A MULTIVARIATE SKEW-GARCH MODEL

Giovanni De Luca, Marc G. Genton and Nicola Loperfido

ABSTRACT

Empirical research on European stock markets has shown that they behave differently according to the performance of the leading financial market identified as the US market. A positive sign is viewed as good news in the international financial markets, a negative sign means, conversely, bad news. As a result, we assume that European stock market returns are affected by endogenous and exogenous shocks. The former raise in the market itself, the latter come from the US market, because of its most influential role in the world. Under standard assumptions, the distribution of the European market index returns conditionally on the sign of the one-day lagged US return is skew-normal. The resulting model is denoted Skew-GARCH. We study the properties of this new model and illustrate its application to time-series data from three European financial markets.

1. INTRODUCTION

The pioneering work of Engle (1982) has represented the starting point of a tremendous scientific production with the aim of modeling and forecasting

Econometric Analysis of Financial and Economic Time Series/Part A
Advances in Econometrics, Volume 20, 33–56

ISSN: 0731-9053/doi:10.1016/S0731-9053(05)20002-6

the volatility of financial time series. The AutoRegressive Conditional Heteroskedasticity (ARCH) model has been dealt with in depth. Many variants have been proposed. Among them, we emphasize its most popular generalizations, the Generalized ARCH model (Bollerslev, 1986) and the Exponential GARCH model (Nelson, 1991) allowing for the inclusion of the asymmetric effect of volatility. Moving from a univariate to a multivariate perspective, the multivariate GARCH model is quite interesting because it can shed light on the common movements of the volatilities across markets (Bollerslev, Engle, & Wooldridge, 1988; Bollerslev, 1990; Engle, 2002).

When the analysis focuses on one or more markets, the possible relevance of an external leading market is usually ignored. Nonetheless, it is an important point which can help explaining some empirically detected features. Actually, a wide literature has dealt with the issue of the international transmission of stock markets movements. Eun and Shim (1989) stressed the most influential role of the US stock market. Innovations in the US market are transmitted to the other markets. Conversely, none of the other markets can affect the US market movements. The time horizon of the transmission is very short: the other stock exchanges' responses rapidly decrease after one day. The conclusion of Eun and Shim (1989) suggests to consider the US market as *the most important producer of information affecting the world stock market.* The contemporaneous and lead/lag relationships among stock markets are also studied in Koch and Koch (1991).

The analysis of univariate and multivariate GARCH models has traditionally neglected this aspect, with few exceptions. For instance, Lin, Engle, and Ito (1994) carried out an empirical investigation of the relationship between returns and volatilities of Tokyo and New York markets. The peculiarity of the study is the use of a decomposition of daily return into two components: the daytime return and the overnight return. The conclusion is the existence of cross-market interdependence in returns and volatilities. Karolyi (1995) detected the interdependence between the US and Canadian markets through a bivariate GARCH model.

Our analysis starts from the empirical detection of the different behavior of three European markets, according to the performance of the leading market identified as the US market. From a statistical perspective, we assume that European stock market returns are affected by endogenous and exogenous shocks. The former raise in the market itself, the latter come from the US market, defined as the leading market because of its most influential role in the world. Moreover, the flow of information from the US market to the European markets is asymmetric in its direction as well as in its effects. We recognize a negative (positive) performance of the US market as a proxy for bad (good) news for the world stock market. European financial markets returns behave in a different way according to bad or good news. Moreover, they are more reactive to bad news than to good ones.

The European and the US markets are not synchronous. When European markets open on day t, the US market is still closed; when European markets close on the same day, the US market is about to open or has just opened (depending on the European country). This implies a possible causal relationship from the US market return at time $t-1$ to the European returns at time t.

The distribution of European returns changes according to the sign of the one-day lagged performance of the US market. Average returns are negative (positive) in presence of bad (good) news and they are very similar in absolute value. Volatility is higher in presence of bad news. Skewness is negative (positive) and more remarkable in presence of bad (good) news. In both cases, a high degree of leptokurtosis is observed. Finally, bad news involves a stronger correlation between present European returns and the one-day lagged US return.

Allowing for a GARCH structure for taking into account the heteroskedastic nature of financial time series, under standard assumptions, the distribution of the European returns conditionally on news (that is, on the sign of the one-day lagged US return) and past information turns out to be skew-normal (Azzalini, 1985). This is a generalization of the normal distribution with an additional parameter to control skewness. The two conditional distributions are characterized by different features according to the type of news (bad or good). In particular, the skewness can be either negative (bad news) or positive (good news). The resulting model is denoted Skew-GARCH (henceforth SGARCH). The theoretical features of the model perfectly match the empirical evidence.

The basic idea can be extended to a multivariate setting. The international integration of financial markets is more remarkable in presence of a geographical proximity. The European markets tend to show common movements. Under standard assumptions, the joint distribution of European stock market returns conditionally on the sign of the one-day lagged US market return and past information is a multivariate skew-normal distribution (Azzalini & Dalla Valle, 1996), whose density is indexed by a location vector, a scale matrix and a shape vector. Finally, unconditional (with respect to the performance of the US market) returns have some features in concordance with empirical evidence.

The paper is organized as follows. Section 2 describes the theory of the skew-normal distribution. In Section 3, the multivariate SGARCH model is presented. In Section 4, the conditional distribution and related moments are obtained. Some special cases are described, including the univariate model when the dimensionality reduces to one. Section 5 refers to the unconditional distribution. Section 6 exhibits the estimates of the univariate and multivariate models applied to three small financial markets in Europe: Dutch, Swiss and Italian. The results show the relevance of the performance of the leading market supporting the proposal of the SGARCH model. Section 7 concludes. Some proofs are presented in the appendix.

2. THE SKEW-NORMAL DISTRIBUTION

The distribution of a random vector z is multivariate skew-normal (*SN*, henceforth) with location parameter ζ, scale parameter Ω and shape parameter α, that is $z \sim SN_p(\zeta, \Omega, \alpha)$, if its probability density function (pdf) is

$$f(z; \zeta, \Omega, \alpha) = 2\phi_p(z-\zeta; \Omega)\Phi\left[\alpha^{\mathrm{T}}(z-\zeta)\right], \quad z, \zeta, \alpha \in \mathbb{R}^p,\ \Omega \in \mathbb{R}^{p\times p}$$

where $\Phi(\cdot)$ is the cdf of a standardized normal variable and $\phi_p(z-\zeta;\Omega)$ is the density function of a p-dimensional normal distribution with mean ζ and variance Ω. For example, $Z \sim SN_1(\zeta, \omega, \alpha)$ denotes a random variable whose distribution is univariate *SN* with pdf

$$f(z; \zeta, \omega, \alpha) = \frac{2}{\sqrt{\omega}}\phi\left(\frac{z-\zeta}{\sqrt{\omega}}\right)\Phi[\alpha(z-\zeta)]$$

where $\phi(\cdot)$ is the pdf of a standard normal variable.

