

A discusión

FORECASTING THE CONDITIONAL COVARIANCE MATRIX OF A PORTFOLIO UNDER LONG-RUN TEMPORAL DEPENDENCE*

Trino-Manuel Níguez and Antonio Rubia**

WP-AD 2003-34

Correspondence to: Trino-Manuel Níguez, Universidad de Alicante, Departamento de Fundamentos de Análisis Económico, Campus de San Vicente del Raspeig, 03071 Alicante, Spain, E-mail: trino@merlin.fae.ua.es.

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.
Primera Edición Octubre 2003
Depósito Legal: V-4509-2003

IVIE working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

* We are grateful to Carol Alexander, Eva Ferreira and Juan Mora for helpful comments and discussions. The authors are responsible for any remaining errors. Financial support from the Spanish Ministry of Science and Technology through projects BEC2002-03097 and BEC2002-03797 is gratefully acknowledged.

** T.M. Níguez: Dpto. Fundamentos de Análisis Económico, Universidad de Alicante; A. Rubia: Dpto. de Economía Financiera, Universidad de Alicante.

FORECASTING THE CONDITIONAL COVARIANCE MATRIX OF A PORTFOLIO UNDER LONG-RUN TEMPORAL DEPENDENCE

Trino-Manuel Níguez and Antonio Rubia

ABSTRACT

Long-range persistence in volatility is widely modelled and forecasted in terms of the so-called fractional integrated models. These models are mostly applied in the univariate framework, since the extension to the multivariate context of assets portfolios, while relevant, is not straightforward. We discuss and apply a procedure which is able to forecast the multivariate volatility of a portfolio including assets with long-memory. The main advantage of this model is that it is feasible enough to be applied on large-scale portfolios, solving the problem of dealing with extremely complex likelihood functions which typically arises in this context. An application of this procedure to a portfolio of five daily exchange rate series shows that the out-of-sample forecasts for the multivariate volatility are improved under several loss-functions when the long-range dependence property of the portfolio assets is explicitly accounted for.

Key Words: Exchange Rates, Fractional Integration, Long Memory, MGARCH models, PCA.

JEL Classification: C22, C52, C53, F31.

1 Introduction

The dynamics of the time-varying volatility of financial assets play a main role in diverse fields, such as derivative pricing and risk management. Consequently, the literature focused on estimating and forecasting conditional variance is vast. The most popular method for modelling volatility belongs to the family of GARCH models (see Bollerslev *et al.* 1992 for a review of this topic), although other alternatives (such as stochastic volatility models) also provide reliable estimates. The success of GARCH processes is unquestionably tied to the fact that they are able to fit the stylized features exhibited by volatility in a fairly parsimonious and convincing way, through quite a feasible method. The seminal models developed by Engle (1982) and Bollerslev (1986) were rapidly generalized in an increasing degree of sophistication to reflect further empirical aspects of volatility.

One of the more complex features that univariate GARCH-type models have attempted to fit is the so-called long-memory property. The volatility of many financial assets exhibits a strong temporal dependence which is revealed through a slow decay to zero in the autocorrelation function of the standard proxies of volatility (usually squared and absolute valued returns) at long lags. The basic GARCH model does not succeed in fitting this pattern because it implicitly assumes a fast, geometric decay in the theoretical autocorrelations. Engle and Bollerslev (1986) were the first concerned with this fact and suggested an integrated GARCH model (IGARCH) by imposing unit roots in the conditional variance. The theoretical properties of IGARCH models, however, are not entirely satisfactory in fitting actual financial data, so further models were later developed to face temporal dependence. Ballie, Bollerslev and Mikkelsen (1996) proposed the so-called fractionally integrated GARCH models (FIGARCH) for volatility in the same spirit as fractional ARIMA models which were evolved for modelling the mean of time series (see Baillie, 1996). These models imply an hyperbolic rate of decay in the autocorrelation function of squared residuals, and generalize the basic framework by still using a parsimonious parameterization.

There has been a great interest in modelling the temporal dependence in the volatility of financial series, mostly in the univariate framework¹. The analysis of the long-memory property in the multivariate framework, however, has received much less attention, even though the estimation of time-varying covariances between asset returns is crucial for risk management, portfolio selection, optimal hedging and other important applications. The main reason is that modelling conditional variance in

¹An alternative approach for modelling long-memory through GARCH-type models is based on the family of stochastic volatility (see Breidt, Crato and de Lima, 1998). An extension of FIGARCH models has been considered in Ding, Granger and Engle (1993).

the multivariate GARCH framework is much harder than in the univariate case. The two main problems related to this issue are that the number of parameters to be estimated is a polynomial function of the system's dimension, and a large number of restrictions must be added to ensure a positive definite estimation for all possible realizations. They both complicate the parameter estimation considerably, and hence the long-memory property and even other relatively simpler features (like asymmetric and non-linear responses) are usually left out from multivariate GARCH models for the sake of tractability and despite their potential importance.

It is clear that embodying this empirical feature in the multivariate GARCH models would lead to gains in statistical efficiency in parameter estimation and, furthermore, it could improve the accuracy of covariance matrix forecasts and lead to better decisions on optimal hedging and portfolio allocation. Bollerslev and Mikkelsen (1996) early outlined the importance of long-run dependence in making such decisions. The price to be paid for this improvement, however, lies in a much higher technical difficulty, so any feasible procedure would be worthy of consideration. Consequently with this idea, the aim of this paper is to discuss a relatively simple model to forecast the conditional covariance of asset portfolios, accounting for long-run dependences in the conditional second moment. As this property has been observed mainly in foreign exchange rates, we focus on a portfolio formed by such assets for an empirical application of the model. The method used here combines the multivariate model by Alexander (2001) – based on the principal component analysis, and the univariate hyperbolic GARCH (HYGARCH) model recently proposed by Davidson (2003), which is intended to fit temporal dependences in quite a general setting. We therefore consider an extension of the so-called Orthogonal multivariate GARCH model to the context in which a portfolio is formed by assets exhibiting long-run temporal dependences in volatility.

As in the univariate framework, some improvements over GARCH processes should be expected when modelling explicitly the long-range dependences. We assess the quality of the multivariate model by its ability to make short-term forecast of *(i)* the daily conditional covariance matrix of the portfolio, and *(ii)* the daily overall variance of an equally-weighted portfolio over a period of 200 days. The alternative (baseline) model to this procedure employs the simpler GARCH model instead of the HYGARCH process, so no long-term dependences are acknowledged. In this way we can measure the relative importance of temporal dependences for forecasting the conditional covariance matrices. The Orthogonal GARCH models have been used in Byström (2000), among others.

The remainder of the paper is organized as follows. Section 2 discusses the multivariate model and the univariate specifications for the conditional variance. Section 3 presents the data and the analysis procedure. Section

4 reports the out-of-sample forecasting performance. Finally, concluding remarks are summarized in Section 5.

2 The multivariate modelling of long-memory

Although long-memory has been observed in the volatility of a wide range of assets, the literature on the topic is mainly focused on foreign exchange rate time series (FX hereafter). There exists a great deal of empirical literature focused on modelling and forecasting the volatility of exchange-rate returns in terms of the FIGARCH models in the univariate framework. An exhaustive review of the literature is beyond the aim of this paper. Some recent empirical works on this issue can be found in Vilasuso (2002) and Beine *et al.* (2002). On the other hand, the literature dealing with the multivariate case is scarce.

