The Relation between Conditionally Heteroskedastic Factor Models and Factor GARCH Models

Enrique Sentana

CEMFI

CEMFI, Casado del Alisal 5, 28014 Madrid, Spain.
Tel: 341 4290551, fax: 341 4291056, http://www.cemfi.es.

Abstract

The factor GARCH model of Engle (1987) and the latent factor ARCH model of Diebold and Nerlove (1989) have become rather popular multivariate volatility parameterizations due to their parsimony, and the commonality in volatility movements across different financial series. Nevertheless, there is some confusion in the literature between them. The purpose of this note is to make clear their similarities and differences by providing a formal nesting of the two models, which can be exploited to analyze their statistical features in a more general context. At the same time, their differences may be important in the interpretation of empirical results.

1   Introduction

Most of the econometric and finance literature on time-varying volatility models is concerned with univariate time series. However, many issues in finance can only be fully addressed within a multivariate framework. Although multivariate generalizations of the existing univariate models are straightforward in theory, their empirical applications have been hampered by the sheer number of parameters involved. Motivated by the commonality in volatility movements across different financial time series, several parsimonious alternatives have been proposed. Within the ARCH class, two such parameterizations have become rather popular, namely the factor GARCH model of Engle (1987) and the conditionally heteroskedastic latent factor model introduced by Diebold and Nerlove (1989) and extended by King, Sentana and Wadhwani (1994). Broadly speaking, Engle's factor GARCH model specifies the time-varying part of the covariance matrix as a function of a few linear combinations of the observed random variables, but leaves its constant part fully unrestricted. On the other hand, Diebold and Nerlove's model is a traditional statistical factor analysis model, with a diagonal idiosyncratic covariance matrix, in which the variances of the common factors are parameterized as univariate ARCH models.

Given that the two models are closely related, it is perhaps not surprising that there is some confusion in the literature on the similarities and differences between them, and in some cases they are treated as equivalent. Nevertheless, there are at least two important differences that distinguish them. Firstly, while the covariance matrix of a factor GARCH model is measurable by construction with respect to an information set that contains only past values of the observed variables, this is not the case for Diebold and Nerlove's original model, which in fact can be regarded as a stochastic volatility model (see e.g. Andersen (1992) and Shephard (1996)). A second, less well known distinctive feature, is that the implicit definition of the

factors is significantly different between the two. While in Diebold and Nerlove's model the factors capture the comovements between the observed series, in Engle's model they are directly related to those linear combinations of the series which summarize the comovements in their conditional variances.

The purpose of this note is to make clear their similarities and differences by providing a formal nesting of the two models. This is important for at least two reasons. First, while many properties of these models have already been studied in detail, it is not always entirely clear whether they are specific to the model analyzed, or apply to a broader class. Second, as we shall see, in some cases of substantive interest the interpretation of the empirical results may be sensitive to the specific nature of the model used.

The rest of the paper is organized as follows. We define the general class of conditionally heteroskedastic models in section 2, and discuss some of their properties. Then in section 3, we introduce the factor ARCH model, summarize its main properties, and present our main result. Finally, section 4 contains a discussion of our results in relation to empirical tests of multi-beta asset pricing theories. Two auxiliary lemmas are included in an appendix.

2 Conditionally Heteroskedastic Factor Models

Consider the following multivariate model:

$$\mathbf{x}_t = \mathbf{C}\mathbf{f}_t + \mathbf{w}_t \tag{1}$$

$$\mathbf{f}_t = \mathbf{\Lambda}_t^{1/2}\mathbf{f}_t^* \tag{2}$$

$$\begin{pmatrix} \mathbf{f}_t^* \\ \mathbf{w}_t \end{pmatrix} \sim i.i.d. \left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma} \end{pmatrix} \right] \tag{3}$$

where $\mathbf{x}_t$ is a $N \times 1$ vector of observable random variables, $\mathbf{f}_t$ is a $k \times 1$ vector of unobserved common factors, $\mathbf{C}$ is the $N \times k$ matrix of associated factor loadings,

with $N \geq k$ and rank $(\mathbf{C}) = k$, $\mathbf{w}_t$ is a $N \times 1$ vector of idiosyncratic noises, which are conditionally orthogonal to $\mathbf{f}_t$, $\mathbf{\Gamma}$ is a $N \times N$ positive semidefinite (p.s.d.) matrix of constant idiosyncratic variances, and $\mathbf{\Lambda}_t$ is a $k \times k$ positive definite (p.d.) matrix of time-varying factor variances.

Let us define the information sets $I_{t-1} = \{\mathbf{x}_{t-1}, \mathbf{f}_{t-1}, \mathbf{x}_{t-2}, \mathbf{f}_{t-2}, \ldots\}$ and $J_{t-1} = \{\mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \ldots\}$, with $J_{t-1} \subset I_{t-1}$. Our assumptions imply that the distribution of $\mathbf{x}_t$ conditional on I_{t-1} has zero mean, and covariance matrix $\mathbf{\Sigma}_t = \mathbf{C}\mathbf{\Lambda}_t\mathbf{C}' + \mathbf{\Gamma}$. For this reason, we shall refer to the data generation process specified by (1-3) as a multivariate conditionally heteroskedastic factor model.

