International Journal of Forecasting 28 (2012) 70–93

Contents lists available at SciVerse ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

A conditionally heteroskedastic independent factor model with an application to financial stock returns

Antonio García-Ferrer[a,*], Ester González-Prieto[b], Daniel Peña[b]

[a] Dpto. de Análisis Económico: Economía Cuantitativa, Universidad Autónoma de Madrid, Spain

[b] Dpto. de Estadística, Universidad Carlos III de Madrid, Spain

ARTICLE INFO

Keywords:
ICA
Multivariate GARCH
Factor models
Forecasting volatility

ABSTRACT

We propose a new conditionally heteroskedastic factor model, the GICA-GARCH model, which combines independent component analysis (ICA) and multivariate GARCH (MGARCH) models. This model assumes that the data are generated by a set of underlying independent components (ICs) that capture the co-movements among the observations, which are assumed to be conditionally heteroskedastic. The GICA-GARCH model separates the estimation of the ICs from their fitting with a univariate ARMA-GARCH model. Here, we will use two ICA approaches to find the ICs: the first estimates the components, maximizing their non-Gaussianity, while the second exploits the temporal structure of the data. After estimating and identifying the common ICs, we fit a univariate GARCH model to each of them in order to estimate their univariate conditional variances. The GICA-GARCH model then provides a new framework for modelling the multivariate conditional heteroskedasticity in which we can explain and forecast the conditional covariances of the observations by modelling the univariate conditional variances of a few common ICs. We report some simulation experiments to show the ability of ICA to discover leading factors in a multivariate vector of financial data. Finally, we present an empirical application to the Madrid stock market, where we evaluate the forecasting performances of the GICA-GARCH and two additional factor GARCH models: the orthogonal GARCH and the conditionally uncorrelated components GARCH.

1. Introduction

Since Engle (1982) introduced the ARCH model and Bollerslev (1986) generalized it to the GARCH representation, the interest in modelling volatilities has grown considerably. In multivariate time series, researchers are interested in understanding not only the co-movements of the volatilities of financial assets, but also the co-movements of financial returns. For these purposes, a multivariate modelling approach is required. Multivariate GARCH (MGARCH) models should be able to explain the structure of the covariance matrix of large financial datasets, and also represent the dynamics of their conditional variances and covariances. Depending on the parametrization of the conditional covariance matrix, different specifications for MGARCH models have been proposed in the literature (see for example the survey by Bauwens, Laurent, & Rombouts, 2006). Two popular MGARCH specifications are the VEC model (Bollerslev, Engle, & Wooldridge, 1988), which is an extension of the univariate GARCH model (see Engle, Granger, & Kraft, 1984, for an ARCH version), and the BEKK model (Engle & Kroner, 1995), which can be seen as a restricted version of the VEC model. However, the number of parameters requiring es-

* Corresponding author. Tel.: +34 914974811.
E-mail addresses: antonio.garcia@uam.es (A. García-Ferrer), ester.gonzalez@uc3m.es (E. González-Prieto), daniel.pena@uc3m.es (D. Peña).

0169-2070/$ – see front matter
doi:10.1016/j.ijforecast.2011.02.010

timation in most of these developments can be very large, and the restrictions for guaranteeing the positive definiteness of the conditional covariance matrix are difficult to implement.

Factor models are an alternative for achieving dimensionality reduction in large datasets. They are based on the idea of the existence of a few underlying components that are the driving forces for large datasets. In finance, there are many empirical applications for factor models with conditional heteroskedasticity. For example, asset pricing models usually assume that the dynamics of the prices of different assets can be explained by a small number of underlying dynamic factors that are conditionally heteroskedastic.

There are two branches of literature relating to factor GARCH models, depending on whether the factor structure refers to the conditional or unconditional distribution of the data. On the one hand, the FACTOR-ARCH model (Engle, 1987) exploits the conditional distribution of the data by using common factors to model the conditional covariance matrix of the observations. The factors, which follow GARCH-type processes, are given by the linear combinations of the data that summarize the co-movements in their conditional variances. Some applications of the FACTOR-ARCH parametrization include: modelling the term structure of interest rates (Engle, Ng, & Rothschild, 1990; Ng, Engle, & Rothschild, 1992), investigating whether international stock markets have the same volatility processes (Engle & Susmel, 1993), and modelling the common persistence in the conditional variance (Bollerslev & Engle, 1993). Particular models which are related to the FACTOR-ARCH model are the orthogonal models. They assume that the data conditional covariance matrix is generated by some underlying factors that follow univariate GARCH processes. Examples of this class of models are the orthogonal GARCH (O-GARCH) model (Alexander, 2001), the generalized orthogonal GARCH (GO-GARCH) model (van der Weide, 2002), the generalized orthogonal factor GARCH (GOF-GARCH) model (Lanne & Saikkonen, 2007), and the conditional uncorrelated component GARCH (CUC-GARCH) model (Fan, Wang, & Yao, 2008). In addition, the full factor GARCH (FF-GARCH) model proposed by Vrontos, Dellaportas, and Politis (2003) and extended by Diamantopoulos and Vrontos (2010) to allow for multivariate Student-t distributions is also nested in the FACTOR-ARCH approach. On the other hand, the latent factor ARCH model (Diebold & Nerlove, 1989) applies the factor structure in the unconditional distribution of the data, and can be seen as a traditional latent factor model where the factors display strong evidence of an ARCH structure. In this model, the factors represent the co-movements among the observations, and it is assumed that the commonalities in the volatilities among observations are due to the ARCH effect of such common latent factors. Harvey, Ruiz, and Sentana (1992) extended the Diebold and Nerlove model to allow for general dynamics in the mean, and provided a modified version of the Kalman filter for unobserved components models with GARCH disturbances. King, Sentana, and Wadhwani (1994), who consider a multifactor model for aggregate stock returns, and Doz and Renault (2004), who present a conditionally heteroskedastic factor model where the common factors represent conditionally orthogonal influences, also extend the Diebold and Nerlove model. The dynamic factor GARCH (DF-GARCH) model (Alessi, Barigozzi, & Capasso, 2006) is another example of this branch of the literature. It can be seen as a generalized dynamic factor model where both the dynamic common factors and the idiosyncratic components are conditionally heteroskedastic.

In this paper we propose a multivariate conditionally heteroskedastic factor model, known as the GICA-GARCH model. The GICA-GARCH model is a new method for explaining the conditional covariance matrix of large datasets using a small number of factors with GARCH effects. It is based on the intuition that financial markets are driven by a few latent factors that represent the co-movements of financial variables. These factors are estimated by independent component analysis (ICA). ICA can be seen as a factor model (Hyvärinen & Kano, 2003) where the unobserved components are non-Gaussian and mutually independent. Previous researchers, such as Back and Weigend (1997), Cha and Chan (2000), Kiviluoto and Oja (1998), Mălăroiu, Kiviluoto, and Oja (2000) among others, have applied ICA to financial data. Furthermore, ICA can be considered as a generalization of principal component analysis (PCA) (Hyvärinen, Karhunen, & Oja, 2001), and seems to be, a priori, more suitable than PCA for explaining the non-Gaussian behavior of financial data (Wu & Yu, 2005).

The GICA-GARCH model assumes that observations are generated by a set of underlying factors that are independent and conditionally heteroskedastic. Once the ICs have been estimated, they are sorted in terms of the total explained variability, in order to choose the few components which represent the co-movements of financial variables. The GICA-GARCH model then assumes a factor structure in the unconditional distribution of the data. Furthermore, due to the statistical assumption on the ICs, the GICA-GARCH model fits a univariate ARMA-GARCH model to each of them, and the conditional covariance matrix of the ICs is then allowed to be diagonal. Thus, the GICA-GARCH model transforms the complexity associated with the estimation of a multivariate ARMA-GARCH model into the estimation of a small number of univariate ARMA-GARCH models, and approximates the conditional covariance matrix of the data by a linear combination of the conditional variances of a few ICs. The GICA-GARCH model therefore also applies the factor structure to the conditional distribution of the data.

The rest of this paper is organized as follows. In Section 2 we present the ICA model, describe the three ICA algorithms used to estimate the unobserved components and explain a procedure for sorting the ICA components in terms of their explained variability. Furthermore, the relationship between ICA and the dynamic factor model (DFM) is analyzed. In Section 3 we introduce the GICA-GARCH model for explaining and forecasting the conditional covariance matrix of a vector of stock returns from the univariate conditional variances of a small number of components. Furthermore, we analyze the relationship between the GICA-GARCH model and other factor GARCH models proposed in the literature. Next, Section 4 presents some simulation experiments that illustrate the ability of the GICA-GARCH model to estimate the underlying components of conditionally heteroskedastic data. An empirical application to a real-time dataset is shown in Section 5. Finally, Section 6 gives some concluding remarks.

2. The ICA model

In this section we introduce the concept of ICA. First, we present the basic ICA model according to the formal definition given by Common (1994). Then, we briefly describe the three algorithms which we use to estimate the ICA components. As the definition of ICA implies that there is no ordering of the ICs, a procedure for weighting and sorting them is explained next. Finally, we formulate the ICA model as a particular DFM and analyze the relationship between the two models.

2.1. Definition of ICA

ICA assumes that the observed data are generated by a set of unobserved components that are independent. Let $\mathbf{x}_t = (x_{1t}, x_{2t}, \ldots, x_{mt})'$ be the m-dimensional vector of stationary time series, with $E[\mathbf{x}_t] = \mathbf{0}$ and $E\left[\mathbf{x}_t\mathbf{x}_t'\right] = \Gamma_{\mathbf{x}}(0)$ being positive definite. It is assumed that $\mathbf{x}_t$ is generated by a linear combination of $r\,(r \leq m)$ latent factors. That is,

$$\mathbf{x}_t = \mathbf{A}\mathbf{s}_t, \quad t = 1, 2, \ldots, T, \tag{1}$$

where $\mathbf{A}$ is an unknown $m \times r$ full rank matrix, with elements a_{ij} that represent the effect of s_{jt} on x_{it}, for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, r$, and $\mathbf{s}_t = (s_{1t}, s_{2t}, \ldots, s_{rt})'$ is the vector of unobserved factors, which are called independent components (ICs). It is assumed that $E[\mathbf{s}_t] = \mathbf{0}$, $\mathbf{\Gamma}_{\mathbf{s}}(0) = E\left[\mathbf{s}_t\mathbf{s}_t'\right] = \mathbf{I}_r$, and that the components of $\mathbf{s}_t$ are statistically independent. Let $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T)$ be the observed multivariate time series. The problem is to estimate both $\mathbf{A}$ and $\mathbf{s}_t$ from only $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T)$. That is, ICA looks for an $r \times m$ matrix, $\mathbf{W}$, such that the components given by

$$\widehat{\mathbf{s}}_t = \mathbf{W}\mathbf{x}_t, \quad t = 1, 2, \ldots, T, \tag{2}$$

are as independent as possible. However, previous assumptions are not sufficient to enable us to estimate $\mathbf{A}$ and $\mathbf{s}_t$ uniquely, and it is required that no more than one IC be normally distributed. From Eq. (1) we have:

$$\begin{aligned} \mathbf{\Gamma}_{\mathbf{x}}(0) &= E[\mathbf{x}_t\mathbf{x}_t'] = \mathbf{A}\mathbf{A}', \\ \mathbf{\Gamma}_{\mathbf{x}}(\tau) &= E[\mathbf{x}_t\mathbf{x}_{t-\tau}'] = \mathbf{A}\mathbf{\Gamma}_{\mathbf{s}}(\tau)\,\mathbf{A}', \quad \tau \geq 1. \end{aligned} \tag{3}$$

All of the dynamic structure of the data therefore comes through the unobserved components, and if they are uncorrelated, then $\mathbf{\Gamma}_{\mathbf{s}}(\tau) = E[\mathbf{s}_t\mathbf{s}_{t-\tau}']$ is a diagonal matrix for all $\tau \geq 1$.

Note that, in spite of previous assumptions, ICA cannot determine either the sign or the order of the ICs. In the following, we focus on the most basic form of ICA, which considers that the number of observed variables is equal to the number of unobserved factors, i.e., $m = r$.

2.2. Procedures for estimating the ICs

Both ICA and PCA obtain the latent factors as linear combinations of the data. However, their aims are slightly different. On the one hand, PCA tries to get uncorrelated factors, and, for this purpose, it requires the matrix $\mathbf{W}$ to be such that $\mathbf{W}\mathbf{W}' = \mathbf{I}$, and the rows of $\mathbf{W}$ are the projection vectors that maximize the variance of the estimated unobserved factors, $\widehat{\mathbf{s}}_t$. On the other hand, ICA tries to obtain independent factors, and the most commonly used methods for estimating the ICs impose the restriction that the rows of $\mathbf{W}$ are the directions that maximize the independence of $\widehat{\mathbf{s}}_t$.

Three main ICA algorithms have been proposed: JADE, FastICA and SOBI. JADE (Cardoso & Souloumiac, 1993) and FastICA (Hyvärinen, 1999; Hyvärinen & Oja, 1997) are based on the non-Gaussianity of the ICs, while SOBI (Belouchrani, Abed Meraim, Cardoso, & Moulines, 1997) is based on the temporal uncorrelatedness between components. Before any of these algorithms are applied, it is useful to standardize the data. Thus, we search for a linear transformation of $\mathbf{x}_t$, $\mathbf{z}_t = \mathbf{M}\mathbf{x}_t$, where $\mathbf{M}$ is an $m \times m$ matrix such that the m-dimensional vector $\mathbf{z}_t$ has an identity covariance matrix. This multivariate standardization is carried out as follows. From Eq. (3) we have

$$\mathbf{\Gamma}_{\mathbf{x}}(0) = \mathbf{A}\mathbf{A}' = \mathbf{E}\mathbf{D}\mathbf{E}', \tag{4}$$

where $\mathbf{E}_{m\times m}$ is the orthogonal matrix of eigenvectors, and $\mathbf{D}_{m\times m}$ the diagonal matrix of eigenvalues. Then $\mathbf{M} = \mathbf{D}^{-1/2}\mathbf{E}'$, and Eq. (1) in terms of $\mathbf{z}_t$, calling $\mathbf{U} = \mathbf{M}\mathbf{A}$, is

$$\mathbf{z}_t = \mathbf{U}\mathbf{s}_t, \tag{5}$$

where $\mathbf{U}'\mathbf{U} = \mathbf{I}_r$ and $\mathbf{U}\mathbf{U}' = \mathbf{I}_m$. However, in order to avoid identification problems, the ICA algorithms look for an $m \times m$ orthogonal matrix $\mathbf{U}$, as the multivariate standardization of the original data guarantees the orthogonality of the loading matrix.

2.2.1. Joint approximate diagonalization of eigen-matrices: JADE

JADE (Cardoso & Souloumiac, 1993) estimates the ICs by maximizing their non-Gaussianity. After whitening the observed data, JADE looks for a matrix $\mathbf{U}'$ such that the components given by

$$\widehat{\mathbf{s}}_t^J = \mathbf{U}'\mathbf{z}_t \tag{6}$$

are maximally non-Gaussian distributed. Note that under the non-Gaussianity assumption, the information provided by the covariance matrix of the data, $\Gamma_{\mathbf{z}}(0) = \mathbf{I}$, is not sufficient to compute Eq. (6), and higher-order information is needed. Cardoso and Souloumiac (1993) use cumulants, which are the coefficients of the Taylor series expansion of the logarithm of the characteristic function. In practice, it is enough to take into account fourth-order cumulants, which are defined as

$$\begin{aligned} &\mathrm{cum}_4(z_{it}, z_{jt}, z_{ht}, z_{lt}) \\ &\quad = E\{z_{it}z_{jt}z_{ht}z_{lt}\} - E\{z_{it}z_{jt}\}E\{z_{ht}z_{lt}\} \\ &\qquad - E\{z_{it}z_{ht}\}E\{z_{jt}z_{lt}\} - E\{z_{it}z_{lt}\}E\{z_{jt}z_{ht}\}, \end{aligned} \tag{7}$$

and the fourth-order cumulant tensor associated with $\mathbf{z}_t$ is a $m \times m$ matrix which is given by

$$[Q_{\mathbf{z}}(\mathbf{Q})]_{ij} = \sum_{k,l=1}^{r} \mathrm{cum}_4\left(z_{it}, z_{jt}, z_{kt}, z_{lt}\right) q_{kl},$$

where $\mathbf{Q} = (q_{kl})_{k,l=1}^{m}$ is an arbitrary $m \times m$ matrix, and $\mathrm{cum}_4\left(z_{it}, z_{jt}, z_{kt}, z_{lt}\right)$ is as in Eq. (7). It is easy to see that random vectors are independent if all of their cross-cumulants of order higher than two are equal to zero. In particular, $\widehat{\mathbf{s}}_t^J$ will be independent if its associated fourth order cumulant tensor, $Q_{\widehat{\mathbf{s}}_t^J}(\cdot)$, is diagonal. Cardoso and Souloumiac (1993) show that, given a set of $m \times m$ matrices $\Im = \left\{\mathbf{Q}_1, \ldots, \mathbf{Q}_q\right\}$, there exists an orthogonal transformation $\mathbf{V}$ such that the matrices $\left\{\mathbf{V}' Q_{\mathbf{z}}(\mathbf{Q}_i)\mathbf{V}\right\}_{\mathbf{Q}_i \in \Im}$ are approximately diagonal. Then we can choose $\mathbf{V} = \mathbf{U}'$ and estimate the latent factors using Eq. (6). JADE uses an iterative process of Jacobi rotations to solve the joint diagonalization of several fourth-order cumulant matrices. It is a very efficient algorithm in low dimensional problems, but when the dimension increases, it has a high computational cost.