The modelling of long-memory in the multivariate framework was firstly studied by Teyssière (1997), who implemented several long memory volatility processes in a bivariate context, focusing on daily FX time series. He used an approach initially based on the multivariate constant conditional correlation model (Bollerslev, 1990), which allows for long-memory ARCH dynamics in the covariance equation. He also weakened the assumption of constant correlations and estimated time-varying patterns. Teyssière (1998) estimated several trivariate FIGARCH models on some intraday FX rate returns. This author finds a common degree of long-memory in the marginal variances, while the covariances do not share the same level of persistence with the conditional variances. More recently, Pafka and Mátyás (2001) analyzed a multivariate diagonal FIGARCH model on three FX time-series through quite a complex computational procedure. The multivariate modelling on other time series has focused on the crude oil returns (Brunetti and Gilbert, 2001). A bivariate constant correlation FIGARCH model is fitted on these data to test for fractional cointegration in the volatility of the NYMEX and IPE crude oil markets². To our knowledge, there is no other literature concerned with modelling temporal dependences in the multivariate context.

The previous research affords a valuable contribution to the better understanding of long-run dependences in multivariate volatility. A major shortcoming in applying these approaches in practice, however, lies in the overwhelming computational burden involved, which simply makes the straightforward extension of these methods to large portfolios unfeasible (note that only two or three assets are considered in the empirical applications of these methods). The procedure we shall discuss is specifically

²NYMEX is the New York Mercantile Exchange; the IPE is the International Petroleum Exchange located in London.

intended to model and forecast the time-varying covariance matrix of large portfolios.

2.1 The orthogonal multivariate model

We firstly introduce notation and terminology. Consider a portfolio of K financial assets and denote by $\mathbf{r}_t = (\mathbf{r}_{1t}, \mathbf{r}_{2t}, \dots, \mathbf{r}_{Kt})'$, $t = 1, \dots, T$, a weakly-stationary random vector with each component representing the return of each portfolio asset at time t . Denote by \mathcal{F}_t the set of relevant information up to time t , and define the conditional covariance matrix of the process by $\mathbb{E}(\mathbf{r}_t \mathbf{r}_t' | \mathcal{F}_{t-1}) = \mathbb{E}_{t-1}(\mathbf{r}_t \mathbf{r}_t') = \mathbf{H}_t$. Denote as $\mathbb{E}(\mathbf{r}_t \mathbf{r}_t') = \Omega$ the (finite) unconditional second order moment of the random vector. Note that only second-order stationarity is required, which is the basic assumption in the literature concerned with estimating covariance matrices of asset returns. Other procedures proposed for estimating the covariance matrix require much stronger assumptions (see, for instance, Ledoit and Wolf, 2003), as the existence of higher-order moments and even iid-ness in the driving series.

As the covariance matrix Ω is positive definite, it follows by the spectral decomposition that $\Omega = \mathbf{P}\Lambda\mathbf{P}'$, where \mathbf{P} is an orthonormal $K \times K$ matrix of eigenvectors, and Λ is a diagonal matrix with the corresponding eigenvalues of Ω in its diagonal. Lastly, assume that the columns of \mathbf{P} are ordered by size of the eigenvalues of Λ , so the first column is the one related to the highest eigenvalue, and so on.

The orthogonal model by Alexander is based on applying the principal component analysis (PCA) to generate a set of uncorrelated factors from the original series³. The PCA analysis is a well-known method widely used in practice, and several investment consultants, such as Advanced Portfolio Technologies, use procedures based on principal components. The basic strategy in the Alexander model consists of linearly transforming the original data into a set of uncorrelated latent factors so-called principal components whose volatility can then be modelled in the univariate framework. With these estimations, the conditional matrix \mathbf{H}_t is easily obtained by the inverse map of the linear transformation.

The set of principal components, $\mathbf{y}_t = (\mathbf{y}_{1t}, \mathbf{y}_{2t}, \dots, \mathbf{y}_{Kt})'$, is simply defined through the linear application $\mathbf{y}_t = \mathbf{P}'\mathbf{r}_t$. It follows easily that $\mathbb{E}(\mathbf{y}_t) = \mathbf{0}$ and $\mathbb{E}(\mathbf{y}_t \mathbf{y}_t') = \Lambda$ by the orthogonal property of \mathbf{P} . The columns of the matrix \mathbf{P} were previously ordered according to the corresponding eigenvalues size, so that ordered principal components have a decreasing ability to explain the total variability and the main sources of variability

³This procedure belongs to the field of the factor GARCH models introduced by Engle, Ng and Rothschild (1990), and takes the principal components as factors. See also Ding (1994), Alexander and Chibumba (1998).

can be identified⁴. Under the crucial assumption that the conditional covariance of \mathbf{y}_t , say \mathbf{V}_t , is also a diagonal matrix, this could be approached by $\mathbb{E}_{t-1}(\mathbf{y}_t\mathbf{y}_t') = \mathbf{P}'\mathbf{H}_t\mathbf{P}$, so it immediately follows that the conditional covariance matrix of the original data can be estimated by using the inverse map of the linear application and hence $\hat{\mathbf{H}}_t = \hat{\mathbf{P}}\hat{\mathbf{V}}_t'\hat{\mathbf{P}}$. It is worth recalling that $\hat{\mathbf{V}}_t$ is determined by estimating the conditional variances of each orthogonal factor \mathbf{y}_{it} in a univariate setting, and $\hat{\mathbf{P}}$ is characterized by the spectral decomposition of the sample estimator of Ω .

This methodology has several great advantages. Firstly, it allows us to circumvent the computational burden that arises in the optimization of a multivariate log-likelihood function, since all estimations are actually made in the univariate context. Moreover, the covariance matrix forecasts are obtained in a simple fashion by making projections from the inverse map, $\mathbb{E}_T(\mathbf{H}_{T+s}) = \hat{\mathbf{P}}\mathbb{E}_T(\hat{\mathbf{V}}_{T+s})\hat{\mathbf{P}}'$. Secondly, the number of parameters to be estimated, which typically grows at an exponential rate, becomes here a linear function in K . Finally, this procedure is especially indicated for large portfolios because system reduction is a straightforward possibility under PCA. In this case, a sufficient number of principal factors should be included to explain a sufficiently high proportion of common variance. On the other hand, the orthogonal method is not free of shortcomings. As is usual in the PCA, it is generally not possible to provide an economical meaning for the principal components; furthermore, assuming that \mathbf{V}_t is a diagonal matrix is actually a very strong assumption, which does not necessarily have to be fulfilled in practice⁵. Loosely speaking, all the multivariate GARCH models usually assume strong restrictions which are not satisfied by the actual behaviour of the underlying series, but it is necessary to impose them to guarantee feasibility in the analysis.

The Orthogonal procedure is completed when a suitable univariate model is used for modelling the volatility of every factor. In the basic framework, the GARCH(1,1) model is used and the whole methodology is called Orthogonal GARCH (OGARCH). Recall that the GARCH(1,1) equation, applied on the i -th conditional variance ($i = 1, \dots, K$), is given by

$$h_{it} = \omega_i + \alpha_i y_{it-1}^2 + \beta_i h_{it-1}, \quad \omega_i > 0, \quad \alpha_i, \beta_i \geq 0, \quad \alpha_i + \beta_i < 1 \quad (1)$$

⁴If $\Lambda = \{\lambda_i\}_{i=1}^K$ is the diagonal matrix of eigenvalues and \mathbf{P} the related matrix of eigenvectors, the i -th principal component in the set $\mathbf{Y}_t = \mathbf{P}'\mathbf{r}_t$ explains a proportion of total variance given by $\lambda_i/\text{trace}(\Lambda)$.