Such a formulation nests several models widely used in the empirical literature, which typically assume that the unobserved factors follow dynamic heteroskedastic processes, but differ in the exact parameterization of $\mathbf{\Lambda}_t$ and $\mathbf{\Gamma}$. For instance, in the latent factor model with ARCH effects on the underlying factors introduced by Diebold and Nerlove (1989), the conditional variances of the factors are parameterized as univariate *strong* ARCH models, in the sense of Drost and Nijman (1993). In particular, for a GARCH (p,q) formulation,

$$\lambda_{jt} = \phi_j + \sum_{s=1}^{q} \alpha_{js} f_{jt-s}^2 + \sum_{r=1}^{q} \beta_{jr} \lambda_{jt-r} \qquad (4)$$

Notice that in this case, $\lambda_{jt} \in I_{t-1}$ but $\lambda_{jt} \notin J_{t-1}$, which has important inference implications (see Harvey, Ruiz and Sentana (1992), Fiorentini and Sentana (1997) and Shephard (1996)). Notice also that a significant characteristic of a model like (1-4) is that it is dimension independent, in the sense that it remains valid for any subset of $\mathbf{x}_t$.

Another frequently made assumption is that the common factors represent (conditionally) orthogonal influences, which implies that $\mathbf{\Lambda}_t$ is diagonal. Otherwise, we say that the factor model is oblique. Similarly, it is often assumed that $\mathbf{\Gamma}$ is diagonal, in which case the factor structure is termed strict or exact. However, in some applications, diagonality of $\mathbf{\Gamma}$ may be thought to be too restrictive.

3

For instance, it will not be preserved by linear transformations. For that reason, Chamberlain and Rothschild (1983) introduced the concept of approximate factor structures, in which the idiosyncratic terms may be correlated, but only up to a certain degree. Alternatively, one could assume that $\mathbf{\Gamma}$ has reduced rank. In fact, the rank of $\mathbf{\Gamma}$ is related to the observability of the factors. If rank$(\mathbf{\Gamma}) = N - k$, the factors would be fully revealed by the $\mathbf{x}_t$ variables, otherwise they are only partially revealed (see King, Sentana and Wadhwani (1994)).

A non-trivial advantage of these models is that they automatically guarantee a p.s.d. covariance matrix for $\mathbf{x}_t$ once we ensure that the covariance matrix of the factors is p.s.d. Moreover, using Lemma 1 in the Appendix, it is straightforward to derive necessary and sufficient conditions that guarantee that $\mathbf{\Sigma}_t$ is actually p.d. even when $\mathbf{\Gamma}$ is of reduced rank.

But the most distinctive feature of factor models is that they provide a parsimonious specification of the (dynamic) cross-sectional dependence of a vector of observable random variables. In this case in particular, the factor structure, together with the constancy of $\mathbf{\Gamma}$, implies that the time-variation of $\mathbf{\Sigma}_t$ is of reduced rank (see Engle, Ng and Rothschild (1990)). More formally, if $\ddot{\mathbf{E}}$ denotes a $N \times (N - k)$ matrix of full column rank such that $\ddot{\mathbf{E}}'\mathbf{C} = \mathbf{0}$, and $\bar{\mathbf{E}} = (\dot{\mathbf{E}}, \ddot{\mathbf{E}})$ is a $N \times N$ matrix of full rank, with $\dot{\mathbf{E}}$ arbitrary, then, the only time-varying component in the covariance matrix of $\bar{\mathbf{x}}_t' = \mathbf{x}_t'\bar{\mathbf{E}} = (\mathbf{x}_t'\dot{\mathbf{E}}, \mathbf{x}_t'\ddot{\mathbf{E}}) = (\dot{\mathbf{x}}_t', \ddot{\mathbf{x}}_t')$ is the covariance matrix of $\dot{\mathbf{x}}_t$. In this respect, Gourieroux, Monfort and Renault (1991) provide a particularly attractive choice of $\bar{\mathbf{E}}$, which makes $\dot{\mathbf{x}}_t$ and $\ddot{\mathbf{x}}_t$ orthogonal, and allows us to express any conditionally heteroskedastic factor model as an oblique factor model with time-varying conditional variances, constant conditional covariances, and a singular idiosyncratic covariance matrix. In particular, if we assume for simplicity that $\mathbf{\Gamma}$ is non-singular, we can write:

$$\mathbf{x}_t = \mathbf{C}\mathbf{f}_t^G + \mathbf{w}_t^G \tag{5}$$

4

$$\begin{pmatrix} \mathbf{f}_t^G \\ \mathbf{w}_t^G \end{pmatrix} \mid I_{t-1} \sim \left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{\Lambda}_t + (\mathbf{C}'\mathbf{\Gamma}^{-1}\mathbf{C})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma} - \mathbf{C}(\mathbf{C}'\mathbf{\Gamma}^{-1}\mathbf{C})^{-1}\mathbf{C}' \end{pmatrix} \right]$$

where

$$\mathbf{f}_t^G = (\mathbf{C}'\mathbf{\Gamma}^{-1}\mathbf{C})^{-1}\mathbf{C}'\mathbf{\Gamma}^{-1}\mathbf{x}_t = \mathbf{f}_t + (\mathbf{C}'\mathbf{\Gamma}^{-1}\mathbf{C})^{-1}\mathbf{C}'\mathbf{\Gamma}^{-1}\mathbf{w}_t$$

are the Generalized Least Squares (GLS) estimates of the common factors and $\mathbf{w}_t^G = \left[\mathbf{I} - (\mathbf{C}'\mathbf{\Gamma}^{-1}\mathbf{C})^{-1}\mathbf{C}'\mathbf{\Gamma}^{-1}\right]\mathbf{x}_t$.[1] Notice that even when $\mathbf{\Lambda}_t$ is diagonal, the elements of $\mathbf{f}_t^G$ are contemporaneously correlated, unless $(\mathbf{C}'\mathbf{\Gamma}^{-1}\mathbf{C})$ is itself diagonal. Also note that if $\mathbf{\Lambda}_t$ is given by a strong GARCH model such as (4), the results in Nijman and Sentana (1996) imply that these factor scores will follow weak GARCH processes.