2.2.2. Fast fixed-point algorithm: FastICA

FastICA is a fixed-point algorithm which was proposed by Hyvärinen and Oja (1997). It estimates

$$\widehat{\mathbf{s}}_t^F = \mathbf{U}'\mathbf{z}_t \tag{8}$$

by maximizing their univariate kurtosis. Thus, FastICA searches for the directions of projection that maximize the absolute value of the kurtosis of $\widehat{\mathbf{s}}_t^F$. As the kurtosis is very sensitive to outliers, FastICA is not a robust algorithm. Hyvärinen (1999) proposes a more robust version of FastICA using an approximation of negentropy instead of kurtosis to measure the non-Gaussianity of the ICs. Negentropy is the normalized version of the entropy, given by:

$$J\left(\widehat{\mathbf{s}}_t^F\right) = H\left(\widehat{\mathbf{s}}_t^G\right) - H\left(\widehat{\mathbf{s}}_t^F\right),$$

where $\widehat{\mathbf{s}}_t^G$ is a Gaussian vector of the same correlation matrix as $\widehat{\mathbf{s}}_t^F$, and $H(\cdot)$ is the entropy of a random vector defined as $H\left(\widehat{\mathbf{s}}_t^F\right) = -E[\log p_{\widehat{\mathbf{s}}_t^F}(\xi)]$, where $p_{\widehat{\mathbf{s}}_t^F}(\cdot)$ is the density function of $\widehat{\mathbf{s}}_t^F$. Negentropy is a good index for non-Gaussianity because it is always non-negative and it is zero *iff* the variable is Gaussian distributed. Therefore, the ICs, given by Eq. (8), are estimated as the projections of the data in the directions which maximize the negentropy of $\widehat{\mathbf{s}}_t^F$. The main advantage of FastICA is that it converges in a small number of iterations.

2.2.3. Second-order blind identification: SOBI

Belouchrani et al. (1997) extended the previous work of Tong, Liu, Soon, Huang, and Liu (1990), and proposed the SOBI algorithm. SOBI requires that the ICs, given by

$$\widehat{\mathbf{s}}_t^S = \mathbf{U}'\mathbf{z}_t, \tag{9}$$

will be mutually uncorrelated for a set of time lags. That is, the matrix $\mathbf{U}'$ is obtained so that a set of $\mathbf{K}$ time delayed covariance matrices of $\widehat{\mathbf{s}}_t^S$,

$$\mathit{\Gamma}_{\mathbf{s}}(\tau) = E\left\{\widehat{\mathbf{s}}_t^S \widehat{\mathbf{s}}_{t-\tau}^{S\prime}\right\}, \quad \tau \in J = \{1, \ldots, K\}, \tag{10}$$

should be diagonal. Thus, SOBI searches for an orthogonal transformation that jointly diagonalizes Eq. (10). This algorithm also applies whitening as a preprocessing procedure, and the covariance structure of the whitened data model (Eq. (5)) is given by:

$$\mathbf{\Gamma}_{\mathbf{z}}(\tau) = \mathbf{U}\mathbf{\Gamma}_{\mathbf{s}}(\tau)\mathbf{U}', \quad \tau \geq 1, \tag{11}$$

where $\mathbf{U}$ is an orthogonal matrix. Therefore,

$$\mathbf{\Gamma}_{\mathbf{s}}(\tau) = \mathbf{U}'\mathbf{\Gamma}_{\mathbf{z}}(\tau)\mathbf{U}, \quad \tau \geq 1. \tag{12}$$

Thus, SOBI searches for an orthogonal transformation that will be the joint diagonalizer of the set of time delayed covariance matrices, $\left\{\mathbf{\Gamma}_{\mathbf{s}}\left(\tau_q\right)\right\}_{\tau_q \in J}$. The optimization problem is to minimize

$$F(\mathbf{U}) = \sum_{\tau_q \in J} \mathbf{off}\left(\mathbf{U}'\mathbf{\Gamma}_{\mathbf{z}}(\tau)\mathbf{U}\right),$$

where '**off**' is a measure of the non-diagonality of a matrix, which is defined as the sum of the squares of their off-diagonal elements. SOBI solves this problem using Jacobi rotation techniques. Belouchrani et al. (1997) show that this problem has a unique solution: if there exist two different ICs that have different autocovariances for at least one time-lag, then the joint diagonalizer, $\mathbf{U}$, exists and is unique. That is, if, for all $1 \leq i \neq j \leq r$, there exists any $q = 1, \ldots, K$ such that $\gamma_{\mathbf{s}_i}\left(\tau_q\right) \neq \gamma_{\mathbf{s}_j}\left(\tau_q\right)$, then the components of $\widehat{\mathbf{s}}_t^S$ can be separated; they are unique and lagged uncorrelated. Note that SOBI cannot obtain the ICs if they have identical autocovariances for the lags considered.

2.2.4. Weighting the ICs

After estimating the components, we should decide which of them are most important for explaining the underlying structure of the data. Note that the PCs are sorted in terms of variability, but the ICs are undetermined with respect to the order. Following Back and Weigend (1997), we will sort the ICs in terms of their explained variability. According to model (1), the ith observed variable is given by $x_{it} = \sum_{j=1}^{m} a_{ij}s_{jt}$, and its variance is

$$var(x_{it}) = \sum_{j=1}^{m} a_{ij}^2, \quad i = 1, \ldots, m. \tag{13}$$

For each x_{it}, with $i = 1, \ldots, m$, Back and Weigend (1997) define the weighted ICs in terms of the elements of the ith row of $\mathbf{A}$ as $\mathbf{s}_t^{w(i)} = \mathrm{diag}\left(a_{i1}, a_{i2}, \ldots, a_{im}\right)\mathbf{s}_t$. That is, for each x_{it}, the jth weighted IC is given by $s_{jt}^{w(i)} = a_{ij}s_{jt}$ for $j = 1, \ldots, m$, and its variance is

$$\mathrm{var}\left(s_{jt}^{w(i)}\right) = a_{ij}^2, \quad i, j = 1, \ldots, m. \tag{14}$$

Therefore, from Eqs. (13) and (14), the variance of x_{it} which is explained by $s_{jt}^{w(i)}$ is computed as:

$$\nu_j^i = \frac{a_{ij}^2}{\sum_{j=1}^{m} a_{ij}^2}, \quad i, j = 1, \ldots, m, \tag{15}$$

and the total variance of $\mathbf{x}_t$ explained by the jth IC is given by:

$$\vartheta_j = \frac{\sum_{i=1}^{m} \nu_j^i}{\sum_{j=1}^{m}\left(\sum_{i=1}^{m} \nu_j^i\right)}, \quad j = 1, \ldots, m.$$

Thus, after we have obtained $\{\vartheta_1, \vartheta_2, \ldots, \vartheta_m\}$, we know how much of the total variance is explained by each IC, and can sort them in terms of the variability. Thus, the most important ICs will be those which explain the maximum variance of $\mathbf{x}_t$.

2.3. ICA and the dynamic factor model

Suppose that there are r non-Gaussian components in the basic ICA model, $\mathbf{s}_t^{(1)} = (s_{1t}, \ldots, s_{rt})'$ with $r < m$, representing the common dynamic of the time series, but that the other $m - r$ components, $\mathbf{s}_t^{(2)} = (s_{r+1t}, \ldots, s_{mt})'$, are Gaussian. We can then split the matrix $\mathbf{A} = [\mathbf{A}_1 : \mathbf{A}_2]$ accordingly, and write

$$\mathbf{x}_t = \mathbf{A}_1\mathbf{s}_t^{(1)} + \mathbf{A}_2\mathbf{s}_t^{(2)}. \quad (16)$$

Calling $\mathbf{n}_t = \mathbf{A}_2\mathbf{s}_t^{(2)}$ to the vector of Gaussian noise, we have $\mathbf{x}_t = \mathbf{A}_1\mathbf{s}_t^{(1)} + \mathbf{n}_t$, which is similar to the DFM studied by Peña and Box (1987) and generalized by Peña and Poncela (2006). However, there are two main differences between these models. First, in the factor model, the r common factors $\mathbf{s}_t^{(1)}$ are assumed to be Gaussian and linear, whereas here they are non-Gaussian. Second, in the standard factor model, the covariance matrix of the noise is of full rank, whereas here it will have a rank equal to $m - r$. This last constraint can be relaxed by assuming that the ICA model has been contaminated by some Gaussian error model, as in $\mathbf{x}_t = \mathbf{A}\mathbf{s}_t + \mathbf{u}_t$, where $\mathbf{u}_t$ is Gaussian. Note that the latent factors of the DFM can be estimated consistently by PCA when both the number of series and the sample size (m and T respectively) go to infinity (see for example Stock & Watson, 2002).

3. The GICA-GARCH model

This section presents the GICA-GARCH model as a new multivariate conditionally heteroskedastic factor model. From now on, let $\mathbf{x}_t = (x_{1t}, x_{2t}, \ldots, x_{mt})'$ be the vector of m financial time series. First, we introduce the GICA-GARCH model, give its mathematical formulation, and describe the structure of the ICA components. Next, we explain how this model can be used to forecast the conditional variances of a vector of financial data from the univariate conditional variances of a set of common ICs. Finally, we relate the GICA-GARCH model to the factor GARCH models.

3.1. The model

Let us assume that $\mathbf{x}_t$ is a linear combination of a set of independent factors given by equation (1). Because series of stock returns are characterized by the presence of clusters of volatility, some of the underlying factors will follow conditionally heteroskedastic processes. In the literature, GARCH models are the most popular specifications for modelling the conditional variance of the stock returns. In addition, from empirical finance, it is common to admit that the stock returns could exhibit low order temporal dependence on the conditional mean, which can be explained by an ARMA model. Therefore, as there could also be temporal structure on the conditional mean of the latent factors, it seems reasonable to propose an ARMA-GARCH specification for modeling the underlying factor given by Eq. (1). Then, we assume that the vector of unobserved components, $\mathbf{s}_t$, follows an r-dimensional ARMA(p, q) model with GARCH $\left(p', q'\right)$ disturbances:

$$\mathbf{s}_t = \sum_{i=1}^{p} \mathbf{\Phi}_i \mathbf{s}_{t-i} + \sum_{l=0}^{q} \mathbf{\Theta}_l \mathbf{e}_{t-l}, \quad (17)$$

where $\mathbf{\Phi}_i = \operatorname{diag}\left(\phi_i^{(1)}, \ldots, \phi_i^{(r)}\right)$ with $|\phi_i^{(j)}| < 1 \forall j$; $\mathbf{\Theta}_l = \operatorname{diag}\left(\theta_l^{(1)}, \ldots, \theta_l^{(r)}\right)$ with $\mathbf{\Theta}_0 = \mathbf{I}_r$ and $|\theta_l^{(j)}| < 1 \forall j$; and $\mathbf{e}_t$ is an r-dimensional vector of conditionally heteroskedastic errors given by:

$$\mathbf{e}_t = \mathbf{H}_t^{1/2} \boldsymbol{\varepsilon}_t, \quad (18)$$

where $\boldsymbol{\varepsilon}_t \sim$ i.i.d. $(\mathbf{0}, \mathbf{I}_r)$ and $\mathbf{H}_t^{1/2} = \operatorname{diag}(\sqrt{h_{jt}})$ is an $r \times r$ positive definite diagonal matrix such that

$$h_{jt} = \alpha_0^{(j)} + \sum_{i=1}^{p'} \alpha_i^{(j)} e_{jt-i}^2 + \sum_{l=1}^{q'} \beta_l^{(j)} h_{jt-l}, \quad \text{for } j = 1, \ldots, r, \quad (19)$$

where h_{jt} is a stationary process, is independent of ε_{jt}, and represents the conditional variance of the jth IC: $h_{jt} = V\left(e_{jt}|\mathbf{I}_{t-1}\right) = V\left(s_{jt}|\mathbf{I}_{t-1}\right)$, where $\mathbf{I}_{t-1}$ is the past information available up to time $t - 1$. In order to ensure a positive $h_{jt} > 0$, $\forall j$, it is assumed that $\alpha_0^{(j)} > 0$, $\alpha_i^{(j)} \geq 0$, $\beta_i^{(j)} \geq 0$, and $\sum_{i=1}^{\max(p',q')}\left(\alpha_i^{(j)} + \beta_i^{(j)}\right) < 1$ (see Bollerslev, 1986).

Focusing on forecasting the volatility of the observed financial data, from Eq. (1), we know that the conditional covariance matrix of $\mathbf{x}_t$ is:

$$\mathbf{\Omega}_t = V\left(\mathbf{x}_t|\mathbf{I}_{t-1}\right) = \mathbf{A}\mathbf{H}_\mathbf{t}\mathbf{A}', \quad (20)$$

where $\mathbf{H}_t = \operatorname{diag}(h_{1t}, \ldots, h_{rt})$ is the $r \times r$ conditional covariance matrix of $\mathbf{s}_t$ at time t. In order to guarantee the diagonality of $\mathbf{H}_t$, we should assume that the conditional correlations of the ICs are zero. This assumption allows us to achieve our purpose: explaining and forecasting the conditional covariances of the observations from the univariate conditional variances of the set of conditionally heteroskedastic components that represents the co-movements of the stock returns. In the GICA-GARCH model, it is assumed that the number of conditionally heteroskedastic common ICs is small relative to the dimension of the dataset. Then, the GICA-GARCH reduces the number of parameters to be estimated considerably, but at the cost of obtaining conditional covariances matrices with a reduced rank. Furthermore, note that the GARCH structure of $\mathbf{x}_t$ is ensured because each IC is generated by an independent GARCH process, and the linear combination of r independent GARCH processes will be a weak GARCH process (see Nijman & Sentana, 1996).

3.2. Fitting the model

The model is fitted in two steps. First, we use ICA to identify the underlying independent components and the loading matrix. Second, univariate GARCH models are fitted to the components. We describe these two steps in what follows. All of the previous ICA algorithms standardize the data as a preprocessing step, and solve the basic ICA model for the normalized data, which is given by Eq. (5). Thus, JADE, FastICA, and SOBI will estimate the orthogonal loading matrix and the m ICs, defined by Eqs. (6), (8) and (9), respectively. After estimating the model, we should choose the common ICs that we will take into account for forecasting the conditional variances of the financial variables. For this purpose, we weight the ICs according to the procedure explained in Section 2.2.4: we sort the ICs in terms of their explained total variability, and split the vector of ICs as $\mathbf{s}_t = [\mathbf{s}_t^{(1)}\mathbf{s}_t^{(2)}]$, where $\mathbf{s}_t^{(1)} = (s_{1t}, \ldots, s_{rt})'$ are the r ICs (with $r < m$) which we choose to represent the co-movements of the data, and $\mathbf{s}_t^{(2)} = (s_{r+1t}, \ldots, s_{mt})'$ are the $m - r$ ICs which we consider as noise. This splitting is done by testing whether the $m-r$ ICs are white noise. As an alternative, we can fit ARMA (p, q) models to $\mathbf{s}_t$ and $\mathbf{s}_t^2$ and check that the order selected using the BIC is ARMA(0, 0) in both cases. From now on, we focus on the r selected ICs that are conditionally heteroskedastic, and fit a univariate ARMA(p, q)-GARCH($p\prime, q\prime$) to each of them. According to the corresponding model, we estimate the univariate conditional variance of each IC and generate the conditional covariance matrix of $\mathbf{s}_t^{(1)}$, $\mathbf{H}_t$. Finally, we get the conditional covariance matrix of the observed data from Eq. (20), and its ith diagonal term, $\gamma_{it}^2 = \sum_{j=1}^r h_{jt} a_{ij}^2$, is the conditional variance of x_{it}, for $i = 1, 2, \ldots, m$.

Note that the performance of the GICA-GARCH model depends on the method used to estimate the ICs. In what follows, we will investigate the usefulness of the three algorithms presented in Section 2. Since they use different estimation principles (JADE and FastICA use non-Gaussianity, and SOBI uses dynamic uncorrelatedness) the performance of the algorithms is expected to depend on the features of the data. If the data have excess kurtosis and do not have a significant autocorrelation structure, FastICA and JADE would work better than SOBI. However, for data with large autocorrelation coefficients, SOBI may be the most appropriate algorithm for estimating the ICs.

3.3. The GICA-GARCH model and related factor GARCH models

In this section, we investigate the relationship between the GICA-GARCH model and other factor GARCH models such as the latent factor ARCH model (Diebold & Nerlove, 1989), the dynamic factor GARCH (DF-GARCH) model (Alessi et al., 2006), the factor GARCH model (Engle, 1987; Engle et al., 1990), and several orthogonal models.

