The Impact of News on Higher Moments

Eric Jondeau   Michael Rockinger

First version: November 2006
Current version: November 2006

This research has been carried out within the NCCR FINRISK project on “Financial Econometrics for Risk Management”
The Impact of News on Higher Moments

Eric Jondeau\textsuperscript{a} and Michael Rockinger\textsuperscript{b}

November 2006

Abstract

In this paper, we extend the concept of News Impact Curve developed by Engle and Ng (1993) to the higher moments of the multivariate returns’ distribution, thereby providing a tool to investigate the impact of shocks on the characteristics of the subsequent distribution. For this purpose, we present a new methodology to describe the joint distribution of returns in a non-normal setting. This methodology allows to gain a better understanding of the temporal evolution of the returns’ distribution and can be used to analyze the behavior of the optimal portfolio distribution. We apply our methodology to provide stylized facts on the four largest international stock markets. In particular, we document the persistence in 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.

Keywords: Volatility, Skewness, Kurtosis, GARCH model, Multivariate skewed Student t distribution, Stock returns.

JEL classification: C22, C51, G12.

\textsuperscript{a}Corresponding author. Swiss Finance Institute and HEC Lausanne. University of Lausanne, Institute of Banking and Finance, CH 1015 Lausanne, Switzerland. E-mail: eric.jondeau@unil.ch.

\textsuperscript{b}Swiss Finance Institute, HEC Lausanne and CEPR. University of Lausanne, Institute of Banking and Finance, CH 1015 Lausanne, Switzerland. E-mail: michael.rockinger@unil.ch.

Both authors acknowledge help from the Swiss National Science Foundation through NCCR (Financial Valuation and Risk Management). We benefited from comments of and discussions with Tim Bollerslev, Ales Cerny, René Garcia, Eric Jacquier, Nour Meddahi, and Raman Uppal. The usual disclaimer applies.
1 Introduction

The analysis of the distributional and dynamic properties of asset returns is a very active area of research in theoretical as well as empirical finance. First, abundant empirical evidence has shown that asset returns are non-normal, observing that extreme returns occur too often to be consistent with normality and that crashes are found to occur more often than booms.\(^1\) This evidence suggests that returns are driven by asymmetric and fat-tailed distributions. Second, asset returns have 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, as described by the so-called GARCH models.\(^2\) Engle (1982) has shown in this context that the non-normality of returns and the time-variability of volatility are two related phenomenons, because the latter contributes to the former. While it has been found to be sufficient for restoring the normality of the innovation process at low frequency, its distribution is still non-normal at daily frequency. In the quest of the best suited conditional distribution, several distributions have been proposed: the \(t\) distribution (Bollerslev, 1987), the entropy distribution (Jondeau and Rockinger, 2002), various forms of asymmetric \(t\) distributions (Hansen, 1994, Harvey and Siddique, 1999, Jondeau and Rockinger, 2003). In addition, several recent papers have provided evidence that the characteristics of this conditional distribution vary over time.\(^3\) This finding implies that the probability of a return of a given size depends on recent events.

Regarding the modeling of the joint behavior of asset returns, most of the recent contributions focused on the dynamics of the conditional covariance (as in first-


generation multivariate GARCH models, see Bollerslev, Engle, and Wooldridge, 1988, and Engle and Kroner, 1995) or on the dynamics of the conditional correlation (as in the more recent DCC models, see Engle, 2002, Cappiello, Engle, and Sheppard, 2003). In this multivariate setting, some authors have explicitly introduced multivariate 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 the skewed Student $t$ distribution (Sahu, Dey, and Branco, 2003, Fiorentini, Sentana, and Calzolari, 2003, and Bauwens and Laurent, 2005). Some desirable features of a multivariate model are the time variability of variances and correlations, and the asymmetry and fat-tailedness of the joint distribution. The ability to reproduce these properties of the empirical multivariate distribution is important, since they are related to the well-documented contagion issue.\footnote{This issue has also been analyzed in the context of the extreme value theory, by Longin and Solnik (2001) and Poon, Rockinger, and Tawn (2004).} Another issue that has not been addressed so far is the time-variability of the conditional distribution and in particular how the joint distribution is affected by past shocks.