⁵Note, however, that Alexander (2001) found that the degree of accuracy which is lost by assuming conditional diagonality is usually not high in empirical applications. Engle (2002) shows through Monte Carlo experimentation that OGARCH procedure produces similar results to those from fitting full multivariate GARCH models. The assumption of conditional diagonality is crucial and cannot be weakened, because it allows us to generate large GARCH-type covariance matrices, which is the whole point of this procedure.

where the parameter restrictions are needed to ensure both the positivity of the conditional variance and covariance stationarity. It should be noted that the basic framework of this approach could straightforwardly be extended by using more complex models if the data features make it advisable. As we are interested in modelling the dynamics of assets which are known to display long temporal dependences, we need to focus on a more general specification than the GARCH(1,1) model.

2.2 Fitting long-memory dynamics: The HYGARCH model

FIGARCH models are usually used for modelling volatility persistence in the univariate context. Nevertheless, they are not methodologically suitable in the framework of PCA, since they directly imply the inexistence of the unconditional second-order moment⁶. Only models allowing for covariance-stationarity in their parameter range can be embedded in the orthogonal method and hence proper parameter restrictions are needed in using this procedure⁷. Note that the GARCH(1,1) model would be only suitable under the restriction $\alpha_i + \beta_i < 1$. However, this imposes a geometric decay and hence only allows for short-term temporal dependences, so the basic OGARCH procedure is unable to forecast the patterns properly due to long-memory in volatility.

HYGARCH models allows for stationarity and simultaneously long-memory under some constraints, still using a parsimonious parameterization. Furthermore, they generalize the basic GARCH process as well as the FIGARCH model, so embedding this in the Orthogonal procedure is an extension of the basic methodology which may properly accommodate long-run dependences in the conditional second moment of a portfolio. We discuss briefly the main features of the HYGARCH model below, although a deeper discussion and an empirical analysis can be found in Davidson (2003). Some technical aspects are further discussed in Appendix A.

Consider the general formulation of a GARCH(p,q) model for the conditional variance process of a time-series ε_t ,

$$h_t = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)h_t \quad (2)$$

with L being the usual lag operator $\{L^j z_t \equiv z_{t-j}\}$ and $\omega > 0$. If the polynomial $\beta(L)$ has all its roots inside the unit circle, the former equation

⁶FIGARCH models are weakly stationary only in the trivial case of $d = 0$, corresponding, hence, with the stationary GARCH model.

⁷Although there exist a vast literature focused on testing for unit roots in the driving process of the conditional mean, this subject is complicated in the context of volatility, since it is never observed directly. Therefore, suitable testing has to be carried out with parametric models. Of course, this procedure has severe limitations, because the veracity of conclusions are conditional to the correct specification of the model. Integration or even explosive patterns when fitting GARCH-type model often suggest model misspecification.

can be arranged for the sake of convenience as follows,

$$h_t = \omega[1 - \beta(1)]^{-1} + \left([1 - \beta(L)]^{-1} \alpha(L) \right) \varepsilon_t^2 \quad (3)$$

The process ε_t is weekly stationary when the characteristic polynomial $\lambda(L) = 1 - \alpha(L) - \beta(L)$ has all its roots inside the unit circle. If there are unit roots, the equation would correspond to the IGARCH(p, q) model and it could be rewritten as,

$$h_t = \omega[1 - \beta(1)]^{-1} + [1 - \phi(L) [1 - \beta(L)]^{-1} (1 - L)] \varepsilon_t^2 \quad (4)$$

where $\phi(L) = \lambda(L) (1 - L)^{-1}$. The integration of the conditional variance implies that the long run predictions tend towards infinity following a straight line with a positive slope, which is not a very realistic assumption for the volatility of financial assets. The FIGARCH(p, d, q) model generalizes the idea of integer unit roots allowing for fractional unit roots. This is achieved by replacing the difference operator $(1 - L)$ by the fractional polynomial $(1 - L)^d$, $0 \leq d \leq 1$, in the above equation,

$$h_t = \omega[1 - \beta(1)]^{-1} + \left(1 - \phi(L) [1 - \beta(L)]^{-1} (1 - L)^d \right) \varepsilon_t^2 \quad (5)$$

It should be remarked that FIGARCH(p, d, q) models are not covariance-stationary, although their theoretical properties differ from those of IGARCH models. Specifically, FIGARCH processes are characterized through theoretical autocorrelations decaying toward zero at a polynomial rate. This decay is so slow that the autocorrelations are not absolutely summable and, therefore, the unconditional variance is not well-defined. Davidson (2003) proposed the HYGARCH equation as a generalization of the FIGARCH, allowing for a (possible) faster non-geometric rate of decay for which weakly stationarity would still be possible. The model is given by

$$h_t = \frac{\omega}{1 - \beta(1)} + \left(1 - \phi(L) [1 - \beta(L)]^{-1} \left[1 + \theta[(1 - L)^d - 1] \right] \right) \varepsilon_t^2 \quad (6)$$

If $0 \leq \theta < 1$ and the GARCH component of the model observes the usual covariance stationary restriction, which imply $\phi(L) / \beta(L) > 0$, the process is covariance stationary and their cumulative impulse response weights decrease quicker towards zero than the FIGARCH model.

3 Data and empirical procedure

The set of data analyzed in this paper consists of five currencies versus of the U.S. dollar. It includes daily quotes of the Deutsch mark (DEM),

the Japanese yen (JPY), the Britain pound (GBP), the Swiss franc (SWF) and the Swedish krona (SDK). The original data is transformed into daily returns by taking the usual logarithm differences, $r_{it} = \ln(P_{it}/P_{it-1})$. Some usual descriptive statistical information is reported in Table 1.

It can be seen that the main features of the return series include a similar degree of unconditional univariate variance, thick-tailed distributions and some degree of autocorrelation. Although the empirical literature on FX typically finds no linear temporal dependences, some serial autocorrelation is nevertheless detected in the sample⁸. A deeper study of the returns time-series shows not only the presence of serial autocorrelation but also first-order cross-correlation. We shall therefore take this multivariate pattern into account as a previous step of the main procedure. The whole sample is split into two sub-samples. The out-of-sample period includes the last 210 observations (roughly a whole year of trading days) and the in-sample period takes the first T=4672 observations. The in-sample period is then used to calibrate parameters and make forecasts following a rolling scheme over the out-of-sample period. The whole procedure employed to forecast the portfolio conditional covariance matrix is sketched as follows:

- i) First, the cross-correlation component observed in the multivariate mean is filtered by using a vector autoregressive (VAR) regression. This well-known method is intended to fit the univariate structure of each series as well as any possible interdependence among them. After comparing several alternatives through the standard information criterion, a VAR(1) model proves to be sufficient to fit the predictable component of the system⁹. The original series are hence decomposed into two terms, mutually, orthogonal as follows,

$$\mathbf{r}_t = \hat{\boldsymbol{\mu}}_t + \boldsymbol{\varepsilon}_t; \quad \boldsymbol{\mu}_t = \Phi \mathbf{r}_{t-1} \quad (7)$$

where the first term $\boldsymbol{\mu}_t$ denotes a cross-predictable term following a drift-less autoregressive process, and $\boldsymbol{\varepsilon}_t$ denotes the unpredictable component of the returns. Note that volatility refers to the (conditional) variance of the unpredictable term, so we focus hereafter on the residuals of this regression through this approach.

- ii) Second, the PCA is applied on the residuals $\hat{\boldsymbol{\varepsilon}}_t$ from the former regression. The linear transformation of the Orthogonal method is then applied to generate the set of orthogonal factors, and both the

⁸The same fact is observed in Beine et al., (2002) in using a similar data base as the one used here.

⁹The results regarding VAR modelling, estimation and checking are not presented for the sake of space, but they are available from the authors upon request.

