Laboratory of Economics and Management

Sant'Anna School of Advanced Studies
Piazza Martiri della Libertà, 33 - 56127 PISA (Italy)
Tel. +39-050-883-343 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://www.lem.sssup.it/

LEM

Working Paper Series

Generalized Dynamic Factor Model + GARCH Exploiting multivariate information for univariate prediction

Lucia Alessi*
Matteo Barigozzi*
Marco Capasso*

*Scuola Superiore Sant'Anna

2006/13 July 2007

ISSN (online) 2284-0400

Generalized Dynamic Factor Model + GARCH Exploiting multivariate information for univariate prediction

Lucia Alessi[*] Matteo Barigozzi[†] Marco Capasso[‡]

Laboratory of Economics and Management (LEM)

Sant'Anna School of Advanced Studies, Pisa

Abstract

We propose a new model for volatility forecasting which combines the Generalized Dynamic Factor Model (GDFM) and the GARCH model. The GDFM, applied to a large number of series, captures the multivariate information and disentangles the common and the idiosyncratic part of each series of returns. In this financial analysis, both these components are modeled as a GARCH. We compare GDFM+GARCH and standard GARCH performance on two samples up to 171 series, providing one-step-ahead volatility predictions of returns. The GDFM+GARCH model outperforms the standard GARCH in most cases. These results are robust with respect to different volatility proxies.

Keywords: Dynamic Factors, GARCH, volatility forecasting
JEL-classification: C32, C52, C53

[*] *E-mail* : lucia.alessi@sssup.it
[†] *E-mail* : matteo.barigozzi@gmail.com
[‡] *E-mail* : marco.capasso@gmail.com

We thank Carlo Bianchi, Giorgio Calzolari, Mario Forni, Marc Hallin, Daniel Peña, Esther Ruiz, all the participants to the Italian Congress on Time Series SER 2006 (Villa Mondragone - Roma) and to the XXVI International Symposium on Forecasting 2006 (Santander) for helpful comments and suggestions. Special thanks to Marco Lippi and Gabriele Fiorentini. Of course, the responsibility for any error is entirely our own.

1 Introduction

Forecasting volatility of future returns by exploiting multivariate information is a major challenge for financial econometrics. The advantages of using multivariate models versus univariate ones, i.e. univariate GARCH or any kind of univariate generalization, as well as univariate Stochastic Volatility (SV) models, are enormous. Being able to exploit information on covariances of return series yields predictions which are necessarily at least as good as univariate ones; common sense suggests that they are strictly better because of the existence of relations across assets and markets which univariate techniques ignore. These both contemporaneous and lagged relations across stocks are important, which ultimately implies that multivariate models are of great advantage with respect to univariate ones.

The generalization of univariate models to multivariate ones, however, is far from trivial. The main pitfall of multivariate GARCH models in most specifications is the very large number of parameters, which rapidly makes the estimation unfeasible as the number of series grows; those specifications which bypass this problem, on the other hand, pay the price in terms of a severe loss of generality[1]. Neither multivariate SV models, although relatively more parsimonious, are able to handle more than a few number of series because of their complexity of estimation[2]. For both streams of literature, the key for dimensionality reduction stands in the idea of the existence of a few latent variables, the so called *factors*, as driving forces for the whole dataset. Models as CAPM explain theoretically why we may speak of factors in the market. Indeed, the use of factor models allows to disentangle within each stock the component which is directly linked to these common forces and the component which is peculiar to the stock itself. Doing this way, the factor analysis makes use of co-movements across stocks in order to improve forecasts.

Here we focus on the GARCH side of the story[3]. Diebold and Nerlove [1989] develop a static factor model on return series where the covariance matrix of factors is conditionally heteroskedastic, while the conditional covariance of the idiosyncratic part is homoskedastic. Given that the number of factors is small, the factor model reduces dramatically the number of parameters to be estimated with respect to the multivariate GARCH model. Engle et al. [1990] propose a model in which the decomposition in factors is at the level of conditional variance; Sentana [1998] proves that this model is nested in the previous by Diebold and Nerlove. More recently, the Orthogonal GARCH model by Alexander [2000], typically used for Value-at-Risk modeling, and the PC-GARCH by Burns [2005] retrieve the factors of the system by means of standard principal component analysis, while the GO-GARCH model by van der Weide [2002] generalizes the Orthogonal-GARCH approach within the boundaries of the static framework.

The novelty of our approach stands in the introduction of *dynamics*. By applying the Generalized Dynamic Factor Model (GDFM) by Forni et al. [2000] we are able to handle a very large number of series and capture all the multivariate information not only in the cross dimension but also in the time dimension. The GDFM model generalizes on the one hand the dynamic factor model proposed by Sargent and Sims [1977] and Geweke [1977] by allowing for mildly correlated idiosyncratic components; on the other hand the approximate factor model by Chamberlain [1983] and Chamberlain and Rothschild [1983] which is static. In the same stream of literature, Stock and Watson [2002] deal with forecasting issues, although in a macroeconomic context, by means of an approximate dynamic factor model which is estimated

[1]See Bauwens et al. [2006].

[2]See Harvey et al. [1994].

