

# ON THE UNIVARIATE REPRESENTATION OF BEKK MODELS WITH COMMON FACTORS

**Alain Hecq,**

Department of Quantitative Economics  
Maastricht University  
The Netherlands

**Sébastien Laurent**

Aix-Marseille University (Aix-Marseille School of Economics),  
CNRS & EHESS,  
Aix-Marseille Graduate School of Management – IAE,  
France

**Franz C. Palm \***

Department of Quantitative Economics  
Maastricht University  
The Netherlands

February 16, 2015

---

\*Corresponding author: Franz Palm, Department of Quantitative Economics, Maastricht University, P.O.Box 616, 6200 MD Maastricht, The Netherlands. E-mail: f.palm@maastrichtuniversity.nl. This paper is a revised version of our 2011 METEOR discussion paper "On the Univariate Representation of Multivariate Volatility Models with Common Factors". We would like to thank an anonymous referee for the useful comments on the previous version. The usual disclaimer applies.

## Abstract

Simple low order multivariate GARCH models imply marginal processes with a lot of persistence in the form of high order lags. This is not what we find in many situations however, where parsimonious univariate GARCH(1,1) models for instance describe quite well the conditional volatility of some asset returns. In order to explain this paradox, we show that in the presence of common GARCH factors, parsimonious univariate representations can result from large multivariate models generating the conditional variances and conditional covariances/correlations. The diagonal model without any contagion effects in conditional volatilities gives rise to similar conclusions though.

Consequently, after having extracted a block of assets representing some form of parsimony, remains the task of determining if we have a set of independent assets or instead a highly dependent system generated with a few factors. To investigate this issue, we first evaluate a reduced rank regressions approach for squared returns (Engle and Marcucci, 2006) that we extend to cross-returns. Second we investigate a likelihood ratio approach, where under the null the matrix parameters have a reduced rank structure (Lin, 1992). It emerged that the latter approach has quite good properties enabling us to discriminate between a system with seemingly unrelated assets (e.g. a diagonal model) and a model with few common sources of volatility.

JEL: C10, C32

Keywords: Common GARCH, factor models, BEKK, Final equations.

## 1 Motivation

Parsimonious univariate volatility models, such as the GARCH(1,1) model, are often detected in empirical work, especially in macro-finance. This paper shows that this feature might arise in the presence of few common factors generating the joint volatility of multiple asset returns. We get this result on volatilities using the final equation representation developed for VARMA models (see Zellner and Palm, 1974, 1975, 2004) under additional reduced rank restrictions (Cubadda *et al.*, 2009). This framework allows us to derive the marginal univariate GARCH representation for the conditional variances and conditional covariances of the multivariate GARCH model (e.g. Nijman and Sentana, 1996 in the unrestricted case) but in the presence of common GARCH factors. However, the simple diagonal multivariate model (Bollerslev *et al.*, 1988) for conditional volatilities delivers similar parsimonious marginal representations. In order to discriminate between those radically different multivariate models, we propose, after having detected a block of assets with a parsimonious univariate representation, to test for the existence of common volatility. As an example, among the bunch of fifty daily stock returns that we analyze in this paper, six of them are compatible with a GARCH(1,1) specification. Then we will investigate whether there is a factor structure behind this feature or simply an absence of causality from the past covariances to the variances (diagonal model).

Our common feature modeling as well as our testing strategy is different from two types of modeling commonly found in this literature. First, contrary to the usual factor models in finance (e.g. King and Wadhvani, 1990), only the common factor component (but not the idiosyncratic part) will be time-varying. In this sense we are closer to Diebold and Nerlove (1989) although in our framework the number of factors will be tested. Second, with the volatility being unobservable, we do not directly test using a LM framework for potential combinations of heteroskedastic series that are conditionally homoskedastic (see Engle and Susmel, 1993; Engle and Kozicki 1993, Ruiz, 2009, Arshanapalli *et al.* 1997). Instead we translate the task to detect the presence of common volatility in asset returns into an analysis of common cyclical features (Vahid and Engle, 1993) in the dynamics of the squared returns and cross-returns. Our framework can be seen as a generalization of the Engle and Marcucci (2006) pure variance model although they do not consider the covariances in their analysis. Indeed, Engle and Marcucci (2006) propose a theoretical framework in which the true model is a DCC (Engle, 2002) with a pure variance specification for the logarithms of conditional variances. They apply a canonical correlation procedure on the log squared returns and therefore assume their conditional normality. This strategy does not perform well however if the multivariate model is not an exponential GARCH with non Gaussian conditional errors. Moreover, also considering the cross-products in addition to the squared residuals makes the use of the logarithm transformation inadequate. In that respect, our paper also studies the small sample performance of the canonical correlation test on squared returns proposed by Engle and Marcucci (2006) when the data generating process is not the one in their study.

The common volatility factors extracted by our strategy in a multivariate GARCH framework and in particular the BEKK model (Baba, Engle, Kraft and Kroener, 1989)<sup>1</sup> are quite intuitive since they represent the conditional variance of a portfolio composed by the series involved in the analysis. Applying our common cycle approach and hence testing for the presence of commonalities in the *vec* or the *vecd* of the conditional covariance matrix is hazardous however. Indeed, even under the null, the combination that annihilates the temporal dependence in squared returns and cross-returns is a martingale difference sequence. We show that this leads to large size distortions if one uses canonical correlation test statistics as proposed in Engle and Marcucci (2006). Moreover, accounting for the presence of heteroskedasticity using a robustified version of the reduced rank tests similar to the one proposed by White for single equation analyses (Hecq and Issler, 2012) still produces size distortions because of the non-normality (and especially asymmetry) of the disturbances. It turns out that, although the representation of multivariate systems in terms of squared returns and cross-returns is crucial to obtain the orders of the marginal models, only a full system maximum likelihood strategy consisting in estimating a multivariate GARCH for the returns under the reduced rank null hypothesis (and in our case with a conditional multivariate

---

<sup>1</sup>We adopt the multivariate BEKK GARCH volatility models as the specification ensures fairly easily that the conditional covariance matrices are almost surely positive (semi-) definite. Moreover this model is frequently used in empirical macro-finance and it continues to be the object of research in econometric theory. See Section 3 of our paper for a more formal representation.

Gaussian distribution) has good size and power properties. Also note that our goal is not to provide a glossary of marginal representations obtained from every multivariate system<sup>2</sup> but to provide tools to understand the underlying behavior of a parsimonious block of assets.

The rest of this paper is as follows. Section 2 sets up the notations and derives the general results for the final equation representation of multivariate GARCH models. We propose in Section 3 some multivariate models accounting for co-movements in volatility and show that the implied marginal volatility processes are of low order. We introduce the pure portfolio common GARCH volatility model based on the factor GARCH specification proposed by Lin (1992) for the BEKK model. We develop in Section 4 several testing strategies that we evaluate in Section 5 using a Monte Carlo exercise. In Section 6 we apply our preferred test to six US stock return series. In particular we are able to discriminate between a diagonal model and a general multivariate framework, with correlated conditional variances and contagion effects, driven by a small number of common factors in volatility. The final section concludes.

## 2 The final equation representation of multivariate GARCH models

### 2.1 General results

To illustrate the main features of our paper and before investigating in details the multivariate GARCH models, let us consider a set of  $n$  stationary time series  $y_t = (y_{1t}, \dots, y_{nt})'$ , following the linear one factor model

$$y_t = \omega + \beta f_t + v_t,$$

where  $\beta$  is a  $n \times 1$  vector of factor loadings for the common factor  $f_t$ . We assume  $f_t$  to be generated by an ARMA processes  $\phi(L)f_t = \theta(L)u_t$  of order  $(p, q)$ , with  $u_t$  a white noise process.  $v_t$  denotes an  $n \times 1$  vector of disturbances (or idiosyncratic components) independent of  $u_t$ . The elements of  $v_{it}$  of the idiosyncratic disturbance vector  $v_t$  are assumed to follow univariate ARMA processes of order  $(\tilde{p}_i, \tilde{q}_i)$  such as  $\tilde{\phi}_i(L)v_{it} = \tilde{\theta}_i(L)w_{it}$  with  $w_{it}$  being white noise. Given that the  $v_{it}$ 's are individual effects, we assume them to be contemporaneously uncorrelated. Of course, one could allow for contemporaneous correlation between the  $w_{it}$ 's and thus between the  $v_{it}$ 's. As the multivariate process for  $v_t$  is diagonal

---

<sup>2</sup>Let us just name the diagonal model, the constant conditional correlation (CCC), the dynamic conditional correlation (DCC, see Engle, 2002), the dynamic equicorrelation (DECO, see Engle and Kelly, 2008), the approach by Baba, Engle, Kraft and Kroner (1999, hereafter BEKK), the orthogonal GARCH, the factor GARCH, as well as some of their block versions such as the BLOCK-DCC (Billio, Caporin and Gobbo, 2006) and BLOCK-DECO (Engle and Kelly, 2008) as examples proposed in the literature to restrict the multivariate setting towards a manageable size as well as to impose the positive definitiveness of the covariance matrix (see *inter alia* the surveys by Bauwens, Laurent and Rombouts, 2006 and Silvennoinen and Terasvirta, 2009).