Table 1
Descriptive statistics

Returns	DM	GBP	JPY	SWF	SWK
Mean	-0.0002	-0.0001	0.0002	-0.0004	0.0006
Std. Dev.	0.067	0.068	0.063	0.073	0.062
Skewness	-0.128	-0.506	-0.013	-0.198	0.284
Kurtosis	4.839	7.251	6.232	4.671	8.258
$\rho(1)$	0.05 (14.1)	0.03 (7.5)	0.03 (5.2)	0.07 (26.0)	-0.03 (4.3)
$\rho(2)$	-0.01 (15.2)	0.00 (7.9)	0.02 (8.4)	0.00 (26.4)	-0.00 (4.4)
$\rho(3)$	-0.02 (18.2)	0.00 (8.1)	-0.01 (9.1)	-0.02 (28.5)	0.00 (4.4)
$\rho(4)$	-0.00 (18.2)	0.00 (8.3)	0.00 (9.3)	0.01 (29.7)	0.00 (4.5)
$\rho(5)$	-0.29 (22.2)	0.01 (9.9)	-0.01 (9.8)	0.02 (33.3)	0.00 (5.0)

The table reports descriptive statistics for the daily FX returns of the Deutsch mark (DEM), the Japanese yen (JPY), the Britain pound (GBP), the Swiss franc (SWF) and the Swedish krona (SDK) in relation to the US dollar. The statistics $\rho(i)$ denote the i -th autocorrelation (Ljung-Box Q-statistics in parenthesis).

GARCH(1,1)¹⁰ and the HYGARCH(1, d ,1) models are then estimated by (quasi) maximum likelihood (QML) under the assumption of conditional normality¹¹. Stationarity constraints are imposed in the optimization procedure. The method combining the multivariate methodology and the HYGARCH model shall be referred to as OHYGARCH throughout the remainder of the paper.

- iii) Once the estimation stage is completed, the conditional covariance matrix is forecasted for the s th-step-ahead day ($s = 1, 2, \dots, 10$) from the last day of the in-sample period¹². This procedure is then repeated 200 times by using a rolling-scheme. At each iteration, the in-sample window is updated one day and the oldest observation is removed, so the window length always remains constant over the process. Models are completely reestimated and forecasts are made again from the last observation in the in-sample period through an iterative process. We denote by $\mathbb{E}_{T+n}^s(\mathbf{H}_{T+n+s})$ the forecast for the s th-step-ahead day made at day $T+n$, ($s = 1, \dots, 10$; $n = 0, \dots, 199$).

As is standard procedure in the forecasting framework, we also consider the ‘naive’ forecasts from an unconditional model. These forecasts are constructed following a random walk scheme on the basis of the unconditional variance from the in-sample period, i.e., Ω .

- iv) Several symmetric and asymmetric loss-functions are then considered in order to assess the forecasting performance of models. The forecast error for each day, say \mathbf{U}_{T+n}^s , is given by the difference between the matrix forecast, $\mathbb{E}_{T+n}^s(\mathbf{H}_{T+n+s})$, and the realized covariance matrix in the corresponding day, which is usually approached by the cross product of the realized errors, $\hat{\varepsilon}_{T+n+s}\hat{\varepsilon}_{T+n+s}'$.
- v) Finally, when the forecast construction stage is ended, the forecasting performance of models is assessed in both the multivariate and the univariate framework. We consider the daily forecasts for the multivariate conditional covariance matrix, on the one hand, and the overall variance of an equally-weighted portfolio, on the other hand, to compare the out-of-sample forecast accuracy of the models involved.

¹⁰It is worth noting that, recently, Engle (2002) has shown the relative good performance of the OGARCH methodology to estimate correlations over a wide variety of processes, including: Two MGARCH models, two integrated dynamic conditional correlations (DCC) models, one mean reverting DCC process, the exponential smoother from Riskmetrics and the familiar 100-day moving average model.

¹¹It should be noted that there is not a formal proof of consistency of the QML for long-memory processes. The polynomial expansion related to the fractional integration $(1-L)^d$ is expanded with a truncation lag equal to 1000, as it is the standard procedure.

¹²Empirical work has shown that the forecasting performance of GARCH models decreases considerably for periods longer than ten-day horizons (see Christoffersen, Diebold and Schuermann, 1998).

Specifically, we consider several matrix norms (Fobrenius and p -norms) to assess the ability of the different procedures in forecasting the whole conditional covariance matrix. Under such metrics, all elements in the forecasted matrix are equally important, so that models are penalized if, say, they have a greater ability for forecasting the on-diagonal elements but a poor performance for accurately predicting covariances. Besides, we consider ordinary univariate measures (mean squared error, mean absolute error and several asymmetric absolute mean errors) to gauge errors when forecasting the overall variance of an equally-weighted portfolio. Appendix B provides a comprehensive description of these measures.

Finally, the forecasting performance of the OGARCH and the OHY-GARCH is compared in statistical terms by means of the test proposed by Diebold and Mariano [DM] (1995). This test assumes no differences between the loss-functions of two alternative models under the null hypothesis. The null is rejected for large values of the statistic $DM_s = \bar{x} / \sqrt{2\pi f_x(\omega = 0) / N}$, where \bar{x} is the sample mean of the differences in the forecast errors, $f_x(\omega = 0)$ is the spectral density function of the forecast error differences evaluated at the zero frequency (long run variance) and $N = 200$ is the number of forecasts over the out-of-sample period. This statistic is asymptotically distributed as a standard normal random variable under the null. We have computed $f_x(\omega = 0)$ using the heteroskedasticity and autocorrelation consistent estimator of Newey and West (1987) with bandwidth parameter $s - 1$ for forecasting horizons $s > 1$, $s = 1, \dots, 10$.

4 Forecasting performance

The main results of the PCA analysis are shown in Table 2. It can be seen that the sorted principal components are able to explain about 73%, 13%, 6%, 5% and 1% of common variance, respectively. It is worth remarking the importance of the first factor, which reveals a large amount of common variance underlying the time series. We shall consider, later, the system reduction based on considering solely this factor, since a good approach could be expected from such a high predictive power. A casual look at the descriptive statistics of the factors shows that they become more leptokurtic as their descriptive ability gets lower. It is clear that the assumption of normality does not hold on these series and, furthermore, it is hard to assume that the same data-generating process underlies every factor. Nevertheless, as the usual procedure in the Orthogonal methodology applies the same specification to each factor, we proceed in such a spirit.

Table 3 reports the QML estimates from the OGARCH and OHY-GARCH models in the in-sample period. The results show a marked improvement in the in-sample fit of the OHYGARCH over the GARCH model,

Table 2
Principal Component Analysis

Factors	F ₁	F ₂	F ₃	F ₄	F ₅
$\widehat{\Omega}$	0.0046	0.0024	0.0031	0.0046	0.0033
		0.0047	0.0018	0.0027	0.0018
			0.0040	0.0032	0.0025
				0.0054	0.0033
					0.0038
Eigenvalue	0.0165	0.0030	0.0015	0.0012	0.0003
Proportion	73.02%	13.29%	6.76%	5.46%	1.47%
Cumulative	73.02%	86.31%	93.07%	98.53%	100%
Mean	0.0002	0.0001	0.0000	0.0000	0.0000
Std. Dev.	0.128	0.054	0.039	0.035	0.018
Skewness	-0.168	0.375	-0.243	1.842	0.115
Kurtosis	4.874	6.917	6.240	37.184	6.112

The table presents the results of the Principal Component Analysis for the daily exchange-rate returns. The values of the in-sample unconditional covariance matrix ($\widehat{\Omega}$) are presented as well as several information regarding the results of the principal component analysis. The rows *proportion* refers to the amount of common variance explained by each factor; *cumulative* is the accumulative proportion of common variance explained up to the i-th factor. Finally some usual statistical information on the resulting factors is presented.