[3]For multivariate SV models within the factor approach, see Chib et al. [2006].

in a static way.
We combine the GDFM and the GARCH in a two step procedure: in the first step we apply the GDFM to the series of returns in order to split each of the series in its common part and its idiosyncratic part; in the second step we model both components as a GARCH, allowing for different GARCH orders and different values of parameters across series. The predicted one-step-ahead conditional variance is then obtained by summing up the one-step-ahead predictions for common part and idiosyncratic part. Finally, results are compared with predictions generated by a standard univariate GARCH applied to each series of returns as such. The GDFM+GARCH model outperforms the standard GARCH in most cases.

The paper is structured as follows. Section 2 outlines the GDFM+GARCH model and the estimation procedure along the lines of Forni et al. [2006]. Section 3 overviews the literature on volatility proxies in the context of a more general discussion of the issues related to the prediction of volatility. In section 4 we present the results of the empirical analysis, that we run on two different samples respectively of 140 and 171 series. The comparison between the GDFM+GARCH model and the benchmark is carried out by means of Mincer-Zarnowitz regressions, RMSE evaluation, and the prediction accuracy test by Clark and West [2007]. Section 5 concludes and provides an outlook on future developments.

2 The model

We denote as $x_t = (x_{1t} \dots x_{Nt})'$ the N-dimensional vector process of standardized stock returns. Each of the series is stationary and second order moments $\gamma_{ik} = E[x_{it}x'_{it-k}]$ exist finite. As in the Generalized Dynamic Factor Model (GDFM) proposed by Forni et al. [2000] we assume that each series x_{it} can be written as the sum of two mutually orthogonal unobservable components, the *common component* χ_{it} and the *idiosyncratic component* ξ_{it}. The common component is driven by a small number q of dynamic common factors u_{jt} with $j = 1, \dots, q$, which are loaded with possibly different coefficients and lags. Formally, we assume:

$$x_{it} = \chi_{it} + \xi_{it} = b_{i1}(L)u_{1t} + b_{i2}(L)u_{2t} + \dots + b_{iq}(L)u_{qt} + \xi_{it} \qquad i = 1, \dots, N \tag{1}$$

The q-dimensional vector process $u_t = (u_{1t} \dots u_{qt})'$ is an orthonormal white noise. The N-dimensional vector process $\xi_t = (\xi_{1t} \dots \xi_{Nt})'$ has zero mean and is stationary. Moreover ξ_{it} is orthogonal to u_{jt-k} for all k, i and j. The polynomials in the lag operator $b_{i1}(L) \dots b_{iq}(L)$ are square-summable, one-sided filters of order s, that is to say that $r = q(s + 1)$ static factors are loaded contemporaneously.
In order to move to the frequency domain we need to assume that the process x_t admits a Wold representation $x_t = \sum_{k=0}^{+\infty} C_k w_{t-k}$ where innovations have finite fourth order moment and the entries of the matrices C_k satisfy $\sum_{k=0}^{+\infty} |C_{ij,k}|k^{1/2}$. We denote the spectral density matrices of the common part and the idiosyncratic part respectively as $\Sigma^{\chi}(\theta)$ and $\Sigma^{\xi}(\theta)$, with $\theta \in [-\pi, \pi]$, and assume that the q largest eigenvalues of $\Sigma^{\chi}(\theta)$ diverge almost everywhere as the number of series goes to infinity, while all the eigenvalues of $\Sigma^{\xi}(\theta)$ are bounded. This last condition, in other words, relaxes the assumption of mutual orthogonality of idiosyncratic components by allowing for a limited amount of cross-sectional correlation.
We assume that both the common component and the idiosyncratic component of each of the series can be modeled as a GARCH (p,z) process with possibly different coefficients. Formally,

a generic univariate GARCH model is written as:

$$\begin{aligned} y_t &= m(y_{t-1}, \ldots, y_{t-k}) + a_t \\ a_t &= \epsilon_t \sigma_t \\ \sigma_t^2 &= \omega + \sum_{j=1}^{z} \alpha_j a_{t-j}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \end{aligned} \tag{2}$$

In our context this model is applied for every series x_{it} to both the common component χ_{it} and to the idiosyncratic component ξ_{it}, while the conditional mean $m(y_{t-1}, \ldots, y_{t-k})$ can be either modeled as an ARMA process or set equal to zero. The conditional variance is obtained as the sum of the conditional variances for the common part and for the idiosyncratic part, which is a legitimate procedure given that the two components are orthogonal by definition.
The problem of contemporaneous aggregation of GARCH processes has already been faced by Nijman and Sentana [1996], who found out that the sum of two (strong) GARCH processes gives rise to a weak GARCH process, a process originally introduced by Drost and Nijman [1993] for the case of temporal aggregation of GARCH processes. Following the notation in (2) (without taking into account the mean evolution part), Nijman and Sentana [1996] show that the sum of two strong GARCH (1,1) processes y_1 and y_2 evolves as:

$$(y_{1t} + y_{2t})^2 = d_1 + d_2 + [1 - (\alpha_1 + \beta_1) L]^{-1} [1 - \beta_1 L] g_{1t} + [1 - (\alpha_2 + \beta_2) L]^{-1} [1 - \beta_2 L] g_{2t} + 2 y_{1t} y_{2t}, \tag{3}$$