ARMA, this would however not affect the degrees of the implied univariate ARMA processes for the  $v_{it}$ 's.

Substituting the specifications for  $f_t$  and for  $v_{it}$  into the model for  $y_t$  leads to the following univariate ARMA model for element  $i$  of  $y_{it}$  :

$$\phi(L)\tilde{\phi}_i(L)y_{it} = c_i + \beta_i\tilde{\phi}_i(L)\theta(L)u_t + \phi(L)\tilde{\theta}_i(L)w_{it}$$

where  $c_i = \phi(1)\tilde{\phi}_i(1)\omega_i$ , that is  $y_{it}$  is ARMA( $p + \tilde{p}_i, \max[\tilde{p}_i + q, p + \tilde{q}_i]$ ). Several cases emerges: (i) if both the common factor and the idiosyncratic elements are generated by low order ARMA models, the univariate processes for the  $y_{it}$ 's will also be of low order; (ii) in absence of a factor structure, i.e., if  $\beta = 0$ ,  $y_t$  is generated by a diagonal multivariate ARMA process of low orders if all  $\tilde{p}_i$  and  $\tilde{q}_i$  are small digits; (iii) if the idiosyncratic vector  $v_t$  is white noise, the univariate processes for  $y_{it}$  implied by the common factor model are  $\phi(L)y_{it} = c_i + \beta_i\theta(L)u_t + \phi(L)w_{it}$ , where  $c_i = \phi(1)\tilde{\phi}_i(1)\omega_i$ , that is the univariate model for  $y_{it}$  is ARMA( $p, \max[p, q]$ ) for all  $i$ . The orders of the univariate models will be low, provided  $p$  and  $q$  are low; (iv) the presence of additional factors (a two or a three factors model) increases the order of the ARMA-processes for  $y_{it}$  using the same algebra.

In the presence of a common factor, the autoregressive polynomials of the marginal processes for  $y_{it}$  should have the roots of  $\phi(L)$  in common. Moreover, for the pure common factor model, the implied orders and the parameters of the marginal ARMA schemes are identical for all elements of  $y_t$ . The moving average polynomials could exhibit some parameter heterogeneity due to the heterogeneity of the factor loadings  $\beta_i$ . Testing these implications could provide the grounds in the presence of parsimonious univariate ARMA schemes to decide which one of the two basic models contributes to explaining features of the data (and to what extent each of them - rather than only one of them- contributes to the explanation of the parsimony of the univariate processes). And by nesting the two basic models in a more general model, one avoids the problems with testing non-nested hypotheses.

The investigation of the problem is more demanding for implied univariate models in the case of factors in volatility and in particular concerning the orders of the marginal ARMA-processes in the case of more than one factor. In this paper we have chosen a particular structure for the factors as proposed by Engle and Marcucci (2006) such that the factor is a combination of past volatilities. The diagonal model of Bollerslev *et al.* (1988) will be a simple representation where each volatility and cross-products depend only on their own past. Both models are nested within an unrestricted VAR for the volatility and the cross product.

## 2.2 The MGARCH specification

To state the notation for univariate excess returns,<sup>3</sup>  $\varepsilon_{1t}$  is such that  $\varepsilon_{1t} = u_{1t}\sqrt{h_{11t}}$ , where  $u_{1t}$  has any centered parametric distribution with unit variance and the conditional variance of  $\varepsilon_{1t}$  follows a GARCH( $p, q$ ) with  $h_{11t} = \omega_1 + \sum_{j=1}^p \beta_{1,j}h_{11t-j} + \sum_{i=1}^q \alpha_{1,i}\varepsilon_{1t-i}^2$ . Consequently here,  $p$  refers to the GARCH terms and  $q$  is the order of the moving average ARCH term. The error in the squared returns is obtained using  $v_{1t} = \varepsilon_{1t}^2 - h_{11t}$ . As an example a GARCH(1, 2) can be rearranged for the squared returns such that  $\varepsilon_{1t}^2 = \omega_1 + (\alpha_{1,1} + \beta_{1,1})\varepsilon_{1t-1}^2 + \alpha_{1,2}\varepsilon_{1t-2}^2 + v_{1t} - \beta_{1,1}v_{1t-1}$  (see *inter alia* Bollerslev, 1996). In general a GARCH( $p, q$ ) has got an ARMA( $\max(p, q), p$ ) representation for the squared returns.

For multivariate modeling, we denote by  $\varepsilon_t = H_t^{1/2}z_t$ ,  $t = 1, \dots, T$ , the  $n \times 1$  vector of excess returns of financial assets observed at the time period  $t$ .  $T$  is the number of observations and  $z_t \sim i.i.d.(0, I_n)$ . Consequently we have  $\varepsilon_t|\Omega_{t-1} \sim D(0, H_t)$  with  $\Omega_{t-1}$  the past information set and  $H_t$  the conditional covariance matrix of the  $n$  assets that is measurable with respect to  $\Omega_{t-1}$ ;  $D(., .)$  is some arbitrary multivariate distribution. Similarly to the univariate case, and using half-vectorization operators for the  $vech(H_t)$ , let us denote  $v_t = vech(\varepsilon_t\varepsilon_t') - vech(H_t)$  which by definition is a martingale difference sequence with respect to the past of  $vech(\varepsilon_t\varepsilon_t')$ .

Therefore, a multivariate GARCH(0,1) (MGARCH(0,1) hereafter) can be written as a VAR(1) for observed squared returns and covariances. For instance a bivariate MGARCH(0,1) with  $vech(H_t) = \{h_{11t}, h_{12t}, h_{22t}\}'$  and  $vech(\varepsilon_t\varepsilon_t') = \{\varepsilon_{1t}^2, \varepsilon_{1t}\varepsilon_{2t}, \varepsilon_{2t}^2\}'$  gives

$$\begin{pmatrix} \varepsilon_{1t}^2 \\ \varepsilon_{1t}\varepsilon_{2t} \\ \varepsilon_{2t}^2 \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t-1}^2 \\ \varepsilon_{1t-1}\varepsilon_{2t-1} \\ \varepsilon_{2t-1}^2 \end{pmatrix} + \begin{pmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{pmatrix}, \quad (1)$$

where usual non-negativity and stationarity restrictions on the parameters are assumed to be met. Let us now generalize the system (1) for  $n$  assets and  $q$  lags, with  $vech(\varepsilon_t\varepsilon_t') = \varepsilon_t$  a  $N = n(n+1)/2$  dimensional vector, such that

$$A(L)\varepsilon_t = \omega + v_t, \quad (2)$$

with the lag operator  $L$  such that  $Lz_t = z_{t-1}$ .  $A(L) = (I - A_1L - \dots - A_qL^q)$  is a matrix polynomial of degree  $q$  in this example, where  $A_i$  are coefficient matrices;  $\omega$  and  $v_t$  are  $N$ -dimensional vectors. Let us now premultiply both sides of (2) by the adjoint of the matrix polynomial  $A(L)$  to obtain  $|A(L)|\varepsilon_t = \omega^* + Adj\{A(L)\}v_t$  where  $|A(L)| = \det\{A(L)\}$ , i.e. the determinant of the matrix polynomial  $A(L)$ , is a scalar polynomial in  $L$ ;  $Adj\{A(L)\}$  denotes the adjoint (or the adjugate) of  $A(L)$  and  $\omega^* = Adj\{A(L)\}\omega$ . This can be rewritten as a system of univariate weakly ARMA models in  $\varepsilon_t$  with autoregressive polynomial  $|A(L)|$  and scalar moving average polynomial  $\theta_i(L)$  and a white noise disturbance  $\zeta_{it}$

$$|A(L)|\varepsilon_{it} = \omega_i^* + \theta_i(L)\zeta_{it}, \quad i = 1, \dots, n. \quad (3)$$

---

<sup>3</sup>By excess return we mean that the conditional mean (a constant or an ARMA model for instance) has been subtracted from the returns  $r_t$ , say  $\varepsilon_t = r_t - E(r_t|\Omega_{t-1})$ , where  $\Omega_{t-1}$  denotes the information set up to and including  $t-1$ .

