COINTEGRATION AND ASSET ALLOCATION: A NEW ACTIVE HEDGE FUND STRATEGY

Carol Alexander, Ian Giblin and Wayne Weddington III

ABSTRACT

Models that are based on mean-variance analysis seek portfolio weights to minimise the variance of the portfolio for a given level of return. The portfolio variance is measured using a covariance matrix that represents the volatility and correlation of asset returns. However these matrices are notoriously difficult to estimate and ad hoc *methods often need to be applied to limit or smooth the mean-variance efficient allocations that are recommended by the model. Moreover the mean-variance criterion has nothing to ensure that tracking errors are stationary. Although the portfolios will be efficient, the tracking errors will in all probability be random walks. Therefore the replicating portfolio can drift very far from the benchmark unless it is frequently re-balanced.*

A significant difference between traditional hedge fund strategies and the model presented in this paper is that portfolio optimization is based upon the cointegration of asset prices rather than the correlation of asset returns. We show that it is possible to devise allocations that always have stationary tracking errors. Moreover, efficient market neutral long short hedge strategies may be achieved with relatively few stocks and with much lower turnover rates compared to traditional market neutral strategies.

Financial Risk and Financial Risk Management, Volume 16, pages 65–89.

ISBN: 0-7623-0858-3

1. INTRODUCTION

Investors recognize that traditional investment vehicles have become increasingly risky, and the managers of traditional long only investment styles are handicapped in that they can take advantage of only one side of the investment process. If they do not like an asset, then they simply do not buy it, whereas the hedge fund manager can actively sell those assets he does not like, presenting a "double alpha" opportunity. No matter how well the traditional manager selects his portfolio, he will still have significant market exposure. Beta-neutrality, dollar-neutrality, and sector-neutrality are all intended to make the hedged portfolio more predictable by eliminating systemic risk.

An unfortunate consequence of hedge funds' present popularity is that "hedge fund" can refer to many different investment styles. Traditionally, a hedge was meant to define a market opportunity in which the risk of the overall market was eliminated from the source of return. But it has become a term used to describe any fund that is not conventional, including long-only bonds, long-only equities (mutual funds), or long-only money markets. Some consider all absolute-return funds to fall within the hedge fund definition, even if these funds do not typically sell short.[1] Many funds are not hedged at all but use leverage to multiply their active market exposure. The returns can be high, but so can the losses, as leveraged directional investments (that are not hedged) tend to have a large impact on performance. It is noteworthy that most hedge fund databases still restrict their searches to absolute return.

The fundamental hedge fund proposition is that pure alpha depends upon dynamic asset selection rather than market direction. The hedge manager's expanded tool set provides greater opportunity to exercise relative value assessments: the manager can fully utilize a stock selection model and, most importantly, focus upon relative valuation rather than absolute valuation. The latter can be particularly profitable when 'all' stocks are over-valued, as the S&P100 stocks became in the quarter ending March 2000. Three types of hedge fund may be distinguished:

1. Directional

Directional hedge funds buy undervalued securities and sell short over valued securities for the purpose of making a profit. Profits are realised on long positions in undervalued assets by selling the asset at a price higher than their buying price; on short positions in overvalued assets profits are made by buying them back at a price lower than their sell price. If leverage is used the speculator is exposed to risk that is greater than the underlying security.

2. Arbitrage

Arbitrage funds operate by the simultaneous buying and selling of a security at two different prices in two different markets, resulting in profits without risk. Perfectly efficient markets present no arbitrage opportunities.

3. Classic Hedge

These funds aim to minimize the market risk by offsetting otherwise risky positions. They are non-directional and often maintain long and short positions for 'dollar' or 'market' neutrality.

Most hedge fund categories did well last year. The SP500-stock index was down 9.1%, the Nasdaq down 39%, while the average U.S. hedge fund was up 4.85% according to the CSFB/Tremont Hedge Fund Index. It is precisely these periods – when volatility is high or in bear markets – that hedge funds become 'hot.' It is amusing that hedge funds are relatively unpopular in strong equity markets, only to regain popularity when the market 'corrects.' Of course, hedge funds may not outperform the market during euphoric upswings. In their simplest market-neutral form, equity hedge funds only seek to provide a return greater than the short-term risk free rate that is, technically, their benchmark. When the markets are down, however, the performances of hedge funds are notable.

The truth is that hedge funds have consistently outperformed the U.S. mutual funds on a risk-adjusted basis (as opposed to absolute return). As investors become increasingly sophisticated regarding the value-at-risk of their investments, hedge funds cannot help but benefit. While the public's fascination with hedge funds has been on-again, off-again depending upon market conditions, each "on" cycle brings incrementally higher interest and higher capital allocations.

This paper presents a classic hedge fund strategy: an investment vehicle whose key objective is to minimize investment risk in an attempt to deliver profits under all market circumstances. Such a strategy may be described as "non-directional" or "market-neutral": the hedge is designed to have minimal correlation with the market and, irrespective of market direction, the fund seeks to generate positive alpha. Market-neutral hedge strategies are presently in great demand, though the true condition of market-neutrality remains elusive to even the most well known hedge funds. Certain high profile market-neutral managers were very market-neutral . . . until they weren't. And in some cases, they were shown to be not-so-market-neutral in spectacular fashion. But, as a group, market neutral managers have done well in this era of high volatility: equity market-neutral funds in the U.S. returned 14.99% in 2000.

The outline of the paper is as follows: Section 2 introduces the concept of cointegrated time series and explains the relationship between cointegrated prices and correlated returns with examples from different types of financial markets. Section 3 describes the methodology for tracking a benchmark using a basket of assets with a price that is cointegrated with the benchmark. Section 4 expands this methodology to a long-short hedge strategy that is based on cointegration. Section 5 reports the results of backtesting such a model for a long-short equity hedge of the S&P100 index, and Section 6 summarizes and concludes.

2. COINTEGRATION

The strategy that is described in this paper uses historical price patterns to project the future performance of a stock. A significant difference between this model and more traditional hedge fund strategies is that portfolio optimization is based upon the cointegration of prices rather than the correlation of returns. We shall see that this affects the character of portfolio performance: In particular turnover will be relatively low for a market-neutral strategy (approximately 2% per trading day) and each leg of the hedge will be independently convex to the market.[2]

Financial markets by their very nature are highly co-dependent. It is, however, unfortunate that many market practitioners still base their analysis of the relationships between assets on the very limited concept of correlation. Trying to model the complex inter-dependencies between financial assets with so restrictive a tool is like trying to surf the internet with an IBM AT. Therefore it is gratifying that more sophisticated dynamic models based on multivariate time series analysis are now being applied to analyze the complex relationships between financial assets.

Cointegration refers not to co-movements in returns, but to co-movements in asset prices, exchange rates or yields. If spreads are mean-reverting, asset prices are tied together in the long-term by a common stochastic trend, and we say that the prices are ‘cointegrated’. Since the seminal work of Engle and Granger (1987) cointegration has become the prevalent tool of time series econometrics.[3] Cointegration has emerged as a powerful technique for investigating common trends in multivariate time series, and provides a sound methodology for modelling both long-run and short-run dynamics in a system.

Cointegration is a two step process: first any long-run equilibrium relationships between prices are established and then a dynamic correlation model of returns is estimated. This error correction model (ECM), so-called because short-term deviations from equilibrium are corrected, reveals the Granger causalities that must be present in a cointegrated system. Thus cointegration may be a sign of market inefficiency, but it can also be the result of market efficiency, as for example is the cointegration between spot and futures prices.

