfolio results

<i>Filter</i>	<i>No filter</i>	<i>Threshold</i>	<i>Correlation</i>	<i>Hybrid</i>
<i>In-sample</i>				
Return	10.25%	10.33%	8.56%	8.30%
SD	6.20%	5.21%	5.65%	4.59%
MaxDD	−8.72%	−6.48%	−7.44%	−5.65%
Sharpe	1.6524	1.9836	1.5139	1.8080
Lev. factor	1.1757	1.5941	1.1507	1.4705
<i>Out-of-sample</i>				
Return	6.55%	7.54%	4.30%	5.25%
SD	5.67%	5.36%	5.48%	5.07%
MaxDD	−6.19%	−5.44%	−9.48%	−5.65%
Sharpe	1.1557	1.4054	0.7850	1.0369
Lev. factor	1.0586	1.3849	0.4536	0.9293

Traditional regression analysis portfolio

It is evident from Table 6 that the out-of-sample portfolio does not produce returns sufficient to cover transactions costs without the use of a filter. It is also evident that the use of filtering methodologies is vindicated. From the in-sample results the correlation filter would have been chosen (decision based on highest leverage factor); the out-of-sample results also show the correlation filter to be the best performer.

NNR portfolio

It is evident from Table 7 that the out-of-sample portfolio does produce returns sufficient to cover transactions costs even without the use of a filter. It is also evident that the use of filtering methodologies is vindicated. From the in-sample

Table 5 MACD portfolio results

<i>Filter</i>	<i>None</i>	<i>Threshold</i>	<i>Correlation</i>	<i>Hybrid</i>
<i>In-sample</i>				
Return	9.92%	9.12%	9.82%	9.08%
SD	11.27%	11.16%	11.21%	11.10%
MaxDD	−5.58%	−5.58%	−5.58%	−5.58%
Sharpe	0.8802	0.8176	0.8762	0.8178
Lev. factor	1.7785	1.6353	1.7616	1.6276
<i>Out-of-sample</i>				
Return	11.07%	13.54%	10.81%	13.28%
SD	11.04%	10.98%	11.03%	10.97%
MaxDD	−2.55%	−2.55%	−2.55%	−2.55%
Sharpe	1.0024	1.2338	0.9801	1.2114
Lev. factor	4.3421	5.3126	4.2411	5.2116

results the hybrid filter would have been chosen (decision based on highest leverage factor); the out-of-sample results also show the hybrid filter to be the best-performing model type.

CONCLUSIONS

It can be concluded from these results that the use of the various filtering techniques; the hybrid filter in particular is vindicated. This is evident as the selection decision we make is proven to be incorrect only once. This is the case of the MACD model for which the unfiltered model is chosen based on the in-sample leverage factor; however, the out-of-sample statistics show that a better filter model would have been the threshold filter.

In all other models, the selection of the best in-sample filtering technique leads to the best out-of-sample performance. The hybrid filter

Table 6 Traditional regression analysis portfolio results

Filter	None	Threshold	Correlation	Hybrid
<i>In-sample</i>				
Return	3.24%	3.37%	8.01%	4.03%
SD	10.60%	4.84%	8.72%	6.82%
MaxDD	−40.97%	−11.21%	−18.38%	−24.35%
Sharpe	0.3058	0.6949	0.9197	0.5913
Lev. factor	0.0792	0.3003	0.4360	0.1657
<i>Out-of-sample</i>				
Return	−5.36%	0.39%	3.56%	1.98%
SD	9.44%	5.64%	8.64%	6.92%
MaxDD	−19.73%	−9.67%	−12.89%	−8.97%
Sharpe	−0.5674	0.0687	0.4119	0.2854
Lev. factor	−0.2716	0.0401	0.2761	0.2203

was selected on one of these occasions, and proved to be the best out-of-sample performer.

It can also be concluded that the use of the fair value cointegration model is vindicated returning an out-of-sample leverage factor of 1.38. The best model type for trading spreads, based on this evidence, seems to be the NNR model, which returns an out-of-sample leverage factor of above 7 and an out-of-sample Sharpe ratio of above 1.5. It can therefore be concluded that the various filters shown here, and in particular the hybrid filter, can provide the spread trader with an advantage over plain trading decision models.

Furthermore, drawing on the results of the study by Salcedo,¹⁰ it can be concluded that spread portfolios offer diversification benefits. Salcedo¹⁰ shows a maximum out-of-sample leverage factor of 4.6, for the MACD model; here the out-of-sample leverage factor has been increased to 7.3 – this time for the NNR model.

Table 7 NNR portfolio results

Filter	None	Standard	Correlation	Hybrid
<i>In-sample</i>				
Return	15.97%	11.57%	15.62%	8.89%
SD	9.82%	4.60%	8.14%	3.83%
MaxDD	−26.48%	−3.93%	−17.37%	−2.06%
Sharpe	1.6256	2.5137	1.9187	2.3232
Lev. factor	0.6029	2.9453	0.8993	4.3138
<i>Out-of-sample</i>				
Return	10.71%	13.03%	12.17%	10.76%
SD	11.10%	7.09%	9.82%	6.90%
MaxDD	−23.42%	−2.38%	−20.38%	−1.52%
Sharpe	0.9652	1.8379	1.2390	1.5601
Lev. factor	0.4573	5.4807	0.5970	7.0964