The GICA-GARCH model assumes that the observations are given by a linear combination of a set of underlying components that are independent and conditionally heteroskedastic. Let us assume that r of these components, $\mathbf{s}_t^{(1)} = (s_{1t}, \ldots, s_{rt})'$, with $r < m$, explain the co-movements between the observations, and the other $m-r$ components, $\mathbf{s}_t^{(2)} = (s_{r+1t}, \ldots, s_{mt})'$, are the noisy ones. Splitting the matrix $\mathbf{A} = [\mathbf{A}_1 : \mathbf{A}_2]$ properly, the GICA-GARCH is given by

$$\mathbf{x}_t = \mathbf{A}_1\mathbf{s}_t^{(1)} + \mathbf{n}_t, \tag{21}$$

where $\mathbf{n}_t = \mathbf{A}_2\mathbf{s}_t^{(2)}$ is the noise vector. By assumption, both the common and noisy components are conditionally heteroskedastic and distributed as

$$\begin{pmatrix} \mathbf{s}_t^{(1)} \\ \mathbf{n}_t \end{pmatrix} \Bigg| \mathbf{I}_{t-1} \sim D\left\{ \begin{pmatrix} \boldsymbol{\mu}_t^{(1)} \\ \boldsymbol{\mu}_t^{(n)} \end{pmatrix}, \begin{pmatrix} \mathbf{H}_t & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma}_t \end{pmatrix} \right\}$$

where $\mathbf{H}_t$ is a $r \times r$ conditional covariance matrix of the vector of common factors, and $\boldsymbol{\Gamma}_t$ is a $m \times m$ conditional covariance matrix of the noise vector with rank$(\boldsymbol{\Gamma}_t) = m - r$. Note that the GICA-GARCH model assumes that the vector of common components and the noise vector are conditionally uncorrelated, and allows for the possibility that the common factors and the noise have a non-zero conditional mean (the GICA-GARCH model assumes that each IC could fit a univariate ARMA-GARCH model, see Eqs. (17)–(19)). Furthermore, due to the independence assumption on the underlying components, both $\mathbf{H}_t$ and $\boldsymbol{\Gamma}_t$ are diagonal matrices: $\mathbf{H}_t = \mathrm{diag}(h_{1t}, \ldots, h_{rt})$ and $\boldsymbol{\Gamma}_t = \mathrm{diag}(0, \ldots, 0, h_{r+1t}, \ldots, h_{mt})$. According to these assumptions, the GICA-GARCH model assumes a factor structure in both the unconditional distribution of the data,

$$\boldsymbol{\Gamma}_{\mathbf{x}}(0) = \mathbf{A}_1\boldsymbol{\Gamma}_{\mathbf{s}^{(1)}}(0)\mathbf{A}_1' + \boldsymbol{\Gamma}_{\mathbf{n}}(0), \tag{22}$$

and the conditional distribution

$$\boldsymbol{\Omega}_t = \mathbf{A}_1\mathbf{H}_t\mathbf{A}_1' + \boldsymbol{\Gamma}_t, \tag{23}$$

where $\mathbf{H}_t = \mathrm{diag}(h_{1t}, \ldots, h_{rt})$ and h_{jt} is the conditional variance of the jth component of $\mathbf{s}_t^{(1)}$ given by Eq. (19).

In practice, the GICA-GARCH model approximates the data conditional covariance matrix as

$$\boldsymbol{\Omega}_t = \mathbf{A}_1\mathbf{H}_t\mathbf{A}_1' = \sum_{i=1}^{r} \mathbf{a}_{(1)i}\mathbf{a}_{(1)i}' h_{it}, \tag{24}$$

with an accuracy that depends on the number of chosen common components, r, and where $\mathbf{a}_{(1)i} = (a_{1i}, \ldots, a_{mi})'$. Plugging Eq. (19) into Eq. (24), we have:

$$\boldsymbol{\Omega}_t = \sum_{i=1}^{r} \mathbf{a}_{(1)i}\mathbf{a}_{(1)i}' \left(\alpha_0^{(i)} + \sum_{l=1}^{p'} \alpha_l^{(i)} e_{it-l}^2 + \sum_{l=1}^{q'} \beta_l^{(i)} h_{it-l} \right), \quad \text{for } i = 1, \ldots, r, \tag{25}$$

where $e_{it} = s_{it}^{(1)} - \mu_{it}^{(1)}$ for $i = 1, \ldots, r$. Note that $s_{it}^{(1)} = \mathbf{w}_{(1)i}'\mathbf{x}_t$, where $\mathbf{w}_{(1)i}$ is the ith row vector of $\mathbf{W}_1$

($\mathbf{W}' = [\mathbf{W}_1 : \mathbf{W}_2]'$ is such that $\mathbf{AW}' = \mathbf{W}'\mathbf{A} = \mathbf{I}_m$). Then,

$$\mathbf{\Omega}_t = \sum_{i=1}^{r} \mathbf{a}_{(1)i}\mathbf{a}'_{(1)i}\left(\alpha_0^{(i)} + \sum_{l=1}^{p'} \alpha_l^{(i)}(\mathbf{w}'_{(1)i}\mathbf{x}_{t-l} - \mu^{(1)}_{it-l})^2 + \sum_{l=1}^{q'} \beta_l^{(i)}(\mathbf{w}'_{(1)i}\mathbf{\Omega}_{t-l}\mathbf{w}_{(1)i})\right), \quad \text{for } i = 1, \dots, r, \tag{26}$$

and it is clear that the data conditional covariance matrix, which is estimated by the GICA-GARCH model, is measurable with respect to the information set that contains only past values of the observations.

In what follows, we analyze the relationship between the GICA-GARCH model and other factor GARCH models. We distinguish between the two branches of the literature about factor GARCH models based on whether the factor structure refers to the unconditional or conditional distribution of the data.

3.3.1. Factor structure in the unconditional distribution of the data

Here, we analyze the relationship between the GICA-GARCH model and the latent factor GARCH (Diebold & Nerlove, 1989) and DF-GARCH (Alessi et al., 2006) models.

First, the GICA-GARCH model can be seen as a latent factor model with GARCH effects (Diebold & Nerlove, 1989). As with the GICA-GARCH model, the latent factor ARCH model (Diebold & Nerlove, 1989) assumes that there are a few common latent factors (in particular, $r = 1$ in Diebold and Nerlove's model) that explain the co-movements among the observations and evolve according to univariate GARCH models ($\mathbf{H}_t = h_{1t}$). However, whereas the GICA-GARCH model assumes a factor structure in both the unconditional and conditional distributions of the data, the Diebold and Nerlove model only assumes a factor structure in the unconditional covariance matrix of the dataset. Consequently, the latent factor GARCH model assumes that the commonalities in the volatilities among observations are due to the ARCH effect of the common factor. That is, in Diebold and Nerlove's model, the conditional covariance matrix of the observations is given by:

$$\mathbf{\Omega}_t = \mathbf{a}_{(1)1}\mathbf{a}'_{(1)1}h_{1t} + \mathbf{\Gamma}, \tag{27}$$

where $\mathbf{\Gamma}$ is a diagonal matrix whose elements correspond to the constant conditional variances of the noisy components. Furthermore, in Eq. (27), h_{1t} is the conditional variance of the common factor that is not unobservable. Then, $\mathbf{\Omega}_t$ is not measurable when the information set contains only past values of the observations (it should contain past values of the latent factor too).

The GICA-GARCH model can also be seen as a parsimonious version of the DF-GARCH model (Alessi et al., 2006). Both models exploit the unconditional information contained in the entire dataset in order to estimate the conditional covariance matrix of the observations. The main difference between the two models is the parametrization of the common factors conditional covariance matrix. While the GICA-GARCH model, due to the statistical independence of the unobserved components, fits a univariate ARMA(p, q)-GARCH(p', q') model to each of them and assumes that $\mathbf{H}_t$ is diagonal, the DF-GARCH model assumes that the common factors have a zero-conditional mean and evolve according to a MGARCH model that is parameterized as a BEKK model:

$$\mathbf{H}_t = \mathbf{C}_0\mathbf{C}'_0 + \mathbf{C}'_1\mathbf{s}^{(1)}_{t-1}\mathbf{s}^{(1)'}_{t-1}\mathbf{C}_1 + \mathbf{C}'_2\mathbf{H}_{t-1}\mathbf{C}_2, \tag{28}$$

where $\mathbf{C}_i$ are matrices of constant parameters. Therefore, whereas the conditional covariance matrix of the dataset in the GICA-GARCH model depends on the conditional variances of the r common components, the DF-GARCH model estimates the conditional covariance matrix of the observations, taking into account both the conditional variances and covariances among the common latent factors. For both models, the GICA-GARCH and the DF-GARCH, the noise components, which represent the idiosyncratic part in the DF-GARCH model, follow univariate ARMA-GARCH models. The conditional covariance matrix, $\mathbf{\Gamma}_t$, is then diagonal for both models, but it is a full rank matrix in the DF-GARCH model, whereas in the GICA-GARCH model it will have a rank equal to $m - r$.

3.3.2. Factor structure in the conditional distribution of the data

In this section, we analyze the relationship between the GICA-GARCH model, Engle's model and some orthogonal models.

From Eqs. (24)–(26), it is clear that the GICA-GARCH model is related to the FACTOR-ARCH model (Engle, 1987). Both models assume that the data conditional covariance matrix is given by a linear combination of the conditional variances of some portfolios (factors) of the observations. Therefore, $\mathbf{\Omega}_t$ is measurable when the information set contains only past values of the observations. Engle's factor GARCH model assumes that $\mathbf{\Gamma}_t$ is a constant matrix that does not play any role in the model.

The GICA-GARCH model is also related to several orthogonal models, such as the O-GARCH (Alexander, 2001), the GO-GARCH (van der Weide, 2002), the GOF-GARCH (Lanne & Saikkonen, 2007), and the CUC-GARCH (Fan et al., 2008). All of these models assume that the data are generated by a linear combination of several factors that follow univariate GARCH models. The GICA-GARCH model can be seen as an extension of the O-GARCH model where the estimates of the factors are given by the ICs instead of the principal components (PCs). Both the GICA-GARCH and O-GARCH models approximate the data conditional covariance matrix by the univariate conditional variances of a few factors (the most risky factors), and transform the problem of estimating a MGARCH model into the estimation of a small number of univariate volatility models. The cost of reducing the dimensionality is that the factor conditional covariance matrices have reduced ranks. One extension of the O-GARCH model is the GO-GARCH model (van der Weide, 2002), which does not reduce the dimension and considers $r = m$. A restricted version of the model where only a subset of the underlying factors has a time-varying conditional variance has recently been analyzed by Lanne and Saikkonen (2007). This model, called the GOF-GARCH model, parameterizes the factor conditional covariance matrix as

$$\mathbf{H}_t = \mathrm{diag}(\mathbf{V}_t : \mathbf{I}_{m-r}) \tag{29}$$

where $\mathbf{V}_t = \mathrm{diag}(v_{1t}, \ldots, v_{rt})$ is the conditional covariance matrix of the heteroskedastic components. The GOF-GARCH model is then similar to the GICA-GARCH model when the noisy components of the GICA-GARCH are homoskedastic ($\mathbf{\Gamma}_t \equiv \mathbf{\Gamma}$ is a constant matrix). Thus, the GOF-GARCH model estimates the data conditional covariance matrix as:

$$\mathbf{\Omega}_t = \mathbf{A}\mathbf{H}_t\mathbf{A}' = \mathbf{A}_1\mathbf{V}_t\mathbf{A}_1' + \mathbf{\Gamma}, \tag{30}$$

where $\mathbf{\Gamma} = \mathbf{A}_2\mathbf{A}_2'$. Therefore, the GOF-GARCH model is also related to Engle's model, but, assuming that $\mathbf{\Gamma}$ plays a specific role, it is the conditional covariance matrix of the homoskedastic components. Finally, the GICA-GARCH model is related to the work proposed by Fan et al. (2008) that models multivariate volatilities through conditionally uncorrelated components. Both the GICA-GARCH and CUC-GARCH models separate the estimation of the unobserved components from fitting a univariate GARCH model for each one of them, and they estimate the components by looking for an orthogonal matrix that is the solution of a non-linear optimization problem. However, the GICA-GARCH model requires the components to be statistically independent, while the CUC-GARCH model imposes the weaker assumption of conditional uncorrelatedness.

Table 1 summarizes the main features of all of the models considered in this section.

4. Simulation experiments

In this section we compare the performances of the GICA-GARCH, O-GARCH, and CUC-GARCH models. The main differences among the three models are related to the properties assumed for the latent factors: the O-GARCH model assumes unconditionally uncorrelated factors which are estimated by PCA; the CUC-GARCH model assumes conditionally uncorrelated components which follow extended GARCH(1, 1) models and are estimated by quasi-maximum likelihood; and the GICA-GARCH model generalizes the previous models, assuming independent underlying factors which are estimated by ICA. It would therefore be interesting to analyze the performances of the three models in order to identify the conditionally heteroskedastic components.

We present three simulation experiments to demonstrate the effectiveness of ICA and CUC versus PCA for identifying unobserved components that have the main features of financial assets: excess kurtosis and non-Gaussian conditional distributions. In each of the experiments we generate six components of 1000 observations each, and standardize them to have a zero mean and unit variance. We then generate a 6×6 random loading matrix $\mathbf{A}$, mix the components according to Eq. (1), apply the three procedures (the GICA-GARCH, O-GARCH, and CUC-GARCH models) to the vector of observations, $\mathbf{x}_t$, and obtain the ICs, PCs, and CUCs respectively.

In the first experiment, we consider the case where the excess kurtosis in the data comes from different standard ARMA-GARCH specifications, and, in addition to Gaussian innovations, we include the Student's t distribution (Bollerslev, 1987), the Laplace distribution (Granger & Ding, 1995), and the generalized error distribution (GED) (Nelson, 1991). The second experiment considers conditionally heteroskedastic factors without temporal dependencies on the conditional mean. In the third experiment, we explore the case where the different excess kurtosis of the latent factors comes from different conditional distributions, and distinguish between two cases: the Student's t distribution with different degrees of freedom and the GED with different values for the shape parameter.

In order to analyze the performances of the three models, we compute the correlation coefficient between each original component and its estimation. Moreover, we compute the mean square error (MSE) between the original and the estimated components as $\mathrm{MSE}(s_j, \widehat{s}_j^{(\cdot)}) = 1/T(\sum_{t=1}^{T}\left(s_{jt} - \widehat{s}_{jt}^{(\cdot)}\right)^2)$, for $j = 1, \ldots, r$, where $\widehat{s}_{jt}^{(\cdot)}$ is the jth estimated component by the corresponding method.

In the first simulation experiment, we generate the components as defined in Table 2.

Note that the conditional distribution of the ARMA-GARCH components depends on the conditional distribution of ε_{jt}, $\forall j = 1, 2, 3, 4$. We consider four possible distributions for the innovations. First, we generate the factors defined in Table 2 assuming that ε_{jt} is conditionally Gaussian $\forall j = 1, 2, 3, 4$. We repeat this procedure three more times, assuming that the conditional distribution of ε_{jt}, $\forall j = 1, 2, 3, 4$, is Student's t (t_6), Laplace, and GED ($\kappa = 1.5$). Table 3 presents the average results for the correlation coefficients and the MSE between the original and the corresponding estimated components.

According to the results shown in Table 3, we can see that the average of the correlation coefficients and the MSE take almost identical values along the four conditional distributions we have considered here. Independently of which conditional distribution we take into account, the GICA-GARCH model that estimates the ICs applying FastICA or JADE provides the most reliable identification of the unobserved ARMA-GARCH components. On the other hand, PCA performs worst for all distributions. SOBI is the ICA algorithm that has the worst performance, although it is slightly better than CUC. This is to be expected, as conditionally heteroskedastic components have excess kurtosis and small correlation coefficients.

In the second experiment we generate components which have a constant conditional mean but are conditionally heteroskedastic, as given in Table 4.

As in the first experiment, we generate the factors defined in Table 4 assuming that ε_{jt} is conditionally Gaussian $\forall j = 1, 2, 3$. We then repeat the procedure twice, assuming a Student's t (t_6) distribution and the GED ($\kappa = 1.3$) for ε_{jt}, $\forall j = 1, 2, 3$. We compute the correlation coefficients and the MSEs between each original and the corresponding estimated component. The results (average measures) are shown in Table 5, and are very similar to those from the first experiment. This result is not surprising, and we conclude that imposing an ARMA structure on the conditional mean does not change the results at all.

Table 1
Taxonomy of some factor GARCH models.

