Journal of Financial Econometrics, 2009, Vol. 7, No. 2, 77–105

The Impact of Shocks on Higher Moments

ERIC JONDEAU
Swiss Finance Institute and University of Lausanne

MICHAEL ROCKINGER
Swiss Finance Institute, University of Lausanne, and CEPR

ABSTRACT

In this paper, we extend the concept of the news impact curve of volatility developed by Engle and Ng (1993) to the higher moments and co-moments of the multivariate generalized autoregressive conditional heteroskedasticity (GARCH) model with non-normal innovations. For this purpose, we present a new methodology to describe the joint distribution of GARCH processes in a non-normal setting. Then, we provide expressions for the response of the moments of the subsequent distribution to a shock. This tool enhances the understanding of the temporal evolution of the joint distribution. We use our methodology to provide stylized facts for the four largest international stock markets. In particular, we document the persistence of large (positive or negative) daily returns. In a multivariate setting, we find that foreign holdings provide a good hedge against changes in domestic volatility after good shocks but a bad hedge after crashes. Finally, using generalized impulse responses, we show that the effect of shocks on the higher moments of the distribution is short-lasting. (*JEL*: C22, C51, G12)

KEYWORDS: GARCH model, non-normality, kurtosis, skewness, stock returns, volatility

Both of us acknowledge help from the Swiss National Science Foundation through NCCR FINRISK (Financial Valuation and Risk Management). We benefited from comments and discussions with Tim Bollerslev, Ales Cerny, René Garcia, Eric Jacquier, Nour Meddahi, and Raman Uppal. We are grateful to three referees as well as to the associate editor for their insightful comments. The usual disclaimer applies. A Technical Appendix containing the proofs of our propositions and additional results is available on the webpage "http://www.hec.unil.ch/ejondeau/Publications/Publications.htm."
Address correspondence to Michael Rockinger, Swiss Finance Institute, University of Lausanne, and CEPR. Faculty of Business and Economics, University of Lausanne. CH 1012 Lausanne, Switzerland, or e-mail: michael.rockinger@unil.ch.

doi: 10.1093/jjfinec/nbn017
Advance Access publication December 9, 2008

Analyzing the distributional and dynamic properties of asset returns is a very active area of research in theoretical as well as empirical finance. Early empirical evidence has shown that extreme returns occur too often to be consistent with normality and that large, negative shocks (crashes) occur more often than large, positive shocks (booms). This evidence suggests that returns are driven by asymmetric and leptokurtic distributions. Asset returns have also been found to be time-dependent. Although the predictability of returns is still a matter of debate, the predictability of volatility is acknowledged as a major feature of return dynamics, often described by the so-called generalized autoregressive conditional heteroskedasticity (GARCH) models (Bollerslev 1986; Bollerslev, Chou, and Kroner 1992). In this context, Engle (1982) has shown that return non-normality and volatility time-variability are two related phenomena because the latter contributes to the former. At a low frequency (say monthly), GARCH models may allow the recovery of normally distributed innovations. However, at a high frequency (say daily), the distribution of the innovation process generally remains highly non-normal. In the quest for the best suited conditional distribution, several candidates have been proposed: the Student *t* distribution (Bollerslev 1987), the entropy distribution (Jondeau and Rockinger 2002), or various forms of asymmetric *t* distributions (Hansen 1994; Harvey and Siddique 1999). In addition, several recent papers have provided evidence that the characteristics of this conditional distribution vary over time.[1] This finding implies that the probability distribution of the innovation depends on recent events.

Regarding the modeling of the joint behavior of asset returns, most of the recent contributions have focused on the dynamics of either the conditional covariance matrix (as in the first-generation multivariate GARCH models; see Bollerslev, Engle, and Wooldridge 1988 and Engle and Kroner 1995) or the conditional correlation matrix (as in the more recent dynamic conditional correlation (DCC) models; see Engle 2002 and Cappiello, Engle, and Sheppard 2006). In this multivariate setting, some authors have explicitly introduced non-normal distributions, such as the mixture of normal densities (Vlaar and Palm 1993), the Student *t* distribution (Bollerslev and Wooldridge 1992; Harvey, Ruiz, and Sentana 1992) or, more recently, various forms of skewed Student *t* distribution (Sahu, Dey, and Branco 2003; Fiorentini, Sentana, and Calzolari 2003; Bauwens and Laurent 2005; Mencia and Sentana 2005). Some desirable features of a multivariate model are the time-variability of variances and correlations as well as the asymmetry and leptokurticity of the joint distribution. The ability to reproduce these properties of the empirical multivariate distribution is important because they are related to the well-documented contagion effect.[2] Another issue that has not been addressed so far is the time-variability of the conditional distribution, particularly how the joint distribution is affected by past shocks.

[1]The sequence of contributions by Hansen (1994), Harvey and Siddique (1999), Jondeau and Rockinger (2002, 2003), and Brooks et al. (2005) provides univariate models describing time variation in volatility, skewness, and kurtosis.

[2]This effect has also been analyzed, by means of the extreme value theory, by Longin and Solnik (2001) and Poon, Rockinger, and Tawn (2004).

As a first contribution, we present a new methodology to investigate, both in univariate and multivariate settings, the effect of shocks on the conditional distribution of asset returns. We consider two asymmetric multivariate specifications for the covariance matrix (the asymmetric BEKK model of Kroner and Ng (1998) and the asymmetric DCC model of Cappiello, Engle, and Sheppard (2006)), in which innovations are drawn from an asymmetric and leptokurtic distribution. Realizing the difficulty of understanding the complex dynamics generated by a negative shock, we follow Engle and Ng (1993) and develop a graphical tool to summarize the impact of past shocks on the subsequent characteristics of the returns' distribution. In the univariate setting, this leads us to introduce the concept of news impact curve (NIC) of skewness and kurtosis, which extends the well-known NIC of volatility developed by Engle and Ng (1993). In the multivariate setting, we obtain an analogous tool for the joint distribution, namely the news impact surface (NIS) for given higher moments and co-moments.[3] We describe how the dynamics of the response to shocks can be analyzed through impulse response functions.

In terms of empirical contribution, we investigate the actual patterns of responses to shocks in the context of daily index returns for the four largest international markets. We show that, after a large negative (or positive) shock, the subsequent conditional distribution tends to have fatter tails and be negatively (or positively) skewed. This result suggests that large shocks of a given sign are positively correlated. In the multivariate framework, we observe a similar phenomenon: after joint negative (or positive) shocks, the probability of subsequent joint negative (or positive) shocks increases.

The outline of the paper is as follows. In Section 1, we present the multivariate statistical model describing the evolution of returns. In Section 2, we provide the main theoretical results concerning NIC and NIS. In Section 3, we present the data and comment on the estimation of the model. In Section 4, we discuss some stylized facts on the evolution of returns that we infer from the NIC and NIS. In Section 5, we go one step further and describe how to construct generalized impulse responses in this context. Section 6 concludes.

1 A MULTIVARIATE TIME-VARYING CONDITIONAL DISTRIBUTION

Our multivariate conditional setting incorporates most statistical features required for modeling stock market returns. First, it accounts for the well-known time-dependence properties, namely volatility clustering (Engle 1982) and persistence in correlations (Engle 2002). Second, it is well suited to capture both the asymmetry and the leptokurticity often found in the distribution of market returns. After presenting the general setup, we will then briefly describe these two components.

Let $r_t = (r_{1,t}, \ldots, r_{n,t})'$, for $t = 1, \ldots, T$, be a time series of n asset returns. It is convenient to split the data generating process of r_t into three components:

$$r_t = \mu_t(\theta | I_{t-1}) + \varepsilon_t, \tag{1}$$

[3]Depending on the context, NIC and NIS may be singular or plural.

$$\varepsilon_t = \Sigma_t (\theta | I_{t-1})^{1/2} z_t, \quad \text{and} \tag{2}$$

$$z_t \sim g(z_t | \eta_t). \tag{3}$$

Equation (1) decomposes the return at time t as the sum of the $n \times 1$ vector of conditional means, $\mu_t \equiv \mu_t(\theta|I_{t-1}) = E_{t-1}[r_t]$, and the $n \times 1$ vector of unexpected returns, ε_t, where E_{t-1} denotes the expectation conditional on the information available at date $t-1$, denoted I_{t-1}.

Equation (2) indicates that unexpected returns ε_t are a combination of n independent innovations z_t. The conditional covariance matrix is denoted by $\Sigma_t \equiv \Sigma_t(\theta|I_{t-1}) = E_{t-1}[(r_t - \mu_t)(r_t - \mu_t)']$. We denote by $\Sigma_t^{1/2}$ a matrix such that $\Sigma_t = \Sigma_t^{1/2}\Sigma_t^{1/2\prime}$. Given the way that nonsynchronicity is dealt with in our empirical application, we primarily focus on the Cholesky decomposition. An alternative way to construct the "square root" of the covariance matrix is to use a spectral decomposition, such that $\Sigma_t = V_t \Lambda_t V_t'$, where V_t is the $n \times n$ matrix of eigenvectors, and Λ_t is the diagonal matrix of eigenvalues. In this case, one simply has $\Sigma_t^{1/2} = V_t \Lambda_t^{1/2} V_t'$.[4] The vector θ contains all the parameters associated with the conditional mean vector and the conditional covariance matrix.

The vector of innovations, $z_t = \Sigma_t^{-1/2}(r_t - \mu_t)$, has zero mean and identity covariance matrix and is a martingale difference. Equation (3) specifies that innovations are drawn from a conditional distribution g with, possibly time-varying, shape parameters η_t.

1.1 Dynamics of First and Second Moments

To capture possible serial autocorrelation in returns, we assume an AR(p) structure for returns, implying the following conditional mean:

$$\mu_t = \mu + \varphi_1 r_{t-1} + \cdots + \varphi_p r_{t-p}, \tag{4}$$

where μ is an $n \times 1$ vector, and φ_k are $n \times n$ diagonal matrices, $k = 1, \ldots, p$.

We consider two alternative specifications for the covariance matrix Σ_t. The first model is the asymmetric version of the BEKK model proposed by Kroner and Ng (1998). This approach allows a very general specification for the covariance matrix, yet it requires the estimation of a large number of parameters. The second model is the asymmetric version of the DCC model proposed by Cappiello, Engle, and Sheppard (2006). This model is more restrictive, though it captures persistence and asymmetry in conditional correlations.

The asymmetric BEKK (ABEKK) specification is given by

$$\Sigma_t = \bar{\Omega} + B'\Sigma_{t-1}B + A'\varepsilon_{t-1}\varepsilon_{t-1}'A + G'\zeta_{t-1}\zeta_{t-1}'G, \tag{5}$$

[4] In the Technical Appendix, we investigate the consequences of changing the ordering of the variables in the Cholesky decomposition and obtain essentially the same results as those reported in Section 3. We also report and comment on results obtained with the spectral decomposition. Mencia and Sentana (2005) use a location scale mixture of normals to obtain a parameterization that is independent of the decomposition of the covariance matrix. Such a parameterization does not hold for the multivariate distribution we consider in this paper.

where $\bar{\Omega}$ is an $n \times n$ positive definite and symmetric matrix, A, B, and G are $n \times n$ matrices, and $\zeta_{t-1} = \varepsilon_{t-1}\ 1_{\{\varepsilon_{t-1} \leq 0\}}$ captures the possible asymmetric effect of past shocks on the variances and covariances. Because the last three terms on the right-hand side of Equation (5) are expressed in quadratic form, the conditional covariance matrix is positive definite provided $\bar{\Omega}$ is positive definite. This specification involves $[n(n+1)/2] + 3n^2$ unknown parameters. The constant term matrix $\bar{\Omega}$ can be estimated from sample moments as $\text{vec}(\bar{\Omega}) = (I_{n^2} - (A \otimes A)' - (B \otimes B)')\text{vec}(\bar{\Sigma}) - (G \otimes G)'\text{vec}(\bar{M})$, where $\bar{\Sigma}$ and $\bar{M}$ denote the unconditional covariance matrices of ε_t and ζ_t, respectively.