In this paper, we present a new methodology to investigate, both in the univariate and multivariate setting, the effect of shocks on the volatility, skewness, and kurtosis of asset returns. Our first contribution is to develop a model well suited for the complex multivariate dynamics of asset returns. This model is an extension of Engle’s (2002) DCC model where innovations are drawn from a skewed and fat-tailed distribution. Second, realizing the difficulty of understanding the complex dynamics by simply inspecting parameters, we develop, following Engle and Ng (1993) and Cappiello, Engle, and Sheppard (2003), a graphical tool to summarize the impact of past shocks on subsequent characteristics of the returns’ distribution. In the univariate setting, this leads us to introduce News Impact Curves (NIC) of skewness and kurtosis that extend 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 Surfaces (NIS). These tools should enhance the understanding of the behavior of the joint distribution of asset returns.
In an empirical study, involving daily index returns for the four largest international markets, we investigate the actual patterns of responses to shocks. We show that after a large negative (positive) shock, the subsequent conditional distribution tends to have fatter tails and to lean leftward (resp. rightward). This suggests that large negative (positive) shocks are positively correlated. Another finding, for the US and UK, is that positive shocks are more persistent than negative shocks, because the distribution is more positively skewed after a good shock than negatively skewed after a bad shock. Turning to the multivariate analysis, we observe a similar phenomenon: After a joint negative (positive) news, the probability of subsequent joint negative (positive) news increases. Also joint positive shocks are more likely to be followed by shocks of the same sign. This pattern of responses suggests that the data is compatible with an over-reaction of asset returns to news.

The non-normality and the time-variability of the multivariate distribution of asset returns are very likely to have dramatic consequences from the risk and portfolio management perspectives. It is well recognized that the Value-at-Risk of a portfolio has to be computed in a dynamic way 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 like to be able to allocate her portfolio on the basis of the distributional properties and in particular on the basis of the recent characteristics of the asset returns’ distribution.

To address this issue, we use our methodology to investigate how the characteristics of the optimal portfolio change after some shocks on assets returns. For the US-Japan pair, we evaluate the optimal portfolio allocation for all possible pairs of shocks and compute the NIS of the various moments of this portfolio. We show that, given the characteristics of market returns, an allocation based on recent developments greatly improves the distributional properties of the portfolio as compared to the static strategy.
2 A Multivariate Auto-Regressive Conditional Distribution

In this section, we describe a multivariate conditional setting that 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 fat-tailedness often found in the distribution of financial returns. After presenting the general set-up, we focus separately on 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:

\[
\begin{align*}
    r_t & = \mu_t(\theta) + \varepsilon_t, \\
    \varepsilon_t & = \Sigma_t(\theta)^{1/2} z_t, \\
    z_t & \sim g(z_t | \eta_t).
\end{align*}
\]

Equation (1) decomposes the return at time \( t \) as the sum of the \( n \times 1 \) vector of conditional means, \( \mu_t \equiv \mu_t(\theta) = E_{t-1}[r_t] \), and the \( n \times 1 \) vector of unexpected returns, \( \varepsilon_t \), where \( E_t \) denotes the expectation conditional on the information available at date \( t \). Equation (2) indicates that unexpected returns \( \varepsilon_t \) are defined as the product of independent innovations \( z_t \) and the conditional covariance matrix of returns, \( \Sigma_t \equiv \Sigma_t(\theta) = E_{t-1}[(r_t - \mu_t)(r_t - \mu_t)'] \). We denote by \( \Sigma_t^{1/2} \) a matrix such that \( \Sigma_t = \Sigma_t^{1/2} \Sigma_t^{1/2} \). Typically \( \Sigma_t^{1/2} \) will be a Choleski decomposition. The \( n \times 1 \) vector of independent innovations, \( z_t = \Sigma_t^{-1/2}(r_t - \mu_t) \), has zero mean and identity covariance matrix. The vector \( \theta \) contains all the parameters associated with the conditional mean and the conditional variance equations. Finally, equation (3) specifies that innovations are drawn from a conditional distribution \( g \) with, possibly time-varying, shape parameters \( \eta_t \). When the conditional distribution is normal, there is no shape parameter since the normal distribution is entirely characterized by its mean and variance. In more general cases, shape parameters \( \eta_t \) may involve parameters capturing asymmetry and fat-tailedness of the distribution.
2.1 Dynamics of first and second moments

