Correlation in Crude Oil and Natural Gas Markets

Carol Alexander

Correlation is a measure of dependency between stationary time series. In this chapter the stationary time series will be the $\tau$-period log returns on crude oil and natural gas futures or forwards, denoted either $x_t = \ln(P_t/P_{t-\tau})$ or $x(t) = \ln(S(t)/S(t-\tau))$ depending on whether we are considering a discrete time or a continuous time application. We consider only the simple linear pair-wise correlation $\rho_{xy} = \sigma_{xy}/\sigma_x\sigma_y$ where $\sigma_{xy}$ denotes the covariance between two log returns $x$ and $y$ (i.e. the first moment about the mean of their joint density) and $\sigma_x$ and $\sigma_y$ denote the standard deviations of each marginal returns density.

Two broad categories for applications of correlation to energy markets may be distinguished:

- Risk assessment and risk control of trading operations
- Pricing of multi-asset derivative products and the hedging of portfolios

The first category includes the assessment of portfolio risk as given by a statistical forecast of the volatility of a portfolio, or the portfolio’s value-at-risk (VaR), over a pre-defined risk horizon. The portfolio risk is assessed using the volatility and correlations between the instruments in the portfolio, or the risk factors of the portfolio and the factor sensitivities of the instruments, if a factor model is used. These volatilities and correlations are summarized in the covariance matrix: a square, symmetric positive definite matrix with variances along the diagonal and covariances on the off-diagonals. Portfolio risk assessments are aggregated using assumptions on the broad correlations between different types and locations of traded instruments, to achieve a total risk assessment covering all trading operations. The individual risk assessments are used to guide the global positioning of risks by senior management, who undertake a risk budgeting process across the different trading units in which risk limits are related to portfolio risk assessments.

The second category concerns the pricing and hedging of path dependent multi-asset derivative products, where option values are based on correlated price diffusions. A forecast of the average volatility of an underlying instrument over the lifetime of a single asset option may be implied from its current market price. In some cases (e.g. spread options) implied correlations may also be inferred from market prices. Again it is convenient to summarise these in a covariance matrix; what distinguishes this matrix from the matrix used in the first category of applications is that, when possible and if implied correlations are not too unstable, the parameters in the covariance matrix are calibrated to current prices of European put and call options, rather than forecasted using historical data on the underlying returns.

---

1 Professor of Risk Management and Director of Research, ISMA Centre, University of Reading Business School, Berks, UK. Email: c.alexander@ismacentre.rdg.ac.uk
The main distinction between these two broad applications is the mathematical approach: one of discrete
time series analysis in the first category and continuous time diffusion models in the second. In this
chapter we consider both types of applications.

**TIME SERIES MODELS FOR COVARIANCE MATRICES**

In this section we consider statistical forecasts of covariance matrices based on discrete time series
models.\(^a\) These are used for

- *Assessing the volatility of a linear portfolio*: this is given by \( \sqrt{w'Vw} \) where \( V \) is the covariance matrix of
  returns to the instruments in the portfolio and \( w \) is the vector of portfolio weights. If the
  assumption of normal returns is made, then the volatility is directly translated into the portfolio
  VaR on multiplication by a critical value of the standard normal density.\(^iii\)

- *Monte Carlo VaR of a portfolio of options*: the covariance matrix is used to simulate correlated
  movements in the underlying instruments over the risk horizon and the VaR estimate is a lower
  percentile of the density of price changes based on the simulated terminal price density.

- *Multi-asset option valuation*: whilst volatilities are often calibrated to current option values, when
  correlations are calibrated to the market they can be very unstable. For this reason many authors
  advocate the use of historical data to estimate correlations between diffusion processes.

Examples to illustrate the different time series models for covariance matrices will be based on the crude
oil and natural gas prices that were kindly provided by the editor, Vince Kaminski. For the period from 4\(^{th}\)
January 1993 until 20\(^{th}\) November 2003 we have daily closing New York Mercantile Exchange (NYMEX)
futures prices, from 2 to 12 months out, on West Texas Intermediate (WTI) light sweet crude oil and on
natural gas. These prices are shown in Figures 1 and 2.

<Figures 1 & 2 here>

The crude oil futures shown in Figure 1 normally display a downwards sloping term structures
(backwardation). However during periods when prices were trending downwards, such as May 1993 –
May 1994 and during the whole of 1998 into early 1999, an upwards sloping term structure (contango) is
apparent. Clearly long term futures prices are less volatile than short term futures prices, which respond
more to current market demand and are less influenced by expectations. Nevertheless crude oil futures
form a very highly correlated system with only a few independent sources of information influencing their
movements.

A different picture emerges on examination of the natural gas prices in Figure 2. There is no systemic
backwardation or contango in this market and the futures returns display lower and less stable correlations
than the crude oil futures. Instead there are significant seasonal effects with the short term future
responding most to fluctuations in demand and supply. Storage also plays an important role and, if filled to capacity, long term futures prices may be little influenced by short term fluctuations in demand.

### Moving Average Models

Simple time series models for covariance matrices may be based on moving averages, being either:

- **Equally weighted using a fixed number of historical returns**: also called ‘historical’ volatilities and correlations, these are calculated from the simple variance (the average of the past \( n \) squared returns) and covariance (the average of the past \( n \) cross products of returns). Although simple to calculate, these are known to have the problem of ‘ghost features’ where a single outlier can give the false impression of a high volatility or high correlation for exactly \( n \) days after its occurrence. This model is not as useful for short term forecasting as other models, but it can provide some useful information about the average volatility and correlation over a longer time horizon.

- **Exponentially weighted using a ‘smoothing constant’ \( \lambda \)**: in this approach, popularised by the RiskMetrics group during the mid 1990’s, volatility and correlation declines exponentially following a large price jumps, so that no ‘ghost features’ arise. Though more useful than the equally weighted average for forecasting over the short to medium term, it has two significant drawbacks: first, there is no optimal method for estimating the value of the smoothing constant; and secondly, the model merely provides an estimate of the current volatility or correlation. If used as a forecast, the assumption is that volatility and correlation are constant.