Although empirical models of cointegrated financial time series are common place in the academic literature, the practical implementation of these models into systems for investment analysis or portfolio risk is still in its early stages. This is because the traditional starting point for portfolio allocation and risk management is a correlation analysis of returns. In standard risk-return models the price data are differenced before the analysis is even begun, and differencing removes a-priori any long-term trends in the data. Of course these trends are implicit in the returns data, but any decision based on common trends in the price data is excluded in standard risk-return modelling. The aim of cointegration analysis, on the other hand, is to detect any common stochastic trends in the price data, and to use these common trends for a dynamic analysis of correlation in returns.

Correlation is based only on return data, which are stationary, denoted I(0). Cointegration is based on the raw price, rate or yield data as well as the return data. Price, rate or yield data are not normally stationary, in fact they are usually a random walk, or at least integrated of order 1, denoted I(1).[4] Since it is normally the case that log prices will be cointegrated when the actual prices are cointegrated it is standard, but not necessary, to perform the cointegration analysis on log prices. A set of I(1) series are termed 'cointegrated' if there is a linear combination of these series that is stationary. In the case of just two integrated series:

$$x \text{ and } y \text{ are cointegrated if } x, y \sim I(1) \text{ but there exists } \alpha$$

$$\text{such that } z = x - \alpha y \sim I(0).$$

The definition of cointegration given in Engle and Granger (1987) is far more general than this, but the basic definition presented here is sufficient for the purposes of this paper.

Cointegration and Correlation

Cointegration and correlation are related, but different concepts. High correlation does not imply high cointegration, and neither does high cointegration imply high correlation. In fact cointegrated series can have correlations that are quite low at times. For example a large and diversified portfolio of stocks in an equity index, where allocations are determined by their weights in the index, should be cointegrated with the index. Although the portfolio should move in line with the index in the long-term, there will be periods when stocks that are not in the portfolio have extreme price movements. Following this the empirical correlations between the portfolio and the index may be quite low for a time.

Conversely, high correlation of returns does not necessarily imply high cointegration in prices or rates either. An example is given in Fig. 1, with 8 years of daily data on U.S. dollar spot exchange rates of the German Mark (DEM) and the Dutch Guilder (NLG) from 1986 to 1992. Their returns are very highly correlated, in fact the unconditional correlation coefficient over the whole period is 0.9642. The rates themselves also appear to be moving together. The spread is very stable indeed and in fact they appear to be cointegrated, which is highly unusual for two exchange rates (Alexander & Johnston, 1992, 1994).

Now suppose that a very small incremental stochastic return is added to the spread, to create the NLG 'plus' series that is also shown in Fig. 1. The NLG 'plus' is clearly not cointegrated with DEM. They are not tied together by a stationary spread, in fact they are diverging. However the correlation between the returns to NLG 'plus' and the DEM is virtually unchanged, at 0.9620.

Thus high correlations may occur when there is cointegration, or when there is no cointegration. That is, correlation tells us nothing about the long-term

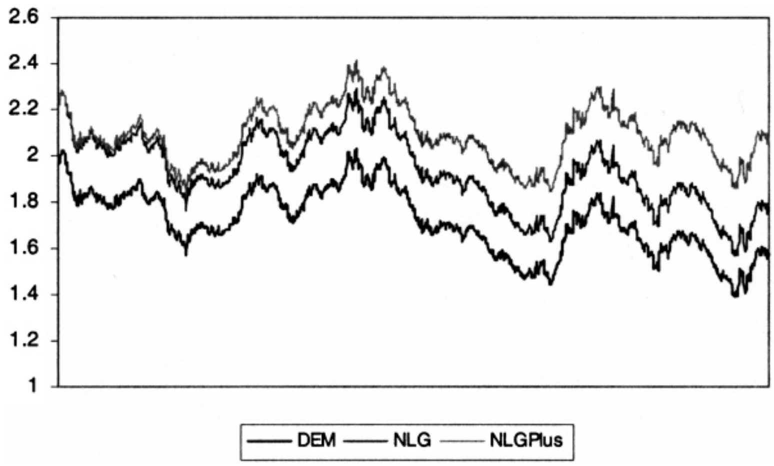

Fig. 1. DEM and NLG Daily FX Rates, Jan 1986–Dec 1992.

behaviour between two markets: they may or may not be moving together over long periods of time, and correlation is not an adequate tool for measuring this.

Correlation reflects co-movements in returns, which are liable to great instabilities over time. In fact, correlation is intrinsically a short-run measure because it is based on returns, which are short memory processes. Therefore classic hedge fund strategies that are based only correlations cannot guarantee long term performance. The model has no mechanism to ensure the reversion of the hedge to the underlying, there is nothing to prevent the tracking error from behaving in the unpredictable manner of a random walk and, consequently, correlation based hedge strategies will normally require frequent re-balancing. To avoid the erosion of profits by high transaction costs, sometimes there is limited re-balancing in direction only, or the optimal allocations are smoothed over time, or Bayesian methods are used to impose priors on the covariance matrix. Each of these strategies has the effect of increasing the fund volatility in a more or less *ad hoc* fashion.

Since high correlation is not sufficient to ensure the long-term performance of hedges, there is a need to augment standard risk-return modelling methodologies to take account of common long-term trends in prices. This is exactly what cointegration provides. Cointegration measures long-run co-movements in prices, which may occur even through periods when static correlations appear low. Therefore hedging methodologies based on cointegrated financial assets should be more effective in the long term. Moreover, the cointegration methodology loses none of the traditional methods. It merely augments the basic correlation model to include a preliminary stage in which the multivariate price data are analyzed and then extends the correlation model to include a dynamic analysis of the lead-lag behaviour between returns.

Common Trends and Long-Run Equilibria

When asset price time series are random walks, over a period of time they may have wandered virtually anywhere, because a random walk has infinite unconditional variance. There is little point in modelling them individually since the best forecast of any future value is the just value today plus the drift. However when two or more asset prices are cointegrated a multivariate model will be worthwhile because it reveals information about the long-run equilibrium in the system. For example, if a spread is found to be mean-reverting we know that, wherever one series is in several years time, the other series will be right there along with it.

Cointegrated log asset prices have a common stochastic trend (Stock & Watson, 1988). They are 'tied together' in the long-run even though they might drift apart in the short-run because the spread or some other linear combination is mean-reverting. A simple example[5] that illustrates why cointegrated series have a common stochastic trend is

$$\begin{aligned} x_t &= w_t + \varepsilon_{xt} \\ y_t &= w_t + \varepsilon_{yt} \\ w_t &= w_{t-1} + \varepsilon_t \end{aligned} \tag{1}$$

where all the errors are i.i.d. and independent of each other. In the example (1) the I(1) variables x and y are cointegrated because $x - y \sim I(0)$. They also have a common stochastic trend given by the random walk component w. Note that the correlation between Δx and Δy will be less than 1 and when the variances of ε_{xt} and/or ε_{yt} are large the correlation can be quite low.[6] Thus, as already mentioned above, cointegration does not imply high correlation.

The linear combination of I(1) variables that is stationary is denoted z. It is called the 'disequilibrium term' because it captures deviations from the long run equilibrium in the error correction model (for further details see chapter 12 of Alexander, 2001). The expectation of z gives the long-run equilibrium relationship between x and y and short-term periods of disequilibrium occur as the observed value of z varies around its expected value. The 'cointegrating vector' is the vector of weights in z. In the case of two I(1) variables x and y where $x - \alpha y \sim I(0)$ the cointegrating vector is $(1, -\alpha)$. When only two integrated series are considered for cointegration, there can be at most one cointegrating vector, because if there were two cointegrating vectors the original series would have to be stationary.