Despite the presence of an additional parameter, the *SN* distribution resembles the normal one in several ways, formalized through the following properties:

Inclusion: The normal distribution is an *SN* distribution with shape parameter equal to zero:

$$z \sim SN_p(\zeta, \Omega, 0) \Leftrightarrow z \sim N_p(\zeta, \Omega)$$

Greater norms of α imply greater differences between the density of the multivariate $SN_p(\zeta, \Omega, \alpha)$ and the density of the multivariate $N_p(\zeta, \Omega)$.

Linearity: The class of *SN* distributions is closed with respect to linear transformations. If A is a $k \times p$ matrix and $\mathrm{b} \in \mathbb{R}^k$, then

$$z \sim SN_p(\zeta, \Omega, \alpha) \Rightarrow Az + \mathrm{b} \sim SN_k\left(A\zeta + b, A\Omega A^{\mathrm{T}}, \bar{\alpha}\right)$$
$$\bar{\alpha} = \frac{\left(A\Omega A^{\mathrm{T}}\right)^{-1} A\Omega\alpha}{\sqrt{1 + \alpha^{\mathrm{T}}\Omega\left(\Omega^{-1} - A^{\mathrm{T}}\left(A\Omega A^{\mathrm{T}}\right)^{-1} A\right)\Omega\alpha}}$$

It follows that the *SN* class is closed under marginalization: subvectors of *SN* vectors are *SN*, too. In particular, each component of an *SN* random vector is univariate *SN*.

Invariance: The matrix of squares and products $(\mathrm{z}-\zeta)(\mathrm{z}-\zeta)^{\mathrm{T}}$ has a Wishart distribution:

$$(\mathrm{z}-\zeta)(\mathrm{z}-\zeta)^{\mathrm{T}} \sim W(\Omega, 1)$$

Notice that the distribution of $(\mathrm{z}-\zeta)(\mathrm{z}-\zeta)^{\mathrm{T}}$ does not depend on the shape parameter. In the univariate case, it means that $Z \sim SN_1(\zeta, \omega, \alpha)$ implies $(Z-\zeta)^2/\omega \sim \chi_1^2$.

All moments of the *SN* distribution exist and are finite. They have a simple analytical form. However, moments of the *SN* distribution differ from the normal ones in several ways:

- Location and scale parameters equal mean and variance only if the shape parameter vector α equals zero.
- Moments are more conveniently represented through the parameter $\delta = \Omega\alpha/\sqrt{1+\alpha^{\mathrm{T}}\Omega\alpha}$.
- Tails of the *SN* distribution are always heavier than the normal ones, when the shape parameter vector α differs from zero.

Table 1 reports the expectation, variance, skewness and kurtosis of the *SN* distribution, in the multivariate and univariate cases.

Multivariate skewness and kurtosis are evaluated through Mardia's indices (Mardia, 1970). Notice that in the univariate case Mardia's index of kurtosis equals the fourth moment of the standardized random variable. On the other hand, in the univariate case, Mardia's index of skewness equals the square of the third moment of the standardized random variable. In the following sections, when dealing with skewness, we shall refer to Mardia's index in the multivariate case and to the third moment of the standardized random variable in the univariate case.

Table 1. Expectation, Variance, Skewness and Kurtosis of the *SN* Distribution.

	$z \sim SN_p(\zeta, \Omega, \alpha)$	$Z \sim SN_1(\zeta, \omega, \alpha)$
Expectation	$\zeta + \sqrt{\frac{2}{\pi}}\delta$	$\zeta + \sqrt{\frac{2}{\pi}}\delta$
Variance	$\Omega - \frac{2}{\pi}\delta\delta^{\mathrm{T}}$	$\omega - \frac{2}{\pi}\delta^2$
Skewness	$2(4-\pi)^2\left(\frac{\delta^{\mathrm{T}}\Omega^{-1}\delta}{\pi - 2\delta^{\mathrm{T}}\Omega^{-1}\delta}\right)^3$	$\sqrt{2}(4-\pi)\left(\frac{\delta}{\sqrt{\pi\omega - 2\delta^2}}\right)^3$
Kurtosis	$8(\pi-3)\left(\frac{\delta^{\mathrm{T}}\Omega^{-1}\delta}{\pi - 2\delta^{\mathrm{T}}\Omega^{-1}\delta}\right)^2$	$8(\pi-3)\left(\frac{\delta^2}{\pi\omega - 2\delta^2}\right)^2$

3. THE SGARCH MODEL

Let Y_t be the *leading* (US) market return at time *t.* A simple GARCH(p,q) model is assumed. Then we can write

$$\begin{aligned} Y_t &= \eta_t \varepsilon_t \\ \eta_t^2 &= \delta_0 + \sum_{i=1}^{q} \delta_i (\eta_{t-i}\varepsilon_{t-i})^2 + \sum_{j=q+1}^{q+p} \delta_j \eta_{t+q-j}^2 \end{aligned} \tag{1}$$

where $\{\varepsilon_t\}\sim$i.i.d. $N(0,1)$. $\{\varepsilon_t\}$ is the innovation (or shock) of the US market and is hypothesized to be Gaussian. In order to ensure the positivity of η_t^2,δ_0 has to be positive and the remaining parameters in (1) non-negative. After denoting $Z^{t-1} = [z_{t-1}, z_{t-2}, \ldots]$, it turns out that

$$Y_t | Y^{t-1} \sim N(0, \eta_t^2)$$

Let x_t be the $p \times 1$ return vector of the European markets at time *t*. We assume that returns at time t depend both on an endogenous (local) shock and an exogenous (global) shock. The endogenous shocks do have relationships with each other (common movements are usually observed in neighboring markets). The $p \times 1$ local shock vector is denoted ζ_t. The exogenous or global shock is an event that has an influence across more markets. For the European markets we identify the global shock as the innovation of the US market one-day before, that is ε_{t-1}. The lag is due to the mentioned nonsynchronicity of the markets.

The function $\mathrm{f}(\varepsilon_{t-1})$, specified below, describes the relationship between the return vector x_t and ε_{t-1}.

The local and the global shocks are assumed to be independent and to have a joint $(p+1)-$ dimensional normal distribution,

$$\begin{pmatrix} \varepsilon_{t-1} \\ \zeta_t \end{pmatrix} \sim N_{p+1}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0^T \\ 0 & \Psi \end{pmatrix}\right) \tag{2}$$

where ψ is a correlation matrix with generic off-diagonal entry ρ_{ij}. If the hypothesis of Gaussianity is far from true for returns, it appears to be consistent for shocks, even if some authors propose more general distributions (e.g. Bollerslev, 1987). Moreover, we assume that the variances have a multivariate GARCH structure. We can then write

$$\mathrm{x}_t = D_t[\mathrm{f}(\varepsilon_{t-1}) + \zeta_t] \tag{3}$$

$$\mathrm{f}(\varepsilon_{t-1}) = \sqrt{\frac{2}{\pi}}\gamma + \beta\varepsilon_{t-1} - \gamma|\varepsilon_{t-1}| \tag{4}$$

where

$$D_t = \begin{pmatrix} \sigma_{1t} & 0 & & 0 \\ 0 & \sigma_{2t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & & 0 & \sigma_{pt} \end{pmatrix} \tag{5}$$

$$\sigma_{kt}^2 = \omega_{0k} + \sum_{i=1}^{q} \omega_{ik}\left(\sigma_{k,t-i}\zeta_{k,t-i}\right)^2 + \sum_{j=q+1}^{q+p} \omega_{jk}\sigma_{k,t+q-j}^2 + \omega_{q+p+1,k}\eta_{t-1}^2 \tag{6}$$

The last term takes into account the possible volatility spillover from the US to the European markets. If at time $t-1$ the US stock exchange is closed, then $\eta_{t-1}^2 = \eta_{t-2}^2$ The positivity of σ_{kt}^2 is ensured by the usual constraints on the parameters.