where

$$d_i = \omega_i (1 - \alpha_i - \beta_i)^{-1}; \tag{4}$$

$$g_{it} = \left(\epsilon_{it}^2 - 1\right) \sigma_{it}^2. \tag{5}$$

In other words, the sum of two independent strong GARCH (1,1) processes is weak GARCH (2,2). The presence of the cross-product term in the right-hand side of (3) represents the practical difference between a weak GARCH (2,2) and a strong GARCH (2,2), as it complicates the derivation of the weak GARCH parameters for the aggregate series. However, the estimation of these parameters is still consistent both by exploiting the autocorrelation of $(a_{1t} + a_{2t})^2$ and by Quasi Maximum Likelihood estimation. Simulation results obtained by Nijman and Sentana [1996] confirm that QML estimations of a weak GARCH process (that is ML estimations of a strong GARCH process) onto the aggregate series may often yield values which are very similar to the true weak GARCH parameters of the aggregate series, especially for the case of a large dimension of the observed time series. For this reason, in the existing literature, GARCH models have been estimated for the (log) returns in the Deutsche mark/US dollar exchange rate, the US dollar/Japanese yen exchange rate, and the Deutsche mark/Japanese yen rate, where the returns on the third exchange rate are simply the sum of the returns on the first two exchange rates.
Our hypothesis of a factor structure governing our dataset drives us to the idea that, for forecasting purposes, we might model separately the conditional variances of the common part and of the idiosyncratic part of each series, in order to get better conditional variance predictions of the aggregate series than a ML estimated strong GARCH applied directly to the aggregate series. In the empirical part of our work, we get rid of all the problems related to the orders of the GARCH processes, by always choosing the smallest possible order that eliminates the serial correlation of the standardized (squared and not squared) residuals.

The state of the art as far as volatility forecast is concerned basically exploits high-frequency data to build various volatility proxies and finally get a forecast of future values of these proxies themselves. Our aim however is different: although here we are interested in volatility prediction, we want a model that has the possibility of predicting both levels and volatility of returns at once - which is what the market needs as a first best. Therefore we choose to stick to the world-wide used GARCH model. An alternative approach would be to run the GDFM factor decomposition directly on volatility proxies[4], however we preferred to act at the return level because predicting both first and second moment allows for the construction of interval predictions, which is of great interest although beyond the purpose of this exercise[5].

The estimation of the model follows the two-step procedure proposed in Forni et al. [2006] for the GDFM part. In the first step the spectral density matrix of x_t, $\Sigma^x(\theta)$, is estimated by applying the Fourier transform to the sample auto-covariance matrices $\hat{\Gamma}_k$. Then the dynamic principal component decomposition is applied, thereby selecting the first q largest eigenvalues of $\hat{\Sigma}^x(\theta)$ and the corresponding eigenvectors. Calling $P(\theta)$ the matrix with eigenvectors as columns and $\Lambda(\theta)$ the matrix with eigenvalues on the diagonal, the estimated spectral density matrix of χ_t is computed as: $\hat{\Sigma}^\chi(\theta) = P(\theta)\Lambda(\theta)P(\theta)'$. It's worth noticing at this point the key difference between this dynamic approach and the static principal component method used by Stock and Watson [2002]: while the first exploits the information contained in lagged covariance matrices, the latter makes use of contemporaneous covariances only. By applying to $\hat{\Sigma}^\chi(\theta)$ the inverse Fourier transform we retrieve estimates of the covariance matrices of the common component, $\hat{\Gamma}^\chi_k$; the estimate of the covariance matrices of ξ_t, $\hat{\Gamma}^\xi_k$, is obtained by difference. To overcome the problem of bilateral filters, in the second step of the procedure we move to a static representation of the model in which we estimate the first r generalized eigenvectors of $\hat{\Gamma}^\chi_k$ with respect to $\hat{\Gamma}^\xi_k$. The first generalized eigenvector solves the following maximization problem:

$$\begin{cases} z^{(1)} & = \underset{a\in\mathbb{R}^n}{\operatorname{argmax}}\ a'\hat{\Gamma}^\chi_0 a \\ \text{s.t.} & a'\hat{\Gamma}^\xi_0 a = 1 \end{cases} \tag{6}$$

We collect the first r generalized eigenvectors in the matrix $Z = (z^{(1)} \dots z^{(r)})$ and by means of such matrix and of the contemporaneous covariance matrix, estimated in the first step, we are able to estimate the common component as:

$$\hat{\chi}_t = \hat{\Gamma}^\chi_0 Z(Z'\hat{\Gamma}_0 Z)^{-1}Z'x_t \qquad \forall\ t = 1,\dots,T \tag{7}$$

We obtain the idiosyncratic component simply as difference between the original series x_t and $\hat{\chi}_t$. Indeed, the one-sided estimator allows to forecast the common component at $T+h$ by substituting the estimated lagged covariances $\hat{\Gamma}^\chi_h$ to the contemporaneous covariance $\hat{\Gamma}^\chi_0$ in (7).

3 Mean modeling prediction

The decision of predicting both mean and variance of returns exerts an important influence not only on the theoretical model used, but also on the volatility measures employed for the

[4]Such an approach requires the existence of the fourth moment of the returns, which is still an issue under discussion in the literature.

[5]See Corradi and Swanson [2004].

prediction accuracy measurement. The relation between the aims of the model, the structure of the model, and the out-of-sample performance measurement is deep and complex, and has characterized the evolution of research during the last twenty years.

