6. PERSPECTIVES ON RISK MEASUREMENT: A CRITICAL ASSESSMENT OF PC-GARCH AGAINST THE MAIN VOLATILITY FORECASTING MODELS

Marius MATEI[1]

Abstract

The paper makes a critical assessment of the Principal Components-GARCH (PC-GARCH) model and argues why, when dealing with hundreds or thousands of variables, this model comes up as the most appropriate to be used. The suitability originates from the perspective of quality/cost ratio of volatility forecasts, allowing for a trade-off between quality and costs when computational efforts are significant. PC-GARCH not only provides a method that allows for simpler volatility modeling, reducing significantly the computational time and getting rid of any problem that may arise from complex data manipulations, but also improves the modeling process quality by ensuring a stricter control of noise due to more stable correlation estimates.

Keywords: GARCH models, volatility forecasting, econometric models, evaluating forecasts, nonlinear time series

JEL Classification: C32, C53, C58

1. Introduction

The paper offers a critical assessment of the Principal Components-GARCH (PC-GARCH) model and support for the rationale behind one idea: the PC-GARCH model is the most appropriate model to use when evaluating the volatility of the returns of very large groups of stocks, containing hundreds or even thousands of variables. The appropriateness of the model is seen from the perspective of the quality/cost ratio of volatility forecast provided by PC-GARCH when compared to any other alternative model. Although an empirical study will follow to present how PC-GARCH works and to reveal the strengths of such a method, the test will not be used in order to compare

[1]*ESADE Business School, Ramon Llull University, Spain and National Institute for Economic Research, The Romanian Academy.Email: marius.matei@esade.edu, mateimar@gmail.com*

PC-GARCH directly with other methods, for which there would have to be tested hundreds of variables for each model considered, reason why direct empirical comparisons with similar models will not be undertaken. Conclusions regarding PC-GARCH will stem from the theoretical content described, and from the procedure itself as revealed, even as presented with only seven sets of variables; while extensive empirical exercises with more models tested, which would empirically quantify the gains in both forecast accuracy and cost (time, resources etc.) of PC-GARCH, are left for subsequent studies. However, the conclusions of the paper enforce the idea that PC-GARCH reveals its superiority when working with hundreds of variables, or even with thousands of variables. Such a conclusion can be reached by balancing two factors: i) the quality of the results, understood as the model's capacity to grasp the relationship between the exogenous variables and the endogenous ones, by taking into account the autocorrelations and interaction effects that may exist within the data[2]; and ii) the time or the amount of computational efforts needed to obtain such results.

Some previous papers on this topic evaluated the benefits of using the principal component analysis in orthogonal models. Alexander (2000) described an analysis developed within an Orthogonal GARCH context, but without any methodology offered for principal components in a GARCH model. Burns' (2005) paper offered such a methodology for PC-GARCH, but without any empirical implementation. However, none of the papers has emphasized the cost factor component of such methods.

The present study addresses two issues not fully explored previously. It attempts to provide a ranking/benchmarking analysis of volatility forecasting models looking at the cost factor and also by putting in balance the amount of computational efforts needed and the quality of the results, and applies, in support of the proposed solution, a method (principal component) to a multivariate GARCH, not previously empirically implemented (although described in his methodology by Burns, 2005). As a further contribution, the implementation that follows includes elements of GARCH testing that, to my knowledge, have not been discussed in any previous papers. Due to their complexity and to the size of the panel of data taken into account some models need to estimate too large a number of parameters. In this case, the model estimation may take too long, and the quality of the results does not necessarily make up for the length of the time when that is considerable. Sometimes, a trade-off between output (represented by the quality of the results) and costs (measured by the amount of time spent to obtain such results, and other computational efforts that may exist) may prove useful. In other words, one may find useful to obtain results that weight in terms of accuracy about eighty percent but with the computational time reduced to one third.

For the first type of factors, the one that concerns the quality of the results, we will assert that the GARCH models clearly outperform more basic models in terms of predictive accuracy. However, the analysis will go further and the PC-GARCH model will be highlighted as an effective solution when also addressing the second type of factors, namely costs.

---

[2] *The quality factor measures preciseness by comparing the forecast with the real (historical) values. A comment on the methodology used for that will follow.*

The remainder of the paper is structured as follows: first we present the main issues regarding the ranking of the performances achieved by volatility forecasting models and we show that there is no consensus as to the accuracy of the forecasts, according to the literature reviewed. Then, a new criterion will be proposed, that of quality/cost ratio, and according to it the PC-GARCH will be justified as an ideal tool to be used for large data matrices. To serve those ends, a presentation of the PC-GARCH model follows. We offer a brief resume of the Principal Component Analysis and PC-GARCH model. However, the PC-GARCH discussion will continue in a subsequent paper that will provide a link between the theoretical discussion and an empirical setting. Both the theoretical and the empirical parts, from both papers, are developed in order to offer a complete understanding of how PC-GARCH works. They will both contribute to the main conclusion of the research that will state the superiority of PC-GARCH in terms of quality/cost criterion as against any alternative models when one deals with large portfolios of data.

2. Assessing the quality of the volatility forecasting techniques

2.1. Advances in volatility modeling

Various techniques designed to obtain reliable volatility forecasts have been continuously produced during the last three decades. They range from extremely simplistic models that employ the so-called “naive” (random walk) assumptions to relatively complex conditional heteroskedastic models of the ARCH group (until GARCH and derivatives of it). The most debated univariate volatility models are the Autoregressive Conditional Heteroskedastic (ARCH) model compiled by Engle (1982) and the Generalized ARCH (GARCH) model compiled by Bollerslev (1986). Numerous extensions of them gained importance, such as the exponential GARCH (EGARCH) model of Nelson (1991) or the conditional heteroskedastic autoregressive moving average (CHARMA) model proposed by Tsay (1987). Other models used for volatility forecasting were the random coefficient autoregressive (RCA) model of Nicholls and Quinn (1982), and the stochastic volatility (SV) models compiled by Melino and Turnbull (1990), Taylor (1994), Harvey, Ruiz and Shephard (1994), and Jacquier, Polson and Rossi (1994), among others.

Comprehensive reviews of the literature have been written by Bollerslev, Chou and Kroner (1992), Bera and Higgins (1993), Bollerslev, Engle and Nelson (1994) and more recently Andersen, Bollerslev, Christoffersen and Diebold (2005). Each model brings with it strengths and weaknesses, and given the coexistence of such a large number of models, designed to serve to similar purposes, it is essential to compare and rank them properly in order to assess those with superior predictive capacities.

In such a large panel of models compounded for volatility forecasting purposes, a general consensus on classifying models in terms of forecast accuracy has not yet been reached. This is due to the fact that the literature contains contradictory evidence as regards the quality of volatility forecasts. Subjectivism arises from various sources, starting with the fact that conditional evidence is unobserved and there is no natural and intuitive way to model the conditional heteroskedasticity, so that each model will try to capture features considered important by the author and, ultimately, from the fact that models with poor forecasting capacities in all empirical tests have not been yet identified.