Assumptions (2–6) compound the SGARCH model. The function $\mathrm{f}(\varepsilon_{t-1})$ models the effect of the exogenous shock ε_{t-1} on the vector x_t and $\{\zeta_t\}$ is a sequence of serially independent random vectors. The parameter vectors β and γ are constrained to be non-negative. Moreover, $\beta \geq \gamma \geq 0$. The former describes the direct effect of the past US innovations on x_t, the latter the feedback effect. Volatility feedback theory (Campbell & Hentschel, 1992) implies that news increases volatility, which in turn lowers returns. Hence the direct effect of good (bad) news is mitigated (strengthened) by the feedback effect. A point in favor of the SGARCH model is the formalization of the two effects.

The conditional distribution of the return vector is

$$\mathrm{x}_t|I_{t-1} \sim N_p(D_t\mathrm{f}(\varepsilon_{t-1}), D_t\Psi D_t)$$

where I_{t-1} denotes the information at time $t-1$.

The SGARCH model does not involve a conditional null mean vector. Instead, the mean does depend on the volatility. Moreover, returns are more reactive to bad news than good news. In fact,

$$\left.\frac{\partial \mathrm{x}_t}{\partial \varepsilon_{t-1}}\right|_{\varepsilon_{t-1}>0} = D_t(\beta - \gamma)$$
$$\left.\frac{\partial \mathrm{x}_t}{\partial \varepsilon_{t-1}}\right|_{\varepsilon_{t-1}<0} = D_t(\beta + \gamma)$$

4. CONDITIONAL DISTRIBUTIONS AND RELATED MOMENTS

We are interested in the p-variate distribution of x_t conditional on D_t and on news from the US market, that is on the sign of Y_{t-1}.

Theorem 1. Under the SGARCH model's assumptions, the following distributions hold:

Good news

$$\left(D_t^{-1}\mathrm{x}_t | Y_{t-1} > 0\right) \sim SN_p\left(\gamma\sqrt{\frac{2}{\pi}}, \Omega_+, \alpha_+\right)$$

where

$$\Omega_+ = \Psi + \delta_+\delta_+^{\mathrm{T}}, \quad \alpha_+ = \frac{\Psi^{-1}\delta_+}{\sqrt{1+\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}} \quad \text{and} \quad \delta_+ = \beta - \gamma$$

Bad news

$$\left(D_t^{-1}\mathrm{x}_t | Y_{t-1} < 0\right) \sim SN_p\left(\gamma\sqrt{\frac{2}{\pi}}, \Omega_-, \alpha_-\right)$$

where

$$\Omega_- = \Psi + \delta_-\delta_-^{\mathrm{T}}, \quad \alpha_- = \frac{-\Psi^{-1}\delta_-}{\sqrt{1+\delta_-^{\mathrm{T}}\Psi^{-1}\delta_-}} \quad \text{and} \quad \delta_- = \beta + \gamma$$

The proof is given in the appendix.

Applying the results of linear transformations of a multivariate *SN* distribution, we can obtain the conditional distributions of returns $(\mathrm{x}_t|D_t, Y_{t-1} > 0)$ and $(\mathrm{x}_t|D_t, Y_{t-1} < 0)$. Their formal properties are found to be consistent with empirical findings. In more detail:

Expectation: On average, good (bad) news determines positive (negative) returns. The effect of both kinds of news is equal in absolute value. Moreover, it is proportional to the direct effect (modeled by the parameter β) mitigated or amplified by the volatility structure (modeled by the diagonal matrix D_t),

$$E(\mathrm{x}_t|D_t, Y_{t-1} > 0) = \sqrt{\frac{2}{\pi}} D_t\beta$$

$$E(\mathrm{x}_t|D_t, Y_{t-1} < 0) = -\sqrt{\frac{2}{\pi}} D_t\beta$$

This result could be interpreted as the presence of an arbitrage opportunity in the market with implications regarding market efficiency. Actually, our analysis is carried out using returns computed from two close prices (close-to-close). During the time from the closing to the opening of a European exchange, there is the closing of the US exchange and its effect is reflected mainly in the open prices of the European exchanges. If we used open-to-close returns, this apparent arbitrage opportunity would disappear.

Variance: Variance is higher (lower) in the presence of bad (good) news. More precisely, the elements on the main diagonal of the covariance matrix are greater (smaller) when previous day's US market returns were negative (positive)

$$V(x_t|D_t, Y_{t-1} > 0) = D_t\left(\Psi + \frac{\pi-2}{\pi}\delta_+\delta_+^{\mathrm{T}}\right)D_t$$

$$V(\mathrm{x}_t|D_t, Y_{t-1} < 0) = D_t\left(\Psi + \frac{\pi-2}{\pi}\delta_-\delta_-^{\mathrm{T}}\right)D_t$$

Skewness: Symmetry of conditional returns would imply that news from the US are irrelevant. In this framework, univariate skewness is negative (positive) in presence of bad (good) news and higher (smaller) in absolute value. On the other hand, multivariate skewness is always positive but its level cannot be related to the kind of news. Skewness of x_t when $Y_{t-1}>0$ can be either lower or higher than skewness of x_t when $Y_{t-1}<0$, depending on the parameters. The two indices are

$$S(\mathrm{x}_t|D_t, Y_{t-1}>0) = 2(4-\pi)^2\left(\frac{\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}{\pi+(\pi-2)\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}\right)^3$$

$$S(\mathrm{x}_t|D_t, Y_{t-1}<0) = 2(4-\pi)^2\left(\frac{\delta_-^{\mathrm{T}}\Psi^{-1}\delta_-}{\pi+(\pi-2)\delta_-^{\mathrm{T}}\Psi^{-1}\delta_-}\right)^3$$

Kurtosis: In the SGARCH model, relevant news (good or bad) always lead to leptokurtotic returns. Again, there is no relationship between the kind of news (good or bad) and multivariate kurtosis (high or low). However, SGARCH models imply that a higher kurtosis is related to a higher skewness. The two indices are

$$K(\mathrm{x}_t|D_t, Y_{t-1}>0) = 8(\pi-3)\left(\frac{\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}{\pi+(\pi-2)\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}\right)^2$$

$$K(\mathrm{x}_t|D_t, Y_{t-1}<0) = 8(\pi-3)\left(\frac{\delta_-^{\mathrm{T}}\Psi^{-1}\delta_-}{\pi+(\pi-2)\delta_-^{\mathrm{T}}\Psi^{-1}\delta_-}\right)^2$$