More generally cointegration exists between n integrated series if there exists at least one cointegrating vector, that is at least one linear combination of the I(1) series that is stationary. Each stationary linear combination acts like 'glue' in the system, and so the more cointegrating vectors found the more the coherence and co-movements in the prices. Cointegration can be thought of as a form of factor analysis similar to principal component analysis[7] so it is not surprising that cointegration analysis often works very well on the futures or interest rate term structures that are so successfully modelled by a principal component analysis. Most yield curves have very high cointegration. Often each of the $n - 1$ independent spreads is mean reverting, so there are $n - 1$ cointegrating vectors, the maximum possible number.

Cointegration in Financial Markets

Cointegration has been the subject of extensive research in many financial markets: within term structures, between spot and futures prices, and between international equity and bond market indices.[8] Research on cointegration of stock prices has been more limited, although findings of cointegration were recorded at least a decade ago (Cherchi & Havenner, 1988; Pindyck & Rothenberg, 1992).

This paper concerns the cointegration between a basket of stocks and the stock index. Since the index is, by definition, a linear combination of the constituents, there should be some basket of stocks that is cointegrated with the index. Assuming that the basket is sufficiently large and the index weights do not change too much over time, the tracking error will be stationary, that is, the basket will cointegrated with the index. What is, perhaps, surprising, is that a cointegrating basket can normally be found that contains relatively few stocks. Therefore transactions costs can be substantially reduced when this method is applied to track an index.

3. THE METHODOLOGY

When portfolios are constructed on the basis of mean-variance returns analysis, frequent re-balancing is usually necessary to keep the portfolio in line with the index. One of the reasons for this is that the portfolio variance will normally be measured using a covariance matrix, but these matrices are notoriously difficult to estimate. Unless they are based on very long term averages, which will not respond to current market conditions by definition, they often lack robustness. When the covariance matrix changes considerably from day to day, so will the efficient frontier and the corresponding recommendation for the optimal portfolio. Moreover the mean-variance criterion has nothing to ensure that tracking errors are stationary: indeed the tracking errors will in all probability be random walks. Therefore the replicating portfolio can drift arbitrarily far from the benchmark unless it is frequently re-balanced, and *ad hoc* adjustments are often employed to avoid onerous transactions costs.

The previous section has explained why the cointegration methodology can form the basis of a very powerful tool for investment analysis. When the allocations in a portfolio are designed so that the portfolio is cointegrated with the index, it will track the index over the long term. The portfolio and the index will deviate, but only in the short term, and over the longer term they will be tied together. This property, combined with the fact that cointegrating portfolios can often be formed using relatively few stocks, leads to the construction of optimal portfolios that have less risk, less turnover and lower transaction costs than the traditional mean-variance optimal portfolios.

Selection and Allocation

The criteria that are used in cointegration analysis are to maximise the stationarity and to minimise the variance of the tracking error. Thus, using cointegration it is possible to devise optimal portfolios that are tied to the benchmark and also have minimum risk tracking errors. A linear regression of log prices is employed: the dependent variable will be the log index price or some other benchmark, such as LIBOR, that is used to evaluate the performance of the portfolio;[9] the explanatory variables will be the log prices of the assets in the tracking portfolio; and the residuals are the tracking errors.[10]

There are two parts to the problem: first select the assets, and then optimise the portfolio weights. The asset selection process is perhaps the hardest but most important part, and can be approached in a number of ways. Selection methods range from a 'brute force' approach, such as when the number of assets is fixed and then linear models are fitted for all possible portfolios with this number of assets, to methods that are tailored to investors preferences over various types of stocks, or proprietory technical analysis.

The optimal allocation process uses least squares regression analysis: allocations are made according to a cointegrating regression, so that the fitted portfolio will be cointegrated with the benchmark and the tracking error will be stationary. Suppose a benchmark with log price index y is to be tracked with a number of assets with log prices $x_1, \dots, x_n$. The Engle-Granger cointegration method is to regress y on a constant and $x_1, \dots, x_n$, and then to test the residuals for stationarity.[11] The coefficients $\alpha_1, \dots, \alpha_n$ in the Engle-Granger regression

$$y_1 = \alpha_0 + \alpha_1 x_{1,t} + \dots + \alpha_n x_{n,t} + \varepsilon_t \qquad (2)$$

are normalized to sum to one, thereby giving the portfolio weights. Thus the problem of finding the optimal replicating portfolio can be solved by finding the best assets with log prices $x_1, \dots, x_n$ to use in the cointegrating regression, and then defining allocations to give the maximum stationarity in the tracking error ε. The more stationary the tracking error, the greater the cointegration between the benchmark and the candidate portfolio. In practice, a very high degree of cointegration can be found between the benchmark and the tracking portfolio, so the standard augmented Dickey-Fuller (ADF) unit root test will be sufficient to compare different portfolio specifications and choose those with the most stationary tracking errors.

When there are a large number of potential assets that could be used in a replicating portfolio it is not at all a trivial problem to test all possible portfolios to find the one that has the most stationary tracking error. If there are N assets in total one has to test $N!/n!(N - n)!$ portfolios for every n less than or equal to N.

This strategy can be extended for global asset management models where the benchmark may be a global index such as the Morgan Stanley World Index. In this case there will be two-stages to the selection – allocation process. First select the country indices to track the global index and assign optimal country allocations, and then either buy/sell the country index futures (if available) or repeat the process for tracking the individual country indices with individual stocks. A single country model could also be approached in two stages: first select the industrial sectors and assign weights optimally, then select the stocks within each industry sector and optimize portfolios to track the industry indices.

Constrained Allocations

Examples of constrained allocations include:

- A fund manager may wish to go long-short in exactly twelve different countries, with the EAFE index as benchmark. The problem then becomes one of selecting the basket of twelve countries that are currently most highly cointegrated with the EAFE index.
- A small asset management company might seek a benchmark return of 5% per annum above the S&P 100 index, so in this case the benchmark index will be the S&P 100 'plus'.
- Assets may be selected according to quite specific preferences of investors. For example, 50% of the fund may have to be allocated to the U.K., or no more than 5% of capital can be allocated to any single asset.

Equality constraints on allocations, such as 40% in technology related stocks, are simple to implement. The dependent variable just becomes $y - \omega_j x_j$, where a fraction ω_j of the fund must be assigned to the jth asset; the other log asset prices are used as regressors with the constraint that the sum of the weights is $1 - \omega_j$. Similarly if more than one asset has a constrained allocation, the dependent variable becomes y – the weighted sum of the constrained log asset prices, and the remaining log asset prices are used as regressors.

Inequality constraints are more difficult to implement. How should one deal with the constraint of no short sales, $\omega_j > 0$ for all j? First perform an unconstrained estimation of the model by ordinary least squares (OLS) because if no constraint is violated there will be no problem. Suppose the constraints $\omega_j > 0$ for some j are violated. Then the model is restricted so that all these ω_j are set to zero, and re-estimated to ensure that no other coefficients that were originally positive have now become negative. If that is the case the resulting constrained OLS estimator is obtained, but it will of course be biased. That it is more efficient than the original estimator because it reflects the value of further information may be little compensation.