Although the "clustering" effect of the returns' volatility has long been recognized, it seems that only since the GARCH model was proposed by Bollerslev (1986) could such temporal dependencies be formally modeled using econometric models. This boosted the empirical success of the GARCH-class of models, numerous papers reporting their success in modeling in-sample volatility of asset prices. However, many other papers have suggested the little success of standard volatility models to explain ex post squared returns (Cumby, Figlewski and Hasbrouck, 1993; Figlewski, 1997; Jorion, 1995), recommending the simple moving averages technique for such a purpose. Soon after, a few papers have addressed such problem and restated the usefulness of the GARCH models in providing accurate forecasts (Andersen and Bollerslev, 1998; Andersen, Bollerslev and Lange, 1999). They addressed the latent character of volatility, or inherently unobserved, stochastically evolving through time. Stock volatility consists of intraday volatility and variation between days. Unlike price, which may be described as a flow variable and which can be measured instantaneously, volatility is a stock variable and cannot be directly measured since it is not a directly observable variable. The non-observable character of volatility was a persistent problem posed by econometricians, which makes impossible its exact measurement, but rather its estimation. The non-observability of volatility of stock return results in difficulties related to the assessment of accuracy of different forecasting models. As such, the latent character of volatility converts the problem of volatility assessment into a filtering one in which the "true" volatility is not determined exactly, but only extracted out of a time series with some degree of error. However, the volatility estimates provided by the models will be compared with the "true" underlying volatility. The errors will be then a factor of choosing the model that provides the forecasts or of how the true volatility has been estimated. The previously mentioned papers highlighted a novel point of understating the possible sources of conflicting findings with respect to the models' performance ranking. Therefore, the reference authors underlined the fact that the failure of some GARCH-class of models to provide the most accurate forecasts was not a failure of the GARCH-class of models itself, but rather a failure to specify correctly and accurately the true volatility against which the forecasting performance has been measured. They observe that the standard method of employing ex post daily squared returns as measures of "true" volatility of the daily forecasts is flawed because an estimate compounded in this way comprises a large and noisy independent zero mean constant variance error term that is unrelated to the actual volatility. Andersen and Bollerslev were among the first who proposed that instead of daily squared returns, cumulative squared-returns from intra-day data be used as a more accurate way to express the "true" underlying volatility. Such a measure, called "integrated volatility", paves the way towards a more meaningful and accurate volatility forecast evaluation. This represents a significant step forward in the forecasting debate, because it opens the perspective towards investigating high frequency data for a meaningful, improved way of modeling the daily volatility in empirical estimations.

The above-mentioned literature review reveals the ongoing debate about which models rank best from the point of view of the quality of their forecasts (the upper term of the quality/cost perspective on which we attempt to rank models). However, as reasoned by Matei (2009), the GARCH models may still be considered ultimately as better forecasters, in line with the common sense belief existing both in academia and business. Papers that showed GARCH models' supremacy in any context have not yet been written, although most of them rank GARCH among the better models. We take the study of Matei (2009) as the evidence according to which we will consider that GARCH gives better forecasts and continue the discussion with the lower term of the criterion, that involving costs. In the next section, we will justify the cost-cutting role of the PC-GARCH models, and their capacity of improving forecastibility of the classic GARCH models, which make them an ideal solution in applications involving large data matrices, containing hundreds or thousands of variables.

3. Why PC-GARCH?

GARCH splits the variance forecasts into two components - autocorrelations, or volatility in the past, and innovations, or exogenous shocks in the volatility of returns. Using GARCH(1,1) leads us immediately to the question of how much of the innovation is truly "exogenous" and how much is it explained by "other factors" not considered in the model. To improve the model, we could begin by considering other explanatory variables that could influence the volatility of our estimate (in other words, to endogenise some of the exogeneity). However, adding explanatory variables leads us to a particular weakness of the GARCH: the parameter estimation problem. Due to correlations (usually not zero) between the variables used in the GARCH, the problem requires substantial amounts of data and computational power to come up with a reasonably robust estimate. Thus, we aim to improve the volatility forecast of an asset compared to that obtained with GARCH, but using a more tractable method that handles multiple independent variables. This is accomplished by using PC-GARCH.

In what follows, we discuss the issues of multivariate GARCH estimation uncovered in the previous sections. We know that the number of parameters in a multivariate GARCH increases at the rate of the square of the number of variables. For example, using n variables will necessitate estimation of $\frac{n(n+1)}{2}$ parameters; this is because each additional variable brings with it terms of correlation with the other variables, and each of these correlation terms has its own parameter. The dimensionality of the problem and, hence, the computational power requirement is rather large. Further, a robust parameter estimation imposes demanding data requirements. Apart from estimation problems, there are practical issues of stability of prediction: a large number of parameters as inputs to the model would frequently result in unstable estimates. Due to the inherent data-fitting nature of every statistical procedure, there may be noise in the estimation period that is captured as signals into this model.

One of the methods proposed to make the problem tractable is the PC-GARCH (another algorithm that also uses Principal Components but is different in its implementation is that called Orthogonal GARCH). In this study, a simple model will be used to illustrate the power of this method; in particular, the power of the Principal

Component Analysis (PCA) used in conjunction with GARCH to solve the problems stated above.

As noted earlier, the increased dimensionality of the multivariate GARCH is due to the large number of covariances between the independent variables that enter the parameter space. Therefore, making these covariances zero reduces the dimension of the problem to *n* (or we will have to estimate only *n* parameters, each for the GARCH and ARCH). Thus, PCA is the tool to be used to simplify the problem and make it tractable. PCA is a method of transforming original independent variables into orthogonal factors. Thus, using *n* (possibly correlated) independent variables and applying PCA reduces the number of parameters to be estimated to *2n+1* instead of $\frac{n(n+1)}{2}$ (a linear instead of a quadratic increase in the number of parameters to be estimated). Thus, the PCA method helps us reduce the modeling problem into *n* univariate GARCH models. The methodology for the analysis to be followed in the paper is that developed by Burns (2005). There are alternative methods developed in the literature that use PCA in conjunction with GARCH; such examples are Alexander (2000) and van der Weide (2002).

Briefly stating the problem in mathematical terms, we have the variable *y* which is dependent on *k* independent variables. The *n* historical observations of each of these *k* independent variables are arranged in a matrix **X** of dimension $n \times k$, and the n historical observations of the dependent variable are arranged in a $n \times 1$ matrix, **Y**. In very general terms, we wish to find the function *f* that maps the independent variables onto the dependent variable: **Y** = *f*(**X**). To summarize, the problem of finding this general function is that

1. even a small increase in *k* makes the problem computational and data intensive and,
2. some of the independent variables are correlated: they contain common information, and we wish to coalesce similar information into a single variable that represents that information and have uncorrelated independent explanatory variables.

3.1 Principal Component Analysis (PCA): A brief introduction to the method

Principal Component Analysis is an algorithm used in Factor Analysis. Factor Analysis is a generic method given to a class of multivariate statistical methods that has as its main goal to identify the underlying structure in a data matrix. Specifically, the Factor Analysis has two primary uses: summarization and data reduction. Summarization results from describing the data with a much smaller number of variables, while reduction comes from transforming the data matrix into a score matrix, in which each column stands for a factor.

The Principal Component Analysis is a method used for extracting uncorrelated sources of information in the data. From a set of *k* stationary returns, up to *k* orthogonal stationary variables will return, called Principal Components (PCs) or variates. PCA is a classical technique to derive such uncorrelated variates. An output of the method also states how much of the total variation in the original data is explained by each PC.

Due to the high sensitivity of the results to re-scaling data, before proceeding to the analysis, the standard procedure is to normalize the data. Thus, we assume that each column in the stationary matrix has mean zero and variance one, after previously subtracting the sample mean and dividing by the sample standard deviation.