Finally, notice that if $\mathbf{f}_t$ is conditionally homoskedastic, the model in (1-3) reduces to the standard (i.e. static) factor analysis model (see e.g. Johnson and Wichern (1992)). But even if $\mathbf{f}_t$ is conditionally heteroskedastic, provided that it is covariance stationary, it also implies an unconditional k factor structure for $\mathbf{x}_t$. That is, the unconditional covariance matrix, $\mathbf{\Sigma}$, can be written as:

$$\mathbf{\Sigma} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}' + \mathbf{\Gamma} \tag{6}$$

where $V(\mathbf{f}_t) = E(\mathbf{\Lambda}_t) = \mathbf{\Lambda}$. This property makes the models considered here compatible with traditional factor analysis.

3 Factor GARCH Models as Conditionally Heteroskedastic Factor Models

The factor GARCH model was originally introduced as a parsimonious special case of the multivariate GARCH(p,q) model considered in Bollerslev, Engle and

---

[1]These factor scores are different from the minimum (conditional) mean square error estimates, $\mathbf{f}_{t|t} = E(\mathbf{f}_t \mid \mathbf{x}_t, I_{t-1})$, but closely related as $\mathbf{f}_t^G = \left[\mathbf{I} + (\mathbf{C}'\mathbf{\Gamma}^{-1}\mathbf{C})^{-1}\mathbf{\Lambda}_t^{-1}\right]\mathbf{f}_{t|t}$.

Wooldridge (1988):

$$vech(\mathbf{\Sigma}_t) = \mathbf{\Phi} + \sum_{s=1}^{q} \mathbf{A}_s vech(\mathbf{x}_{t-s}\mathbf{x}_{t-s}') + \sum_{r=1}^{q} \mathbf{B}_r vech(\mathbf{\Sigma}_{t-r})$$

where $\mathbf{A}_s$ and $\mathbf{B}_r$ are square matrices of order $m(m+1)/2$, and $\mathbf{\Phi}$ is a vector of the same order. In fact, it could be regarded as a prototype of the positive definite BEEK parameterization in Engle and Kroner (1995), which is the most general covariance specification proposed so far within this class (see Kroner (1987) and Lin (1992)). Formally, Engle's (1987) k factor GARCH(p,q) model is characterized by the following data generation process:

$$\mathbf{x}_t = \mathbf{\Sigma}_t^{1/2}\mathbf{x}_t^*, \mathbf{x}_t^* \sim i.i.d. \; (\mathbf{0}, \mathbf{I}_N)$$

$$\mathbf{\Sigma}_t = \mathbf{\Psi} + \sum_{j=1}^{k} \mathbf{c}_j\mathbf{c}_j' \left[\sum_{s=1}^{q} \alpha_{js}(\mathbf{d}_j'\mathbf{x}_{t-s})^2 + \sum_{r=1}^{p} \beta_{jr}(\mathbf{d}_j'\mathbf{\Sigma}_{t-r}\mathbf{d}_j) \right] =$$

$$\mathbf{\Psi} + \sum_{j=1}^{k} \mathbf{c}_j\mathbf{c}_j'[\alpha_j(L)(\mathbf{d}_j'\mathbf{x}_t)^2 + \beta_j(L)(\mathbf{d}_j'\mathbf{\Sigma}_t\mathbf{d}_j)] \quad (7)$$

where $\mathbf{\Psi}$ is a $N \times N$ symmetric p.s.d. matrix of rank $N_1 \leq N$, $\mathbf{C} = (\mathbf{c}_1|\ldots|\mathbf{c}_k)$ and $\mathbf{D} = (\mathbf{d}_1|\ldots|\mathbf{d}_k)$ are $N \times k$ coefficient matrices of full column rank satisfying $\mathbf{D}'\mathbf{C} = \mathbf{I}_k$, with each column of $\mathbf{D}$ (or $\mathbf{C}$) normalized to avoid scale indeterminacy, and $\alpha_j(L) = \sum_{s=1}^{q} \alpha_{js}L^s$, $\beta_j(L) = \sum_{r=1}^{p} \beta_{jr}L^r$ are polynomials in the lag operator, with the roots of $1 - \beta_j(L)$ outside the unit circle.

The name factor GARCH stems from the fact that the time variation in the conditional variance can be summarized by k linear combinations of $\mathbf{x}_t$ which are univariate strong GARCH. In this respect, it can be regarded as the conditional variance counterpart to reduced rank vector autoregressive models. These linear combinations, $\mathbf{d}_j'\mathbf{x}_t$, referred to as "factor representing portfolios" by Ng, Engle and Rothschild (1992), are univariate strong GARCH(p,q) processes with conditional variance

$$\delta_{jt} = \vartheta_{jj} + \sum_{s=1}^{q} \alpha_{js}(\mathbf{d}_j'\mathbf{x}_{t-s})^2 + \sum_{r=1}^{p} \beta_{jr}\delta_{jt-r} = \frac{\vartheta_{jj}}{1 - \beta_j(1)} + \frac{\alpha_j(L)}{1 - \beta_j(L)}(\mathbf{d}_j'\mathbf{x}_t)^2$$

where $\vartheta_{jj} = \mathbf{d}'_j \mathbf{\Psi} \mathbf{d}_j$. Note, however, that in general $cov_{t-1}(\mathbf{d}'_i \mathbf{x}_t, \mathbf{d}'_j \mathbf{x}_t) = \mathbf{d}'_i \mathbf{\Psi} \mathbf{d}_j = \vartheta_{ij} \neq 0$ for $i \neq j$, so that the constant part of $V_{t-1}(\mathbf{D}' \mathbf{x}_t)$ is not usually diagonal (see Lin (1992)).