Roughly speaking, the problem faced by the finance researcher may be described as follows. If we think that the mean of a stock return cannot be predicted, then the most important moment we are interested in is the variance; therefore we might use volatility proxies both in-sample for estimating the parameters of the model, and out-of-sample to evaluate the accuracy of the prediction. Starting from this intuition, Andersen et al. [2001] studied the distribution and evolution of volatility, and Andersen et al. [2003] investigated the prediction of volatility. From a technical point of view, this choice simplifies the researcher's task, because she can operate at just one "level", and so she can apply traditional ARMA processes, or long memory and multivariate modifications of them, on the chosen volatility proxy.
This line of reasoning sounds perfect as long as we give up mean predicting. Such a withdrawal, although justifiable on a scientific basis, is difficult to digest for a financial world in which risk management needs the coupling of at least the first two moments of a return distribution. This is one of the reasons why GARCH models, and in general all the models that may take into account two "levels", cannot be ignored, even if researchers tend to sacrifice this feature and use GARCH models with a zero-mean assumption, when good mean predictors are missing. We might hope that the conditional mean is always constant and equal to zero. However, this could only be an approximation, coming out from our difficulties in modeling the conditional mean evolution. Therefore we prefer to apply an ARMA + GARCH model to the return series, both directly (as a benchmark) and indirectly (when using our model's splitting of the original series into a common and an idiosyncratic part), in order to have the possibility of a better prediction of the conditional mean whenever the data set allows it.

We now outline the algorithm used for estimating GARCH models, with or without ARMA component. For each of the return series, we run the algorithm on the return series x_t as such, on the series χ_t representing the common part, and on the series ξ_t representing the idiosyncratic part, i.e. we estimate $3N$ models for each sample.

- We begin by estimating the ARMA part of the process, in the cases in which we assume non-zero conditional mean. We start by fitting an ARMA(0,0), then perform a Ljung-Box test on the residuals at the 0.05 significance level, including 4 lags, thus setting 4 degrees of freedom for the chi-square distribution. If the ARMA(0,0) fails the test, i.e. residuals are serially autocorrelated, we increase the AR order by one unit and run the Ljung-Box test again. If the ARMA(1,0) fails the test, we estimate an ARMA(2,0). If necessary, we increase then the order of the MA part by one unit at a time up to 2.

- Next, we verify the presence of ARCH effects in the series by performing an Engle's test on the ARMA residuals with 0.05 significance level and 4 lags included. If this is the case, we estimate the GARCH model starting from an ARCH(1) and perform again the test. If the ARCH(1) fails the test, we move to an ARCH(2) and if necessary to a GARCH(2,1). The highest order we allow for is GARCH(2,2).

4 Empirics

4.1 Preliminary analysis

The dataset we use for the empirical investigation includes 475 return series of stocks traded on the NYSE (we arbitrarily choose all the stocks for which options are also traded). Each series goes from 8th March 1995 to 30th April 1999 (1045 daily observations). Series have been cleaned from outliers[6]. For each trading day we also have the highest and the lowest price at which each stock has been traded. From these we obtain the range, as defined in (10). We run the analysis on two different subsamples: the first contains stocks belonging to the financial sector and the second represents the electronics sector[7]. We reduce the dimension of the dataset by considering sectoral samples in order to study how the GDFM+GARCH performance changes once only the most correlated series are left in, which might improve the prediction results, as already highlighted in Boivin and Ng [2003]. However, the estimator is consistent for the cross and time dimensions going to infinity. Indeed, as shown in the tables in this section, running the analysis sector by sector does improve the factor decomposition and thus the results. The prelimiary analysis is conducted with 515 and with 1030 observations, aiming to study the properties of both cases.

Firstly, we verify that our dataset does fulfill GDFM assumptions on the eigenvalues $\lambda_i(\theta)$ of the spectral density matrix of x_t. According to Brillinger [1981], we define the variance explained by the i^{th} factor as:

$$EV_i = \frac{\int_{-\pi}^{\pi} \lambda_i(\theta) d\theta}{\sum_{j=1}^{N} \int_{-\pi}^{\pi} \lambda_j(\theta) d\theta} \tag{8}$$

We require that, as $N \longrightarrow \infty$:

$$\begin{cases} EV_i \longrightarrow \infty & \text{for} \quad i = 1, \dots, q \\ \exists\, M \in \mathbb{R}^+ \quad \text{s.t.} \quad EV_i \leq M & \text{for} \quad i = q+1, \dots, N \end{cases} \tag{9}$$

Indeed, as shown for example in figure 1 for the short finance sample, this is the case. The subsequent figure shows the cumulated explained variance relative to the first q eigenvalues for the same sample.

For all samples we keep a number q of dynamic factors corresponding to the number of dynamic eigenvalues of the spectral density matrix which explain more than 5% of total variance each. In all the three samples, the chosen number of dynamic factors is much higher when considering a shorter time horizon. A value of q less or equal to 4, i.e. the maximum number of dynamic factors usually found in this kind of analysis on macroeconomic data, is reached here only when dealing with about four years of daily data. A tentative economic interpretation of this fact relies in the nature of the forces leading the market in the short term, which may be reasonably thought to be a larger number than those few driving the economy in the long run. We set the number of lags to $s = 4$, i.e. we consider one trading week. Table 1 summarizes the results of the dynamic factor decomposition, while table 2 presents descriptive statistics on the distribution of the variance of the common part over the total variance of each series.