We start with the matrix **X** with columns $(x_1, x_2, \ldots, x_k)$, where $\{x_i, 1 \le i \le k\}$ is such that **X'X** is a k x k symmetric matrix, having one on its diagonal. **Ω** = **X'X** is the variance-covariance matrix of the variables in **X**, and thus is positive and definite. Each principal component will be then a combination of these columns:

$$p_m = a_{1,m}x_1 + a_{2,m}x_2 + \ldots + a_{k,m}x_k \,,\ 1 \le m \le k \quad (1)$$

In a matrix form, (1) can be written as **P** = **XA**, where **A** is called the matrix of the eigenvectors of **Ω.** The weights $a_{i,m}$ for each x_i are chosen from the set of eigenvectors of the correlation matrix **Ω** such that:

1. The Principal Components (PCs) are orthogonal. Thus, we impose the orthogonality condition to the matrix **P** of the principal components (as this is the main property of such PCs) and, accordingly, we have to find the matrix **A** of weights that fulfils this condition. In other words, we want to know which are the $a_{i,m}$'s such that their matrix, multiplied by an **X** matrix of observations, gives an orthogonal matrix.
2. The first principal component explains the maximum amount of total variation in **X,** the second component explains the maximum amount of the remaining variation, and so on.

We know from matrix algebra that if we choose the matrix **A** to be composed of orthogonal unit eigenvectors of **X'X**, then the resulting PCs are orthogonal. It means, then, the only condition for **P** to be orthogonal is that the columns of **A** are orthogonal.

We next order the columns of **A** in descending order. Thus, if **A**($a_{i,j}$), $i, j \in \{1,\ldots,k\}$ then the m^{th} column of **A**, denoted by $a_m = (a_{1,m}, a_{2,m}, \ldots, a_{k,m})'$ is the $(k\ x1)^{th}$ eigenvector corresponding to the eigenvalue λ_m, and the column must be ranked so that $\lambda_1 > \lambda_2 > \ldots > \lambda_k > 0$.

We now define a new matrix **Λ** as the diagonal matrix of the eigenvalues of **Ω** (with λ_i on the diagonal), and we note that **X'XA=AΛ,** from where it results that **Λ** =**A'X'XA** = **A'ΩA.** Also, we have **(XA)'XA = A'X'XA = Λ.** The above relationship then becomes

$$\Lambda = P'P = A'X'XA = A'\Omega A \quad (2)$$

Since **Λ** is a diagonal matrix, and it is the variance-covariance matrix of **P**, this implies that the components (columns) of **P** are uncorrelated. Since **A** is orthogonal, **A'**=**A^{-1}** and **P'P**=**Λ. A'**=**A^{-1}** is equivalent to **X**=**PA'** that is $X_i=w_{i1}P_1+ w_{i2}P_2+\ldots+ w_{ik}P_k$ where X_i and P_i denote the columns of **X** and **P**, respectively. Thus, each data vector is a linear combination of the principal components. The proportion of the total variation in **X,** explained by the m^{th} principal component, is λ_m/(sum of the eigenvalues).

Thus, the operation of scaling the original variables with the matrix of orthogonal unit eigenvectors **A** gives us uncorrelated components (PCs) that we could use to reduce the earlier multivariate GARCH problem to a set of univariate GARCH problems.

3.2 Methodology

The results of Matei (2009) will be used and, accordingly, the superiority of GARCH models will be assumed. However, as noted earlier, there are significant costs attached to using GARCH. Because of this concern, at this point we address the second factor that supports second part's main conclusion, using Principal Component Analysis (PCA) may be an effective and at hand solution. PCA does two things that improve the model: one is that it reduces the dimensionality of the problem, and the other is that it excludes autocorrelations in the data. The only subjective point in the problem is the cut point the user has to choose. In other words, how much of the preciseness should be sacrificed for how much time saved. This ability of choosing the output to time report gives the user of the model flexibility, allowing for tailored options according to activities and companies' specific features.

4. Experimental study

4.1 Data setting

The task is to estimate the volatility of the return of a particular portfolio formed of inter-correlated stocks (Adobe, Apple, Autodesk, Cisco, Dell, Microsoft and 3M) using PCA in conjunction with the GARCH model. The selection of these seven stocks has been driven by the fact that PCA works best when there is a reasonable amount of correlation between the variables; there is good reason to suppose that the chosen seven US stock returns would be correlated.

We are thus in a position to argue that Microsoft, for instance, is not influenced only by its own past, but also by the past of the other shares included in this selected portfolio. As a hypothetical example, a lot of volatility in Microsoft returns could signal the uncertainty in the technology sector; the next day Adobe would take into account[3] the uncertainty in the technology sector induced by Microsoft and extrapolate that into the uncertainty forecasts of its activity. While limiting my study to the seven stock return series mentioned above, we do not suggest that these are the only shares that matter - this study is simply a means to demonstrate the power of the technique and claim that this model is "the best" in forecasting volatility of portfolios with large intercorrelated time series.

The data range between February 16th, 1990, and June 18th, 2009. This gives us a total of 5044 return observations. As discussed above, the choice of these equity returns is determined by their high (as expected) correlations, which makes their cases ideal for applying PCA. However, the seven stocks are fundamentally different too, which seems interesting to isolate the effects of their composition. First, there will be a review of the data used.

[3] *Through the aggregation of the trades of market participants. No particular information transmission mechanism is supposed here, only that some such mechanisms hold.*

4.2 Preparing data for the PCA

The selected data consists of n = 5044 observations of returns for each of Adobe, Apple, Autodesk, Cisco, Dell, Microsoft and 3M. We want to find the principal components. Since each component is a linear combination of the centered variables, we must first obtain these centered variables by subtracting the mean of each x_i. Thus, the mean of each of these stock returns will be calculated (thus, Adobe, Apple, Autodesk, Cisco, Dell, Microsoft and 3M have their mean returns over the sample period) and this will be then subtracted from x. We thus obtain the matrix of the centered variables. To obtain the matrix of the principal components, we multiply **X** by **A** matrix, where $\mathbf{A}$ is the matrix of loadings. We have then $\mathbf{XA} = \mathbf{P} = (p_1, p_2, \ldots, p_7)$. We want to find a matrix **A** that, when multiplied with the matrix of the centered variables, **X**, gives us an orthogonal matrix **P** with which we can work (in each cell of **P** we shall have a principal component that will be a linear combination between the centered variables, *x*'s and *a*'s). If we decide to use a number of PCs smaller than the original number of variables, we would lose some information, but keep data **P** uncorrelated, which still can explain **Y** (see Alexander, 2000, for details). Thus, to find **P** we must find **A** that solves **XA** = **P** and we impose the orthogonality condition for **P**, that is $Var(\mathbf{P}) = \begin{pmatrix} var1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & var7 \end{pmatrix}$.

4.2.1 Solving orthogonality

By definition, the variance-covariance matrix **P** is equal to **P'P**. Using the property that (**AB**)' = **B'A'**, we find that Var(**P**) = **P'P**= (**XA**)**'XA** = **A'X'XA**, and **X'X** = Var(**X**). We call **X'X** = Var(**X**) =**Ω**, from which it results that Var(**P**) = **P'P** = **A' Ω A**.

Since one of our initial problems was that some elements were correlated, we want a **P** such that it is composed of orthogonal elements. Thus, next we impose the orthogonality condition on **P**. From a larger matrix of data **X**, we want to obtain the matrix **P** of smaller or equal dimension that has only uncorrelated values, each element of **P** being a linear combination of the elements of **X**.

To see the meaning of the term "uncorrelated elements of **P**", let us call $\mathbf{P} = (p_1, p_2, \ldots, p_7)$, then $\mathbf{P}' = (p_1, p_2, \ldots, p_7)'$, and $\mathbf{P}'\mathbf{P} = Var(\mathbf{P})$, which is symmetric. Orthogonality of **P** means that $cov(p_i, p_j) = 0$. This implies that $Var(\mathbf{P}) = diag(\sigma(p_1), \ldots, \sigma(p_7))$. Thus, we see that the variance-covariance matrix of a matrix of orthogonal elements is a diagonal matrix. From matrix algebra, we use the result that the matrix **A** is the matrix of orthogonal unit eigenvectors of **Ω**.

