A Primer on the Orthogonal GARCH Model

Professor Carol Alexander
ISMA Centre, The Business School for Financial Markets,
University of Reading

First Version: April 1999
This Version: March 2001

Keywords: *Principal component analysis, covariance matrix, orthogonal factor model, Value-at-Risk, generalised autoregressive conditional heteroscedasticity, exponentially weighted moving average.*

1. Introduction

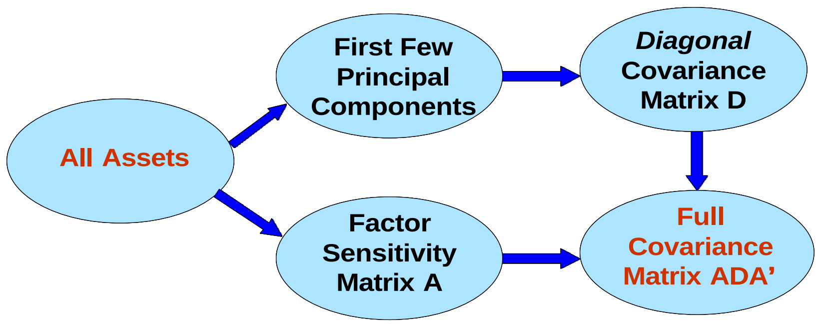
The generation of large, positive semi-definite covariance matrices has presented a great challenge to finance practitioners for many years. Since the 1996 Amendment to the 1988 Basle accord, where the principals of internal models for the calculation of market risk capital were outlined, it has been a major problem to generate the covariance matrices that are necessary to calculate firm-wide value-at-risk measures.[1] Large covariance matrices also have a major role to play in investment analysis, because portfolio risk is normally determined by the covariance matrix of all the assets in the portfolio. In very large portfolios a risk factor model may be employed, but it is still necessary to have a covariance matrix of all the risk factors of the portfolio. The need for large covariance matrices is not just confined to the middle office of a large investment bank. The traders in the front office also require these matrices to price and hedge their option portfolios.

Calculation of a large positive definite covariance matrix is a complex problem and often very simple measures of volatility and correlation are used in this covariance matrix. For example the RiskMetrics methodologies designed by JP Morgan use either equally weighted moving averages or exponentially weighted moving averages with the same smoothing constant for all returns. The RiskMetrics methods have limitations, and these have been outlined in (link to website).

The purpose of this paper is to show how orthogonal factor models can be used to simplify the process of producing these large covariance matrices on a daily basis. Its central idea is the orthogonal factorisation by principal component analysis of the assets or risk factors that the covariance matrix represents. The paper focuses on how these principal components can be used in conjunction with standard volatility estimation methods, such as exponentially weighted moving averages (EWMA) or generalised autoregressive conditional heteroscedasticity (GARCH), to produce large positive semi definite covariance matrices.

[1] Of the three methods that are in standard use (the covariance method for linear portfolios, and Monte Carlo and historical simulation) only the historical simulation method requires no covariance matrix. And even with historical simulation, covariance matrices are normally used in stress testing and scenario analysis. See Alexander (2001).

The method is computationally very simple: it takes the univariate volatilities of the first few principal components of a system of risk factors and the factor weights matrix of the principal components representation to produce a full covariance matrix for the original system. A schematic of the method is shown below. Here the matrix **A** is a matrix of re-scaled factor weights and the diagonal matrix **D** is a matrix of either GARCH or EWMA volatilities of the principal components.



The orthogonal method for generating covariance matrices has many advantages:

- The computational burden is much lighter when all of the k(k+1)/2 volatilities and correlations are simple matrix transformations of just 2 or 3 EWMA or GARCH variances;
- The matrices are always positive semi-definite;
- Very few constraints are imposed on the movements in volatility and correlation;[2]
- If the method is applied with EWMA the effective value of λ for a returns series is determined by its correlation in the system. There is no need to impose the same value of the smoothing constant on all returns, as there is in the RiskMetrics data sets.
- If the method is applied with univariate GARCH one can generate term structures of covariance matrix term structures that are mean reverting. That is, there is no necessity to apply the 'square-root-of-time' rule and assume that volatilities and correlations are constant. Instead the usual GARCH analytic formulae for computing the term structure of volatility and correlation are applied so that the n-day covariance matrix converges to the long term average as n increases;
- By using only the first few principal components to represent the system, the correlations estimates become more stable and less influenced by variation that would be better ascribed to 'noise' in the data.
- Data may be difficult to obtain directly, particularly on new issues or on financial assets that are not heavily traded. When data are sparse or unreliable on some of the variables in the system a direct estimation of volatilities and correlation may be very difficult. But if there is sufficient

[2] For example, it is not necessary to impose the constraint that all volatility and correlation estimates have the same persistence to market shocks. This constraint is necessary in the RiskMetrics EWMA matrices.

information to infer their factor weights in the principal component representation, their volatilities and correlations may be obtained using the orthogonal method;

In a highly correlated system only a few principal components are required to represent the system variation to a very high degree of accuracy. For example if 3 principal components were used to represent a term structure with n maturities, only 3 volatilities will need to be calculated: a far simpler method than calculating all of the n(n+1)/2 volatilities and correlations directly. The covariance matrix that is constructed using the principal components method is guaranteed to be positive semi-definite. Also by tailoring the number of principal components to capture only the main variations, rather than the small and insignificant movements that might better be ascribed to 'noise', correlation estimates will be more stable than if they were estimated directly.

The first part of the paper deals with the orthogonalization of risk factors in a multi-factor model using principal component analysis. Sections 2 and 3 cover the algebra of the method for generating a full nxn covariance matrix, which will always be positive semi-definite, from the volatilities of the principal components alone. The basic algebra is illustrated by an example that uses equally weighted moving average estimates of the volatilities of the principal components.

The next section extends the basic model to allow exponentially weighted moving average variances of the principal components. Exponentially weighted moving averages of the squares and cross products of returns are a standard method for generating covariance matrices. But a limitation of this type of direct application of exponentially weighted moving averages is that the covariance matrix is only guaranteed to be positive semi-definite if the same smoothing constant is used for all the data. That is, the reaction of volatility to market events and the persistence in volatility must be assumed to be the same in all the markets that are represented in the covariance matrix. A major advantage of the orthogonal factor method described here is that it allows exponentially weighted moving average methods to be used without this unrealistic constraint. In fact the smoothing constant that defines the exponential volatility of any particular asset or market will be given by its factor weights in the principal components representation. Put another way, the volatility persistence and market reaction of a particular returns series will not be the same as that of the other variables in the system, but instead it is related to its correlation with those variables.

Having applied the orthogonal method with exponentially weighted moving averages in section 4, it is a small step to replace exponentially weighted moving average variances with fully fledged GARCH variances as described in section 5. Multivariate GARCH models have been the subject of extensive academic research (in particular see Engle and Kroner, 1993). One of the many good reasons for this is that a GARCH model normally gives mean-reverting term structures of volatility and correlation with a simple analytic form. Bollerslev, Engle and Nelson (1994) provide a good review of most of the earlier literature on GARCH model and see Alexander (2001) for a review of the more recent work. There has been considerable research on different ways of parameterizing multivariate GARCH models so that the GARCH covariance matrices are positive definite. Unfortunately the computational aspects become more and more problematic as the dimension increases and at the moment there is no chance that multivariate GARCH models can be used to estimate directly the very large covariance matrices that are required to net all the risks in a large trading book. The importance of the method that is explained in this 'primer' is that it will allow large multivariate GARCH matrices to be generated from univariate GARCH models.

Engle (2000) describes the alternative methods for estimating multivariate GARCH covariance matrices from univariate GARCH models and shows that the orthogonal GARCH model performs extremely well according to three out of the four diagnostics that he has chosen for assessing the accuracy of correlation forecasts.

The orthogonal GARCH model is validated empirically using data on commodity futures, interest rates and equity indices. A main focus of section 5 is on the calibration of volatilities and correlations that are generated using the orthogonal GARCH model with those that are generated directly using standard multivariate GARCH parameterisations. These examples have been included to show that the model calibration will need much more care when the system is not so highly correlated. But although the initial calibration of the orthogonal GARCH model may require some care, once the model has been calibrated in this way it may be used on a daily basis without recalibration.

Orthogonal GARCH has a number of advantages over direct multivariate GARCH: Since only very few univariate GARCH models are required to generate the large covariance matrix, convergence problems of the optimization routines will be rare (whereas they are common place with the application of direct multivariate GARCH to large systems); There need be no constraints on dimensionality of the original system (whereas direct multivariate GARCH models can only really cope with single figure dimensions); The orthogonal GARCH method gives one the option of cutting out any 'noise' in the data that would otherwise make correlation estimates unstable; also the orthogonal method allows one to generate estimates for volatilities and correlations of variables in the system even when data are sparse and unreliable, for example in illiquid markets.

Section 6 shows how the orthogonal factor method can be applied to generate very large covariance matrices. It suggests how one should divide risk factors into different categories before the application of principal component analysis, and demonstrates how best to 'splice' together different blocks into a covariance matrix of the original large system.

This is a long paper but the main points are quite simple. First, in the world of financial markets where there is so much uncertainty, it makes sense to distill the important information into a few factors that influence all the variables in the system to a greater or lesser extent. Much of market variability can be just put down to 'noise' and models that do not know how to filter that out may lack robustness. And second, the orthogonal method allows one to use GARCH models to generate large covariance matrices for many types of systems, from equities or foreign exchange rates to all types of term structures.

This primer begins by showing how the orthogonal model may be applied with exponentially weighted moving average variance, but there are very good reasons to prefer GARCH models to exponentially weighted moving average models. Perhaps the most important reason is the convergence of GARCH volatility and correlation forecasts to their long-term average levels, whereas the exponentially weighted moving average model has a constant term structure of forecasts. The paper has been written in response to the many requests that I have received since first explaining these ideas (Alexander and Chibumba, 1996 and Alexander, 2000, 2001).[3] First and foremost it is written as a primer on the method, and so I have

[3] Zhuanxin Ding has recently pointed out to me that in his 1994 PhD dissertation with Rob Engle at UCSD he discussed about 23 different possible forms of multivariate GARCH model that are all guaranteed to be positive definite. The Principal Component Multivariate ARCH Model is listed as model #23, but it did not perform so well because in order to ensure strict positive definiteness Ding took all principal components in the model. The great advantage of the orthogonal GARCH model that is explained in this paper, is that only a few, uncorrelated key risk

provided the examples and supporting programs as free downloads from the Wilmott website (link). References within this paper will take you to the TSP programs with the zip file of data provided. A demo version of TSP is available free from their website (www.tspintl.com) but it has restricted memory capabilities so some of the example programs presented here could not be run on the demo version of TSP. A rudimentary excel add-in for orthogonal GARCH is available from www.chrisleigh.co.uk and commercial software for orthogonal GARCH is also available from Algorithmics Inc (www.algorithmics.com).

2. Principal Component Analysis

The ideas of this section are illustrated by generating covariance matrices for two sets of returns (a) a term structure of crude oil futures, and (b) a small set of French equities. Each of these systems is more or less correlated: the crude oil futures far more than the equities. If one were to generate a covariance matrix for each system by applying some simple weighted average measures of variance and covariance to the returns, it would certainly be found that the crude oil futures correlations are higher and more stable over time than the equity correlations. The crude oil futures are highly collinear because there are only a few important sources of information in the data, which are common to many variables.

Principal component analysis (PCA) is a method for extracting the most important uncorrelated sources of information in the data. From a set of k stationary returns it will give up to k orthogonal stationary variables which are called the principal components. At the same time PCA states exactly how much of the total variation in the original data is explained by each principal component. The results of PCA are sensitive to re-scaling of the data, and so it is standard practice to normalise the data before the analysis. We therefore assume that each column in the stationary data matrix **X** has mean zero and variance 1, having previously subtracted the sample mean and divided by $\sqrt{T}$ times the sample standard deviation. Other forms of normalisation are occasionally applied, which explains why statistical packages may give different results.