[6]Outliers have been dropped and replaced with an average of previous and following returns.

[7]Following the SIC classification we identify the finance sector with the 1-digit SIC code 6 and include in the electronics sector all 2-digit SIC codes between 35 and 38 (included).

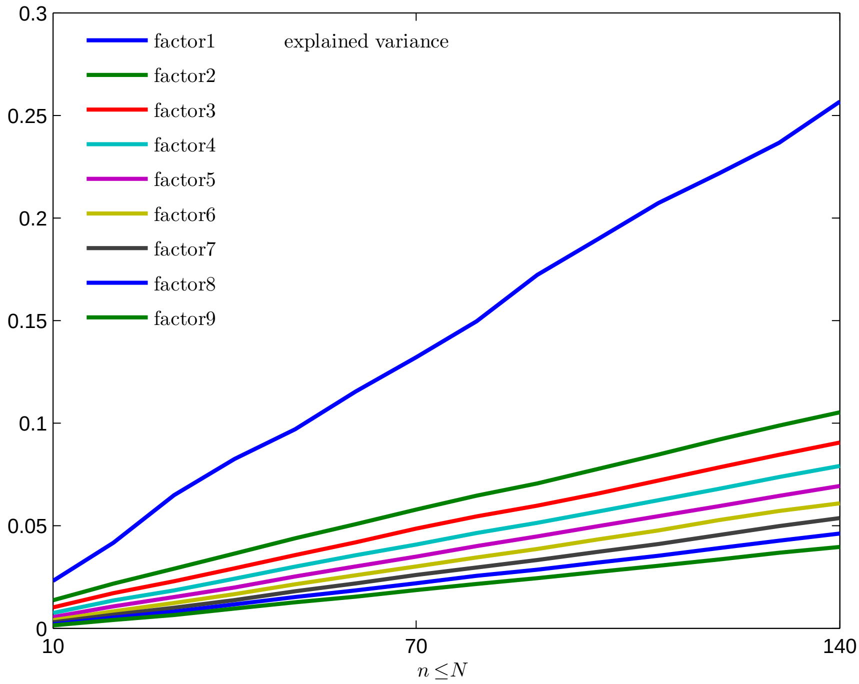

Figure 1: Finance, 515 observations. Explained variance.

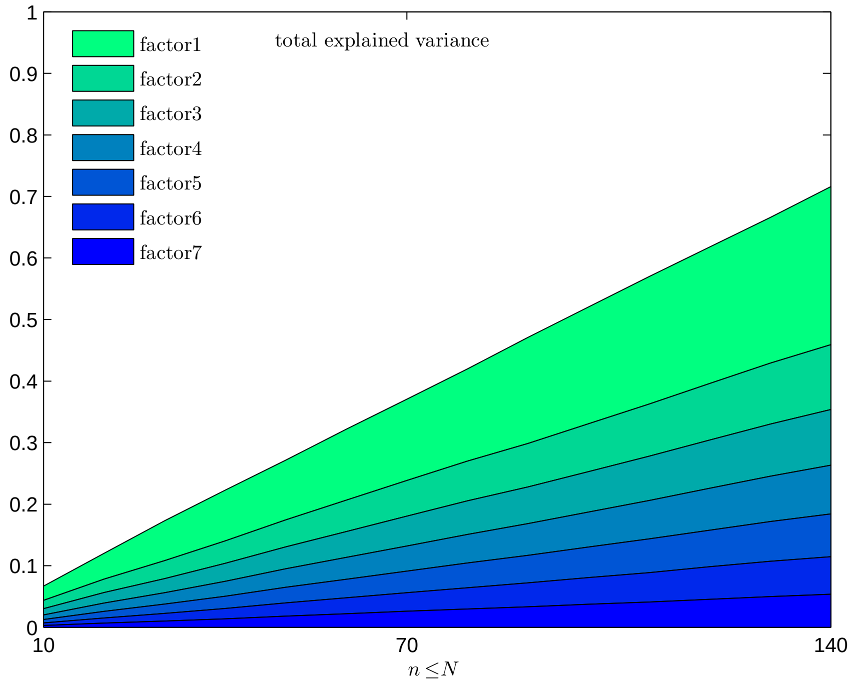

Figure 2: Finance, 515 observations. Cumulated explained variance.

Sample	Number of series	Length of insample	Number of dynamic factors q	Number of static factors r	Variance explained by the first q eigenvalues
Random	475	515	8	40	72%
		1030	3	15	34%
Finance	140	515	7	35	72%
		1030	3	15	46%
Electronics	171	515	8	40	74%
		1030	4	20	40%

Table 1: Dynamic factor decomposition.

Sample	No. of series	Length of insample	Variance of the common part over total			
			average	max	min	std
Random	475	515	18%	45%	6%	7%
		1030	18%	50%	2%	9%
Finance	140	515	36%	60%	17%	11%
		1030	32%	62%	7%	15%
Electronics	171	515	31%	59%	15%	9%
		1030	21%	48%	4%	9%

Table 2: Variance of the common part.