Problems arise when imposing the constraints causes more constraints to be violated, so that other coefficients that were positive in the unconstrained model become negative in the constrained model. The only feasible solution is to put those coefficients to zero, re-estimate a further constrained model, and to keep shooting coefficients to zero until a purely long portfolio of assets is obtained. Clearly this can cause severe bias in results: the more constraints that have to be imposed the further the model will be from a true underlying market equilibrium. Therefore the model developer needs to approach this exercise with caution, and to validate his or her judgment by thorough back testing.

Parameter Selection

The basic cointegration index tracking model can be defined in terms of certain parameters:

- Any ‘alpha’ return over and above the index;
- The time-span of daily data that is used in the cointegrating regression (2);
- The number of assets in the portfolio;[12]
- Any constraints on allocations that are defined by the preferences of the investor.

The optimal parameter values are chosen by recording a number of in-sample and post-sample performance measures for each set of parameters. The optimal parameter set is that which gives the ‘best’ performance measures and for the purposes of this paper the most important in-sample performance measures are:

- *Tracking error stationarity*: The standard ADF test is used to test the level of cointegration between the portfolio the benchmark on the historic data: the larger and more negative the ADF statistic, the greater the level of cointegration and the more stationary the tracking error;[13]
- *Standard error of the regression*: The in-sample tracking error will be stationary if the portfolio is cointegrated with the benchmark, so it cannot deviate from the benchmark for too long. However this does not imply that the short-term deviations between the portfolio and the benchmark are necessarily small. It is also important to choose a portfolio for which the in-sample tracking error has a low volatility, and this is measured by the standard error of the regression.
- *Turnover*: Only those portfolios showing realistic turnover projections as the model is rolled over the back test period should be considered. Typically

turnover projections from cointegration based strategies will be much lower than those based on mean-variance analysis.

Having specified the selections and the allocations on the in-sample ‘training period’, a fixed period of data immediately following the in-sample data is used to analyze the out-of-sample performance of the portfolio. These post-sample data are called the ‘testing period’. If the strategy requires monthly re-balancing then it is normal to use a testing period of one month or two months for the post-sample diagnostics. Some typical post-sample diagnostics are:

- *Tracking error variance*: This is the variance of the daily tracking errors during the testing period. The tracking error variance is equivalent to the root mean square forecast error if it is measured as an equally weighted average;
- *Differential return:* The difference between the portfolio return and the benchmark return over the testing period;
- *Information ratio:* The ratio between the mean daily tracking error and the standard deviation of the daily tracking error over the testing period. In-sample information ratios are zero by design (because the residuals from ordinary least squares regression have zero mean) but a high positive post-sample information ratio is very important as a risk adjusted performance measure.

Consider a simple example of how to decide which parameters are optimal. The problem is to track the Morgan Stanley European, Asian and Far Eastern (EAFE) index with a one year buy-and-hold strategy. The alpha over the EAFE index is fixed at 3% per annum and there are no constraints on allocations. Thus there are only two model parameters to be chosen, the number of country indices in the portfolio (at the time of optimization the maximum was 23) and the training period for the model. Figure 2 shows the 12 month out-of-sample information ratios that are obtained as the number of assets selected varies from 5 to 15 and the length of training period varies from 10 to 130 months. From the figure it seems that the highest information ratio of 3.8 occurs when the training period is between 100 and 115 months and the number of assets is between 7 and 11.

Instead of fixing the alpha over an index – or indeed under an index – it may be preferable to fix the number of assets in the portfolio. In that case this type of two-dimensional ‘heat map’ can be used to determine the optimal choice of two other important parameters: the alpha over the index and the length of training period. Examples of such heat maps are given in the next section.

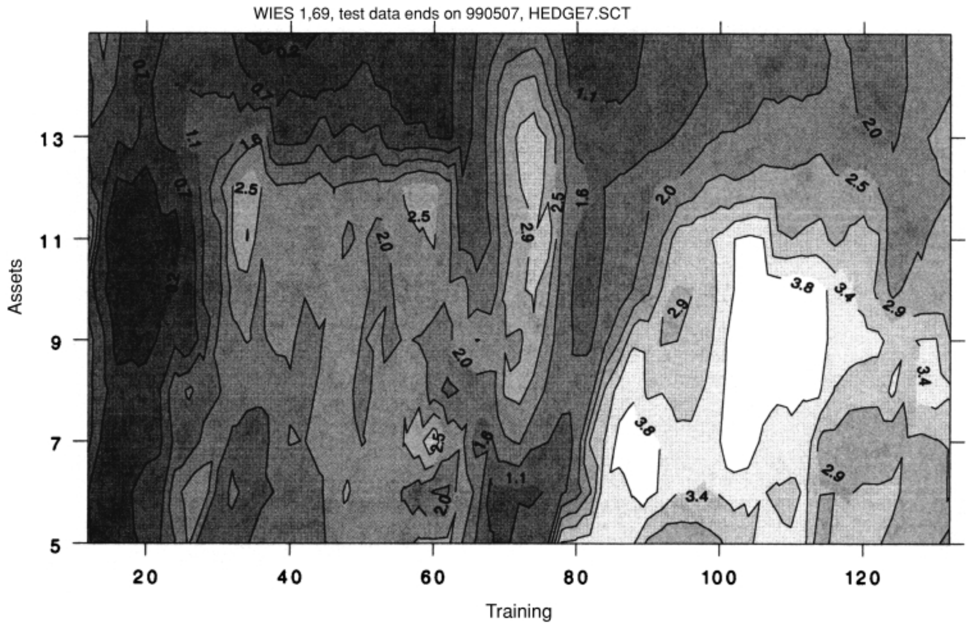

Fig. 2. Alpha 0.03 (outperformance) 12 month Information Ratio.

4. THE HEDGE FUND STRATEGY

The hedge strategy consists of two legs, a long and a short portfolio. Each leg of the long-short equity hedge consists of a portfolio of 75 stocks that will be used to track an S&P 100 index 'plus' benchmark. The model parameter specification stage begins with a heat map, similar to that shown in Fig. 2, that is generated by finding the 75 asset portfolio that is most highly cointegrated with the index plus alpha percent per annum. In this case the parameters to choose will be the alpha over the index (or under, if it is negative) and the length of training period.

Each time the alpha and the training period are changed the choice of assets and the allocations in the portfolio will change. These allocations are not recorded at this stage: all that will be stored are the in-sample and out-of-sample diagnostics that have been described above, for each parameter vector. Figure 3a shows the one month out-of-sample information ratios, and Fig. 3b shows the one month out-of-sample differential returns for a 75 stock portfolio in the S&P 100 index that is being optimised at the end of February 2000. Figures 3c and 3d are similar to Figs 3a and 3b but for the two month information ratio. The maps are colour coded to indicate the regions where better diagnostic test results are obtained.

These heat maps show a clear 'hot spot' when the alpha is negative but no more than about –7% per annum, and the training period is between 28 and 48 months. Another region that gives promising out-of-sample diagnostics is for a high, positive alpha and a very long training period. However the highest differential return and information ratio are in fact obtained within the 'hot spot' when the alpha is approximately –5% and the training period is about 3 years.

The heat maps in Fig. 3 also have a 'cold spot', that is a region where the parameter choices give rise to rather bad performance measures. In particular when the alpha is –12% and the training period is 72 months, the one-month and two-month out-of-sample information ratios are negative, as are the differential returns. For this parameter vector, the out-of-sample performance

Table 1. Long and Short Portfolio Parameter Choices, February 2000.