Let the columns of **X** be $\mathbf{x}_1, \ldots, \mathbf{x}_k$ so that **X'X** is a kxk symmetric matrix with 1's along the diagonal, of correlations between the variables in **X**. Each principal component is a linear combination of these columns, where the weights are chosen from the set of eigenvectors of the **X'X** correlation matrix so that (a) the first principal component explains the maximum amount of the total variation in **X**, the second component explains the maximum amount of the remaining variation, and so on; and (b) the principal components are uncorrelated with each other.

Denote by **W** the matrix of eigenvectors of **X'X**. Thus

$$\mathbf{X'X\,W = W\,\Lambda}$$

factors are used to represent the system. By taking only a few principal components much of the 'noise' that makes correlation estimates so unstable is controlled. The covariance matrix will not be strictly positive definitive but it *will* always be positive semi-definite.

where **Λ** is the diagonal matrix of eigenvalues of **X'X**. Order the columns of **W** according to size of corresponding eigenvalue. Thus if **W** = (w_{ij}) for i,j = 1, ...k, then the mth column of **W**, denoted $\mathbf{w}_m = (w_{1m}, \ldots, w_{km})'$ is the kx1 eigenvector corresponding to the eigenvalue λ_m and the column labelling has been chosen so that $\lambda_1 > \lambda_2 > \ldots > \lambda_k$. Then define the mth principal component of the system by

$$\mathbf{p}_m = w_{1m}\,\mathbf{x}_1 + w_{2m}\,\mathbf{x}_2 + \ldots + w_{km}\,\mathbf{x}_k$$

where $\mathbf{x}_i$ denotes the ith column of **X,** or in matrix notation

$$\mathbf{p}_m = \mathbf{X}\mathbf{w}_m$$

Each principal component is a time series of linear combinations of the **X** variables, and if these are placed as the columns of a full Txk matrix **P** of principal components we have

$$\mathbf{P} = \mathbf{X}\mathbf{W} \qquad (1)$$

To see that this procedure leads to uncorrelated components, note that

$$\mathbf{P'P} = \mathbf{W'X'XW} = \mathbf{W'W}\boldsymbol{\Lambda}.$$

But **W** is an orthogonal matrix, that is $\mathbf{W'} = \mathbf{W}^{-1}$ and so $\mathbf{P'P} = \boldsymbol{\Lambda}$. Since this is a diagonal matrix the columns of **P** are uncorrelated. Since $\mathbf{W'} = \mathbf{W}^{-1}$ (1) is equivalent to $\mathbf{X} = \mathbf{PW'}$, that is

$$X_i = w_{i1}P_1 + w_{i2}P_2 + \ldots\ldots + w_{ik}P_k \qquad (2)$$

where X_i and P_i denote the ith columns of **X** and **P** respectively. Thus each data vector in **X** is a linear combination of the principal components. The proportion of the total variation in **X** that is explained by the mth principal component is λ_m/(sum of eigenvalues).

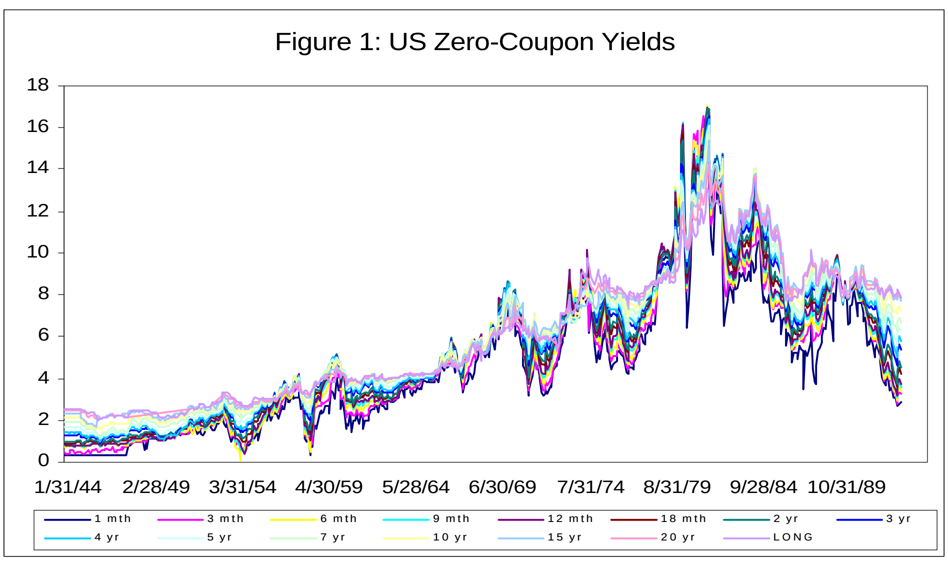


Figure 1: US Zero-Coupon Yields

The principal components method is first illustrated by a standard example: an analysis of US semi-annualised zero coupon rates using monthly data from 1944 to 1992[4], shown in figure 1. The input from these data to a PCA is the returns correlation matrix **X'X** (table 1). For the US data this matrix exhibits the typical structure of the yield curve: correlations tend to decrease with the spread, and the 1mth rate and long rate have lower correlations with other rates since they are most influenced by fiscal and monetary policy.

<u>Table1: Correlation Matrix for US Zero Coupon Rates</u>

	1mth	3mth	6mth	9mth	12mth	18mth	2yr	3yr	4yr	5yr	7yr	10yr	15yr	long
1mth	1													
3mth	0.78739	1												
6mth	0.72919	0.93306	1											
9mth	0.69303	0.88567	0.96762	1										
12mth	0.65619	0.83888	0.92607	0.99126	1									
18mth	0.63125	0.80718	0.90502	0.96856	0.9767	1								
2yr	0.60375	0.77336	0.87517	0.93652	0.94421	0.9928	1							
3yr	0.53997	0.71008	0.82236	0.89329	0.90723	0.96247	0.97383	1						
4yr	0.4898	0.6561	0.77169	0.84662	0.86392	0.92133	0.9347	0.99091	1					
5yr	0.47581	0.634	0.74939	0.82487	0.84294	0.90431	0.92026	0.97895	0.9897	1				
7yr	0.43925	0.58092	0.69222	0.76613	0.78478	0.84915	0.86793	0.92848	0.94151	0.97988	1			
10yr	0.39309	0.53476	0.64898	0.72019	0.73871	0.80584	0.82737	0.88376	0.89529	0.94095	0.97211	1		
15yr	0.30855	0.44558	0.55781	0.61993	0.63647	0.69972	0.72162	0.76671	0.77426	0.81206	0.8361	0.93883	1	
long	0.21933	0.35774	0.43355	0.49401	0.51243	0.54815	0.5573	0.61202	0.62941	0.67691	0.71954	0.74812	0.70197	1

The output from PCA is summarised in table 2: The table 2a gives the eigenvalues and corresponding amount of variation in the original system that is explained, for the first three principal components.

<u>Table 2a: Eigenvector Analysis</u>

Component	Eigenvalue	Cumulative R^2
P1	11.01	0.786
P2	1.632	0.903
P3	0.4963	0.938

---

[4] Copyright Thomas S. Coleman, Lawrence Fisher, Roger G. Ibbotson, U.S Treasury Yield Curves, 1993 Edition,Ibbotson Associates, Chicago.

The trace of the **X'X** matrix above is 15, the sum of the diagonal elements, which is the number of variables in the system. The diagonal matrix of its eigenvalues has the same trace as **X'X** (because trace is invariant under similarity transforms) and so the sum of the eigenvalues is also 15. The largest eigenvalue is 11.01, so the proportion of total variation that it explains is 11.01/15, or 78.6%. The second largest eigenvalue, 1.632, explains a further 1.632/15 that is 11.7% and the third largest eigenvalue (0.4963) explains another 3.5% of the total variation. Thus 93.8% of the total variation in the zero coupon bond returns is explained by the linear model with just 3 principal components.

The second part of the output from PCA is a kxk matrix of factor weights, **W** given in table 2b.Only the factor weights corresponding to the first three principal components are shown below, and they exhibit certain stylised facts. First note that the weights on the first principal component w_{I1} are similar, except perhaps for the very short and very long maturities that have lower correlation with the rest of the system. But in general the correlations are quite high, and this is reflected in the similarity of the factor weights w_{I1}. Under perfect correlation **X'X** is simply a matrix of 1's, with rank 1 and the single eigenvector (1,1,….,1)'. For systems with full rank the first eigenvector corresponding to the largest eigenvalue will take values less than 1, but the more highly correlated the variables the larger and more similar are the eigenvector values corresponding to the largest eigenvector. Put another way, the factor weights on the first principal component will be large, and similar for all variables in a highly correlated system. An upwards shift in the first principal component therefore induces a roughly parallel shift in the yield curve, and for this reason the first principal component is called the *trend component* of the yield curve, and in this example it explains 78.6% of the total variation over the period.

Table 2b: Factor Weights

	P1	P2	P3
1mth	0.63451	0.57207	0.34291
3mth	0.80172	0.50173	0.16278
6mth	0.89228	0.37901	0.033712
9mth	0.94293	0.27852	-0.04566
12mth	0.9451	0.21936	-0.08602
18mth	0.97481	0.11973	-0.12606
2yr	0.97181	0.061225	-0.14593
3yr	0.97585	-0.07672	-0.1628
4yr	0.95465	-0.15533	-0.1684
5yr	0.95542	-0.22317	-0.10985
7yr	0.9234	-0.31032	-0.02539
10yr	0.89628	-0.39553	0.056755
15yr	0.79469	-0.4439	0.12832
long	0.65674	-0.48628	0.46605

The factor weights on the second principal component, w_{i2} , are monotonically decreasing from 0.57207 on the 1mth rate to -0.48628 on the long rate. Thus an upward movement in the second principal component induces a change in slope of the yield curve, where short maturities move up but long maturities move down. The second principal component is called the *tilt* and in this example 11.7% of the total variation is attributed to changes in slope.

The factor weights on the third principal component, w_{i3} , are positive for the short rates, but decreasing and becoming negative for the medium term rates, and then increasing and becoming positive again for the longer maturities. So the third principal component influences the *convexity* of the yield curve, and in this example 3.5% of the variation during the data period is due to changes in convexity.

3. Generating a Covariance Matrix for a Single Risk Factor Category

Now consider how principal components may be used to generate a small covariance matrix, such as the covariance matrix for a yield curve, or the covariance matrix for a set of equities, or a set of equity indices. First suppose that the ith asset return is $\mathbf{y}_i$, so that the normalised variables are $\mathbf{x}_i = (\mathbf{y}_i - \mu_i)/\sigma_i$ where μ_i

and σ_i are the mean and standard deviation of $\mathbf{y}_i$ for i = 1, … k. Write the principal components representation as

$$\mathbf{y}_i = \mu_i + \omega^*_{i1}\,\mathbf{p}_1 + \omega^*_{i2}\,\mathbf{p}_2 + ... + \omega^*_{im}\,\mathbf{p}_m + \boldsymbol{\varepsilon}_i \tag{3}$$

where $\omega^*_{ij} = w_{ij}\sigma_i$ and the error term in (3) picks up the approximation from using only m of the k principal components.

Since principal components are orthogonal their covariance matrix is diagonal. The variances of the principal components can be quickly transformed into a covariance matrix of the original system using the factor weights: Taking variances of (3) gives

$$\mathbf{V} = \mathbf{ADA'} + \mathbf{V}_{\boldsymbol{\varepsilon}} \tag{4}$$

where $\mathbf{A} = (\omega^*_{ij})$ is the matrix of normalised factor weights, $\mathbf{D} = \text{diag}(V(P_1), ...\ V(P_m))$ is the diagonal matrix of variances of principal components and $\mathbf{V}_{\boldsymbol{\varepsilon}}$ is the covariance matrix of the errors. Thus the full kxk covariance matrix of asset returns **V is** obtained from a just a few estimates of the variances of the principal components, and the covariances of the errors.