Correlation: Let ρ_+ (ρ_-) be the $p\times 1$ vector whose ith component ρ_{i+} (ρ_{i-}) is the correlation coefficient between the ith European return X_{it} and the previous day's US return Y_{t-1} conditionally on good (bad) news, that is $Y_{t-1}>0$, ($Y_{t-1}<0$), and volatility D_t. The correlation of $X_{it}, Y_{t-1}|D_t, Y_{t-1}>0$ ($X_{it}, Y_t|D_t, Y_{t-1}<0$) is the same under the multivariate and univariate SGARCH model (the former being a multivariate generalization of the latter). Hence, we can recall a property of the univariate SGARCH model (De Luca & Loperfido, 2004) and write

$$\rho_{i+} = \frac{(\beta_i-\gamma_i)\sqrt{\pi-2}}{\sqrt{\pi+(\pi-2)(\beta_i-\gamma_i)^2}}, \quad \rho_{i-} = \frac{(\beta_i+\gamma_i)\sqrt{\pi-2}}{\sqrt{\pi+(\pi-2)(\beta_i+\gamma_i)^2}}$$

for $i=1,\ldots,p$. Implications of the above results are twofold. In the first place, ρ_+ and $\rho-$ are functions of $\delta_+=\beta-\gamma$ and $\delta_-=\beta+\gamma$, respectively. In the second place, a little algebra leads to the inequalities $0\leq\rho_+\leq\rho_-\leq 1_p$, where 1_p is a p-dimensional vector of ones. It follows that bad news strengthen the association between European returns and previous day's US returns.

The two p-variate distributions in Theorem 1 have different characteristics according to the news coming from the US market. As a result, the conditioning appears to be relevant.

It is interesting to consider some special cases. Firstly, if the parameter vector γ is zero, then $\delta_+ = \delta_- = \beta$ and $\Omega_+ = \Omega_- = \Omega$. The two conditional distributions in Theorem 1 differ for the shape parameter which has the same absolute value but a different sign, that is

$$\left(D_t^{-1}\mathrm{x}_t | Y_{t-1} > 0\right) \sim SN_p(0, \Omega, \alpha)$$

where

$$\Omega = \Psi + \beta\beta^{\mathrm{T}}, \quad \alpha = \frac{\Psi^{-1}\beta}{\sqrt{1 + \beta^{\mathrm{T}}\Psi^{-1}\beta}}$$

and

$$\left(D_t^{-1}\mathrm{x}_t | Y_{t-1} < 0\right) \sim SN_p(0, \Omega, -\alpha).$$

In this case, no feedback effect exists. However, it still makes sense to condition on the type of news and to introduce *SN* distributions.

Secondly, if the parameter vector β is zero (implying that also γ is zero), there is no evidence of any (direct or feedback) effect of the US news on European returns. The two markets are independent. The *SN* distributions in Theorem 1 turn out to have a zero shape parameter which shrinks them to the same normal distribution. As a result, the multivariate skewness and kurtosis indices also shrink to zero.

Finally, if the dimensionality parameter p equals one, that is if we move from a multivariate framework to a univariate perspective, the multivariate SGARCH model equals the univariate SGARCH model (De Luca & Loperfido, 2004). In this case, denoting by X_t a European return at time t, we have

$$\frac{X_t}{\sigma_t} = \sqrt{\frac{2}{\pi}}\gamma + \beta\varepsilon_{t-1} - \gamma|\varepsilon_{t-1}| + \zeta_t$$

where β and γ are now scalars and ζ_t is a unidimensional random variable. The main features of the model in a univariate context (parameters of the distribution of X_t/σ_t given the sign of Y_{t-1} and moments of X_t given σ_t and the sign of Y_{t-1}) are summarized in Table 2.

Table 2. Features of the Univariate SGARCH Model.

	Good News ($Y_{t-1} > 0$)	Bad News ($Y_{t-1} < 0$)	Overall
Location	$\sqrt{\frac{2}{\pi}}\gamma$	$\sqrt{\frac{2}{\pi}}\gamma$	—
Scale	$\sqrt{1+(\beta-\gamma)^2}$	$\sqrt{1+(\beta+\gamma)^2}$	—
Shape	$\frac{\beta-\gamma}{\sqrt{1+(\beta-\gamma)^2}}$	$\frac{-(\beta+\gamma)}{\sqrt{1+(\beta+\gamma)^2}}$	—
Expectation	$\sqrt{\frac{2}{\pi}}\beta\sigma_t$	$-\sqrt{\frac{2}{\pi}}\beta\sigma_t$	0
Variance	$\sigma_t^2\left[1+\frac{\pi-2}{\pi}(\beta-\gamma)^2\right]$	$\sigma_t^2\left[1+\frac{\pi-2}{\pi}(\beta+\gamma)^2\right]$	$\sigma_t^2\left[1+\frac{\pi-2}{\pi}\gamma^2+\beta^2\right]$
Skewness	$\frac{4-\pi}{2}\left[\frac{2(\beta-\gamma)^2}{\pi+(\pi-2)(\beta-\gamma)^2}\right]^{1.5}$	$\frac{\pi-4}{2}\left[\frac{2(\beta+\gamma)^2}{\pi+(\pi-2)(\beta+\gamma)^2}\right]^{1.5}$	$\frac{\sqrt{2}\gamma\left(2\pi\gamma^2-3\pi\beta^2-4\gamma^2\right)}{\left[\pi+(\pi-2)\gamma^2+\pi\beta^2\right]^{1.5}}$
Kurtosis	$2(\pi-3)\left[\frac{2(\beta-\gamma)^2}{\pi+(\pi-2)(\beta-\gamma)^2}\right]^2$	$2(\pi-3)\left[\frac{2(\beta+\gamma)^2}{\pi+(\pi-2)(\beta+\gamma)^2}\right]^2$	$3+\frac{8\gamma^2(\pi-3)}{\left[\pi+(\pi-2)\gamma^2+\pi\beta^2\right]^2}$
Correlation with Y_{t-1}	$\frac{(\beta-\gamma)\sqrt{\pi-2}}{\sqrt{\pi+(\pi-2)(\beta-\gamma)^2}}$	$\frac{(\beta+\gamma)\sqrt{\pi-2}}{\sqrt{\pi+(\pi-2)(\beta+\gamma)^2}}$	$\frac{\beta\sqrt{\pi}}{\sqrt{\pi+(\pi-2)\gamma^2+\pi\beta^2}}$

5. UNCONDITIONAL DISTRIBUTIONS

We refer to the expression "unconditional distribution" to indicate the distribution of the vector x_t unconditionally on Y_{t-1}. However, it is still conditional on its own past history.