Model	Factor structure	Factors conditional covariance matrix $\mathbf{H}_t$	Data conditional covariance matrix $\boldsymbol{\Omega}_t$	Noise conditional covariance matrix $\boldsymbol{\Gamma}_t$
Latent factor ARCH (Diebold & Nerlove, 1989)	$r = 1$ common latent factor, s_{1t}, explains the co-movements of the data	$\mathbf{H}_t = h_{1t}$ $h_{1t} = \alpha_0 + \sum_{l=1}^{p} \alpha_l s_{1t-l}^2$	$\boldsymbol{\Omega}_t = \mathbf{a}_{(1)1}\mathbf{a}'_{(1)1}h_{1t} + \boldsymbol{\Gamma}$ Not measurable when $I_{t-1} = \{\mathbf{x}_{t-j}\}_{j=1}^{t-1}$	$\boldsymbol{\Gamma} = \mathrm{diag}(\gamma_1, \ldots, \gamma_m)$ full-rank matrix, γ_i constant conditional variance of the noise
DF-GARCH (Alessi et al., 2006)	$r < m$ common dynamic latent factors, $\mathbf{s}_t^{(1)}$, are assumed in the UD of the data $\mathbf{s}_t^{(1)}$ are conditionally correlated	$\mathbf{H}_t$ non-diagonal and time-varying evolves according to a MGARCH (BEKK model) $\mathbf{H}_t = \mathbf{C}_0\mathbf{C}'_0 + \mathbf{C}'_1\mathbf{s}_{t-1}^{(1)}\mathbf{s}_{t-1}^{(1)'}\mathbf{C}_1 + \mathbf{C}'_2\mathbf{H}_{t-1}\mathbf{C}_2$	$\boldsymbol{\Omega}_t = \mathbf{A}_1\mathbf{H}_t\mathbf{A}'_1 + \boldsymbol{\Gamma}_t$ Not measurable when $I_{t-1} = \{\mathbf{x}_{t-j}\}_{j=1}^{t-1}$	$\boldsymbol{\Gamma}_t = \mathrm{diag}(\gamma_{1t}, \ldots, \gamma_{mt})$ full-rank matrix, γ_{it} conditional variance of n_{it}, $\mathrm{n}_{it} \sim$ ARMA-GARCH model
Factor ARCH model (Engle, 1987; Engle et al., 1990)	$r < m$ common factors ('portfolios'), $\mathbf{s}_t^{(1)} = \mathbf{W}'_1\mathbf{x}_t$, referred to the CD of the data	h_{it} conditional variance of the ith portfolio $h_{it} = \alpha_0^{(i)} + \sum_{l=1}^{p} \alpha_l^{(i)}(\mathbf{w}'_{(1)i}\mathbf{x}_{t-l})^2$	$\boldsymbol{\Omega}_t = \sum_{i=1}^{r} \mathbf{a}_{(1)i}\mathbf{a}'_{(1)i}h_{it} + \boldsymbol{\Gamma}$ Measurable when $I_{t-1} = \{\mathbf{x}_{t-j}\}_{j=1}^{t-1}$	$\boldsymbol{\Gamma}$ full-rank positive definite non-specific role
O-GARCH (Alexander, 2001)	$r < m$ unconditionally uncorrelated factors, $\mathbf{s}_t^{(1)}$, referred to the CD of the data $\mathbf{s}_t^{(1)}$ are the first r PCs	$\mathbf{H}_t = \mathrm{diag}(h_{1t}, \ldots, h_{rt})$ $h_{it} = \alpha_0^{(i)} + \sum_{l=1}^{p} \alpha_l^{(i)}(\mathbf{w}'_{(1)i}\mathbf{x}_{t-l})^2 + \sum_{l=1}^{q} \beta_l^{(i)}h_{it-l}$	$\boldsymbol{\Omega}_t = \mathbf{A}_1\mathbf{H}_t\mathbf{A}'_1 + \boldsymbol{\Gamma}_t$ $\boldsymbol{\Omega}_t \approx \mathbf{A}_1\mathbf{H}_t\mathbf{A}'_1$ Measurable when $I_{t-1} = \{\mathbf{x}_{t-j}\}_{j=1}^{t-1}$	$\boldsymbol{\Gamma}_t = \mathrm{diag}(0, \ldots, 0, h_{r+1t}, \ldots, h_{mt})$ of rank $m - r$, $\{h_{it}\}_{i=r+1}^{m}$ conditional variances of the $m - r$ discarded PCs
GO-GARCH (van der Weide, 2002)	$r = m$ unconditionally uncorrelated factors, $\mathbf{s}_t = \mathbf{W}'\mathbf{x}_t$, explain the CD of the data	$\mathbf{H}_t = \mathrm{diag}(h_{1t}, \ldots, h_{mt})$ $h_{it} = \alpha_0^{(i)} + \alpha_l^{(i)}(\mathbf{w}'_{(1)i}\mathbf{x}_{t-1})^2 + \beta_l^{(i)}h_{it-1}$	$\boldsymbol{\Omega}_t = \mathbf{A}\mathbf{H}_t\mathbf{A}'$ Measurable when $I_{t-1} = \{\mathbf{x}_{t-j}\}_{j=1}^{t-1}$	No noise
GOF-GARCH (Lanne & Saikkonen, 2007)	m unconditionally uncorrelated factors, $\mathbf{s}_t$, referred to the CD of the data $r < m$ heteroskedastic factors, $\mathbf{s}_t^{(1)} = \mathbf{W}'_1\mathbf{x}_t$, $m - r$ homoskedastic factors, $\mathbf{s}_t^{(2)} = \mathbf{W}'_2\mathbf{x}_t$	$\mathbf{H}_t = \mathrm{diag}(\mathbf{V}_t : \mathbf{I}_{m-r})$ $\mathbf{V}_t = \mathrm{diag}(v_{1t}, \ldots, v_{rt})$, $v_{it} = \alpha_0^{(i)} + \alpha_l^{(i)}(\mathbf{w}'_{(1)i}\mathbf{x}_{t-1})^2 + \beta_l^{(i)}v_{it-1}$	$\boldsymbol{\Omega}_t = \mathbf{A}\mathbf{H}_t\mathbf{A}'$ $\boldsymbol{\Omega}_t = \mathbf{A}_1\mathbf{V}_t\mathbf{A}'_1 + \boldsymbol{\Gamma}$ Measurable when $I_{t-1} = \{\mathbf{x}_{t-j}\}_{j=1}^{t-1}$	$\boldsymbol{\Gamma} = \mathbf{A}_2\mathbf{A}'_2$ covariance matrix of the homoskedastic factors
CUC-GARCH (Fan et al., 2008)	$m = r$ conditionally uncorrelated factors, $\mathbf{s}_t = \mathbf{W}'\mathbf{x}_t$, explain the CD of the data	$\mathbf{H}_t = \mathrm{diag}(h_{1t}, \ldots, h_{mt})$	$\boldsymbol{\Omega}_t = \mathbf{A}\mathbf{H}_t\mathbf{A}'$	No noise

Table 1 (*continued*)

Model	Factor structure	Factors conditional covariance matrix $\mathbf{H}_t$	Data conditional covariance matrix $\boldsymbol{\Omega}_t$	Noise conditional covariance matrix $\boldsymbol{\Gamma}_t$
		$h_{it} = \alpha_0^{(i)} + \alpha_l^{(i)}(\mathbf{w}'_{(1)i}\mathbf{x}_{t-1})^2 + \beta_l^{(i)} h_{it-1}$	Measurable when $I_{t-1} = \{\mathbf{x}_{t-j}\}_{j=1}^{t-1}$	
GICA-GARCH	m statistically independent components, $\mathbf{s}_t$, explain the co-movements of the data and imply factor structure in the CD of the data $r < m$ are the most risky ICs, $\mathbf{s}_t^{(1)} = \mathbf{W}'_1\mathbf{x}_t\, m - r$ noisy components, $\mathbf{s}_t^{(2)} = \mathbf{W}'_2\mathbf{x}_t$	$\mathbf{H}_t = \mathrm{diag}(h_{1t}, \ldots, h_{rt})$	$\boldsymbol{\Omega}_t = \mathbf{A}_1\mathbf{H}_t\mathbf{A}'_1 + \boldsymbol{\Gamma}_t$	$\boldsymbol{\Gamma}_t = \mathrm{diag}(0, \ldots, 0, h_{r+1t}, \ldots, h_{mt})$ of rank $m - r$, $\{h_{it}\}_{i=r+1}^{m}$ conditional variances of the $m - r$ discarded ICs
		$h_{it} = \alpha_0^{(i)} + \sum_{l=1}^{p'} \alpha_l^{(i)} e_{it-l}^2 + \sum_{l=1}^{q'} \beta_l^{(i)} h_{it-l}$ e_{it} residuals of the ARMA model for s_{it}	$\boldsymbol{\Omega}_t \approx \mathbf{A}_1\mathbf{H}_t\mathbf{A}'_1$ Measurable when $I_{t-1} = \{\mathbf{x}_{t-j}\}_{j=1}^{t-1}$	
		$h_{it} = \alpha_0^{(i)} + \sum_{l=1}^{p'} \alpha_l^{(i)}(\mathbf{w}'_{(1)i}\mathbf{x}_{t-l} - \mu_{it-l}^{(1)})^2 + \beta_l^{(i)} h_{it-l}$		

Note: UD and CD denote the unconditional and conditional distributions, respectively.

Table 2
Definition of the original factors.

$s_{1t} \sim$ AR(1)-GARCH(1, 1)	$s_{1t} = 0.0289 + 0.7112 s_{1t-1} + a_{1t}$
	$a_{1t} = \sqrt{h_{1t}}\varepsilon_{1t}$; $h_{1t} = 0.0152 + 0.2080 a^2_{1t-1} + 0.7918 h_{1t-1}$
$s_{2t} \sim$ AR(2)-ARCH(1)	$s_{2t} = 1.2 s_{2t-1} - 0.32 s_{2t-1} + a_{2t}$
	$a_{2t} = \sqrt{h_{2t}}\varepsilon_{2t}$; $h_{2t} = 0.2 + 0.7 a^2_{2t-1}$
$s_{3t} \sim$ ARMA(1,1)-GARCH(1, 1)	$s_{3t} = 5 + 0.9 s_{3t-1} + a_{3t} - 0.4 a_{3t-1}$
	$a_{3t} = \sqrt{h_{3t}}\varepsilon_{3t}$; $h_{3t} = 0.0079 + 0.0650 a^2_{3t-1} + 0.9291 h_{3t-1}$
$s_{4t} \sim$ GARCH(1, 3)	$a_{4t} = \sqrt{h_{4t}}\varepsilon_{4t}$; $h_{4t} = 0.241 a^2_{4t-1} + 0.077 h_{4t-1} + 0.430 h_{4t-2} + 0.203 h_{4t-3}$
$s_{5t} \sim U(0, 1)$	$s_{6t} \sim$ GED(0, 1, 1.8)

Note: ε_{jt} is a random noise with a zero mean and unit variance, and is independent of h_{jt}, $\forall j = 1, 2, 3, 4$. We generate four sets of these components by changing the conditional distribution of ε_{jt}: Gaussian, Student's t (t_6), Laplace, and GED ($\kappa = 1.5$).

Table 3
Average values for the correlation coefficients and the MSE between the original and the estimated components.

	Gaussian		Student's (t_6)		Laplace		GED	
	Correlation	MSE	Correlation	MSE	Correlation	MSE	Correlation	MSE
CUC	0.7903	0.4192	0.7824	0.4349	0.7663	0.4672	0.7360	0.5277
FAST	0.9617	0.0766	0.9634	0.0731	0.9571	0.0858	0.9586	0.0828
JADE	0.9591	0.0817	0.9408	0.1184	0.9158	0.1682	0.9554	0.0892
SOBI	0.8353	0.3292	0.7790	0.4419	0.8403	0.3192	0.8076	0.3846
PCA	0.6646	0.6700	0.7035	0.5925	0.6952	0.6091	0.7087	0.5820

Table 4
Definition of the original factors.

$s_{1t} \sim$ ARCH(1)	$s_{1t} = \sqrt{h_{1t}}\varepsilon_{1t}$; $h_{1t} = 0.2 + 0.7 s^2_{1t-1}$
$s_{2t} \sim$ GARCH(1, 1)	$s_{2t} = \sqrt{h_{2t}}\varepsilon_{2t}$; $h_{2t} = 0.021 + 0.073 s^2_{2t-1} + 0.906 h_{2t-1}$
$s_{3t} \sim$ GARCH(1, 2)	$s_{3t} = \sqrt{h_{3t}}\varepsilon_{3t}$; $h_{3t} = 1.692 + 0.245 s^2_{3t-1} + 0.337 h_{3t-1} + 0.310 h_{3t-2}$
$s_{4t} \sim t_9$	$s_{5t} \sim N(0, 1)$ $s_{6t} \sim U(0, 1)$

Note: ε_{jt} is a random noise with zero mean and unit variance, and is independent of h_{jt}, $\forall j = 1, 2, 3$. We generate four sets of these components by changing the conditional distribution of ε_{jt}: Gaussian, Student's t (t_6), and GED ($\kappa = 1.3$).

Table 5
Average values for the correlation coefficients and the MSE between the original and estimated components.

	Gaussian		Student's t		GED	
	Correlation	MSE	Correlation	MSE	Correlation	MSE
CUC	0.7870	0.4257	0.8438	0.3123	0.8523	0.2953
FAST	0.9711	0.0578	0.9850	0.0300	0.9847	0.0306
JADE	0.9796	0.0408	0.9852	0.0296	0.9733	0.0533
SOBI	0.9218	0.1563	0.8495	0.3009	0.9392	0.1215
PCA	0.6994	0.6007	0.7037	0.5920	0.6964	0.6066

Table 6
Average values for the correlation coefficients and the MSEs between the original and estimated components.

	Student's t		GED	
	Correlation	MSE	Correlation	MSE
CUC	0.7113	0.5771	0.7443	0.5111
FAST	0.8039	0.3920	0.9004	0.1991
JADE	0.8868	0.2263	0.8949	0.2101
SOBI	0.8191	0.3616	0.7926	0.4146
PCA	0.7297	0.5400	0.6705	0.6583

In the third experiment we analyze the situation where all of the components follow the same ARMA-GARCH specification, and the different excess kurtosis comes from different conditional distributions, as defined in Table 2. The conditional distribution for $\varepsilon_{jt}, j = 1, 2, 3$, could be: (i) Student's t with different degrees of freedom for each $j = 1, 2, 3$, or (ii) GED with different values for the shape parameter for each $j = 1, 2, 3$. For (i) we generate $\varepsilon_{1t} \sim t_6$, $\varepsilon_{2t} \sim t_9$, and $\varepsilon_{3t} \sim t_{11}$. On the other hand, for (ii) we generate $\varepsilon_{1t} \sim$ GED(0, 1, 1.5), $\varepsilon_{2t} \sim$ GED(0, 1, 2), and $\varepsilon_{3t} \sim$ GED(0, 1, 1.01). The average results obtained for the correlation coefficients and the MSEs are given in Table 6. The results show that when the excess kurtosis comes from different conditional distributions (or, to put it better, from the same conditional distribution with different values for the parameters) any of the ICA methods performs better than either PCA or CUC. If the innovations come from a Student's t conditional distribution with different degrees of freedom, PCA and CUC have similar performances. However, if the conditional distribution is the GED with different shape parameters, PCA performs worse than CUC.

From these simulations, we conclude that the ICA algorithms, especially FastICA and JADE, perform the best for identifying the unobserved conditionally heteroskedastic factors. The performances of the three ICA algorithms are as expected: as FastICA and JADE look for the independence of the ICs maximizing the non-Gaussianity, they

Table 7
Summary statistics for the standardized stock returns.

Stocks	Summary statistics								
	Zero mean stock returns $\mathbf{x}_t$							$\lvert\mathbf{x}_t\rvert$	$\mathbf{x}_t^2$
	Median	Maximum	Minimum	St. Dev	Kurtosis	JB	LB(50)	LB(50)	LB(50)
ACS	0.0003	0.0868	−0.0797	0.0182	5.3619	295.64*	119.23**	1053.39**	596.25**
ACX	−0.0004	0.0923	−0.0998	0.0206	4.8558	182.63*	84.03**	718.48**	425.03**
ALT	0.0003	0.0823	−0.0994	0.0189	5.9571	493.04*	82.19**	1188.20**	667.89**
AMS	−0.0001	0.1421	−0.1502	0.0298	5.3303	282.95*	65.20	985.56**	463.07**
ANA	−0.0003	0.0720	−0.0731	0.0155	5.8949	436.48*	42.75	678.85**	703.10**
BBVA	0.0002	0.0944	−0.0799	0.0217	4.7493	165.20*	87.12**	2411.12**	1871.64**
BKT	0.0003	0.0900	−0.0906	0.0192	5.9641	458.73*	66.99	1488.85**	843.23**
ELE	0.0005	0.0831	−0.0747	0.0175	5.4053	305.54*	75.20	2430.05**	1931.99**
FCC	−0.0005	0.0784	−0.0595	0.0173	5.0625	245.37*	73.40	1008.10**	599.91**
FER	−0.0009	0.0836	−0.0800	0.0194	4.6196	138.63*	62.73	972.33**	625.17**
IBE	−0.0001	0.0567	−0.0592	0.0121	5.3139	282.92*	53.60	655.78**	352.19**
IDR	−0.0002	0.0903	−0.0921	0.0232	4.8257	177.73*	66.29	892.09**	626.85**
NHH	0.0001	0.0872	−0.0845	0.0182	4.7760	164.72*	64.54	334.79**	191.12**
POP	0.0000	0.0722	−0.0601	0.0157	4.9358	199.95*	84.47**	786.48**	498.84**
REP	0.0004	0.0879	−0.0814	0.0180	4.9307	197.25*	77.35**	1927.56**	1033.91**
SAN	0.0002	0.0964	−0.1135	0.0233	5.0546	220.64*	67.92	2524.01**	1783.61**
SGC	0.0004	0.1414	−0.1394	0.0339	4.8686	189.30*	69.27	1394.44**	850.22**
TEF	0.0002	0.1016	−0.0872	0.0235	4.0998	72.24*	61.42	1533.22**	740.86**
TPI	0.0004	0.1402	−0.1305	0.0294	5.5625	342.14*	67.48	1334.54**	682.21**

Notes: JB denotes the Jarque-Bera test statistic for normality and LB is the Ljung-Box test statistic based on 50 lags for the autocorrelation of the rates of return, the absolute and the squared returns.
* Indicates that the null hypothesis of normality is rejected at the 1% level of significance for the absolute returns.
** Indicates that the null of no autocorrelation is rejected at the 1% level of significance for the rates of returns and squared returns.

capture the excess kurtosis of the conditionally heteroskedastic components better than SOBI. PCA performs the worst, so it seems that the orthogonal GARCH models would not be good methods for forecasting the conditional variance of large datasets. According to the results, the GICA-GARCH method seems to outperform the CUC-GARCH and the O-GARCH. We will investigate this contention in the next section.