Assuming daily data, ‘historic’ correlations are calculated by dividing the equally weighted covariance estimate over the last \( n \) days by the square root of the product of the two \( n \)-day variance estimates:

\[
\hat{\rho}_{t} = \frac{\sum_{i=1}^{n} r_{1,t-i} r_{2,t-i}}{\sqrt{\sum_{i=1}^{n} r_{1,t-i}^2 \sum_{i=1}^{n} r_{2,t-i}^2}}
\]

Exponentially weighted moving average (EWMA) correlations are calculated by dividing the covariance estimate

\[
\hat{\sigma}_{12,t} = (1-\lambda) r_{1,t-1} r_{2,t-1} + \lambda \hat{\sigma}_{12,t-1}
\]

by the square root of the product of the two variance estimates based on the same value of \( \lambda \), i.e. each variance estimate is given by a similar recursion:

\[
\hat{\sigma}_{t}^2 = (1-\lambda) r_{t-1}^2 + \lambda \hat{\sigma}_{t-1}^2
\]

Figures 3 and 4 show the historic correlations between the 2mth and the 6mth future based on \( n = 250 \) and the EWMA correlations based on \( \lambda = 0.94 \) (this being the ‘RiskMetrics’ value for the smoothing constant). As expected the crude oil futures are very highly correlated, remaining above 0.95 most of the
sample period and only falling during times of crisis (e.g. the outbreak of war in Iraq); Substantial differences between the EWMA short term correlations estimates and the long term ‘historic’ estimates arise following market crises, where the long term correlation remains too low for too long – for example a single outlier in March 1996 induced a low historic correlation estimate for a whole year, miraculously jumping up to the normal level exactly 250 days after the event, even though nothing happened in the markets on that day in March 1997 – the single outlier simply fell out of the moving average window.

Whilst ‘historic’ long term correlations are biased by single outliers, on the other hand the EWMA short term correlations shown in Figures 3 and 4 jump around too much. The natural gas futures correlations in particular are highly variable (note that the scale of the two graphs is very different). In addition they have a systemic seasonality. When supply is filled to capacity, often between the months of August to November, the 6 month future responds much less than the 2 month future to demand fluctuations and their correlation can be very low. At other times the correlation is normally above 0.9. The marked upwards trend in historic correlations between the natural gas futures is probably a result of increasing liquidity in the natural gas markets.

GARCH Models

Both types of moving average models provide only an estimate of the current volatility and correlation. If used as a forecast the assumption of constant volatility and correlation must be made (i.e. the current estimate is the forecast for all risk horizons) and implicit in this assumption is that returns are independent and identically distributed over time. Given the huge amount of empirical evidence on ‘volatility clustering’ going back as far as Mandlebrot (1973) we know that this assumption is unrealistic. Hence the popularity of the generalised autoregressive conditional heteroscedasticity (GARCH) models of volatility introduced by Engle (1982) and Bollerslev (1986). A simple GARCH model that captures the volatility clustering effect has a conditional variance given by

\[ \sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 \]

where \( r_t \) is the (zero mean) return\(^{\text{vii}} \) and restrictions may be applied to the parameters to guarantee a positive and finite conditional variance. The parameters have a natural interpretation: \( \alpha \) measures the reaction of volatility to market events, \( \beta \) measures the persistence in volatility irrespective of anything happening in the market, and \( \omega \) determines the long term average volatility level – this is given by the square root of \( \omega/(1 - \alpha - \beta) \).

The GARCH model can be viewed as a generalisation of the EWMA model. By adding a constant term \( \omega \) and disengaging the parameters \( \alpha \) and \( \beta \) so that they are not constrained to add to one,\(^{\text{viii}} \) the GARCH model generates term structure forecasts of volatility that converge smoothly to the long term average.
That is, forecasts of the average volatility over the next $n$ days converge to the long term average volatility as $n$ increases, just as implied volatility term structures do. Similarly, GARCH correlation models can provide term structure forecasts for correlation, where short term correlations may be higher or lower than long term correlations, and correlation forecasts over the next $n$ days converge to the long term average correlation as $n$ increases.

Simple moving average models might fix a few points on the correlation term structure forecast, but in an ad hoc way. For instance, two points on the correlation term structure could be a 1 day forecast from the EWMA model (with some value of $\lambda$ chosen arbitrarily) and a 1 year forecast from the historic model (with some fixed number of historic returns, also chosen arbitrarily). On the other hand a multivariate GARCH model provides forecasts of average volatility and correlation for all possible future horizons; over the next day, two days, week, month, year and so forth. Thus GARCH covariance matrices represent a significant advance on the simple moving average covariance matrices. Recently many different multivariate GARCH models have been developed and Bauwens et al. (2003) provide an up to date and extensive survey of these.

In energy markets daily variations in futures prices have a significant amount of ‘noise’ that it is not useful for forecasting volatility and correlation in the future. Because of this the EWMA estimates can be too variable to be of much use in forecasting. They are based on all price movements even those fluctuations that an analyst would wish to ascribe to noise: hence the estimates shown in Figures 3 and 4 jump around too much. For this reason, the orthogonal GARCH model introduced by Alexander (2001a, 2002) is an attractive choice for generating covariance matrices for energy futures. The model is based on GARCH volatility forecasts of a number of principal components in the system. The number of principal components chosen depends on how much of the variation is considered to be irrelevant ‘noise’—i.e. movements that are not important for the perspective of forecasting correlation. Often just the first two or three principal components are required in a large system of highly correlated returns. It is therefore ideally suited to term structures such as energy futures, where principal component analysis is a useful tool that is commonly applied to reduce dimensionality (see Panel 1).

To illustrate the OGARCH model we have estimated univariate GARCH(1,1) models for the first two principal components from the analysis in Panel 1, and hence obtained the full $11 \times 11$ covariance matrix of variances and covariances of all maturity futures in the system. Each covariance matrix contains 55 volatilities and correlations and the model generates term structure covariance matrix forecasts for every day over the next year (or more, if forecasts have not reached their long term average level by then). Moreover these 55 term structures change every time the model is re-estimated, so from just two principal components a remarkably rich structure of correlation forecasts is generated. Figures 5 - 8 show just some
of the 1 day forecasts, i.e. those relating to just the 2mth, 6mth and 12mth futures in each of the two systems.

<Figures 5 – 8 here>

For crude oil the volatility of the 2mth future is consistently 2 – 2.5% higher than that of the 6mth future and 3 – 3.5% higher than that of the 12mth future (see also Table 1 for a comparison of long term average volatilities). Common peaks in all volatility series corresponding to large upwards shifts in the futures term structure are associated with major political and economic crises and these are accompanied by a general increase in correlation. Figure 6 shows that the OGARCH 2mth – 6mth correlation is relatively stable, since it is based only on parallel shift and tilt movements in the term structure (see Panel 1). However the stability of the OGARCH correlation estimates decreases, along with its average level, as the difference in maturity of the futures increases.