Notice that each of the variables  $\zeta_{it}$  is serially uncorrelated, but they are cross-correlated at different lags. In a MGARCH(0,  $q$ ), each component of the vector  $\epsilon_t$  has a weak ARMA( $Nq, (N - 1)q$ ). Next propositions (see also Nijman and Sentana, 1996 for the unrestricted case) summarize the main features of the final equation representation.

**Proposition 1** *For a  $n$ -dimensional MGARCH(0,  $q$ ), each univariate component is a weak GARCH<sup>4</sup> with a univariate ARMA( $Nq, (N - 1)q$ ) representation of the squared returns and cross-returns with identical autoregressive polynomials, where  $N = \frac{n(n+1)}{2}$ . Consequently each component in (3) follows a weak GARCH( $(N - 1)q, Nq$ ). The orders should be taken as upperbounds for the orders of the univariate ARMA and GARCH models.*

**Proof.** The proof is obvious from the definition of the determinant and the adjoint. This well known result is simply due to the fact that in the MGARCH(0,  $q$ ) for instance  $|A(L)|$  contains by construction up to  $L^{Nq}$  terms and the adjoint matrix is a collection of  $\{(N - 1) \times (N - 1)\}$  cofactor matrices, each of the matrix elements can contain the terms  $1, L, \dots, L^q$ . As  $v_t$  is a vector martingale difference sequence it is serially uncorrelated and each element of  $Adj\{A(L)\}v_t$  can be represented as a univariate moving average and therefore it is a weak GARCH process. ■

**Proposition 2** *Proposition 1 generalizes in a straightforward manner to show that for a  $n$ -dimensional MGARCH( $p, q$ ), squared and cross-returns have a univariate ARMA( $N \max\{p, q\}, (N - 1) \max\{p, q\} + p$ ) representation at most with identical autoregressive polynomials. Consequently each component follows a weak GARCH( $(N - 1) \max\{p, q\} + p, N \max\{p, q\}$ ) process at most.*

**Proof.** It is similar to that of Proposition 1. ■

The above outcomes, that apply the usual results of the VAR( $p$ ) and VARMA( $p, q$ ) are generally not in agreement with empirical findings suggesting low order univariate GARCH schemes. Indeed, for  $n = 20$  assets, a MGARCH(0, 2) implies individual ARMA(420, 418) models in squared returns and cross-products and individual GARCH(418, 420) processes. Obviously these orders should be taken as upperbounds. For instance there might exist coincidental situations (Granger and Newbold, 1986) in which there exist “quasi” common roots in the determinant and the adjoint (see Nijman and Sentana, 1996 for an example). Interestingly, Chevillon, Hecq and Laurent (2015) show that long memory can be the result of the marginalization of a multivariate model (e.g. a VAR(1)) when  $n \rightarrow \infty$ . Note also the implied orders of the ARMA models given in Proposition 2 are unchanged if  $Cov(v_t)$  is assumed to be a diagonal matrix.

---

<sup>4</sup>We follow Drost and Nijman (1993, p.911) by defining a (univariate) weak GARCH process as a process, say  $\epsilon_{1t}$ , of which its linear projection on 1, its past values and squared past values is zero and of which the linear projection of  $\epsilon_{1t}^2$  on the unit vector, on the past values and past squared values is given by the GARCH model (3) with  $\zeta_{it}$  being the error white noise error from the linear projection.

A particular case in which there are exact common roots to the implied AR and MA parts is the diagonal model of Bollerslev *et al.* (1988) where in equation (1)  $A_1 = \text{diag}(\alpha_{11}, \alpha_{22}, \alpha_{33})$  such that

$$\begin{pmatrix} \varepsilon_{1t}^2 \\ \varepsilon_{1t}\varepsilon_{2t} \\ \varepsilon_{2t}^2 \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & 0 & 0 \\ 0 & \alpha_{22} & 0 \\ 0 & 0 & \alpha_{33} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t-1}^2 \\ \varepsilon_{1t-1}\varepsilon_{2t-1} \\ \varepsilon_{2t-1}^2 \end{pmatrix} + \begin{pmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{pmatrix}. \quad (4)$$

Hence a diagonal multivariate strong GARCH process is identical to a set of strong GARCH univariate processes with possibly contemporaneous correlated disturbances.<sup>5</sup> Another popular multivariate GARCH model is the BEKK( $p, q$ ) of Baba *et al.* (1989). In this specification there are no common roots between the determinant and the adjoint and consequently the general results of propositions 1 and 2 apply.

Next proposition uses results obtained by Cubadda *et al.* (2009) for the VAR( $p$ ) and applies them to the MGARCH (1) with some reduced rank in  $A(L)$ .

**Proposition 3** *In a  $n$ -dimensional MGARCH(0,  $q$ ) with  $k$  factors and hence  $\text{rank}[A(L)] = k$ , the squared returns and cross-returns have a univariate ARMA( $p^*, q^*$ ) representation with identical autoregressive polynomials. The orders of  $p^*$  and  $q^*$  are at most  $kq$ . When  $k = 1$ , each component of a MGARCH(0, 1) follows a weak GARCH(1, 1) whatever the number of assets jointly considered.*

**Proof.** The proof applies the results in Cubadda *et al.* (2009) for the VAR( $p$ ). ■

This result can be easily generalized (see Appendix) to the  $n$ -dimensional MGARCH( $p, q$ ) when the coefficient matrices of the ARCH and the GARCH part have a nested reduced rank structure or not. As an example, a MGARCH(1,1) with one factor and with the same matrix generating the left null space of the ARCH and the GARCH part, implies univariate GARCH(1,1) models as in the MGARCH(0,1); but GARCH(2,1) models in the absence of the commonality of the left null spaces.

### 3 Pure portfolio common GARCH model

#### 3.1 The single common factor BEKK model

According to the propositions 1 and 2, simple multivariate models (e.g. MGARCH(0,1)) do not imply marginal parsimonious low order univariate GARCH processes. Parsimony of the univariate marginal representations might be obtained for instance from a multivariate diagonal model (4) proposed by Bollerslev *et al.* (1988). Alternatively, the presence of common factors generating the volatility of asset returns in the MGARCH is able to explain those stylized facts. Note that we have illustrated in Subsection 2.1 some alternative specifications lying in between those two polar cases. We only consider those in this paper however.

---

<sup>5</sup>The conditional orthogonal model further assumes that the cross-product term  $\alpha_{22} = 0$  (and then  $\omega_2 = 0$ ).

Factor (G)ARCH models have been proposed and extensively studied in the literature, among others by Vrontos, Dellaportas and Politis (2003), Fiorentini, Sentana and Shephard (2004), Lanne and Saikkonen (2007), Hafner and Preminger (2009) and more recently by García-Ferrer, González-Prieto and Peña (2012). The factor structure developed in Section 2 is not very meaningful in the light of these existing models. For instance the combination that involves both squared returns and cross-returns suffers from mathematical incoherence.

The BEKK model (Baba, Engle, Kraft and Kroner, 1989) however would give a direct link with the common feature literature. Indeed, testing for common GARCH feature (in the sense of Engle and Kozicki, 1993) amounts to look towards  $s$  directions  $\delta' \varepsilon_t$  such that the  $\delta' \varepsilon_t$  combinations of asset returns are conditionally homoskedastic, namely such that  $\delta' \varepsilon_t | \Omega_{t-1} \sim D(0, C)$  where  $C$  does not depend on  $t$ . To be more explicit let us define the following partitioned matrix  $\Lambda$  spanning  $\mathcal{R}^n$

$$\Lambda = \begin{pmatrix} I_s & \delta_{s \times (n-s)}^{*'} \\ 0_{(n-s) \times s} & I_{n-s} \end{pmatrix},$$

where we have normalized the cofeature space  $\delta' = (I_s : \delta_{s \times (n-s)}^{*'})$  for the sake of notation. Let us also denote without loss of generality  $H_t = H + \tilde{H}_t$  as the sum of a constant part and a time-varying part. We define a common volatility feature model such that in  $\Lambda \varepsilon_t | \Omega_{t-1} \sim D(0, \Lambda H \Lambda' + \Lambda \tilde{H}_t \Lambda')$ , we have  $\text{rank}(\Lambda \tilde{H}_t \Lambda') = n - s$  instead of  $n$ .