In the asymmetric DCC (ADCC) model that we consider, each conditional variance, $\sigma^2_{i,t}$, is given by

$$\sigma^2_{i,t} = \omega_i + \beta_i \sigma^2_{i,t-1} + \alpha_i \varepsilon^2_{i,t-1} + \psi_i \zeta^2_{i,t-1}, \quad i = 1, \ldots, n, \tag{6}$$

and the conditional correlation matrix, $\Gamma_t = \{\rho_{ij,t}\}_{i,j=1,\ldots,n}$, is

$$\Gamma_t = (\text{diag}(Q_t))^{-1/2} \cdot Q_t \cdot (\text{diag}(Q_t))^{-1/2}, \quad \text{and} \tag{7}$$

$$Q_t = \tilde{\Omega} + \delta_1 Q_{t-1} + \delta_2(u_{t-1}u'_{t-1}) + \delta_3(n_{t-1}n'_{t-1}), \tag{8}$$

where $u_t = D_t^{-1}\varepsilon_t = \{\varepsilon_{i,t}/\sigma_{i,t}\}_{i=1,\ldots,n}$ is the vector of normalized unexpected returns, and $n_{t-1} = u_{t-1}\ 1_{\{u_{t-1} \leq 0\}}$. The term $\text{diag}(Q_t)$ denotes a matrix with zeros, except for the diagonal that contains the diagonal of Q_t, and $D_t = \{\sigma_{i,t}\}_{i=1,\ldots,n}$ is the $n \times n$ diagonal matrix with the standard deviations on its diagonal and 0 elsewhere. Parameters δ_1, δ_2, and δ_3 are restricted to ensure that the conditional correlation matrix is positive definite. Finally, the covariance matrix Σ_t is simply defined as $\Sigma_t = D_t \Gamma_t D_t$. This specification involves $[n(n+1)/2] + 4n + 3$ unknown parameters.[5] The matrix $\tilde{\Omega}$ can be estimated as $\tilde{\Omega} = [(1 - \delta_1 - \delta_2)\bar{Q} - \delta_3 \bar{N}]$, where $\bar{Q}$ and $\bar{N}$ are the unconditional covariance matrices of u_t and n_t, respectively.

1.2 Conditional Distribution

Although joint normality of innovations is often assumed in the multivariate modeling of asset returns (see, for instance, Kroner and Ng 1998; Ang and Chen 2002; or Cappiello, Engle, and Sheppard 2006), we adopt a non-normal conditional distribution in this paper. Each component $z_{i,t}$ is drawn from an independent skewed t (Sk-t) distribution (Hansen 1994; Fernández and Steel 1998). The joint distribution

[5]When the parameters of the constant terms $\bar{\Omega}$ and $\tilde{\Omega}$ are estimated from sample moments, with $\bar{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_t \varepsilon'_t$, $\bar{M} = \frac{1}{T}\sum_{t=1}^{T} \zeta_t \zeta'_t$, $\bar{Q} = \frac{1}{T}\sum_{t=1}^{T} u_t u'_t$, and $\bar{N} = \frac{1}{T}\sum_{t=1}^{T} n_t n'_t$, the number of parameters that must be estimated numerically is a quadratic function of the number of assets for the ABEKK model, but it is only a linear function for the ADCC model. For a bivariate system, the difference is small (12 for the ABEKK and 11 for the ADCC). For a four-dimensional system, the difference is much more sizeable (48 for the ABEKK and 19 for the ADCC).

of the $n \times 1$ vector of innovations z_t is therefore defined as[6]

$$g(z_t|\eta) = \prod_{i=1}^{n} \frac{2b_i}{\xi_i + \frac{1}{\xi_i}} \frac{\Gamma\left(\frac{\nu_i+1}{2}\right)}{\sqrt{\pi(\nu_i - 2)}\,\Gamma\left(\frac{\nu_i}{2}\right)} \left(1 + \frac{\kappa_{i,t}^2}{\nu_i - 2}\right)^{-\frac{\nu_i+1}{2}}, \tag{9}$$

where

$$\kappa_{i,t} = \begin{cases} (b_i z_{i,t} + a_i)\xi_i, & \text{if } z_{i,t} \leq -a_i/b_i, \\ (b_i z_{i,t} + a_i)/\xi_i, & \text{if } z_{i,t} > -a_i/b_i, \end{cases}$$

and $a_i = M_{i,1}$ as defined in Equation (10) below and $b_i^2 = \xi_i^2 + 1/\xi_i^2 - 1 - a_i^2$. The vector $\eta = (\nu_1, \ldots, \nu_n, \xi_1, \ldots, \xi_n)'$ collects the shape parameters.

In this setup, the dependence between unexpected returns is captured by the covariance matrix through $\varepsilon_t = \Sigma_t^{1/2} z_t$, while the innovations $z_{i,t}$ are distributed independently from each other. This approach allows us to explicitly separate the modeling of the multivariate conditional distribution (through the parameters of the conditional distribution) from the modeling of the multivariate dependence (through the parameters of the covariance matrix).

The marginal distribution of $z_{i,t}$ is a univariate Sk-t distribution $g(z_{i,t}|\nu_i, \xi_i)$, where ν_i and ξ_i correspond to the degree of freedom and the asymmetry parameter, respectively. Each marginal distribution is defined for $2 < \nu_i < \infty$ and $\xi_i > 0$. Moments up to the fourth exist if $\nu_i > 4$. The constants a_i and b_i, which are introduced in the definition of $\kappa_{i,t}$, ensure that $z_{i,t}$ has zero mean and unit variance. This, in turn, ensures that μ_t and Σ_t can be interpreted as the conditional mean vector and the conditional covariance matrix of r_t, respectively. The moment of order r of $z_{i,t}$ is given by

$$M_{i,r} = m_{i,r} \frac{\xi_i^{r+1} + \frac{(-1)^r}{\xi_i^{r+1}}}{\xi_i + \frac{1}{\xi_i}} \quad \text{with} \quad m_{i,r} = \frac{\Gamma\left(\frac{\nu_i - r}{2}\right)\Gamma\left(\frac{r+1}{2}\right)(\nu_i - 2)^{\frac{r+1}{2}}}{\sqrt{\pi(\nu_i - 2)}\,\Gamma\left(\frac{\nu_i}{2}\right)}. \tag{10}$$

Provided that they exist, the skewness and kurtosis of $z_{i,t}$ are then given by

$$sk_i^Z = E\left[Z_i^3\right] = M_{i,3} - 3M_{i,1}M_{i,2} + 2M_{i,1}^3, \text{ and} \tag{11}$$

$$ku_i^Z = E\left[Z_i^4\right] = M_{i,4} - 4M_{i,1}M_{i,3} + 6M_{i,2}M_{i,1}^2 - 3M_{i,1}^4. \tag{12}$$

As the previous equations clearly show, these terms are directly related in a nonlinear way to the degree of freedom ν_i and the asymmetry parameters ξ_i. Note that $sk_i^Z = 0$ when $\xi_i = 1$.

The objective of the paper is to investigate how shocks on asset returns will affect the shape of the subsequent distribution. For this purpose, we use the bijection

[6] Alternative strategies could extend the Student t distribution to multivariate random variables. One may assume that the χ^2, which appears in the definition of the t distribution, is the same for each component. Such an extension has been analyzed by Sahu, Dey, and Branco (2003) and Bauwens and Laurent (2005). In such cases, however, innovations are not independent.

between the parameters of the distribution and its first four moments. We now describe how the moments of unexpected returns, ε_t, are related to the characteristics of the shocks, z_t. Given $\Sigma_t^{1/2} = (\omega_{ij,t})_{i,j=1,\cdots,n}$, we have $\varepsilon_{i,t} = \sum_{r=1}^{n} \omega_{ir,t} z_{r,t}$. Clearly, $E_{t-1}[\varepsilon_t] = 0$ and $V_{t-1}[\varepsilon_t] = \Sigma_t$. Then, we define the conditional co-skewness and co-kurtosis between unexpected returns as

$$sk^{\varepsilon}_{ijk,t} = \frac{\mathcal{S}^{\varepsilon}_{ijk,t}}{\sigma_{i,t}\sigma_{j,t}\sigma_{k,t}} \qquad \text{and} \qquad ku^{\varepsilon}_{ijkl,t} = \frac{\mathcal{K}^{\varepsilon}_{ijkl,t}}{\sigma_{i,t}\sigma_{j,t}\sigma_{k,t}\sigma_{l,t}}, \tag{13}$$

where $S^{\varepsilon}_{ijk,t} = E_{t-1}[\varepsilon_{i,t}\varepsilon_{j,t}\varepsilon_{k,t}]$ and $K^{\varepsilon}_{ijkl,t} = E_{t-1}[\varepsilon_{i,t}\varepsilon_{j,t}\varepsilon_{k,t}\varepsilon_{l,t}]$ are the conditional third and fourth central moments, respectively. Given the properties of z_t, we have

$$\mathcal{S}^{\varepsilon}_{ijk,t} = E_{t-1}\left[\sum_{r=1}^{n} \omega_{ir,t}\omega_{jr,t}\omega_{kr,t}\, z^3_{r,t}\right] = \sum_{r=1}^{n} \omega_{ir,t}\omega_{jr,t}\omega_{kr,t}\, sk^Z_r \tag{14}$$

and

$$\begin{aligned}\mathcal{K}^{\varepsilon}_{ijkl,t} &= E_{t-1}\left[\sum_{r=1}^{n} \omega_{ir,t}\omega_{jr,t}\omega_{kr,t}\omega_{lr,t}\, z^4_{r,t}\right] + E_{t-1}\left[\sum_{r=1}^{n}\sum_{s\neq r} \psi_{rs,t}\, z^2_{r,t} z^2_{s,t}\right] \\ &= \sum_{r=1}^{n} \omega_{ir,t}\omega_{jr,t}\omega_{kr,t}\omega_{lr,t}\, ku^Z_r + \sum_{r=1}^{n}\sum_{s\neq r} \psi_{rs,t}, \end{aligned} \tag{15}$$

where $\psi_{rs,t} = \omega_{ir,t}\omega_{jr,t}\omega_{ks,t}\omega_{ls,t} + \omega_{ir,t}\omega_{js,t}\omega_{kr,t}\omega_{ls,t} + \omega_{is,t}\omega_{jr,t}\omega_{kr,t}\omega_{ls,t}$. These expressions are true for all possible values of i, j, k, and l, such that they also apply to the individual conditional third and fourth central moments. The time-variability of co-skewness and co-kurtosis between unexpected returns clearly has two possible sources. On the one hand, the covariance matrix Σ_t is time varying, such that the $\omega_{ij,t}$ terms are also time varying. On the other hand, individual skewness and kurtosis of innovations may also be time varying, as we will describe in the next section. In absence of asymmetry in the univariate distributions ($sk^Z_r = 0$, $\forall r$), no asymmetry will be found in the multivariate distribution of returns (see Equation (14)). In addition, because the $z_{r,t}$ terms are independent from each other, the co-skewness are equal to zero when the covariance matrix is diagonal. The co-kurtosis is constituted of two blocks. The first one corresponds to the individual kurtosis ku^Z_r. The second one involves products of the form $E_{t-1}[z^2_{r,t} z^2_{s,t}]$, which are equal to 1 for $r \neq s$. We shall provide more interpretations regarding the higher co-moments in Section 4.3, when we comment on the response of these higher co-moments to shocks.