In the factor decomposition for the large dataset with 475 series the average variance explained by the common part is just 18% (versus 36% - 32% for finance and 31% - 21% for electronics). Indeed it seems not to be a good factor decomposition, probably due to the inclusion in the sample of too many heterogeneous series. Therefore we test our model only for the finance and electronics subsamples. Analogously, at each step of our forecasting scheme only 515 working days will be used as in-sample observations.
For both samples, we adopt a rolling scheme with 100 iterations. At each iteration we make one-step-ahead volatility predictions by using the information contained in the previous 515 observations of returns. The benchmark model is the univariate GARCH, which uses a single return series to forecast volatility. Our model exploits all the in-sample return series to predict volatility. In both cases, when we model the conditional mean part of processes, we follow the procedure explained in section 3, and we use the suffix “w mean” in the tables. On the other hand when we do not model the conditional mean we use the suffix “w/o mean”.

Table 3 reports, for the first iteration, the percentages of series of returns, common and idiosyncratic components presenting an ARMA structure, i.e. a significative autocorrelation in the levels.

Finance sample		
515 obs.		
x_t	χ_t	ξ_t
33%	64%	38%
Electronics sample		
515 obs.		
x_t	χ_t	ξ_t
30%	59%	29%

Table 3: Percentage of series containing an ARMA component.

4.2 Volatility proxies

The comparison of volatility prediction accuracy between our model and the benchmark is done with respect to the adjusted range, given the unavailability of high-frequency data in our dataset. In order to improve the robustness of our results we also compare volatility forecasts using squared returns as proxies only for the "w/o mean" predictions.
The idea of a range-based estimation of volatility dates back in time (see e.g. Feller [1951]); we compute the range (actually meaning the intradaily log-range) as:

$$RANGE_t = \log\left(P_{t\max}\right) - \log\left(P_{t\min}\right), \tag{10}$$

where $(P_{t\max})$ and $(P_{t\min})$ are respectively the highest price and the lowest price on day t. However, whereas the existence of a relation between range and volatility seems not to be deniable, there is not a wide consensus about the way of adjusting the range to best approximate volatility. According to Parkinson [1980], we adjust the range as in the following:

$$adj.RANGE_t = \frac{RANGE_t}{\sqrt{\log 16}} \approx 0.6006 \times RANGE_t, \tag{11}$$

where $(adj.RANGE_t)^2$ is an unbiased proxy for the stock volatility at time t, when the stock price follows a random walk without drift. Different adjustments have been suggested, among others, by Rogers and Satchell [1991], Kunitomo [1992] and Yang and Zhang [2000] as consequences of different theoretical assumptions on the data generating process (e.g. random walk with drift). For our empirical purposes, we prefer to use the adjusted range as described in (11), because this proxy has been shown by Brandt and Kinlay [2005] to better approximate realized volatility, and therefore it seems to better mimic conditional variance when dealing with real data.

4.3 Performance evaluation: Mincer-Zarnowitz regressions

Following Andersen et al. [2003], we evaluate the volatility forecasts of our model by running a Mincer-Zarnowitz regression (Mincer and Zarnowitz [1969]). We project ex-post volatility proxies on a constant and the one-step-ahead model forecasts. For each series we run a regression based upon real and predicted conditional standard deviations:

$$(V_{t+1})^{1/2} = b_0 + b_1 \left(\hat{\sigma}^2_{t+1}\right)^{1/2} + e_{t+1} \qquad t = 515, \ldots, 615 \tag{12}$$

Sample	Series number	Model	$\hat{b}_0$	$\hat{b}_1$	R^2
Finance	1	GARCH	**-0.0578**	**2.0131**	**0.1558**
		GDFM+GARCH	-0.0837	2.7571	0.0428
Electronics	1	GARCH	0.0168	0.1648	0.0027
		GDFM+GARCH	**0.0162**	**0.1922**	**0.0034**
Finance	2	GARCH	0.0074	0.2995	0.0322
		GDFM+GARCH	**0.0058**	**0.4319**	**0.0389**
Electronics	2	GARCH	0.0390	**-0.4170**	0.0052
		GDFM+GARCH	**-0.0350**	1.8722	**0.0406**
Finance	3	GARCH	**0.0000**	**0.6674**	**0.0361**
		GDFM+GARCH	0.0049	0.3577	0.0056
Electronics	3	GARCH	**0.0039**	**0.5149**	0.0326
		GDFM+GARCH	-0.0207	1.7097	**0.0549**
...	...	...	...	...	...
Finance	arit. mean	GARCH	0.0050	0.4299	**0.0296**
		GDFM+GARCH	**0.0038**	**0.5162**	0.0236
Electronics	arit. mean	GARCH	0.0113	0.2491	0.0266
		GDFM+GARCH	**0.0039**	**0.5918**	**0.0297**

Table 4: MZ regression - against adjusted range with mean - using conditional standard deviations.

where the volatility proxy V_{t+1} is the squared adjusted observed range, and $\hat{\sigma}^2_{t+1}$ represents the volatility forecast, as predicted at time t. For each subsample, we perform the projection both on GDFM+GARCH forecasts and traditional GARCH forecasts, so that we can compare the results. Should a model be correctly specified, we would obtain values of $\hat{b}_0$ and $\hat{b}_1$ that are close to 0 and 1, respectively. Table 4 summarizes the results of the regression, run by using conditional standard deviations both in the real data, here approximated by the adjusted range, and in the model forecasts, obtained with mean predicting. The GDFM+GARCH outperforms the traditional GARCH both in the parameters and in the R^2 for the electronics sample, and only in the parameters for the finance sample.