Expectation: The expected value of the returns is the null vector, consistently with empirical findings and economic theory,

$$E(\mathrm{x}_t|D_t) = 0$$

Variance: The variance of x_t can be seen as decomposed into the sum of two components, an endogenous component, determined by the market internal structure, and an exogenous component, determined by news from the US market. The latter can be further represented as the sum of a component depending on the direct effect and another component depending on the feedback effect:

$$V(\mathrm{x}_t|D_t) = D_t\left(\Psi + \beta\beta^{\mathrm{T}} + \frac{\pi - 2}{\pi}\gamma\gamma^{\mathrm{T}}\right)D_t$$

Skewness: In the multivariate SGARCH model, the feedback effect determines the asymmetric behavior of returns. More formally, multivariate skewness, as measured by Mardia's index, is zero if and only if the feedback parameter is the zero vector:

$$\gamma = 0 \Rightarrow S(\mathrm{x}_t|D_t) = 0, \quad \gamma \neq 0 \Rightarrow S(\mathrm{x}_t|D_t) > 0$$

The proof of this result is in the appendix.

Kurtosis: In the univariate SGARCH model ($p = 1$), we can show the presence of kurtosis. The same result holds for a linear combination of returns following the model. In the multivariate case we conjecture the existence of multivariate kurtosis as measured by Mardia's index.

Correlation: Let ρ be the $p \times 1$ vector whose ith component ρ_i is the correlation coefficient between the ith European return X_{it} and the previous day's US return Y_{t-1}, conditionally on volatility D_t. It easily follows from the definition of x_t as a function of the US shocks and ordinary properties of covariance that

$$\rho_i = \frac{\beta_i\sqrt{\pi}}{\sqrt{\pi + (\pi - 2)\gamma_i^2 + \pi\beta_i^2}}, \quad i = 1, \ldots, p$$

as reported in the last row of Table 2. This result implies that $0 \leq \rho \leq 1_p$, where 1_p is a p-dimensional vector of ones. It also implies that ρ_i is an increasing function of β_i and a decreasing function of γ_i. It follows that association between European returns and previous day's US returns is directly related to the vector of direct effects β and inversely related to the vector of feedback effects γ.

The same statistics are reported for the univariate case in the last column of Table 2. The complete unconditional distribution could be obtained by simulation.

6. ANALYSIS OF SOME FINANCIAL MARKETS

We focus on a univariate and multivariate analysis of three European financial markets: the Dutch, the Swiss and the Italian market. They are small capitalized markets compared to the US market. The differences in sizes between the US market and the three European exchanges are evident. The weight of the capitalization of the US market on the world capitalization is over 40%, while the weights of each of the latter does not exceed 3%.

The close-to-close log-returns of representative market indexes have been observed in the period 18/01/1995–02/05/2003. The three markets (Dutch, Swiss and Italian) are represented by the AEX, SMI and MIB indexes, respectively. The returns of the US market are represented by the Standard & Poor 500 (S&P), the most popular market index for the New York Stock Exchange.

6.1. Univariate Analysis

The analysis proceeds by looking at the most important features of the returns in absence and in presence of the conditioning on the performance of the one-day lagged US market.

Table 3 reports the most salient statistics for the indexes when no conditioning is taken into account. They describe the typical features of financial returns. The average returns are very close to zero and a certain degree of negative skewness is apparent. Fat tails in the distribution are revealed by the kurtosis indices. Finally, the correlations with the one-day lagged S&P returns are reported. A dynamic linkage from the US market to the European markets does exist.

Then, we divide the entire samples into two subsamples, according to the sign of the one-day lagged S&P return. In order to take into account the

Table 3. Descriptive Statistics for the Three Stock Indexes.

	Average	Standard deviation	Skewness	Kurtosis	Correlation
AEX	0.019	1.575	−0.093	6.549	0.322
SMI	0.027	1.328	−0.134	6.689	0.276
MIB	0.022	1.564	−0.064	4.804	0.210

Table 4. Descriptive Statistics for the Three Stock Indexes According to the Sign of One-Day Lagged S&P Returns.

	Average	Standard deviation	Skewness	Kurtosis	Correlation
AEX					
S&P > 0	0.384	1.498	0.346	8.020	0.178
S&P < 0	−0.378	1.550	−0.412	5.071	0.275
SMI					
S&P > 0	0.252	1.241	0.192	7.517	0.157
S&P < 0	−0.219	1.366	−0.297	6.214	0.286
MIB					
S&P > 0	0.259	1.502	0.028	5.098	0.084
S&P < 0	−0.226	1.584	−0.038	4.494	0.210

differences in closing days of the stock exchanges, some hypotheses have to be made. We assume that if the European exchange is open at time t and the US exchange is closed at time $t-1$, then ε_{t-1} is set to zero (there is no information from the US market). If the European exchange is closed at time t, the US exchange information at time $t-1$ is useless; the next European exchange return (at time $t+1$) is related to ε_t. We compute the same statistics as above in this setting, summarized in Table 4.

The resulting statistics are very interesting. They show a different behavior of the European market indexes according to the sign of the last trading day in the American stock exchange.

In the three markets the average return is positive (negative) when the one-day lagged return of the S&P is positive (negative). The standard deviation of the returns of the European market indexes is always greater in presence of a negative sign coming from the US market. The skewness coefficient is negative and stronger when the American stock exchange return is negative; it is positive in the opposite case. Finally, the relationship between present European returns and past US returns is clearly stronger

Table 5. Maximum Likelihood Estimates (Standard Error) of the Univariate SGARCH Models for the Three Indexes' Returns.

	$\hat{\omega}_0$	$\hat{\omega}_1$	$\hat{\omega}_2$	$\hat{\omega}_3$	$\hat{\beta}$	$\hat{\gamma}$
AEX	0.012	0.115	0.866	0.020	0.343	0.072
	(0.006)	(0.016)	(0.019)	(0.011)	(0.022)	(0.033)
SMI	0.035	0.133	0.840	0.006	0.267	0.121
	(0.011)	(0.019)	(0.021)	(0.009)	(0.024)	(0.034)
MIB	0.083	0.115	0.836	0.020	0.186	0.113
	(0.022)	(0.018)	(0.025)	(0.016)	(0.021)	(0.033)

when the S&P return is negative. All these empirical findings match the theoretical features of the univariate SGARCH model (De Luca & Loperfido, 2004). Only the observed kurtosis is always smaller in presence of a negative US return, which contradicts the model.

The autocorrelation functions of the squared European returns steadily decrease, hinting that $p = q = 1$ in the variance equation. Table 5 contains the maximum likelihood estimates of the parameters of the three univariate models. According to the ratios between estimate and standard error, all the parameters are significant but the volatility spillover coefficients. This involves the advantage of a model including the parameters β and γ. The highest value of the quantity $\beta - \gamma$ refers to the AEX returns involving the major distance between the behaviors of the index conditionally on the signals coming from the US market.

On the whole, the economic interpretation is straightforward: it is relevant to distinguish between bad and good news from the US market. The inclusion of the effect of exogenous news can significantly improve the predictive performance.

In order to check the stability of the coefficients (particularly β and γ) we carried out recursive estimates. The first sample is composed of the observations from 1 to 1000. For each subsequent sample we added an observation. Fig. 1 shows the dynamics of the two parameters. The differences between β and γ are approximately constant for AEX and MIB. For SMI, the distance tends to be slightly more variable.