In Figure 1, we represent bivariate contour plots of the Sk-t distribution. Figures on the left represent the distribution of the (uncorrelated) innovations, while figures on the right represent the distribution of the unexpected returns, assuming a correlation of 0.5. Top figures are obtained for symmetric marginal distributions. In the middle figures, both marginal densities have negative skewness, whereas the densities in the bottom figures have opposite skewness. The figures show that, once innovations' distributions and correlations are combined, very different patterns can be obtained for the distribution of unexpected returns. Middle figures

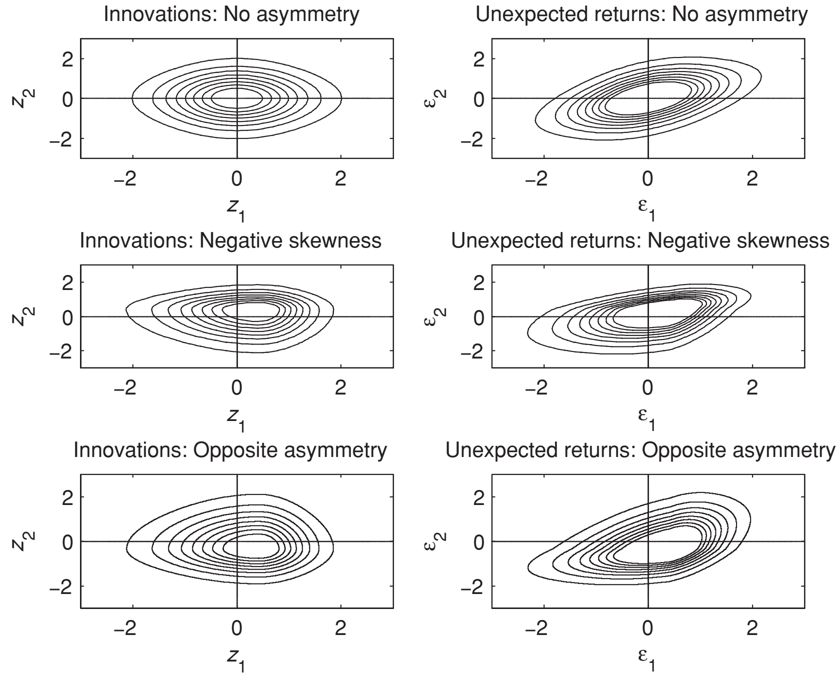

Figure 1 This figure displays contour plots of the Sk-t distribution in the bivariate case. The left figures represent cases where the marginal distributions are uncorrelated. The right figures correspond to a correlation of 0.5. The upper figures are obtained for symmetric marginal distributions ($\xi_1 = \xi_2 = 1$). In the middle figures, both marginal densities have negative skewness ($\xi_1 = \xi_2 = 0.5$), whereas the bottom figures have opposing skewness with the marginal density, which is distributed along the horizontal axis and negatively skewed ($\xi_1 = 0.5$, $\xi_2 = 1.5$). In all cases, the degrees of freedom are $\nu_1 = \nu_2 = 10$.

illustrate the situation most frequently encountered in the empirical part of this paper, namely when both marginal distributions are negatively skewed.

1.3 Dynamics of the Higher Moments

We will now describe how we model the temporal evolution of the conditional distribution's shape parameters, η_t. In the case of the Sk-t distribution, η_t includes the degree of freedom and the asymmetry parameter $(\nu_{i,t}, \xi_{i,t})_{i=1,\ldots,n}$. A natural approach is to render the shape parameters dependent on past shocks, as in $\eta_t = \eta(z_{t-1}, z_{t-2}, \ldots)$. These dynamics should be constrained to ensure that the function g is a well-defined distribution.[7] Once the dynamics of the shape parameters are estimated, higher moments are deduced from Equations (11) and (12).

[7] A similar approach was adopted by Hansen (1994) and Harvey and Siddique (1999). Hansen (1994) was the first one to model the dynamics of conditional higher moments, yielding the concept of autoregressive conditional density (ARCD). Jondeau and Rockinger (2003) discuss several possible specifications for the dynamics of the shape parameters.

We adopt the following asymmetric GARCH-like specifications:

$$(1 - c_{i,2}L)\log(\nu_{i,t} - \underline{\nu}) = c_{i,0} + c_{i,1}^{-}|z_{i,t-1}|N_{i,t-1} + c_{i,1}^{+}|z_{i,t-1}|(1 - N_{i,t-1}), \quad (16)$$

$$(1 - d_{i,2}L)\log(\xi_{i,t}) = d_{i,0} + d_{i,1}^{-}z_{i,t-1}N_{i,t-1} + d_{i,1}^{+}z_{i,t-1}(1 - N_{i,t-1}), \quad (17)$$

where L is the lag operator, and $N_{i,t} = 1_{\{z_{i,t} \leq 0\}}$. The parameter $\underline{\nu}$ is the lower bound for the degree of freedom. To ensure that moments up to the fourth exist, we set $\underline{\nu} = 4$.

Three main features of these specifications are worth emphasizing. First, the degree of freedom $\nu_{i,t}$ is related to the absolute value of lagged innovations, because a large shock $z_{i,t-1}$ is expected to affect the heaviness of the distribution's tails regardless of its sign. In contrast, the dynamics of the asymmetry parameter naturally depends on signed innovations, because $\xi_{i,t}$ is likely to reflect the sign and size of recent shocks. Second, instead of assuming that the impacts of positive and negative shocks are of the same magnitude on the distribution's shape, we allow the shape parameters to react asymmetrically to recent shocks. Finally, Equations (16) and (17) include a lag of the dependent variable to capture possible persistence in the dynamics of the higher moments.

Our specification makes it possible to explore the strengths of the various sources of asymmetry in the model: (i) shocks are allowed to affect variances and correlations asymmetrically; (ii) the conditional distribution is itself asymmetric, such that large shocks of a given sign may be more likely to occur than large shocks of the other sign; and (iii) the extent of the asymmetry and the thickness of the conditional distribution can be altered in an asymmetric way by past shocks, depending on their sign.

1.4 Estimation

Under normality, the DCC model can be estimated in two steps (Engle 2002). In the case of the Sk-t distribution, the parameters pertaining to the first and second moments have to be estimated jointly with the parameters pertaining to the conditional distribution.

The sample log-likelihood function of the multivariate DCC model with Sk-t distribution is therefore

$$\log L\,(r_1, \ldots, r_T|\theta, \eta) = \sum_{t=1}^{T}\left[\log\left(g\left(\Sigma_t(\theta)^{-1/2}(r_t - \mu_t(\theta))|\eta\right)\right) - \frac{1}{2}\log|\Sigma_t(\theta)|\right], \quad (18)$$

where

$$\theta = (\theta_1, \ldots, \theta_n, \delta_1, \delta_2, \delta_3)' \quad \text{with} \quad \theta_i = (\mu_i, \varphi_{i,1}, \ldots, \varphi_{i,p}, \omega_i, \alpha_i, \beta_i, \psi_i)',$$

$$\eta = (\eta_1, \ldots, \eta_n)' \quad \text{with} \quad \eta_i = (c_{i,0}, c_{i,1}^{-}, c_{i,1}^{+}, c_{i,2}, d_{i,0}, d_{i,1}^{-}, d_{i,1}^{+}, d_{i,2})'.$$

Maximizing expression (18) with respect to parameter vectors θ and η yields the maximum-likelihood (ML) estimates. The log-likelihood of the model is very nonlinear in the parameter set, in particular the parameters driving the conditional correlation and the degree-of-freedom and asymmetry parameter dynamics. For this reason, it is not possible to compute the covariance matrix of the parameter estimates using the analytical gradient and Hessian.

2 NEWS IMPACT CURVES AND SURFACES

The NIC was introduced by Engle and Ng (1993) to represent the response of volatility to a shock on asset returns. More precisely, it measures the effect of a shock at date t on the volatility at date $t+1$, while the information dated $t-1$ and earlier is held constant. The NIC has been extended to the response of the conditional correlation to shocks on two asset returns by Kroner and Ng (1998) as well as Cappiello, Engle, and Sheppard (2006). We extend this concept to the response of the conditional distribution to shocks. This is done in two steps. We begin with the NIC of the individual higher moments of innovations, $z_{i,t}$. Then, we construct the NIS of the higher moments of unexpected returns, $\varepsilon_{i,t}$, which also involves the response of the covariance matrix to shocks.

2.1 News Impact Curves for the Marginal Distribution

At date t, a shock z occurs, while all the characteristics of the unexpected returns (mean, variance, and shape parameters) are equal to their unconditional levels. The shock z affects the level of the shape parameters at date $t+1$, which in turn affect the higher moments of the marginal distribution.

The following proposition provides the various moments at date $t+1$ conditioned on a shock of value z at date t.