We begin with a brief description of the dynamics of the first two moments of the return process. The conditional mean of returns is described by an AR(1) process to capture the possible first-order serial correlation in returns:

\[ r_t = \mu + \varphi r_{t-1} + \varepsilon_t, \]  

(4)

where \( \mu \) is an \( n \times 1 \) vector and \( \varphi \) an \( n \times n \) diagonal matrix. The dynamics of the covariance matrix \( \Sigma_t \) is described by the DCC model proposed by Engle (2002), Engle and Sheppard (2001) and Cappiello, Engle, and Sheppard (2003). It is designed to capture both volatility clustering and persistence in correlations.\(^5\) The dynamics of each conditional variance, \( \sigma^2_{i,t} \), is given by the Glosten, Jagannathan, and Runkle (1993) model

\[ \sigma^2_{i,t} = \omega_i + \beta_i \sigma^2_{i,t-1} + \alpha_i \varepsilon^2_{i,t-1} + \gamma_i \varepsilon^2_{i,t-1} 1_{\{\varepsilon_{i,t-1} \leq 0\}}, \quad i = 1, \ldots, n, \]

(5)

while the dynamics of the conditional correlation matrix \( \Gamma_t = \{\rho_{ij}\}_{i,j=1,\ldots,n} \) is given by

\[ \Gamma_t = \left( \text{diag} \left( Q_t \right) \right)^{-1/2} \cdot Q_t \cdot \left( \text{diag} \left( Q_t \right) \right)^{-1/2}, \]

(6)

\[ Q_t = \left[ (1 - \delta_1 - \delta_2) \hat{Q} - \delta_3 \hat{N} \right] + \delta_1 Q_{t-1} + \delta_2 (u_{t-1} u'_{t-1}) + \delta_3 (n_{t-1} n'_{t-1}), \]

(7)

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, \( n_{t-1} = u_{t-1} 1_{\{u_{t-1} \leq 0\}} \) and \( \text{diag}(Q_t) \) is a matrix with zeros, except for the diagonal that contains the diagonal of \( Q_t \). Matrices \( \hat{Q} \) and \( \hat{N} \) are the unconditional covariance matrices of \( u_t \) and \( n_t \) respectively. As shown in equations (6) and (7), the dynamics of the correlation matrix, \( \Gamma_t \), is given by the asymmetric DCC specification developed by Cappiello, Engle, and Sheppard (2003). \( \delta_1, \delta_2, \) and \( \delta_3 \) are restricted to ensure that the conditional correlation matrix is positive definite.

Eventually, the covariance matrix \( \Sigma_t \) is defined as \( \Sigma_t = D_t \Gamma_t D_t \), where

\[ D_t = \begin{pmatrix} \sigma_{1,t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{n,t} \end{pmatrix}. \]

(8)

\(^5\)See also Kroner and Ng (1998) for a complete analysis of the covariance dynamics.
2.2 Conditional distribution

Multivariate analysis of asset returns often assumes the joint normality of innovations. For instance, Kroner and Ng (1998), Cappiello, Engle, and Sheppard (2003), or Ang and Chen (2002) investigate the properties of returns’ correlation assuming joint normality. As a consequence, the asymmetry found in the volatility or correlation dynamics may be partly due to the failure of the model to capture the asymmetry of the conditional distribution. Thus, disentangling the various sources of asymmetry appears to be an important empirical issue.

The \( t \) distribution was introduced to capture the conditional non-normality of asset returns by Bollerslev (1987) in the univariate set-up, and by Bollerslev and Wooldridge (1992) in a multivariate set-up. It is, however, a symmetric distribution that cannot capture the well-known negative skewness of most asset returns. A more general distribution designed to fill this gap is the so-called skewed Student \( t \) (Sk-\( t \)) distribution proposed by Hansen (1994) and Fernández and Steel (1998). It is obtained by introducing an asymmetry parameter in the Student \( t \) distribution, while maintaining the assumption of a zero mean and unit variance. Further extensions of this distribution are by Theodossiou (1998) and Jondeau and Rockinger (2003).