Table 3
QML estimates

OGARCH(1,1)					
	F ₁	F ₂	F ₃	F ₄	F ₅
ω_i	0.000 (3.48)	0.000 (3.65)	0.000 (3.14)	0.000 (3.12)	0.000 (3.09)
α_i	0.060 (7.53)	0.065 (4.41)	0.092 (3.67)	0.154 (4.65)	0.123 (2.54)
β_i	0.914 (71.7)	0.928 (51.7)	0.890 (14.7)	0.845 (65.5)	0.831 (17.8)
AIC _{<i>i</i>}	-1.339	-3.167	-3.770	-4.113	-5.271
OHYGARCH(1, <i>d</i> ,1)					
	F ₁	F ₂	F ₃	F ₄	F ₅
ω_i	0.000 (2.78)	0.000 (2.42)	0.000 (2.56)	0.000 (2.34)	0.000 (2.97)
θ_i	0.986 (115)	0.999 (122)	0.999 (116)	0.999 (125)	0.999 (1505)
β_i	0.743 (7.43)	0.574 (6.29)	0.689 (8.32)	0.916 (8.25)	0.623 (5.75)
ϕ_i	0.261 (2.66)	0.301 (8.83)	0.450 (5.87)	0.402 (2.41)	0.416 (5.42)
d_i	0.544 (2.83)	0.367 (8.63)	0.264 (5.61)	0.793 (2.34)	0.366 (5.51)
AIC _{<i>i</i>}	-1.345	-3.177	-3.785	-4.164	-5.284

The reported coefficients shown in each row of the table are QML estimates of the GARCH(1,1) and HYGARCH(1,*d*,1) models for the principal component series, F_i , on the VAR(1) residuals on original returns. Robust QML t-statistics are given in parenthesis below the parameter estimates. AIC_{*i*} denotes the usual Akaike information criterion statistic.

as measured by the increase in the log-likelihood function and the standard information criteria, so the former model provides a better representation of the underlying dynamics. The sum of the GARCH parameters is very close to one, showing thus a strong degree of persistence in the data involved. The estimated amplitude parameter θ in the HYGARCH model is close to the boundary of the admissible parameter space. This suggests evidence of non-stationarity in the driving process of the asset returns, although it should be noted that if the true value of the parameter is close to the knife-edge limit (i.e., the process is near-nonstationary), inferring whether the process is stationary or not on the basis of point estimates is a really problematic task, as it is a well-documented fact in the unit root literature.

On the other hand, volatility non-stationarity when estimating parametric models is sometimes regarded as a matter of model misspecification. It is likely that the parameter estimates of the HYGARCH model are sensitive to the omission of relevant structures in the data involved, as in the general case of models in the GARCH family. Brooks (2001) showed the presence of highly non-linear effects on the volatility of FX series which were fitted by a double-threshold GARCH model. This sort of structure, which might be present in the multivariate dynamics of the series, is neglected in fitting the above models. In this way, further extensions of the HYGARCH model facing these kinds of effects together with long-run persistence would likely yield better in-sample results, so further improvements on this model are deserved.

Nevertheless, it is clear that the HYGARCH model is much more than an acceptable first-cut approximation, since it improves the in-sample performance of the model through the better fit of the conditional multivariate dynamics. As the main aim of this paper is concerned with the forecasting ability of the models involved, it is therefore worth assessing whether accounting for long-memory through this model leads to more accurate forecasts. Out-of-sample forecast accuracy provides a better and potentially more useful comparison of alternative models (López, 2001).

The out-of-sample forecast behavior is summarized in Tables 4 and 5. The OHYGARCH procedure beats the OGARCH method significantly according to all univariate and multivariate measures, as the forecast error is reduced for all the forecasting horizons. The difference is particularly strong and significant for some metrics (such as the univariate mean absolute error, the asymmetric mean absolute error penalizing harder overpredictions and the Fobrenius norm). In the remaining measures, the evidence supporting an overperformance of the OHYGARCH is not so strong, although there is a clear and significant trend to reduce the forecast errors as the forecasting horizon grows. This is really important, because it is clear that the effect of any long-run dependence on the forecast accuracy should be more remarkable when forecasting for distant periods.

Table 4
Out-of-sample performance

OGARCH(1,1)							
s	MAE_s	MSE_s	$MAEU_s$	$MAEO_s$	$Pr(U)_s$	NF_s	$N1_s$
1	59.06	8.516	100.0	179.2	0.245	0.767	139.9
2	58.85	8.418	98.39	180.6	0.250	0.763	139.8
3	58.68	8.379	96.45	181.9	0.240	0.758	139.2
4	59.82	8.405	98.65	184.0	0.235	0.760	140.1
5	60.57	8.389	99.68	186.4	0.235	0.756	140.3
6	61.10	8.581	100.2	188.0	0.240	0.764	140.3
7	61.69	8.623	100.4	189.8	0.240	0.765	140.7
8	61.92	8.624	101.6	189.1	0.250	0.766	141.3
9	62.06	8.712	100.9	189.4	0.245	0.770	141.6
10	62.20	8.702	101.4	189.5	0.240	0.772	141.9
OHYGARCH(1, d ,1)							
s	MAE_s	MSE_s	$MAEU_s$	$MAEO_s$	$Pr(U)_s$	NF_s	$N1_s$
1	57.89	8.438	100.3	174.5	0.260	0.763	138.6
2	57.65	8.334	98.24	176.3	0.250	0.759	138.4
3	57.53	8.317	97.07	177.3	0.240	0.755	137.9
4	58.28	8.337	99.76	178.2	0.250	0.757	138.5
5	58.92	8.307	99.97	180.5	0.250	0.752	138.7
6	59.43	8.500	100.0	182.5	0.240	0.759	138.6
7	59.94	8.536	100.2	184.0	0.240	0.760	139.0
8	60.21	8.533	101.9	183.4	0.255	0.761	139.6
9	60.37	8.625	101.7	183.6	0.255	0.765	139.9
10	60.47	8.603	102.2	183.2	0.260	0.766	140.1

Out-of-sample forecasting performance of the OGARCH and the OHYGARCH models. Forecasts are made for the s th-step-ahead day over the out-of-sample period. The forecasting ability is compared in terms of predicting the whole covariance matrix (NF_s and $N1_s$ denote the Fobrenius norm and the 1-norm, respectively) and the overall variance of the equally-weighted portfolio [mean absolute error (MAE_s), mean squared error (MSE_s), and some asymmetric measures penalizing harder overpredictions ($MAEO_s$) and underpredictions ($MAEU_s$)]. The column $Pr(U)_s$ denotes the binomial probability of underprediction for the variance of the equally-weighted portfolio. Loss-function values are scaled by 100.

Table 5
DM statistics of the performance of the OHYGARCH against
the OGARCH models

s	MAE_s	MSE_s	$MAEU_s$	$MAEO_s$	NF_s	$N1_s$
1	3.89	2.00	-0.42	5.56	1.67	5.40
2	3.90	2.13	0.26	5.04	2.25	5.88
3	4.22	1.58	-1.01	7.76	1.63	6.43
4	5.81	1.54	-0.52	7.49	1.51	7.57
5	6.11	1.73	-0.15	7.92	1.69	7.89
6	6.01	1.52	0.55	8.40	1.62	8.54
7	6.23	1.55	0.44	8.72	1.69	8.54
8	6.06	1.56	-0.57	8.16	1.63	7.78
9	5.72	1.40	-1.15	8.05	1.62	7.84
10	6.51	1.61	-1.29	8.89	1.75	8.35

The table presents t-statistics from the test of Diebold and Mariano (1995). The test assumes no differences between the forecasting ability of the OHYGARCH and the OGARCH models under each loss-function. The forecasting ability is compared in terms of predicting the whole covariance matrix (NF_s and $N1_s$ denote the Fobrenius norm and the 1-norm, respectively) and the overall variance of the equally-weighted portfolio [mean absolute error (MAE_s), mean squared error (MSE_s), and some asymmetric measures penalizing harder overpredictions ($MAEO_s$) and underpredictions ($MAEU_s$)].