4.2.2 Finding the matrix of Principal Components

Let us sum up the problem: we want to use *x*'s to explain the *y*'s, but the *x*'s are too many (*k* is too large). We chose to make *k* smaller, so that we must pick factors that explain most of the variation (or as much as possible with a *k* that makes the problem tractable). We are looking to find the linear relationship of *x*'s that gives us the orthogonal *p*'s.

In our problem, *p*'s are the new *x*'s, so we have to rearrange the **P** matrix in the descending order of columns to see which *p*'s are the highest. Once we rearrange it, we impose the **P** = **XA** condition (where **A** is the matrix of factor loadings and **X** is the matrix as defined above). According to the matrix notation, this translates into a $n \times 7$ matrix having $\sum_{i=1}^{7}(a_{ki}(x_i - \bar{x}_i))$, $k = 1,...,7$ on the column, and where a_{ij}; $1 \le i, j \le 7$ represent the factor loadings. We have thus linear transformations of the *x*'s that give us *p*'s; in other words we have transformed the *x*'s into orthogonal *p*'s. This means that $p_k = \sum_{i=1}^{7}(a_{ki}(x_i - \bar{x}_i))$, $k = 1,...,7$ ..

We know *x*'s, but we do not know *a*'s. What is left to do is that we have to find the *a*'s that give us the orthogonal factors, since *a*'s signify the weights of each of the *x*'s. For this, because we want orthogonality, we impose the restriction that the resulting covariance matrix is just a diagonal matrix (as done before); after this, we reduce *x*'s to $p_1, p_2, ..., p_7$. Once we enter all the *x*'s and all *y*'s, the software gives us the factor loadings (*a*'s) and the eigenvalues (*λ*'s, that are actually the $\sigma's$) that come from the condition of orthogonality $Var(\mathbf{P}) = diag(\sigma(p_1),...,\sigma(p_7))$. The eigenvectors are actually the columns of **A** (**A**= $\left(\underset{Eigenvector\,1}{\overset{\lambda_1}{\begin{bmatrix} ... \\ ... \\ ... \end{bmatrix}}} \quad \underset{Eigenvector\,2}{\overset{\lambda_2}{\begin{bmatrix} ... \\ ... \\ ... \end{bmatrix}}} \quad ... \quad \underset{Eigenvector\,7}{\overset{\lambda_7}{\begin{bmatrix} ... \\ ... \\ ... \end{bmatrix}}} \right)$). They are orthogonal and are of unit length. After we find the factor loadings and the eigenvalues, we can pick up *p*'s (which now are uncorrelated, since we impose the orthogonality condition) that show the highest variance (that is given by the eigenvalues).

We now have the orthogonal **P** whose values are a linear combination of the independent variables **X**. We now can work with **P** to make forecasts on the variations of **Y**.

5. PC-GARCH implementation - The algorithm

The algorithm uses as reference the work of Burns (2005); it uses five of the seven steps described for handling a multivariate GARCH with univariate estimations, in order to obtain formulations of standardized innovations expressed in terms of principal components. The new formulations will allow a more efficient GARCH dissemination between autocorrelations and innovations, improving the estimation exercise by obtaining independent innovations; this highlights one of the PC-GARCH strengths. As described in detail in earlier sections, PC-GARCH will be used to enable a tractable version of multivariate GARCH. This tractability arises from the lack of correlation among the multiple variables used, reducing the parameter set to a manageable number. In this section, a brief overview of the algorithm will be provided and in the next, we shall advance with further details.

Firstly, a univariate GARCH for each of the seven price returns will be estimated; this step establishes whether a multivariate GARCH is required in the first place. If the univariate GARCH models sufficiently describe "reality", the errors of these models must be uncorrelated. Strong correlation between the errors implies the presence of a common factor that drives the seven return series simultaneously. Instead of simply using the autocorrelations of the same stock, we can exploit the autocorrelations among the various stock time series. In more intuitive terms, a correlation among the errors of the return series implies that 'there is information' in the other returns that can be used to forecast the volatility of each return series. Recall here that GARCH is a technique that splits variances into those caused by autocorrelation (effects of the past) and innovations (errors, defined essentially as the difference between the predicted and the observed values). Thus, a correlation between the errors implies that what the univariate GARCH model presumes to be innovations are not truly innovations, but can be explained by movements in the other stock returns.

Since our test yields that a multivariate GARCH is reasonable in this scenario, the next step is to find the seven uncorrelated factors that drive the price returns. As noted in the theoretical discussion above, since we have seven variables (that are not collinear, even if they are highly correlated), we work in a seven-dimensional environment where each dimension represents the returns of a stock. These seven dimensions are, as we saw, highly correlated; hence not orthogonal. As stressed repeatedly, these non-orthogonal but highly correlated variables result in tractability problems, and thus we want to identify orthogonal (uncorrelated) factors that we could conveniently use.

The PCA is applied on the residuals of the previous GARCH; we try to find seven uncorrelated sources of "errors", these "errors" being the innovations obtained from the earlier univariate GARCH. Intuitively, what we try here is to isolate the seven different factors that drive innovations. Since we try to find the factors that drive stock returns of the same country, we expect to find one factor whose effect on all these returns is large, and six other stock-specific factors.

An output of the principal component analysis is the matrix of coefficients. This matrix will be used to estimate the new residuals due to each PC. For reasons we shall go through in greater detail in the presentation below, we do not drop any of the principal components, but use all seven. These new residuals are thus orthogonal to each other, and running a multivariate GARCH on them is equivalent to running seven separate univariate GARCH models. This reduction of multivariate GARCH to a set of univariate GARCH models is a key reason for the popularity of the PC-GARCH technique. The univariate GARCH models will be duly run on each of these transformed residuals. Intuitively, what was previously "unexplained" will now get "explained" on the basis of the seven orthogonal factors. But clearly, our aim has been to obtain a GARCH model of the stock return volatilities, and not the GARCH model of the transformed residuals. Thus, we need to transform these GARCH models back into the space of return volatilities. This is easy: we note that pre-multiplying by the inverse of the matrix of coefficients and post-multiplying by the matrix of coefficients gives us back our desired original variance-covariance matrix (in this case, this is a set of seven GARCH models). This will be explained in greater detail in the section below.

5.1 Step one: Estimating the dimensions (p,q) of the GARCH models

As discussed earlier, the first step in running a PC-GARCH algorithm is to begin with a univariate GARCH and check the necessity of adding extra variables.

The rationale is rather practical - to use a parsimonious model if it is "good enough", where the validity of the model depends on the user's requirements. Thus, the attempt here should be to use the best possible univariate GARCH model. This means that the coordinates *p* and *q* of GARCH(*p,q*) must be selected in order to optimize the trade-off between the extra parameters and the extra predictive ability achieved. The selection of the variables *p* and *q* is optimized independently of the other models under consideration.

Since the aim is to illustrate the PC-GARCH approach, we simply choose a GARCH(1,1) and fit each of the daily return volatilities. The formula for the GARCH(1,1) model is $y_t = C + \varepsilon_t, \sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \beta_1 \varepsilon_{t-1}^2$.

We may see from the results that we can reject the null hypothesis that α_0 and α_1 are separately equal to zero (since the t-values are outside +/-1.96 interval, thus we are in the rejection region). In other words, it is appropriate to model the time series of volatility as a GARCH(1,1). We pause to consider the "visual effect" of the GARCH(1,1) decomposition; also contrasting this with the decomposition after the PC-GARCH procedure.