Clearly in this presentation there exist  $s$  linear combinations of  $\varepsilon_t$  that have time-invariant conditional distributions. The remaining  $n - s$  combinations generate the time-varying volatility of the system. This definition makes the BEKK model particularly interesting for our common factors investigations although it necessitates some caution when determining the number of factors. Let us denote a BEKK( $p, q$ ) model with  $k$  factors as F-BEKK( $p, q, k$ ) (see Lin, 1992) and in particular a F-BEKK(0, 1, 1) such that

$$\begin{aligned} \varepsilon_t &= H_t^{1/2} z_t, \\ H_t &= \Gamma_0 \Gamma_0' + \Gamma_1 \varepsilon_{t-1} \varepsilon_{t-1}' \Gamma_1', \\ &= \Gamma_0 \Gamma_0' + \gamma \varphi' \varepsilon_{t-1} \varepsilon_{t-1}' \varphi \gamma', \end{aligned}$$

where for  $k = 1$  we have  $\text{rank}(\Gamma_1) = \text{rank}(\varphi \gamma') = 1$ . We can write this system using the half-vec operator  $\text{vech}$  such that  $\text{vech}(H_t) = \text{vech}(\Gamma_0 \Gamma_0') + A_1 \text{vech}(\varepsilon_{t-1} \varepsilon_{t-1}')$  or in terms of squared errors and cross products  $\text{vech}(\varepsilon_t \varepsilon_t') = \text{vech}(\Gamma_0 \Gamma_0') + A_1 \text{vech}(\varepsilon_{t-1} \varepsilon_{t-1}') + v_t$  with the multivariate martingale difference sequence  $v_t = \text{vech}(\varepsilon_t \varepsilon_t') - \text{vech}(H_t)$ .  $A_1 = L_{N \times n^2} (\Gamma_1 \otimes \Gamma_1) S_{n^2 \times N}$  where  $L_{N \times n^2}$  is the selection matrix that eliminates the redundant lower-triangular elements of  $H_t$  and  $S_{n^2 \times N}$  selects columnwise the coefficients of the squared returns and sum the columns of the cross-returns (see e.g. Harville, 1997, p.

357).<sup>6</sup>  $A_1$  can be written in a reduced form with

$$A_1 = \begin{pmatrix} \gamma_1^2 \varphi_1^2 & 2\gamma_1^2 \varphi_1 \varphi_2 & \gamma_1^2 \varphi_2^2 \\ \gamma_1 \gamma_2 \varphi_1^2 & 2\gamma_1 \gamma_2 \varphi_1 \varphi_2 & \gamma_1 \gamma_2 \varphi_2^2 \\ \gamma_2^2 \varphi_1^2 & 2\gamma_2^2 \varphi_1 \varphi_2 & \gamma_2^2 \varphi_2^2 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{\gamma_2}{\gamma_1} \\ (\frac{\gamma_2}{\gamma_1})^2 \end{pmatrix} \begin{pmatrix} \gamma_1^2 \varphi_1^2 & 2\gamma_1^2 \varphi_1 \varphi_2 & \gamma_1^2 \varphi_2^2 \end{pmatrix},$$

such that there exists a normalized cofeature matrix

$$\delta = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{\gamma_1^2}{\gamma_2} & -\frac{\gamma_1}{\gamma_2} \end{pmatrix},$$

with  $\delta' A = 0_{2 \times 3}$ . From the determinant  $\det(I - A_1 L) = 1 - (\gamma_1 \varphi_1 + \gamma_2 \varphi_2)^2 L$  and the adjoint (which is not reported to save space) it emerges that all squared returns and their cross-products are ARMA(1,1) and hence we have GARCH(1,1) specifications for the conditional variances and covariances. This is in accordance with Proposition 3. Importantly enough however, the combination of the series representing the factor, i.e.,  $P_{t-1}^2 = (\gamma_1 \varphi_1 \varepsilon_{1t-1} + \gamma_2 \varphi_2 \varepsilon_{2t-1})^2 = \gamma_1^2 \varphi_1^2 \varepsilon_{1t-1}^2 + 2\gamma_1^2 \varphi_1 \varphi_2 \varepsilon_{1t-1} \varepsilon_{2t-1} + \gamma_2^2 \varphi_2^2 \varepsilon_{2t-1}^2$  is the variance of the portfolio made up by the series whose weights are determined by the reduced rank analysis and the factor structure.

Therefore, each volatility and cross-correlation in such a system only depend on that portfolio. This is the reason why we call this model a pure portfolio common GARCH model given in terms of unobserved volatilities

$$\begin{pmatrix} h_{11t} \\ h_{12t} \\ h_{22t} \end{pmatrix} = \begin{pmatrix} \varpi_1 \\ \varpi_2 \\ \varpi_3 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} P_{t-1}^2, \quad (5)$$

where the  $\beta_i$ 's measure the impact of the volatility of the portfolio on the volatilities and covariances of the underlying assets. Moreover there exist combinations of series such that  $\beta'_\perp \text{vech}(H_t)$  does not depend on the volatility of the portfolio. This approach is also different from the GO-GARCH where the direction for the combination is such that the variance of the returns is maximized using a principal components analysis. In our case we look at the combinations that are the most correlated with the volatility and covariances of the assets.

Note that this interpretation is made easier using the factor BEKK than from an unrestricted MGARCH with a reduced rank in squared returns and cross-returns. Finding a reduced rank in the general model (1) for instance does not ensure that the coefficient of the cross-correlation is numerically two times the coefficients in front of the returns.

---

<sup>6</sup>Note that in the one-factor case,  $A_1$  can also be obtained using  $\varphi \varphi' (\gamma' \varepsilon_{t-1})^2 = \gamma \varphi' \varepsilon_{t-1} \varepsilon_{t-1}' \varphi \gamma'$ . Therefore, the matrix  $\varphi \varphi'$  is of rank one as well as the coefficient matrix  $A_1$ .

### 3.2 Multiple common pure portfolios in the BEKK

Our findings about the parsimony of the implied series generalize in the  $k$  factor case for the BEKK with caution. Indeed the presence of  $k$  factors implies that the matrices  $A_i$  in  $(I - A(L))$  are of rank  $k$  only in the MGARCH model in Equation (1). This is not the case for the BEKK. Let us illustrate this with the BEKK(0,1)  $H_t = \Gamma_0 \Gamma_0' + \Gamma_1 \varepsilon_{t-1} \varepsilon_{t-1}' \Gamma_1'$ , where we apply the vectorization operator (see for instance Lütkepohl, 1996) such that

$$\begin{aligned} \text{vec}(H_t) &= \text{vec}(\Gamma_0 \Gamma_0') + \text{vec}(\Gamma_1 \varepsilon_{t-1} \varepsilon_{t-1}' \Gamma_1') \\ &= \text{vec}(\Gamma_0 \Gamma_0') + (\Gamma_1 \otimes \Gamma_1) \text{vec}(\varepsilon_{t-1} \varepsilon_{t-1}'), \end{aligned}$$

where  $\text{vec}(H_t)$  is of size  $n^2 \times 1$ . Hence when  $\text{rank}(\Gamma_1) = 1$  we also have that  $\text{rank}(\Gamma_1 \otimes \Gamma_1) = \text{rank}(\Gamma_1) \times \text{rank}(\Gamma_1) = 1$ . When there are  $k$  factors in  $\Gamma_1$  however, we have that  $\text{rank}(\Gamma_1) = k$  and  $\text{rank}(\Gamma_1 \otimes \Gamma_1) = k^2$ . This has some consequences for the model in which one eliminates the redundant lines for the cross-correlations in  $\text{vec}(H_t)$  as well as for the pure variance model in which we only focus on the vector of variances  $\varepsilon_t^2$ .

For the *vech*,

$$\text{vech}(H_t) = \gamma_0 + L_{N \times n^2} (\Gamma_1 \otimes \Gamma_1) S_{n^2 \times N} \text{vech}(\varepsilon_{t-1} \varepsilon_{t-1}'), \quad (6)$$

where  $\gamma_0 = L_{N \times n^2} \text{vec}(\Gamma_0 \Gamma_0')$  is a  $N$ -dimensional vector with the intercepts, it can be shown (see Harville (1997, p. 358-9) that

$$\text{rank}[L_{N \times n^2} (\Gamma_1 \otimes \Gamma_1) S_{n^2 \times N}] = \frac{1}{2} [\text{rank}(\Gamma_1)]^2 + \frac{1}{2} [\text{rank}(\Gamma_1)] = \tilde{k}_{\text{vech}}, \quad (7)$$

which means that for  $k = 1, 2, 3, \dots$  in  $\Gamma_1$ , there are respectively  $\tilde{k}_{\text{vech}} = 1, 3, 6, \dots$  factors in the *vech* with reduced rank restrictions. Consequently, one cannot test for  $k = 2$  factors in the VAR representation of squared returns and cross-returns with  $n = 3$  assets.