REFERENCES

- 1 Butterworth, D. and Holmes, P. (2002) Inter-market spread trading: Evidence from UK index futures markets. *Applied Financial Economics* 12(11): 783–791.
- 2 Dunis, C., Laws, J. and Evans, B. (2006) Trading futures spreads: An application of correlation and threshold filters. *Applied Financial Economics* 16(12): 903–914.
- 3 See for example Krueger,⁴ Tucker,⁵ and Ross.⁶
- 4 Krueger, D. (2000) Spread trading 101, <http://www.FuturesMag.com>, October 2000.
- 5 Tucker, S. (2000) Spreading the wealth, <http://www.FuturesMag.com>, January 2000.
- 6 Ross, J. (2003) Getting an edge with reduced margin spreads, <http://www.FuturesMag.com>, Special Issue: Secrets of the Masters 2003.
- 7 See for example Salcedo,⁸ Girma and Paulson,⁹ and Salcedo.¹⁰
- 8 Salcedo, Y. (2004b) Spreads for the fall, <http://www.FuturesMag.com>, September 2004.
- 9 Girma, P.B. and Paulson, S.A. (1998) Seasonality in petroleum futures spreads. *The Journal of Futures Markets* 18(5): 581–598.
- 10 Salcedo, Y. (2004a) Making the most of Ag spreads, <http://www.FuturesMag.com>, June 2004.
- 11 See Figures A1–A3, Max, Min and Std. Dev.
- 12 Working, H. (1949) The theory of price of storage. *American Economic Review* 39: 1254–1262.

- 13 Meland, L. (1981) Futures market liquidity and the technique of spreading. *Journal of Futures Markets* 1: 405–411.
- 14 See for example Mackinlay and Ramaswamy,¹⁵ Yadav and Pope,¹⁶ and Chung,¹⁷ among others.
- 15 Mackinlay, C. and Ramaswamy, K. (1988) Index futures arbitrage and the behaviour of stock index futures prices. *Review of Financial Studies* 1: 137–158.
- 16 Yadav, P.K. and Pope, P.F. (1990) Stock index futures arbitrage: International evidence. *Journal of Future Markets* 10: 573–604.
- 17 Chung, Y.P. (1991) A transactions data test of stock index futures market efficiency and index arbitrage profitability. *Journal of Finance* 46: 1791–1809.
- 18 Notable exceptions include studies by Butterworth and Holmes,¹⁸ Billingsley and Chance,¹⁹ Board and Sutcliffe,²⁰ Butterworth and Holmes,²¹ and Kawaller *et al.*²²
- 19 Billingsley, R. and Chance, D. (1988) The pricing and performance of stock index futures spreads. *Journal of Futures Markets* 8: 303–318.
- 20 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.
- 21 Butterworth, D. and Holmes, P. (2003) Mispricings in Stock Index Futures Contracts: Evidence For FTSE100 and FTSEmid250 Contracts. University of Durham. Working Papers Series in Economics and Finance.
- 22 Kawaller, I.G., Koch, P.D. and Ludan, L. (2002) Calendar spreads, outright futures positions and risk. *Journal of Alternative Investments* (Winter), pp. 59–74.
- 23 Sweeney, R.J. (1988) Some new filter rule tests: Methods and results. *Journal of Financial and Quantitative Analysis* 23: 285–300.
- 24 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.
- 25 Dunis, C. (1989) Computerised technical systems and exchange rate movements. In: M. Feeny (ed.) *Exchange Rate Forecasting*. Cambridge, UK: Probus Publishing Company, pp. 165–205.
- 26 Lukac, L. and Brorsen, B. (1990) A comprehensive test of futures market disequilibrium. *The Financial Review* 25(4): 593–633.
- 27 Sullivan, R., Timmerman, A. and White, H. (1998) Datasnooping, technical trading rule performance and the bootstrap. *Journal of Finance* 55(5): 1647–1691.
- 28 Kaastra, I. and Boyd, M. (1996) Designing a neural network for forecasting financial and economic time series. *Neurocomputing* 10: 215–236.
- 29 Chen, W., Wagner, L. and Lin, C.H. (1996) Forecasting the 30-year US treasury bond with a system of neural networks. *NeuroVeSt Journal* (Jan/Feb).
- 30 Lindemann, A., Dunis, C. and Lisboa, P. (2005b) Level estimation, classification and probability distribution architectures for trading the EUR/USD exchange rate. *Neural Computing and Applications* 14(3): 256–271.
- 31 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 storing this until the futures contract expires. As the costs of storage of both underlyings are very similar, they will offset each other to a large extent.
- 32 It has been decided that for ease of calculation, any round turn commission (~ 0.03 per cent) should be ignored.
- 33 Johansen, S. (1988) Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12: 231–254.
- 34 Lindemann, A., Dunis, C. and Lisboa, P. (2005a) Probability distributions and leveraged strategies: An application of Gaussian mixture models to the Morgan Stanley technology index tracking fund. *Quantitative Finance* 5: 459–474.
- 35 Similar to the intercept for standard regression.
- 36 The leverage factor is calculated as the annualised return divided by the drawdown. This gives some idea of the amount of return per unit risk.
- 37 All net of transactions costs, the formulation of which is described in “Trading Decision Models”.
- 38 Source of bid-ask spreads: <http://www.sucden.co.uk>.

APPENDIX

See Figure A1–A3.

Figure A1: WTI probability distribution function.

Figure A2: Brent probability distribution function.

Figure A3: WTI – Brent probability distribution function.