5. Empirical application

In this section we apply our procedure to a dataset of stock returns. First, we describe the data used; second, we explain the procedure for estimating the components; and, third, we present the results from using the GICA-GARCH, CUC-GARCH, and O-GARCH models to forecast the conditional variances of the stock returns.

The data consist of daily closing prices of the 19 assets which were always included in the IBEX 35 from 2000 to 2004 (see Table 12 in the Appendix for a detailed description of the 19 stocks). The IBEX 35 index is the main stock market index of the Madrid stock market. Its composition is revised twice a year, and comprises the 35 companies on the Madrid stock exchange with the largest trading volume. We apply some preprocessing steps to the data. First of all, to achieve stationarity, we computed the daily stock returns by taking the first differences of the logarithm of daily closing prices: $\mathbf{r}_t = \log(\mathbf{p}_{t+1}) - \log(\mathbf{p}_t)$, $t = 1, \ldots, T = 1250$. Then $\mathbf{r}_t$ is a 19×1250 multivariate vector of stock returns, whose columns are the values of these 19 stocks in the 1250 trading days over the period 2000–2004. There are some extreme observations that correspond to outliers, which are due to known changes such as stock splits or other legal changes; these have been removed. Finally, we also remove the mean from the stock returns, and $\mathbf{x}_t = \mathbf{r}_t - \bar{\mathbf{r}}$ are the data that we analyze.

Table 7 presents a summary of the basic statistics of the data. This table includes the Jarque-Bera statistic and the Ljung-Box statistic computed based on 50 lags of the series, as well as the absolute values and the squares of the stock returns.

The standard deviation of the stock returns, varying from 0.0121 for IBE to 0.0339 for SGC, indicates that there are both high and low volatility stock returns in our dataset. The high values of the kurtosis coefficients (higher than 3 for all of the stock returns) confirm the fat-tailed property of the conditional stock return distribution. Moreover, the Jarque-Bera test statistics are very high, and we clearly reject the null hypothesis of normality at the 1% level of significance. Then, as the conditional distribution of stock returns is far away from Gaussianity, ICA may have the potential to identify the set of latent components that explain the co-movements of the stock returns. According to the Ljung-Box statistics for the stock returns, 13 of the 19 series do not present relevant autocorrelation (the other 6 series have some significant autocorrelation coefficients which can be removed by fitting autoregressive models to the series). For the squares and the absolute values of the stock returns, the high values of the Ljung-Box statistics indicate strong autocorrelation in all series, and suggest the presence of non-linear dependence in the stock returns. These are the empirical results that we expect when dealing with financial data.

We apply the GICA-GARCH, CUC-GARCH, and O-GARCH models to the vector of zero mean stock returns, $\mathbf{x}_t$, and we obtain the corresponding estimates of the 19 unobserved factors. We sort the ICs, the CUCs, and the PCs in terms of the explained total variance. From the results, which are displayed in Table 8, we can quantify how much risk is

Table 8
Sorted components in terms of their explained variability.

CUC	$\%_{\text{CUC}}$	FAST	$\%_{\text{Fast}}$	JADE	$\%_{\text{JADE}}$	SOBI	$\%_{\text{SOBI}}$	PCA	$\%_{\text{PCA}}$
$\widehat{s}_{1t}^{C}$	18.10	$\widehat{s}_{1t}^{F}$	17.72	$\widehat{s}_{1t}^{J}$	11.75	$\widehat{s}_{1t}^{S}$	11.13	$\widehat{s}_{1t}^{P}$	35.30
$\widehat{s}_{2t}^{C}$	16.44	$\widehat{s}_{2t}^{F}$	10.22	$\widehat{s}_{2t}^{J}$	7.29	$\widehat{s}_{2t}^{S}$	9.65	$\widehat{s}_{2t}^{P}$	7.00
$\widehat{s}_{3t}^{C}$	8.04	$\widehat{s}_{3t}^{F}$	6.40	$\widehat{s}_{3t}^{J}$	6.48	$\widehat{s}_{3t}^{S}$	9.16	$\widehat{s}_{3t}^{P}$	5.91
$\widehat{s}_{4t}^{C}$	5.43	$\widehat{s}_{4t}^{F}$	5.92	$\widehat{s}_{4t}^{J}$	6.36	$\widehat{s}_{4t}^{S}$	8.15	$\widehat{s}_{4t}^{P}$	4.78
$\widehat{s}_{5t}^{C}$	5.20	$\widehat{s}_{5t}^{F}$	5.76	$\widehat{s}_{5t}^{J}$	6.17	$\widehat{s}_{5t}^{S}$	7.52	$\widehat{s}_{5t}^{P}$	4.73
$\widehat{s}_{6t}^{C}$	4.35	$\widehat{s}_{6t}^{F}$	4.89	$\widehat{s}_{6t}^{J}$	5.70	$\widehat{s}_{6t}^{S}$	5.42	$\widehat{s}_{6t}^{P}$	4.35
$\widehat{s}_{7t}^{C}$	4.27	$\widehat{s}_{7t}^{F}$	4.65	$\widehat{s}_{7t}^{J}$	5.61	$\widehat{s}_{7t}^{S}$	5.23	$\widehat{s}_{7t}^{P}$	4.24
$\widehat{s}_{8t}^{C}$	3.86	$\widehat{s}_{8t}^{F}$	4.62	$\widehat{s}_{8t}^{J}$	5.52	$\widehat{s}_{8t}^{S}$	4.37	$\widehat{s}_{8t}^{P}$	4.04
$\widehat{s}_{9t}^{C}$	3.58	$\widehat{s}_{9t}^{F}$	4.43	$\widehat{s}_{9t}^{J}$	5.20	$\widehat{s}_{9t}^{S}$	4.11	$\widehat{s}_{9t}^{P}$	3.62
$\widehat{s}_{10t}^{C}$	3.45	$\widehat{s}_{10t}^{F}$	4.15	$\widehat{s}_{10t}^{J}$	5.16	$\widehat{s}_{10t}^{S}$	4.00	$\widehat{s}_{10t}^{P}$	3.60
$\widehat{s}_{11t}^{C}$	3.37	$\widehat{s}_{11t}^{F}$	3.86	$\widehat{s}_{11t}^{J}$	5.12	$\widehat{s}_{11t}^{S}$	3.83	$\widehat{s}_{11t}^{P}$	3.31
$\widehat{s}_{12t}^{C}$	3.33	$\widehat{s}_{12t}^{F}$	3.85	$\widehat{s}_{12t}^{J}$	4.74	$\widehat{s}_{12t}^{S}$	3.73	$\widehat{s}_{12t}^{P}$	3.13
$\widehat{s}_{13t}^{C}$	3.30	$\widehat{s}_{13t}^{F}$	3.69	$\widehat{s}_{13t}^{J}$	4.01	$\widehat{s}_{13t}^{S}$	3.57	$\widehat{s}_{13t}^{P}$	3.03
$\widehat{s}_{14t}^{C}$	3.22	$\widehat{s}_{14t}^{F}$	3.67	$\widehat{s}_{14t}^{J}$	3.85	$\widehat{s}_{14t}^{S}$	3.57	$\widehat{s}_{14t}^{P}$	2.86
$\widehat{s}_{15t}^{C}$	3.05	$\widehat{s}_{15t}^{F}$	3.56	$\widehat{s}_{15t}^{J}$	3.84	$\widehat{s}_{15t}^{S}$	3.57	$\widehat{s}_{15t}^{P}$	2.66
$\widehat{s}_{16t}^{C}$	3.03	$\widehat{s}_{16t}^{F}$	3.47	$\widehat{s}_{16t}^{J}$	3.76	$\widehat{s}_{16t}^{S}$	3.42	$\widehat{s}_{16t}^{P}$	2.56
$\widehat{s}_{17t}^{C}$	2.90	$\widehat{s}_{17t}^{F}$	3.26	$\widehat{s}_{17t}^{J}$	3.51	$\widehat{s}_{17t}^{S}$	3.26	$\widehat{s}_{17t}^{P}$	2.21
$\widehat{s}_{18t}^{C}$	2.83	$\widehat{s}_{18t}^{F}$	2.97	$\widehat{s}_{18t}^{J}$	3.41	$\widehat{s}_{18t}^{S}$	3.22	$\widehat{s}_{18t}^{P}$	1.71
$\widehat{s}_{19t}^{C}$	2.29	$\widehat{s}_{19t}^{F}$	2.89	$\widehat{s}_{19t}^{J}$	2.51	$\widehat{s}_{19t}^{S}$	3.10	$\widehat{s}_{19t}^{P}$	0.93
	100.00		100.00		100.00		100.00		100.00

associated with each component. This fact is crucial, since we would like to calculate the value at risk of a portfolio of the IBEX 35 index, or indeed any other risk management application.

We use Fig. 1, which shows the variability explained by the components estimated using the five algorithms, to determine the optimal number of components for each method. That is, we choose the components that are the most important sources of risk. The results are given in Table 9, which also includes the absolute variability explained by the r selected components.

We are interested in determining which assets are most important for defining each component. From Eq. (2), $\{\widehat{s}_{it}\}_{i=1}^{19}$ can be written as a linear combination of the stock returns, $\widehat{s}_{it} = \sum_{j=1}^{19} w_{ij}x_{jt}$, where w_{ij} represents the effect of the jth stock returns on the ith component, and the largest weights correspond to the most important assets. The ICs, the CUCs, and the PCs each have different interpretations. As an example, we analyze the first components. The first PC is given by a weighted mean of the 19 stock returns, and can be considered as an index of the market. Indeed, if we plot the variation in the variability of the first PC and the IBEX 35 index, considering groups of ten observations, it is clear that the first PC reflects the main movements of the index IBEX 35 (see Fig. 2). Then, if we forecast the volatility of $\mathbf{x}_t$ from the volatility of the first PC, the 19 stock returns will tend to move together.

The results for the ICs are different: they cannot be seen as indexes of the market. The first ICs are mainly associated with electricity, the building industries, and banking (the sectorial economic classification is detailed in the Appendix), and separate the stock returns in terms of the individual explained variability, $\{\nu_1^i\}_{i=1}^{19}$ (see Eq. (15)). As an example, we analyze the first FastICA, $\widehat{s}_{1t}^{F}$. In Fig. 3, which shows the variation in the variability of $\widehat{s}_{1t}^{F}$ and the largest weighted assets on $\widehat{s}_{1t}^{F}$, we see that all of the assets form a cluster of high variability from observations 600 to 750. The assets which are positively weighted only show this period of higher variability, but the negative ones are also volatile at the beginning of the sample.

Table 9
Number of unobserved components and the percentage of the total variability explained.

	CUC	FAST	JADE	SOBI	PCA
r	4	2	2	5	1
% Variability	47.97	27.95	19.04	45.62	35.30

The forecasting performances of the GICA-GARCH, CUC-GARCH, and O-GARCH models are checked as follows:

1. We estimate $\mathbf{A}$ and the unobserved components, for each model, using the whole sample. The components are then sorted and r is fixed.
2. Using the whole sample, we fit an ARMA(p, q) with GARCH(p', q') disturbances for each component $\widehat{s}_{jt}$, with $j = 1, \ldots, r$.
3. The standard ARMA-GARCH processes assume conditionally Gaussian distributions. However, as the stock returns are far away from Gaussianity, the unobserved components should be non-Gaussian too, in which case the standard ARMA-GARCH specification may not be adequate to fit the components. In this paper, we explore alternative conditional distributions, and estimate the parameters of the ARMA(p, q)-GARCH(p', q') model, with a sample of 1000 observations, using the Gaussian, Student's t, and GED distributional models for innovations. Then, for each model, we generate one-step-ahead forecasts for the univariate conditional variance of each $\widehat{s}_{jt}$,

$$\widehat{h}_{j,1001|1000} = V\left[\widehat{s}_{j1001} | \mathbf{I}_{1000}\right], \quad j = 1, \ldots, r. \qquad (31)$$

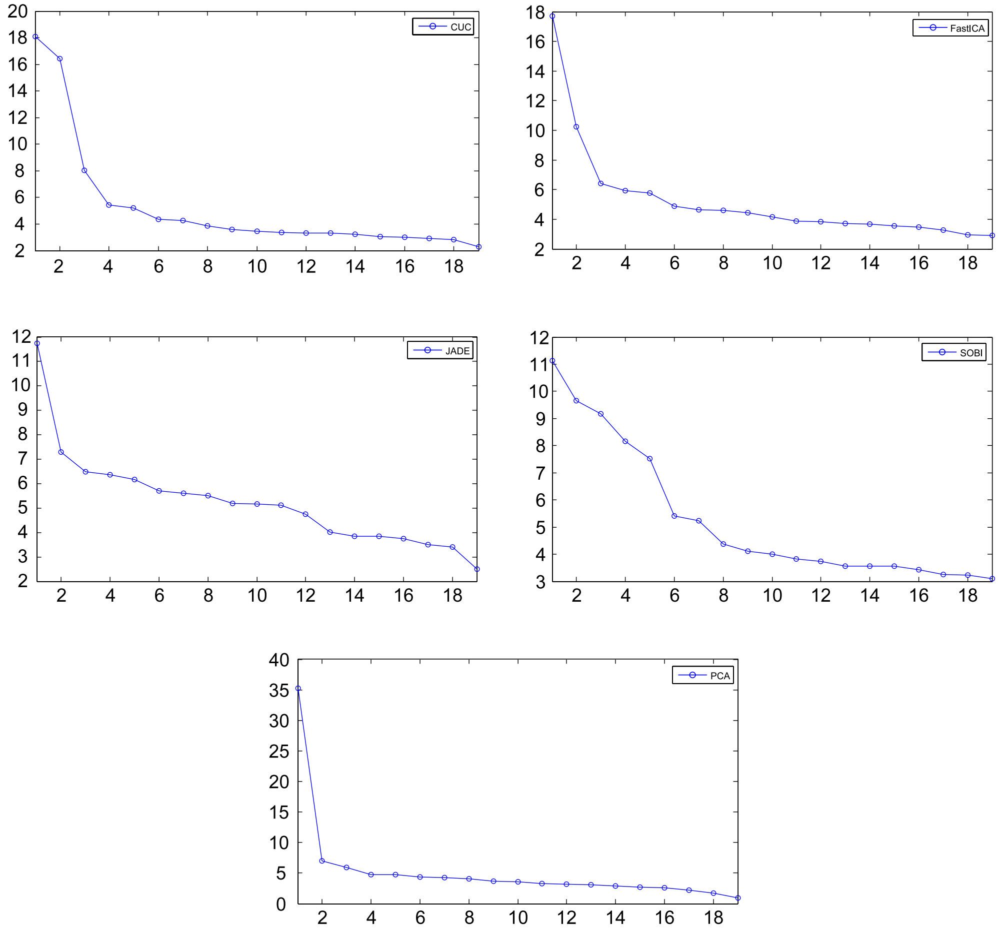

Fig. 1. Total variability explained by the components.

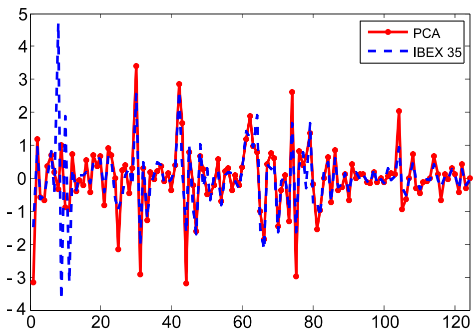

Fig. 2. Variation in the variability of $\widehat{s}^p_{1t}$ and the IBEX 35 index.

Thus, using a rolling prediction for $t = 1001, \ldots, 1250$, we have:

$$\widehat{\mathbf{H}}_{t|t-1} = \text{diag}(\widehat{h}_{1,t|t-1}, \ldots, \widehat{h}_{r,t|t-1}), \quad t = 1001, \ldots, 1250, \tag{32}$$

which is the conditional covariance matrix of $\widehat{\mathbf{s}}_t = (\widehat{s}_{1t}, \ldots, \widehat{s}_{rt})'$ at time t.

4. The conditional variance of $\mathbf{x}_t$ at time t, $\boldsymbol{\Omega}_t$, is computed using Eq. (20). The conditional variance of the ith stock return at time t is then given by the ith diagonal term of $\boldsymbol{\Omega}_t$:

$$\widehat{\gamma}^2_{i,t|t-1} = \sum_{j=1}^{r} \widehat{h}_{j,t|t-1} a^2_{ij}, \quad i = 1, 2, \ldots, 19, \; t = 1001, \ldots, 1250. \tag{33}$$

From this expression and Eq. (19), we can see that $\mathbf{x}_t$, which is generated by a linear combination of a set of ICs, possess a GARCH-type structure. This result is confirmed by the work of Nijman and Sentana (1996), who show that a linear combination of independent GARCH processes will be a weak GARCH process.