Figures 7 and 8 illustrate some of the OGARCH volatilities and correlations for the natural gas futures. Here the 2mth future responds most to variations in supply and demand and is thus more volatile than the longer dated futures. On average its volatility is almost double that of the 6mth future and more than double that of the 12mth future (see also Table 1). Peaks in volatility are seasonal and these are common to all futures though most pronounced in the near term futures. From the year 2000 there has been a general increase in the level of volatility and at the same time the correlation between futures of different maturities also increased. Consequently a positive association between volatility and correlation is more obvious in the natural gas futures than in the crude oil system.

We now compare the OGARCH and EWMA estimates of the same correlation. The OGARCH and EWMA volatilities shown in Figures 9 and 10 and the correlations shown in Figure 11 are all very close to each other. But in Figure 12 the OGARCH correlations for natural gas 2 month – 6month futures are much higher and more stable than the EWMA correlations. This is because, being based on only the first two principal components, the OGARCH correlations are based only on systematic trend and tilt movements in futures prices, the other movements being ignored because they are attributed to ‘noise’. From the PCA results in Panel 1 we know that the OGARCH model with two components is modelling only 87.3% of the variation. The addition of one more principal component to the OGARCH model might be advisable, but a fourth principal component may lead to correlation forecasts that are too unstable.

<Figures 9 – 12 here>

Basket hedging is a very uncertain activity in the natural gas markets because of the unstable correlations between futures of different maturities. Some producers may be unwilling to hedge at all, but there are still many end-users, speculators and arbitrage traders that do hedge using options on baskets of futures. Given the large and growing demand for these and other correlation dependent gas derivatives,
forecasting correlations has become an important activity. We have demonstrated how the OGARCH model provides a very useful tool for cutting through the ‘noise’ in these forecasts.

**Cross Market Correlations**

There are many products on related energy markets that are tailor made for end-users and producers alike. Energy producers that purchase both crude oil and gas will hedge revenues with basket options that are cheaper than buying options on individual futures. ‘Best of’ options allow end users to purchase energy supply at either the gas price or the oil price which ever is better, and in these options price depends on correlation through the volatility of the relative price. Long term swaptions and options on related markets diversify the risks and may be preferred to hedging all costs with derivatives based on a single market. And currency protected products allow the purchaser to hedge the foreign exchange risk.

The prices of these products depend on cross market correlation forecasts and are very difficult to price when correlation is unpredictable. For instance, a basket option on crude oil and natural gas futures should be cheaper than buying two separate options on each underlying because the basket volatility will be less than the sum of individual volatilities unless the underlyings are perfectly correlated. However, the price of a basket option can vary a lot over time if the correlation is unstable.

The results in Table 3 indicate that crude oil and natural gas futures price correlations are very close to zero in ‘normal’ markets and that it is only for the more unusual large price jumps that the correlation becomes significant. On the whole their correlation is low, so basket options on natural gas and crude oil should be relatively cheap. However amongst other factors, differences in settlement dates and procedures across different markets can produce highly unstable short term correlations.

Figure 13 shows that the EWMA correlation between the 2mth futures on crude oil and natural gas has increased during the last few years. However it remains very variable. The correlations shown are based on a very high value for the smoothing constant ($\lambda = 0.99$) and even these can halve or double in the space of a few weeks. So even though they may be cheaper than individual options, market prices of basket options can be highly variable.

<Figure 13 here>
MODELLING NON-NORMAL RETURNS

The use of correlation as a measure of dependency is only appropriate when the multivariate distribution of returns is elliptical and often it is assumed to be normal. But even the marginal returns distributions of crude oil and natural gas futures are highly non-normal. Table 1 summarises the first four moments of the empirical marginal daily returns densities of all maturity futures estimated over the entire sample from 4th January 1993 until 20th November 2003.

A highly significant negative skewness and leptokurtosis is evident in all maturities.ii Whilst these higher moments display no distinct pattern with respect to maturity, both the mean and the volatility of returns decrease with maturity. Clearly the assumption of normality is violated and therefore the use of a single correlation estimate to summarise returns dependency is inappropriate for these markets.

Normal Mixture Densities

In this section we examine the use of normal mixture densities to provide an improved description of the dependent behaviour of returns. A normal mixture density is a probability weighted sum of normal density functions. Thus a mixture of \( n \) normal densities \( \phi_1(x), \ldots, \phi_n(x) \) is the density

\[
p_1\phi_1(x;\mu_1,\sigma_1^2) + \ldots + p_n\phi_n(x;\mu_n,\sigma_n^2)
\]

where \( p_1 + \ldots + p_n = 1 \).

The mean of the mixture distribution is just the average of the means of the component densities and the variance of the mixture consists of two parts, the average of the variances and the variance of the means.

A mixture of normal densities with identical means but different variances gives a symmetric leptokurtic mixture density where the variance is just the average of the component variances. Table 1 indicates that the excess kurtosis in crude oil and natural gas futures is far more significant than the skew, and to simplify the model for the purpose of illustration we shall model marginal returns as mixtures of normal components with identical means. In this case we refer to the mixture as a normal variance mixture density. Also, for transparency and ease of interpretation, only two components are used in the mixture model for each future; this still provides a richer correlation structure than a multivariate normal distribution for futures returns. The interested reader will find McLachlan and Peel (2000) useful to extend these ideas to the more general case.

Setting both means equal to the sample mean, illustrative estimates for the three parameters of the variance mixture of two normal densities for marginal returns can be obtained using simple moment and percentile matching algorithms.xiii For example, for the 2mth crude oil future the parameter estimates are
as follows: the probability of a high volatility component $p = 0.25$, the volatility of that component $\sigma_H = 31\%$ and the volatility of the other component $\sigma_L = 10.56\%$. The density with the higher variance is associated with the observations in the tail. Hence with $p = 0.25$ it is the observations in the 12.5% tails of the density that determine the component with the higher volatility of 31% and the observations in the central 75% ‘core’ of the distribution determine the component with volatility 10.26%. The normal variance mixture representation also has a behavioural interpretation, stemming from the research of Epps and Epps (1976) amongst others. If traders have different perceptions of risk and returns according to which they form expectations and trade, then heterogeneous volatility components can be present in the same market. Here the probability of the high volatility component is capturing the market assessment of the likelihood of a price jump in either direction, and correspondingly a period of high volatility $\sigma_H$. In Table 2 the probability of a price jump has been fixed at that estimated from the normal variance mixture model for nearest future, and then the high and low volatility components for the other futures are estimated.