However **V** may not be positive definite.[5] Although **D** is positive definite because it is a diagonal matrix with positive elements, there is nothing to guarantee that **ADA'** will be positive definite when m < k. To see this write

$$\mathbf{x'ADA'x} = \mathbf{y'Dy}$$

where $\mathbf{A'x} = \mathbf{y}$. Since **y** can be zero for some non-zero **x** , **x'ADA'x** will not be strictly positive for all non-zero **x**. Of course **V** would be positive definite if **ADA'** were positive definite, because $\mathbf{V}_{\boldsymbol{\varepsilon}}$ is positive definite. So if a good approximation has been achieved with m < k principal components there is a reasonable chance that $\mathbf{V}_{\boldsymbol{\varepsilon}}$ will be strictly positive definite. However if it is only positive semi-definite some weights **x** could give zero portfolio variance. But when covariance matrices are based on (4) with m < k, they can always be run through an eigenvalue check to ensure strict positive definiteness. The only way to guarantee strict positive definiteness without having to check is to take all k principal components in the factor model, in which case there is no error term and $\mathbf{V}_{\boldsymbol{\varepsilon}} = \mathbf{0}$.

Just to illustrate the procedure, consider the returns to three stocks in the CAC 40: Paribas, SocGen and Danone using daily data from 1st Jan 1994 to 9th Feb 1999. The direct calculation of their covariance matrix, using equally weighted data over the whole period, is

[5] A symmetric matrix **A** is positive definite if **x'Ax** > 0 for all non-zero **x**. **A** is positive definite if and only if all its eigenvalues are positive (as can be seen by writing $\mathbf{A} = \mathbf{C'\Lambda C}$ where $\boldsymbol{\Lambda}$ is the diagonal matrix of eigenvalues of **A**).

Table 3a: Correlation Matrix

	Paribas	SocGen	Danone
Paribas	1.5728	2.00601	1.28405
SocGen	2.00601	7.39971	2.7741
Danone	1.28405	2.7741	7.50572

The same matrix may be obtained using **ADA'** where **A** is the matrix of rescaled principal components factor weights and **D** is the diagonal matrix of variances of the principal components. To see this, first perform PCA with the full number of components, and this gives the following output:

Table 3b: Eigenvalue Analysis

Component	Eigenvalue	Cumulative R^2
P1	1.897885	0.632628
P2	0.690134	0.862673
P3	0.411982	1

Table 3c: Factor Weights

	P1	P2	P3
Paribas	0.84027	0.29563	0.45447
SocGen	0.83958	0.29946	-0.45325
Danone	0.69781	-0.71628	-0.00192

The matrix **A** is obtained by multiplying each factor weight by the corresponding standard deviation:

	Std Dev
Paribas	1.25411
SocGen	2.72024
Danone	2.73966

So the matrix **A** is:

1.053791	0.370753	0.569955
2.283859	0.814603	-1.23295
1.911762	-1.96236	-0.00526

Since in this case we are just taking equally weighted variance estimates over the whole period, and since the data were normalised before the analysis, the principal components all have unit variance. That is, **D** = **I**, the 3x3 identity matrix. So **ADA'** = **AA'** and the reader may verify that this gives the same covariance matrix as the one calculated directly above.

The example above may be reproduced by the reader using the program ex1.tsp. It has been included simply to illustrate the method, but clearly there is nothing to be gained from the method when equally weighted average variances are employed. Equally weighted averages of the squares are unbiased estimates of the unconditional variance,[6] and each principal component will have a variance estimate of 1 if the estimate is taken over the same data period as the PCA. But suppose exponentially weighted average estimates of the unconditional variance were employed instead? These have the substantial advantage of responding better to current market circumstances, and being less affected by stress events far in the past, than equally weighted averages (see Alexander, 1998). On the other hand they have the disadvantage of there being no one best method for choosing an optimal value of the smoothing constant.

4. Generating Covariance Matrices using Exponentially Weighted Moving Average Variances of the Principal Components.

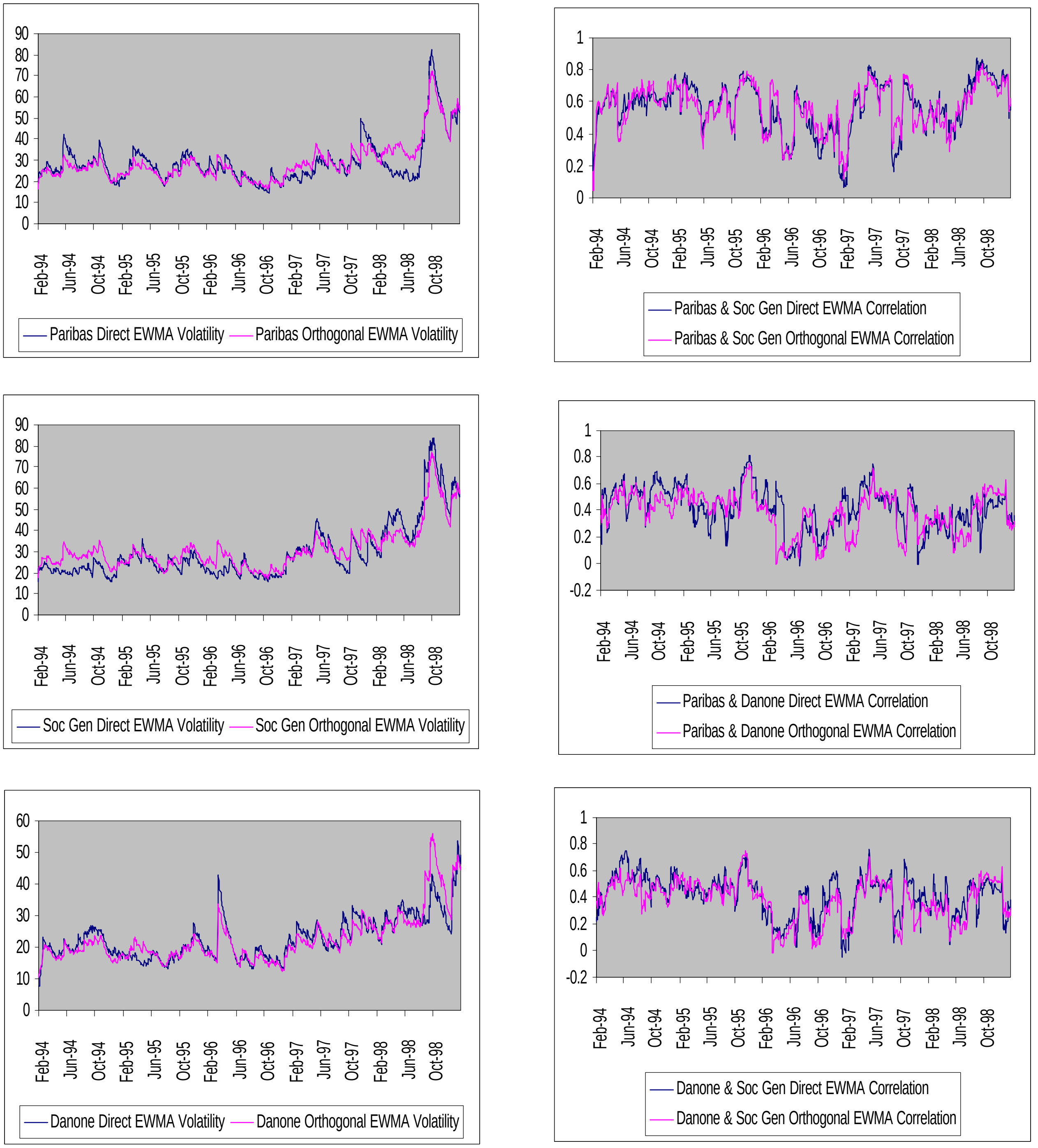
When exponentially weighted moving average (EWMA) volatilities and correlations are estimated directly, the decay factor, as defined by the smoothing constant, must be the same for all series in a large covariance matrix, otherwise it may not be positive semi-definite (see the RiskMetrics Technical Document on www.riskmetrics.com). But when exponentially weighted moving average volatilities and correlations are estimated indirectly using the orthogonal factor method just described, each volatility and correlation will have a different decay factor. Even if the EWMA variances of the principal components all had the same smoothing constant[7] the transformation of these variances using factor weights, by equation (4), will induce different decay rates for the variances and covariances of the variables in the original system. So, provided all principal components are retained, the method provides a simple way to apply EWMA to generate a large positive definite covariance matrix.

The program ex2.tsp uses the same French equity data as ex1.tsp but with exponentially weighted moving averages. Figure 2 plots the volatilities and correlations that are obtained using the orthogonal method with the volatilities and correlations that are obtained using exponentially weighted moving averages directly on the squared returns. This is a very basic example, so the smoothing constant has simply been set as 0.95 for all exponentially weighted moving averages (later examples will use different values of the smoothing constant for the principal components).

[6] One does not normally take squared mean deviations with a bias correction (n-1) in the denominator, since it makes no discernable difference for variance estimators based on daily financial returns.

[7] Choosing identical smoothing constants for all principal components is in fact neither necessary for positive definiteness nor desireable for optimal forecasting. The optimal smoothing constants may be lower for the higher, less important principal components, wheras the volatility of the first, trend component may be the most persistent of the principal component volatilities in a highly correlated system.

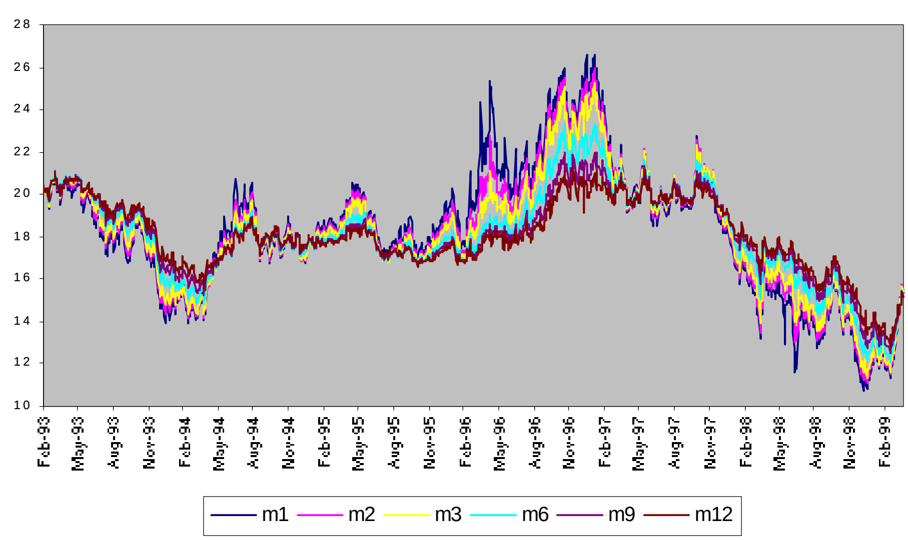
Figure 2: Comparison of Direct and Orthogonal EWMA Volatilities and Correlations



A comparison of these plots is a crucial part of the orthogonal model calibration. If these volatilities and correlations are not similar it will be because (a) the data period used for the PCA is too long, or (b) there are variables that are included in the system that are distorting the volatilities and correlations of other variables computed using the orthogonal method. Both these problems arise if there is insufficient correlation in the system for the method to be properly applied. If one or more of the variables have a low degree of correlation with the other variables over the data period, the factor weights in the PCA will lack robustness over time. The model could be improved by using a shorter data period, and/or omitting the less correlated variables from the system.

Having detailed the method, let us now see its real strength by applying it to a larger and highly correlated system. The program ex3.tsp applies the orthogonal method using just 3 principal components to the WTI crude oil futures data on all monthly maturities from 1 month to 12 months, sampled daily between 4th February 1993 and 24th March 1999. The 1, 2, 3, 6, 9 and 12-month maturity futures prices are shown in figure 3, and see Alexander (1999) for a full discussion of these data and of correlations in energy markets in general.[8]

Figure 3: NYMEX Sweet Crude Prices



[8] Many thanks to Enron for providing these data.