In order to guarantee that $\delta_{jt} \geq 0 \ \forall t$, it is often assumed that $\vartheta_{jj} \geq 0$ and the coefficients in the power expansion of $\alpha_j(L)/(1 - \beta_j(L))$ are non-negative (see Nelson and Cao (1992) and Drost and Nijman (1993)). In practice, it is also convenient to rule out cases in which $\mathbf{d}_j$ is in the nullspace of $\mathbf{\Psi}$, i.e. $\vartheta_{jj} = 0$, since otherwise δ_{jt}, and indeed $\mathbf{d}'_j \mathbf{x}_t$, will converge to zero with probability 1 (cf. Nelson (1990)).

An important property of factor GARCH models is that they are closed with respect to full-rank linear transformations (see Lin (1992)). However, they are not generally closed with respect to (block) marginalization (see Nijman and Sentana (1996)), as in most cases δ_{jt} becomes non-measurable with respect to a smaller information set (see the discussion after expression (4)).

Bollerslev and Engle (1993) show that $\mathbf{x}_t$ is covariance stationary if and only if $\alpha_j(1) + \beta_j(1) < 1 \ \forall j$. In that case, its unconditional covariance matrix is given by the following expression:

$$\mathbf{\Sigma} = \mathbf{\Psi} + \sum_{j=1}^{k} \mathbf{c}_j \mathbf{c}'_j \vartheta_{jj} \frac{\alpha_j(1) + \beta_j(1)}{1 - \alpha_j(1) - \beta_j(1)}$$

For our purposes, it is more convenient to write $\mathbf{\Sigma}_t$ as

$$\mathbf{\Sigma}_t = \mathbf{\Psi} - \sum_{j=1}^{k} \mathbf{c}_j \mathbf{c}'_j \vartheta_{jj} + \sum_{j=1}^{k} \mathbf{c}_j \mathbf{c}'_j \delta_{jt} = \mathbf{\Psi} + \sum_{j=1}^{k} \mathbf{c}_j \mathbf{c}'_j \left[\frac{\vartheta_{jj} \beta_j(1)}{1 - \beta_j(1)} + \frac{\alpha_j(L)}{1 - \beta_j(L)} (\mathbf{d}'_j \mathbf{x}_t)^2 \right]$$

or in matrix form

$$\mathbf{\Sigma}_t = \mathbf{\Psi} + \mathbf{C} \left[\mathbf{\Pi} + \mathbf{\Lambda}_t(\mathbf{0}) \right] \mathbf{C}' = \bar{\mathbf{\Psi}} + \mathbf{C} \mathbf{\Lambda}_t(\mathbf{0}) \mathbf{C}' \tag{8}$$

where

$$\mathbf{\Lambda}_t(\mathbf{0}) = diag \left[\frac{\alpha_1(L)}{1 - \beta_1(L)} (\mathbf{d}'_1 \mathbf{x}_t)^2, \ldots, \frac{\alpha_k(L)}{1 - \beta_k(L)} (\mathbf{d}'_k \mathbf{x}_t)^2 \right]$$

7

$$\mathbf{\Pi} = diag \left[\frac{\vartheta_{11}\beta_1(1)}{1 - \beta_1(1)}, \ldots, \frac{\vartheta_{kk}\beta_k(1)}{1 - \beta_k(1)} \right]$$

and $\bar{\mathbf{\Psi}} = \mathbf{\Psi} + \mathbf{C}\mathbf{\Pi}\mathbf{C}'$.

Equation (8) clearly suggests that Engle's model must be related to the conditionally heteroskedastic factor model in (1-3). In fact, as the following proposition shows, the factor GARCH model is observationally equivalent to a whole family of oblique factor models with constant conditional covariances, whose limiting cases are an orthogonal factor model in which the variances of the factors are given by $\mathbf{\Lambda}_t(\mathbf{0})$, and an oblique factor model like (5) with a singular idiosyncratic covariance matrix:

Proposition 1 *Assume that $\mathbf{\Sigma}_t$ is p.d., and furthermore that $\bar{\mathbf{\Psi}}$ has full rank. Let $\mathbf{-}$ be any $k \times k$ p.s.d. matrix such that the eigenvalues of $\mathbf{-}\, \mathbf{C}'\bar{\mathbf{\Psi}}^{-1}\mathbf{C}$ are less than or equal to 1. Then, the k factor GARCH(p,q) model in (7) is observationally equivalent (up to conditional second moments) to any of the following conditionally heteroskedastic factor models:*

$$\mathbf{x}_t = \mathbf{C}\mathbf{f}_t(\mathbf{-}) + \mathbf{w}_t(\mathbf{-})$$

where

$$V_{t-1}\left(\mathbf{f}_t(\mathbf{-})\right) = \mathbf{\Lambda}_t(\mathbf{-}) = \mathbf{-} + \mathbf{\Lambda}_t(\mathbf{0})$$
$$V_{t-1}\left(\mathbf{w}_t(\mathbf{-})\right) = \mathbf{\Gamma}(\mathbf{-}) = \bar{\mathbf{\Psi}} - \mathbf{C}\mathbf{-}\mathbf{C}'$$
$$cov_{t-1}\left(\mathbf{f}_t(\mathbf{-}), \mathbf{w}_t(\mathbf{-})\right) = \mathbf{0}$$