<Table 2 here>

**Multivariate Normal Mixture Densities**

Although the variance mixture of two normal densities is certainly not the best fitting parametric form – the generalised hyperbolic density will always give a closer fit to the data – it has a pleasing and intuitive interpretation as we have seen above. Moreover it can be extended easily into higher dimensions. For instance, the bivariate normal variance mixture density may be written:

$$g(x_1, x_2) = p_1 p_2 \Phi(x_1, x_2; V_1) + (1 - p_1) p_2 \Phi(x_1, x_2; V_2) + p_1 (1 - p_2) \Phi(x_1, x_2; V_3) + (1 - p_1)(1 - p_2) \Phi(x_1, x_2; V_4)$$

where $\Phi$ denotes the bivariate normal density function and the covariance matrices $V_i$ correspond to ‘tail’, ‘core-tail’, ‘tail-core’ and ‘core’ components for $i = 1, 2, 3, 4$ respectively. More generally a multivariate normal variance mixture in which each marginal is a variance mixture of two normal components may be written as a sum of $m = 2^n$ multivariate normal densities:

$$g(x) = p_1 p_2 \ldots p_n \Phi(x; V_1) + (1 - p_1) p_2 \ldots p_n \Phi(x; V_2) + \ldots + (1 - p_1)(1 - p_2) \ldots (1 - p_n) \Phi(x; V_m)$$

where $x = (x_1, x_2, \ldots, x_n)$, $\Phi$ denotes the $n$-variate normal density function and the $n \times n$ covariance matrices correspond to different tail-core components.

For the systems of crude oil and natural gas futures (even month maturities only) Table 3 presents just the ‘core’ and ‘tail’ correlation matrices based on the entire sample. The pair-wise ‘core’ correlations have been calculated on all data points that lie within the rectangular area $\{X_L < X < X_U \text{ and } Y_L < Y < Y_U\}$ where $X$ and $Y$ are the futures returns and the subscripts “L” and “U” refer to the lower and upper $(1 - p/2)\%$-iles.
of the empirical density. The remaining points are used to estimate the tail correlations. The tail is defined taking $p = 0.25$ for crude oil and $p = 0.275$ for natural gas, to be consistent with the normal mixture densities in Table 2. The number of data points in the tail depends on the two maturities (very approximately, about 750 data points lie in each tail).

In many financial markets correlations are found to be more significant in the tail of the joint density of asset returns than in the central, core part of the joint density (for instance, see Alexander and Pezier, 2003). That is, large price movements are more highly (positively or negatively) correlated than small ones, and ‘stress’ correlation matrices are normally constructed to reflect this stylised fact. Examination of Table 3 shows that this is also the case in the crude oil market. The tail correlations are consistently higher than the core correlations and in many, but not all, cases the difference is significant at 1%. However, the opposite is the case in the natural gas market: correlations are consistently and significantly lower in the tail than in the core. The reason for this may lie in the storage constraints: large price movements in a future of any given maturity may be associated with an expected surplus or deficit in supply at the time corresponding to the maturity of the future, with futures of different maturities being less affected. In the lower part of Table 3 the cross correlations between the natural gas futures and the crude oil futures of maturity 2, 4 and 6 months are indicated in italics. Notice that only the tail cross correlations are significant. For smaller price movements all cross correlations are insignificantly different from zero.

Combining results such as those in Table 3 (and similar results for ‘tail-core’ and ‘core-tail’ type correlation matrices) with the marginal normal variance mixture density parameter estimates such as those in Table 2 gives a crude representation of the multivariate normal variance mixture density. We have also simplified the problem by assuming a symmetric data generating process. However the ideas generalize easily to the asymmetric case by allowing the component normal densities to have different means. Clearly, given the significant skew in crude oil and natural gas markets, this should be done before models of this type are used for risk management or trading. In practice, a general multivariate normal mixture density can be estimated via the expectations maximization (EM) algorithm (see McLachlan and Peel, 2000). Applications include the computation of VaR for portfolios of leptokurtic assets (see Alexander, 2001b) and the arbitrage free pricing of European spread options (see Panel 2).

**Time-Varying Multivariate Normal Mixture Densities**

The EM algorithm can provide an accurate representation of the unconditional multivariate density as a normal mixture, but it has some limitations. First, it must be based on a very long sample of daily data so that sufficient points are observed in the tails. Secondly, there is no term structure in the volatility and correlation forecasts. They are only in-sample estimates over a very long historic period and if used in forecasting the implicit assumption is again that of constant volatility and correlation. Proper term structures for forecasts of the volatility and correlation components in a time-varying multivariate normal
mixture model can, however, be generated by combining the orthogonal GARCH model with the normal mixture GARCH model of Alexander and Lazar (2003). Though beyond the scope of this chapter, the time-varying ‘core’ and ‘tail’ component covariance matrix forecasts provided by the normal mixture orthogonal GARCH model have useful applications to multi-asset option pricing with the lognormal mixture forward rate model, as we shall see in the next section.

**Multi-Asset Option Valuation**

Many options in the oil and gas markets are complex path dependent structures based on a system of correlated underlying instruments that are designed to protect the costs to energy consumers over a long cash flow period. For example, a 1m barrel call at 30$ on WTI over 1 year has a pay-off each month of (‘spot’ – 30)$m if positive and zero otherwise, where ‘spot’ is really an average spot price over several days during the preceding month. The value of this option depends on a sequence of 12 pay-offs, where the $n$th pay-off depends on the evolution of the spot $n$ months in the future.