5.2 Step two: Obtaining residuals of GARCH(1,1) and standardizing them

From the above methodology, we see that the GARCH(1,1) models for Adobe, Apple, Autodesk, Cisco, Dell, Microsoft and 3M are as follows:

ADOBE: $y_t = 12.06 \times 10^{-4} + \varepsilon_t$, $\sigma_t^2 = 3.06 \times 10^{-6} + 0.971\sigma_{t-1}^2 + 0.028\varepsilon_{t-1}^2$

APPLE: $y_t = 18.09 \times 10^{-4} + \varepsilon_t$, $\sigma_t^2 = 52.95 \times 10^{-6} + 0.8681\sigma_{t-1}^2 + 0.084\varepsilon_{t-1}^2$

AUTODESK: $y_t = 14.51 \times 10^{-4} + \varepsilon_t$, $\sigma_t^2 = 62.19 \times 10^{-6} + 0.838\sigma_{t-1}^2 + 0.102\varepsilon_{t-1}^2$

CISCO: $y_t = 19.30 \times 10^{-4} + \varepsilon_t$, $\sigma_t^2 = 9.17 \times 10^{-6} + 0.920\sigma_{t-1}^2 + 0.071\varepsilon_{t-1}^2$

DELL: $y_t = 11.32 \times 10^{-4} + \varepsilon_t$, $\sigma_t^2 = 3.59 \times 10^{-6} + 0.950\sigma_{t-1}^2 + 0.049\varepsilon_{t-1}^2$

MICROSOFT: $y_t = 10.01 \times 10^{-4} + \varepsilon_t$, $\sigma_t^2 = 6.41 \times 10^{-6} + 0.920\sigma_{t-1}^2 + 0.069\varepsilon_{t-1}^2$

3M: $y_t = 3.89 \times 10^{-4} + \varepsilon_t$, $\sigma_t^2 = 2.66 \times 10^{-6} + 0.957\sigma_{t-1}^2 + 0.031\varepsilon_{t-1}^2$.

For each day (of the 5044 days of our sample), we calculate the volatility forecast and call this σ_t. We use this calculated variance forecast to obtain the standardized residuals of the daily returns. In other words, we calculate $\frac{y_t - \bar{y}_t}{\sigma_t}$ as for each *t* we know the return y_t. Thus, we now have a matrix of standardized residuals, **R**. This matrix is of *5044 x 7* dimension (days x number of stocks). If the univariate GARCH(1,1) was an "adequate description" of "reality", we should find that the columns of **R** have zero mean (which they do by our construction), and a variance of unit (which needs not be true, since we use the forecast variance estimate, and not the true variance) and the covariance between the rows should be zero (meaning that there are no "common factors" outside the explanation provided by autocovariance of daily residuals).

As regards the comparison of the residuals, of the conditional standard deviations, and of the returns, we split the variance into variance innovations and conditional standard deviations in order to investigate if the fitted innovations exhibit volatility clustering (see Appendix, Figures 1 to 7). From the visual inspection of the graphs of each stock we can observe volatility clustering in innovations and returns. As previously said, we want to see if the innovations in the seven price returns are uncorrelated, which will show us the necessity of performing a PC-GARCH. As a hint for their correlation, we see in these graphs that innovations vary around approximately identical dates, due to probably common factors that influence all of them. Also, we can observe that for each stock $\beta_1 + \alpha_1$ is very close to 1, which means it is very close to the integrated, non-stationary boundary given by the constraints stated for a standard GARCH model.

We saw that the previous fitted innovations display volatility clustering. However, if we plot the standardized innovations (the innovations divided by their conditional standard deviation), they appear generally stable with little clustering. The existence of GARCH effects and of correlation between innovations that disappears after treating the data makes us conclude that the GARCH model is a suitable model to be used to explain the variances of the seven stocks. Thus, our intuitive choice of the seven stock returns is justified, and we proceed to the next stage.

5.3 Third step: Principal component analysis of standardized residuals

We have seen details of the PCA method above, so we just confine ourselves to reporting the major results here. The matrix of standardized residuals is the matrix on which we will perform the PCA, because we wish to identify the common causes of what the GARCH(1,1) model leaves out as unexplained innovations.

We perform the Principal Component Analysis to the standardized innovations. The PCA gives us seven mutually orthogonal linear combinations of the standardized residuals. The *latent* output gives us the eigenvalues. Accordingly, we can calculate their power, meaning the percentage of variation that each principal component explains.

We see that most of the variance is explained by the first principal component, to which the change in volatility of all seven return series contributes by a very similar magnitude. This accord with the initial intuition that since all seven are stock prices based in the US IT sector, there is a large common factor that moves all of them in the same direction.

We can observe that P1 (=0.3793*Adobe standardized innovation + 0.3639*Apple standardized innovation + 0.3448*Autodesk standardized innovation + 0.4337*Cisco standardized innovation + 0.4022*Dell standardized innovation + 0.4201*Microsoft standardized innovation + 0.2804*3M standardized innovation) is the first principal component that explains almost 43% of the variance of the standardized residuals. The second factor explains about 12% of the variance in standardized residuals - we note that this it is positively weighted by Apple, Cisco, Dell and Microsoft, but negatively weighted by Adobe, Autodesk and 3M. The third factor explains only 11% of the variance in the standardized residuals, and this seems to be positively driven by the Adobe and Autodesk returns, while negatively driven by Apple, Cisco, Dell, Microsoft and 3M. The rest of the principal components weigh less than 10% each.

Thus, we obtain, in decreasing order, seven PCs that drive the standardized excess returns - which the GARCH model earlier called innovations. We see that they are not really all innovations: that most of this innovation is driven by one major factor that drives all seven stocks together. We could now choose, for the sake of parsimony, to keep just this first PC that explains about 43% of the so-called innovations, and leave the rest out. But, a technical issue is that leaving out any of the principal components may occasionally lead to meaningless results since we not guarantee that the resulting variance covariance matrix is positive and definite (see Alexander, 2000). Since in this particular case we do not have too many variables, we can include all the factors to ensure that our results are always meaningful.

5.4 Fourth step: Running GARCH(1,1) on the PCs

Once again, recalling that the GARCH(1,1) model is $y_t = C + \varepsilon_t, \sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \beta_1 \varepsilon_{t-1}^2$, we run this model on the newly obtained PCs. Thus, we have a GARCH model that predicts the volatilities of the seven PCs. However, we began with the aim of obtaining volatility forecasting models for the daily return series. We shall see that this is achieved through a simple linear transformation in the next section.

5.5 Fifth step: Obtaining the GARCH model of the indices

We note that the GARCH(1,1) models that we obtained previously are for the principal components. For instance, $\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \beta_1 \varepsilon_{t-1}^2$ gives us the volatility forecast of each PC. One may see that since these seven PCs are orthogonal to each other, we can write their variance-covariance matrix in a diagonal form. In other words, recalling our earlier discussion, we re-visit equation (2) **Λ** = **P'P** = **A'X'XA** = **A'ΩA**. We now have **Λ,** which consists of the volatility forecasts of the seven PCs. Using the property that **A'**=$\mathbf{A}^{-1}$**,** we see that **AΛA'** = **AA'ΩAA'** = **Ω**. Thus, the simple linear transformation of pre-multiplying the forecasts by the matrix **A** and post-multiplying by **A'** gives us the volatility forecasts of the seven standardized return series. The seven equations we obtain are presented below.