In order to evaluate the performance of the model in out-of-sample forecasting, we computed one-step-ahead forecasts of the volatility, $\sigma^2_{t|t-1}$, for $t = 1601, \ldots,$ using the SGARCH model. We compared them with benchmark forecasts obtained from a standard GARCH(1,1) model. Following Pagan and Schwert (1990), we ran a regression of log squared returns versus log forecasted volatility, and then computed the F-test for the

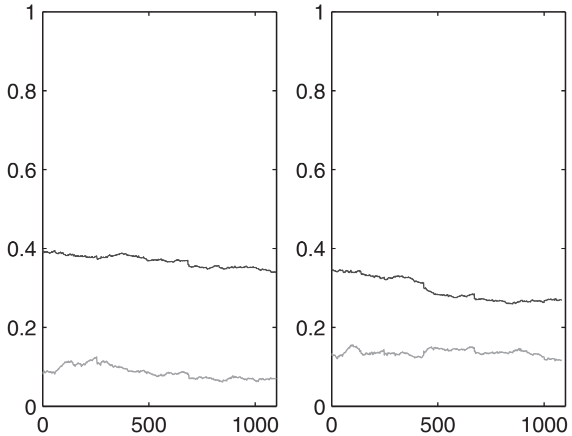

Fig. 1. Recursive Estimates of β (Top Curve) and γ (Bottom Curve) for AEX (Left), SMI (Middle) and MIB (Right).

hypothesis of a null intercept and a unit slope. The *F*-statistics for SGARCH and GARCH, respectively, are 53.14 and 69.24 (AEX), 59.25 and 69.24 (SMI), 70.00 and 72.65 (MIB) and thus favor the SGARCH model. Moreover, we computed how many times the sign of the return had been correctly predicted with the SGARCH model. The percentages of correct negative signs are 64.37 (AEX), 58.98 (SMI) and 56.15 (MIB). For positive sign, they are 56.78 (AEX), 51.33 (SMI) and 49.14 (MIB).

6.2. Multivariate Analysis

When a multivariate model is considered, the focus on daily data poses some problems. In fact there are unavoidably some missing values, due to the different holidays of each country. As an example, on the 1st of May, there is the Labor Holiday in Italy and Switzerland and the stock exchanges do not operate. But on the same day the stock exchange in the Netherlands is open. Deleting the day implies missing a datum. In order to overcome this drawback, we assumed that the Italian and Swiss variance in that day, σ^2_{1stMay}, has the same value as on the last opening day of the stock exchange ($\sigma^2_{30thApril}$ if not Saturday or Sunday). The next day variance, σ^2_{2ndMay} if not Saturday or Sunday, was computed according to a GARCH (1,1) model without considering the past holiday. In general, for the *k*-th market index,

$$\sigma_{kt}^2 = \begin{cases} \omega_{0k} + \omega_{1k}\left(\sigma_{k,t-1}\zeta_{k,t-1}\right)^2 + \omega_{2k}\sigma_{k,t-1}^2 + \omega_{3k}\eta_{t-1}^2, & t-1 = \text{open} \\ \omega_{0k} + \omega_{1k}\left(\sigma_{k,t-2}\zeta_{k,t-2}\right)^2 + \omega_{2k}\sigma_{k,t-1}^2 + \omega_{3k}\eta_{t-1}^2, & t-1 = \text{close} \end{cases}$$

Maximum likelihood estimation has been performed using a Gauss code written by the authors. The algorithm converged after a few iterations. The parameters are again all significant, but the spillover parameters. Their values, reported in Table 6, are not very far from the corresponding estimates in the univariate context. In addition we obtain the estimates of the correlation coefficients, indicated by ρ_{ij}.

The diagnostic of the SGARCH model can be based on the squared norms $\{S_t\}$ of the residuals $\{\mathrm{r}_t\}$, defined as follows:

Table 6. Maximum Likelihood Estimates and Standard Errors of the Multivariate SGARCH Model. The Subscript Letters have the following Meanings: A = AEX, S = SMI, M = MIB.

Parameter	Estimate	Standard Error
ω_{0A}	0.042	0.008
ω_{1A}	0.085	0.010
ω_{2A}	0.865	0.019
ω_{3A}	0.021	0.010
ω_{0S}	0.075	0.012
ω_{1S}	0.100	0.013
ω_{2S}	0.826	0.021
ω_{3S}	0.008	0.006
ω_{0M}	0.148	0.027
ω_{1M}	0.087	0.013
ω_{2M}	0.852	0.023
ω_{3M}	−0.017	0.011
ρ_{AM}	0.665	0.012
ρ_{AS}	0.705	0.011
ρ_{MS}	0.605	0.014
β_A	0.335	0.022
γ_A	0.106	0.035
β_S	0.235	0.021
γ_S	0.148	0.032
β_M	0.192	0.029
γ_M	0.168	0.041

$$S_t = \mathrm{r}_t^{\mathrm{T}} \mathrm{r}_t, \ \mathrm{r}_t = \begin{cases} \hat{\Omega}_+^{-1/2}\left(\hat{D}_t^{-1}\mathrm{x}_t - \hat{\gamma}\sqrt{\frac{2}{\pi}}\right), & Y_{t-1} > 0 \\ \hat{\Omega}_-^{-1/2}\left(\hat{D}_t^{-1}\mathrm{x}_t - \hat{\gamma}\sqrt{\frac{2}{\pi}}\right), & Y_{t-1} < 0 \end{cases}$$

If the model is correctly specified, the following results hold:

1. The squared norms $\{S_t\}$ are i.i.d. and $S_t \sim \chi_p^2$.
2. If n is the number of observed returns, then

$$\sqrt{\frac{n}{2p}}\left(\frac{1}{n}\sum_{t=1}^{n} S_t - p\right) \overset{a}{\sim} N(0,1) \tag{7}$$

The proof of this result is given in the appendix. We computed the statistic in (7). Its value is -0.031, such that the hypothesis of skew-normality of the conditional distributions cannot be rejected.

7. CONCLUSIONS

The multivariate Skew-GARCH model is a generalization of the GARCH model aimed at describing the behavior of a vector of dependent financial returns when an exogenous shock coming from a leading financial market is taken into account. We analyzed returns from three European markets, while the leading market was identified as the US market. It turned out to be significant to consider the effects of the exogenous shock. The distributions of the European returns show different features according to the type of news arriving from the leading market. When the above assumptions are not consistent, the estimation step reveals the drawback. In this case, some parameters of the model are zero and the multivariate Skew-GARCH model shrinks to the simple multivariate GARCH model with constant correlation coefficients. A future extension of our proposed model would be to replace the multivariate skew-normal distribution by a multivariate skew-elliptical distribution, see the book edited by Genton (2004). For example, a multivariate skew-t distribution would add further flexibility by introducing an explicit parameter controlling tail behavior.

REFERENCES

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, *12*, 171–178.

Azzalini, A., & Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, *83*, 715–726.

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

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

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH approach. *The Review of Economics and Statistics*, *72*, 498–505.

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

Campbell, J. Y., & Hentschel, L. (1992). No news is a good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics*, *31*, 281–331.