Engle and Marcucci (2006) only consider commonalities in volatilities in their pure variance framework. We now study the implications in terms of model representations for the variances if the true model has commonalities in a complete system like in (6). In order to first investigate the consequences of ignoring the covariances on the left-hand side of (6), let us now define  $D_{n \times N}$  a selection matrix that selects the subsystem consisting of the rows of  $\text{vech}(H_t)$  corresponding to the variances such that  $D_{n \times N} \text{vech}(H_t) = \text{vecd}(H_t) = \varepsilon_t^2$  where *vecd* denotes the diagonal vectorization operator.

We have that

$$\text{vecd}(H_t) = \tilde{\gamma}_0 + D_{n \times N} L_{N \times n^2} (\Gamma_1 \otimes \Gamma_1) S_{n^2 \times N} \text{vech}(\varepsilon_{t-1} \varepsilon_{t-1}'), \quad (8)$$

where  $\tilde{\gamma}_0 = D_{n \times N} \text{vec}(\Gamma_0 \Gamma_0')$ . Let us finally assume that there is no omitted variable bias by also excluding the cross-products (i.e. no contagion with zero coefficients) from the right-side of (8) or that we have a model that projects past squared returns and cross-returns on squared returns. Using  $K_{N \times n}$  the

matrix that eliminates the columns corresponding to the cross-products we obtain a relationship between  $vecd(H_t)$  and  $vecd(\varepsilon_{t-1}\varepsilon'_{t-1})$ . From the usual inequalities related to the rank of the product of matrices we can easily obtain  $\tilde{k}_{vecd} = \min(n, \tilde{k}_{vech})$ .

Our previous results on the implied univariate models (Proposition 3) must be adapted using  $\tilde{k}_{vech}$  or  $\tilde{k}_{vecd}$  instead of  $k$  (and therefore  $\tilde{k}_{vech}$  and  $\tilde{k}_{vecd}$  are now the number of common volatility factors). For instance, in an  $n$ -dimensional stationary GARCH(0,  $q$ ), the individual ARMA processes for the squared excess returns have both AR and MA orders not larger than  $\tilde{k}_{vech}q$ .

The previous results also show that those obtained by Engle and Marcucci (2006) might be misleading for getting the number of factors  $k$  when the model is unknown. Indeed, they assume a pure variance specification in which the covariances do not play any role. In practice they consider an exponential form and thus take the log-transform to ensure strictly positive squared returns. However, whether this is an appropriate model description of the data or not, they test for the presence of reduced rank between the squared returns and the past squared returns in a multiple regression  $\varepsilon_t^2 = \gamma_0 + \varphi\gamma'\varepsilon_{t-1}^2 + v_t$  for  $\varepsilon_t^2 = (\varepsilon_{1t}^2, \dots, \varepsilon_{nt}^2)'$ . Consequently if the DGP has a BEKK representation, the numbers that one obtains ( $\tilde{k}_{vecd} = \min(n, \tilde{k}_{vech})$ ) must be translated to get back to  $k$ . Finally note that Engle and Marcucci (2006) take the log of the elements of  $\varepsilon_t^2$ . Furthermore, to avoid taking the log of zero squared returns, they add a tiny constant  $\iota$  to the squared returns, (i.e.  $\ln(\varepsilon_t^2 + \iota)$ ) with the undesirable consequence of introducing large negative values and hence artificially making  $\ln(\varepsilon_t^2 + \iota)$  closer to an *i.i.d.* process

## 4 Test statistics

In the previous sections we have studied the implications of the presence of a factor structure in BEKK(0,  $q$ ). Three factor representations have been derived. We now propose and compare two different strategies to detect the presence of these reduced rank structures. A first group of tests building upon the one proposed by Engle and Marcucci (2006) are based on canonical correlation analyses of the VAR representations of models (6), (8) and the misspecified model between the squared returns and their past only. Then we also consider a likelihood ratio test for determining  $k$  in a general BEKK( $p$ ,  $q$ ).

As explained above, the first group of tests is inspired by the one proposed by Engle and Marcucci (2006). For  $\varepsilon_t = H_t^{1/2}z_t$ , the three VAR representations of the generalized versions of our models to the F-BEKK(0,  $q$ ,  $k$ ) are as follows

$$\begin{aligned} \mathcal{M}_1 : \quad & vech(\varepsilon_t\varepsilon'_t) = \gamma_0 + A_1vech(\varepsilon_{t-1}\varepsilon'_{t-1}) + \dots + A_qvech(\varepsilon_{t-q}\varepsilon'_{t-q}) + v_t, \\ \mathcal{M}_2 : \quad & vech(\varepsilon_t\varepsilon'_t) = \tilde{\gamma}_0 + \tilde{A}_1vech(\varepsilon_{t-1}\varepsilon'_{t-1}) + \dots + \tilde{A}_qvech(\varepsilon_{t-q}\varepsilon'_{t-q}) + \tilde{v}_t, \\ \mathcal{M}_3 : \quad & vech(\varepsilon_t\varepsilon'_t) = \check{\gamma}_0 + \check{A}_1vech(\varepsilon_{t-1}\varepsilon'_{t-1}) + \dots + \check{A}_qvech(\varepsilon_{t-q}\varepsilon'_{t-q}) + \check{v}_t, \end{aligned}$$

where  $v_t$  and  $\tilde{v}_t$  are martingale differences of the Wold or VAR representation of  $vech(\varepsilon_t\varepsilon'_t)$  and  $vecd(\varepsilon_t\varepsilon'_t)$ , and  $\check{v}_t$  is a martingale difference if the DGP is a pure variance model. As they are martingale differences,

they might exhibit heteroskedasticity but not autocorrelation. We aim at determining for  $\mathcal{M}_i$ ,  $i = 1, 2, 3$  the number of common factors  $k$  driving  $H_t$  or similarly the  $n - k$  or  $N - k$  vectors annihilating the dynamics. We assume nested reduced rank structures in the  $A$  matrices.<sup>7</sup> In order to determine the rank  $\tilde{k}_{vech}$  or  $\tilde{k}_{vecd}$  corresponding to the number of factors  $k$  we first rely on the canonical correlation analysis (as suggested by Engle and Marcucci, 2006 for a specification similar to  $\mathcal{M}_3$ ), i.e. an analysis of the eigenvalues and the eigenvectors of

$$\hat{\Sigma}_{YY}^{-1} \hat{\Sigma}_{YW} \hat{\Sigma}_{WW}^{-1} \hat{\Sigma}_{WY}, \quad (9)$$

or similarly in the symmetric matrix  $\hat{\Sigma}_{YY}^{-1/2} \hat{\Sigma}_{YW} \hat{\Sigma}_{WW}^{-1} \hat{\Sigma}_{WY} \hat{\Sigma}_{YY}^{-1/2}$ .  $\hat{\Sigma}_{ij}$  are the estimated covariance matrices,  $Y$  is the left hand-side variable of one of the three systems  $\mathcal{M}_i$  and  $W$  the right hand-side explanatory variables. For *iid* normally distributed disturbances, the likelihood ratio test statistic for the null hypothesis that there exist at least  $s$  linear combinations that annihilate  $\tilde{k}_{vecd} = (n - s)$  or  $\tilde{k}_{vech} = (N - s)$  features in common to these random variables is given by

$$\zeta_{LR(s)} = -T \sum_{j=1}^s \ln(1 - \hat{\lambda}_j) \quad s = 1, \dots, n \text{ or } N, \quad (10)$$

where  $\hat{\lambda}_j$  is the  $j$ -th smallest eigenvalue of the estimated matrix (9).<sup>8</sup> In VAR models with Gaussian errors, (10) is the usual likelihood ratio statistic and follows asymptotically a  $\chi_{(df_i)}^2$  distribution under the null for  $\mathcal{M}_i$ ,  $i = 1, 2, 3$  where respectively  $df_1 = sNq - s(N - s)$ ,  $df_2 = snq - s(N - s)$  and  $df_3 = snq - s(n - s)$ . However in our setting, the errors  $v_t$ ,  $\tilde{v}_t$  and  $\check{v}_t$  are neither *i.i.d.* nor Gaussian but highly skewed and at best martingale difference sequences. Consequently  $\zeta_{LR(s)}$  is likely not to be  $\chi_{(df)}^2$  distributed under the null. Nevertheless, we consider this case because it is a direct extension of the approach of Engle and Marcucci (2006) that we want to evaluate. Note that in their theoretical framework of an exponential pure variance model, the log transformation of the squared residuals first renders the residuals normally distributed (justifying the use (10)) and attenuates the heteroskedasticity of the error term. Whether this test is accurate in a F-BEKK(0,  $q$ ,  $k$ ) framework (not imposing an exponential structure and allowing for cross-returns in the factor structure), is an open question that we investigate in the next section by means of a Monte Carlo analysis.<sup>9</sup>

<sup>7</sup>We have also investigated a variant of  $\mathcal{M}_2$  with a factor representation of the variances only. To do so, we concentrate out the effect of the covariances before applying  $\mathcal{M}_3$  on residuals. The behavior of this test was poor and consequently the results are not reported to save space.