Proposition 1. *(1) The NIC of the conditional distribution's shape parameters are given by*

$$\nu_i^Z(z) = \begin{cases} \underline{\nu} + \exp(A_{\nu,i} + c_{i,1}^- |z|), & \text{if } z \leq 0, \\ \underline{\nu} + \exp(A_{\nu,i} + c_{i,1}^+ |z|), & \text{if } z > 0, \end{cases} \tag{19}$$

$$\xi_i^Z(z) = \begin{cases} \exp(A_{\xi,i} + d_{i,1}^- z), & \text{if } z \leq 0, \\ \exp(A_{\xi,i} + d_{i,1}^+ z), & \text{if } z > 0, \end{cases} \tag{20}$$

where $A_{\nu,i} = c_{i,0} + c_{i,2}\log(\bar{\nu}_i - \underline{\nu})$ *and* $A_{\xi,i} = d_{i,0} + d_{i,2}\log(\bar{\xi}_i)$, *and where* $\bar{\nu}_i$ *and* $\bar{\xi}_i$ *denote the unconditional levels of the shape parameters.*

(2) The NIC of the conditional skewness and kurtosis of innovations are given by

$$sk_i^Z(z) = \frac{\xi_i^Z(z)}{1+\xi_i^Z(z)^2}\big[C_{i,4}\big(m_{i,3} - 3m_{i,1}m_{i,2} + 2m_{i,1}^3\big) + C_{i,2}\big(3m_{i,1}m_{i,2} - 4m_{i,1}^3\big)\big], \tag{21}$$

$$\begin{aligned} ku_i^Z(z) = {} & \frac{\xi_i^Z(z)}{1+\xi_i^Z(z)^2}\big[C_{i,5}\big(m_{i,4} - 4m_{i,1}m_{i,3} + 6m_{i,1}^2 m_{i,2} - 3m_{i,1}^4\big) \\ & + C_{i,3}\big(4m_{i,1}m_{i,3} - 12m_{i,1}^2 m_{i,2} + 9m_{i,1}^4\big) \\ & + C_{i,1}\big(6m_{i,1}^2 m_{i,2} + 12m_{i,1}^4\big)\big], \end{aligned} \tag{22}$$

where $m_{i,r}$ is defined in Equation (10), and $C_{i,r} = \xi_i^Z(z)^r - \xi_i^Z(z)^{-r}$. The NIC of skewness and kurtosis of unexpected returns are $sk_i^\varepsilon(z) = sk_i^Z(z)$ and $ku_i^\varepsilon(z) = ku_i^Z(z)$, respectively. (3) The NIC of the conditional third and fourth moments of unexpected returns are given by

$$\mathcal{S}_i^\varepsilon(z) = sk_i^Z(z)\sigma_i(z)^3 \qquad \text{and} \qquad \mathcal{K}_i^\varepsilon(z) = ku_i^Z(z)\sigma_i(z)^4,$$

where the expression for the variance $\sigma_i(z)^2$ depends on the specification adopted for the conditional covariance matrix.

Proof. All proofs are available in the Technical Appendix. The expressions of $\sigma_i(z)^2 = (\Sigma(z))_{ii}$ for the ABEKK and ADCC models are provided in Proposition 2. ■

In Proposition 1, the unconditional levels of the shape parameters ($\bar{\nu}_i$ and $\bar{\xi}_i$) correspond to the values of these parameters that would prevail in the absence of shocks. For specifications (16) and (17), these values can be obtained from

$$(1 - c_{i,2})\log(\bar{\nu}_i - \underline{\nu}) = c_{i,0} + (c_{i,1}^- \bar{N}_i \bar{\xi}_i + c_{i,1}^+ (1 - \bar{N}_i)/\bar{\xi}_i) E[|t_i|],$$
$$(1 - d_{i,2})\log(\bar{\xi}_i) = d_{i,0} + (-d_{i,1}^- \bar{N}_i \bar{\xi}_i + d_{i,1}^+ (1 - \bar{N}_i)/\bar{\xi}_i) E[|t_i|],$$

where $\bar{N}_i = \Pr[Z_i > 0 \mid \bar{\nu}_i, \bar{\xi}_i] = \bar{\xi}_i^2/(1 + \bar{\xi}_i^2)$ and $E[|t_i| \mid \bar{\nu}_i] = m_{i,1}$ is the expected value of the absolute value of a standard t variable with $\bar{\nu}_i$ degrees of freedom. Solving these two expressions numerically provides the estimates of $\bar{\nu}_i$ and $\bar{\xi}_i$.

The NIC of conditional skewness and kurtosis of unexpected returns $\varepsilon_{i,t}$ are defined per unit of standard deviation. We deduce that the higher moments of the unexpected return distribution are independent of the volatility dynamics in a univariate model. In contrast, the NIC of the third and fourth moments of unexpected returns incorporate the additional effect of the volatility dynamics.

2.2 News Impact Surfaces for the Bivariate Distribution

We now consider a set of shocks $z = (z_1, z_2)'$ at date t and evaluate their effect on the covariance, co-skewness, and co-kurtosis matrices of returns at date $t + 1$. We then obtain NIS because each component of these matrices is affected by a combination of shocks. We notice that a set of shocks z at date t translates instantaneously to unexpected returns at date t through $\varepsilon(z) = \bar{\Sigma}^{1/2} z$, where $\bar{\Sigma}$ is the unconditional covariance matrix, i.e., the covariance matrix that would prevail in absence of shocks at date $t - 1$ or earlier.[8] Similarly, the normalized unexpected returns are defined as $u(z) = \bar{D}^{-1}\varepsilon(z)$, where $\bar{D}$ is the matrix with unconditional standard deviations on its diagonal. Proposition 2 gives the expressions for the NIS of the covariance matrix in the ABEKK and ADCC models.

[8]As already mentioned, the "square root" of the unconditional covariance matrix $\bar{\Sigma}$ is not unique. Consequently, the effect of z on $\varepsilon(z)$ is not unique and depends on the decomposition used to define $\bar{\Sigma}^{1/2}$. Remember that the elements of $\Sigma^{1/2}(z)$ are denoted $(\omega_{ij}(z))_{i,j=1,\cdots,n}$.

Proposition 2. *(1) For the asymmetric BEKK model, the NIS of the covariance matrix of unexpected returns is given by* $\Sigma(z) = \{\sigma_{ij}(z)\}$, *where*

$$\Sigma(z) = A_\Sigma + A'\varepsilon(z)\varepsilon(z)'A + G'\zeta(z)\zeta(z)'G, \tag{23}$$

where $A_\Sigma = \bar{\Omega} + B'\bar{\Sigma}B$ *and* $\zeta(z) = \varepsilon(z)1_{\{\varepsilon(z)\leq 0\}}$.
(2) For the asymmetric DCC model, the NIS of the covariance matrix of unexpected returns is given by $\Sigma(z) = \{\sigma_{ij}(z)\}$, *with* $\sigma_i(z)^2 = \sigma_{ii}(z)$ *and*

$$\sigma_i(z)^2 = \begin{cases} A_{\sigma,i} + (\alpha_i + \psi_i)\,\varepsilon_i(z)^2, & \text{if } z_i \leq 0, \\ A_{\sigma,i} + \alpha_i\varepsilon_i(z)^2, & \text{if } z_i > 0, \end{cases} \tag{24}$$

$$\sigma_{ij}(z) = \sigma_i(z)\sigma_j(z)\rho_{ij}(z), \tag{25}$$

where $A_{\sigma,i} = \omega_i + \beta_i\bar{\sigma}_i^2$. *The NIS of the conditional correlation* $\rho_{ij}(z)$ *is provided by Cappiello, Engle, and Sheppard (2006) in their Appendix A.2.*

From Equations (14) and (15), we deduce the following Proposition 3, which gives the NIS of co-skewness and co-kurtosis matrices:

Proposition 3. *(1) The NIS of third and fourth central moments of unexpected returns are given by*

$$\mathcal{S}^{\varepsilon}_{ijk}(z) = \sum_{r=1}^{n} \omega_{ir}(z)\omega_{jr}(z)\omega_{kr}(z)sk_r^Z(z),$$

$$\mathcal{K}^{\varepsilon}_{ijkl}(z) = \sum_{r=1}^{n} \omega_{ir}(z)\omega_{jr}(z)\omega_{kr}(z)\omega_{lr}(z)ku_r^Z(z) + \sum_{r=1}^{n}\sum_{s\neq r} \psi_{rs}(z),$$

where $sk_r^Z(z)$ *and* $ku_r^Z(z)$ *are defined in Equations (21) and (22), respectively, and*

$$\psi_{rs}(z) = \omega_{ir}(z)\omega_{jr}(z)\,\omega_{ks}(z)\omega_{ls}(z) + \omega_{ir}(z)\omega_{js}(z)\omega_{kr}(z)\omega_{ls}(z) \\ +\omega_{is}(z)\omega_{jr}(z)\omega_{kr}(z)\omega_{ls}(z).$$

(2) The NIS of co-skewness and co-kurtosis of unexpected returns are given by

$$sk^{\varepsilon}_{ijk}(z) = \frac{\mathcal{S}^{\varepsilon}_{ijk}(z)}{\sigma_i(z)\sigma_j(z)\sigma_k(z)},$$

$$ku^{\varepsilon}_{ijkl}(z) = \frac{\mathcal{K}^{\varepsilon}_{ijkl}(z)}{\sigma_i(z)\sigma_j(z)\sigma_k(z)\sigma_l(z)}.$$

In the empirical section, we focus on the response of unexpected returns' moments to shocks. This analysis combines the effect of shocks on the covariance matrix and the conditional distribution. The response of the covariance matrix combines the effect of the shocks on the variances and the correlations. Accordingly, the response of the third and fourth moments combines the effect of the shocks on the covariance matrix and the individual skewness and kurtosis.

3 DATA AND ESTIMATION RESULTS

3.1 Data

In the following, we consider the four largest international stock markets, namely the United States, Japan, the United Kingdom, and Germany. For all four countries, we use the reference market index over the period from January 1973 to December 2004, for a total of 8352 daily observations. Indices are the S&P500, the Nikkei 225, the FTSE-100, and the DAX 30, respectively. For the FTSE-100, we spliced the series with the "FTSE—all Shares" before the reference index was established. The return series, r_t, are defined as continuously compounded returns in US dollars. To account for the nonsynchronicity between the US and the other markets, the US market has been lagged one day.[9]

Table 1 displays several sample statistics on market returns. Concerning the higher moments, we notice a rather large dispersion in the magnitude of skewness and kurtosis across markets. All markets except Japan have a significantly negative skewness, meaning that crashes occur more often than booms. In addition, the high level of kurtosis (between 6 and 9) found for all markets is not consistent with the normality assumption. We then test the serial correlation in both returns and squared returns, using the Ljung–Box test and the Lee and King (1993) test, respectively. The latter allows for testing correlation in squared returns even in the presence of serial correlation in returns. Daily returns clearly display both serial correlation and heteroskedasticity.

3.2 Bivariate Model with Time-Varying Higher Moments

In this section, we focus on the estimation of the ADCC model with Sk-t distribution and time-varying shape parameters. The model is defined by Equations (1)–(4), (6)–(8), and (16)–(17). We consider all the pairs involving the US market in combination with one of the other markets. All the parameters of a given bivariate model are estimated simultaneously using the ML technique described in Section 1.4.[10]

Table 2 reports the estimates of the four bivariate models when an AR(1) term is used for the conditional means. As expected, the conditional mean equation displays little serial correlation. Only the first lagged return gives a small, though for some markets significant, autoregressive parameter φ_1.

We tested the validity of the model using the robust conditional moment test procedure, which aims at detecting whether the model fails to capture particular features observed in the data (Wooldridge 1990, 1991). Because our model is designed to capture the dynamics of the first four moments, the moment conditions

[9]Because the October 1987 crash may have a dramatic effect on the shape of the distribution and on the dynamics of the higher moments, we also investigated the consequences of eliminating this observation. The results of this estimation are reported in the Technical Appendix.

[10]The estimation of the model including the four markets is reported and discussed in the Technical Appendix. The main difference with the bivariate models is that the parameters of the conditional correlation are assumed to be the same across the pairs of markets. The results are reported in the Technical Appendix.