Proof. Since $V_{t-1}(\mathbf{x}_t) = \mathbf{\Gamma}(\mathbf{-}) + \mathbf{C}\mathbf{\Lambda}_t(\mathbf{-})\mathbf{C}' = \bar{\mathbf{\Psi}} + \mathbf{C}\mathbf{\Lambda}_t(\mathbf{0})\mathbf{C}' = \mathbf{\Sigma}_t$, we only need to check that $\mathbf{\Gamma}(\mathbf{-})$ and $\mathbf{\Lambda}_t(\mathbf{-})$ are p.s.d. Given that $\bar{\mathbf{\Psi}}$ is p.d., the restrictions on $\mathbf{-}$ guarantee that $\mathbf{\Gamma}(\mathbf{-})$ is p.s.d. in view of Lemma 2. Similarly, since $\delta_{jt} \geq \vartheta_{jj}/[1 - \beta_j(1)] \; \forall t$, $\lambda_{jt}(\mathbf{-}) \geq \lambda_{jt}(\mathbf{0}) = \delta_{jt} - \vartheta_{jj}/[1 - \beta_j(1)] \geq 0$ $\qquad \square$

It is also straightforward to check that the GLS factor representing portfolios of the above models, $\mathbf{f}_t^G(\mathbf{-}) = (\mathbf{C}'\mathbf{\Gamma}^{-1}(\mathbf{-})\mathbf{C})^{-1}\mathbf{C}'\mathbf{\Gamma}^{-1}(\mathbf{-})\mathbf{x}_t$ do not depend on $\mathbf{-}$, and coincide with $\mathbf{f}_t\left((\mathbf{C}'\bar{\mathbf{\Psi}}^{-1}\mathbf{C})^{-1}\right) = (\mathbf{C}'\bar{\mathbf{\Psi}}^{-1}\mathbf{C})^{-1}\mathbf{C}'\bar{\mathbf{\Psi}}^{-1}\mathbf{x}_t$. In this respect, it is worth noting that the "largest" $\mathbf{-}$ compatible with $\mathbf{\Gamma}(\mathbf{-})$ p.s.d.

is precisely $(\mathbf{C}'\bar{\boldsymbol{\Psi}}^{-1}\mathbf{C})^{-1}$. Therefore, if we call $(\mathbf{C}'\bar{\boldsymbol{\Psi}}^{-1}\mathbf{C})^{jj}$ the jth diagonal element of $(\mathbf{C}'\bar{\boldsymbol{\Psi}}^{-1}\mathbf{C})^{-1}$, $\lambda_{jt}(\text{-}\,) \leq \lambda_{jt}\left((\mathbf{C}'\bar{\boldsymbol{\Psi}}^{-1}\mathbf{C})^{-1}\right) = \delta_{jt} + (\mathbf{C}'\bar{\boldsymbol{\Psi}}^{-1}\mathbf{C})^{jj} - \mathbf{d}_j'\bar{\boldsymbol{\Psi}}\mathbf{d}_j \leq \delta_{jt}$ by the Cauchy-Schwartz inequality, with equality if and only if $\mathbf{D}'= (\mathbf{C}'\bar{\boldsymbol{\Psi}}^{-1}\mathbf{C})^{-1}\mathbf{C}'\bar{\boldsymbol{\Psi}}^{-1}$. Except in this unlikely case, $\bar{\boldsymbol{\Psi}} - \mathbf{D}'\bar{\boldsymbol{\Psi}}\mathbf{D}$ is an indefinite matrix, and we cannot interpret $\mathbf{D}'\mathbf{x}_t$ as common factors. As a result, $\mathbf{f}_t(\text{-}\,)$ will generally follow weak, rather than strong, GARCH processes, since $\mathbf{D}'\mathbf{x}_t \neq \mathbf{f}_t(\text{-}\,)$ for all admissible $\text{-}\,$. In any case, we can replace the infinite distributed lags in $\boldsymbol{\Lambda}_t(\text{-}\,)$ by the recursions

$$\lambda_{jt}(\text{-}\,) = [1 - \beta_j(1)]\omega_{jj} + \alpha_j(L)(\mathbf{d}_j'\mathbf{x}_t)^2 + \beta_j(L)\lambda_{jt}(\text{-}\,)$$

Finally, note that $\text{rank}(\bar{\boldsymbol{\Psi}}) = N$ is very slightly stronger than required, as it only excludes those cases in which $\text{rank}(\mathbf{U}_2'\mathbf{C}\boldsymbol{\Pi}^{1/2}) < N - N_1$ but $\text{rank}(\mathbf{U}_2'\mathbf{C}) = \text{rank}(\mathbf{U}_2'\mathbf{C}\left[\boldsymbol{\Pi} + \boldsymbol{\Lambda}_t(0)\right]^{1/2}) = N - N_1$, where the columns of $\mathbf{U}_2$ constitute a basis of the nullspace of $\boldsymbol{\Psi}$. Nevertheless, Proposition 1 can be tediously extended on the basis of Lemma 2b in the appendix to those cases in which $\boldsymbol{\Sigma}_t$ is p.d. but $\bar{\boldsymbol{\Psi}}$ is not, at the cost of making the range of admissible $\text{-}\,'s$ depend on k and the rank of $\bar{\boldsymbol{\Psi}}$. For instance, if $k = 1$, $\text{-}\,$ can be any non-negative scalar such that $\text{-}\,\mathbf{C}'\bar{\boldsymbol{\Psi}}^{-1}\mathbf{C} \leq 1$ if $\text{rank}(\bar{\boldsymbol{\Psi}}) = N$, and 0 if $\text{rank}(\bar{\boldsymbol{\Psi}}) = N - 1$.