<sup>8</sup>Canonical correlation based tests have been extensively used in economics to test for the presence of common factors and to determine their number (see e.g. Anderson and Vahid, 2007, who proposed a version of this test that is robust to the presence of jumps in the observed series).

<sup>9</sup>We have also investigated the use of a multivariate robust Wald counterpart to the canonical correlation approach modifying the Wald approach to test for reduced rank by Christensen, Pagan and Hurn (2011) with the multivariate extension of the White approach proposed in Ravikumar *et al.* (2000) for system of seemingly unrelated regressions. Given the disappointing behaviour of those tests we do not report the results to save space. They can be found in the working paper version of this paper.

For this latter reason and because the above canonical correlation tests are also more difficult to implement in a VARMA context<sup>10</sup>, i.e. for the general F-BEKK( $p, q, k$ ) model, we also propose a standard likelihood ratio test. The Gaussian maximum likelihood (ML) estimator is the one that maximizes the log likelihood function  $L(\vartheta) = \sum_{i=1}^T l_t(\vartheta)$  with

$$l_t(\vartheta) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \ln |H_t(\vartheta)| - \frac{1}{2} \varepsilon_t' H_t^{-1}(\vartheta) \varepsilon_t.$$

Let us denote by respectively  $L(\vartheta^{un})$  and  $L(\vartheta^{res})$  the likelihood values for the unrestricted full-BEKK and the reduced rank BEKK with  $k$  factors. For instance in a BEKK(0, 1), the unrestricted model is  $H_t = \Gamma_0 \Gamma_0' + \Gamma_1 \varepsilon_{t-1} \varepsilon_{t-1}' \Gamma_1'$  and the restricted one  $H_t = \Gamma_0 \Gamma_0' + \alpha \beta_1' \varepsilon_{t-1} \varepsilon_{t-1}' \beta_1 \alpha'$ . Then the likelihood ratio is

$$LR = 2\{L(\vartheta^{un}) - L(\vartheta^{res})\} \sim \chi_{(df)}^2$$

where the number of degrees of freedom is the difference between the number of estimated parameters of the unrestricted and restricted models. As an example, in a BEKK(0,  $q$ ) this difference is  $df = \underbrace{n^2 q}_{unrestricted} - \underbrace{(nk + nkq - k^2)}_{restricted}$ . The advantage of this LR approach over the canonical correlation is that it can be easily generalized to F-BEKK( $p, q, k$ ). From the generalization of Proposition 3 (see Appendix) we favor a model with a same left null space for the ARCH and the GARCH matrix parameters. For instance  $H_t = \Gamma_0 \Gamma_0' + \alpha \beta_1' \varepsilon_{t-1} \varepsilon_{t-1}' \beta_1 \alpha' + \alpha \delta_1' \varepsilon_{t-1} \varepsilon_{t-1}' \delta_1 \alpha'$  in a BEKK(1, 1,  $k$ ). The number of degrees of freedom is computed in the same manner. The main drawback of this approach might be the difficulty to estimate the model for large  $n$ .

## 5 Monte Carlo results

We investigate in this section the small sample behavior of  $\zeta_{LR(s)}$  for  $\mathcal{M}_1$ ,  $\mathcal{M}_2$  and  $\mathcal{M}_3$  and of  $LR$ . The DGP is a factor BEKK(0, 1,  $k$ )  $H_t = \Gamma_0 \Gamma_0' + \Gamma_1 \varepsilon_{t-1} \varepsilon_{t-1}' \Gamma_1'$  with  $k = rank(\Gamma_1) = 1, 2$  factors for  $n = 5$  assets.<sup>11</sup> We take  $T = 1000$  and 2000 observations (sample sizes typically encountered with financial time series) and compute the rejection frequencies when testing the null hypothesis at a 5% significance level in 1000 replications. The parameters of the factor BEKK models have been "calibrated" on estimations obtained from some daily observations of stocks used in the empirical application but here we forced the model to have successively  $k = 1$  and 2 factors. Table 1 shows that the strategy that consists in

<sup>10</sup>See Tiao and Tsay (1983) for reduced rank analyses in the VARMA(p,q).

<sup>11</sup>Fewer size distortions are obtained for smaller  $n$  and hence it was less obvious to discriminate between the different approaches. Also note that many alternative strategies have been investigated but are not reported due to their bad performance. For instance a partial least square approach, similar to the one proposed by Cubadda and Hecq (2011) for the VAR, underestimates the number of factors and concludes to multivariate white noise process in most cases. Also, taking the log of the covariance matrix seems to work for some particular cases but not in general.

Table 1: Rejection frequencies of common GARCH factor tests

		$T = 1000$				$T = 2000$			
	Tests	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$LR$	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$LR$
<b><math>k = 1</math></b>		28.6	23.5	11.5	7.8	30.8	27.5	13.3	6.1
	Tests	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$LR$	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$LR$
<b><math>k = 2</math></b>		7.5	2.0	0.2	7.1	11.6	3.8	0.1	6.1

estimating directly the unrestricted and the restricted BEKK models, i.e.  $LR$ , is preferred. Indeed, the rejection frequencies for  $LR$  are close to the nominal 5% level. On the other hand the frequencies with which test statistics  $\zeta_{LR(s)}$  reject the null hypothesis are far from the 5% nominal significance level. For instance the canonical correlation approach proposed by Engle and Marcucci (2006), i.e.  $\mathcal{M}_1$ , seems to diverge when  $T$  increases. This is the case both for  $k = 1$  and  $k = 2$ . One could think that  $\mathcal{M}_1$  gives rejection frequencies close to the nominal size when  $k = 2$  but this is a pure random circumstance. Results are very different for different  $k$ 's.

In order to investigate the local power of the  $LR$  test that has good size properties, we perturb the reduced rank matrix  $\Gamma_1$  such that

$$\Gamma_1^* = \Gamma_1 + \Theta,$$

where  $\Theta$  is a  $n \times n$  matrix of zeros but the lower right element which is  $\theta = -0.3, -0.25, \dots, 0, \dots, 0.3$ . This allows us to present both the unadjusted power when  $\theta \neq 0$  and the rejection frequency (empirical size) when  $\theta = 0$  on a same picture. We only consider the case  $k = 1$  for  $T = 2000$  observations. Again 1000 replications are used for each  $\theta$ . As expected the power quickly converges to 100% the farther away one gets from the null hypothesis.

## 6 An illustrative example

In this section, the presence of commonalities in volatility within  $m$  stock returns is considered. Wrapping up our strategy, we first estimate  $m$  univariate models for each of these assets. Then we focus on a group of  $n \leq m$  series presenting similarities in terms of parsimony, e.g. GARCH(1,1) specifications. This block of  $n$  assets is more deeply studied in a multivariate setting in order to discriminate between a framework with few factors and something else, for instance a diagonal or an orthogonal model. The pure portfolio common volatility model introduced in this paper is a particular factor model generating homoskedastic portfolios for a subset of  $n$  out of the  $m$  original series (which is by no means riskfree).

To illustrate this approach, the data set we use is obtained from TickData and consists of daily closing transaction prices for fifty large capitalization stocks from the NYSE, AMEX NASDAQ, covering

Figure 1: Power of the likelihood ratio test

the period from January 1, 1999 to December 31, 2008 (2489 trading days).<sup>12</sup> For the conditional mean we have estimated AR(2) models with daily dummies to capture Monday and Friday effects. For the conditional variance we have run four different specifications. These are the GARCH(1,1), the GARCH(1,2) and two long memory models, namely FIGARCH(1,  $d$ , 0) and FIGARCH(1,  $d$ , 1). Out of the 50 series, the GARCH(1,1) model is favoured in six cases using both formal likelihood ratio tests and the Hannan-Quinn information criterion. The six returns with no indication of long memory are ABT (Abbott Laboratories), BMY (Bristol Myers Squibb Co.), GE (General Electric), SLB (Schlumberger N.V.), XOM (Exxon Mobil), XRX (Xerox Corporation). Table 2 reports the value of the estimated parameters in the conditional variance equation, the results on the conditional mean equations being not reported to save space.