Table 1 Summary statistics on market returns

	United States	Japan	UK	Germany
Moments				
Mean	0.030	0.024	0.030	0.036
Std dev.	0.993	1.369	1.167	1.308
Minimum	−8.642	−8.241	−8.417	−13.058
Maximum	5.573	12.571	9.048	9.332
Skewness	−0.190[a]	0.190[a]	−0.071[a]	−0.270[a]
Kurtosis	6.991[a]	7.759[a]	6.578[a]	8.103[a]
JB test	5399.5[a]	7795.4[a]	4356.0[a]	9024.2[a]
Serial correlation				
LB(5)	45.746[a]	7.330	156.492[a]	14.341[b]
LK(5)	25.246[a]	22.153[a]	33.159[a]	24.427[a]
Correlation matrix				
United States	1	0.224	0.234	0.273
Japan	0.224	1	0.102	0.131
UK	0.234	0.229	1	0.452
Germany	0.273	0.131	0.452	1

This table reports summary statistics on stock-market returns, sampled at daily frequency from January 1973 until December 2004, for a total of 8352 observations. Tests for the null hypotheses that the skewness and excess kurtosis are equal to zero are based on the asymptotic distribution under normality. The Jarque–Bera statistics is denoted by JB. The Ljung–Box statistics for serial correlation, corrected for heteroskedasticity, computed with 5 lags is denoted LB(5). Under the null of no serial correlation, it is distributed as a $\chi^2(5)$. The Lee and King (1993) statistics for heteroskedasticity is denoted by LK(5). Under the null of no serial correlation of squared returns, the test statistics are distributed as a $\chi^2(5)$.
A1 Superscripts [a] and [b] indicate that a statistic is significant at the 1% and 5% level, respectively.

rely on these four moments. As this test is rather well known, we relegate the description of how we implemented it and the discussion of the main results to the Technical Appendix. Overall, only eight moment conditions are rejected out of 164, suggesting that our specification provides a good description of the data. Four rejections are in fact due to the model's difficulty in capturing the asymmetry of the conditional volatility.

3.2.1 Dynamics of the covariance matrix. As Table 2 shows, the correlation dynamics are strongly persistent (δ_1 is close to 0.98), and the parameters δ_2 are all significantly positive. Although it is not clear from Equations (7) and (8), because δ_2 plays a role in the correlation dynamics through the numerator and the denominator, a positive value of δ_2 indicates that the correlation increases when the markets are simultaneously affected by shocks of the same sign and decreases when they are affected by shocks of opposite sign. This result is confirmed by the subsequent NIS of the conditional correlation matrix (displayed in Figure 4) and is consistent

Table 2 Parameter estimates of the ADCC model with Sk-t distribution and time-varying shape parameters

	United States		Japan		UK		Germany	
	Parameter	Standard error	Parameter	Standard error	Parameter	Standard error	Parameter	Standard error
Conditional mean								
μ	0.029	(0.012)	0.031	(0.011)	0.028	(0.006)	0.036	(0.012)
φ_1	0.062	(0.012)	0.022	(0.010)	0.067	(0.012)	−0.019	(0.018)
Conditional variance								
ω	0.008	(0.002)	0.020	(0.003)	0.022	(0.004)	0.020	(0.003)
α	0.034	(0.006)	0.037	(0.006)	0.049	(0.007)	0.045	(0.007)
ψ	0.055	(0.008)	0.076	(0.009)	0.039	(0.008)	0.039	(0.008)
β	0.937	(0.005)	0.911	(0.007)	0.909	(0.009)	0.917	(0.007)
Conditional correlation								
δ_1	-		0.9831	(0.0053)	0.9817	(0.0099)	0.9838	(0.0039)
δ_2	-		0.0084	(0.0024)	0.0042	(0.0021)	0.0101	(0.0021)
δ_3	-		0.0001	(0.0019)	0.0015	(0.0022)	0.0001	(0.0018)
Conditional degree of freedom								
$c_0(/100)$	0.040	(0.008)	0.038	(0.009)	0.019	(0.001)	0.165	(0.049)
c_1^-	−0.724	(0.268)	−0.298	(0.034)	0.573	(0.288)	−0.492	(0.174)
c_1^+	−2.374	(0.431)	−0.469	(0.956)	−0.303	(0.312)	−0.574	(0.474)
c_2	0.548	(0.117)	−0.324	(0.262)	0.577	(0.080)	−0.486	(0.140)
Conditional asymmetry								
d_0	0.969	(0.009)	1.005	(0.014)	0.985	(0.009)	0.989	(0.011)
d_1^-	0.030	(0.036)	0.057	(0.021)	0.025	(0.012)	0.028	(0.014)
d_1^+	0.109	(0.023)	0.047	(0.018)	0.036	(0.013)	0.041	(0.019)
d_2	0.615	(0.089)	0.467	(0.192)	0.841	(0.076)	0.788	(0.070)
lnL	-		23472.5		22407.1		23219.9	

This table reports parameter estimates for the multivariate ADCC model with a Sk-t distribution and time-varying higher moments. Figures in parentheses represent standard errors. All bivariate pairs involve the US market. Estimates for the US correspond to the model for the US–Japan pair. The log-likelihood of the sample is denoted lnL. Standard errors are computed using the inverse of the Hessian.

with a number of papers that document that the correlation increases after shocks of the same sign (Ang and Chen 2002; Cappiello, Engle, and Sheppard 2006).

We also find that the asymmetry parameter δ_3 in the DCC model is insignificant in our estimations. At first glance, this finding appears at odds with the evidence reported in Cappiello, Engle, and Sheppard (2006). Their significant estimate of δ_3 suggests that correlation increases more after a co-crash than after a co-boom. In fact, our estimates also indicate that subsequent co-crashes are more likely to occur after a co-crash. Yet, in our model, they are not due to the increase in the conditional correlation, but rather to the persistence in higher moments: a given co-crash induces a more negative skewness, so the probability of occurrence of other co-crashes increases. We conclude that the asymmetry present in returns is rather due to the individual distribution than to the dynamics of the covariance matrix.

3.2.2 Dynamics of the higher moments. Most parameters pertaining to the dynamics of shape parameters are statistically significant. Beginning with the degree of freedom $\nu_{i,t}$, we notice that almost all parameters $c_{i,1}^-$ and $c_{i,1}^+$ are negative. Therefore, independently of its sign, a large shock generates a subsequent distribution with fatter tails. We also notice that $|c_{i,1}^-| < |c_{i,1}^+|$ in the US, suggesting that a positive shock increases the distribution's tails more so than a negative shock does. Turning to the skewness parameter $\xi_{i,t}$, we observe that parameters $d_{i,1}^-$ and $d_{i,1}^+$ are positive for all markets. As a consequence, a large negative shock decreases the subsequent skewness and, therefore, increases the probability of another large negative shock in the subsequent period. Similarly, a large positive shock also tends to increase the subsequent skewness and, therefore, to increase the probability of another large positive shock in the next period.

The big picture that emerges from the parameter estimates is that a shock of a given sign is often followed by another shock with the same sign. The estimated model incorporates two different features of asymmetry that reinforce each other: First, the lagged unexpected return ε_{t-1} has a different effect on the subsequent *volatility* depending on its own sign. Second, the lagged innovation z_{t-1} affects the asymmetry of the subsequent *conditional distribution*. These two features are complementary. The asymmetry in volatility indicates that the distribution of returns will be more dispersed after a negative shock. It does not predict the actual shape of the distribution. This shape is determined by changes in the degree of freedom and the asymmetry parameter, which are themselves functions of past shocks. After a negative shock, the conditional distribution is more negatively skewed. This effect is independent of the level of volatility.

Deducing the behavior of higher moments by contemplating parameter estimates is a difficult task because skewness and kurtosis are jointly related to shape parameters in a highly nonlinear way. In Figure 2, we display the dynamics of the asymmetry parameter, the degree of freedom, and the conditional skewness and kurtosis for the US market return. The dynamics of the shape parameters and the higher moments are smoothed over 4 weeks using a simple moving average. We observe a similar evolution for the other markets. As it appears in the figure,

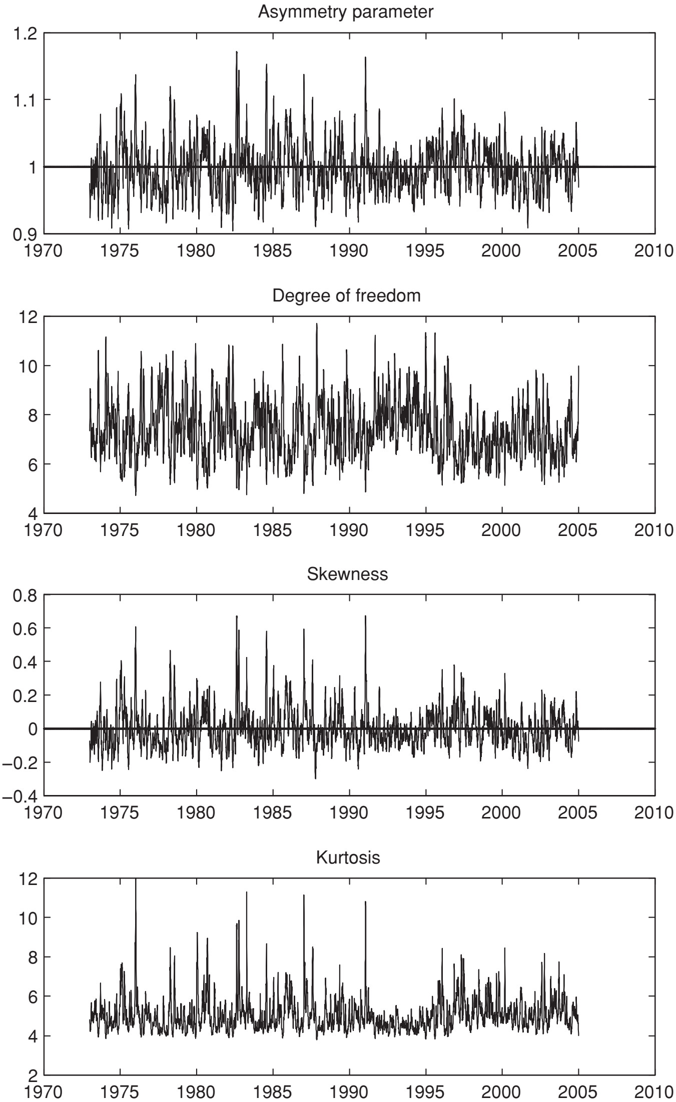

Figure 2 This figure displays the evolution of the asymmetry parameter, the degree of freedom, and the conditional skewness and kurtosis for the US return, resulting from the parameter estimates reported in Table 2. The series are smoothed over 4 weeks using a simple moving average.

the asymmetry parameter and the conditional skewness display similar patterns, while the degree of freedom and the conditional kurtosis display patterns inverse of each other. Conditional skewness and kurtosis take rather reasonable values and vary substantially through time. Over the last 10 years of the sample, changes in skewness and kurtosis tend to be less erratic than over the beginning of the sample. Trends also appear more pronounced. In particular, the skewness of the US return increased between 1994 and 1998 (from −0.2 to 0.3) and then dropped to −0.2 in 2002. This trend may reflect the dynamics of the markets during the development and burst of the Internet bubble.

4 BEHAVIOR OF THE JOINT DISTRIBUTION

As argued before, the behavior of the joint distribution of unexpected returns is determined by the properties of the covariance matrix and of the conditional distribution of innovations. We begin with a discussion of the properties of these two components and then turn to the properties of the higher co-moments between unexpected returns.