This very highly correlated system is ideally suited to the use of PCA. The output from PCA given in table 4 below shows that 99.8% of the variation in the system may be explained by just 3 principal components. In fact just the first principal component explains almost 96% of the variation over the period, and with two principal components over 99% of the variation is explained. Of course the factor weights show that, as with any term structure, the interpretations of the first three principal components are the trend, tilt and curvature components respectively.

Table 4a: Eigenvalue Analysis

Component	Eigenvalue	Cumulative R^2
P1	11.51	0.9592
P2	0.397	0.9923
P3	0.069	0.9981

Table 4b: Factor Weights

	P1	P2	P3
1mth	0.89609	0.40495	0.18027
2mth	0.96522	0.24255	-0.063052
3mth	0.98275	0.15984	-0.085002
4mth	0.99252	0.087091	-0.080116
5mth	0.99676	0.026339	-0.065143
6mth	0.99783	-0.020895	-0.046369
7mth	0.99702	-0.062206	-0.023588
8mth	0.99451	-0.098582	0.00018279
9mth	0.99061	-0.13183	0.020876
10mth	0.98567	-0.16123	0.040270
11mth	0.97699	-0.19269	0.064930
12mth	0.97241	-0.21399	0.075176

The great advantage in using the orthogonal method on term structure data is that all the volatilities and correlations in the system can be derived from just 3 exponentially weighted moving average variances. That is, instead of estimating 78 exponentially weighted moving average volatilities and correlations directly, using the same value of the smoothing constant throughout, only 3 exponentially weighted moving average variances of the trend, tilt and curvature principal components need to be generated. In some term structures, including the crude oil futures term structure used in example 3, only 2 components already explain over 99% of the variation, so adding a 3rd component makes no discernible difference to the covariance results.

All the volatilities and correlation variances of the original system can be recovered using simple transformations of the diagonal matrix of principal component variances, as is done in ex3.tsp. Moreover the principal component variances may use different smoothing constants. In example 3 the default value of 0.95 for the 1st , 2nd and 3rd principal components has been used, but the reader may be interested to experiment with using different smoothing constants for the principal component variances. Even if the principal component variances do all have the same smoothing constant, the volatilities of different maturities in the term structure would have different exponential smoothing properties. This is of course because they have different factor weights in the principal component representation.

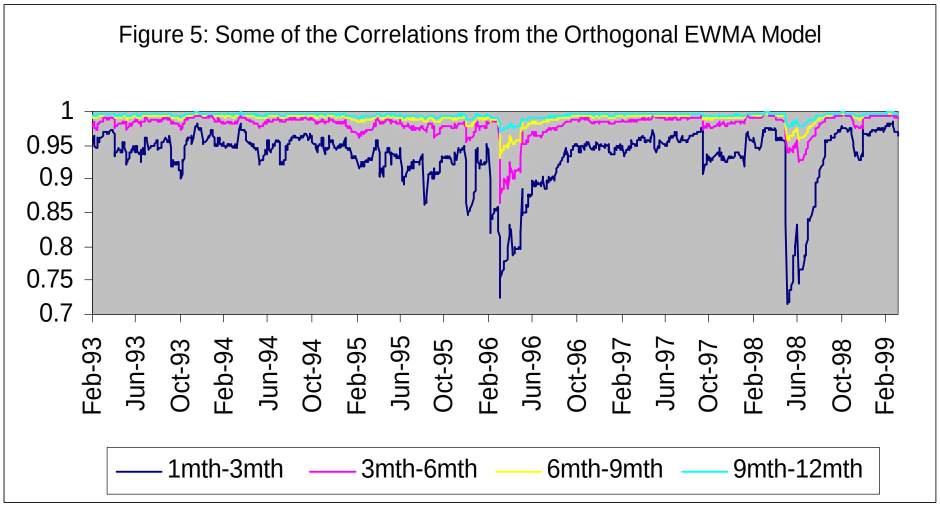
The figures in figure 4 show some of the volatilities that are generated using the orthogonal method compared with directly estimated exponentially weighted moving average volatilities. The coincidence of the results obtained by the orthogonal method with those obtained by direct estimation shows how powerful this method is. From just 2 or 3 exponentially weighted moving averages, the entire 12x12 covariance matrix of the original system is recovered with negligible loss of precision. Figure 5 shows some of the correlations obtained using the orthogonal method for different pairs of maturities. The reader may easily view more of them from the off-diagonal elements of the covariance matrix in ex3.tsp.

The orthogonal EWMA correlations shown in figure 5 are very similar indeed to the correlations generated by direct EWMA. But there is a problem with using exponentially weighted averages at all in the crude oil futures market. Although they contain fewer 'ghost features' and other artificial effects that result from the use of equally weighted moving averages, there is still a disturbing lack of correlation between some of the near maturity futures. This problem will be a point of discussion based on figure 8 below, where the same correlations are measured by the orthogonal GARCH model.

There are many advantages with the orthogonal method for generating covariance matrices, even when it is applied using only exponentially weighted moving average variance estimates. Obviously the computational burden is much lighter when all of the k(k+1)/2 volatilities and correlations are simple matrix transformations of just 2 or 3 exponentially weighted moving average variances. But also data may be difficult to obtain directly, particularly on some financial assets that are not heavily traded. When data are sparse or unreliable on some of the variables in the system a direct estimation of volatilities and correlation may be very difficult. But if there is sufficient information to infer their factor weights in the principal component representation, their volatilities and correlations may be obtained using the orthogonal method.

For example, some bonds or futures may be relatively illiquid for certain maturities, and statistical forecasts of their volatilities may be difficult to generate directly on a daily basis. But the variances of the principal components of the entire term structure can be transformed using the factor weights into a full covariance matrix that generates flexible forecasts of all maturities, including the illiquid ones.

Figure 4: Comparison of Direct and Orthogonal EWMA Volatilities



5. Introducing the Orthogonal GARCH Model

The univariate generalised autoregressive conditional heteroscedasticity (GARCH) models that were introduced by Engle (1982) and Bollerslev (1986) have been very successful for volatility estimation and forecasting in financial markets. The mathematical foundation of GARCH models compares favourably with some of the alternatives used by financial practitioners, and this mathematical coherency makes GARCH models easy to adapt to new financial applications. There is also evidence that GARCH models generate more realistic long-term forecasts, since the volatility and correlation term structure forecasts will converge to the long-term average level (see Alexander, 1998, 2000). As for short-term volatility forecasts, statistical results are mixed (see for example Brailsford and Faff, 1996, Dimson and Marsh, 1990, Figlewski, 1994, Alexander and Leigh (1997)). This is not surprising since the whole area of statistical evaluation of volatility forecasts is fraught with difficulty (see Alexander, 2000). Another test of a volatility forecasting model is in its hedging performance: and there is much to be said for using the GARCH volatility framework for pricing and hedging options (see Duan 1995, 1996). Engle and Rosenberg (1995) provide an operational evaluation of GARCH models in option pricing and hedging, where their superiority to the Black-Scholes methods stems from the fact that stochastic volatility is already built into the model.

Large covariance matrices that are based on GARCH models would, therefore, have clear advantages. But previous research in this area has met with rather limited success. It is straightforward to generalise the univariate GARCH models to multivariate parameterisations, as in Engle and Kroner (1993). But the actual implementation of these models is extremely difficult. With so many parameters, the likelihood function becomes very flat, and so convergence problems are very common in the optimization routine. If the modeller needs to 'nurse' the model for systems with only a few variables, there is little hope of a fully functional implementation of a direct multivariate GARCH model to work on large risk systems.

The idea of using factor models with GARCH is not new. Engle, Ng and Rothschild (1990) use the capital asset pricing model to show how the volatilities and correlations between individual equities can be generated from the univariate GARCH variance of the market risk factor. Their results have a straightforward extension to multi-factor models, but unless the factors are orthogonal a multi-variate GARCH model will be required, with all the associated problems.

A principal components representation is a multi-factor model. In fact the orthogonal GARCH model is a generalisation of the factor GARCH model introduced by Engle, Ng and Rothschild (1990) to a multi-factor model with orthogonal factors. The orthogonal GARCH model allows kxk GARCH covariance matrices to be generated from just m univariate GARCH models. Here k is the number of variables in the system and m is the number of principal components used the represent the system. It may be that m can be much less than k, and quite often one would wish m to be less than k so that extraneous 'noise' is excluded from the data. But since only univariate GARCH models are used it does not really matter: there no dimensional restrictions as there are with the direct parameterisations of multivariate GARCH.[9]

Of course, the principal components are only unconditionally uncorrelated, so a GARCH covariance matrix of principal components is not necessarily diagonal. However the assumption of zero conditional correlations has to be made, otherwise it misses the whole point of the model, which is to generate large GARCH covariance matrices from GARCH volatilities alone.

Before presenting some empirical examples on orthogonal GARCH let us just rephrase the results of section 3 in the framework of stochastic volatility. Thus the mxm diagonal matrix of variances of the principal components is a time-varying matrix denoted $\mathbf{D}_t$ and the time-varying covariance matrix $\mathbf{V}_t$ of the original system is approximated by

$$\mathbf{V}_t = \mathbf{A}\ \mathbf{D}_t\ \mathbf{A}' \qquad (5)$$

where **A** is the kxm matrix of re-scaled factor weights. The representation (5) will give a positive semi-definite matrix at every point in time, even when the number m of principal components is much less that the number k of variables in the system. However the accuracy of the representation (5) depends on the number of principal components used being sufficient to explain a large part of the variation in the system. The method will therefore work well when principal component analysis works well, i.e. on term structures and other highly correlated systems.

The model (5) is called orthogonal GARCH when the diagonal matrix $\mathbf{D}_t$ of variances of principal components is estimated using a GARCH model. In the examples given here the standard 'vanilla' GARCH(1,1) model is used. The conditional variance at time t is defined as:

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \qquad (6)$$

where the 'market reaction' parameter α and the 'volatility persistence' parameter β should sum to less than one (for convergence of term structure volatility forecasts). In the exponentially weighted moving average model these parameters always sum to one, so the volatility term structure will be constant.

The orthogonal GARCH model is particularly useful for term structures where the more illiquid maturities can preclude the direction estimation of GARCH volatilities. When market trading is rather thin there may be little autoregressive conditional heteroscedasticity in the data, and what is there may be rather unreliable.

---

[9] The 1994 PhD thesis of Zhuanxin Ding with Rob Engle at UCSD first introduced the idea of using principal component analysis with univariate GARCH on all principal components. The important difference between Ding's model and the orthogonal GARCH is the use of only a small number of principal components; this is crucial for correlations to be stable and volatilities to be robust to umimpotant changes in the risk factors.

The orthogonal GARCH model has the advantage that the volatilities of such assets, and their correlations with other assets in the system, are derived from the principal component volatilities that are common to all assets and the factor weights that are specific to that particular asset.