Typically the values of path dependent options such as the 1m barrel call are derived using the multivariate lognormal diffusion model for forward rates introduced by Brace, Gatarek & Musiela (1997) and Miltersen, Sandermann & Sondermann (1997). The lognormal forward rate model may be written

$$\frac{df_i(t)}{f_i(t)} = \mu_i(t) \, dt + \sigma_i(t) \, dW_i \quad [i = 1, \ldots, m; \ 0 < t < t_i]$$

where $dW_1, \ldots, dW_m$ are correlated Brownian motions, i.e. $E[dW_i \, dW_j] = \rho_{ij}(t) \, dt$ and the drift term in each diffusion is given by

$$\mu_i(t) = \sigma_i(t) \sum_{j=m(t)}^i \frac{g_{ij}(t) \sigma_j(t) f_j(t)}{1 + f_j(t)}$$

**Calibration of the Lognormal Forward Rate Model**

Calibration of the model requires using current market and/or historical data to estimate the parameters $\sigma_i(t)$ and $\rho_{ij}(t) \ [i, j = 1, \ldots, m]$. That is, to estimate the covariance matrices of the forward rates relevant to each pay-off. Although these covariance matrices have many parameters, parsimonious parameterisations have been suggested by several authors. For example Rebonato and Joshi (2001) parameterize each correlation matrix with a ‘circulant’ form.
so that $\rho_{ij} = |I - i|$.

$$
\begin{pmatrix}
1 & \rho & \rho^2 & \ldots & \rho^{2m-1} \\
\rho & 1 & \rho^2 & \ldots & \rho^{2m-2} \\
\rho^2 & \rho & 1 & \ldots & \rho^{2m-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\rho^{2m-1} & \ldots & \ldots & \ldots & 1
\end{pmatrix}
$$

Authors such as Rebonato (1999a), Rebonato and Joshi (2001) Hull and White (1999, 2000) and Logstaff, Santa-Clara and Schwartz (1999) advocate use of market data to calibrate the volatilities in the covariance matrices but base correlations on historical data because they are more stable than those calibrated directly to option smiles. In particular, they reduce the rank of forward rate correlation matrices by setting all but the largest eigenvalues to zero, and use historic data to calibrate the their eigenvectors. The implication of zeroing eigenvalues is a transformation of the lognormal forward rate model where each forward rate is driven by a few orthogonal factors that are derived from a principal component analysis.

For instance, given the results in Panel 1, the crude oil forward rate dynamics may be expressed in terms of only two uncorrelated stochastic processes that are common to all forward rates:

$$
df_i(t)/ f_i(t) = \mu_i(t) \, dt + \lambda_{i,1}(t) \, dZ_1 + \lambda_{i,2}(t) \, dZ_2$$

where $dZ_1$ and $dZ_2$ are uncorrelated Brownian motions. Then

$$
\sigma_i(t) \, dW_i = \lambda_{i,1}(t) \, dZ_1 + \lambda_{i,2}(t) \, dZ_2
$$

so $\sigma_i(t)^2 = \lambda_{i,1}(t)^2 + \lambda_{i,2}(t)^2$ and $\rho_{ij}(t) = [\lambda_{i,1}(t) \lambda_{j,1}(t) + \lambda_{i,2}(t) \lambda_{j,2}(t)] / \sigma_i(t) \sigma_j(t)$. Thus the forward rate covariances are completely determined by two volatility ‘components’ for each forward rate, $\lambda_{i,1}(t)$ and $\lambda_{i,2}(t)$.

If implied forward rate volatilities are calibrated to market prices of options, then the volatility components can be determined from a principal component analysis of the empirical correlation matrix of the m forward rates: denote by $\Lambda_1(t)^2$ and $\Lambda_2(t)^2$ the two largest eigenvalues of this matrix and denote their eigenvectors by $w_1$ and $w_2$ (see Panel 1). Set

$$
M(t) = \sigma_i(t)/ [w_{i1}^2\Lambda_1(t)^2 + w_{i2}^2\Lambda_2(t)^2]
$$

Then the volatility components are given as
\[ \lambda_{i,1}(t) = M(t)w_{i1} \Lambda_1(t) \quad \text{and} \quad \lambda_{i,2}(t) = M(t)w_{i2} \Lambda_2(t). \]

**Extensions of the Lognormal Forward Rate Model**

One obvious limitation of the lognormal forward rate model is that empirical distributions of futures returns are too skewed and heavy tailed to be normal. For instance, the results in Table 1 showed that historical empirical futures returns in the crude oil and natural gas markets have highly leptokurtic and skewed unconditional densities.

The inappropriate assumption of lognormal price densities is an important reason why implied volatility smile/skew effects are observed in most options markets. To explain at least part of the smile/skew with a univariate option pricing model, some authors introduce a stochastic volatility process as in Hull and White (1987) or Heston (1993). Time variation in instantaneous volatility gives terminal price densities that are not lognormal. However, although stochastic volatility models can be ‘smile consistent’, stochastic volatility alone is not sufficient to ensure multi-asset option prices are consistent with implied correlation ‘frowns’. For this, the correlation structure itself must be enriched (see Panel 2).

Following Brigo and Mercurio (2001, 2002) a multi-dimensional extension by Brigo, Mercurio and Rapisarda (2003) provides smile consistent arbitrage free prices for multi-asset options. Remaining in a complete markets setting, they broaden the assumption of lognormality in the forward rate model to assume lognormal mixture dynamics for each of the underlying forward rates. That is, the risk neutral density of the log prices of the forwards is assumed to be a multivariate normal mixture density. The beauty of this approach is that the local volatility function is uniquely determined by the risk neutral price density. No other local volatility parameterisation offers such an intuition.\textsuperscript{v} A limitation is that – in common with all local volatility models – it says nothing about the movements in the smile surface over time. It may be a ‘complete markets model’ but the model itself is incomplete because if the smile moves during the hedging period there will be residual hedging uncertainty.

Using a two dimensional lognormal mixture diffusion model, Alexander and Scourse (2003) obtain risk neutral smile consistent and correlation frown consistent spread option values. Joint densities of the two underlying assets in the spread are assumed to be bivariate normal mixtures, similar to those introduced in the previous section of this chapter. More details are given in Panel 2.

**ACKNOWLEDGEMENTS**

Many thanks to Vince Kaminski for providing the data and to my PhD student Anca Dimitriu for invaluable help with the empirical work.
**Panel 1: Principal Component Analysis of Crude Oil and Natural Gas Futures**

The objective of principal component analysis (PCA) is to reduce dimensionality by taking only the first $m$ principal components in each of the representations

$$X_i = w_{i1}P_1 + w_{i2}P_2 + \ldots + w_{in}P_n$$

for $i = 1, \ldots, n$. Here $X$ denotes a time series of returns in a correlated system of $n$ returns and the ‘principal components’ $P_1, \ldots, P_n$ are also time series, but are uncorrelated with each other. That is, they are orthogonal linear transformations of the original time series $X_1, \ldots, X_n$. They will be orthogonal when the $n \times n$ ‘factor weights’ matrix $W = (w_{ij})$ is taken as the matrix of eigenvectors of $XX'$ where the $T \times n$ matrix $X$ has columns $X_1, \ldots, X_n$.