4 Summary and Discussion

In this note, we discuss the relation between Engle's (1987) factor GARCH model, and a general class of conditionally heteroskedastic factor models, which includes the latent factor ARCH model of Diebold and Nerlove (1989) as a special case. We formally introduce both models, discuss some of their properties, and provide a precise nesting of the two.

Such a nesting can be fruitfully exploited to analyze the statistical features of the factor GARCH model in the more general context of model (1-3). For example,

9

Demos and Sentana (1997) exploit this relationship to adapt the EM algorithm to this class of models, while Sentana and Fiorentini (1997) analyze the identifiability of the factor loadings matrix $\mathbf{C}$ within a unified framework. Similarly, it can also be employed to extend some of the inference procedures developed for the factor GARCH model to the general class of models in (1-3). This would be the case, for instance, of the common features test in Engle and Kozicki (1993), under our maintained assumption of constant idiosyncratic variances.

Nevertheless, one should always remain aware of their differences, especially in empirical applications. For instance, an implication of our main result is that if we interpret the factor GARCH model as a traditional factor model in which the variances of the factors change over time, we would not generally be able to separately identify the contributions to the unconditional covariance matrix of the idiosyncratic terms and common factors, as we can transfer variance between them (see Sentana and Fiorentini (1997)). This problem is particularly relevant in testing some of the implications of multi-beta asset pricing theories, such as the integration of financial markets (see King, Sentana and Wadhwani (1994)).

Similarly, even if we avoid the aforementioned observational equivalence, we will not be able to differentiate the contribution to the covariance matrix of conditionally homoskedastic common factors (see Engle, Ng and Rothschild (1990)). For instance, if for identification purposes we go to the extreme, and choose the parameters so that the idiosyncratic covariance matrix has rank $N-k$, then model (7) can also be written as a model with N common factors and no idiosyncratic noise, in which the first k factors have time-varying variances but constant covariances, and the remaining $N-k$ factors are conditionally homoskedastic, as well as orthogonal to each other and to the first k ones. Obviously, the corresponding asset pricing implications are rather different.

References

Andersen, T.G. (1992): "Volatility", Finance Working Paper 144, Kellog Graduate School of Management, Northwestern University.

Bollerslev, T. and Engle, R.F. (1993): "Common persistence in conditional variances", *Econometrica* 61, 166-187.

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

Chamberlain, G. and Rothschild, M. (1983): "Arbitrage, factor structure, and mean-variance analysis on large asset markets", *Econometrica* 51, 1281-1304.

Demos, A. and Sentana, E. (1997): "An EM algorithm for conditionally heteroskedastic factor models", forthcoming in *Journal of Business and Economic Statistics*.

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

Drost, F.C. and Nijman, T.E. (1993): "Temporal aggregation of GARCH processes", *Econometrica* 61, 909-927.

Engle, R.F. (1987): "Multivariate ARCH with factor structures - Cointegration in variance", mimeo, University of California at San Diego.

Engle, R.F. and Kozicki, S. (1993): "Testing for common features", *Journal of Business and Economic Statistics* 11, 369-380.

Engle, R.F. and Kroner, K. F. (1995): "Multivariate simultaneous generalized ARCH", *Econometric Theory* 11, 122-150.

Engle, R.F., Ng, V.M. and Rothschild, M. (1990): "Asset pricing with a factor-ARCH structure: Empirical estimates for Treasury Bills", *Journal of Econometrics* 45, 213-237.

11

Fiorentini, G. and Sentana, E. (1997): "Exact maximum likelihood estimation of conditionally heteroskedastic factor models", mimeo, CEMFI.

Gourieroux, C., Monfort, A. and Renault, E. (1991): "A general framework for factor models", mimeo, INSEE.

Harvey, A.C., Ruiz, E. and Sentana, E. (1992): "Unobservable component time series models with ARCH disturbances", *Journal of Econometrics* 52, 129-157.

Johnson, R.A. and Wichern, D.W. (1992): *Applied Multivariate Statistical Analysis*, 3rd edition, Prentice-Hall.

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

Kroner, K.F. (1987): "Estimating and testing for factor GARCH", mimeo, University of California at San Diego.

Lin, W.L. (1992): "Alternative estimators for factor GARCH models: A Monte Carlo comparison", *Journal of Applied Econometrics* 7, 259-279.

Nelson, D.B. (1990): "Stationarity and persistence in the GARCH(1,1) model", *Econometric Theory* 6, 318-334.

Nelson, D.B. and Cao, C.Q. (1992): "Inequality constraints in the univariate GARCH model", *Journal of Business and Economic Statistics* 10, 229-235.

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

Nijman, T. and Sentana, E. (1996): "Marginalization and contemporaneous aggregation of multivariate GARCH processes", *Journal of Econometrics* 71, 71-87.

Sentana, E. and Fiorentini, G. (1997): "Identification, estimation and testing of conditionally heteroskedastic factor models", CEMFI Working Paper 9709.

Shephard, N. (1996): "Statistical aspects of ARCH and stochastic volatility", in *Time Series Models with Econometric, Finance and Other Applications*, D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielson, eds., Chapman and Hall.

Appendix

The following lemmas are useful to derive the relationship between the k factor GARCH in (7), and conditionally heteroskedastic factor models in (1-3).