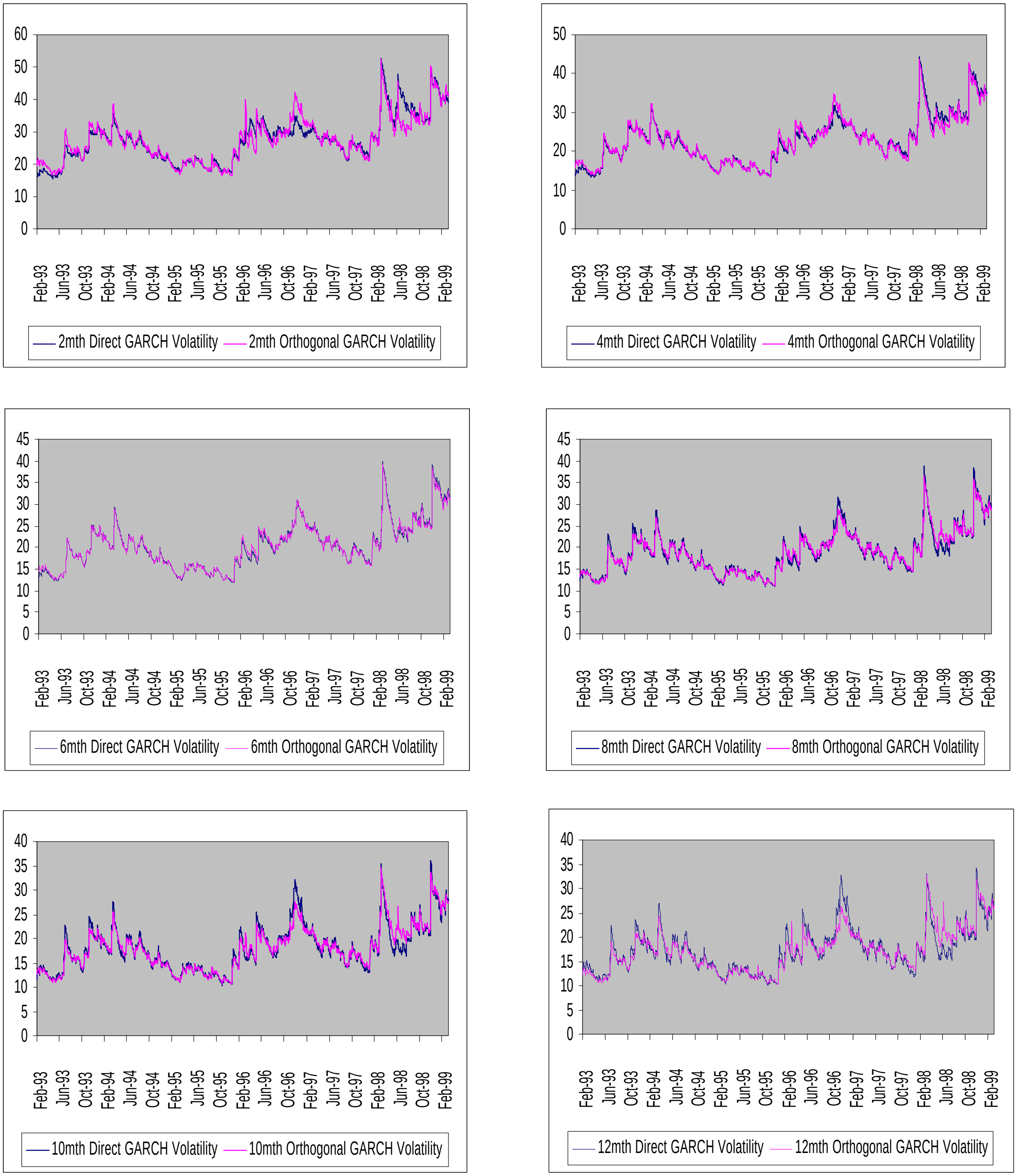
The first example of the orthogonal GARCH model is a straightforward extension of example 3, using GARCH(1,1) variances of principal components in place of exponentially weighted moving averages. The ex4.tsp uses the same crude oil futures term structure data as example 3 and table 5 reports the estimated coefficients in GARCH(1,1) models of the first 2 principal components.

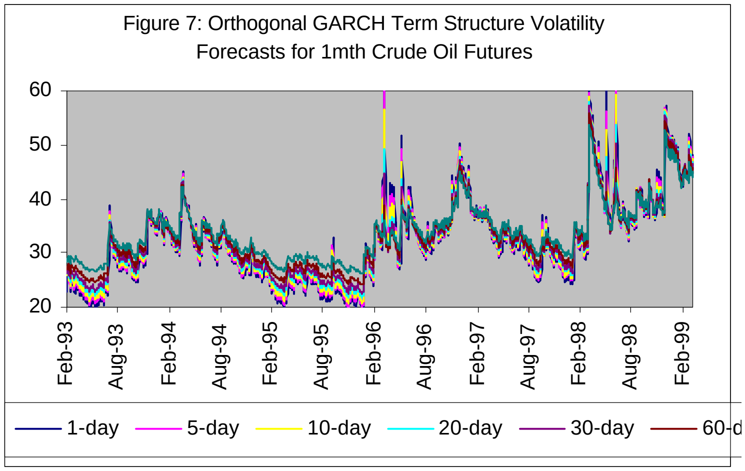
Table 5: GARCH(1,1) Models of the First and Second Principal Component

	1st Principal Component		2nd Principal Component	
	Coefficient	t-stat	Coefficient	t-stat
constant	.650847E-02	.304468	.122938E-02	.066431
ω	.644458E-02	3.16614	.110818	7.34255
α	.037769	8.46392	.224810	9.64432
β	.957769	169.198	.665654	21.5793

Figure 6 is similar to figure 4, the only difference being that GARCH(1,1) models have been used to generate figure 6 wherever exponentially weighted moving averages were used for figures 4. The orthogonal and direct volatilities that are compared in figure 6 are very close indeed. In fact, they are almost identical to the EWMA volatilities illustrated in figure 4. Why bother with GARCH then? There are two important reasons. The first is that EWMA volatility term structure forecasts do not converge to the long-term average, but GARCH forecasts do, provided $\alpha + \beta < 1$. In fact the orthogonal GARCH model can be extended quite easily to provide forecasts of the average volatility over the next n days, for any n (see Alexander 1998, 2000). The ex5.tsp does precisely this for the crude oil term structure data, producing series of term structure volatility forecasts that converge to a long-term average. Volatility terms structures for the 1mth future are shown in figure 7.

Figure 6: Comparison of Direct and Orthogonal GARCH Volatilities



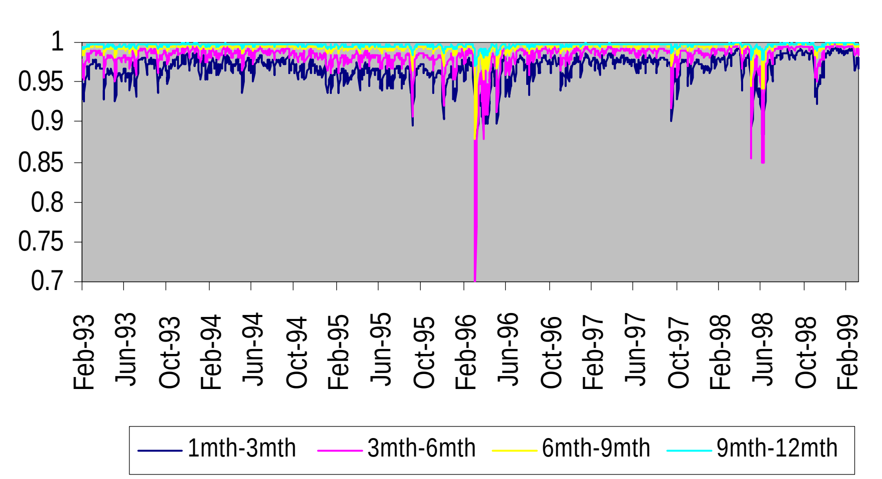


The second good reason to use orthogonal GARCH rather than orthogonal EWMA is that the orthogonal GARCH correlations will more realistically reflect what is happening in the market. As already mentioned, the correlations shown in figure 5 that were generated by the orthogonal EWMA are a little worrying. One would expect correlations between commodity futures to be more or less perfect most of the time, but the EWMA correlations between the 1 mth futures and other futures, and between other pairs at short maturities, can be considerably below 1 for long periods of time. For example during long periods of 1996 and 1998 the EWMA correlations are nearer to 0.8 than 1.

The reason for this is the smoothing constant of 0.95, which is an appropriate choice for the volatilities (as we know from the comparison of figure 4 with the optimised GARCH models in figure 6) but is clearly too large for the correlations. Unfortunately if one were to reduce the values of the smoothing constants used in the orthogonal EWMA model, so that the correlations were less persistent, so also would the volatilities be less persistent. One can only guess by trial and error what are the approriate values for the smoothing constants, and it may be that there is no clear answer to this question.

In the crude oil futures market price decoupling only occurs over very short time spans so correlations may deviate below 1, but only for a short time. Now, if the orthogonal model were to be used with just one principal component (which is possible since the results from section 4 indicate that this trend component explains over 95% of the variation) the correlations would of course be unity. So all the variation in the orthogonal GARCH correlations is coming from the movements in the second principal component. This second principal component is the tilt component, and it explains about 4% of the movement (see section 4). The GARCH(1,1) models of the first two principal components of this term structure, given in the table above, indicate that the second principal component has a lot of reaction (α is about 0.22) but little persistence (β is about 0.66). In other words these tilt movements in the term structure of futures prices are intense but short-lived. So one would expect the correlations given by the orthogonal GARCH model in figure 8 to be more accurately reflecting real market conditions than the orthogonal EWMA correlations in figure 5.

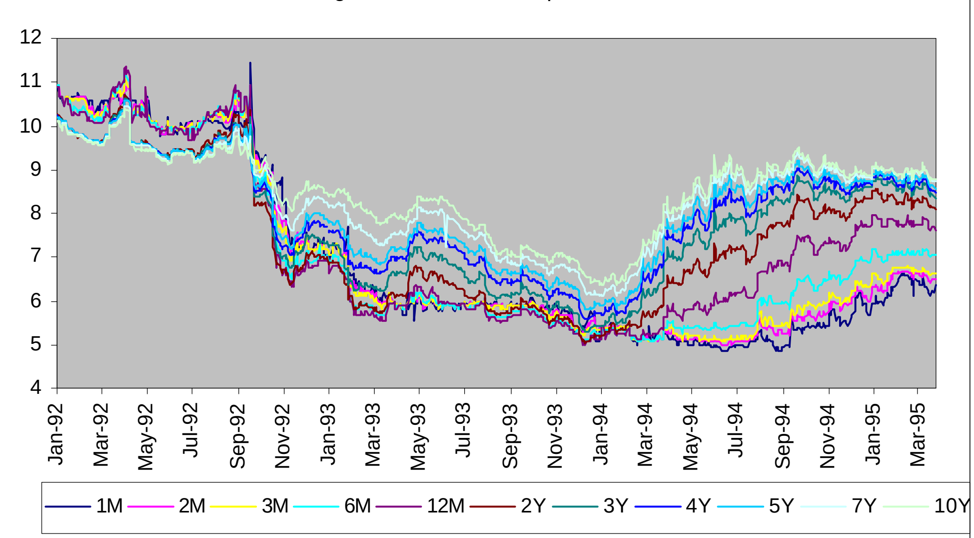
Figure 8: Some of the Correlations from the Orthogonal GARCH Model



To summarise the results so far, this paper has shown how 78 different volatilities and correlations of the term structure of crude oil futures between 1 mth and 12mths can be generated, very simply and very accurately, from just two univariate GARCH models of the first two principal components. It has also shown how volatility forecasts of different maturities can also be generated as simple transformations of these two basic GARCH variances.

Now let us now step up a little with the complexity of the data. Still a term structure, but rather a difficult one. The program given in ex6.tsp has been trained on daily zero coupon yield data in the UK with 11 different maturities between 1mth and 10 years from 1st Jan 1992 to 24th Mar 1995, shown in figure 9.

Figure 9: UK Zero-Coupon Yields



It is not an easy task to estimate univariate GARCH models on these data directly because yields may remain relatively fixed for a number of days. Particularly on the more illiquid maturities, there may be insufficient conditional heteroscedasticity for GARCH models to converge well. The reader that uses ex6.tsp will see how problematic is the direct estimation of GARCH models on these data. So the orthogonal GARCH volatilities in figure 10 have been compared instead with exponentially weighted moving average volatilities (with a smoothing constant of 0.9). The orthogonal GARCH volatilities are not as closely aligned with the exponentially weighted moving average volatilities as they were in the previous example, but there is sufficient agreement between them to place a fairly high degree of confidence in the orthogonal GARCH model. Again two principal components were used in the orthogonal GARCH, but the principal component analysis below shows that these two components only account for 72% of the total variation (as opposed to over 99% in the crude oil term structure).