4.4 Performance evaluation: root mean square errors

Series by series, we take the prediction of the two models and compute one-step-ahead root mean square errors (RMSE) against the real value of the volatility proxy. We compute the RMSE as follows:

$$RMSE_i = \sqrt{\frac{1}{100}\sum_{t=515}^{615}\left(\hat{\sigma}^2_{it+1} - V_{it+1}\right)^2} \qquad i = 1, \ldots, N \tag{13}$$

where $\hat{\sigma}^2_{it+1}$ is the one-step-ahead volatility forecast of the considered model for series i. The proxy used are squared returns and squared adjusted range. In table 5 we report the RMSE for the first series and an average, according to which the GDFM+GARCH performs better

Sample	Series number	Model	RMSE adj. range w mean	RMSE adj. range w/o mean	RMSE squared returns
Finance	1	GARCH	0.0235	0.0240	0.0255
		GDFM+GARCH	**0.0231**	**0.0083**	**0.0243**
Electronics	1	GARCH	0.0137	0.0137	0.0197
		GDFM+GARCH	**0.0135**	**0.0135**	**0.0196**
Finance	2	GARCH	0.0104	0.0099	0.0152
		GDFM+GARCH	**0.0094**	**0.0070**	**0.0148**
Electronics	2	GARCH	0.0154	0.0155	0.0232
		GDFM+GARCH	**0.0139**	**0.0139**	**0.0220**
Finance	3	GARCH	0.0065	**0.0065**	**0.0098**
		GDFM+GARCH	**0.0063**	0.0091	0.0099
Electronics	3	GARCH	0.0104	0.0105	0.0148
		GDFM+GARCH	**0.0099**	**0.0099**	**0.0144**
...	...	...	...	...	...
Finance	arit. mean	GARCH	0.0081	0.0077	0.0121
		GDFM+GARCH	**0.0073**	**0.0073**	**0.0118**
Electronics	arit. mean	GARCH	0.0121	0.0113	0.0173
		GDFM+GARCH	0.0107	0.0108	0.0169

Table 5: RMSE - against adjusted range with mean, adjusted range without mean and squared returns.

than the univariate GARCH for both samples and all proxies. Results are also summarized in table 6 by means of two statistics:

- P corresponds to the percentage of series for which the GDFM+GARCH outperforms the univariate GARCH, i.e. the percentage of the cases for which

$$\frac{RMSE_i(GDFM + GARCH)}{RMSE_i(GARCH)} < 1 \qquad i = 1, \ldots, N \tag{14}$$

  In both samples and for all proxies the GDFM+GARCH outperforms the GARCH model for more than 80% of the series.

- Q is the geometric mean of the RMSE ratios:

$$Q = \left(\prod_{i=1}^{N} \frac{RMSE_i(GDFM + GARCH)}{RMSE_i(GARCH)} \right)^{\frac{1}{N}} \tag{15}$$

  In other words, the quantity $(1 - Q)$ is a measure of the average gain obtained by using our model. For both samples and all proxies Q is slightly smaller than 1.

In order to test the significance of the difference between the RMSEs of two models when one of the models nests the other, Clark and West [2007] show that a correction is needed on the

Sample	Adj. range w mean		Adj. range w/o mean		Squared	
	P	Q	P	Q	P	Q
Finance	85.00%	0.91	86.43%	0.95	82.86%	0.98
Electronics	88.89%	0.90	89.47%	0.95	92.40%	0.97

Table 6: One-step-ahead results.

	Finance	Electronics
t-ratio for series 1	0.9608	2.5476**
t-ratio for series 2	3.6514**	6.1044**
t-ratio for series 3	1.9237**	1.5143*
t-ratio for series 4	2.7354**	11.280**
...	...	...
number of series for which GDFM+GARCH outperforms GARCH at 10%	121 (86.43% of total)	151 (88.30% of total)
number of series for which GDFM+GARCH outperforms GARCH at 5%	106 (75.71% of total)	140 (81.87% of total)

Table 7: Clark-West test results against adjusted range with mean.

RMSE of the nested model (in our case the GDFM+GARCH). In particular, the following difference must be computed for each time t and each series i:

$$\hat{f}_{it+1} = \left(V_{it+1}^{1/2} - \hat{\sigma}_{it+1G}\right)^2 - \left[\left(V_{it+1}^{1/2} - \hat{\sigma}_{iFt+1}\right)^2 - \left(\hat{\sigma}_{it+1G} - \hat{\sigma}_{it+1F}\right)^2\right] , \tag{16}$$

where V_{it+1} represents the volatility proxy at time $t+1$ and $\hat{\sigma}^2_{it+1}$ represents the standard deviation forecast at time $t+1$, as predicted at time t by the simple GARCH (subscript G) or the GDFM+GARCH (subscript F). We then test for equal mean square prediction error by regressing each series $\hat{f}_{it}$ on a constant and using the resulting t-statistic for a zero coefficient. For each series i, GDFM+GARCH proves to work better than the traditional GARCH whenever the t-ratio is greater than $+1.282$ (for a one sided 0.10 test) or $+1.645$ (for a one sided 0.05 test). In table 7 we show our results of the test by Clark and West [2007] for both samples and all proxies. At both levels of significance, GDFM+GARCH performs better than univariate GARCH in the great majority of cases. Since the correction consists in subtracting a positive quantity from the nested model's RMSE, in case the test does not reject the hypothesis of the two RMSEs being different from each other it is possible that the percentage of series for which the GDFM+GARCH outperforms the GARCH actually increases with respect to the P statistic as in the case of the finance sample.