In our conditional set-up, the dependence between unexpected returns is introduced using the covariance matrix through \( \varepsilon_t = \Sigma_t^{1/2} z_t \). Therefore, the innovations \( z_{i,t} \) have to be distributed independently from each other. The main advantage of this approach is that it explicitly separates the modeling of the multivariate conditional distribution (through the parameters of the Sk-\( t \) distribution) and the modeling of the multivariate dependence (through the parameters of the covariance matrix). We therefore adopt a multivariate Sk-\( t \) distribution with independent components.\(^6\)

\(^6\)A similar multivariate distribution has been analyzed by Brooks, Burke, and Persand (2005) yet in the context of a symmetric \( t \) distribution. Alternative strategies could be followed to extend the Student \( t \) distribution for multivariate random variables. One may assume that the \( \chi^2 \), 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 would not be independent anymore.
We assume that the $n \times 1$ vector of innovations $z_t$ is drawn from the multivariate Sk-$t$ distribution defined as

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

where $\eta = (\nu_1, \cdots, \nu_n, \xi_1, \cdots, \xi_n)'$ denotes the vector of shape parameters,

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

and

$$
a_i = \frac{\Gamma\left(\frac{\nu_i - 1}{2}\right) \sqrt{\nu_i - 2}}{\sqrt{\pi} \Gamma\left(\frac{\nu_i}{2}\right)} \left(\xi_i - \frac{1}{\xi_i}\right),
$$

$$
b_i^2 = \xi_i^2 + \frac{1}{\xi_i^2} - 1 - a_i^2.
$$

Focusing on a given component $i$, parameters $\nu_i$ and $\xi_i$ correspond to the individual degree of freedom and the asymmetry parameter respectively. The marginal distribution of $z_{i,t}$ is a univariate Sk-$t$ distribution $g(z_{i,t}|\nu_i, \xi_i)$ as described in Hansen (1994) and Fernández and Steel (1998). The marginal distribution is defined for $2 < \nu_i < \infty$ and $\xi_i > 0$. Parameters $a_i$ and $b_i$ are required to center and scale the asymmetric distribution so that $z_{i,t}$ has zero mean and unit variance.

Higher moments of $z_{i,t}$ are easily deduced from those of the symmetric $t$ distribution $t(\cdot|\nu_i)$. If the $r$-th moment of a random variable with distribution $t(\cdot|\nu_i)$ exists, then the associated variable $z_{i,t}$ with distribution $g(\cdot|\nu_i, \xi_i)$ has a finite $r$-th moment, defined as

$$
M_{i,r} = m_{i,r} \frac{\xi_i^{r+1} + (-1)^r}{\xi_i + 1},
$$

where

$$
m_{i,r} = 2E[Z_i^r | Z_i > 0] = \frac{\Gamma\left(\frac{\nu_i - r}{2}\right) \Gamma\left(\frac{\nu_i + 1}{2}\right) \left(\nu_i - 2\right)^{\frac{r+1}{2}}}{\sqrt{\pi(\nu_i - 2)} \Gamma\left(\frac{\nu_i}{2}\right)},
$$

is the $r$-th moment of $t(\cdot|\nu_i)$ truncated to the positive real values. Note that for a given degree of freedom, the skewness of a random variable drawn from $g(\cdot|\nu_i, \xi_i)$ is minus the skewness of a random variable drawn from $g(\cdot|\nu_i, 1/\xi_i)$. We also have
$E[Z_t^3] = 0$ when $\xi_t = 1$. Provided that they exist, the skewness and kurtosis of $z_{i,t}$ are then obtained as

$$sk_{i,t}^Z = E[Z_t^3] = M_{i,3} - 3M_{i,1}M_{i,2} + 2M_{i,1}^3,$$

and

$$ku_{i,t}^Z = E[Z_t^4] = M_{i,4} - 4M_{i,1}M_{i,3} + 6M_{i,2}M_{i,1}^2 - 3M_{i,1}^4.$$  

(11)  

(12)

As it appears clearly, they are directly related, in a non-linear way, to the degree of freedom and asymmetry parameters $\nu_i$ and $\xi_i$.