	Alpha	Training Months
Long	–5%	36
Short	–12%	72

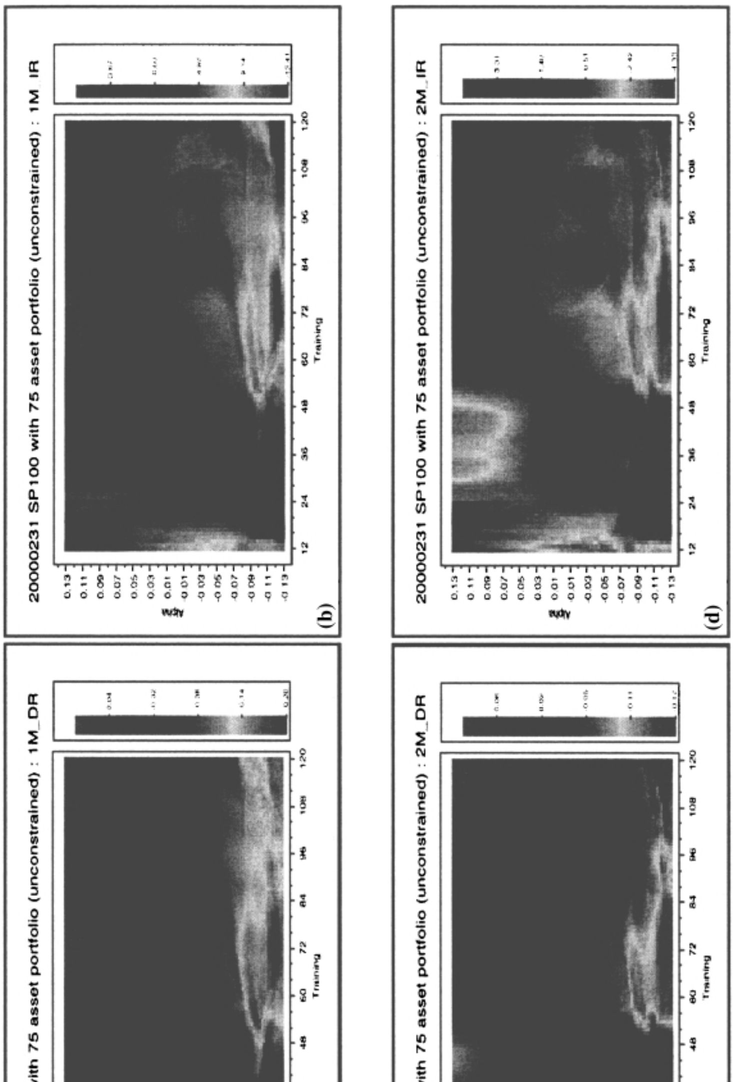

Fig. 3. Heat Maps for Model Parameter Optimization.

Table 2. Optimal Parameter Choices for S&P 100.

	Long		Short			Long		Short	
	Alpha	Training	Alpha	Training		Alpha	Training	Alpha	Training
Jan-95	–0.1	54	–0.06	24	Jan-98	0.04	24	–0.1	78
Feb-95	–0.12	30	0.09	30	Feb-98	0.12	90	–0.1	72
Mar-95	0.08	78	–0.05	72	Mar-98	0.12	90	–0.05	30
Apr-95	0	18	–0.05	66	Apr-98	0.12	102	–0.03	12
May-95	0.1	66	–0.05	18	May-98	0.12	84	–0.1	24
Jun-95	0.1	66	–0.1	102	Jun-98	0.12	84	–0.1	24
Jul-95	0.12	30	–0.12	102	Jul-98	0.12	12	–0.1	24
Aug-95	0.12	30	–0.12	96	Aug-98	0.12	84	–0.12	102
Sep-95	0.12	42	–0.12	36	Sep-98	0.12	30	–0.12	54
Oct-95	0.1	66	–0.07	12	Oct-98	–0.11	102	–0.1	12
Nov-95	0.12	60	–0.1	102	Nov-98	0.09	12	–0.1	12
Dec-95	0.12	54	–0.11	102	Dec-98	0.05	12	–0.04	18
Jan-96	–0.12	48	–0.12	102	Jan-99	0.12	78	–0.11	78
Feb-96	0.08	78	–0.12	102	Feb-99	0.12	102	–0.12	66
Mar-96	0.12	48	–0.05	90	Mar-99	0.03	48	–0.11	54
Apr-96	0.11	60	–0.08	96	Apr-99	0.06	102	0.04	36
May-96	0.02	72	–0.1	78	May-99	0.05	102	0.11	18
Jun-96	0.11	78	–0.09	78	Jun-99	–0.03	12	–0.08	30
Jul-96	0.12	78	–0.11	102	Jul-99	0.12	30	–0.09	30
Aug-96	0.12	78	–0.12	96	Aug-99	0.12	78	–0.12	96
Sep-96	0.01	42	0	96	Sep-99	0.12	78	–0.12	84
Oct-96	0.11	84	0	96	Oct-99	0.12	36	–0.11	102
Nov-96	0.12	78	–0.12	90	Nov-99	0.12	54	–0.1	36
Dec-96	–0.12	66	0.12	66	Dec-99	0.12	54	–0.11	96
Jan-97	0.01	24	0.11	66	Jan-00	0.12	16	–0.12	88
Feb-97	0.01	102	–0.12	54	Feb-00	0.12	6	–0.11	12
Mar-97	0.01	102	–0.12	102	Mar-00	0.12	64	–0.12	92
Apr-97	0.09	66	–0.12	30	Apr-00	0.12	64	–0.1	60
May-97	0.05	72	–0.12	12	May-00	0.12	16	–0.08	40
Jun-97	0.11	90	0	102	Jun-00	0.06	12	–0.08	28
Jul-97	0.11	96	–0.1	102	Jul-00	0.12	48	–0.1	88
Aug-97	0.11	96	–0.12	102	Aug-00	0.12	120	–0.13	20
Sep-97	–0.11	102	–0.12	36	Sep-00	0.12	120	–0.12	20
Oct-97	0.12	12	–0.11	12	Oct-00	0.11	108	–0.13	88
Nov-97	–0.08	12	0.12	12	Nov-00	–0.11	90	–0.12	24
Dec-97	0	18	0.12	96	Dec-00	–0.09	20	–0.09	26

of the portfolio is returning considerably less than the index. Therefore it would be possible to make money by going short this portfolio.

Note that this 'short' portfolio will itself contain long and short positions, unless the constraint of no short sales has been applied. Similarly the 'long' portfolio, the one that has the highest information ratio and differential return, will typically also consist of long and short positions. Then a hedged portfolio is obtained by matching the amount invested in the long portfolio with the same amount being shorted with the short portfolio.

When this type of long-short strategy is used with a 75 stock portfolio from the S&P 100, Fig. 3 indicates that the optimal parameter choices for February 2000 are shown in Tables 1 and 2.

Typically the optimal parameter choices will be different every month, and will depend on any allocation constraints. Table 2 shows the parameter choices that were actually used for the 75 asset long portfolio and a 75 asset short portfolio in the S&P 100 index. Note that in February 2000 the parameters are different from those shown in Table 1. This is because in Table 2 allocations were constrained so that the total (long + short) allocation to each asset will be no more than 5% of the fund.[14] Of course some of the same assets will be chosen in both portfolios, and the net position in these assets will be determined by the difference of their weight (positive or negative) in the long portfolio minus their weight (positive or negative) in the short portfolio.