Therefore we consider these six series and we apply our proposed tests. On the six series we consider both a BEKK(0,1) and a BEKK(1,1) where for the latter we impose the same left null space generating the ARCH and the GARCH parts. For these two models we compute the likelihood of the full-BEKK as well as the factor BEKK for  $k = 0, 1, 2, 3, 4, 5$ ; the F-BEKK with  $k = 6$  being the unrestricted full-BEKK. Using the  $LR$  test statistics we reject the null of any rank reduction ( $p$ -values are all smaller

---

<sup>12</sup>To save space, we do not report company names but only the well known ticker symbols. There are AAPL, ABT, AXP, BA, BAC, BMY, BP, C, CAT, CL, CSCO, CVX, DELL, DIS, EK, EXC, F, FDX, GE, GM, HD, HNZ, HON, IBM, INTC, JNJ, JPM, KO, LLY, MCD, MMM, MOT, MRK, MS, MSFT, ORCL, PEP, PFE, PG, QCOM, SLB, T, TWX, UN, VZ, WFC, WMT, WYE, XOM, XRX.

Table 2: MLE of GARCH(1,1) on the 6 retained series

	ABT	BMV	GE	SLB	XOM	XRX
$\omega$	0.0139 (0.0036)	0.0076 (0.0027)	0.0028 (0.0015)	0.0284 (0.0115)	0.0253 (0.0071)	0.0162 (0.0060)
$\alpha_1$	0.0439 (0.0048)	0.0617 (0.0058)	0.0433 (0.0049)	0.0428 (0.0051)	0.0726 (0.0067)	0.0612 (0.0053)
$\beta_1$	0.9512 (0.0054)	0.9405 (0.0046)	0.9580 (0.0046)	0.9515 (0.0065)	0.9180 (0.0082)	0.9413 (0.0054)
$Q^2(20)$	0.72	0.96	0.38	0.50	0.28	0.98

Note: Robust standard errors are reported in brackets.  $Q^2(20)$  is the p-value of the Ljung-Box test on squared standardized residuals with 20 lags.

than 0.001 and are therefore not reported to save space). Consequently, the parsimony observed for these six returns is more likely due to independent behavior (diagonal or orthogonal model) than to the presence of factors. We also estimated diagonal BEKK(0,1) and a BEKK(1,1) and found that on these 6 series, likelihood ratio tests favour the diagonal models compared to the full-BEKK ones. A diagonal BEKK(1,1) is therefore retained for these series.

Interestingly, we also applied the reduced rank tests to the VAR representation for the squared returns and cross-returns, i.e.  $\mathcal{M}_1$ ,  $\mathcal{M}_2$  and  $\mathcal{M}_3$ . Unlike the *LR* test, these tests point out the presence of commonalities in volatility which is very likely due to the strong size distortion of these tests as highlighted in our Monte Carlo simulations.

This result has some important implications for practitioners. Indeed, to solve the curse of dimensionality in multivariate systems, either strong assumptions/restrictions of diagonality of the parameter matrices are imposed on the model (e.g. diagonal BEKK versus full BEKK) or common factor structures are tested for. Both our simulation study and empirical analysis suggest that standard canonical correlation tests (in the spirit of the one of Engle and Marcucci, 2006) tend to over-reject the null of no common factors compared to a standard likelihood ratio test when applied to BEKK models. Results also suggest that data favor a diagonal BEKK model in which the conditional variance of asset  $i$  only depends on past (squared) shocks on asset  $i$  and not shocks on other series while the conditional covariance of assets  $i$  and  $j$  only depends on past shocks on both assets. This model is very much in line with the DCC of Engle (2002).

## 7 Conclusions

This paper studies the orders of the univariate weak GARCH processes implied by multivariate GARCH models. We recall that except in some coincidental situations, the implied marginal models are generally non-parsimonious. However, the presence of common features in volatility leads to a large decrease of these implied theoretical orders. We emphasize two radically different structures that may give rise to similar parsimonious univariate representations. These are diagonal models on the one hand (Bollerslev *et al.*, 1988) that could result when idiosyncratic volatility is the dominating source of volatility, and models with reduced rank matrices resulting from the presence of common factors on the other hand.

Consequently we propose different strategies to detect the presence of such GARCH co-movements in a multivariate setting. We believe that this is preferred to assuming non-contagion of asset returns from the outset or imposing a factor framework when it is not present. We find that reduced rank test statistics in squared returns (and cross-returns) display severe size distortions while our proposed likelihood ratio test is correctly sized and has good power properties.

Our results plead for looking at individual series prior to multivariate modelling. For instance, this would help to discover and estimate separately blocks of assets sharing the same sort of dynamic behavior.

In our application, we detected six series following GARCH(1,1) specifications. We applied our proposed likelihood ratio test and did not find any evidence of commonalities in volatility.

Extensions of this paper are numerous. For instance one could have used bootstrapped versions of some of the test statistics presented in this paper. This approach has been used by Hafner and Herwartz (2004) to test for causality in the full-BEKK model but with little improvement compared to the asymptotic versions of their test. Next we can further decompose the set of  $m$  assets into  $K$  blocks, each containing  $n_i$  assets, with  $m = n_1 + \dots + n_K$ . There might exist a group of series with long memory having similarities in their dynamics.

We could also account for a permanent, possibly long memory, component as Engle and Lee (1999) propose for a univariate setting if such a component is present in some series. The methodology that we propose here would then apply to the transitory components of the volatility of those assets which exhibit permanent and transitory components volatilities and to the volatilities themselves of those assets which do not have a permanent common volatility.

## 8 References

ANDERSON, H.M., AND F. VAHID (2007), Forecasting the Volatility of Australian Stock Returns: Do Common Factors Help?, *Journal of Business and Economic Statistics* 25, 76-90.

ARSHANAPALLI, B, DOUKAS, J. AND L. LANG (1997), Common volatility in the industrial structure

of global capital markets, *Journal of International Money and Finance* 16, 189-209.

BABA, Y., R.F. ENGLE, D.F. KRAFT, AND K.F. KRONER (1989), Multivariate Simultaneous Generalized ARCH, DP 89-57 San Diego.

BAUWENS, L., S. LAURENT AND J.V.K. ROMBOUTS (2006), Multivariate GARCH Models: a Survey, *Journal of Applied Econometrics* 21, 79–109.

BOLLERSLEV, T., ENGLE, R. AND J. WOOLDRIDGE (1988), A capital asset pricing model with time-varying covariances, *Journal of Political Economy*, 96, 116-131.

BILLIO M., M. CAPORIN AND M. GOBBO (2006), Flexible Dynamic Conditional Correlation Multivariate GARCH for Asset Allocation, *Applied Financial Economics Letters*, 2, 123–130.

CANDELON, B., HECQ, A. AND W. VERSCHOOR (2005), Measuring Common Cyclical Features During Financial Turmoil, *Journal of International Money and Finance* 24, 1317-1334.

CHEVILLON, G., HECQ, A., LAURENT, S. (2015), Long Memory Through Marginalization of Large Systems and Hidden Cross-Section Dependence, mimeo.

GARCÍA-FERRER, A., GONZÁLEZ-PRIETO, E. AND D. PEÑA (2012), A conditionally heteroskedastic independent factor model with an application to financial stock returns, *International Journal of Forecasting* 28, 70-93.

CUBADDA, G., HECQ A. AND F.C. PALM .(2008), Macro-panels and Reality, *Economics Letters* 99, 537-540.

CUBADDA, G., HECQ A. AND F.C. PALM (2009), Studying Co-movements in Large Multivariate Models Prior to Modeling, *Journal of Econometrics* 148, 25-35.

CUBADDA, G. AND A. HECQ (2011), Testing for Common Autocorrelation in Data Rich Environments, *Journal of Forecasting* 30, 325-335 .

DROST, F.C. AND T.E. NIJMAN (1993), Temporal Aggregation of GARCH Processes, *Econometrica* 61, 909-927.

ENGLE, R. F. (2002), Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models, *Journal of Business and Economic Statistics* 20, 339-350.

ENGLE, R. F., V. NG AND M. ROTHSCHILD (1990), Asset Pricing With a Factor ARCH Covariance Structure: Empirical Estimates of Treasury Bills, *Journal of Econometrics* 45, 213-238.

ENGLE, R. F. AND B. KELLY (2008), Dynamic Equicorrelation, Stern School of Business, WP NYU.

ENGLE, R. F. AND S. KOZICKI (1993), Testing for Common Features (with Comments), *Journal of Business and Economic Statistics* 11, 369-395.

ENGLE, R. F. AND G.J. LEE (1999), A Permanent and Transitory Component Model of Stock Return Volatility, in Engle and White (eds.) Cointegration, Causality and Forecasting: A Festschrift in Honor of Clive W. J. Granger, Oxford University Press, Oxford, 475-497.