With regard to the asymmetric error measures, the OHYGARCH model strongly outperforms the OGARCH approach when the overprediction is penalized harder, while there are no significant differences when underprediction is penalized. Therefore, the upward bias is higher in the GARCH model as it finds more problems to adapt to the effects caused by extreme observations. Finally, either the OGARCH or the OHYGARCH beats easily the naive forecasts based on unconditional estimations (see Table 6) when forecasting in distant periods¹³. For the sake of space, we only present the comparison for the OHYGARCH case, although all results are available from the authors upon request. Overall, the results presented in this section allow for a clearer discrimination among the competing models, providing evidence supporting the forecasting superiority of the OHYGARCH model.

4.1 System reduction

As there exists a great deal of common variance which is successfully captured by very few factors, it is valuable to consider the system reduction. The first factor is able to explain over 73% of common variance, which seems high enough to provide a good approach (Alexander, 2001). We therefore consider a single factor in the same forecasting scheme as above¹⁴. Since only forecasts on a sole time-series are needed, the procedure is considerably faster than before. The forecasted conditional covariance matrices are now given by

$$\mathbf{H}_t^{(1)} = h_{1t} \mathbf{P} \mathbf{B} \mathbf{P}' \quad (8)$$

where \mathbf{B} is a ($K \times K$) matrix of zeros with its 1,1-element being equal to one, and h_{1t} is the conditional variance forecast of the first factor.

The results on this stage are reported in Tables 7 and 8. They show again the outperformance of the HYGARCH method over the GARCH model, although forecast errors are higher than when the whole system is considered. It should be noted that the predominance of the OHYGARCH over the OGARCH model in this context means that there exists a strong degree of common long-run dependence among the series and, furthermore, that it is important enough to belong to the first factor.

¹³The naive model can outperform the conditional-based model in the one-ahead predictions. This is not very surprising and it is due to the similarity between the estimation of the realized covariance matrix and the forecasted one. The mechanism is nearly the same as it is based on the OLS estimations of the VAR model, as it only differs in two observations. Therefore, the one-step-ahead estimation is really close to the proxy of realized volatility, but the analysis in the longer periods shows that it is really an artifact as the validity of the naive benchmark is rapidly and strongly rejected.

¹⁴We have also considered system reduction based on the two most important factors covering over 86% of total variance. The results are not qualitatively different from that shown here, so they are not presented. They are available upon request.

Table 6
Out-of-sample performance of the OHYGARCH against the naive models

Naive						
s	MAE _s	MSE _s	MAEU _s	MAEO _s	NF _s	N1 _s
1	0.001	0.000	1.598	1.523	0.000	2.283
3	162.2	26.31	162.3	402.5	1.156	197.8
5	324.4	105.2	324.6	569.8	4.628	278.4
7	486.7	322.4	486.1	697.8	10.46	341.5
10	730.1	533.6	730.0	854.2	23.43	418.5
DM Test						
s	MAE _s	MSE _s	MAEU _s	MAEO _s	NF _s	N1 _s
1	-11.4	-3.26	-10.1	-25.3	-3.88	-32.1
3	21.7	7.16	7.46	34.8	2.04	12.8
5	55.8	39.9	23.4	66.6	20.1	27.4
7	79.2	104.1	38.3	98.2	45.7	36.3
10	114.2	158.2	49.0	139.2	93.8	47.7

Out-of-sample forecasting performance of the OHYGARCH model against a naive procedure. Forecasts are made for the s th-step-ahead day over the out-of-sample period. The forecasting ability is compared in terms of predicting the whole covariance matrix (NF_s and N1_s denote the Fobrenius norm and the 1-norm, respectively) and the overall variance of the equally-weighted portfolio [mean absolute error (MAE_s), mean squared error (MSE_s), and some asymmetric measures penalizing harder overpredictions (MAEO_s) and underpredictions (MAEU_s)]. The lower part of the table presents the t -statistics from the test of Diebold and Mariano (1995). The test assumes no differences between the forecasting ability of the OHYGARCH model and the naive method under each loss-function. Loss-function values are scaled by 100.

Table 7
Out-of-sample performance with the first factor

OGARCH(1,1)							
s	MAE_s	MSE_s	$MAEU_s$	$MAEO_s$	$Pr(U)_s$	NF_s	$N1_s$
1	59.72	8.551	100.1	181.6	0.245	1.352	150.0
2	59.51	8.456	98.01	183.1	0.240	1.357	149.8
3	59.37	8.419	95.92	184.4	0.225	1.358	149.7
4	60.55	8.445	98.53	186.8	0.220	1.365	150.1
5	61.29	8.431	99.74	188.8	0.235	1.368	150.4
6	61.82	8.630	100.3	190.3	0.240	1.382	150.7
7	62.42	8.674	100.3	192.0	0.235	1.389	150.9
8	62.64	8.674	101.4	191.5	0.240	1.397	151.5
9	62.80	8.765	100.7	191.8	0.240	1.412	152.0
10	62.97	8.752	101.3	192.0	0.235	1.419	152.5

OHYGARCH(1, d ,1)							
s	MAE_s	MSE_s	$MAEU_s$	$MAEO_s$	$Pr(U)_s$	NF_s	$N1_s$
1	58.50	8.465	100.0	176.9	0.255	1.301	148.4
2	58.30	8.364	98.14	179.0	0.240	1.302	148.3
3	58.19	8.348	97.12	179.8	0.240	1.300	147.9
4	58.93	8.368	98.64	180.6	0.245	1.304	148.5
5	59.60	8.339	99.68	183.0	0.240	1.303	148.3
6	60.12	8.538	100.2	184.7	0.240	1.315	148.6
7	60.63	8.576	100.3	186.2	0.240	1.318	148.8
8	60.85	8.572	101.9	185.5	0.255	1.324	149.4
9	61.03	8.666	101.6	185.7	0.255	1.336	149.8
10	61.05	8.641	101.8	185.2	0.255	1.341	150.3

Out-of-sample forecasting performance of the OGARCH and the OHYGARCH models. Forecasts are made for the s th-step-ahead day over the out-of-sample period. The forecasting ability is compared in terms of predicting the whole covariance matrix (NF_s and $N1_s$ denote the Frobenius norm and the 1-norm, respectively) and the overall variance of the equally-weighted portfolio [mean absolute error (MAE_s), mean squared error (MSE_s), and some asymmetric measures penalizing harder overpredictions ($MAEO_s$) and underpredictions ($MAEU_s$)]. The column $Pr(U)_s$ denotes the binomial probability of underprediction for the variance of the equally-weighted portfolio. Loss-function values are scaled by 100.