5. THE RESULTS

This section describes two types of model back tests. The first type of back test is of a fixed parameter set that is optimal according to 'heat maps' of the type just described and the test is of its performance over an historic period. A simple snap-shot of portfolio performance at one instance in time as in Fig. 3, may not provide sufficient evidence that parameter choices are optimal. Dynamic performance measures can be obtained by running the model over time, for example month by month. Each month a new set of assets will be chosen and new allocations will be made, but the set of parameters remains fixed. Table 3 reports the in-sample ADF, the turnover percent, and the one-month, two-month and three-month out of sample information ratios for the long and the short portfolio parameters that were optimal in October 2000, given in Table 2. The rest of the table indicates how these parameter choices would have performed since January 1995. For example the first line of the table shows that the portfolio parameter choice of alpha 5% (long) and –8% (short) and training 30 months would not have been a good choice at all in

!! Author: Table 2 OK !!

Table 3. Back Testing the October 2000 Parameter Choices.

	Long					Short				
	ADF	Turnover	1mth IR	2mth IR	3mth IR	ADF	Turnover	1mth IR	2mth IR	3mth IR
Jan-95	–10.57	0.78	–4.66	–2.47	–1.29	–10.59	0.77	9.09	2.54	2.14
Feb-95	–11.12	0.33	–4.32	–4.38	–2.80	–11.47	0.58	–0.22	4.16	1.69
Mar-95	–11.11	0.42	–0.63	–2.35	–3.31	–12.61	0.71	1.86	0.67	3.47
Apr-95	–10.93	0.47	0.13	–0.25	–1.47	–12.35	0.61	–3.22	–1.04	–0.76
May-95	–10.51	0.57	–0.32	–0.10	–0.27	–12.19	0.47	1.49	–0.85	–0.13
Jun-95	–10.74	0.36	3.79	1.39	0.93	–11.95	0.38	–4.91	–2.19	–2.50
Jul-95	–11.27	0.30	5.16	4.53	2.60	–12.16	0.64	–0.40	–2.80	–1.60
Aug-95	–10.88	0.38	3.97	4.61	4.37	–12.60	0.58	–4.74	–2.22	–3.26
Sep-95	–11.54	0.56	–2.13	0.42	1.86	–12.27	0.98	0.29	–2.29	–1.50
Oct-95	–11.57	0.39	3.50	0.75	1.57	–11.97	0.49	–6.04	–2.90	–3.54
Nov-95	–11.76	0.35	2.79	3.21	1.39	–12.33	0.99	2.26	–1.50	–0.96
Dec-95	–12.08	0.25	3.81	3.10	3.26	–12.06	0.70	1.75	2.05	–0.47
Jan-96	–11.52	0.26	3.36	3.54	3.43	–12.38	0.54	–3.32	–1.84	–0.25
Feb-96	–11.25	0.44	0.78	2.01	2.41	–12.26	0.82	–1.25	–2.46	–1.67
Mar-96	–11.94	0.42	3.77	2.06	2.52	–12.64	0.37	–0.68	–0.92	–1.81
Apr-96	–12.06	0.26	–0.35	1.56	1.27	–12.41	1.19	0.24	–0.26	–0.54
May-96	–12.65	0.38	2.68	1.15	1.96	–12.05	0.51	–2.08	–0.92	–0.84
Jun-96	–12.33	0.26	3.03	2.88	1.84	–11.57	0.87	3.85	0.83	0.64
Jul-96	–12.02	0.28	0.29	1.49	1.80	–11.62	0.60	–1.80	0.11	–0.44
Aug-96	–12.97	0.39	2.87	1.37	1.89	–12.99	1.08	–1.68	–1.74	–0.44
Sep-96	–12.28	0.53	–1.63	1.03	0.69	–12.68	0.82	–0.49	–1.18	–1.44
Oct-96	–12.19	0.26	–3.27	–2.57	–0.43	–13.18	1.03	1.38	0.60	–0.24
Nov-96	–12.26	0.23	–2.67	–3.00	–2.62	–12.40	1.19	2.79	2.04	1.32
Dec-96	–12.89	0.29	2.26	0.12	–0.89	–11.94	0.57	6.42	4.79	3.58
Jan-97	–12.00	0.31	–2.57	–0.40	–0.83	–12.07	0.76	–4.45	0.49	0.69
Feb-97	–12.48	0.17	5.54	0.10	0.80	–12.25	0.69	2.82	–0.33	1.33
Mar-97	–11.85	0.33	3.99	4.61	1.34	–12.19	0.46	–4.38	–0.54	–1.65
Apr-97	–12.79	0.26	3.06	3.50	3.93	–12.22	0.37	–2.78	–3.58	–1.30
May-97	–11.54	0.46	–0.98	1.54	2.35	–13.04	0.52	0.55	–1.23	–2.27
Jun-97	–12.32	0.26	3.40	1.38	2.09	–13.16	0.83	–0.08	0.20	–0.82
Jul-97	–11.97	0.16	1.00	2.03	1.21	–12.72	0.73	–3.94	–2.14	–1.39
Aug-97	–11.97	0.64	–3.27	–0.78	0.48	–12.26	1.01	2.46	–0.90	–0.65
Sep-97	–11.36	0.47	–2.58	–2.76	–1.52	–12.12	0.66	3.12	2.82	0.34
Oct-97	–10.67	0.64	–4.70	–3.58	–3.46	–12.62	1.50	0.96	1.87	2.06
Nov-97	–11.89	0.45	2.92	–1.13	–1.68	–12.76	1.13	–0.40	0.42	1.24
Dec-97	–11.37	0.37	2.26	2.55	0.19	–12.69	1.00	1.45	0.72	0.82

Table 3. Continued.