In general, unexpected returns $\varepsilon_{i,t}$ are correlated with each other through the covariance matrix $\Sigma_t$. The properties of the joint distribution of the unexpected returns $\varepsilon_i$ will be useful in the following to derive the NIS of the unexpected returns’ joint distribution. Since unexpected returns are defined as $\varepsilon_t = \sum_{r=1}^{n} \omega_{ir,t}z_{r,t}$, the first four moments can be computed analytically. We first observe that $E_{t-1} \{ \varepsilon_t \} = 0$ and $V_{t-1} \{ \varepsilon_t \} = \Sigma_t$. Denoting $\Sigma_t^{1/2} = (\omega_{ij,t})_{i,j=1,...,n}$, we also have $\varepsilon_{i,t} = \sum_{r=1}^{n} \omega_{ir,t}z_{r,t}$. Then, we can define the conditional co-skewness and co-kurtosis between unexpected returns as

$$sk_{ijk,t}^\varepsilon = S_{ijk,t}\sigma_i\sigma_j\sigma_k,$$

and

$$ku_{ijkl,t}^\varepsilon = K_{ijkl,t}\sigma_i\sigma_j\sigma_k\sigma_l,$$

(13)  

(14)

where $S_{ijk,t} = E_{t-1} \{ \varepsilon_{i,t}\varepsilon_{j,t}\varepsilon_{k,t} \}$ and $K_{ijkl,t} = E_{t-1} \{ \varepsilon_{i,t}\varepsilon_{j,t}\varepsilon_{k,t}\varepsilon_{l,t} \}$ are the conditional third and fourth central moments. Given the properties of $z_t$, these expressions immediately yield

$$S_{ijk,t}^\varepsilon = E_{t-1} \left[ \sum_{r=1}^{n} \omega_{ir,t}\omega_{jr,t}\omega_{kr,t} z_{r,t}^3 \right] = \sum_{r=1}^{n} \omega_{ir,t}\omega_{jr,t}\omega_{kr,t} s_{k,r},$$  

and

$$K_{ijkl,t}^\varepsilon = E_{t-1} \left[ \sum_{r=1}^{n} \omega_{ir,t}\omega_{jr,t}\omega_{kr,t}\omega_{lr,t} z_{r,t}^4 \right] + E_{t-1} \left[ \sum_{r=1}^{n} \sum_{s\neq r} \psi_{rs,t} z_{r,t}^2 z_{s,t}^2 \right] = \sum_{r=1}^{n} \omega_{ir,t}\omega_{jr,t}\omega_{kr,t}\omega_{lr,t} k_{u,r} + \sum_{r=1}^{n} \sum_{s\neq r} \psi_{rs,t,},$$

(15)  

(16)

where $\psi_{rs,t} = \omega_{ir}\omega_{jr}\omega_{ks}\omega_{ls} + \omega_{ir}\omega_{js}\omega_{ks}\omega_{lr} + \omega_{is}\omega_{jr}\omega_{kr}\omega_{ls} - 3$. As it clearly appears from


these expressions, the time-variability of co-skewness and co-kurtosis between unexpected returns may have two sources. On the one hand, the covariance matrix between unexpected returns is time varying, so that the $\omega_{ij,t}$ are time-varying. On the other hand, individual skewness and kurtosis of innovations may be themselves time varying, as we will describe in the next section.

We observe that in the absence of asymmetry in the univariate distributions ($sk_{r,t}^Z = 0, \forall r$), no asymmetry will be found in the multivariate distribution of returns (see equation (15)). In addition, since the $z_{i,t}$ are independent from each other, the co-skewness are equal to zero when the covariance matrix is diagonal. Concerning the co-kurtosis, it is constituted of two blocks. The first one corresponds to the individual kurtosis $ku_{r}^Z$. The second term involves products of the form $E_{t-1} [z_{r,t}^2 z_{s,t}^2]$, which are equal to one for $r \neq s$.

In Figure 1, we represent bivariate contour plots of the Sk-t distribution. Left-hand-side figures represent the distribution of (uncorrelated) innovations, while right-hand-side figures represent the distribution of unexpected returns assuming a correlation of 0.5. Top figures are obtained for symmetric marginal distributions ($\xi_1 = \xi_2 = 1$). In middle figures, both marginal densities have negative skewness ($\xi_1 = \xi_2 = -0.5$), whereas the densities in bottom figures have opposite skewness ($\xi_1 = -0.5, \xi_2 = 0.5$). In all cases, the degrees of freedom are $\nu_1 = \nu_2 = 10$. The figures show that very different patterns can be obtained for the distribution of unexpected returns, once innovations’ distributions and correlations are combined. Middle figures illustrate the situation most frequently encountered in the empirical part of the paper with both marginal distributions being negatively skewed.