Formula 1: Multivariate PC-GARCH model for ADOBE daily return volatility.

$$\sigma_t^2 = 37.66\times10^{-2} + \begin{pmatrix} 14.21\times10^{-2} \\ 0.07\times10^{-2} \\ 6.36\times10^{-2} \\ 12.01\times10^{-2} \\ 0 \\ 7.45\times10^{-2} \\ 0.65\times10^{-2} \end{pmatrix}' \begin{pmatrix} (\sigma_{t-1}^{p_1})^2 \\ (\sigma_{t-1}^{p_2})^2 \\ (\sigma_{t-1}^{p_3})^2 \\ (\sigma_{t-1}^{p_4})^2 \\ (\sigma_{t-1}^{p_5})^2 \\ (\sigma_{t-1}^{p_6})^2 \\ (\sigma_{t-1}^{p_7})^2 \end{pmatrix} + \begin{pmatrix} 0.14\times10^{-2} \\ 0.02\times10^{-2} \\ 0.18\times10^{-2} \\ 0.93\times10^{-2} \\ 1.47\times10^{-2} \\ 0.69\times10^{-2} \\ 0.07\times10^{-2} \end{pmatrix}' \begin{pmatrix} (\varepsilon_{t-1}^{p_1})^2 \\ (\varepsilon_{t-1}^{p_2})^2 \\ (\varepsilon_{t-1}^{p_3})^2 \\ (\varepsilon_{t-1}^{p_4})^2 \\ (\varepsilon_{t-1}^{p_5})^2 \\ (\varepsilon_{t-1}^{p_6})^2 \\ (\varepsilon_{t-1}^{p_7})^2 \end{pmatrix}$$

Formula 2: Multivariate PC-GARCH model for APPLE daily return volatility.

$$\sigma_t^2 = 28.26\times10^{-2} + \begin{pmatrix} 13.08\times10^{-2} \\ 2.70\times10^{-2} \\ 6.06\times10^{-2} \\ 29.02\times10^{-2} \\ 0 \\ 4.73\times10^{-2} \\ 0.04\times10^{-2} \end{pmatrix}' \begin{pmatrix} (\sigma_{t-1}^{p_1})^2 \\ (\sigma_{t-1}^{p_2})^2 \\ (\sigma_{t-1}^{p_3})^2 \\ (\sigma_{t-1}^{p_4})^2 \\ (\sigma_{t-1}^{p_5})^2 \\ (\sigma_{t-1}^{p_6})^2 \\ (\sigma_{t-1}^{p_7})^2 \end{pmatrix} + \begin{pmatrix} 0.13\times10^{-2} \\ 0.59\times10^{-2} \\ 0.17\times10^{-2} \\ 2.24\times10^{-2} \\ 0.50\times10^{-2} \\ 0.44\times10^{-2} \\ 0.44\times10^{-4} \end{pmatrix}' \begin{pmatrix} (\varepsilon_{t-1}^{p_1})^2 \\ (\varepsilon_{t-1}^{p_2})^2 \\ (\varepsilon_{t-1}^{p_3})^2 \\ (\varepsilon_{t-1}^{p_4})^2 \\ (\varepsilon_{t-1}^{p_5})^2 \\ (\varepsilon_{t-1}^{p_6})^2 \\ (\varepsilon_{t-1}^{p_7})^2 \end{pmatrix}$$

Formula 3: Multivariate PC-GARCH model for AUTODESK daily return volatility.

$$\sigma_t^2 = 36.49\times10^{-2} + \begin{pmatrix} 11.74\times10^{-2} \\ 0.23\times10^{-2} \\ 34.16\times10^{-2} \\ 0.49\times10^{-2} \\ 0 \\ 0.18\times10^{-2} \\ 0.03\times10^{-2} \end{pmatrix}' \begin{pmatrix} (\sigma_{t-1}^{p_1})^2 \\ (\sigma_{t-1}^{p_2})^2 \\ (\sigma_{t-1}^{p_3})^2 \\ (\sigma_{t-1}^{p_4})^2 \\ (\sigma_{t-1}^{p_5})^2 \\ (\sigma_{t-1}^{p_6})^2 \\ (\sigma_{t-1}^{p_7})^2 \end{pmatrix} + \begin{pmatrix} 0.11\times10^{-2} \\ 0.05\times10^{-2} \\ 0.95\times10^{-2} \\ 0.04\times10^{-2} \\ 0.90\times10^{-2} \\ 0.02\times10^{-2} \\ 0.38\times10^{-4} \end{pmatrix}' \begin{pmatrix} (\varepsilon_{t-1}^{p_1})^2 \\ (\varepsilon_{t-1}^{p_2})^2 \\ (\varepsilon_{t-1}^{p_3})^2 \\ (\varepsilon_{t-1}^{p_4})^2 \\ (\varepsilon_{t-1}^{p_5})^2 \\ (\varepsilon_{t-1}^{p_6})^2 \\ (\varepsilon_{t-1}^{p_7})^2 \end{pmatrix}$$

Formula 4: Multivariate PC-GARCH model for CISCO daily return volatility.

$$\sigma_t^2 = 10.12\times10^{-2} + \begin{pmatrix} 18.58\times10^{-2} \\ 1.03\times10^{-2} \\ 0.37\times10^{-2} \\ 4.15\times10^{-2} \\ 0 \\ 2.28\times10^{-2} \\ 50.33\times10^{-2} \end{pmatrix}' \begin{pmatrix} (\sigma_{t-1}^{p_1})^2 \\ (\sigma_{t-1}^{p_2})^2 \\ (\sigma_{t-1}^{p_3})^2 \\ (\sigma_{t-1}^{p_4})^2 \\ (\sigma_{t-1}^{p_5})^2 \\ (\sigma_{t-1}^{p_6})^2 \\ (\sigma_{t-1}^{p_7})^2 \end{pmatrix} + \begin{pmatrix} 0.18\times10^{-2} \\ 0.22\times10^{-2} \\ 0.01\times10^{-2} \\ 0.32\times10^{-2} \\ 0.10\times10^{-2} \\ 0.21\times10^{-2} \\ 5.60\times10^{-2} \end{pmatrix}' \begin{pmatrix} (\varepsilon_{t-1}^{p_1})^2 \\ (\varepsilon_{t-1}^{p_2})^2 \\ (\varepsilon_{t-1}^{p_3})^2 \\ (\varepsilon_{t-1}^{p_4})^2 \\ (\varepsilon_{t-1}^{p_5})^2 \\ (\varepsilon_{t-1}^{p_6})^2 \\ (\varepsilon_{t-1}^{p_7})^2 \end{pmatrix}$$

Formula 5: Multivariate PC-GARCH model for DELL daily return volatility.

$$\sigma_t^2 = 16.17\times10^{-2} + \begin{pmatrix} 15.98\times10^{-2} \\ 2.37\times10^{-2} \\ 5.08\times10^{-2} \\ 5.78\times10^{-2} \\ 0 \\ 38.39\times10^{-2} \\ 2.91\times10^{-2} \end{pmatrix}' \begin{pmatrix} (\sigma_{t-1}^{p_1})^2 \\ (\sigma_{t-1}^{p_2})^2 \\ (\sigma_{t-1}^{p_3})^2 \\ (\sigma_{t-1}^{p_4})^2 \\ (\sigma_{t-1}^{p_5})^2 \\ (\sigma_{t-1}^{p_6})^2 \\ (\sigma_{t-1}^{p_7})^2 \end{pmatrix} + \begin{pmatrix} 0.15\times10^{-2} \\ 0.51\times10^{-2} \\ 0.14\times10^{-2} \\ 0.45\times10^{-2} \\ 0.16\times10^{-2} \\ 3.57\times10^{-2} \\ 0.32\times10^{-2} \end{pmatrix}' \begin{pmatrix} (\varepsilon_{t-1}^{p_1})^2 \\ (\varepsilon_{t-1}^{p_2})^2 \\ (\varepsilon_{t-1}^{p_3})^2 \\ (\varepsilon_{t-1}^{p_4})^2 \\ (\varepsilon_{t-1}^{p_5})^2 \\ (\varepsilon_{t-1}^{p_6})^2 \\ (\varepsilon_{t-1}^{p_7})^2 \end{pmatrix}$$