4.1 Response of Individual Higher Moments to Shocks

Figure 3 displays the NIC of the asymmetry parameter, the degree of freedom, and the conditional skewness and kurtosis for the US and Japanese daily innovations. We first notice that a shock of a given sign is followed by a subsequent skewness of the same sign. This pattern reveals that the probability of another, subsequent shock with the same sign increases and, therefore, that large shocks of a given sign tend to cluster. As already mentioned for the US market, the response of the skewness is asymmetric, because skewness increases more after a large positive shock than after a large negative shock. This suggests that large positive shocks are more likely to be serially correlated than large negative shocks.

Regarding the NIC of conditional kurtosis, we observe a U-shape pattern for the Japanese market, indicating that kurtosis increases after large shocks of either sign. This means that after a first large shock, the probability of occurrence of another large (negative or positive) shock increases. We obtain a J-shape pattern for the US market, revealing that the increase in the kurtosis in the case of a positive shock is much more pronounced than for a negative shock.

The asymmetry observed in the response of skewness and kurtosis after a negative or positive shock appears to be a feature of the US market. These patterns are confirmed by simply estimating the skewness and kurtosis of returns after a large shock. For our sample, we computed the skewness and kurtosis only for realizations that followed a large (negative or positive) return (beyond one standard deviation). We calculated that the US skewness is significantly lower after a large negative return than after a large positive return (−0.3 and 0.8, respectively). In addition, the kurtosis is significantly lower after a large negative return than after a large positive return (6.6 and 9, respectively). We do not observe such an asymmetry for the Japanese market.

It is worth noticing that these results are partly related to the frequency of the data. There are some indications that the effects on skewness and kurtosis are short-lived. First the analysis of the impulse response functions performed in Section 5 indicates that the initial response is offset after a few days. In addition,

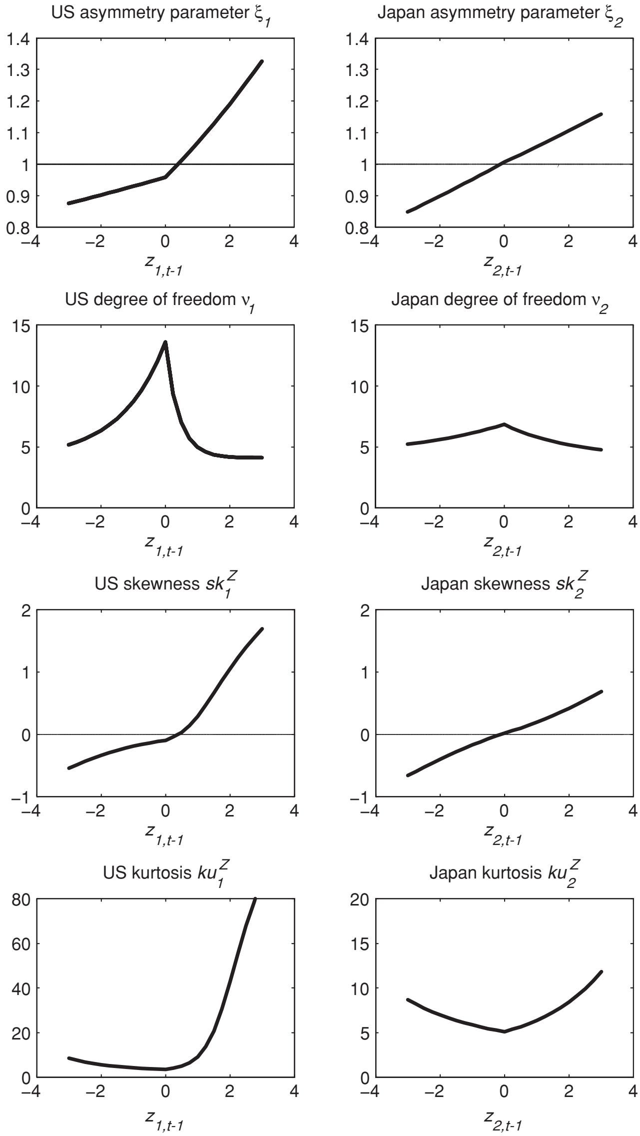

Figure 3 This figure displays the news impact curves of the asymmetry parameter, the degree of freedom, and the conditional skewness and kurtosis for the US and Japanese returns.

re-estimating the model on weekly data shows that a shock of a given sign is generally followed by a skewness of the opposite sign (see Jondeau and Rockinger 2008). This additional evidence suggests that the response of higher moments to shocks can be interpreted as an over-reaction phenomenon.

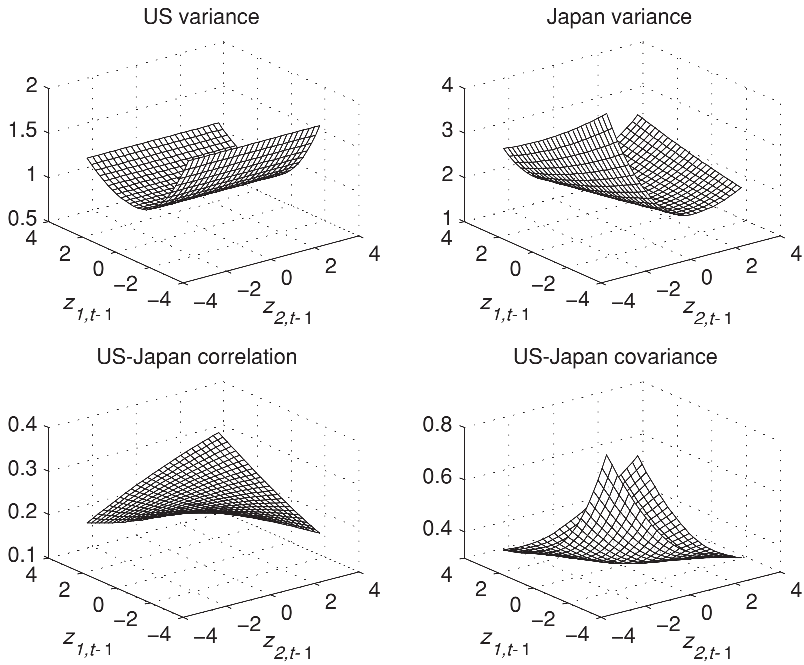

Figure 4 These figures display the news impact surfaces of the second moments for the US and Japanese returns.

4.2 Response of the Covariance Matrix to Shocks

Figure 4 displays the NIS of the covariance matrix between US and Japanese daily returns. Given that the US market closes first, we use a Cholesky decomposition of the covariance matrix, which implies that the US unexpected return cannot be affected by a shock on the Japanese market. Therefore, $z_{1,t-1}$ denotes the shock on the US market return, while $z_{2,t-1}$ denotes the component of the shock on the Japanese market return that is orthogonal to $z_{1,t-1}$. The NIS of the US variance provides the same amount of information as the standard NIC of the variance in a univariate model. In contrast, the NIS of the Japanese variance provides some interesting insight on the effect of a shock on both markets. In particular, the figure reveals that the response of the Japanese variance to a negative Japanese shock is much more pronounced if there is also a negative shock on the US market. In case of a positive Japanese shock, the response of the Japanese variance is likewise more pronounced if there is also a positive shock on the US market.

Inspecting the correlation surface reveals that the correlation increases after large shocks of the same sign, whatever the sign but decreases after large shocks of opposite signs. Now, if we consider the effect on the covariance, we observe that covariance increases after large shocks regardless of sign. The reason is that the shape of the NIC of variances dominates. Eventually, if the covariance is at its average level of 0.3 and if we also assume a shock with a magnitude of 3 standard deviations for both markets, then the subsequent covariance is equal to 0.87 if the two shocks are negative, 0.57 if the two shocks are positive, but only 0.34 if the two shocks are of opposite signs. We observe similar patterns for all other pairs of market returns.[11]

4.3 Response of the Higher Co-moments to Shocks

We now describe how changes in the correlation matrix and the conditional distribution combine to affect the higher moments of the joint distribution of unexpected returns. Figures 5 and 6 display the NIS of the third and fourth central moments for the US–Japan pair.

We begin with the individual third central moments $S^{\varepsilon}_{iii}(z)$, which are closely related to the standard measure of skewness. The response of the US third central moment $S^{\varepsilon}_{111}(z)$ was already discussed in Section 4.1. For other countries, the shape of the third moment is affected by the sign and size of the US shock because the covariance between the two markets intervenes in the computation of $S^{\varepsilon}_{222}(z)$. In particular, for a given shock on the Japanese return, the subsequent third moment is higher, the larger the shock on the US market. For instance, for $z_2 = -3$, the Japanese third moment $S^{\varepsilon}_{222}(z)$ is equal to -5.6 for $z_1 = -3$ and to -2.7 for $z_1 = 3$. This effect suggests that the probability of another large negative event is higher in the case of a simultaneous crash on the two markets.

We now turn to the third central co-moments of the form $S^{\varepsilon}_{iij}(z)$.[12] They indicate if the market in a given country j provides a good hedge against adverse volatility changes in country i. A positive value implies that the return in country j also goes up if the volatility in country i goes up, thereby providing a good hedge against volatility. As Figure 5 shows, co-moments $S^{\varepsilon}_{112}(z)$ and $S^{\varepsilon}_{122}(z)$ are higher when the shocks on the two markets are higher: After two large negative shocks (-3 for the two markets), the subsequent moment $S^{\varepsilon}_{112}(z)$ is equal to -0.63, while it is as high as 1.07 after two large positive shocks. Our figures document that, subsequent to negative shocks, the two markets are bad hedges against volatility in the other country. Inversely, subsequent to positive shocks, the two markets are good hedges against volatility in the other country.

The NISs of the fourth central moments are depicted in Figure 6. The individual fourth moments $K^{\varepsilon}_{iiii}(z)$ display the following pattern: In all markets, a large (negative or positive) domestic shock implies an increase in the subsequent fourth moment. The north-east surface in the figure shows that the tail fatness of the Japanese return is also strongly affected by the US shock: The response of $K^{\varepsilon}_{2222}(z)$ to two large negative shocks is 155.2, while the response to a negative Japanese

[11] To test if these responses are significantly different from each other, we used Monte Carlo simulations to evaluate their finite-sample distribution. For this purpose, we followed the approach developed by Koop, Pesaran, and Potter (1996), which will be briefly described in Section 5. We found that the responses reported in this and the following section are quite precisely estimated. We reject the null hypothesis that the responses are equal to each other for different sets of shocks.

[12] Because we consider bivariate models, terms such as $S^{\varepsilon}_{ijk}(z)$ do not appear. Similarly, terms such as $K^{\varepsilon}_{iijk}(z)$ and $K^{\varepsilon}_{ijkl}(z)$ do not appear.