2.3 Dynamics of the higher moments

We now describe the temporal evolution of conditional distribution’s shape parameters $\eta_t$. In the case of the Sk-t distribution, $\eta_t$ includes the degree-of-freedom and the asymmetry parameters $(\nu_{i,t}, \xi_{i,t})_{i=1,\ldots,n}$. A natural approach is to render the dynamics of the shape parameters dependent on past shocks as in $\eta_t = \eta (z_{t-1}, z_{t-2}, \cdots)$. An important issue is to constrain the dynamics in order to ensure that the function
$g$ is a well-defined distribution.\footnote{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 Auto-Regressive Conditional Density (ARCD). A similar approach was adopted by Jondeau and Rockinger (2003), who discuss several possible specifications for the dynamics of the shape parameters.}

We adopt the following asymmetric GARCH-like model:

$$ \log \left( \nu_{i,t} - \nu \right) = c_{i,0} + c_{i,1} |z_{i,t-1}| N_{i,t-1} + c_{i,1}^+ |z_{i,t-1}| P_{i,t-1} $$

$$ + c_{i,2} \log \left( \nu_{i,t-1} - \nu \right), $$

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

where $N_{i,t} = 1_{\{z_{i,t} \leq 0\}}$ and $P_{i,t} = 1 - N_{i,t}$. Parameter $\nu$ is the lower bound for the degree of freedom to ensure that the distribution is well defined. Three main features of these specifications are worth emphasizing. First, the degree-of-freedom parameter $\nu_{i,t}$ is expected to be related to the absolute value of lagged innovations, since 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, since $\xi_{i,t}$ is likely to reflect the sign and size of the recent shocks. Second, instead of assuming that positive and negative shocks have the same impact on the distribution’s shape, we allow the shape parameters to react asymmetrically to recent shocks. Finally, equations (17) and (18) include a lag of the dependent variable, in order to capture the possible persistence in the dynamics of the higher moments. Once the dynamics of the shape parameters is estimated, higher moments are deduced from equations (11) and (12).

Our specification makes it possible to explore the strength of the various sources of asymmetry in the model. Indeed, shocks are allowed to affect variances and correlations asymmetrically. The conditional distribution is itself asymmetric, so that shocks of a given sign are more likely to occur than shocks of the other sign. Last, 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. An important issue is whether the asymmetry found in the volatility dynamics and the
asymmetry of the conditional distribution may affect each other. As pointed out by Meddahi and Renault (2004), the two concepts are strongly related: the asymmetry found at a low frequency may alternatively result from asymmetry in the volatility process or in the conditional distribution at high frequency.

2.4 Estimation

Under normality, the estimation of the DCC model can be performed in several steps, because the log-likelihood can be written into blocks involving different sets of parameters (Engle, 2002). In the case of the Sk-\( t \) distribution, such a decomposition of the log-likelihood is not available, and the estimation of the DCC model requires estimating all the parameters simultaneously. In the model described above, 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

\[
\log L (r_1, \cdots, r_T | \theta, \eta) = \sum_{t=1}^{T} \log \left[ g \left( \Sigma_t (\theta)^{-1/2} (r_t - \mu_t (\theta)) | \eta \right) \right], \tag{19}
\]

where

\[
\theta = (\theta_1, \cdots, \theta_n, \delta_1, \delta_2, \delta_3)' \quad \text{with} \quad \theta_i = (\mu_i, \varphi_i, \omega_i, \alpha_i, \beta_i, \gamma_i)',
\]

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

Maximizing expression (19) with respect to parameter vectors \( \theta \) and \( \eta \) yields the maximum-likelihood (ML) estimates. Although the joint estimation of all parameters is quite time-consuming, it remains reasonable for moderate-size systems.\(^9\)

\(^9\)The estimation is performed using MATLAB. The Gradient and the Hessian are computed numerically. Since we do not claim that the Sk-\( t \) distribution is necessarily the true distribution, we report systematically robust standard errors.
3 News impact curves and surfaces