The principal components are ordered so that

$$P_i = w_{i1}X_1 + w_{i2}X_2 + \ldots + w_{in}X_n$$

and $w_i = (w_{i1}, \ldots, w_{in})$ is the eigenvector corresponding to the largest eigenvalue of $XX'$; similarly $P_2$ is generated by the eigenvector belonging to the second largest eigenvalue of $XX'$ and so forth. The proportion of variation captured by each principal component is given by the ratio of its eigenvalue to the sum of all eigenvalues. Often we normalize the returns to have mean zero and variance 1, so $XX'$ is the returns correlation matrix and the sum of its eigenvalues is $n$. Alternatively without normalization $XX'$ is proportional to the covariance matrix and the eigenvectors will be dominated by the more volatile returns in the system.

PCA is a very useful technique for reducing the dimensions in highly correlated systems, such as futures of different maturities on the same underlying commodity. There are only a few independent sources of variation so most of this can be explained by movements in just a few principal components. It is common to take only the first two or three principal components to represent the system, because between them they can explain a very large fraction of the variation.

Moreover, when the system is maturity ordered the first few principal components have a natural interpretation. If there is a high correlation in the system as a whole the elements of the first eigenvector will be roughly equal, so a movement in the first principal component represents a roughly parallel shift in all variables; the elements of the second eigenvector will have values that increase (or decrease) more or less linearly with maturity, so the second eigenvector represents a ‘tilt’ in the term structure; similarly the third component represents a change in the curvature of the term structure.

*Principal Component Analysis of 2mth – 12mth Futures*

<table>
<thead>
<tr>
<th>Crude Oil</th>
<th>Natural Gas</th>
</tr>
</thead>
</table>

The table above reports the results of applying PCA to the correlation matrix of (i) crude oil returns and (ii) and natural gas returns to all futures between 2 months and 12 months over the whole sample. Thus there are 11 variables in each system. In crude oil the first principal component (generated by the eigenvector $w_1$) explains $10.732/11 = 97.6\%$ of the variation. The first eigenvector is more or less constant, so the first principal component captures a parallel shift in all maturities. Movements in the second component account for a further $2.2\%$ of the variation and this component captures a ‘tilt’ in the futures term structure, as can be seen from the downwards trending values of the second eigenvector. The higher principal components are not important since together they explain only $0.2\%$ of the movements and these can be ascribed to ‘noise’. Thus there are only two orthogonal sources of information that affect crude oil futures prices and a representation with two principal components is adequate (with only one the futures would be assumed to have perfect correlation).

In the natural gas system the principal components again have the ‘trend – tilt – curvature’ interpretation but the system is less correlated as a whole than the crude oil system so more components are required to capture most of the variation. The trend component explains only $8.696/11 = 79\%$ of the variation, the tilt a further $8.3\%$ and the third and fourth components explain $5.3\%$ and $3.4\%$ of the variation respectively. With a four component representation of the system, the remaining $4\%$ of the variation would be attributed to ‘noise’.

The reduction in dimensionality achieved by a principal component representation can greatly facilitate calculations. Transformations are applied to the principal components and then the factor weights are used to relate these transformations to the original system. Thus PCA has many important applications, for instance to construct large GARCH covariance matrices, for pricing derivative products and in scenario analysis.
As spread risk inherently involves correlation risk, spread options are an important tool in hedging and trading correlation. Hedging with spread options can involve buying and selling calls and puts with European, Bermudan or American exercise conditions. In general no analytic formulae exist for the pricing of these multivariate contingent claims and, despite considerable research in this area, there is no universally prevalent pricing framework. A very detailed and informed survey of recent research on the valuation of European spread options is given in Carmona and Durrleman (2003a). As well as highlighting the theoretical and computational problems associated with pricing equity and interest rate spread options, Carmona and Durrleman place particular emphasis on commodity spread options and include several example applications to energy markets.

It is unrealistic to base spread option values on the assumption that the spread itself has dynamics governed by a univariate diffusion process, because in that case the distribution of the spread would be independent of the correlation between the underlying assets. Most of the European spread option valuation models that have been considered in detail in the academic literature have adopted the assumption of two correlated lognormal diffusions (Ravindran 1993, Shimko 1994, Kirk 1995, James, 2002 and others) and we call this the ‘2GBM’ model. An important extension of the 2GBM model is to include stochastic volatility, as in Dempster and Hong (2000) so that the individual asset price dynamics are consistent with their implied volatility smiles.

Implied correlation is the correlation that is implicit in the 2GBM model for its price to be consistent with an observed market price of a spread option. It varies with the moneyness of the spread option, just as the Black-Scholes implied volatility varies with the moneyness of a single asset option. In a symmetric volatility smile, such as those commonly observed in currency option markets, implied volatility of in-the-money (ITM) and out-of-the-money (OTM) options is greater than the at-the-money (ATM) implied volatility. Hence the term ‘volatility smile’. It arises because traders perceive a greater probability of large price changes than is assumed in the BS model. The perceived leptokurtic asset return density leads to market prices of ITM and OTM options that are greater than BS prices and, all other variables being fixed in the BS model, the only way that the BS model can explain these market prices is to increase the volatility. Because of the positive relationship between volatility and option price, increasing the volatility will increase the BS price.

Similarly, if traders perceive a joint density for asset returns that has heavier tails than the bivariate normal, then market prices of ITM and OTM spread options will be greater than those based on the 2GBM assumption. The correlation between asset returns has a negative relationship with the price of a spread option. That is, the lower (or more negative) the correlation, the more valuable is the spread option. Therefore, to match these higher market prices within the 2GBM framework, the implied correlation for
OTM and ITM spread options must be lower than the implied correlation used for ATM spread options. Hence it is more appropriate to call the variation of implied correlation with moneyness a correlation ‘frown’ rather than a correlation ‘smile’.