5. To evaluate the forecasting performances of the GICA-GARCH, CUC-GARCH, and O-GARCH models, we need to compare the predicted volatility and the real one. As the population volatility is not observed, the literature proposes the substitution of a proxy for the real volatility. Initially, the squares of the stock returns were used as a proxy for the conditional variance (see for example Franses & van Dijk, 1996). However, it has been shown that the squared returns form a noisy proxy for the conditional variance and perform very poorly (Andersen & Bollerslev, 1998). Furthermore, Hansen and Lunde (2006) show that an evaluation based on squared returns can induce an inconsistent ranking of

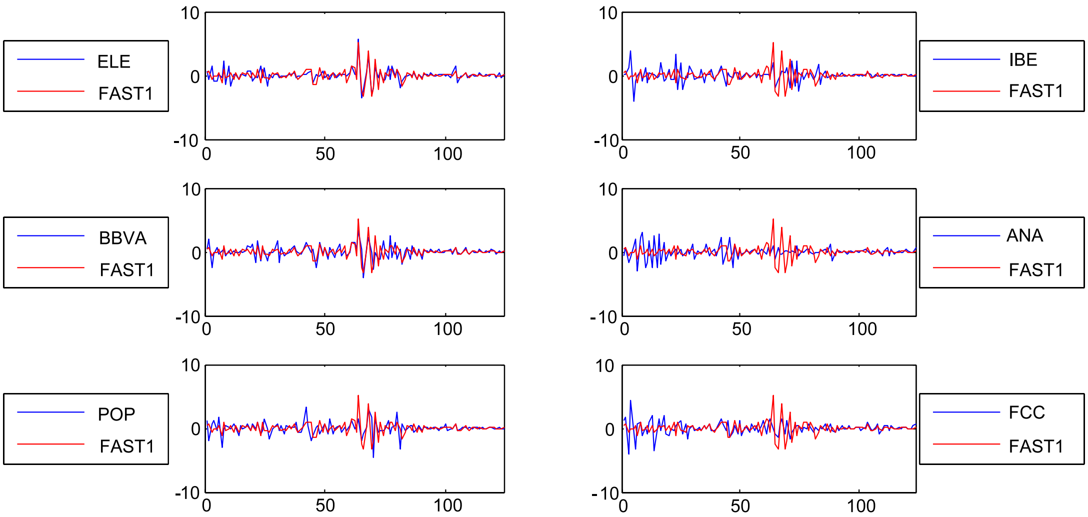

Fig. 3. Variation in the variability of $\widehat{s}^{f}_{1t}$ and the stock returns with the largest weights: the positive ones on the left, and the negative ones on the right.

volatility models, and may select an inferior model as the 'best' with a probability that goes to one when the sample size increases. To avoid such inconsistencies, we follow Hansen and Lunde's approach and estimate the conditional variance using the realized variance (RV), which is constructed by taking the sum of squared intraday returns (for more details, see Hansen & Lunde, 2006). Assuming that at day t we have f intraday observations of the ith stock return, the RV at time t is defined as:

$$\mathrm{RV}_{it} = \sum_{l=1}^{f} x^2_{i,t,l},$$
$$i = 1, \ldots, 19, \ t = 1, \ldots, T = 1250. \tag{34}$$

In our empirical analysis, we construct the intraday stock returns artificially, as follows. For a given trading day t, we use the part of the day that the Madrid stock market is open (9:00–17:30), and generate artificial five-minute returns per day ($f = 102$) by a linear interpolation method. We then have $x^2_{i,t,l}$, and we compute RV_{it}, for $i = 1, \ldots, 19$, $t = 1001, \ldots, T = 1250$, as in Eq. (34). Once we have computed the RV, we need to define the proxy for the true volatility. Following Hansen and Lunde (2006), we employ three different proxies for the conditional variance: $\mathrm{Proxy1}_{it} = \widehat{c}\mathrm{RV}_{it}$, where $\widehat{c} = T^{-1}\sum_{t=1}^{T} x^2_{it}/\mathrm{RV}_{it}$, $\mathrm{Proxy2}_{it} = \mathrm{RV}_{it} + (p^{\mathrm{open}}_t - p^{\mathrm{close}}_{t-1})^2$, and $\mathrm{Proxy3}_{it} = x^2_{it}$. Then, substituting each proxy for the unobserved conditional variance, the one-step-ahead volatility forecast error is given by:

$$\epsilon_{it} = \mathrm{Proxy}_{it} - \widehat{\gamma}^2_{i,t|t-1},$$
$$i = 1, 2, \ldots, 19, \ t = 1001, \ldots, 1250. \tag{35}$$

6. To evaluate the accuracy of the model, we compare the prediction error (Eq. (35)) with a benchmark. This benchmark is obtained by predicting the volatility of the stock returns by their marginal variance. Then, we define the relative forecast error by:

$$RE_{it} = \frac{\epsilon_{it}}{\epsilon^*_{it}},$$
$$i = 1, 2, \ldots, 19, \ t = 1001, \ldots, 1250, \tag{36}$$

where ϵ^*_{it} is the forecast error of the ith stock return obtained by the benchmark method, computed by

$$\epsilon^*_{it} = \mathrm{Proxy}_{it} - \widehat{\sigma}^2_i,$$
$$i = 1, 2, \ldots, 19, \ t = 1001, \ldots, 1250, \tag{37}$$

where $\widehat{\sigma}^2_i$ is the marginal variance of the ith stock return at time t. To minimize the impact of outliers when we analyze the volatility forecasting performances of the GICA-GARCH, CUC-GARCH, and O-GARCH models, we use the Median Relative Absolute Error (MdRAE) criterion (see for example Hyndman & Koehler, 2006, for a complete review of measures of forecast accuracy):

$$\mathrm{MdRAE}(RE_{it}) = \mathrm{median}(|RE_{it}|).$$

In addition, we can also use the ratio of the corresponding measure for the ICA and the CUC methods with respect to the PCA:

$$\mathrm{RelMdRAE} = \frac{\mathrm{MdRAE_{ICA}}}{\mathrm{MdRAE_{PCA}}}. \tag{38}$$

Our purpose here is to compare the forecasting performances of the GICA-GARCH, CUC-GARCH, and O-GARCH models when the latent factors are conditionally Gaussian, Student t, and GED distributed. We propose to make this comparison following two approaches. In the first approach, we fit a univariate ARMA-GARCH model for each component, as we have explained above. In the second approach, even though the CUC-GARCH model assumes that all components follow GARCH(1, 1) processes, and it is common to use this specification for modelling stock returns (see, for example Hansen & Lunde, 2005), we decide to analyze the forecasting performance

Table 10
Comparison of the overall forecasting performance of the GARCH(1, 1) modelling approach. The entries represent the average values of the RelMdRAE criterion, measured over the 19 stock returns, when a univariate GARCH(1, 1) model is fitted to each component.

	Gaussian distribution														
	RelMdRAE (Proxy 1)					RelMdRAE (Proxy 2)					RelMdRAE (Proxy 3)				
	CUC	FAST	JADE	SOBI	PCA	CUC	FAST	JADE	SOBI	PCA	CUC	FAST	JADE	SOBI	PCA
$r = 1$	0.729	0.750	0.754	0.729	1.000	0.732	0.755	0.760	0.734	1.000	0.735	0.750	0.760	0.727	1.000
$r = 2$	0.838	0.794	0.690	0.770	1.000	0.840	0.797	0.696	0.773	1.000	0.832	0.799	0.702	0.759	1.000
$r = 3$	0.964	0.810	0.679	0.781	1.000	0.968	0.814	0.685	0.784	1.000	0.959	0.816	0.683	0.773	1.000
$r = 4$	0.966	0.847	0.660	0.854	1.000	0.967	0.851	0.663	0.857	1.000	0.966	0.849	0.661	0.846	1.000
$r = 5$	1.004	0.863	0.693	0.859	1.000	1.006	0.864	0.694	0.860	1.000	1.007	0.867	0.695	0.856	1.000
	Student's t distribution														
	RelMdRAE (Proxy 1)					RelMdRAE (Proxy 2)					RelMdRAE (Proxy 3)				
	CUC	FAST	JADE	SOBI	PCA	CUC	FAST	JADE	SOBI	PCA	CUC	FAST	JADE	SOBI	PCA
$r = 1$	0.735	0.766	0.769	0.741	1.000	0.740	0.772	0.777	0.747	1.000	0.743	0.768	0.776	0.741	1.000
$r = 2$	0.831	0.821	0.701	0.768	1.000	0.831	0.823	0.708	0.775	1.000	0.830	0.826	0.712	0.764	1.000
$r = 3$	0.944	0.828	0.683	0.782	1.000	0.946	0.833	0.689	0.785	1.000	0.942	0.837	0.688	0.777	1.000
$r = 4$	0.939	0.861	0.667	0.835	1.000	0.940	0.864	0.670	0.838	1.000	0.935	0.858	0.669	0.830	1.000
$r = 5$	0.992	0.885	0.704	0.845	1.000	0.992	0.886	0.706	0.846	1.000	0.988	0.885	0.708	0.840	1.000
	GED distribution														
	RelMdRAE (Proxy 1)					RelMdRAE (Proxy 2)					RelMdRAE (Proxy 3)				
	CUC	FAST	JADE	SOBI	PCA	CUC	FAST	JADE	SOBI	PCA	CUC	FAST	JADE	SOBI	PCA
$r = 1$	0.721	0.745	0.752	0.723	1.000	0.724	0.750	0.759	0.728	1.000	0.733	0.751	0.760	0.726	1.000
$r = 2$	0.818	0.795	0.685	0.755	1.000	0.822	0.799	0.693	0.761	1.000	0.820	0.802	0.696	0.752	1.000
$r = 3$	0.938	0.808	0.668	0.771	1.000	0.939	0.809	0.673	0.774	1.000	0.934	0.814	0.675	0.762	1.000
$r = 4$	0.937	0.842	0.654	0.829	1.000	0.937	0.844	0.657	0.833	1.000	0.935	0.840	0.659	0.821	1.000
$r = 5$	0.985	0.863	0.688	0.836	1.000	0.985	0.863	0.690	0.838	1.000	0.983	0.865	0.694	0.833	1.000

Table 11
Comparison of the overall forecasting performance of the ARMA-GARCH modelling approach. The entries represent the average values of the RelMdRAE criterion, measured over the 19 stock returns, when a univariate ARMA-GARCH model is fitted to each component.

	Gaussian distribution														
	RelMdRAE (Proxy 1)					RelMdRAE (Proxy 2)					RelMdRAE (Proxy 3)				
	CUC	FAST	JADE	SOBI	PCA	CUC	FAST	JADE	SOBI	PCA	CUC	FAST	JADE	SOBI	PCA
$r = 1$	0.729	0.750	0.754	0.731	1.000	0.732	0.755	0.760	0.735	1.000	0.750	0.750	0.760	0.728	1.000
$r = 2$	0.842	0.799	0.693	0.769	1.000	0.843	0.801	0.698	0.773	1.000	0.799	0.799	0.703	0.757	1.000
$r = 3$	0.968	0.814	0.676	0.780	1.000	0.970	0.817	0.682	0.783	1.000	0.818	0.818	0.677	0.774	1.000
$r = 4$	0.881	0.773	0.596	0.753	1.000	0.883	0.776	0.598	0.758	1.000	0.776	0.777	0.603	0.754	1.000
$r = 5$	0.920	0.792	0.630	0.761	1.000	0.923	0.792	0.631	0.762	1.000	0.792	0.794	0.633	0.758	1.000
	Student's t distribution														
	RelMdRAE (Proxy 1)					RelMdRAE (Proxy 2)					RelMdRAE (Proxy 3)				
	CUC	FAST	JADE	SOBI	PCA	CUC	FAST	JADE	SOBI	PCA	CUC	FAST	JADE	SOBI	PCA
$r = 1$	0.735	0.766	0.769	0.743	1.000	0.740	0.772	0.777	0.748	1.000	0.743	0.768	0.776	0.743	1.000
$r = 2$	0.831	0.822	0.702	0.767	1.000	0.831	0.824	0.709	0.773	1.000	0.830	0.826	0.713	0.763	1.000
$r = 3$	0.944	0.830	0.685	0.780	1.000	0.946	0.835	0.691	0.785	1.000	0.942	0.838	0.688	0.774	1.000
$r = 4$	0.939	0.786	0.610	0.753	1.000	0.940	0.788	0.613	0.755	1.000	0.935	0.788	0.614	0.752	1.000
$r = 5$	0.992	0.813	0.647	0.764	1.000	0.992	0.815	0.650	0.769	1.000	0.988	0.813	0.651	0.762	1.000
	GED distribution														
	RelMdRAE (Proxy 1)					RelMdRAE (Proxy 2)					RelMdRAE (Proxy 3)				
	CUC	FAST	JADE	SOBI	PCA	CUC	FAST	JADE	SOBI	PCA	CUC	FAST	JADE	SOBI	PCA
$r = 1$	0.721	0.745	0.752	0.724	1.000	0.724	0.750	0.759	0.729	1.000	0.733	0.751	0.760	0.727	1.000
$r = 2$	0.823	0.799	0.689	0.756	1.000	0.825	0.802	0.696	0.761	1.000	0.822	0.803	0.699	0.751	1.000
$r = 3$	0.942	0.809	0.671	0.770	1.000	0.942	0.811	0.675	0.771	1.000	0.938	0.815	0.677	0.763	1.000
$r = 4$	0.862	0.774	0.603	0.749	1.000	0.862	0.775	0.604	0.752	1.000	0.865	0.775	0.608	0.748	1.000
$r = 5$	0.910	0.797	0.635	0.756	1.000	0.910	0.798	0.637	0.758	1.000	0.906	0.795	0.639	0.754	1.000

Table 12
Components of the IBEX 35 from 2000 to 2004, classified by sectors.

Consumption		
Other goods of consumption	ALT	Altadis
Consumption services		
Leisure time/Tourism/Hotel industry	AMS	Amadeus
	NHH	NH Hoteles
Mass media/Publicity	SGC	Sogecable
	TPI	Telefónica Publicidad e Información
Financial services/Estate agencies		
Banking	BBVA	Banco Bilbao Vizcaya Argentaria
	BKT	Bankinter
	POP	Banco Popular
	SAN	Banco Santander Central Hispano[a]
Oil and energy		
Oil	REP	Repsol
Electricity and gas	ELE	Endesa
	IBE	Iberdrola
Materials/Industry/Building		
Minerals / Metals	ACX	Acerinos
Building	ACS	Grupo ACS
	ANA	Acciona
	FCC	Fomento de Construcciones y Contratas S.A.
	FER	Grupo Ferrovial
Technology/Telecommunications		
Telecommunications and others	TEF	Telefónica
Electronic and software	TPI	Indra

[a] Known as SCH from 01/01/2000 to 31/10/2001.

by fitting univariate GARCH(1, 1) processes to each IC, CUC, and PC.

The estimates of the parameters when we fit a univariate model to each component are shown in the Appendix (see Table 13 for the GARCH(1, 1) specifications and Tables 14–16 for the ARMA-GARCH specifications). From these four tables, we can see that the GARCH parameters are significant for both the GARCH(1, 1) and ARMA-GARCH approaches, and for the three conditional distributions. This then indicates the time-varying volatility phenomenon of the components. Moreover, from Tables 14–16, we can see that the ARMA parameters are also statistically significant. Thus, it seems that fitting a univariate ARMA model to the conditional mean of the components is reasonable. This result is corroborated by the fact that the values of the likelihood function for the ARMA-GARCH models are larger than the corresponding ones for the GARCH(1, 1) specifications. Moreover, the values of the likelihood under the assumption of conditional Student's t innovations are the largest ones (and the GED distribution outperforms the Gaussian one). Under the Student's t distribution, the degrees of freedom parameter ν is very similar for the two approaches. According to both the GARCH(1, 1) and ARMA-GARCH specifications, the estimates for ν vary from 5.12 to 32.73, indicating heavy tails and excess kurtosis. A similar result is obtained with the shape parameter of the GED distribution, which varies from 1.29 to 1.93. According to previous conclusions, the ARMA-GARCH specifications with conditional Student's t innovations seem to provide the most appropriate approach to fitting the underlying conditionally heteroskedastic components.

To evaluate the forecasting performances of the GICA-GARCH, CUC-GARCH, and O-GARCH models, we take into account the two modelling approaches mentioned before. Moreover, in order to analyze the effect of increasing the number of components, when we evaluate the forecasting performances of the three models, we vary r from 1 to 5. The average results of the RelMdRAE, measured over the 19 stock returns, are displayed in Tables 10 (GARCH(1, 1) specifications) and 11 (ARMA-GARCH processes). To avoid having the choice of the proxy affecting our evaluation, we compute the RelMdRAE criterion using the three proxies proposed by Hansen and Lunde (2006). From Tables 10 and 11, we can see that, due to the use of a relative measure, the RelMdRAE, the values of the criterion do not differ very much for the different proxies. For both the GARCH(1, 1) and ARMA-GARCH modelling approaches, we obtain robust results, and JADE is chosen as the best method for estimating the underlying components, independently of the proxy and conditional distribution we use.