Dalla Valle, A. (2004). The skew-normal distribution. In: M. G. Genton (Ed.), *Skew-elliptical distributions and their applications: A journey beyond normality* (pp. 3–24). Boca Raton, FL: Chapman & Hall/CRC.

De Luca, G., & Loperfido, N. (2004). A Skew-in-Mean GARCH model. In: M. G. Genton (Ed.), *Skew-elliptical distributions and their applications: A journey beyond normality* (pp. 205–222). Boca Raton, FL: Chapman & Hall/CRC.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica*, *50*, 987–1008.

Engle, R. F. (2002). Dynamic conditional correlation. A simple class of multivariate GARCH models. *Journal of Business and Economic Statistics*, *20*, 339–350.

Eun, C. S., & Shim, S. (1989). International transmission of stock markets movements. *Journal of Financial and Quantitative Analysis*, *24*, 241–256.

Genton, M. G. (Ed.) (2004). *Skew-elliptical distributions and their applications: A journey beyond normality*. Boca Raton: Chapman & Hall/CRC.

Genton, M. G., & Loperfido, N. (2005). Generalized skew-elliptical distributions and their quadratic forms. *Annals of the Institute of Statistical Mathematics*, *57*, 389–401.

Karolyi, G. A. (1995). A multivariate GARCH model of international transmissions of stock returns and volatility: The case of United States and Canada. *Journal of Business and Economic Statistics*, *13*, 11–25.

Koch, P. D., & Koch, T. W. (1991). Evolution in dynamic linkages across daily national stock indexes. *Journal of International Money and Finance*, *10*, 231–251.

Lin, W., Engle, R., & Ito, T. (1994). Do bulls and bears move across borders? International transmission of stock return and volatility. *The Review of Financial Studies*, *7*, 507–538.

Loperfido, N. (2004). Generalized skew-normal distributions. In: M. G. Genton (Ed.), *Skew-elliptical distributions and their applications: A journey beyond normality* (pp. 65–80). Boca Raton: Chapman & Hall/CRC, FL.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*, 519–530.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns. *Econometrica*, *59*, 347–370.

Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics*, *45*, 267–290.

APPENDIX

Proof of Theorem 1. We shall prove the theorem for $Y_{t-1} > 0$ only. The proof for $Y_{t-1} < 0$ is similar. By assumption, the joint distribution of random shocks is

$$\begin{pmatrix} \varepsilon_{t-1} \\ \zeta_t \end{pmatrix} \sim N_{p+1}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0^{\mathrm{T}} \\ 0 & \Psi \end{pmatrix}\right)$$

By definition, $\Omega_+ = \Psi + \delta_+\delta_+^{\mathrm{T}}$, so that

$$\begin{pmatrix} \varepsilon_{t-1} \\ \gamma\sqrt{2/\pi} + \delta_+\varepsilon_{t-1} + \zeta_t \end{pmatrix} \sim N_{p+1}\left(\begin{pmatrix} 0 \\ \gamma\sqrt{2/\pi} \end{pmatrix}, \begin{pmatrix} 1 & \delta_+^{\mathrm{T}} \\ \delta_+ & \Omega_+ \end{pmatrix}\right)$$

The conditional distribution of $\gamma\sqrt{2/\pi} + \delta_+\varepsilon_{t-1} + \zeta_t$ given $\varepsilon_{t-1} > 0$ is multivariate *SN* (Dalla Valle, 2004):

$$\left(\gamma\sqrt{\frac{2}{\pi}} + \delta_+\varepsilon_{t-1} + \zeta_t|\varepsilon_{t-1} > 0\right) \sim SN_p\left(\gamma\sqrt{\frac{2}{\pi}}, \Omega_+, \frac{\Omega_+^{-1}\delta_+}{\sqrt{1 - \delta_+^{\mathrm{T}}\Omega_t^{-1}\delta_+}}\right)$$

By definition, $\delta_+ = \beta - \gamma$, so that standard properties of absolute values lead to

$$\left(\gamma\sqrt{\frac{2}{\pi}} + \beta\varepsilon_{t-1} - \gamma|\varepsilon_{t-1}| + \zeta_t\big|\varepsilon_{t-1} > 0\right) \sim SN_p\left(\gamma\sqrt{\frac{2}{\pi}}, \Omega_+, \frac{\Omega_+^{-1}\delta_+}{\sqrt{1 - \delta_+^{\mathrm{T}}\Omega_+^{-1}\delta_+}}\right)$$

By definition, US returns and European returns are $Y_{t-1} = \eta_{t-1}\varepsilon_{t-1}$ and $\mathrm{x}_t = D_t(\gamma\sqrt{2/\pi} + \beta\varepsilon_{t-1} - \gamma|\varepsilon_{t-1}| + \zeta_t)$. Hence,

$$\left(D_t^{-1}\mathrm{x}_t|Y_{t-1} > 0\right) \sim SN_p\left(\gamma\sqrt{\frac{2}{\pi}}, \Omega_+, \frac{\Omega_+^{-1}\delta_+}{\sqrt{1 - \delta_+^{\mathrm{T}}\Omega_+^{-1}\delta_+}}\right)$$

Apply now the Sherman-Morrison formula to the matrix $\Omega_+ = \Psi + \delta_+\delta_+^{\mathrm{T}}$:

$$\Omega_+^{-1} = \Psi^{-1} - \frac{\Psi^{-1}\delta_+\delta_+^{\mathrm{T}}\Psi^{-1}}{1 + \delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}$$

A little algebra leads to the following equations:

$$\Omega_+^{-1}\delta_+ = \left(\Psi^{-1} - \frac{\Psi^{-1}\delta_+\delta_+^{\mathrm{T}}\Psi^{-1}}{1+\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}\right)\delta_+$$

$$= \Psi^{-1}\delta_+\left(1 - \frac{\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}{1+\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}\right)$$

$$= \frac{\Psi^{-1}\delta_+}{1+\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}$$

$$1 - \delta_+^{\mathrm{T}}\Omega_+^{-1}\delta_+ = 1 - \frac{\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}{1+\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+} = \frac{1}{1+\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}$$

Recall now the definition of the vector α_+:

$$\frac{\Omega_+^{-1}\delta_+}{\sqrt{1-\delta_+^{\mathrm{T}}\Omega_+^{-1}\delta_+}} = \frac{\Psi^{-1}\delta_+/\left(1+\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+\right)}{1/\sqrt{1+\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}}$$

$$= \frac{\Psi^{-1}\delta_+}{\sqrt{1+\delta_+^{\mathrm{T}}\Psi^{-1}\delta_+}} = \alpha_+$$