	Long					Short				
	ADF	Turnover	1mth IR	2mth IR	3mth IR	ADF	Turnover	1mth IR	2mth IR	3mth IR
Jan-98	–12.11	0.37	1.26	2.67	3.19	–13.27	0.41	–7.46	–1.17	–1.47
Feb-98	–12.55	0.34	3.03	2.01	2.75	–13.43	0.31	–2.87	–5.25	–1.59
Mar-98	–11.95	0.39	1.47	2.13	1.82	–12.95	1.08	–2.24	–2.45	–3.81
Apr-98	–11.99	0.43	8.43	4.79	4.33	–12.64	1.26	–2.26	–2.28	–2.39
May-98	–12.44	0.32	1.63	5.14	3.85	–12.48	1.12	–4.42	–3.24	–2.88
Jun-98	–11.98	0.56	2.47	2.11	4.19	–12.08	1.08	–5.76	–5.21	–4.23
Jul-98	–11.88	0.28	3.63	3.01	2.60	–11.83	1.72	–4.47	–5.11	–4.91
Aug-98	–10.76	0.65	–4.43	–1.59	–0.47	–11.43	1.61	5.82	1.82	0.08
Sep-98	–10.45	0.42	–2.89	–3.77	–2.00	–11.39	0.93	1.31	3.82	1.67
Oct-98	–11.85	0.28	6.13	2.00	–0.42	–12.27	1.10	–3.34	–1.24	1.35
Nov-98	–11.44	0.40	–2.66	3.14	1.07	–11.96	0.71	2.33	–1.22	–0.37
Dec-98	–11.75	0.25	3.81	1.16	3.34	–11.71	0.70	–4.85	–1.31	–2.19
Jan-99	–12.48	0.47	–3.93	–1.05	–1.88	–11.06	0.66	0.81	–2.63	–1.96
Feb-99	–12.37	0.32	2.10	–1.16	–0.07	–10.90	0.97	2.54	1.60	–1.02
Mar-99	–11.78	0.43	–3.73	–0.90	–2.00	–11.53	1.46	2.80	2.68	1.96
Apr-99	–11.78	0.41	1.82	–0.70	0.14	–11.51	1.26	1.71	1.99	2.14
May-99	–11.96	0.30	2.61	2.20	0.33	–11.48	0.51	–2.80	0.30	0.92
Jun-99	–12.27	0.49	3.28	2.98	2.55	–11.51	0.88	–5.27	–4.17	–1.14
Jul-99	–11.92	0.46	–0.64	1.36	1.80	–10.98	1.27	2.71	–2.03	–2.30
Aug-99	–12.59	0.27	3.16	1.35	2.00	–11.54	0.90	–2.05	–0.23	–2.05
Sep-99	–12.30	0.17	7.87	4.97	2.97	–11.44	0.72	–7.83	–4.16	–2.24
Oct-99	–11.48	0.43	–1.66	1.52	2.07	–12.27	1.20	–3.30	–4.86	–3.83
Nov-99	–12.01	0.45	–0.35	–1.08	0.91	–12.08	0.71	–0.10	–1.93	–3.37
Dec-99	–11.49	0.54	1.15	0.63	–0.02	–12.07	0.97	–9.05	–5.21	–4.60
Jan-00	–12.03	0.35	2.65	0.95	1.07	–11.82	1.04	4.99	–0.32	–0.83
Feb-00	–12.15	0.42	–5.15	–1.31	–1.15	–11.31	0.89	–2.00	1.88	–0.82
Mar-00	–13.12	0.50	–1.55	–3.04	–1.41	–11.48	0.95	1.40	0.13	1.69
Apr-00	–13.33	0.32	2.57	0.63	–0.79	–10.94	1.39	2.67	2.01	1.02
May-00	–12.65	0.27	1.77	2.18	0.88	–11.01	1.06	3.37	3.04	2.45
Jun-00	–12.07	0.44	–0.52	0.60	1.42	–11.68	1.45	–5.59	–0.97	0.25
Jul-00	–12.48	0.45	–1.18	–0.88	–0.08	–11.68	1.06	0.65	–2.79	–0.53
Aug-00	–12.01	0.36	–3.85	–1.94	–1.46	–10.97	1.23	–0.79	–0.12	–2.13
Sep-00	–11.80	0.25	4.05	1.64	0.73	–11.40	1.13	2.28	0.68	0.67
Oct-00	**–12.45**	**0.38**	**5.62**	**4.63**	**2.85**	**–11.40**	**1.36**	**–2.54**	**–0.40**	**–0.53**

January 1995. In fact with these choices the long portfolio under performed and the short portfolio out performed the index!

The object of this exercise is to check the robustness of the portfolio over time: this is fundamental to cointegration. Consistency between the 1 month, 2 month and 3 month information ratios is paramount: Table 3 indicates that if the 1 month information ratio is high so also, on the whole, are the 2 month and 3 month information ratios. Thus, if a portfolio starts well for the first month, it tends to perform well over several months. Similarly if the portfolio does not perform well during the first month, this tends to continue for subsequent months. Since the parameter choice for the current month is made on the basis of last month's performance, this autocorrelation in information ratios is a crucial performance indicator. Another robustness check for the portfolio is to ensure high ADFs and relatively low turnover projections when the portfolio is rebalanced using the same fixed parameters each month.[15]

It is evident from Table 2 that the same parameter selection is not usually optimal for two consecutive months. Therefore the most important back testing of the hedge fund strategy is to report the returns that are obtained when the alpha and training periods are re-optimized every month. Table 4 gives the consolidated returns from applying the long-short hedge strategy to the S&P 100 during three years that have been chosen as representative of difficult market conditions (1987, 1990 and 1993) and then continuously from 1995–2000.[16]

Note the large annual rates of return in 1998 and 1999: these were significantly strong years for the stock markets. Due to the inherent convexity

Table 4. Consolidated Returns from the Long-Short Strategy.

	1987	1990	1993	1995	1996	1997	1998	1999	2000
Compound Return									
SP100	8.9%	–4.4%	8.7%	37.2%	24.1%	29.8%	34.3%	33.7%	–13.4%
L-S HEDGE	14.3%	7.3%	15.8%	18.3%	13.6%	11.6%	46.4%	82.5%	34.9%
Fund Daily Returns:									
Maximum	4.6%	2.5%	2.9%	2.2%	2.0%	3.2%	3.8%	5.9%	4.3%
Minimum	–6.0%	–3.0%	–1.5%	–2.0%	–1.9%	–3.0%	–4.3%	–4.3%	–4.1%
Average	0.1%	0.0%	0.1%	0.1%	0.1%	0.0%	0.2%	0.2%	0.2%
Median	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%	0.2%	0.1%	0.0%
Volatility (St Dev)	1.0%	0.7%	0.6%	0.7%	0.6%	0.9%	1.1%	1.5%	1.5%
Max 30-DayPeak-to-Trough	7.8%	10.2%	4.9%	7.3%	5.0%	12.6%	14.4%	10.1%	7.4%
Max 30-Day Trough-to-Peak	9.7%	9.8%	6.9%	8.7%	5.4%	13.9%	22.9%	24.0%	13.8%

Table 5. Monthly Consolidated Returns 1987, 1993 and 1998.

1987	SP100	L-S HEDGE	1993	SP100	L-S HEDGE	1998	SP100	L-S HEDGE
Jan-87	13.06%	1.72%	Jan-93	1.26%	–0.33%	Jan-98	2.04%	6.61%
Feb-87	3.89%	–0.37%	Feb-93	1.56%	4.54%	Feb-98	6.88%	–0.23%
Mar-87	3.31%	3.08%	Mar-93	1.69%	4.17%	Mar-98	5.30%	2.89%
Apr-87	0.68%	0.02%	Apr-93	–1.72%	–0.78%	Apr-98	1.49%	4.95%
May-87	0.68%	5.64%	May-93	2.59%	0.16%	May-98	–1.34%	3.33%
Jun-87	4.86%	0.30%	Jun-93	–0.17%	1.33%	Jun-98	4.83%	13.79%
Jul-87	4.26%	0.57%	Jul-93	–0.52%	1.84%	Jul-98	–0.54%	4.94%
Aug-87	4.50%	0.33%	Aug-93	3.31%	–0.89%	Aug-98	–15.16%	–8.83%
Sep-87	–3.09%	–1.91%	Sep-93	–1.53%	4.67%	Sep-98	5.10%	2.91%
Oct-87	–19.77%	2.27%	Oct-93	1.54%	–2.14%	Oct-98	8.78%	6.53%
Nov-87	–9.87%	–0.75%	Nov-93	–0.55%	0.42%	Nov-98	7.33%	–2.27%
Dec-87	6.00%	2.54%	Dec-93	0.88%	1.63%	Dec-98	4.73%	3.54%
Simple Return	8.51%	13.41%		8.34%	14.63%		29.46%	38.14%
Compound Return	8.88%	14.35%		8.70%	15.75%		34.26%	46.43%
Standard Deviation	8.10%	1.95%		1.52%	2.16%		6.08%	5.23%
Excess Return (Compound vs Index)		5.47%			7.05%			1.7%
Sharpe Ratio (Over Risk Free Rate 5%)		1.38			1.44			2.29

of the "long" side of the hedge (a long-biased combination of longs and shorts) the fund was able to capture market upswings. This did not compromise the downside protection of the "short" side of the hedge (a short-biased combination of shorts and longs) as shown in years in which the market retreated (1990, 2000). At all times the composite portfolio was dollar-neutral.