The existence of a correlation ‘frown’ implies that tail probabilities are underestimated in the 2GBM framework. That is, the 2GBM assumption is not consistent with the observed market prices of OTM call and put spread options because asset returns have leptokurtic densities. So, when we speak of a ‘frown consistent’ valuation model, we mean a model with leptokurtic asset price densities, with a single process correlation, and for which a correlation frown that is similar to the market frown appears in the 2GBM implied correlations that are ‘backed-out’ from model prices.

A deficiency of the 2GBM spread option valuation model is the simplicity of the correlation assumption. Even the addition of stochastic volatility only makes the model smile consistent, not frown consistent. Ignoring the correlation frown can lead to substantial mispricing, because it is likely that the probability of a large spread movement will be under-estimated when using a bivariate normal distribution for the log asset prices.

Few ‘frown consistent’ spread option pricing models have been developed, with two notable exceptions: the jump diffusion model of Carmona and Durrleman (2003b) and the bivariate normal mixture (BNM) model of Alexander and Scourse (2003). In the latter, arbitrage free option values are obtained as weighted sums of different 2GBM values, each based on two correlated lognormal diffusions, but the volatilities and correlation in each pair of lognormal diffusions is different. Thus, although the model has a single ‘overall’ correlation process, this process has a rich structure, for example it has both ‘core’ and ‘tails’ components.

The advantage of the BNM model for spread options is that their values are consistent with both the volatility smiles for each asset and the implied correlation ‘frown’. Alexander and Scourse (2003) show that when the joint log price density is a mixture with four bivariate normal components – such as that explained in this chapter – the ‘core’ correlations determine the height of the correlation frown, whilst the ‘tail’ correlations influence its steepness. Differences between BNM and 2GBM spread option values, which can be attributed to six second order option sensitivities, are greatest for OTM option values and are usually positive but can be negative (i.e. BNM spread option values can be less than 2GBM values) for near ATM options.
REFERENCES


### Table 1: Moments of Marginal Returns Densities

<table>
<thead>
<tr>
<th>Maturity</th>
<th>2mth</th>
<th>3mth</th>
<th>4mth</th>
<th>5mth</th>
<th>6mth</th>
<th>7mth</th>
<th>8mth</th>
<th>9mth</th>
<th>10mth</th>
<th>11mth</th>
<th>12mth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Crude Oil</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Annual Return</td>
<td>2.03%</td>
<td>1.95%</td>
<td>1.85%</td>
<td>1.77%</td>
<td>1.69%</td>
<td>1.62%</td>
<td>1.55%</td>
<td>1.49%</td>
<td>1.44%</td>
<td>1.39%</td>
<td>1.36%</td>
</tr>
<tr>
<td>Average Volatility</td>
<td>13.67%</td>
<td>12.44%</td>
<td>11.60%</td>
<td>10.95%</td>
<td>10.44%</td>
<td>10.03%</td>
<td>9.69%</td>
<td>9.42%</td>
<td>9.19%</td>
<td>8.98%</td>
<td>8.82%</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.25</td>
<td>-0.22</td>
<td>-0.26</td>
<td>-0.31</td>
<td>-0.32</td>
<td>-0.31</td>
<td>-0.29</td>
<td>-0.28</td>
<td>-0.27</td>
<td>-0.27</td>
<td>-0.28</td>
</tr>
<tr>
<td>Excess Kurtosis</td>
<td>4.11</td>
<td>3.41</td>
<td>3.81</td>
<td>4.07</td>
<td>4.06</td>
<td>3.84</td>
<td>3.61</td>
<td>3.42</td>
<td>3.26</td>
<td>3.17</td>
<td>3.11</td>
</tr>
<tr>
<td><strong>Natural Gas</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maturity</td>
<td>2mth</td>
<td>3mth</td>
<td>4mth</td>
<td>5mth</td>
<td>6mth</td>
<td>7mth</td>
<td>8mth</td>
<td>9mth</td>
<td>10mth</td>
<td>11mth</td>
<td>12mth</td>
</tr>
<tr>
<td>Mean Annual Return</td>
<td>4.70%</td>
<td>4.72%</td>
<td>4.64%</td>
<td>4.35%</td>
<td>4.31%</td>
<td>4.25%</td>
<td>4.20%</td>
<td>3.93%</td>
<td>3.64%</td>
<td>3.26%</td>
<td>3.41%</td>
</tr>
<tr>
<td>Average Volatility</td>
<td>21.93%</td>
<td>19.15%</td>
<td>16.47%</td>
<td>14.46%</td>
<td>13.14%</td>
<td>12.52%</td>
<td>12.21%</td>
<td>11.97%</td>
<td>11.52%</td>
<td>10.99%</td>
<td>10.61%</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.27</td>
<td>-0.93</td>
<td>-0.51</td>
<td>-0.30</td>
<td>-0.13</td>
<td>0.02</td>
<td>-0.08</td>
<td>-0.39</td>
<td>-0.59</td>
<td>-0.60</td>
<td>-0.67</td>
</tr>
<tr>
<td>Excess Kurtosis</td>
<td>4.55</td>
<td>11.21</td>
<td>4.44</td>
<td>3.33</td>
<td>3.29</td>
<td>4.05</td>
<td>5.19</td>
<td>6.63</td>
<td>7.86</td>
<td>6.65</td>
<td>8.02</td>
</tr>
</tbody>
</table>