Table 6a: Eigenvalue Analysis

Component	Eigenvalue	Cumulative R^2
P1	5.9284117	0.53894652
P2	1.9899323	0.71984946
P3	0.97903180	0.80885235

Table 6b: Factor Weights

	P1	P2	P3
1mth	0.50916	0.60370	0.12757
2mth	0.63635	0.62136	-0.048183
3mth	0.68721	0.57266	-0.10112
6mth	0.67638	0.47617	--0.10112
12mth	0.83575	0.088099	-0.019350
2yr	0.88733	-0.21379	0.033486
3yr	0.87788	-0.30805	-0.033217
4yr	0.89648	-0.36430	0.054061
5yr	0.79420	-0.37981	0.14267
7yr	0.78346	-0.47448	0.069182
10yr	0.17250	-0.18508	-0.95497

Clearly the lower degree of accuracy from a 2 principal component representation is one reason for the observed differences between the orthogonal GARCH volatilities and the EWMA volatilities. Another is that the 10yr yield has a very low correlation with the rest of the system, as reflected by its factor weight on the 1st principal component, which is quite out of line with the rest of the factor weights on this component. The fit of the orthogonal model is good, but could be improved further if the 10yr bond were excluded from the system.

The GARCH(1,1) model estimates for the first two principal components are given in table 7 below. This time the second principal component has a better-conditioned GARCH model. So the tilts in the UK yield curve are less temporary and more important than they are in the crude oil term structure discussed above. One consequence of this is the orthogonal GARCH correlations will be less jumpy and more stable than the correlations in figure 8.

Table 7: GARCH(1,1) Models of the First and Second Principal Component

	1st Principal Component		2nd Principal Component	
	Coefficient	t-stat	Coefficient	t-stat
constant	.769758E-02	.249734	.033682	1.09064
ω	.024124	4.50366	.046368	6.46634
α	.124735	6.46634	.061022	9.64432
β	.866025	135.440	.895787	50.8779

Figure 10: Comparison of Direct EWMA with Orthogonal GARCH Volatilities

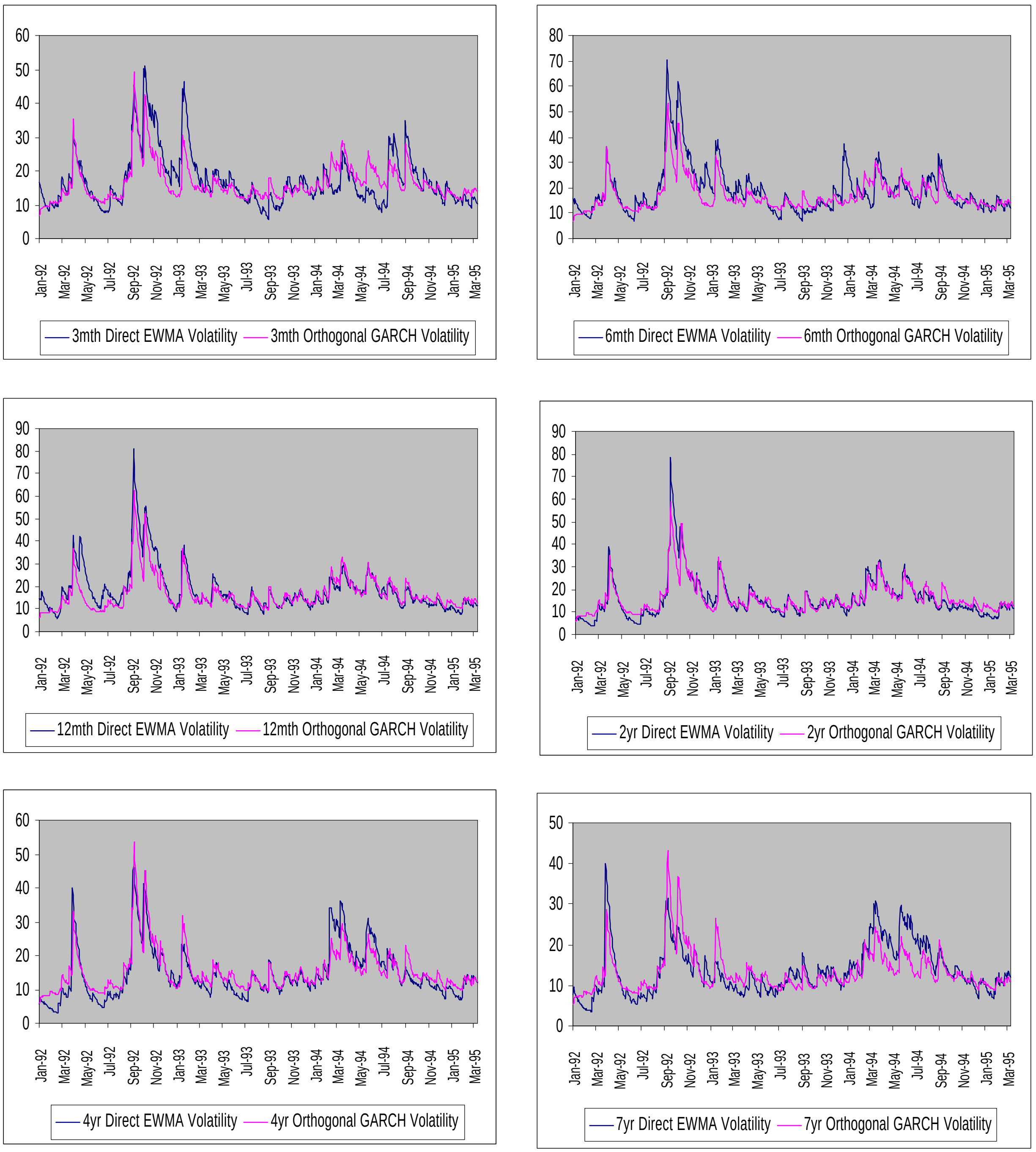
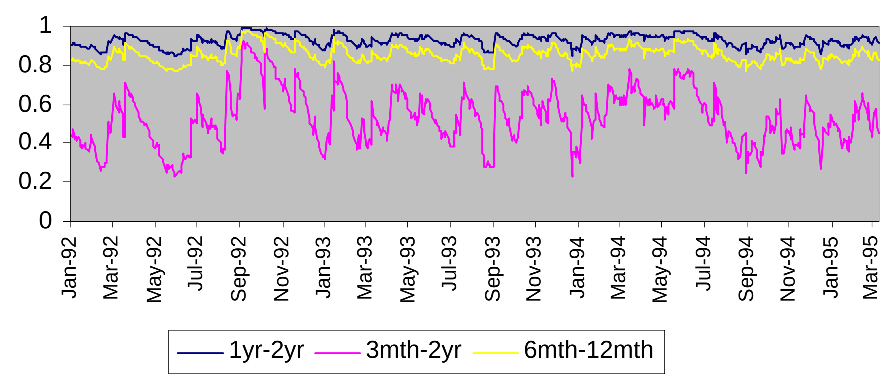


Figure 11 shows some of the orthogonal GARCH correlations for the UK zero coupon yields. Not only does the orthogonal method provide a way of estimating GARCH volatilities and volatility term structures that may be difficult to obtain by direct univariate GARCH estimation.[10] They also give very sensible GARCH correlations, which would be very difficult indeed to estimate using direct multi-variate GARCH. And all these are obtained from just 2 principal components that are representing the important sources of information - all the rest of the variation is ascribed to 'noise' and is not included in the model.

Figure 11: Orthogonal GARCH Correlations of UK Zero-Coupon Yields



A useful technique for parameterising multivariate GARCH models is to compare the GARCH volatility estimates from the multivariate GARCH with those obtained from direct univariate GARCH estimation. Similarly, when calibrating an orthogonal GARCH model one could compare the volatility and correlation estimates with those obtained from other models, such as an EWMA correlation model or other multivariate GARCH models. There are of course problems with this. What choice of smoothing constants should be made when the comparison is with exponentially weighted moving average volatilities? If the system is large convergence problems may very well be encountered, so how sure can one be about the validity of the diagonal vech or BEKK multivariate GARCH parameter estimates?

Since multivariate GARCH is not easy to use for large systems, a relatively small system of six European equity indices is used. Some care must be taken with the initial calibration of orthogonal GARCH. The factors that must be taken into account are:

- *The assets that are included in the system.* Principal component analysis works best in a highly correlated system. An asset that has very idiosyncratic properties compared to other assets in the system (such as the 10yr bond in ex6.tsp) will corrupt the volatilities and correlations of the other assets in the system, because they are all based on the principal components that are common to all assets.

[10] These are not shown but may easily be generated by adapting the ex5.tsp to these data

- *The time period used for estimation.* The GARCH volatilities of the principal components change over time, but it is only their current values that matter for forecasting the covariance matrix. However the factor weights are also used in this forecast, and these are constants that take different values depending on the estimation period. So changing the time period for estimation affects current forecasts of the covariance matrix primarily because it affect the factor weights matrix **A**, rather than the principal component volatilities in $\mathbf{D}_t$.

Univariate GARCH(1,1) models are estimated for each principal component, giving a time-varying diagonal covariance matrix $\mathbf{D}_t$. Then the orthogonal GARCH variances and covariances are obtained from the matrix $\mathbf{V}_t$ given by

$$\mathbf{V}_t = \mathbf{A}\ \mathbf{D}_t\ \mathbf{A}'$$

where **A** is the normalized matrix of factor weights that re-scales each element of **W** by multiplication by the appropriate standard deviation, as described in section 3. The results are the 4 volatilities and 6 correlation graphs shown in figure 12. These graphs compare the orthogonal GARCH volatilities and correlations with those estimated from two other models (see Engle and Kroner, 1993). First there is the diagonal Vech model, which in two dimensions takes the form:

$$\begin{aligned} s_{1,t}^2 &= w_1 + a_1 e_{1,t-1}^2 + b_1 s_{1,t-1}^2 \\ s_{2,t}^2 &= w_2 + a_2 e_{2,t-1}^2 + b_2 s_{2,t-1}^2 \\ s_{12,t} &= w \ \ + a_3 e_{1,t-1} e_{2,t-1} + b_3 s_{12,t-1} \end{aligned}$$

or, in matrix notation

$$\text{vech}\,(\mathbf{V}_t) \ = \ \mathbf{A} \ + \ \mathbf{B}\ \text{vech}\,(\xi_{t-1}\xi_{t-1}') \ + \ \mathbf{C}\ \text{vech}\,(\mathbf{V}_{t-1}) \qquad (7)$$

where $\mathbf{V}_t$ is the conditional covariance matrix at time t, $\text{vech}(\mathbf{V}_t) = (\sigma_{1t}^2 , \sigma_{2t}^2 , \sigma_{12t})'$, $\xi_t = (\varepsilon_{1t}, \varepsilon_{2t})'$, $\mathbf{A} = (\omega_1, \omega_2, \omega_3)'$, $\mathbf{B} = \text{diag}(\alpha_1, \alpha_2, \alpha_3)$ and $\mathbf{C} = \text{diag}(\beta_1, \beta_2, \beta_3)$.

But although it may be relatively easy to estimate, the diagonal vech parameterisation (7) places severe cross-equation restrictions on the parameters. A more general parameterisation of multivariate GARCH is the BEKK model

$$\mathbf{V}_t \ = \mathbf{A}'\mathbf{A} \ + \ \mathbf{B}'\,\xi_{t-1}\xi_{t-1}'\,\mathbf{B} \ + \ \mathbf{C}'\,\mathbf{V}_{t-1}\mathbf{C} \qquad (8)$$

where **A, B** and **C** are nxn matrices and **A** is triangular. The parameterisation (8) guarantees positive definiteness, but the number of parameters to estimate grows rapidly with the dimension of the system, so likelihood surfaces become very flat indeed. Consequently the BEKK model often encounters convergence problems unless some sort of restrictions are placed on the parameters.