BARRA have performed a verification analysis of the 1999 returns.[17] They have shown that the strategy derives its excess returns primarily from risk assessment items such as earnings yield, earnings variation, momentum, size, and, as one would expect, some leverage. Both "Value" and "Growth" assessment are actually negative contributors to the return, which somewhat distinguishes the strategy from the status quo.

Monthly returns for some of the most difficult years for equity markets (1987, 1993 and 1998) are given in Table 5 above. The long-short strategy performs relatively well during the market crashes of October 1987 and August 1998. The returns are also much less volatile than the S&P 100 index returns, and during each of these years the fund outperformed the index considerably.

Over the entire back-test [1987, 1990, 1993, and 1995–2000]:

- the correlation between the hedge strategy and the S&P 100 was only –15.2%;
- the average annual Sharpe Ratio was 1.51;
- the average leverage was approximately 1.5 on both the long and the short legs of the hedge;
- the average annual returns were 27.2% and the average annual risk (annualized volatility from the monthly standard deviation of returns) was 11.4%.

6. SUMMARY AND CONCLUSIONS

This paper has described a long short hedge strategy that is based on cointegration between asset prices. Traditional strategies will not guarantee that the tracking error is stationary and will therefore require frequent rebalancing for the hedge to remain tied to the benchmark. The cointegration strategy, on the other hand, is based on the criterion that the hedge is mean-reverting to the benchmark; tracking errors are designed to be stationary and this may be achieved with relatively few stocks and with much lower turnover rates.

The strategy will accommodate investor's preferences for the alpha as well as flexible constraints on allocations. A sophisticated training and testing methodology has been described for the selection of model parameters. Extensive back testing results were reported and these have demonstrated the ability of the model to capture market upswings whilst not compromising the downside protection.

Hedge funds are clearly the future: not as a replacement to traditional investment techniques, but as an alternative investment tool. There is a growing consensus that the best hedge funds can deliver risk-adjusted returns that are superior to those from traditional "long only" strategies: they can augment a portfolio manager's risk adjusted returns because they are relatively uncorrelated with the returns from the traditional assets or funds in the portfolio. Transparency, liquidity and performance analytics offered by sophisticated on-line platforms will provide security for institutional investors to allocate in size. Indeed European pension funds, with less than 1% of their assets currently in hedge funds, are ready to more than quintuple their investment in alternative asset classes over the next three years. These developments will continue to fuel the growth in market neutral hedge fund strategies such as the one that has been described in this paper.

NOTES

1. TASS identifies 11 basic investment styles in hedge funds: Long/short Equity; Equity Market Neutral; Equity Trading; Event Driven; Convertible Arbitrage; Fixed Income Relative Value/Arbitrage; Global Macro; Short Sellers; Emerging Markets; Managed Futures; Funds of Funds.

2. Convex to the market means that if the market goes up by 10% the fund will increase by more than 10%; if the market goes down 10% the fund will decrease by less than 10%. In other words, the ‘up market’ beta is greater than the ‘down market’ beta; in fact the down market beta might even be negative. Convexity suggests market exposure, and indeed the empirical results described in this paper show that there is some degree of market exposure in certain years.

3. Every modern econometrics text covers the statistical theory necessary to master the practical application of cointegration, Hamilton (1994), Enders (1995) and Hendry (1996) being amongst the best sources.

4. A process is integrated of order 1 if it is not stationary, but becomes stationary after first differencing. Thus random walks are integrated of order 1, but not all I(1) series are random walk because there may be autocorrelation after first differencing.

5. Of course this example is very theoretical. It is unlikely cointegrated series will conform to this model in practice, but it useful for illustration.

6. This follows since $V(\Delta x) = \sigma^2 + 2\sigma_x^2$, $V(\Delta x) = \sigma^2 + 2\sigma_y^2$, and $COV(\Delta x, \Delta y) = \sigma^2$

7. The connection between these two methodologies is that a principal component analysis of cointegrated variables will yield the common stochastic trend as the first principal component. But the outputs of the two analyses differ: principal components gives two or three series which can be used to approximate a much larger set of series (such as the yield curve); cointegration gives all possible stationary linear combinations of a set of random walks. See Gouriereux et al. (1991).

8. Details of the extensive research in these areas are given, with many references, in Alexander (1999a, b and 2001).

9. In the case of tracking an index ‘plus’ alpha percent per annum, the dependent variable will be the log of the index ‘plus’ series (this is defined as the index price plus a small increment that amounts to alpha percent over the year).

10. Note the definition of ‘tracking error’ in this paper. Contrary to standard, but confusing, terminology, ‘tracking error’ here is *not* defined the variance or volatility of the difference between the portfolio and the benchmark.

11. Classical regression assumes the dependent and independent variables are stationary, so that the error term will be stationary by definition. However in a cointegrating regression the dependent and independent variables are integrated, therefore the error will only be stationary under special circumstances. Indeed, the error term will be stationary if and only if the dependent variable is cointegrated with the explanatory variables. Thus a statistical test for cointegration is to perform such a ‘cointegrating regression’ and then test the residuals for stationarity.

12. In fact the number of non-zero allocations need not be specified. Instead the number of assets chosen can depend on a bound that is set for the tracking error variance.

13. The 1% critical value of the ADF statistic is approximately –3.5, although much greater values than this are normally experienced in practice as for example in Table 3 and in Fig. 3d.

14. More information about the operation of this hedge fund strategy is available from www.pennoyercapital.com

15. The figures in Table 3 represent percentage turnover: that is, 1$ long in security X converted to 1$ short in the same security represents a 200% turnover.

16. The returns stated include transactions costs but no other fees. Daily closing prices on the S&P 100 stocks we taken from the University of Chicago Research in Securities Prices (CRISP) database. The prices were adjusted for splits, dividends and mergers.

17. While the data was correct with respect to stock splits and mergers, Pennoyer's analysis did not include dividend effects. BARRA's confirmation did, which most often resulted in slightly higher calculated returns from BARRA. Nonetheless, the difference was not significant.

REFERENCES

Alexander, C. O. (1999a). Optimal Hedging Using Cointegration. *Philosophical Transactions of the Royal Society*, *A357*, 2039–2058.

Alexander, C. O. (1999b). Cointegration and correlation in energy markets. In: *Managing Energy Price Risk* (2nd ed., Chap. 15, pp. 291–304). RISK Publications.

Alexander, C. O. (2001). *Market Models: A Guide to Financial Data Analysis*. John Wiley.

Alexander, C. O., & Johnson, A. (1992). Are foreign exchange markets really efficient? *Economics Letters*, *40*, 449–453.

Alexander, C. O., & Johnson, A. (1994). Dynamic links. *RISK*, *7*(2), 56–61.

Cerchi, M., & Havenner, A. (1988). Cointegration and stock prices. *Journal of Economic Dynamics and Control*, *12*, 333–346.

Covey, T., & Bessler, D. A. (1992). Testing for Granger's full causality. *Review of Economics and Statistics*, 146–153.

Enders, W. (1995). *Applied Dynamic Econometrics*. John Wiley.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*, *55*(2), 251–276.

Gourieroux, C., Monfort, A., & Renault, E. (1991). A general framework for factor models. *Institut National de la Statistique et des Etudes Economiques*, No. 9107.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

Hendry, D. F. (1996). *Dynamic Econometrics*. Oxford University Press.

Pindyck, R. S., & Rothemberg, J. J. (1992). The comovement of stock prices. *Quarterly Journal of Economics*, 1073–1103.

Stock, J. H., & Watson, M. W. (1988). Testing for common trends. *Journal of the American Statistical Association*, *83*(404), 1097–1107.