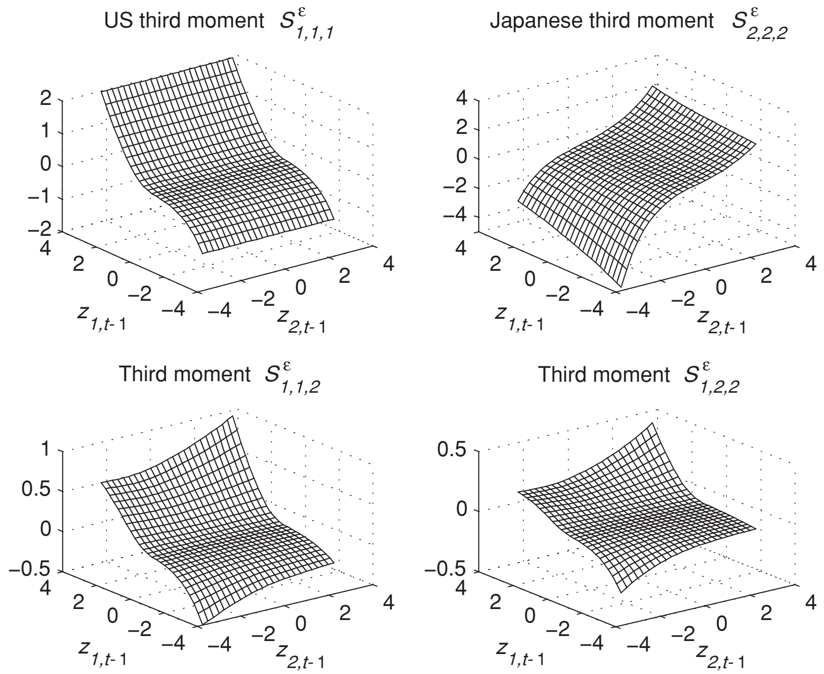

Figure 5 These figures display the news impact surfaces of the third moments for the US and Japanese returns.

shock and a positive US shock is only 47.2. Similarly, the response to two large positive shocks is 77.5, while the response to a positive Japanese shock and a negative US shock is only 60.8. Therefore, the US shock tends to reinforce the dominant effect of the domestic shock.

The fourth co-moment $K^{\varepsilon}_{1122}(z)$ can be interpreted as a measure of the strength between the two variances. A large value of $K^{\varepsilon}_{1122}(z)$ indicates that the volatilities in both markets move together, so both markets provide a bad hedge against high volatility in the other market. The NIS reveals that, subsequent to large shocks with the same sign in the two countries, the variances are much more closely related. This result means that diversification provides the worst hedge during periods of high volatility in both markets. Such a phenomenon relates to the evidence provided by Ang and Bekaert (2002).

Last, we turn to fourth co-moments of the form $K^{\varepsilon}_{iiij}(z)$. Intuitively, a large value of this measure implies that the distribution of market i becomes more negatively skewed when the return in market j is lower than expected. Therefore, market i would be a bad hedge against a fall in market j. We determine that $K^{\varepsilon}_{1112}(z)$ as well as $K^{\varepsilon}_{1222}(z)$ increase after a (positive or negative) US shock. However, after a negative US shock, they decrease with the Japanese shock, while after a positive US shock, they increase with the Japanese shock. This result suggests that when two large positive shocks or two large negative shocks occur in both markets, the

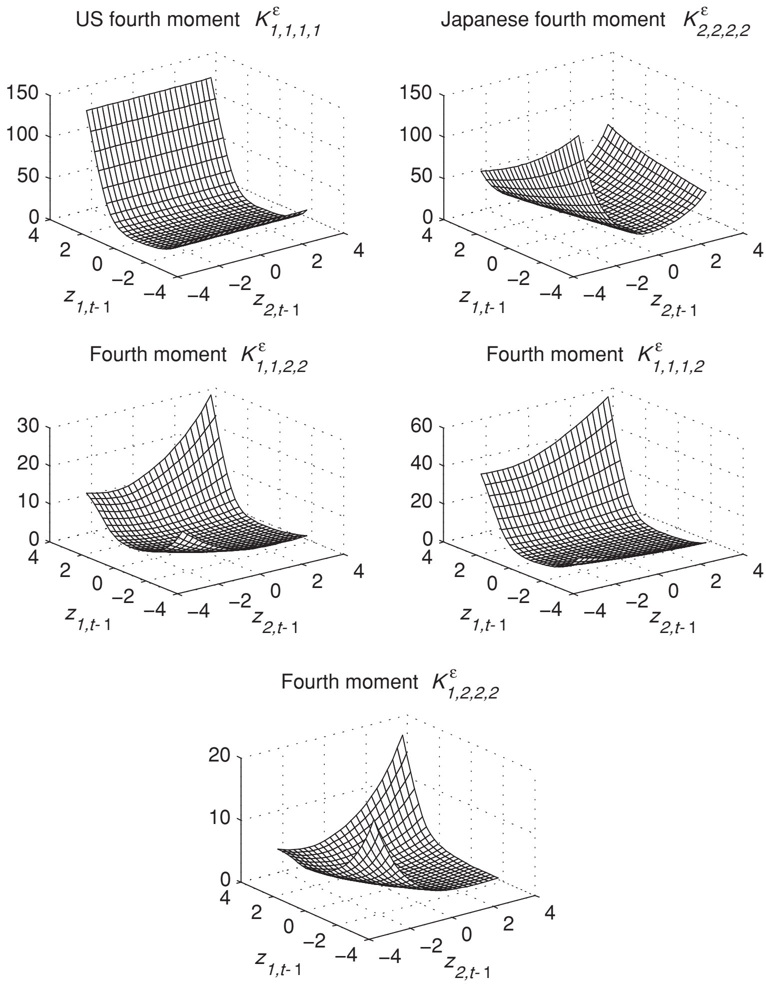

Figure 6 These figures display the news impact surfaces of the fourth moments for the US and Japanese returns.

subsequent skewness of a given market (say, Japan) is more correlated with the shock on the other market (say, the US). For example, after a crash on both markets, the subsequent distribution of the Japanese market will lean either to the left if the US return is negative or to the right if the US return is positive. As a consequence, the likelihood of another event with the same sign on both markets increases.

5 IMPULSE RESPONSE FUNCTIONS

To gain further insight about the behavior of the conditional distribution after a shock, we extend the analysis of the NIC and NIS to impulse response functions. While the NIC and NIS indicate how moments of returns are affected during the period just after a given shock, we consider now how these moments vary over time after a shock. As is obvious from Section 1, the relations between an innovation $z_{i,t}$ and the various higher moments are highly nonlinear. Therefore, it is rather difficult to analytically compute impulse response functions.[13] Instead, we adopt and generalize the approach developed by Koop, Pesaran, and Potter (1996), who introduce the concept of generalized impulse response (GIR) in the case of nonlinear models. In our context, the GIR is the difference between the expected value of returns after a shock v_t at date t and the expected value of returns without this shock, for a given history. It is defined for a given horizon h as

$$GIR(h, v_t, \omega_{t-1}) = E[r_{t+h}|v_t, \omega_{t-1}] - E[r_{t+h}|\omega_{t-1}], \quad h = 0, 1, \ldots, \tag{26}$$

where $\omega_{t-1} = \{z_{t-1}, z_{t-2}, \ldots\}$ denotes the history of shocks.

In a nonlinear model, defining the GIR raises several difficulties. First, the shock experiment has to be carefully designed because the GIR may be highly affected by the level of shocks and their possible correlation. Second, the GIR is conditional on the history because the realization of past shocks is likely to affect the subsequent trajectory. Given the complexity of solving some nonlinear models, Koop, Pesaran, and Potter (1996) recommend the use of Monte Carlo integration to compute the conditional expectations in Equation (26). The main advantage of this simulation approach in our context is that the GIR for the covariance matrix and the higher moments can be very easily computed because they are explicitly modeled as functions of past innovations z_{t-1}. Therefore, after a given shock, v_t, it is possible to recover the subsequent variances, covariances, skewness, and kurtosis in a way similar to Equation (26). A precise description of our simulation procedure is provided in the Technical Appendix.

Figures 7 and 8 display the GIR for the covariance matrix and the higher co-moments of the conditional distribution of unexpected returns for the US and Japanese markets, respectively. We study the persistence of a shock using a 30-day window. In the following, we comment on two experiments, corresponding to a negative and to a positive shock on the US market. If we denote $v_t = (v_{1,t}, v_{2,t})'$ the vector of shocks, the shocks are $(-2\%, 0)$ and $(2\%, 0)$, respectively.

We begin with a negative shock on the US market (Figure 7). It is well known that the variance and covariance decrease only very slowly after a shock, given the high persistence found in GARCH models (see Hafner and Herwatz 2006). A -2% shock on the US market implies an increase of 0.2 of the US variance the day after the shock. After 30 days, the variance is still 0.15 higher than it would have been without the shock. About 100 days are necessary for the variance to return to its average level.

In contrast, the response of the higher moments to a shock is short-lasting. A -2% shock induces a decrease in the third central moment of the US return of

[13] This analytical computation has been partly performed, for instance, by Hafner and Herwartz (2006) for the volatility impulse responses in a multivariate GARCH model.

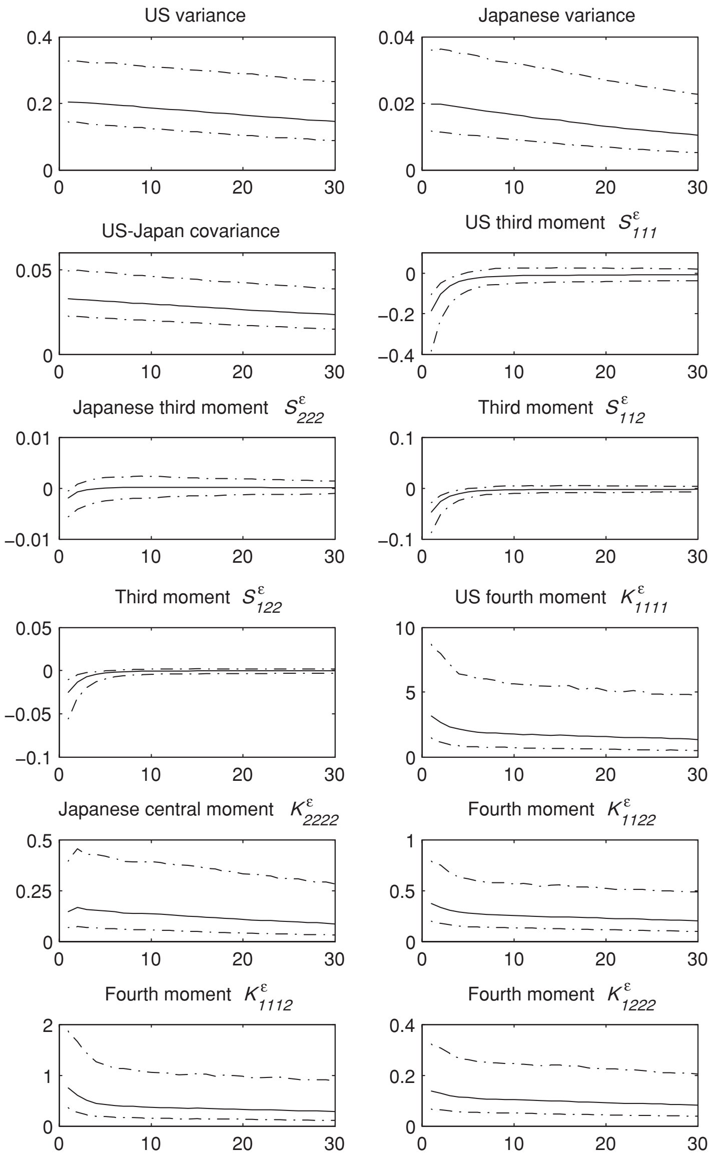

Figure 7 These figures display the generalized impulse response functions for the US and Japanese daily returns, after a −2% shock on the US return. The reported confidence interval corresponds to ± one standard deviation.