Table 8
DM statistics of the performance of the OHYGARCH against
the OGARCH models using the first factor

s	MAE _s	MSE _s	MAEU _s	MAEO _s	NF _s	N1 _s
1	3.98	2.14	-0.13	5.31	6.86	7.73
2	3.92	1.76	0.13	4.90	5.86	6.91
3	4.25	1.75	1.48	8.13	8.87	9.93
4	6.05	1.70	0.19	8.17	8.61	9.84
5	6.20	1.89	-0.15	8.15	8.23	9.60
6	6.05	1.67	-0.28	8.51	7.88	9.43
7	6.28	1.70	0.08	8.82	7.81	9.50
8	6.17	1.72	0.69	8.20	7.74	9.66
9	5.87	1.56	1.30	7.94	7.45	10.1
10	6.73	1.85	0.97	8.28	7.46	9.94

The table reports t-statistics from the test of Diebold and Mariano (1995). The test assumes no differences between the forecasting ability of the OHYGARCH and the OGARCH models under each loss-function. The forecasting ability is compared in terms of predicting the whole covariance matrix (NF_s and N1_s denote the Fobrenius norm and the 1-norm, respectively) and the overall variance of the equally-weighted portfolio [mean absolute error (MAE_s), mean squared error (MSE_s), and some asymmetric measures penalizing harder overpredictions (MAEO_s) and underpredictions (MAEU_s)].

5 Concluding remarks

Long-run temporal dependences in the volatility of financial time-series has been much emphasized in the recent econometric literature. Univariate modelling shows large forecast improvements over standard GARCH models when accounting for such dynamics. The study in the multivariate context in which assets with long-memory volatility are integrated into portfolios is not straightforward, but there is no doubt that it is of great interest for financial applications. The major problem under the multivariate context is related to the ‘dimensionality curse’, which makes most of standard techniques available for univariate time-series unfeasible. Previous empirical work on this topic has been limited to small portfolios composed of a reduced number of assets. Furthermore, little is known about the consequences of skipping long-memory effects in forecasting conditional covariance matrices of asset portfolios.

The main goal of this paper has been to show a tractable and theoretically admissible model in order to forecast the conditional variance of large portfolios of assets which exhibit this sort of pattern. Furthermore, as far as we know, this is the first paper concerned with long-memory effects when forecasting the conditional covariance of portfolios in the out-of-sample period. The Orthogonal GARCH procedure proposed by Alexander (2001) is generalized here to cope with the effects of long-run persistence in volatility, by considering the model developed by Davidson (2003). We have applied this procedure to a portfolio of five exchange rates in order to assess the possible gains in the forecasting accuracy after modelling long-run dependences. The procedure proves to be adequate for forecasting purposes since the out-of-sample forecasting analysis shows a clear outperformance over the simpler GARCH-based procedure. It is shown, therefore, that long-run temporal dependences in the volatility prove to be sufficiently important to enhance the short-term forecasts of the conditional covariances matrices significantly when they are included in multivariate GARCH models, so a greater effort in this direction should be considered in further research. The findings of this paper are consistent with those presented in the univariate literature.

This procedure is suitable for practical risk applications concerning large portfolios of long-memory assets, as it provides better forecasts of the conditional covariance matrix than other simpler alternatives. However, it is likely that more complex models accounting simultaneously for long-memory as well as other stylized features yield better results than those showed in this paper. Also, it would be interesting to consider economic loss-functions instead of purely-statistical measures (e.g., those based on hedging ratios, Value-at-Risk measures), which could be more relevant to financial risk managers. All of these aspects remain interesting topics for further research.

References

- [1] Alexander, C., 2001. *Market Models: a Guide to Financial Data Analysis*. John Willey & Sons.
- [2] Alexander, C., A. M. Chibumba, 1998. Orthogonal GARCH: An empirical validation on equities, foreign exchange and interest rates. Unpublished manuscript.
- [3] Baillie, R. T., 1996. Long-memory processes and fractional integration in econometrics. *Journal of Econometrics* 73, 5-59.
- [4] Baillie, R. T., T. Bollerslev and H. Mikkelsen, 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74, 3-30.
- [5] Beine, M., S. Laurent and C. Lecourt, 2002. Accounting for conditional leptokurtosis and closing day effects in FIGARCH models of daily exchange rates. *Applied Financial Economics* 12, 589-600.
- [6] Bollerslev, T., 1990. Modeling the coherence in the short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics* 72, 498-505.
- [7] Bollerslev, T., R. Y. Chou and K. F. Kroner, 1992. ARCH modelling in finance. *Journal of Econometrics* 52, 5-59.
- [8] Bollerslev, T., R. F. Engle and D. B. Nelson, 1994. *ARCH Models*. Handbook of Econometrics, chapter 49, volume 4. North-Holland.
- [9] Bollerslev, T., H. Mikkelsen, 1996. Modelling and pricing long memory in stock market volatility. *Journal of Econometrics* 73, 151-184.
- [10] Brailsford, J. T., R. W Faff, 1996. An evaluation for forecasting volatility techniques. *Journal of Banking and Finance* 20, 419-438.
- [11] Brooks, C., 2001. A double-threshold GARCH model for the French franc/deutschmark exchange rate. *Journal of Forecasting* 20, 135-143.
- [12] Breidt, J. F., N. Crato and P. de Lima, 1998 The detection and estimation of long-memory in stochastic volatility. *Journal of Economics* 83, 325-348.
- [13] Byström, H., 2000. Orthogonal GARCH and covariance matrix forecasting in an stress scenario. Working paper. Lund University.
- [14] Christoffersen, P. F., F. X. Diebold and T. Schuermann, 1998. Horizon problems and extreme events in financial risk management. *Economic Policy Review FRBNY* 4.

- [15] Davidson, J., 2003. Moment and memory properties of linear conditional heteroskedasticity models, and a new model. Forthcoming in *Journal of Business and Economic Statistics*.
- [16] Diebold, F. X., R. S. Mariano, 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13(3), 253-263.
- [17] Ding, Z., 1994. *Time Series Analysis of Speculative Returns*. Ph.D. Thesis. UCSD.
- [18] Ding, Z., Granger, C. W. J. and R. F. Engle, 1993. A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1, 83-106.
- [19] Engle, R.F., T. Bollerslev, 1986. Modelling the persistence of conditional variances. *Econometric Reviews* 5, 1-50.
- [20] Engle, R. F., V. Ng and M. Rothschild, 1990. Asset pricing with a factor ARCH covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics* 45, 213-238.
- [21] Engle, R. F., 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* 20(3), 339-351.
- [22] Geweke, J., S. Porter-Hudak, 1983. The estimation and application of long-memory time series models. *Journal of Time Series Analysis* 4, 221-238.
- [23] Hosking, J. R. M., 1980. The multivariate Portmanteau statistic. *Journal of the American Statistical Association* 75, 602-608.
- [24] Ledoit, O., M. Wolf, 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance* 10, 603-621.
- [25] López, J. A., 2001. Evaluating the predictive accuracy of volatility models. *Journal of Forecasting* 20, 87-109.
- [26] Newey, W., K. West, 1987. A simple positive semi-definite, heteroskedastic and autocorrelation consistent covariance matrix. *Econometrica* 55, 703-708.
- [27] Pafka, S., L. Mátyás, 2001. Multivariate diagonal FIGARCH: Specification, estimation and application to modelling exchange rate volatility. Unpublished manuscript.
- [28] Teyssière, G., 1997. Modelling exchange rate volatility with multivariate long-memory ARCH processes. Working paper 97-B03. GREQAM.

- [29] Teyssière, G., 1998. Multivariate long-memory ARCH modelling for high frequency foreign exchange rates. *Proceedings of the Second High-frequency Data in Finance* (HFDF-II Conference). Olsen & Associates.
- [30] Vilasuso, J., 2002. Forecasting exchange rate volatility. *Economics Letters* 76, 59-64.

Appendix A: The HYGARCH model

We discuss some aspects related to cumulative impulse-response function (CIRF) of the HYGARCH(1, d ,1) process. An in-depth discussion of several theoretical properties of this model can be found in Davidson (2003).