Lemma 1 *Let $\mathbf{Z}$ be a $N \times k$ matrix of rank k, $\boldsymbol{\Gamma}$ a $k \times k$ p.s.d. matrix with rank $k_1 \leq k$, and $\boldsymbol{\Psi}$ a $N \times N$ p.s.d. matrix with rank $N_1 \leq N$, where the spectral decompositions of $\boldsymbol{\Gamma}$ and $\boldsymbol{\Psi}$ are given by*

$$\boldsymbol{\Gamma} = \mathbf{V}\boldsymbol{\Phi}\mathbf{V}' = \begin{pmatrix} \mathbf{V}_1 & \mathbf{V}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\Phi}_1 & \\ & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{V}_1' \\ \mathbf{V}_2' \end{pmatrix} = \mathbf{V}_1\boldsymbol{\Phi}_1\mathbf{V}_1'$$

$$\boldsymbol{\Psi} = \mathbf{U}\boldsymbol{\Delta}\mathbf{U}' = \begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\Delta}_1 & \\ & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{U}_1' \\ \mathbf{U}_2' \end{pmatrix} =$$

$$\begin{pmatrix} \mathbf{U}_1\boldsymbol{\Delta}_1^{1/2} & \mathbf{U}_2 \end{pmatrix} \begin{pmatrix} \mathbf{I}_{N_1} & \\ & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Delta}_1^{1/2}\mathbf{U}_1' \\ \mathbf{U}_2' \end{pmatrix} = \tilde{\mathbf{U}}\tilde{\boldsymbol{\Delta}}\tilde{\mathbf{U}}'$$

with $\boldsymbol{\Phi}_1, \boldsymbol{\Delta}_1 > \mathbf{0}$, $\mathrm{rank}(\boldsymbol{\Phi}_1) = k_1$ and $\mathrm{rank}(\boldsymbol{\Delta}_1) = N_1$, so that $\boldsymbol{\Psi}^+ = \mathbf{U}_1\boldsymbol{\Delta}_1^{-1}\mathbf{U}_1'$ is the Moore-Penrose inverse of $\boldsymbol{\Psi}$. Then,
a) If $\mathrm{rank}(\boldsymbol{\Psi}) = N, \mathrm{rank}(\boldsymbol{\Psi} + \mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}') = \mathrm{rank}(\boldsymbol{\Psi})$
b) If $\mathrm{rank}(\boldsymbol{\Psi}) < N, \mathrm{rank}(\boldsymbol{\Psi} + \mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}') = \mathrm{rank}(\boldsymbol{\Psi}) + \mathrm{rank}(\mathbf{U}_2'\mathbf{Z}\mathbf{V}_1)$

Proof. We only prove b), as a) is trivial. Given that $\mathbf{U}$ is orthogonal, we can write

$$\boldsymbol{\Psi} + \mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}' = \mathbf{U}(\boldsymbol{\Delta} + \mathbf{U}'\mathbf{Z}\mathbf{V}_1\boldsymbol{\Phi}_1\mathbf{V}_1'\mathbf{Z}'\mathbf{U})\mathbf{U}' = \mathbf{U}\boldsymbol{\Upsilon}\mathbf{U}'$$

with

$$\boldsymbol{\Upsilon} = \begin{pmatrix} \boldsymbol{\Upsilon}_{11} & \boldsymbol{\Upsilon}_{12} \\ \boldsymbol{\Upsilon}_{12}' & \boldsymbol{\Upsilon}_{22} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Delta}_1 + \mathbf{U}_1'\mathbf{Z}\mathbf{V}_1\boldsymbol{\Phi}_1\mathbf{V}_1'\mathbf{Z}'\mathbf{U}_1 & \mathbf{U}_1'\mathbf{Z}\mathbf{V}_1\boldsymbol{\Phi}_1\mathbf{V}_1'\mathbf{Z}'\mathbf{U}_2 \\ \mathbf{U}_2'\mathbf{Z}\mathbf{V}_1\boldsymbol{\Phi}_1\mathbf{V}_1'\mathbf{Z}'\mathbf{U}_1 & \mathbf{U}_2'\mathbf{Z}\mathbf{V}_1\boldsymbol{\Phi}_1\mathbf{V}_1'\mathbf{Z}'\mathbf{U}_2 \end{pmatrix}$$

so that $\mathrm{rank}(\boldsymbol{\Psi} + \mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}') = \mathrm{rank}(\boldsymbol{\Upsilon})$. But since $\mathrm{rank}(\boldsymbol{\Upsilon}_{11}) = N_1$, $\boldsymbol{\Upsilon}$ has the same rank as

$$\begin{pmatrix} \mathbf{I}_{N_1} & \mathbf{0} \\ -\boldsymbol{\Upsilon}_{12}'\boldsymbol{\Upsilon}_{11}^{-1} & \mathbf{I}_{N-N_1} \end{pmatrix} \boldsymbol{\Upsilon} \begin{pmatrix} \mathbf{I}_{N_1} & -\boldsymbol{\Upsilon}_{11}^{-1}\boldsymbol{\Upsilon}_{12} \\ \mathbf{0} & \mathbf{I}_{N-N_1} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Upsilon}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Upsilon}_{22} - \boldsymbol{\Upsilon}_{12}'\boldsymbol{\Upsilon}_{11}^{-1}\boldsymbol{\Upsilon}_{12} \end{pmatrix}$$