−0.2. This is a sizeable effect because it is the order of magnitude of the unconditional third moment. However, the effect is offset after 5 days. We observe a similar behavior in the other third central moments. Regarding fourth central moments, we notice a sharp increase of 4 in the US fourth moment $K^{\varepsilon}_{1111,t}$ (while the

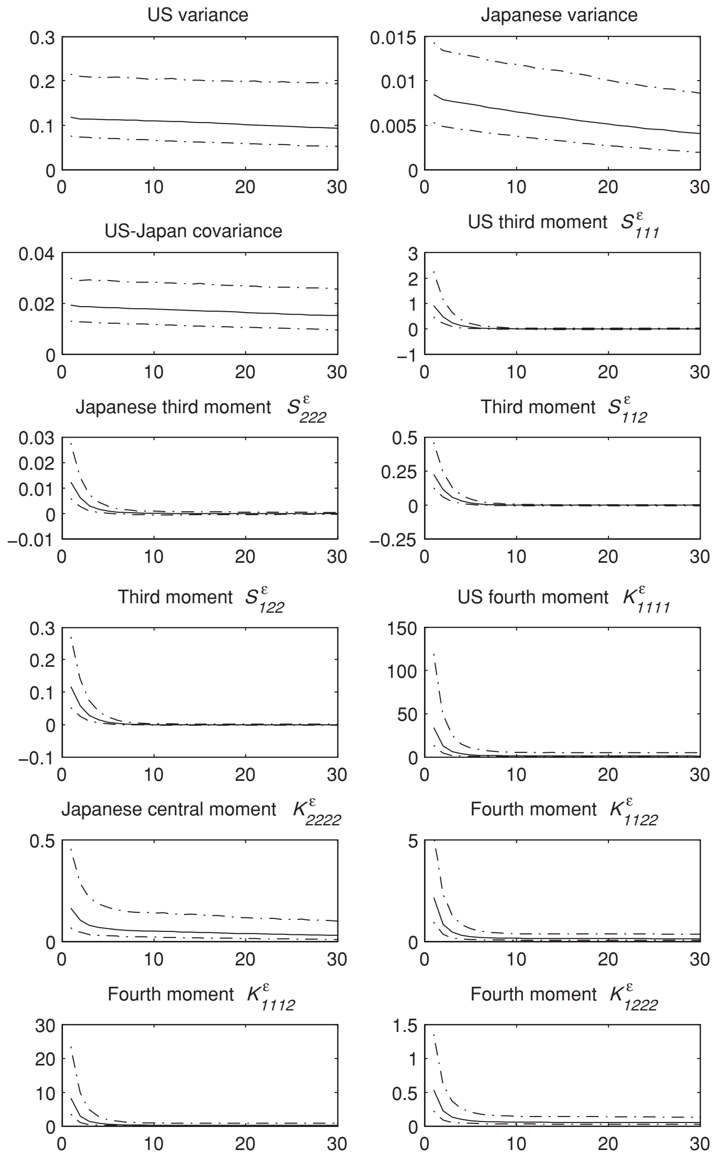

Figure 8 These figures display the generalized impulse response for the US and Japanese daily returns, after a +2% shock on the US return. The reported confidence interval corresponds to ± one standard deviation.

unconditional moment is 7). Similar though less pronounced patterns are observed for the other fourth central moments. Therefore, the contagion effects documented in the existing literature, for example, Longin and Solnik (2001) or Poon, Rockinger, and Tawn (2004), are short-lived contemporaneous effects.

If we consider a positive shock on the US return (Figure 8), we notice three main changes as compared to a negative shock. First, as expected, the magnitude of the response of the covariance matrix is smaller than after a negative shock, reflecting the well-known asymmetric effect of shocks on volatility. Second, the response of third moments is now positive, suggesting that the multivariate distribution tends to lean to the right side. Finally, the responses of the third and fourth moments are significantly larger. This final result was expected from the parameter estimates, but the figures provide a tool for measuring the temporal evolution of these higher moments. Typically, a 2% shock on the US market implies an increase of 0.5 and 20 of the US third and fourth moments, respectively. These numbers must be compared with −0.2 and 4, respectively, in the case of a −2% shock.

6 CONCLUSION

In this paper, we propose two methodological contributions. Our first contribution is a multivariate model for asset returns, in which shocks have a feedback effect not only on the covariance matrix but also on the higher moments (co-skewness and co-kurtosis) of the joint distribution. For this purpose, we extend the DCC model of Engle (2002) to the case of innovations drawn from a Sk-*t* distribution with shape parameters that are function of past shocks. We show that this model fits daily international stock market returns very well.

Our second contribution is the design of a graphical tool that extends the concept of the NIC to the shape of the distribution. In a univariate setting, this leads us to the NIC of skewness and kurtosis. In a multivariate setting, we obtain the NIS of the various co-moments, thus allowing a better characterization of the joint distribution of returns. We find that a large shock is likely to be followed by another large shock because it results in an increase of the subsequent kurtosis. In addition, a large shock of a given sign generally results in a subsequent skewness of the same sign, so it is likely to be followed by another large shock of the same sign. In a multivariate conditional setting, we establish some stylized features. In particular, we document that past foreign shocks have little impact on the current co-skewness or co-kurtosis beyond that information contained in US past shocks. Finally, we investigate the temporal evolution of the response of the higher moments and co-moments to shocks. This impulse response analysis reveals that in all cases the effects are significant but short-lasting.

The non-normality and time-variability of the multivariate distribution of asset returns are very likely to have dramatic consequences from risk and portfolio management perspectives. It is well recognized that the value-at-risk of a portfolio has to be computed dynamically to account for changes in the distribution of returns. Typically, after a shock on market returns, a bank has to reevaluate the VaR of its portfolio accordingly. Also such changes affect the distributional properties of portfolio returns. A risk-averse investor would probably reallocate her portfolio on the basis of the distributional properties and, in particular, on the basis of the recent characteristics of the asset returns' distribution.

Received November 6, 2007; revised September 3, 2008; accepted October 16, 2008.

REFERENCES

Ang, A., and G. Bekaert. 2002. International Asset Allocation with Regime Shifts. *Review of Financial Studies* 15: 1137–1387.

Ang, A., and J. Chen. 2002. Asymmetric Correlations of Equity Portfolios. *Journal of Financial Economics* 63: 443–494.

Bauwens, L., and S. Laurent. 2005. A New Class of Multivariate Skew Densities, with Application to GARCH Models. *Journal of Business and Economic Statistics* 23: 346–354.

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. *Review of Economics and Statistics* 69: 542–547.

Bollerslev T., R. Chou, and K. F. Kroner. 1992. ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics* 52: 5–59.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge. 1988. A Capital Asset Pricing Model with Time-varying Covariance. *Journal of Political Economy* 96: 116–131.

Bollerslev, T., and J. M. Wooldridge. 1992. Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances. *Econometric Reviews* 11: 143–172.

Brooks, C., S. P. Burke, S. Heravi, and G. Persand. 2005. Autoregressive Conditional Kurtosis. *Journal of Financial Econometrics* 3: 399–421.

Cappiello, L., R. F. Engle, and K. Sheppard. 2006. Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns. *Journal of Financial Econometrics* 4: 537–572.

Engle, R. F. 1982. Auto-Regressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* 50: 987–1007.

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., and K. F. Kroner. 1995. Multivariate Simultaneous GARCH. *Econometric Theory*11: 122–150.

Engle, R. F., and V. K. Ng. 1993. Measuring and Testing the Impact of News on Volatility. *Journal of Finance* 48: 1749–1778.

Fernández, C., and M. F. J. Steel. 1998. On Bayesian Modelling of Fat Tails and Skewness. *Journal of the American Statistical Association* 93: 359–371.

Fiorentini, G., E. Sentana, and G. Calzolari. 2003. Maximum Likelihood Estimation and Inference in Multivariate Conditionally Heteroscedastic Dynamic Regression Models with Student t Innovations. *Journal of Business and Economic Statistics* 21: 532–546.

Hafner, C. M., and H. Herwatz. 2006. Volatility Impulse Response Functions for Multivariate GARCH Models: An Exchange Rate Illustration. *Journal of International Money and Finance* 25: 719–740.

Hansen, B. E. 1994. Autoregressive Conditional Density Estimation. *International Economic Review* 35: 705–730.

Harvey, A. C., E. Ruiz, and E. Sentana. 1992. Unobservable Component Time Series Models with ARCH Disturbances. *Journal of Econometrics* 52: 129–158.

Harvey, C. R., and A. Siddique. 1999. Autoregressive Conditional Skewness. *Journal of Financial and Quantitative Analysis* 34: 465–487.

Jondeau, E., and M. Rockinger. 2002. Entropy Densities with an Application to Autoregressive Conditional Skewness and Kurtosis. *Journal of Econometrics* 106: 119–142.

Jondeau, E., and M. Rockinger. 2003. Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements. *Journal of Economic Dynamics and Control* 27: 1699–1737.

Jondeau, E., and M. Rockinger. 2008. Time Variability in Higher Moments is Important for Asset Allocation. Swiss Finance Institute Research Paper.

Koop, G., M. H. Pesaran, and S. M. Potter. 1996. Impulse Response Analysis in Nonlinear Multivariate Models. *Journal of Econometrics* 74: 119–147.

Kroner, K. F., and V. K. Ng. 1998. Modeling Asymmetric Comovements of Asset Returns. *Review of Financial Studies* 11: 817–844.

Lee, J. H. H., and M. L. King. 1993. A Locally Most Mean Powerful Based Score Test for ARCH and GARCH Regression Distrubances. *Journal of Business and Economic Statistics* 11: 17–27.

Longin, F. M., and B. Solnik. 2001. Extreme Correlation of International Equity Markets. *Journal of Finance* 56: 649–676.

Mencia, J., and E. Sentana. 2005. Estimation and Testing of Dynamic Models with Generalized Hyperbolic Innovations. CEPR Discussion Paper No. 5177.

Poon, S. H., M. Rockinger, and J. Tawn. 2004. Extreme Value Dependence in Financial Markets: Diagnostics, Models, and Financial Implications. *Review of Financial Studies* 17: 581–610.

Sahu, S. K., D. K. Dey, and M. D. Branco. 2003. A New Class of Multivariate Skew Distributions with Applications to Bayesian Regression Models. *Canadian Journal of Statistics* 31: 129–150.

Vlaar, P. J. G., and F. C. Palm. 1993. The Message in Weekly Exchange Rates in the European Monetary System: Mean Reversion, Conditional Heteroskedasticity and Jumps. *Journal of Business and Economic Statistics* 11: 351–360.

Wooldridge, J. M. 1990. A Unified Approach to Robust, Regression-Based Specification Tests. *Econometric Theory* 6: 17–43.

Wooldridge, J. M. 1991. On the Application of Robust, Regression-Based Diagnostics to Models of Conditional Means and Conditional Variances. *Journal of Econometrics* 47: 5–46.