5 Conclusions

In this paper we have proposed a new model for multivariate analysis of large financial datasets which combines one of the latest developments in factor analysis, the Generalized Dynamic Factor Model, with the the world-wide used GARCH model. The GDFM+GARCH exploits a dynamic factor decomposition in order to retrieve the common part and the idiosyncratic part of each return series. These components are assumed to present ARCH effects: being ruled out the use of a multivariate GARCH model because of the large number of parameters, we solve this problem by estimating $2N$ univariate GARCH models. Despite the impossibility of estimating conditional covariances, we have the big advantage that, by exploiting the multivariate information embodied in sample covariances, we take into account all the dynamic relations between and within series. In the empirical part of the work we have compared the GDFM+GARCH predictive performance against the performance of the standard univariate GARCH, proxying out-of-sample conditional variance with squared returns and squared adjusted range. Results on two sectoral samples are encouraging and robust: the GDFM+GARCH outperforms the standard GARCH most of the time.

References

C. Alexander. A primer on the orthogonal GARCH model. Technical report, ISMA Center, University of Reading, 2000. Mimeo.

T. G. Andersen, T. Bollerslev, F. X. Diebold, and H. Ebens. The distribution of realized stock return volatility. *Journal of Financial Economics*, 61:43–76, 2001.

T. G. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys. Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–626, 2003.

L. Bauwens, S. Laurent, and J. V. Rombouts. Multivariate GARCH models: a survey. *Journal of Applied Econometrics*, 21(1), 2006. Forthcoming.

J. Boivin and S. Ng. Are more data always better for factor analysis? NBER Working Papers 9829, National Bureau of Economic Research, Inc, July 2003.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31: 307–327, 1986.

M. Brandt and J. Kinlay. Estimating historical volatility. Research article, Investment Analytics, 2005.

D. Brillinger. *Time Series: Data Analysis and Theory.* Holden-Day, Inc., San Francisco, CA, 1981.

P. Burns. Multivariate GARCH with only univariate estimation, 2005. http://www.burns-stat.com/.

G. Chamberlain. Funds, factors, and diversification in arbitrage pricing models. *Econometrica*, 51: 1281–1304, 1983.

G. Chamberlain and M. Rothschild. Arbitrage, factor structure and mean-variance analysis in large asset markets. *Econometrica*, 51:1305–1324, 1983.

S. Chib, F. Nardari, and N. Shephard. Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics*, 2006. Forthcoming.

T. E. Clark and K. D. West. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 2007. Forthcoming.

V. Corradi and N. R. Swanson. Predictive density and conditional confidence interval accuracy tests. *Journal of Econometrics*, 2004.

F. X. Diebold and M. Nerlove. The dynamics of exchange rate volatility: A multivariate latent factor arch model. *Journal of Applied Econometrics*, 4(1):1–21, 1989.

F. C. Drost and T. E. Nijman. Temporal aggregation of Garch processes. *Econometrica*, 61(4): 909–927, 1993.

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

R. F. Engle, V. K. Ng, and M. Rothschild. Asset pricing with a factor-ARCH covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics*, 45(1-2):213–237, 1990.

W. Feller. The asymptotic distribution of the range of sums of random variables. *Annals of Mathematical Statistics*, 22:427–32, 1951.

M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic factor model: identification and estimation. *The Review of Economics and Statistics*, 82:540–554, 2000.

M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic factor model: one-sided estimation and forecasting. *Journal of the American Statistical Association*, 2006. Forthcoming.

J. Geweke. The dynamic factor analysis of economic time series. In D.J. Aigner and A.S. Golberger, eds., Latent Variables in Socio-Economic Models, 1977.

A. Harvey, E. Ruiz, and N. Shephard. Multivariate stochastic variance models. *Review of Economic Studies*, 61(2):247–64, 1994.

N. Kunitomo. Improving the Parkinson method of estimating security price volatilities. *Journal of Business*, 64(2):295–302, 1992.

J. Mincer and V. Zarnowitz. *The Evaluation of Economic Forecasts*. Columbia University Press, 1969.

T. Nijman and E. Sentana. Marginalization and contemporaneous aggregation in multivariate GARCH processes. *Journal of Econometrics*, 71:71–87, 1996.

M. Parkinson. The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1):61–65, 1980.

L. C. G. Rogers and S. E. Satchell. Estimating variance from high, low and closing prices. *Annals of Applied Probability*, 1:504–512, 1991.

T. Sargent and C. Sims. Business cycle modeling without pretending to have too much a-priori economic theory. In C. Sims et al., eds., New Methods in Business Cycle Research, 1977.

E. Sentana. The relation between conditionally heteroskedastic factor models and factor GARCH models. *The Econometrics Journal*, 1:1–9, December 1998.

W. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.

J. H. Stock and M. W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20:147–162, 2002.

R. van der Weide. GO-GARCH: a multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics*, 17(5):549–564, 2002.

D. Yang and Q. Zhang. Drift independent volatility estimation based on high, low, open and close prices. *Journal of Business*, 73:477–491, 2000.