Consider the covariance-stationary HYGARCH(1, d ,1) equation through its infinite ARCH representation,

$$h_t = \frac{\omega}{1-\beta} + \left(1 - (1-\phi L)(1-\beta L)^{-1} \left[1 + \theta((1-L)^d - 1)\right]\right) \varepsilon_t^2 \quad (\text{A.1})$$

The CIRF measures the cumulative impact of past shocks of the volatility process over time, and hence it is defined as the coefficients $\{\varphi_k\}_{k=0}^{\infty}$ of the infinite ARCH representation, that is

$$\varphi(L) = 1 - (1-\phi L)(1-\beta L)^{-1} \left\{1 + \theta[(1-L)^d - 1]\right\} = \sum_{k=0}^{\infty} \varphi_k L^k \quad (\text{A.2})$$

where the fractional polynomial $(1-L)^d$ is usually expanded as

$$(1-L)^d = F(-d, 1; 1; L) = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} L^k \quad (\text{A.3})$$

with $\Gamma(\cdot)$ denoting the Gamma function.

The cumulative impulse-response weights (CIRW) sequence $\{\varphi_k\}_{k=0}^{\infty}$ of the HYGARCH(1, d ,1) model is then given by

$$\varphi_k = \begin{cases} 0, & k = 0 \\ \theta d + \phi - \beta, & k = 1 \\ \theta \left[\beta \varphi_{k-1} + \left(\frac{k-1-d}{k} - \phi\right) \delta_{k-1} \right] - (1-\theta)(\beta^k - \phi\beta^{k-1}), & k \geq 2 \end{cases}$$

where the coefficients δ_k can be obtained through an iterative scheme given by

$$\begin{aligned} \delta_1 &= d \\ \delta_k &= \delta_{k-1} \frac{k-1-d}{k}, \quad k > 1 \end{aligned} \quad (\text{A.4})$$

For the HYGARCH(1, d ,1) model to be well-defined, all the CIRW must be positive. From the recursions $\{\varphi_k\}_{k=1}^{\infty}$ it follows easily that for the covariance-stationary case the inequality constraints

$$\begin{aligned} \beta - \theta d &\leq \phi \leq \frac{2-d}{3} \\ (1-\theta)(\beta^2 - \phi\beta) &\leq \theta \left[\beta \varphi_1 + \left(\frac{1-d}{2} - \phi\right) d \right] \end{aligned} \quad (\text{A.5})$$

ensure that the coefficients in the lag polynomial $\varphi(L)$ are all non-negative. Note that when the amplitude parameter, θ , tends to one, the model converges to the FIGARCH(1, d ,1) and so the inequality constraints above are the same as those in Bollerslev and Mikkelsen (1996). On the other hand, when $\theta > 1$ it is easy to see that the constraints

$$\beta - \theta d \leq \phi \leq \frac{2-d}{3}, \quad d\left(\phi - \frac{1-d}{2}\right) \leq \beta(\phi - \beta + \theta d) \quad (\text{A.6})$$

are sufficient to guarantee that the coefficients in the lag polynomial $\varphi(L)$ are all non-negative.

Appendix B: Loss-functions

Assessing the forecast error in the multivariate context is worthwhile because the covariance matrix is the straightforward result of the methodology involved. We use matrix norms to measure the forecast error in the matrix space, i.e., the distance between forecasted and realized matrices. Under such metrics, all elements in the forecasted matrix are equally important, so that models are penalized if, say, they have a greater ability for forecasting the on-diagonal elements but a poor performance for accurately predicting covariances. Specifically, we use the Frobenius norm and the so-called matrix p -norms. The Frobenius norm of the forecast error matrix, denoted as $\|\mathbf{U}_{T+n}^s\|_F$, is the Euclidean distance between the forecasted and the realized matrices at any time, which is computed as

$$\|\mathbf{U}_{T+n}^s\|_F = \sqrt{\sum_{i=1}^K \sum_{j=1}^K |u_{ij,T+n}^s|^2} \quad (\text{B.1})$$

with $u_{ij,T+n}^s$ denoting the i, j -th element of \mathbf{U}_{T+n}^s .

The matrix p -norms are defined in terms of the corresponding vector p -norms. There are many possible norms in this category but, as in the case of vector p -norms, the most important are the 1-norm and ∞ -norm. Both norms actually yield the same value when applied to symmetrical matrices, so we denote by $\|\mathbf{U}_{T+n}^s\|_1$ the 1-norm (∞ -norm) of the forecast error matrix. Intuitively, this metric measures the maximum absolute cumulative error when forecasting the variance and the covariances of each asset in the portfolio. It is computed as

$$\|\mathbf{U}_{T+n}^s\|_1 = \max_{1 \leq j \leq K} \sum_{i=1}^K |u_{ij,T+n}^s| \quad (\text{B.2})$$

After defining valid measures in the matrix space, we consider a loss-function for each metric defined as the average value of the forecasted error norms over the out-of-sample period, so that

$$N_F^s = \frac{1}{200} \sum_{n=0}^{199} (\|\mathbf{U}_{T+n}^s\|_F)^2 \quad (\text{B.3})$$

$$N_1^s = \frac{1}{200} \sum_{n=0}^{199} \|\mathbf{U}_{T+n}^s\|_1 \quad (\text{B.4})$$

Along with the multivariate framework, it is also of interest to evaluate the forecast errors under the univariate framework. The overall (conditional) variance of an equally-weighted portfolio is given by the scalar $\mathbf{K}^{-2} \mathbf{1}'_k \mathbb{E}_{T+n}^s (\mathbf{H}_{T+n+s}) \mathbf{1}_k$, where $\mathbf{1}_k$ is a vector of ones in \mathcal{R}^K . This is nothing but the weighted sum of the variances and covariances among the assets of the portfolio. It follows easily that the overall variance forecast error is proportional to the quadratic form $q_{T+n}^S = \mathbf{1}'_k \mathbf{U}_{T+n}^s \mathbf{1}_k$ when holding a steady portfolio over time. As those values are univariate measures, forecast accuracy can be judged based on the standard statistical loss-functions, such as the mean squared error (MSE),

$$MSE_s = \frac{1}{200} \sum_{n=0}^{199} (q_{T+n}^S)^2 \quad (\text{B.5})$$

and the mean absolute error (MAE),

$$MAE_s = \frac{1}{200} \sum_{n=0}^{199} |q_{T+n}^S| \quad (\text{B.6})$$

In addition, two asymmetric measures proposed in Brailsford and Faff (1996) are also considered to assess forecast ability in the univariate framework. They provide a mean absolute error after more heavily penalizing either underpredictions (MAEU) or overpredictions (MAEO). They are defined, respectively, as follows

$$MAEU_s = \frac{1}{200} \left(\sum_{n=1}^{N_U} \sqrt{|q_{T+n}^S|} + \sum_{n=1}^{N_O} |q_{T+n}^S| \right) \quad (\text{B.7})$$

$$MAEO_s = \frac{1}{200} \left(\sum_{n=1}^{N_U} |q_{T+n}^S| + \sum_{n=1}^{N_O} \sqrt{|q_{T+n}^S|} \right) \quad (\text{B.8})$$

where N_U is the number of underpredictions and N_O is its complementary. The binomial probability of underprediction can be estimated by $N_U/200$, values of this estimator further than 0.5 provide evidence against the hypothesis of unbiased forecasts. As GARCH-type forecast are known to be fairly sensitive to the effect of additive outliers (recall that they are based on ARMA-type models), it could be of interest, for instance, to penalize harder overpredictions as outliers force forecasts to be abnormally high. Also, forecasts are needed to compute Value-at-Risk measures, which are known to be non-symmetric to forecasts errors, so it is worthwhile to consider asymmetric measures in the forecasting analysis.