### Table 2: Normal Mixture Representations of the Marginal Returns Densities

<p>| Probability (p) of High Volatility Component | 0.25 | 0.275 |
| High Volatility ( \sigma_H ) | 23.78% | 18.13% | 14.23% | 39.92% | 21.52% | 20.08% |
| Low Volatility ( \sigma_L ) | 7.78% | 5.98% | 6.02% | 10.85% | 7.91% | 1.58% |</p>
<table>
<thead>
<tr>
<th></th>
<th>m2</th>
<th>m4</th>
<th>m6</th>
<th>m8</th>
<th>m10</th>
<th>m12</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tail</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m4</td>
<td>0.990</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m6</td>
<td>0.975</td>
<td>0.996</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m8</td>
<td>0.959</td>
<td>0.988</td>
<td>0.997</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m10</td>
<td>0.941</td>
<td>0.975</td>
<td>0.990</td>
<td>0.998</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>m12</td>
<td>0.918</td>
<td>0.959</td>
<td>0.980</td>
<td>0.992</td>
<td>0.998</td>
<td>1</td>
</tr>
<tr>
<td><strong>Core</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m4</td>
<td>0.958</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m6</td>
<td>0.900</td>
<td>0.976</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m8</td>
<td>0.837</td>
<td>0.931</td>
<td>0.984</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m10</td>
<td>0.781</td>
<td>0.883</td>
<td>0.954</td>
<td>0.988</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>m12</td>
<td>0.729</td>
<td>0.830</td>
<td>0.914</td>
<td>0.962</td>
<td>0.989</td>
<td>1</td>
</tr>
<tr>
<td><strong>Natural Gas</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tail</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m4</td>
<td>0.859</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m6</td>
<td>0.759</td>
<td>0.858</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m8</td>
<td>0.705</td>
<td>0.719</td>
<td>0.830</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m10</td>
<td>0.644</td>
<td>0.700</td>
<td>0.707</td>
<td>0.793</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>m12</td>
<td>0.619</td>
<td>0.641</td>
<td>0.713</td>
<td>0.641</td>
<td>0.739</td>
<td>1</td>
</tr>
<tr>
<td><strong>Core</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m4</td>
<td>0.894</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m6</td>
<td>0.797</td>
<td>0.924</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m8</td>
<td>0.726</td>
<td>0.841</td>
<td>0.941</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m10</td>
<td>0.676</td>
<td>0.772</td>
<td>0.865</td>
<td>0.946</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>m12</td>
<td>0.662</td>
<td>0.701</td>
<td>0.789</td>
<td>0.874</td>
<td>0.937</td>
<td>1</td>
</tr>
<tr>
<td><strong>Gas – Oil</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tail</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m4</td>
<td>0.859</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m6</td>
<td>0.759</td>
<td>0.858</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m2</td>
<td>0.237</td>
<td>0.272</td>
<td>0.257</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m4</td>
<td>0.263</td>
<td>0.281</td>
<td>0.263</td>
<td>0.990</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>m6</td>
<td>0.264</td>
<td>0.283</td>
<td>0.266</td>
<td>0.975</td>
<td>0.996</td>
<td>1</td>
</tr>
<tr>
<td><strong>Core</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m4</td>
<td>0.894</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m6</td>
<td>0.797</td>
<td>0.924</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m2</td>
<td>0.091</td>
<td>0.088</td>
<td>0.092</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m4</td>
<td>0.088</td>
<td>0.075</td>
<td>0.095</td>
<td>0.958</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>m6</td>
<td>0.078</td>
<td>0.067</td>
<td>0.081</td>
<td>0.900</td>
<td>0.976</td>
<td>1</td>
</tr>
</tbody>
</table>
i So the joint density is assumed stationary – see Alexander (1999) for more details

ii For reasons of length the treatment of all the models described in this section has been greatly simplified. Readers wishing to gain a more thorough understanding of the mathematical concepts used here are referred to Alexander (2001b) where there are also many illustrative examples. The moving average models are easy to implement in Excel. Both moving average and normal GARCH models have in-built routines in all the major statistical software packages such as E-Views, TSP, RATS, SAS and S-Plus. Principal component routines are also available in these packages but this can also be done in Excel (see the software CD with Alexander, 2001b). By far the best and most flexible econometric package for GARCH modelling is OX: both the normal mixture GARCH and the orthogonal GARCH models are easily programmed in this environment. For more details see http://www.timberlake.co.uk

iii For example, if the volatility is 50% then the standard deviation of 10-day returns is 0.1 assuming 250 risk days per year. If the 1% standard normal critical value of 2.33 is applied, then the 1% 10 day VaR is 2.33 x 0.1 per dollar invested. So for a portfolio with a nominal of 1m$ the 1% 10 day VaR is 233,000$.

iv Consider the example of writing an option with a maturity of one year or more when no similar options are priced in the market. Then a statistical forecast of long-term volatility is required and the equally weighted average estimate based on all returns over the last few years is one classical method of obtaining such a forecast. Having said this, the trader’s prior beliefs about the future price movements should also play an important role in updating this historical estimate into a posterior forecast of the volatility.

v Some practitioners apply a lower value of $\lambda$ (such as 0.94, the ‘RiskMetrics’ value) for shorter term forecasts than for longer term forecasts, but this is very ad hoc.

vi Assuming zero means is simpler and there is no convincing empirical evidence that this degrades the quality of correlation estimates and forecasts in financial time series. This assumption is used for the RiskMetrics volatility and correlation forecasts, including the EWMA forecasts given below. RiskMetrics EWMA covariance matrices require the use of the same value of the smoothing constant for all assets and early versions of RiskMetrics produced covariance matrices that were sometimes non-positive definite.

vii Strictly speaking, it is the unexpected return from the conditional mean equation.

viii This is important because otherwise the model would have no long-term level of volatility $[\omega(1 - \alpha - \beta) = 0/0]$. Clearly we must have $\alpha + \beta < 1$ otherwise the long term volatility would be infinite.

ix Another advantage of GARCH models is that parameters are estimated in a optimal manner, using maximum likelihood estimation, so there is no need to fix parameters arbitrarily as one does for the smoothing constant $\lambda$ in a EWMA model, or the number of observations $n$ in the ‘historic’ model.

x For an exact method of determining how many components to use, the eigenvalues of the correlation matrix can be compared with those of a random correlation matrix; see Plerou et. al. (2002).

xi The basket volatility is related to the volatility of individual options as

$$\sigma_{x+y} = \sqrt{[\sigma_x + \sigma_y]^2 - 2(1 - \rho)\sigma_x\sigma_y}$$

xii Approximate standard errors for skewness and excess kurtosis estimates are $\sqrt{6/n}$ and $\sqrt{24/n}$ respectively where $n$ is the sample size. With our sample of 2721 data points these are 0.047 and 0.094 respectively.
Such algorithms are not recommended when fitting normal mixture densities. The objective in moment matching is a complex non-linear function with multiple local optima, requiring sophisticated numerical methods. The best algorithm for fitting normal mixture densities is the ‘Expectations Maximization’ (EM) algorithm – see McLachlan and Peel (2000).

In the NMOGARCH model univariate normal mixture GARCH models are estimated on the first few principal components of the system.

For instance, one can parameterise the local volatility using Hermite polynomials, hyperbolic functions, or whatever you regard as giving best fit to the smile, and once calibrated the local volatility function may be used to price path dependent options using numerical methods. However, this is a purely pragmatic approach – one has no idea about the properties of the risk neutral price density.