Formula 6: Multivariate PC-GARCH model for MICROSOFT daily return volatility.

$$\sigma_t^2 = 13.55\times10^{-2} + \begin{pmatrix} 17.43\times10^{-2} \\ 0.34\times10^{-2} \\ 0.60\times10^{-2} \\ 8.44\times10^{-2} \\ 0 \\ 22.47\times10^{-2} \\ 23.38\times10^{-2} \end{pmatrix}' \begin{pmatrix} (\sigma_{t-1}^{p_1})^2 \\ (\sigma_{t-1}^{p_2})^2 \\ (\sigma_{t-1}^{p_3})^2 \\ (\sigma_{t-1}^{p_4})^2 \\ (\sigma_{t-1}^{p_5})^2 \\ (\sigma_{t-1}^{p_6})^2 \\ (\sigma_{t-1}^{p_7})^2 \end{pmatrix} + \begin{pmatrix} 0.17\times10^{-2} \\ 0.07\times10^{-2} \\ 0.02\times10^{-2} \\ 0.65\times10^{-2} \\ 0.220\times10^{-2} \\ 2.09\times10^{-2} \\ 2.60\times10^{-2} \end{pmatrix}' \begin{pmatrix} (\varepsilon_{t-1}^{p_1})^2 \\ (\varepsilon_{t-1}^{p_2})^2 \\ (\varepsilon_{t-1}^{p_3})^2 \\ (\varepsilon_{t-1}^{p_4})^2 \\ (\varepsilon_{t-1}^{p_5})^2 \\ (\varepsilon_{t-1}^{p_6})^2 \\ (\varepsilon_{t-1}^{p_7})^2 \end{pmatrix}$$

Formula 7: Multivariate PC-GARCH model for 3M daily return volatility.

$$\sigma_t^2 = 38.49\times10^{-2} + \begin{pmatrix} 7.77\times10^{-2} \\ 34.09\times10^{-2} \\ 4.28\times10^{-2} \\ 0.02\times10^{-2} \\ 0 \\ 0.04\times10^{-2} \\ 0.09\times10^{-2} \end{pmatrix}' \begin{pmatrix} (\sigma_{t-1}^{p_1})^2 \\ (\sigma_{t-1}^{p_2})^2 \\ (\sigma_{t-1}^{p_3})^2 \\ (\sigma_{t-1}^{p_4})^2 \\ (\sigma_{t-1}^{p_5})^2 \\ (\sigma_{t-1}^{p_6})^2 \\ (\sigma_{t-1}^{p_7})^2 \end{pmatrix} + \begin{pmatrix} 0.07\times10^{-2} \\ 7.39\times10^{-2} \\ 0.12\times10^{-2} \\ 0.15\times10^{-4} \\ 0.03\times10^{-2} \\ 0.36\times10^{-4} \\ 0.98\times10^{-4} \end{pmatrix}' \begin{pmatrix} (\varepsilon_{t-1}^{p_1})^2 \\ (\varepsilon_{t-1}^{p_2})^2 \\ (\varepsilon_{t-1}^{p_3})^2 \\ (\varepsilon_{t-1}^{p_4})^2 \\ (\varepsilon_{t-1}^{p_5})^2 \\ (\varepsilon_{t-1}^{p_6})^2 \\ (\varepsilon_{t-1}^{p_7})^2 \end{pmatrix}$$

These equations are used in the following manner:

1. Use the matrix of principal components to calculate the seven PCs from the daily returns.
2. Calculate the volatility and innovation in the returns on the PCs.
3. Substitute the values calculated above in the appropriate multivariate GARCH model to obtain the volatility forecasts.

6. Conclusions on using PC-GARCH

We saw that PC-GARCH is a useful tool to reduce the dimensionality of the multivariate GARCH problem and to obtain robust and stable estimates using orthogonal PCs. While we have mentioned its many benefits, we would like to conclude with visual evidence of how the "innovation" claimed by the GARCH(1,1) is not really innovation. We present the decomposition of Adobe, Apple, Autodesk, Cisco, Dell, Microsoft and 3M volatilities after the PC-GARCH in Figures 8 to 14 in the Appendix. We see a marked difference between the graphs of innovations displayed previously. What is extremely noticeable is that volatility peaks that occurred at the same time (especially the high volatilities during 1993, 1995, 2006, 2008 and 2009) are not considered to be "innovations" anymore, but are considered explained by the simultaneous rise in the innovations of the others. Thus, the innovations are "truly innovations", which perhaps could be explained by other factors. However, while every model can be improved, the improvement usually comes at a cost. One of the costs is that of over-fitting the model to the sample data, which makes out-of-sample model performance crucial for understanding which model to use. We leave this, as we should, in the hands of the user.

The scope of the empirical part was to reveal the superiority of PC-GARCH in terms of quality of results/costs involved when dealing with large samples of data. It has empirically proved how large GARCH correlation matrices can be obtained by using only univariate GARCH estimation techniques on principal components of the original return series. The advantages of such method are as follows:

- It minimizes computational efforts (by transforming multivariate GARCH models into univariate ones), reducing significantly the computational time and getting rid of any problem that may arise from complex data manipulations;
- It ensures a tight control of the amount of “noise” due to reducing the number of variables to fewer principal components. This may prove beneficial since it may result in more stable correlation estimates;
- Such method produces volatilities and correlations for all variables in the system, including those for which direct GARCH estimation is computationally difficult.

GARCH forecasting techniques offer some key advantages, like flexibility and accuracy, for which practitioners find them effective and easy to use, especially in activities like back-office risk management and front-office trading systems. However, this may be put at risk if a feasible method that helps the manipulation of large covariance data matrices is not also implemented. Given the considerable difficulties in data estimation that may arise when dealing with such large GARCH covariance matrices, and given the need for using mean-reverting covariance forecasts in value-at-risk-models, PC-GARCH contribution is notable. Designed to capture variability of a returns sample by few orthogonal casual factors, and assigning the rest of variation to “noise” factors, the use of principal components analysis permits the transformation of optimization procedures into univariate time series. This enables reduction in computational density, as the whole matrix of variances and covariances can be derived out of simple linear transformations of factor variances. Used in several real-world settings, in no case has the PC-GARCH been found defective. Its superiority has been found in all cases, starting from bivariate or trivariate settings with hundreds of variables, up to multivariate ones dealing with several thousands of time series.

The current paper could be extended by estimating different univariate and multivariate GARCH models with Bayesian statistics techniques. Further research could also include extensions to other multivariate GARCH models, like VEC, BEKK, CCC, TVC and DCC models, and the employment of new developments like semi-parametric estimation, more flexible DCC and factor models, finite mixtures of GARCH models, and incorporation of high frequency data in multivariate modeling.

Acknowledgements

This paper is supported by the Sectorial Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number SOP HRD/89/1.5/S/62988

References

Alexander, C. O., 1998. Volatility and correlation: Methods, models and applications. *Risk management and analysis: Measuring and modeling financial risk*. Wileys.