The NIC was introduced in finance by Engle and Ng (1993) to represent the response of volatility to a shock on the asset return. More precisely, it measures the effect of a shock at date $t$ on volatility at date $t + 1$, while the information dated $t - 1$ and earlier is held constant. It has been extended to measure the response of the conditional correlation to shocks on two asset returns by Kroner and Ng (1998) as well as Cappiello, Engle, and Sheppard (2003). We now extend this concept to the response of the conditional distribution to shocks on returns. This analysis is done in two steps. We begin with the NIC of the individual higher moments of innovations $z_{i,t}$. This step relies on the univariate distribution only, since the $z_{i,t}$ are independent from each other. Then, we construct the NIS of the higher moments of unexpected returns $\varepsilon_{i,t}$, which involves in addition the response of the covariance matrix to shocks.

3.1 News impact curves for the univariate distribution

A shock $z$ at time $t$ affects the level of the shape parameters at date $t + 1$ that in turn affect the higher moments of the distribution. The following proposition gives the expression of the NIC of volatility, shape parameters and higher moments.

Proposition 1 (1) The NIC of volatility of unexpected returns is given by

$$\sigma_i(z)^2 = \begin{cases} A_{\sigma,i} + (\alpha_i + \gamma_i) \bar{\sigma}_i^2 z^2, & \text{if } z \leq 0, \\ A_{\sigma,i} + \alpha_i \bar{\sigma}_i^2 z^2, & \text{if } z > 0, \end{cases}$$

where $\bar{\sigma}_i^2$ is the unconditional variance of $\varepsilon_{i,t}$ and $A_{\sigma,i} = \omega_i + \beta_0 \bar{\sigma}_i^2$.

(2) The NIC of conditional distribution's shape parameters are given by

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

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

(20)
where \( A_{\nu,i} = c_{i,0} + c_{i,2} \log (\bar{\nu}_i - \nu) \) 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.

(3) The NIC of conditional skewness and kurtosis of innovations are given by\(^{10}\)

\[
\begin{align*}
sk^Z_i (z) &= \frac{\xi_i}{1 + (\xi_i^Z)^2} \left[ C_{i,4} \left( m_{i,4} - 3m_{i,1}m_{i,2} + 2m_{i,3}^2 \right) \\
&+ C_{i,2} \left( 3m_{i,1}m_{i,2} - 4m_{i,4} \right) \right], \\
ku^Z_i (z) &= \frac{\xi_i}{1 + (\xi_i^Z)^2} \left[ C_{i,5} \left( m_{i,4} - 4m_{i,1}m_{i,3} + 6m_{i,1}^2m_{i,2} - 3m_{i,1}^4 \right) \\
&+ C_{i,3} \left( 4m_{i,1}m_{i,3} - 12m_{i,1}m_{i,2} + 9m_{i,1}^4 \right) \\
&+ C_{i,1} \left( 6m_{i,1}^2m_{i,2} + 12m_{i,1}^4 \right) \right],
\end{align*}
\]

where \( m_{i,r} \) is defined in equation (10) and \( C_{i,r} = (\xi_i^Z)^r - (\xi_i^Z)^{-r} \). The NIC of conditional skewness and kurtosis of unexpected returns are also \( sk^\varepsilon_i (z) = sk^Z_i (z) \) and \( ku^\varepsilon_i (z) = ku^Z_i (z) \).

(4) The NIC of conditional third and fourth moments of unexpected returns are given by

\[
\begin{align*}
\mathcal{S}_i^\varepsilon (z) &= sk_i^Z (z) \sigma_i (z)^3, \quad \text{and} \quad \mathcal{K}_i^\varepsilon (z) = ku_i^Z (z) \sigma_i (z)^4.
\end{align*}
\]

Proof: For all proofs, see Appendix 1.

In a univariate model, 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 contrast, the NIC of the third and fourth moments of unexpected returns incorporate the additional effect of the volatility dynamics.

In Figure 2, we display various patterns for the NIC of conditional skewness \( sk_i^Z \) (top) and kurtosis \( ku_i^Z \) (bottom). We consider three different cases, which are of particular interest for empirical analysis. The first two cases are symmetric ones. For the first one, the occurrence of an extreme shock increases the probability of a subsequent extreme in the same direction (with \( c_{i,1}^+ = c_{i,1}^- = -0.5 \) and \( d_{i,1}^+ = d_{i,1}^- = 0.5 \), Case 1). This case is consistent with a positive correlation of large events.