The last example of orthogonal GARCH model calibration uses 6 European equity indices: France (CAC40), Germany (DAX30), Holland (AEX), Spain (IBEX), Sweden (OMX), and the UK (FTSE100).[11] A principal component analysis on daily return data from Morgan Stanley Index prices, from 1st April 1993 to 31st December reveals rather lower correlations with IBEX and OMX than between the other indices. So it is better to divide the 6 indices into two groups: AEX, CAC, DAX, and FTSE, in one group and IBEX and OMX in the other. For the first group the principal component factor weights are

Table 8a: Factor Weights

	AEX	CAC	DAX	FTSE
P1	0.866	0.834	0.755	0.818
P2	0.068	-0.238	0.615	-0.397
P3	0.224	-0.496	-0.027	0.294
P4	0.441	0.036	-0.226	-0.296

The weights on the first principal component are comparable, and quite high. Since this is the trend component the system is, on the whole, moving together, and the eigenvalue analysis below indicates that common movements in the trend explain 67% of the total variation over the 4-year period:

Table 8b: Eigenvalue Analysis

Component	Eigenvalue	Cumulative R^2
P1	2.686141	0.671535
P2	0.596853	0.820749
P3	0.382549	0.916386
P4	0.334456	1

The 3rd and 4th principal component are often relatively more important in equity systems than in term structures, and this case is no exception. So all 4 principal components have been used in the orthogonal GARCH model, that is, the matrix $\mathbf{D}_t$ is a 4x4 diagonal matrix. Table 9 below reports results from estimating univariate GARCH(1,1) models on each of the four principal components, to give the elements of $\mathbf{D}_t$ at each point in time.[12]

[11] Many thanks to Dr. Aubrey Chibumba for producing these results as part of his M.Phil thesis at Sussex University.

[12] Asymmetric GARCH(1,1) models would be more appropriate because of the leverage effect in equity markets. There is no problem with using asymmetric univerate GARCH(1,1) models in the orthogonal GARCH, but convergence difficulties were encountered when trying to estimate asymmetric GARCH models with both multivariate parameterisations.

Table 9: GARCH(1,1) Models of the Principal Components

	1st PC		2nd PC		3rd PC		4th PC	
	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
constant	0.00613	0.19446	0.00262	0.09008	-0.0801	-0.26523	0.00267	0.087511
ω	0.032609	1.90651	0.066555	3.10594	0.089961	2.12915	0.203359	1.80057
α	0.033085	2.69647	0.086002	4.57763	0.067098	2.92511	0.070417	2.00423
β	0.934716	35.9676	0.846648	25.9852	0.841618	14.4038	0.726134	5.22888

This system has only 4 variables, so there are no convergence problems with either multivariate GARCH model. The 4-dimensional diagonal vech model has 10 equations, each with 3 parameters. The 30 parameter estimates and their t-statistics (in italics) are reported in table 10. The 4-dimensional BEKK model has 42 parameters and the estimates of the matrices **A**, **B,** and **C** are given in table 11.

Table 10: Diagonal Vech Parameter Estimates

Variance Equations				Covariance Equations					
AEX	CAC	DAX	FTSE	AEX-CAC	AEX-DAX	AEX-FTSE	CAC-DAX	CAC-FTSE	DAX-FTSE
ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
$5.8\text{x}10^{-6}$	$3.4\text{x}10^{-6}$	$5.0\ \text{x}10^{-6}$	$1.8\ \text{x}10^{-6}$	$1.9\ \text{x}10^{-6}$	$9.3\ \text{x}10^{-6}$	$1.8\ \text{x}10^{-6}$	$8.6\ \text{x}10^{-6}$	$3.0\ \text{x}10^{-6}$	$1.6\ \text{x}10^{-6}$
3.11	*2.19*	*1.71*	*2.11*	*2.45*	*1.63*	*2.25*	*3.24*	*1.58*	*2.28*
α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
.054900	0.028889	0.028264	0.024601	0.021826	0.031806	0.028069	0.059739	0.022341	0.028377
4.14	*3.15*	*2.28*	*2.68*	*3.75*	*2.40*	*3.68*	*3.56*	*2.18*	*2.88*
β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
0.82976	0.89061	0.83654	0.91227	0.95802	0.74503	0.926414	0.829753	0.861387	0.934363
18.23	*21.17*	*9.58*	*25.2*	*79.49*	*5.20*	*36.36*	*17.93*	*10.67*	*38.74*

Table 11: BEKK Parameter Estimates

	AEX	CAC	DAX	FTSE
A	0.00160	0	0	0
	0.00008	-0.00176	0	0
	0.00094	0.00197	-0.00087	0
	0.00142	-0.00003	-0.00051	$2.5x10^{-6}$
B	0.22394	-0.04156	0.019373	0.04785
	-0.07147	0.18757	-0.05247	0.031895
	-0.06286	-0.04764	0.29719	0.07003
	-0.016277	-0.027589	-0.017405	0.178563
C	0.951805	0.027231	-0.050236	0.026130
	0.033141	0.9615723	0.023822	0.013623
	0.067985	0.053024	0.844291	0.005211
	0.022278	0.029257	-0.014482	0.948453

Figure 12: Comparison of Orthogonal GARCH with Multivariate GARCH Models

Figure 12a: GARCH Models Volatility Comparison (AEX)

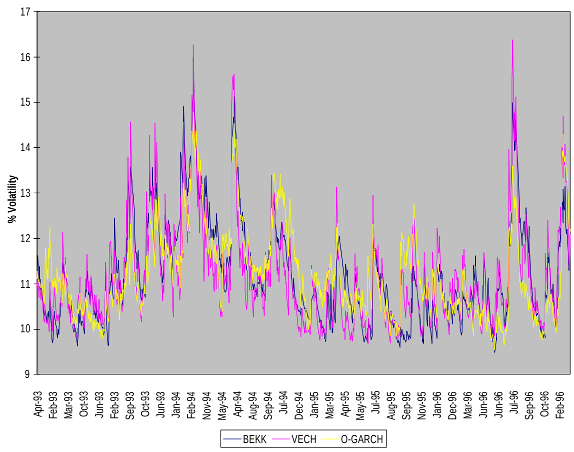


Figure 12b: GARCH Models Volatility Comparison (CAC)

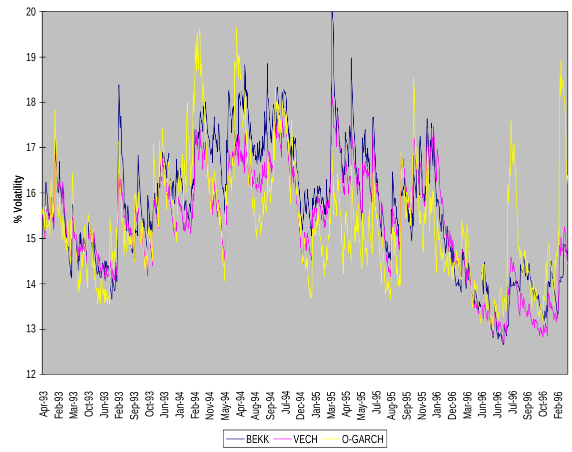


Figure 12c: GARCH Models Volatility Comparison (DAX)

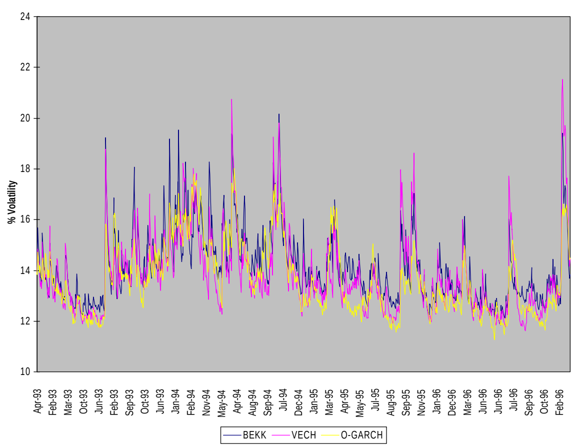


Figure 12d: GARCH Models Volatility Comparison (FTSE)

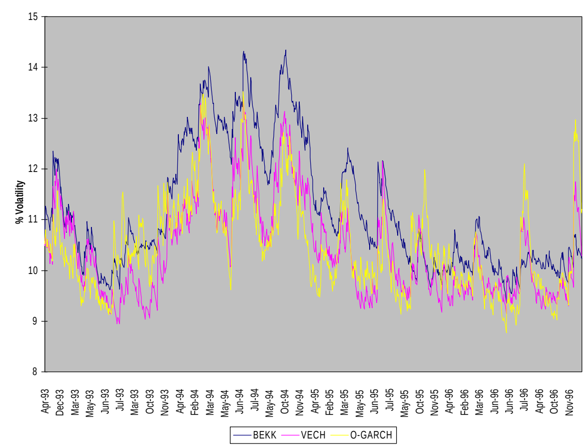


Figure 12e: GARCH Models Correlation Comparison (AEX-CAC)

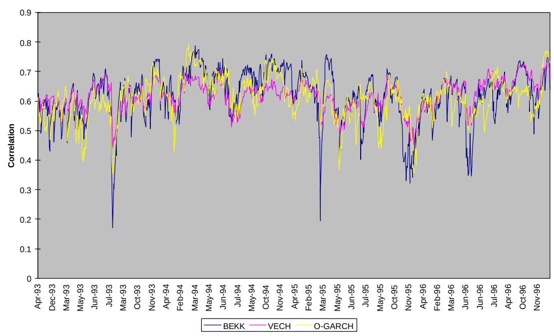


Figure 12f: GARCH Models Correlation Comparison (AEX-DAX)

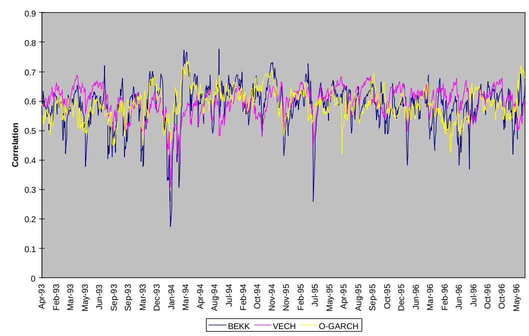


Figure 12g: GARCH Models Correlation Comparison (AEX-FTSE)

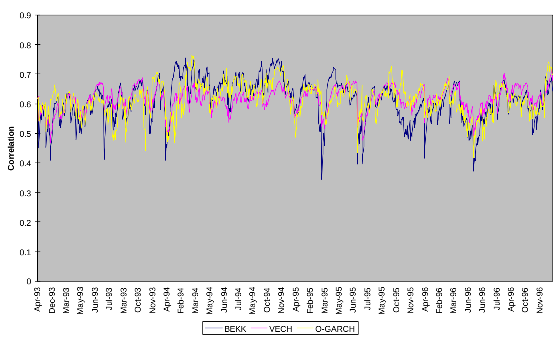


Figure 12h: GARCH Models Correlation Comparison (CAC-DAX)

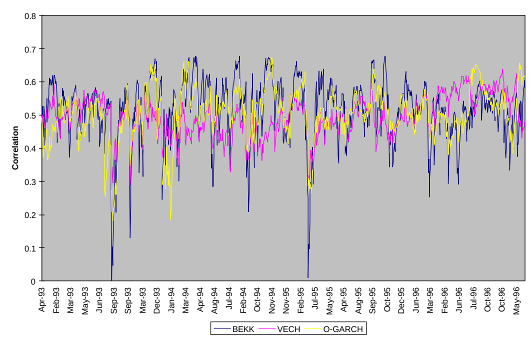


Figure 12i: GARCH Models Correlation Comparison (CAC-FTSE)

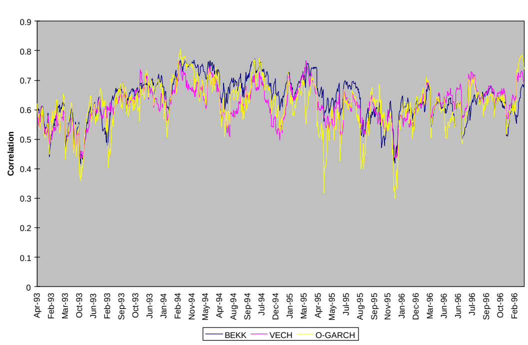
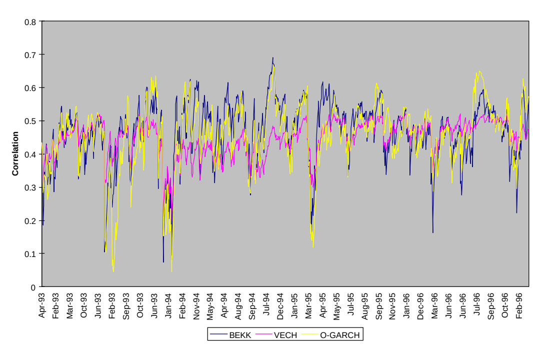


Figure 12j: GARCH Models Correlation Comparison (DAX-FTSE)



It is important to realise that the 10 graphs in figure 12 come from the models reported in tables 9 to 11 above. So the BEKK volatilities and correlations come from the BEKK model estimated in table11. Similarly there is only one diagonal vech model generating the VECH series for the graphs (table 10) and one orthogonal GARCH model for the orthogonal GARCH graphs (tables 9). However there are cases that the orthogonal GARCH volatilities coincide quite closely with the BEKK volatilities but not the vech volatilities (graphs f, h, and j). In some case the orthogonal GARCH are more similar to the vech volatilities than the BEKK volatilities (graphs d, e and i) and in some cases the volatilities differ noticeably (graphs a, b, c, and g). Having said this, there is not a huge difference between the three models in any of the graphs. Given how volatility and correlation estimates *can* differ when different models are used, these graphs are nothing abnormal.