We can then write

$$\left(D_t^{-1}\mathrm{x}_t | Y_{t-1} > 0\right) \sim SN_p\left(\gamma\sqrt{\frac{2}{\pi}}, \Omega_+, \alpha_+\right)$$

and the proof is complete. □

Proof of the positivity of the unconditional skewness. We shall prove the theorem for $\mathrm{z}_t = D_t^{-1}\mathrm{x}_t$ only: the proof for x_t is similar. When $\gamma = 0$ the distribution of z_t is multivariate normal, so that it suffices to show that $\gamma \neq 0$ implies that $S(\mathrm{z}_t) > 0$. Let the vectors w_t, u_t, g_1, g_2 and the matrix Γ be defined as follows:

$$\begin{array}{ll} V(\mathrm{z}_t) = \Sigma, & \Gamma = \Sigma^{-1} \otimes \Sigma^{-1} \otimes \Sigma^{-1} \\ \mathrm{w}_t \sim (\mathrm{z}_t | Y_{t-1} > 0), & \mathrm{g}_1 = E(\mathrm{w}_t \otimes \mathrm{w}_t \otimes \mathrm{w}_t) \\ \mathrm{u}_t \sim (\mathrm{z}_t | Y_{t-1} < 0), & \mathrm{g}_2 = E(\mathrm{u}_t \otimes \mathrm{u}_t \otimes \mathrm{u}_t) \end{array}$$

The distribution of z_t is the mixture, with equal weights, of the distributions of w_t and u_t. Ordinary properties of mixtures lead to

$$E(\mathrm{z}_t \otimes \mathrm{z}_t \otimes \mathrm{z}_t) = \frac{1}{2}E(\mathrm{w}_t \otimes \mathrm{w}_t \otimes \mathrm{w}_t) + \frac{1}{2}E(\mathrm{u}_t \otimes \mathrm{u}_t \otimes \mathrm{u}_t)$$
$$= \frac{1}{2}\left(\mathrm{g}_1 + \mathrm{g}_2\right)$$

Hence the multivariate skewness of the vector z_t can be represented as follows:

$$S(z_t) = \frac{1}{4}\left(\mathrm{g}_1 + \mathrm{g}_2\right)^{\mathrm{T}}\Gamma\left(\mathrm{g}_1 + \mathrm{g}_2\right)$$

The distribution of w_t equals that of $-\mathrm{u}_t$ only when $\gamma = 0$. By assumption $\gamma \neq 0$, so that

$$\mathrm{g}_1 + \mathrm{g}_2 \neq 0 \Rightarrow S(z_t) = \frac{1}{4}\left(\mathrm{g}_1 + \mathrm{g}_2\right)^{\mathrm{T}}\Gamma\left(\mathrm{g}_1 + \mathrm{g}_2\right) > 0$$

The last inequality follows from Σ being positive definite and from properties of the Kronecker product. The proof is then complete. □

Proof of the asymptotic distribution (7). Let the random variables $\{\bar{\mathrm{S}}_t\}$ and the random vectors $\{\bar{\mathrm{r}}_t\}$ be defined as follows:

$$\bar{S}_t = \bar{\mathrm{r}}_t^{\mathrm{T}}\bar{\mathrm{r}}_t, \quad \bar{\mathrm{r}}_t = \begin{cases} \Omega_+^{-1/2}\left(D_t^{-1}\mathrm{x}_t - \gamma\sqrt{\dfrac{2}{\pi}}\right) & Y_{t-1} > 0 \\ \Omega_-^{-1/2}\left(D_t^{-1}\mathrm{x}_t - \gamma\sqrt{\dfrac{2}{\pi}}\right) & Y_{t-1} < 0 \end{cases}$$

From Section 3 we know that the distributions of $D_t^{-1}\mathrm{x}_t | Y_{t-1} < 0$ and $D_t^{-1}\mathrm{x}_t | Y_{t-1} > 0$ are skew-normal:

$$\left(D_t^{-1}\mathrm{x}_t | Y_{t-1} > 0\right) \sim SN_p\left(\gamma\sqrt{\frac{2}{\pi}}, \Omega_+, \alpha_+\right)$$
$$\left(D_t^{-1}\mathrm{x}_t | Y_{t-1} < 0\right) \sim SN_p\left(\gamma\sqrt{\frac{2}{\pi}}, \Omega_-, \alpha_-\right)$$

Apply now linear properties of the *SN* distribution:

$$(\bar{\mathrm{r}}_t | Y_{t-1} > 0) \sim SN_p\left(0, I_p, \lambda_+\right), \quad \lambda_+ = \Omega_+^{1/2}\alpha_+$$
$$(\bar{\mathrm{r}}_t | Y_{t-1} < 0) \sim SN_p\left(0, I_p, \lambda_-\right), \quad \lambda_- = \Omega_-^{1/2}\alpha_-$$

First notice that the vectors $\{\bar{\mathrm{r}}_t\}$ are i.i.d. Moreover, the distribution of $\bar{\mathrm{r}}_t$ is the mixture, with equal weights, of two p-dimensional *SN* distributions with zero location parameter and the identity matrix for the scale parameter, that is:

$$f(\bar{\mathrm{r}}_t) = \frac{1}{(2\pi)^{p/2}} \exp\left(-\frac{\bar{\mathrm{r}}_t^{\mathrm{T}}\bar{\mathrm{r}}_t}{2}\right)\left[\Phi\left(\lambda_-^{\mathrm{T}}\bar{\mathrm{r}}_t\right) + \Phi\left(\lambda_+^{\mathrm{T}}\bar{\mathrm{r}}_t\right)\right]$$

The above density can be represented as follows:

$$f(\bar{\mathrm{r}}_t) = 2\phi_p\left(\bar{\mathrm{r}}_t; I_p\right)\omega(\bar{\mathrm{r}}_t)$$

where $\phi_p(\cdot;\Sigma)$ denotes the density of $N_p(0,\Sigma)$ and $w(\cdot)$ is a function such that $0 \leq w(-\bar{\mathrm{r}}_t) = 1 - w(\bar{\mathrm{r}}_t) \leq 1$. Hence the distribution of $\bar{\mathrm{r}}_t$ is generalized *SN* with the zero location parameter and the identity matrix for the scale parameter (Loperfido, 2004; Genton & Loperfido, 2005). It follows that the pdf of even functions $g(\bar{\mathrm{r}}_t) = g(-\bar{\mathrm{r}}_t)$ does not depend on $w(\cdot)$. As an immediate consequence, we have:

$$\bar{S}_t = \bar{\mathrm{r}}_t^{\mathrm{T}}\bar{\mathrm{r}}_t \sim \chi_p^2 \Rightarrow E\left(\bar{S}_t\right) = p, \quad V\left(\bar{S}_t\right) = 2p$$

A standard application of the Central Limit Theorem leads to

$$\sqrt{\frac{n}{2p}}\left(\frac{1}{n}\sum_{t=1}^{n}\bar{S}_t - p\right) \overset{a}{\sim} N(0,1)$$

When the sample size is large, the distribution of $\{\mathrm{r}_t\}$ approximates the distribution of $\{\bar{\mathrm{r}}_t\}$. Hence the squared norms $\{\mathrm{S}_t\}$ are approximately independent and identically distributed according to a χ_p^2 distribution. Moreover,

$$\sqrt{\frac{n}{2p}}\left(\frac{1}{n}\sum_{t=1}^{n}S_t - p\right) \overset{a}{\sim} N(0,1)$$

and this completes the proof. □