Tables 10 and 11 also show that the values of the RelMdRAE criterion are smaller when we adopt the ARMA-GARCH modelling approach, assuming conditional GED innovations. Then, it seems that the GICA-GARCH model with the underlying components estimated by JADE, and modelled according to univariate ARMA-GARCH models, produces the best forecasting performance. Furthermore, note that, independently of our scenario, all of the ICA algorithms perform better than CUC and PCA. Therefore, the GICA-GARCH model seems to be a good method for

Table 13
Estimates of the parameters when a univariate GARCH(1, 1) model is fitted to each component. The standard errors are in parentheses.

		GARCH parameter estimates (Gaussian)				GARCH parameter estimates (Student's t)					GARCH parameter estimates (GED)				
		α_0	α_1	β_1	L	α_0	α_1	β_1	ν	L	α_0	α_1	β_1	κ	L
CUC	1st	0.01227	0.07732	0.90871	−1285.70	0.00478	0.06068	0.93429	10.89320	−1277.02	0.00761	0.06785	0.92379	1.61184	−1279.61
		(0.00011)	(0.00111)	(0.00169)		(0.00002)	(0.00034)	(0.00045)	(1.25590)		(0.00005)	(0.00058)	(0.00081)	(0.01417)	
	2nd	0.01332	0.13065	0.85452	−1115.61	0.01010	0.11359	0.87552	9.41743	−1106.76	0.01172	0.12276	0.86422	1.51608	−1106.29
		(0.00008)	(0.00207)	(0.00245)		(0.00005)	(0.00158)	(0.00189)	(1.58004)		(0.00007)	(0.00180)	(0.00217)	(0.00996)	
	3rd	0.02650	0.10639	0.86589	−1298.55	0.01442	0.08448	0.90113	8.24801	−1285.39	0.02008	0.09770	0.88168	1.49041	−1286.95
		(0.00053)	(0.00255)	(0.00480)		(0.00026)	(0.00246)	(0.00390)	(1.36669)		(0.00041)	(0.00264)	(0.00463)	(0.00945)	
	4th	0.00533	0.05540	0.93840	−1237.02	0.00535	0.05506	0.93784	13.72002	−1225.95	0.00489	0.05557	0.93874	1.63661	−1230.79
		(0.00001)	(0.00014)	(0.00016)		(0.00001)	(0.00018)	(0.00020)	(1.48350)		(0.00001)	(0.00015)	(0.00016)	(0.03297)	
	5th	0.00680	0.04572	0.94533	−1242.28	0.00544	0.04100	0.95196	8.29869	−1227.34	0.00606	0.04411	0.94794	1.47294	−1229.22
		(0.00005)	(0.00046)	(0.00077)		(0.00006)	(0.00071)	(0.00120)	(3.85762)		(0.00005)	(0.00051)	(0.00085)	(0.01028)	
FAST	1st	0.00524	0.08521	0.91036	−1266.91	0.00643	0.08987	0.90446	18.07498	−1264.81	0.00603	0.08760	0.90700	1.72765	−1264.54
		(−0.00002)	(−0.00057)	(−0.00061)		(0.00002)	(0.00059)	(0.00064)	(0.67388)		(0.00002)	(0.00058)	(0.00063)	(0.01239)	
	2nd	0.01685	0.05415	0.91887	−1153.00	0.02416	0.06498	0.89605	9.14226	−1140.59	0.02094	0.05994	0.90615	1.53997	−1143.29
		(0.00039)	(0.00111)	(0.00393)		(0.00133)	(0.00277)	(0.01202)	(0.63534)		(0.00073)	(0.00174)	(0.00697)	(0.01250)	
	3rd	0.00360	0.04397	0.95215	−1251.49	0.00187	0.02219	0.97463	5.12131	−1212.83	0.00240	0.02994	0.96599	1.30333	−1219.60
		(0.00005)	(0.00118)	(0.00162)		(0.00001)	(0.00030)	(0.00038)	(0.57865)		(0.00003)	(0.00080)	(0.00110)	(0.00741)	
	4th	0.00279	0.04287	0.95329	−1273.14	0.00114	0.03507	0.96295	22.14731	−1271.33	0.00224	0.04022	0.95654	1.84167	−1272.33
		(0.00002)	(0.00051)	(0.00063)		(0.00001)	(0.00029)	(0.00034)	(1.33957)		(0.00001)	(0.00042)	(0.00050)	(0.01534)	
	5th	0.00401	0.04428	0.94883	−1081.87	0.00445	0.04691	0.94566	5.88473	−1050.94	0.00392	0.04389	0.94887	1.33466	−1055.15
		(0.00001)	(0.00043)	(0.00062)		(0.00002)	(0.00068)	(0.00090)	(1.09484)		(0.00001)	(0.00043)	(0.00060)	(0.00798)	
JADE	1st	0.00346	0.07403	0.92268	−1260.39	0.00288	0.07342	0.92418	32.73407	−1259.62	0.00329	0.07367	0.92328	1.92725	−1260.22
		(0.00001)	(0.00042)	(0.00046)		(0.00001)	(0.00042)	(0.00046)	(3.91472)		(0.00001)	(0.00042)	(0.00046)	(0.01528)	
	2nd	0.00136	0.02519	0.97268	−1327.85	0.00118	0.02878	0.96949	11.61326	−1320.81	0.00132	0.02677	0.97118	1.65921	−1323.07
		(0.00001)	(0.00010)	(0.00012)		(0.00001)	(0.00015)	(0.00018)	(3.32136)		(0.00001)	(0.00011)	(0.00014)	(0.01176)	
	3rd	0.00089	0.02107	0.97717	−1245.98	0.00342	0.03006	0.96380	6.88107	−1203.32	0.00273	0.02901	0.96645	1.37756	−1216.31
		(0.00001)	(0.00064)	(0.00084)		(0.00001)	(0.00011)	(0.00016)	(2.41987)		(0.00001)	(0.00023)	(0.00031)	(0.01811)	
	4th	0.00124	0.00791	0.98948	−1184.77		0.01036	0.98901	8.83368	−1160.25	0.00017	0.00845	0.99066	1.49322	−1169.71
		(0.00001)	(0.00001)	(0.00003)		(0.00000)	(0.00001)	(0.00001)	(1.57673)		(0.00000)	(0.00001)	(0.00002)	(0.01734)	
	5th	0.00000	0.12415	0.87585	−1257.95	0.00827	0.08109	0.91028	12.65770	−1242.84	0.00912	0.09136	0.89997	1.68202	−1244.06
		(0.00000)	(0.00095)	(0.00098)		(0.00003)	(0.00097)	(0.00108)	(1.61714)		(0.00003)	(0.00091)	(0.00097)	(0.01074)	
SOBI	1st	0.00333	0.04761	0.94871	−1278.83	0.00259	0.04569	0.95154	12.43569	−1273.65	0.00292	0.04726	0.94957	1.61668	−1273.50
		(0.00001)	(0.00021)	(0.00024)		(0.00001)	(0.00018)	(0.00022)	(2.04489)		(0.00001)	(0.00019)	(0.00023)	(0.01247)	
	2nd	0.01170	0.09669	0.89109	−1248.43	0.00292	0.05750	0.93899	10.41371	−1234.26	0.00513	0.06952	0.92511	1.58657	−1240.27
		(0.00035)	(0.00796)	(0.01040)		(0.00001)	(0.00053)	(0.00059)	(1.93100)		(0.00005)	(0.00198)	(0.00231)	(0.01882)	
	3rd	0.01962	0.07834	0.89444	−1209.84	0.01224	0.07779	0.90655	13.00792	−1202.14	0.01676	0.07808	0.89916	1.67812	−1205.51
		(0.00014)	(0.00030)	(0.00068)		(0.00005)	(0.00033)	(0.00050)	(3.15802)		(0.00009)	(0.00031)	(0.00057)	(0.01911)	
	4th	0.00378	0.04269	0.95273	−1241.83	0.00351	0.03986	0.95514	7.43635	−1217.77	0.00314	0.03855	0.95707	1.38849	−1220.54
		(0.00013)	(0.00478)	(0.00629)		(0.00001)	(0.00045)	(0.00058)	(3.14261)		(0.00002)	(0.00074)	(0.00096)	(0.01457)	

Table 13 (*continued*)

		GARCH parameter estimates (Gaussian)				GARCH parameter estimates (Student's t)					GARCH parameter estimates (GED)				
		α_0	α_1	β_1	L	α_0	α_1	β_1	ν	L	α_0	α_1	β_1	κ	L
	5th	0.00642	0.10249	0.88864	−1057.76	0.00570	0.09197	0.90013	9.35770	−1049.60	0.00639	0.09734	0.89325	1.49605	−1047.64
		(0.00002)	(0.00088)	(0.00097)		(0.00002)	(0.00099)	(0.00110)	(3.81094)		(0.00002)	(0.00091)	(0.00102)	(0.00864)	
PCST	1st	0.01532	0.11528	0.87109	−1281.88	0.00898	0.09332	0.89955	11.26723	−1275.69	0.01232	0.10488	0.88439	1.63892	−1277.09
		(0.00005)	(0.00086)	(0.00083)		(0.00003)	(0.00054)	(0.00054)	(1.22038)		(0.00004)	(0.00067)	(0.00065)	(0.01142)	
	2nd	0.01142	0.09827	0.88501	−1127.17	0.00983	0.08571	0.89938	12.46776	−1122.40	0.01079	0.09356	0.89044	1.60094	−1121.45
		(0.00005)	(0.00105)	(0.00142)		(0.00004)	(0.00089)	(0.00124)	(1.67765)		(0.00005)	(0.00099)	(0.00136)	(0.01073)	
	3rd	0.00249	0.05268	0.94404	−1167.42	0.00174	0.05147	0.94655	12.45832	−1162.00	0.00202	0.05324	0.94440	1.63752	−1162.63
		(0.00001)	(0.00035)	(0.00039)		(0.00001)	(0.00027)	(0.00028)	(1.68052)		(0.00001)	(0.00032)	(0.00034)	(0.01256)	
	4th	0.00498	0.06273	0.93065	−1172.73	0.00757	0.08404	0.90725	9.52976	−1163.11	0.00636	0.07343	0.91853	1.54978	−1164.24
		(0.00004)	(0.00122)	(0.00172)		(0.00005)	(0.00156)	(0.00209)	(0.65577)		(0.00005)	(0.00140)	(0.00197)	(0.01008)	
	5th	0.00756	0.08021	0.91151	−1174.15	0.00241	0.04144	0.95541	7.51165	−1149.85	0.00269	0.03848	0.95715	1.41907	−1156.06
		(0.00015)	(0.00845)	(0.00959)		(0.00001)	(0.00143)	(0.00150)	(0.40320)		(0.00001)	(0.00077)	(0.00086)	(0.01282)	

Table 14
Estimates of the parameters when a univariate ARMA-GARCH model, with conditionally Gaussian innovations, is fitted to each component. The standard errors are in parentheses.

		Conditional mean estimates (Gaussian)			GARCH parameter estimates (Gaussian)					
		ϕ_1	ϕ_2	θ_1	α_0	α_1	β_1	α_2	β_2	L
CUC	1st	–	–	–	0.01227	0.07732	0.90871	–	–	−1285.70
					(0.00011)	(0.00111)	(0.00169)			
	2nd	–	–	–	0.01332	0.13065	0.85452	–	–	−1115.61
					(0.00008)	(0.00207)	(0.00245)			
	3rd	–	–	–	0.02650	0.10639	0.86589	–	–	−1298.55
					(0.00053)	(0.00255)	(0.00480)			
	4th	–	–	–	0.00533	0.05540	0.93840	–	–	−1237.02
					(0.00001)	(0.00014)	(0.00016)			
	5th	–	–	–	0.00680	0.04572	0.94533	–	–	−1242.28
					(0.00005)	(0.00046)	(0.00077)			
FAST	1st	–	–	–	0.00524	0.08521	0.91036	–	–	−1266.91
					(−0.00002)	(−0.00057)	(−0.00061)			
	2nd	−0.17807	−0.10131	–	0.01387	0.04792	0.92942	–	–	−1137.32
		(0.03873)	(0.01395)		(0.00014)	(0.00048)	(0.00149)			
	3rd	–	–	–	0.00360	0.04397	0.95215	–	–	−1251.49
					(0.00005)	(0.00118)	(0.00162)			
	4th	–	–	–	0.00421	0.04991	0.02731	0.00000	0.91670	−1272.69
					(0.00006)	(0.00052)	(0.00063)	(0.00000)	(0.00222)	
	5th	−0.22140	–	–	0.00411	0.04469	0.94823	–	–	−1080.46
		(0.01912)			(0.00002)	(0.00043)	(0.00064)			
JADE	1st	–	–	–	0.00346	0.07403	0.92268	–	–	−1260.39
					(0.00001)	(0.00042)	(0.00046)			
	2nd	–	–	–	0.00136	0.02519	0.97268	–	–	−1327.85
					(0.00001)	(0.00010)	(0.00012)			
	3rd	0.26230	–	−0.3133	0.00095	0.02224	0.97604	–	–	−1239.33
		(0.00939)		(0.02586)	(0.00001)	(0.00054)	(0.00070)			
	4th	–	–	–	0.00124	0.00791	0.98948	–	–	−1184.77
					(0.00001)	(0.00001)	(0.00003)			
	5th	–	–	–	0.00000	0.12415	0.87585	–	–	−1257.95
					(0.00000)	(0.00095)	(0.00098)			
SOBI	1st	–	−0.19241	–	0.00340	0.04671	0.94937	–	–	−1275.42
			(0.00558)		(0.00001)	(0.00019)	(0.00024)			
	2nd	0.66035	–	−0.694639071	0.00988	0.08780	0.90194	–	–	−1242.97
		(0.10792)		(0.06644)	(0.00036)	(0.00854)	(0.01137)			
	3rd	–	–	–	0.01962	0.07834	0.89444	–	–	−1209.84
					(0.00014)	(0.00030)	(0.00068)			
	4th	–	−0.25338	–	0.00094	0.02297	0.97499	–	–	−1217.08
			(0.03571)		(0.00001)	(0.00086)	(0.00105)			
	5th	0.13886	–	–	0.00443	0.08276	0.90974	–	–	−1051.31
		(0.01712)			(0.00001)	(0.00064)	(0.00073)			
PCST	1st	–	–	–	0.01532	0.11528	0.87109	–	–	−1281.88
					(0.00005)	(0.00086)	(0.00083)			
	2nd	0.10370	–	–	0.01046	0.09091	0.89308	–	–	−1123.96
		(0.00849)			(0.00005)	(0.00141)	(0.00185)			
	3rd	−0.27392	–	–	0.00248	0.05245	0.94428	–	–	−1165.67
		(0.04905)			(0.00001)	(0.00035)	(0.00038)			
	4th	−0.50852	−0.12906	–	0.54059	0.28527	–	–	–	−1235.24
		(0.07742)	(0.01737)		(0.00213)	(0.00425)	()			
	5th	–	–	–	0.00756	0.08021	0.91151	–	–	−1174.15
					(0.00015)	(0.00845)	(0.00959)			

forecasting the conditional covariance matrix of large datsets.[1]

6. Concluding remarks

We have proposed a new framework for modelling and forecasting large conditional covariance matrices of stock returns using a few underlying factors with conditional heteroskedasticity. Our model, called the GICA-GARCH model, assumes that the co-movements of a vector of financial data are driven by a few independent components which evolve according to univariate ARMA-GARCH models. In our model, the conditional covariance matrix of the factors is assumed to be diagonal. Therefore, the GICA-GARCH provides a parsimonious representation of the conditional covariance matrix of the data, and reduces the number of parameters to be estimated. Our estimation procedure consists of two steps: in the first step, we exploit the unconditional distribution of the data

[1] Evaluating the forecasting performance of the model using the Relative Geometric Mean Relative Absolute Error (RelGMRAE) gives similar results, which are available from the authors upon request.

Table 15
Estimates of the parameters when a univariate ARMA-GARCH model, with conditional Student's t innovations, is fitted to each component. The standard errors are in parentheses.