Alexander, C. O., 2000. *A primer on the Orthogonal GARCH model.* Available at:< www.ismacentre.rdg.ac.uk/pdf/orthogonal.pdf>[Accessed on 15 March 2009].

Andersen, T. G. and Bollerslev, T., 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts, *International Economic Review*, 39(4), pp. 885-905.

Andersen, T. G., Bollerslev, T., Christoffersen, P. F. and Diebold, F. X., 2005. Volatility Forecasting, *PIER (Penn Institute for Economic Research) Working Paper 05-01*, pp. 1-110.

Andersen, T. G., Bollerslev, T. and Lange, S., 1999. Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon, *Journal of Empirical Finance,* 6(5), pp. 457-477.

Bera, A. K. and Higgins, M. L., 1993. ARCH Models: Properties, Estimation and Testing, *Journal of Economic Surveys*, 7(4), pp. 305-62.

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

Bollerslev, T., Chou, R. Y. and Kroner, K. F., 1992. ARCH Modeling in Finance, *Journal of Econometrics*, 52, pp. 5-59.

Bollerslev, T., Engle, R. F. and Nelson, D.B., 1994. ARCH Models. *Handbook of Econometrics.* Amsterdam: Elsevier Science B. V.

Burns, P.,2005.*Multivariate GARCH with only univariate estimation*. Available at:<www.burns-stat.com>. [Accessed 20 June 2009].

Cumby, R., Figlewski, S. and Hasbrouck, J., 1993. Forecasting volatility and correlations with EGARCH models, *Journal of Derivatives*, 1(Winter), pp. 51-63.

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

Figlewski, S.,1997. Forecasting volatility, *Financial Markets, Institutions and Instruments,* 6(2), pp. 1-88.

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

Jacquier, E., Polson, N. G. and Rossi, P. E., 1994. Bayesian Analysis of Stochastic Volatility Models, *Journal of Business and Economic Statistics*, 12(4), pp. 371-89.

Jorion, P., 1995. Predicting volatility in the foreign exchange market, *Journal of Finance*, 50(2), pp.507-528.

Matei, M., 2009. Assessing volatility forecasting models: Why GARCH models take the lead, *Romanian Journal for Economic Forecasting*, 10 (4).

Matei, M., 2009. Risk analysis in the evaluation of the international investment opportunities. Advances in modeling and forecasting volatility for risk assessment purposes, *Working Papers of the Macroeconomic Modeling Seminar*, Bucharest: Institute for Economic Forecasting, Romanian Academy. Available at:<http://www.ipe.ro/wpiecf.htm>

Matei, M., 2011. Non-Linear volatility modeling of economic and financial time series using high frequency data, *Romanian Journal for Economic Forecasting*, 14 (2), pp. 116-141.

Melino, A. and Turnbull, S.M., 1990. Pricing foreign currency options with stochastic volatility, *Journal of Econometrics*, 45, pp. 239-265.

Nelson, D. B.,1991. Conditional heteroskedasticity in asset returns: A new approach, *Econometrica*, 59, pp. 347-370.

Nicholls, D. F. and Quinn, B. G., 1982. *Random Coefficient Autoregressive Model: An introduction.* New York: Springer-Verlag.

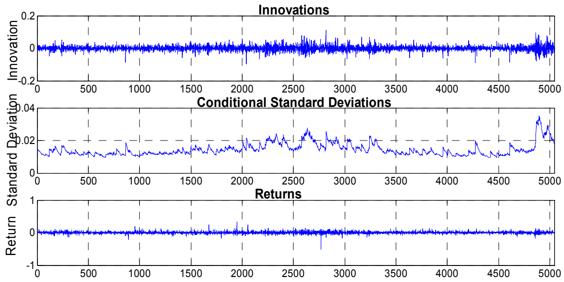
Taylor, S. J., 1994. Modeling stochastic volatility: A review and comparative study, *Mathematical Finance*, 4(2), pp. 183-04.

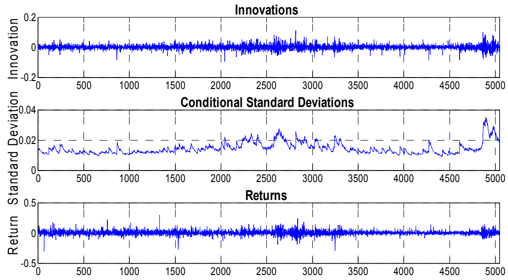
Tsay, R. S.,1987. Conditional heteroskedastic time series models, *Journal of American Statistical Association*, 82, pp. 590-604.

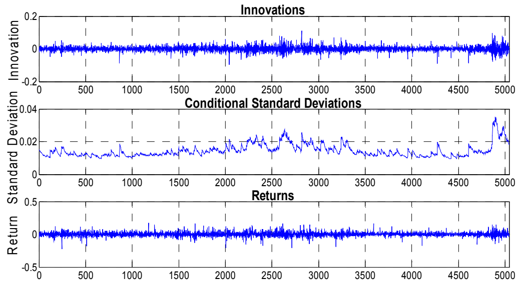
van der Weide, R.,2002. GO-GARCH: A multivariate generalized Orthogonal GARCH model, *Journal of Applied Econometrics*, 17, pp. 549-64.

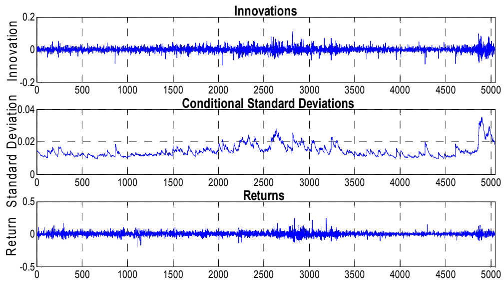
Appendix

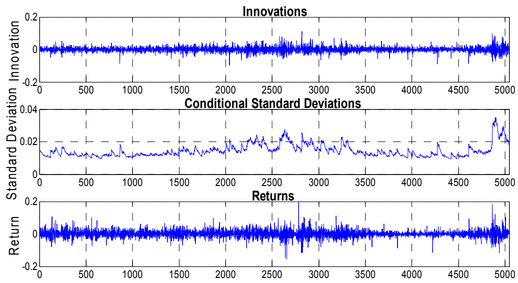
Figures 1 to 7: GARCH decomposition of Adobe, Apple, Autodesk, Cisco, Dell, Microsoft and 3M daily returns over the sample period. Data source: Datastream.

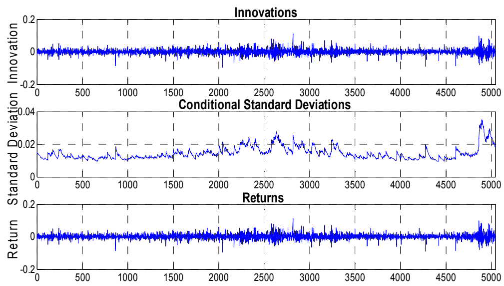


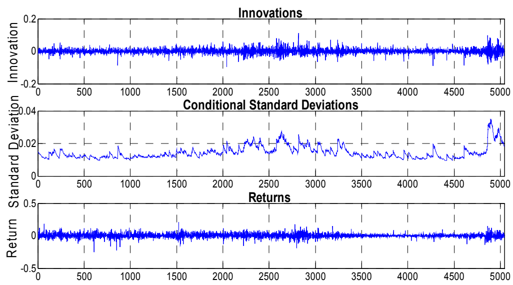












Figures 8 to 14: GARCH decomposition of Adobe, Apple, Autodesk, Cisco, Dell, Microsoft and 3M standardized residuals over the sample period. Data source: Datastream.