6. Generating a Large Covariance Matrix across All Risk Factor Categories

This section describes the use of principal component analysis to generate large dimensional covariance matrices. The method avoids many of the known problems with the RiskMetrics data: positive definiteness can be assured without using the same smoothing constant for all markets; there is no need to reduce dimensions by using interpolation along the yield curve; and the method will conform to regulators requirements on historic data if at least one year of data are used to compute the principal components and their factor weights.

The risk factors - equity market indices, exchange rates, commodities, government bond and money market rates and so on - are first divided into reasonably highly correlated categories, according to geographic locations and instrument types. Principal component analysis is then used to orthogonalise each sub-system of risk factors and an exponentially weighted moving average is applied to each of the principal components to obtain the diagonal covariance matrix. Then the factor weights from the principal component analysis are used to ‘splice’ together a large covariance matrix for the original system.

The method is explained for just two risk factor categories, then the generalisation to any number of risk factor categories is straightforward. Suppose there are n risk factors in the first system, say it is European equity indices, so risk factors are the equity market indices in n different European countries. Let the second system, European exchange rates say, have m risk factors, being the domestic/foreign exchange rates for the same n countries (so m = n or m = n-1 in this example depending on whether 'domestic' is a European country or not). More generally n and m can be anything; it is not the dimensions that matter. What does matter is that each system of risk factors is suitably co-dependent, so that it justifies the categorisation as a separate and coherent risk factor sub-system.

The first step is to find the principal components of each system, $\mathbf{P} = (P_1 , \dots P_r)$, and separately $\mathbf{Q} = (Q_1 ,\dots Q_s)$ where r and s are number of principal components that are used in the representation of each system. Denote by $\mathbf{A}$ (nxr) and $\mathbf{B}$ (mxs) the normalized factor weights matrices obtained in the PCA of the European equity and exchange rate systems respectively. Then the 'within factor' covariances, i.e. the covariance matrix for the equity system, and for the exchange rate system separately, are given by $\mathbf{AD}_1\mathbf{A}'$ and $\mathbf{B\,D}_2\mathbf{B}'$ respectively. Here $\mathbf{D}_1$ and $\mathbf{D}_2$ are the diagonal matrices of the variances of the principal components of each system. They may be estimated using exponentially weighted moving averages as described above, or GARCH models as described below.

The cross factor covariances are $\mathbf{ACB}'$ where $\mathbf{C}$ denotes the rxs matrix of covariances of principal components across the two systems, that is

$$\mathbf{C} = \{\mathrm{COV}(P_i, Q_j)\}$$

Then the full covariance matrix of the system of European equity and exchange rate risk factors is:

$$\begin{pmatrix} \mathbf{AD}_1\mathbf{A}' & \mathbf{ACB}' \\ (\mathbf{ACB}')' & \mathbf{BD}_2\mathbf{B}' \end{pmatrix}$$

The within factor covariance matrices $\mathbf{AD}_1\mathbf{A}'$ and $\mathbf{BD}_2\mathbf{B}'$ will always be positive semi-definite. But it is not always possible to guarantee positive semi-definiteness of the full covariance matrix of the original system, unless the off diagonal blocks $\mathbf{ACB}'$ are set to zero. This is not necessarily a silly thing to do; in fact it may be quite sensible in the light of the huge instabilities often observed in cross-factor covariances. If the risk model required non-zero cross-factor covariances, it would be possible to estimate the covariance between principal components of different risk factor sub-systems using exponentially weighted moving averages or bivariate orthogonal GARCH, giving the required estimate for **C**. But the full matrix would need to be checked for positive semi-definiteness, using a standard eigenvalue routine.

7. Summary and Conclusions

The main focus of this paper is to explain and empirically validate a new method of obtaining large GARCH correlation matrices using only univariate GARCH estimation techniques on principal components of the original return series. Empirical examples on commodity futures, interest rates and on international equity indices have been presented, and used to explain how best to employ the method in different circumstances. It is found that when the systems are suitably tailored, the orthogonal method compares very favourable to the general multivariate GARCH models. In many cases the divergence between the orthogonal GARCH estimates and the BEKK estimates is far less than between the VECH and the BEKK estimates.

The examples presented in this paper illustrate some of the many advantages of this methodology:

- Covariance matrices will always be positive semi-definite: there is no need to impose parameter constraints as in the RiskMetrics data or the standard multivariate GARCH models;
- Computational difficulties are kept to a minimum (since only univariate GARCH models are necessary) and computational time is very significantly reduced;
- Choosing fewer principal components to represent the system can control the amount of 'noise', and this can be advantageous in producing more stable correlation estimates;
- The method will produce volatilities and correlations for all variables in the system, including those for which direct GARCH estimation is computationally difficult;

The flexibility and accuracy of GARCH forecasting techniques place them in a unique position to fulfil many of the needs of risk management and traders. But without a feasible method for computing large covariance matrices using GARCH techniques, this potential will not be realised. Given the insurmountable problems in direct estimation of large GARCH covariance matrices, but given also the need for providing mean-reverting covariance forecasts for use in value-at-risk models and portfolio risk analysis, the modelling methodology presented in this paper is of significance.

The orthogonal GARCH model is based on one good idea: simply to capture the variability in a system of returns by a few uncorrelated causal factors and ascribe the rest of the variation to 'noise'. The use of orthogonal factors allows the GARCH optimization routines to be run only on univariate time series. This will considerably reduce computational complexity, and the full GARCH variances and covariances of the system are derived as simple transformations of the factor variances. The empirical examples presented here show that orthogonal GARCH may be calibrated quite easily so that variances and covariances coincide closely with those obtained from other models.

Whilst this paper provides a thorough empirical validation of orthogonal GARCH models for equities, commodities and interest rates it does not attempt to provide any theoretical results on the statistical properties of orthogonal GARCH models. It is hoped that the work presented here will motivate some readers towards more theoretical work on the model.

Professor Carol Alexander
Chair in Risk Management
ISMA Centre, The Business School for Financial Markets
Reading University,
Whiteknights, PO BOX 242
Reading, RG6 6BA

c.alexander@ismacentre.reading.ac.uk
www.ismacentre.reading.ac.uk
Direct Line: 00 44 1189 31 6431
General Line: 00 44 1189 31 6675

Bibliography

Alexander, C.O. (1992) "The world according to GARCH" RISK **5** no.8 120-125

Alexander C.O. (1996) "Evaluating the use of RiskMetrics™ as a risk measurement tool for your operation : What are its advantages and limitations" *Derivatives Use, Trading and Regulation* **2** No. 3 pp 277-285

Alexander, C.O. and C. Leigh (1997) "On the covariance matrices used in VaR models" *Journal of Derivatives* **4** No. 3 (Spring) pp 50-62

Alexander, C.O. (1998) "Volatility and correlation: Methods, models and applications" in *Risk Management and Analysis*: *Measuring and Modelling Financial Risk* (C. O. Alexander, Ed) Wileys

Alexander, C.O. (1999) "Correlation and cointegration in energy markets" in *Managing Energy Price Risk* (2nd Edition) RISK Publications pp291-304

Alexander, C.O. (2000) "Orthogonal methods for generating large positive semi-definite covariance matrices" *ISMA Centre Discussion Papers in Finance 2000-06* available from www.ismacentre.rdg.ac.uk

Alexander, C.O. (2001) " Orthogonal GARCH" in *Mastering Risk Volume 2* (C.O. Alexander, Ed.) Financial Times-Prentice Hall

Alexander, C.O. and A. Chibumba (1996) "Multivariate Orthogonal Factor GARCH" *University of Sussex Discussion Papers in Mathematics*

Baillie, R.T. and T. Bollerslev (1989) "The message in daily exchange rates: a conditional-variance tale". *Journal of Business and Economic Statistics* 7:3 pp 297-305.

Baillie, R.T. and T. Bollerslev (1990) "Intra-day and inter-market volatility in foreign exchange rates". *Review of Economic Studies* 58 pp 565-85.

Bollerslev, T. (1986) "Generalized autoregressive conditional heteroskedasticity". *Journal of Econometrics* 31 pp 307-327.

Bollerslev, T., R.F. Engle and D.B. Nelson (1994) "ARCH models" in *Handbook of Econometrics* vol. 4, ed. R.F. Engle and D.L. McFaddan, North-Holland

Brailsford, T.J. and Faff, R.W. (1996) "An evaluation of volatility forecasting techniques" *Journal of Banking and Finance* , 20:3 pp419-438

Dimson, E. and Marsh, P. (1990) "Volatility forecasting without data snooping" *Journal of Banking and Finance* , 14 pp399-421

Duan, JC (1995) "The GARCH option pricing model" *Mathematical Finance* 5:1, pp. 13-32.

Duan, JC (1996) 'Cracking the smile' *RISK* 9 pp 55- 59

Engle, R.F. (1982) "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation". *Econometrica* 50:4 pp 987-1007.

Engle, R.F. (2000) "Dynamic conditional correlation - A simple class of multivariate GARCH models" pdf version available from http://weber.ucsd.edu/~mbacci/engle/cv.html

Engle, R.F. and T.Bollerslev (1986) "Modelling the persistence of conditional variances". *Econometric Reviews* **5**:1 pp 1-50.

Engle, R.F. , Ng, V. and Rothschild, M. (1990) "Asset pricing with a factor ARCH covariance structure : Empirical estimates for treasury bills" *Journal of Econometrics* , 45 pp 213-238

Engle, R.F. and K.F. Kroner (1993) "Multivariate Simultaneous Generalized ARCH" *Econometric Theory* 11 pp122-150

Engle, R.F. and G.G.J. Lee (1993) "Long Run Volatility Forecasting for Individual Stocks in a One Factor Model" University of California, San Diego, Department of Economics Working Paper: 93-30, July 1993, p. 20.

Engle, R.F. and Mezrich J (1995) "Grappling with GARCH" *RISK*, 8:9 pp 112-117

Engle, R.F. and J. Rosenberg (1995) "GARCH gamma" *Journal of Derivatives* 2 pp47-59

Figlewski, S. (1994) "Forecasting volatility using historical data" New York University Salomon Center (Leonard N. Stern School of Business) Working Paper Series no. S-94-13

Kroner, K.F. and S. Claessens (1991) "Optimal dynamic hedging portfolios and the currency composition of external debt". *Journal of International Money and Finance* 10 pp 131-48.