		Conditional mean estimates (Student's t)			GARCH parameter estimates (Student's t)						
		ϕ_1	ϕ_2	θ_1	α_0	α_1	β_1	α_2	β_2	ν	L
CUC	1st	–	–	–	0.00478	0.06068	0.93429	–	–	10.89320	−1277.02
					(0.00002)	(0.00034)	(0.00045)			(1.25590)	
	2nd	–	–	–	0.01010	0.11359	0.87552	–	–	9.41743	−1106.76
					(0.00005)	(0.00158)	(0.00189)			(1.58004)	
	3rd	–	–	–	0.01442	0.08448	0.90113	–	–	8.24801	−1285.39
					(0.00026)	(0.00246)	(0.00390)			(1.36669)	
	4th	–	–	–	0.00535	0.05506	0.93784	–	–	13.72002	−1225.95
					(0.00001)	(0.00018)	(0.00020)			(1.48350)	
	5th	–	–	–	0.00544	0.04100	0.95196	–	–	8.29869	−1227.34
					(0.00006)	(0.00071)	(0.00120)			(3.85762)	
FAST	1st	–	–	–	0.00643	0.08987	0.90446	–	–	18.07498	−1264.81
					(0.00002)	(0.00059)	(0.00064)			(0.67388)	
	2nd	−0.16450	−0.09279	–	0.01685	0.04987	0.92217	–	–	8.89697	−1124.68
		(0.02138)	(0.03832)		(0.00045)	(0.00105)	(0.00437)			(1.53093)	
	3rd	–	–	–	0.00187	0.02219	0.97463	–	–	5.12131	−1212.83
					(0.00001)	(0.00030)	(0.00038)			(0.57865)	
	4th	–	–	–	0.00146	0.05707	0.00000	0.16346	0.77648	20.89172	−1270.65
					(0.00002)	(0.00064)	(0.00000)	(0.01313)	(0.01492)	(1.31786)	
	5th	−0.20820	–	–	0.00467	0.04830	0.94390	–	–	5.90092	−1049.51
		(0.02066)			(0.00002)	(0.00075)	(0.00101)			(1.10583)	
	1st	–	–	–	0.00288	0.07342	0.92418	–	–	32.73407	−1259.62
JADE					(0.00001)	(0.00042)	(0.00046)			(3.91472)	
	2nd	–	–	–	0.00118	0.02878	0.96949	–	–	11.61326	−1320.81
					(0.00001)	(0.00015)	(0.00018)			(3.32136)	
	3rd	0.24290	–	−0.31926	0.00308	0.02822	0.96609	–	–	6.70582	−1195.76
		(0.01905)		(0.02896)	(0.00001)	(0.00010)	(0.00013)			(2.16043)	
	4th	–	–	–		0.01036	0.98901	–	–	8.83368	−1160.25
					(0.00000)	(0.00001)	(0.00001)			(1.57673)	
	5th	–	–	–	0.00827	0.08109	0.91028	–	–	12.65770	−1242.84
					(0.00003)	(0.00097)	(0.00108)			(1.61714)	
SOBI	1st	–	−0.21951	–	0.00272	0.04557	0.95143	–	–	13.29084	−1271.02
			(0.06128)		(0.00001)	(0.00017)	(0.00021)			(2.51943)	
	2nd	0.66536	–	−0.68480		0.07030	0.92970	–	–	29.23345	−1235.73
		(0.09307)		(0.05207)	(0.00000)	(0.00123)	(0.00108)			(3.06391)	
	3rd	–	–	–	0.01224	0.07779	0.90655	–	–	13.00792	−1202.14
					(0.00005)	(0.00033)	(0.00050)			(3.15802)	
	4th	–	−0.24597	–	0.00243	0.03053	0.96535	–	–	6.60787	−1189.80
			(0.04729)		(0.00001)	(0.00042)	(0.00057)			(1.90747)	
	5th	0.13256	–	–	0.00436	0.08017	0.91330	–	–	9.64494	−1040.07
		(0.01894)			(0.00001)	(0.00069)	(0.00078)			(2.37447)	
PCST	1st	–	–	–	0.00898	0.09332	0.89955	–	–	11.26723	−1275.69
					(0.00003)	(0.00054)	(0.00054)			(1.22038)	
	2nd	0.11600	–	–	0.00769	0.07146	0.91625	–	–	11.32251	−1116.85
		(0.00906)			(0.00003)	(0.00060)	(0.00082)			(1.37506)	
	3rd	−0.27877	–	–	0.00175	0.05163	0.94643	–	–	12.37339	−1160.80
		(0.04874)			(0.00001)	(0.00027)	(0.00029)			(1.65436)	
	4th	−0.51948	−0.12175	–	0.52641	0.36565	–	–	–	5.23611	−1207.89
		(0.07128)	(0.01586)		(0.00223)	(0.00562)	()			(0.68776)	
	5th	–	–	–	0.00241	0.04144	0.95541	–	–	7.51165	−1149.85

in order to estimate the ICs, sort them in terms of their variability and disentagle the common and idiosyncratic components of the financial data; in the second step, we estimate the conditional covariance matrix of the data as a linear combination of the conditional variances of the common components, which are modelled according to univariate ARMA-GARCH models.

The advantage of the GICA-GARCH model with respect to the existing literature lies in the potential of ICA to identify the underlying components of a vector of financial data. In this paper, we have proposed three simulation experiments to test the potential of ICA (using three different algorithms), CUC, and PCA to identify the conditionally heteroskedastic components when they have different excess kurtosis. We have analyzed the performances of the three models in terms of both the correlation coefficients and the mean square errors between each original component and its estimation. The results show that, regardless of whether the excess kurtosis comes from different GARCH specifications or from different conditional distributions, the ICA methods perform better than either CUC or PCA for identifying the conditionally heteroskedastic components. Furthermore, the results for the ICA algorithms are as expected: both

Table 16
Estimates of the parameters when a univariate ARMA-GARCH model, with conditionally GED innovations, is fitted to each component. The standard errors are in parentheses.

		Conditional mean estimates (GED)			GARCH parameter estimates (GED)						
		ϕ_1	ϕ_2	θ_1	α_0	α_1	β_1	α_2	β_2	κ	L
CUC	1st	–	–	–	0.00761	0.06785	0.92379	–	–	1.61184	−1279.61
					(0.00005)	(0.00058)	(0.00081)			(0.01417)	
	2nd	–	–	–	0.01172	0.12276	0.86422	–	–	1.51608	−1106.29
					(0.00007)	(0.00180)	(0.00217)			(0.00996)	
	3rd	–	–	–	0.02008	0.09770	0.88168	–	–	1.49041	−1286.95
					(0.00041)	(0.00264)	(0.00463)			(0.00945)	
	4th	–	–	–	0.00489	0.05557	0.93874	–	–	1.63661	−1230.79
					(0.00001)	(0.00015)	(0.00016)			(0.03297)	
	5th	–	–	–	0.00606	0.04411	0.94794	–	–	1.47294	−1229.22
					(0.00005)	(0.00051)	(0.00085)			(0.01028)	
FAST	1st	–	–	–	0.00603	0.08760	0.90700	–	–	1.72765	−1264.54
					(0.00002)	(0.00058)	(0.00063)			(0.01239)	
	2nd	−0.18356	−0.11151	–	0.01575	0.04940	0.92441	–	–	1.53683	−1127.43
		(0.05594)	(0.01395)		(0.00024)	(0.00064)	(0.00235)			(0.01162)	
	3rd	–	–	–	0.00240	0.02994	0.96599	–	–	1.30333	−1219.60
					(0.00003)	(0.00080)	(0.00110)			(0.00741)	
	4th	–	–	–	0.00360	0.06065	0.00602	0.16333	0.76453	1.83612	−1271.89
					(0.00004)	(0.00077)	(0.00065)	(0.01029)	(0.01134)	(0.01527)	
	5th	−0.21700	–	–	0.00410	0.04477	0.94767	–	–	1.33268	−1053.63
		(0.01551)			(0.00001)	(0.00046)	(0.00065)			(0.00807)	
JADE	1st	–	–	–	0.00329	0.07367	0.92328	–	–	1.92725	−1260.22
					(0.00001)	(0.00042)	(0.00046)			(0.01528)	
	2nd	–	–	–	0.00132	0.02677	0.97118	–	–	1.65921	−1323.07
					(0.00001)	(0.00011)	(0.00014)			(0.01176)	
	3rd	0.25026	–	−0.30598	0.00243	0.02743	0.96836	–	–	1.37199	−1208.88
		(0.02849)		(0.02275)	(0.00001)	(0.00017)	(0.00023)			(0.01695)	
	4th	–	–	–	0.00017	0.00845	0.99066	–	–	1.49322	−1169.71
					(0.00000)	(0.00001)	(0.00002)			(0.01734)	
	5th	–	–	–	0.00912	0.09136	0.89997	–	–	1.68202	−1244.06
					(0.00003)	(0.00091)	(0.00097)			(0.01074)	
SOBI	1st	–	−0.20611	–	0.00297	0.04644	0.95022	–	–	1.63470	−1270.72
			(0.06716)		(0.00001)	(0.00018)	(0.00023)			(0.01239)	
	2nd	0.64005	–	−0.68167	0.00459	0.06619	0.92914	–	–	1.60530	−1235.73
		(0.08910)		(0.05644)	(0.00004)	(0.00162)	(0.00188)			(0.02064)	
	3rd	–	–	–	0.01676	0.07808	0.89916	–	–	1.67812	−1205.51
					(0.00009)	(0.00031)	(0.00057)			(0.01911)	
	4th	–	−0.26120	–	0.00191	0.02850	0.96796	–	–	1.38617	−1194.25
			(0.06143)		(0.00002)	(0.00085)	(0.00112)			(0.01050)	
	5th	0.13034	–	–	0.00468	0.08160	0.91062	–	–	1.50754	−1040.56
		(0.01330)			(0.00001)	(0.00065)	(0.00075)			(0.01259)	
PCST	1st	–	–	–	0.01232	0.10488	0.88439	–	–	1.63892	−1277.09
					(0.00004)	(0.00067)	(0.00065)			(0.01142)	
	2nd	0.10278	–	–	0.00902	0.08209	0.90390	–	–	1.56961	−1116.80
		(0.00911)			(0.00004)	(0.00098)	(0.00131)			(0.01197)	
	3rd	−0.27098	–	–	0.00201	0.05319	0.94448	–	–	1.63389	−1160.80
		(0.00509)			(0.00001)	(0.00033)	(0.00035)			(0.01251)	
	4th	−0.52010	−0.12053	–	0.51991	0.32528	–	–	–	1.29398	−1207.77
		(0.06942)	(0.01486)		(0.00189)	(0.00445)	()			(0.00551)	
	5th	–	–	–	0.00269	0.03848	0.95715	–	–	1.41907	−1156.06
					(0.00001)	(0.00077)	(0.00086)			(0.01282)	

FastICA and JADE, which estimate the ICs by maximizing their non-Gaussianity, capture the excess kurtosis of the conditionally heteroskedastic factors better than SOBI. Therefore, the GICA-GARCH model seems to provide a more reliable identification of the unobserved components than either the O-GARCH or CUC-GARCH models.

We have tested the GICA-GARCH model empirically on a vector of stock returns of the Madrid stock market. After applying the three ICA algorithms to the identification of the unobserved components and fitting a univariate model to each one of them, the empirical results show that the most appropriate specification for fitting each IC is the ARMA-GARCH model with conditional Student's *t* innovations. Furthermore, as accurate volatility forecasts are a crucial issue, we have evaluated the forecasting performance of our model. We have implemented a rolling window scheme to compare the relative ability to predict the one-step-ahead volatility of the GICA-GARCH, CUC-GARCH, and O-GARCH models. In terms of the average RelMdRAE results, and independently of the proxy used to substitute the real volatility, our model provides more accurate volatility forecasts than either the CUC-GARCH or O-GARCH models for the stock returns of the IBEX 35 index. In particular, according to the empirical results, the volatility forecasts obtained using the JADE algorithm are more accurate than those generated by using any other ICA algorithm.

Designing an alternative procedure for sorting the ICs and choosing the optimal number of factors may be challenges for the future. Moreover, we are interested in comparing the performance of our model with the perfomances of other multivariate GARCH models, such as the dynamic factor GARCH, and extending the GICA-GARCH model to other applications.

Acknowledgements

Financial support from Projects SEJ2006-04957, SEJ2007-64500, and ECO2009-10287 by the Spanish Government (Ministerio de Educación y Ciencia) is gratefully acknowledged. We thank Pilar Poncela, the anonymous referees, and the editors, Esther Ruiz and Nuno Crato, for helpful comments and suggestions. The responsibility for any errors remains ours.

Appendix

See Tables 12–16.

References

Alessi, L., Barigozzi, M., & Capasso, M. (2006). *Dynamic factor GARCH: multivariate volatility forecast for a large number of series*. LEM papers series 2006/25. Laboratory of Economics and Management. LEM. Sant'Anna School of Advanced Studies. Pisa, Italy.

Alexander, C. (2001). *Market models: a guide to financial data analysis*. Chichester, UK: Wiley.

Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review*, *39*, 885–905.

Back, A. D., & Weigend, A. S. (1997). A first application of independent component analysis to extracting structure from stock returns. *International Journal of Neural Systems*, *8*, 473–484.

Bauwens, L., Laurent, S., & Rombouts, J. V. K. (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics*, *21*, 79–109.

Belouchrani, A., Abed Meraim, K., Cardoso, J.-F., & Moulines, E. (1997). A blind source separation technique based on second order statistics. *IEEE Transactions on Signal Processing*, *45*, 434–444.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, *31*, 307–327.

Bollerslev, T. (1987). A conditional heteroscedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, *69*, 542–547.

Bollerslev, T., & Engle, R. F. (1993). Common persistence in conditional variances. *Econometrica*, *61*, 167–186.

Bollerslev, T., Engle, R. F., & Wooldridge, J. W. (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, *96*, 116–131.

Cardoso, J.-F., & Souloumiac, A. (1993). Blind beamforming for non-Gaussian signals. *IEE Proceedings F (Communications, Radar and Signal Processing)*, *140*, 362–370.

Cha, S.-M., & Chan, L.-W. (2000). Applying independent component analysis to factor model in finance. In *Proceedings of the international symposium on intelligent data engineering and learning* (pp. 538–544).

Common, P. (1994). Independent component analysis—a new concept? *Signal Processing*, *36*, 287–314.

Diamantopoulos, K., & Vrontos, I. D. (2010). A Student-*t* full factor multivariate GARCH model. *Computational Economics*, *35*, 63–83.

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

Doz, C., & Renault, E. (2004). *Conditionally heteroskedastic factor models: identification and instrumental variables estimation*. CIRANO working paper.

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

Engle, R. F. (1987). *Multivariate ARCH with factor structures—cointegration in variance*. Discussion paper 87-2. UCSD. Department of Economics.

Engle, R. F., Granger, C. W. J., & Kraft, D. (1984). Combining competing forecasts of inflation using a bivariate ARCH model. *Journal of Economic Dynamics and Control*, *8*, 151–165.

Engle, R. F., & Kroner, K. (1995). Multivariate simultaneous GARCH. *Econometric Theory*, *11*, 122–150.

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

Engle, R. F., & Susmel, R. (1993). Common volatility in international equity markets. *Journal of Business and Economic Statistics*, *11*, 167–176.

Fan, J., Wang, M., & Yao, Q. (2008). Modelling multivariate volatilities via conditionally uncorrelated components. *Journal of the Royal Statistical Society, Series B*, *70*, 679–702.

Franses, P. H., & van Dijk, D. (1996). Forecasting stock market volatility using nonlinear GARCH models. *Journal of Forecasting*, *15*, 229–235.

Granger, C., & Ding, Z. (1995). Some properties of absolute return, an alternative measure of risk. *Annales d'Economie et de Statistique*, *40*, 67–91.

Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, *20*, 873–889.

Hansen, P. R., & Lunde, A. (2006). Consistent ranking of volatility models. *Journal of Econometrics*, *131*, 97–121.

Harvey, A., Ruiz, E., & Sentana, E. (1992). Unobserved component time series models with ARCH disturbances. *Journal of Econometrics*, *52*, 129–157.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*, 679–688.

Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, *10*, 626–634.

Hyvärinen, A., & Kano, K. (2003). Independent component analysis for non-normal factor analysis. In H. Yanai, et al., (Eds.), *New developments in psychometrics* (pp. 649–656). Tokyo: Springer Verlag.

Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent component analysis*. New York: Wiley.

Hyvärinen, A., & Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. *Neural Computation*, *9*, 1483–1492.

King, M. A., Sentana, E., & Wadhwani, S. B. (1994). Volatility and links between national stock markets. *Econometrica*, *62*, 901–933.

Kiviluoto, K., & Oja, E. (1998). Independent component analysis for parallel financial time series. In *Proceedings of the international conference on neural information processing. Tokyo. Vol. 2* (pp. 895–898).

Lanne, M., & Saikkonen, P. (2007). A multivariate generalized orthogonal factor GARCH model. *Journal of Business and Economic Statistics*, *25*, 61–75.

Mălăroiu, S., Kiviluoto, K., & Oja, E. (2000). *Time series prediction with independent component analysis*. Technical report. Helsinki University of Technology. Helsinki.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, *59*, 347–370.

Ng, V., Engle, R. F., & Rothschild, M. (1992). A multi-dynamic-factor model for stock returns. *Journal of Econometrics*, *52*, 245–266.

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

Peña, D., & Box, G. (1987). Identify a simplifying structure in time series. *Journal of the American Statistical Association*, *82*, 836–843.

Peña, D., & Poncela, P. (2006). Nonstationary dynamic factor analysis. *Journal of Statistical Planning and Inference*, *136*, 1237–1257.

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

Tong, L., Liu, R. -W., Soon, V. C., Huang, Y. -F., & Liu, R. (1990). AMUSE: a new blind identification algorithm. In *Procedings of the IEEE international symposium on circuits and systems. New Orleans* (pp. 1784–1787).

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

Vrontos, I. D., Dellaportas, P., & Politis, D. N. (2003). A full-factor multivariate GARCH model. *Econometrics Journal*, *6*, 311–333.

Wu, E. H. C., & Yu, P. L. H. (2005). Volatility modelling of multivariate financial time series by using ICA-GARCH models. In *Proceedings of the international symposium on intelligent data engineering and learning* (pp. 571–579).