13

Hence, $\operatorname{rank}(\boldsymbol{\Upsilon}) = \operatorname{rank}(\boldsymbol{\Upsilon}_{11}) + \operatorname{rank}(\boldsymbol{\Upsilon}_{22} - \boldsymbol{\Upsilon}'_{12}\boldsymbol{\Upsilon}_{11}^{-1}\boldsymbol{\Upsilon}_{12})$. Finally, repeated application of the Woodbury formula for $\boldsymbol{\Upsilon}_{11}$ gives after some simplification

$$\boldsymbol{\Upsilon}_{22} - \boldsymbol{\Upsilon}'_{12}\boldsymbol{\Upsilon}_{11}^{-1}\boldsymbol{\Upsilon}_{12} = \mathbf{U}'_2\mathbf{Z}\mathbf{V}_1(\boldsymbol{\Phi}_1^{-1} + \mathbf{V}'_1\mathbf{Z}'\boldsymbol{\Psi}^+\mathbf{Z}\mathbf{V}_1)^{-1}\mathbf{V}'_1\mathbf{Z}'\mathbf{U}_2$$

□

As a simple example, consider the case of $\operatorname{rank}(\boldsymbol{\Psi}) < N$ and $k = 1$, so that $\mathbf{V}_1 = 1$ if $\text{-} \neq 0$. Then $\operatorname{rank}(\boldsymbol{\Psi} + \mathbf{Z}\text{-}\,\mathbf{Z}') = \operatorname{rank}(\boldsymbol{\Psi}) + 1$ unless $\mathbf{U}'_2\mathbf{Z} = 0$, in which case $\mathbf{Z}$ is in the column space of $\mathbf{U}_1$.

Lemma 2 *Let $\mathbf{Z}, \text{-}$ and $\boldsymbol{\Psi}$ be defined as in Lemma 1. Then:*
a) If $\operatorname{rank}(\boldsymbol{\Psi}) = N$, $\boldsymbol{\Psi} - \mathbf{Z}\text{-}\,\mathbf{Z}'$ remains p.s.d. if and only if the eigenvalues of $\text{-}\,(\mathbf{Z}'\boldsymbol{\Psi}^{-1}\mathbf{Z})$ are all less than or equal to 1.
b) If $\operatorname{rank}(\boldsymbol{\Psi}) < N$, $\boldsymbol{\Psi} - \mathbf{Z}\text{-}\,\mathbf{Z}'$ remains p.s.d. if and only if $\mathbf{U}'_2\mathbf{Z}\mathbf{V}_1 = \mathbf{0}$ and the eigenvalues of $\text{-}\,(\mathbf{Z}'\boldsymbol{\Psi}^+\mathbf{Z})$ are all less than or equal to 1.

Proof. We only prove b), as a) is proved along similar lines. Since $\tilde{\mathbf{U}}$ has full rank, we can always write

$$\boldsymbol{\Psi} - \mathbf{Z}\text{-}\,\mathbf{Z}' = \tilde{\mathbf{U}}(\tilde{\boldsymbol{\Delta}} - \tilde{\mathbf{U}}^{-1}\mathbf{Z}\text{-}\,\mathbf{Z}'\tilde{\mathbf{U}}'^{-1})\tilde{\mathbf{U}}' = \tilde{\mathbf{U}}\tilde{\boldsymbol{\Upsilon}}\tilde{\mathbf{U}}'$$

Therefore, the definiteness of $\boldsymbol{\Psi} - \mathbf{Z}\text{-}\,\mathbf{Z}'$ is the same as the definiteness of $\tilde{\boldsymbol{\Upsilon}}$. In this case, $\tilde{\boldsymbol{\Upsilon}}$ is

$$\begin{pmatrix} \mathbf{I}_{N_1} - \boldsymbol{\Delta}_1^{-1/2}\mathbf{U}'_1\mathbf{Z}\mathbf{V}_1\boldsymbol{\Phi}_1\mathbf{V}'_1\mathbf{Z}'\mathbf{U}_1\boldsymbol{\Delta}_1^{-1/2} & -\boldsymbol{\Delta}_1^{-1/2}\mathbf{U}'_1\mathbf{Z}\mathbf{V}_1\boldsymbol{\Phi}_1\mathbf{V}'_1\mathbf{Z}'\mathbf{U}_2 \\ \mathbf{U}'_2\mathbf{Z}\mathbf{V}_1\boldsymbol{\Phi}_1\mathbf{V}'_1\mathbf{Z}'\mathbf{U}_1\boldsymbol{\Delta}_1^{-1/2} & -\mathbf{U}'_2\mathbf{Z}\mathbf{V}_1\boldsymbol{\Phi}_1\mathbf{V}'_1\mathbf{Z}'\mathbf{U}_2 \end{pmatrix}$$

which is p.s.d. if and only if $\mathbf{U}'_2\mathbf{Z}\mathbf{V}_1 = \mathbf{0}$ and $\mathbf{I}_{N_1} - \boldsymbol{\Delta}_1^{-1/2}\mathbf{U}'_1\mathbf{Z}\mathbf{V}_1\boldsymbol{\Phi}_1\mathbf{V}'_1\mathbf{Z}'\mathbf{U}_1\boldsymbol{\Delta}_1^{-1/2}$ is p.s.d. But this matrix has $N_1 - k_1$ unit eigenvalues, plus another k_1 eigenvalues which are 1 minus the eigenvalues of $\text{-}\,\mathbf{Z}'\mathbf{U}_1\boldsymbol{\Delta}_1^{-1/2}\boldsymbol{\Delta}_1^{-1/2}\mathbf{U}'_1\mathbf{Z} = \text{-}\,\mathbf{Z}'\boldsymbol{\Psi}^+\mathbf{Z}$ □