- ENGLE, R. F. AND J. MARCUCCI (2006), A Long-run Pure Variance Common Features Model for the Common Volatilities of the Dow Jones, *Journal of Econometrics* 132, 7-42.
- ENGLE, R. F. AND R. SUSMEL (1993), Common Volatility in International Equity Markets, *Journal of Business and Economic Statistics* 11, 167-176.
- FIORENTINI, G., E. SENTANA AND N. SHEPHARD (2004), Likelihood-Based Estimation of Latent Generalized ARCH Structures, *Econometrica* 72, 1481-1517.
- FRANCQ, C. AND J.-M. ZAKOIAN (2007), HAC Estimation and Strong Linearity Testing in Weak ARMA Models, *Journal of Multivariate Analysis* 98, 114-144.
- HAFNER, C.M. AND H. HERWARTZ (2004), Testing for causality in variance using multivariate GARCH models. Econometric Institute Report 20, Erasmus University Rotterdam.
- HAFNER, C.M. AND A. PREMINGER (2009), Asymptotic theory for a factor GARCH model, *Econometric Theory* 25, 336-363.
- HECQ, A. AND J.V. ISSLER (2012), A Common-Feature Approach for Testing Present-Value Restrictions with Financial Data, Maastricht University, METEOR RM/12/006.
- LANNE, M. AND P. SAIKKONEN (2007), A Multivariate Generalized Orthogonal Factor GARCH Model, *Journal of Business and Economic Statistics* 25, 61-75.
- LIN, W.-L. (1992), Alternative Estimators for Factor GARCH Models - A Monte Carlo Comparison, *Journal of applied Econometrics* 7, 259-279.
- NIJMAN, T. AND E. SENTANA (1996), Marginalization and Contemporaneous Aggregation in Multivariate GARCH Processes, *Journal of Econometrics* 71, 71-87.
- RAVIKUMAR, B., RAY, S. AND E. SAVING (2000), Robust Wald tests in SUR systems with adding-up restrictions, *Econometrica* 68, 715-720.
- RUIZ, I.(2009), Common volatility across Latin American foreign exchange markets, *Applied Financial Economics* 19, 1197-1211.
- SILVENNOINEN, A. AND T. TERASVIRTA (2009), *Multivariate GARCH Models*, Handbook of Financial Time Series, Ed. Andersen, T.G. and Davis, R.A. and Kreiss, J.P. and T. Mikosch, Springer.
- TIAO, G. C. AND R. S. TSAY (1989), Model Specification in Multivariate Time Series (with Comments), *Journal of Royal Statistical Society, Series B* 51, 157-213.
- ZELLNER, A. AND F.C. PALM (1974), Time Series Analysis and Simultaneous Equation Econometric Models, *Journal of Econometrics* 2, 17-54.
- ZELLNER, A. AND F.C. PALM (1975), Time Series and Structural Analysis of Monetary Models of the US Economy, *Sanhya: The Indian Journal of Statistics, Series C* 37, 12-56.
- ZELLNER, A. AND F.C. PALM (2004), *The Structural Econometric Time Series Analysis Approach*, Cambridge University Press.

## 9 Result for the MGARCH(p,q)

In a MGARCH(p,q) if there exists a rank reduction in the ARCH part  $\Gamma_i$  only (and not in the GARCH  $G_j$ ), each component of the  $vech(\varepsilon_t \varepsilon_t')$  follows a weak ARMA( $(N-s) \max\{p, q\}, (N-s) \max\{p, q\} + p$ ) as in the general VARMA case with a reduced rank structure (see Proposition 2 and 3). This means that the moving average part of the ARMA representation is inflated by an additive factor  $p$ . Now if the coefficient matrices of the ARCH and the GARCH part share the same left null space, i.e. if there exists a matrix  $\delta$  such that  $\delta' \Gamma_i = \delta' G_j = 0$  for all  $i = 1$  to  $q$  and for all  $j = 1$  to  $p$ , the results must be adapted accordingly.

In the VARMA representation of the  $vech$  such that

$$Q(L)x_t = MG(L)M^{-1}Mv_t,$$

where  $x_t = Mvech(\varepsilon_t \varepsilon_t')$ ,  $e_t = Mv_t$ ,  $Q(L) = M\Phi(L)M^{-1}$ ,  $M' \equiv [\delta : \delta_\perp]$ ,  $\delta_\perp$  is the orthogonal complement of  $\delta$  with  $\text{span}(\delta_\perp) = \text{span}(\varphi)$ . Given that  $x_t$  is a non-singular linear transformation of  $vech(\varepsilon_t \varepsilon_t')$ , the maximum AR and MA orders of the univariate representation of the elements of  $vech(\varepsilon_t \varepsilon_t')$  must be the same as those of the elements of  $x_t$ . Since  $M^{-1} = [\bar{\delta} : \bar{\delta}_\perp]$ , where  $\bar{\delta} = \delta(\delta'\delta)^{-1}$ , and  $\bar{\delta}_\perp = \delta_\perp(\delta'_\perp \delta_\perp)^{-1}$ . In order to show the implications of this modeling, let us consider the adjoint of the block triangular matrices  $Q(L)$ ,  $Adj\{Q(L)\} = Q(L)^{-1} \det(Q(L))$  such that

$$Adj\{Q(L)\} = \begin{bmatrix} I_s & 0_{s \times (N-s)} \\ -(\delta'_\perp \Phi(L) \bar{\delta}_\perp)^{-1} \delta'_\perp \Phi(L) \bar{\delta} & (\delta'_\perp \Phi(L) \bar{\delta}_\perp)^{-1} \end{bmatrix} \det[\delta'_\perp \Phi(L) \bar{\delta}_\perp].$$

we have that

$$MG(L)M^{-1} = \begin{bmatrix} I_s & 0_{s \times (N-s)} \\ \delta'_\perp G(L) \bar{\delta} & \delta'_\perp G(L) \bar{\delta}_\perp \end{bmatrix}$$

and consequently

$$\begin{aligned}
& Adj\{Q(L)\}MG(L)M^{-1} \\
&= \det[\delta'_\perp \Phi(L)\bar{\delta}_\perp] \begin{bmatrix} I_s & 0_{s \times (N-s)} \\ -(\delta'_\perp \Phi(L)\bar{\delta}_\perp)^{-1} \delta'_\perp \Phi(L)\bar{\delta} & (\delta'_\perp \Phi(L)\bar{\delta}_\perp)^{-1} \end{bmatrix} \begin{bmatrix} I_s & 0_{s \times (N-s)} \\ \delta'_\perp G(L)\bar{\delta} & \delta'_\perp G(L)\bar{\delta}_\perp \end{bmatrix} \\
&= \det[\delta'_\perp \Phi(L)\bar{\delta}_\perp] \begin{bmatrix} I_s & 0_{s \times (N-s)} \\ -(\delta'_\perp \Phi(L)\bar{\delta}_\perp)^{-1} \delta'_\perp \Phi(L)\bar{\delta} + (\delta'_\perp \Phi(L)\bar{\delta}_\perp)^{-1} \delta'_\perp G(L)\bar{\delta} & (\delta'_\perp \Phi(L)\bar{\delta}_\perp)^{-1} \delta'_\perp G(L)\bar{\delta}_\perp \end{bmatrix} \\
&= \underbrace{\det[\delta'_\perp \Phi(L)\bar{\delta}_\perp]}_{(N-s) \max(q,p)} \times \\
&\quad \begin{bmatrix} - \underbrace{(\delta'_\perp \Phi(L)\bar{\delta}_\perp)^{-1}}_{(N-s-1) \max(q,p) - (N-s) \max(q,p)} \underbrace{\delta'_\perp \Phi(L)\bar{\delta}}_{\max(q,p)} + \underbrace{(\delta'_\perp \Phi(L)\bar{\delta}_\perp)^{-1}}_{(N-s-1) \max(q,p) - (N-s) \max(q,p)} \underbrace{\delta'_\perp G(L)\bar{\delta}}_p \\ \underbrace{(\delta'_\perp \Phi(L)\bar{\delta}_\perp)^{-1}}_{(N-s-1) \max(q,p) - (N-s) \max(q,p)} \underbrace{\delta'_\perp G(L)\bar{\delta}_\perp}_p \end{bmatrix},
\end{aligned}$$

where the lag polynomial orders are given below each element in the last expression. If we focus on the

maximum order bounds, it turns out that we obtain the final equation representation of orders

$$ARMA(\overbrace{p^* \leq (N-s) \max\{p, q\}}^{AR \text{ part}}, \overbrace{q^* \leq \max[(N-s) \max\{q, p\}, (N-s-1) \max\{q, p\} + p]}^{MA \text{ part}}).$$