---

\(^{10}\)To keep notations simple, we omit the dependency of the right-hand-side terms with respect to \( z \).
The second case corresponds to the opposite situation, in which the occurrence of an extreme shock decreases the probability of a subsequent extreme, but if it occurs, it is more likely to be in the opposite direction (with $c_{i,1}^- = c_{i,1}^+ = 0.5$ and $d_{i,1}^- = d_{i,1}^+ = -0.5$, Case 2). This case is consistent with a negative correlation of large events. In the last case, we introduce the following asymmetry pattern in the dynamics of higher moments: a large positive shock is very likely to be followed by another large positive shock, while an large negative shock is less likely to be followed by another large negative shock (with $c_{i,1}^- = -0.2$, $c_{i,1}^+ = -0.6$, $d_{i,1}^- = 0.2$, and $d_{i,1}^+ = 0.6$, Case 3). As we will see in Section 5, this case corresponds rather closely to our estimation of the US model.

3.2 News impact surfaces for the bivariate distribution

We now turn to the response of co-moments to news in a bivariate setting. We therefore 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 $t + 1$. This gives rise to NIS, since each component of these matrices is affected by a combination of shocks. In Proposition 2, we give the expression for the NIS of the covariance matrix.

Proposition 2 (1) The NIS of the covariance matrix of unexpected returns is given by $\Sigma(z) = \{\sigma_{ij}(z)\}$, with $\sigma_i(z)^2 = \bar{\sigma}_i^2$ where

\[
\sigma_i(z)^2 = \begin{cases} 
A_{\sigma,i} + (\alpha_i + \gamma_i) \bar{\sigma}_i^2 z_i^2, & \text{if } z_i \leq 0, \\
A_{\sigma,i} + \alpha_i \bar{\sigma}_i^2 z_i^2, & \text{if } z_i > 0,
\end{cases} \tag{24}
\]

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

where $\bar{\sigma}_i^2$ is the unconditional variance of $\varepsilon_{i,t}$ and $A_{\sigma,i} = \omega_i + \beta_i \bar{\sigma}_i^2$.

This issue has also been addressed by Cappiello, Engle, and Sheppard (2003) in the context of the asymmetric DCC model. They provide (in their Appendix B.2) the expression for the NIS of the bivariate correlation matrix (Part 1 of our Proposition 2).
(2) The NIS of the conditional correlation $\rho_{ij}(z)$ is given by

$$
\rho_{ij}(z) = \frac{A_{\rho,ij} + (\delta_2 + \delta_3) \bar{\sigma}_{ij} \bar{z}_i \bar{z}_j}{\sqrt{(A_{\rho,ii} + (\delta_2 + \delta_3) \bar{\sigma}_{ii}^2 \bar{z}_i^2)(A_{\rho,jj} + (\delta_2 + \delta_3) \bar{\sigma}_{jj}^2 \bar{z}_j^2)}}, \quad \text{if } z_i, z_j \leq 0,
$$

$$
= \frac{A_{\rho,ij} + \delta_2 \bar{\sigma}_{ij} \bar{z}_i z_j}{\sqrt{(A_{\rho,ii} + \delta_2 \bar{\sigma}_{ii}^2 \bar{z}_i^2)(A_{\rho,ij} + \delta_2 \bar{\sigma}_{jj}^2 \bar{z}_j^2)}}, \quad \text{if } z_i, z_j > 0,
$$

$$
= \frac{A_{\rho,ij} + \delta_2 \bar{\sigma}_{ij} \bar{z}_i z_j}{\sqrt{(A_{\rho,ii} + (\delta_2 + \delta_3) \bar{\sigma}_{ii}^2 \bar{z}_i^2)(A_{\rho,ij} + (\delta_2 + \delta_3) \bar{\sigma}_{jj}^2 \bar{z}_j^2)}}, \quad \text{if } z_i \leq 0, z_j > 0,
$$

$$
= \frac{A_{\rho,ij} + \delta_2 \bar{\sigma}_{ij} \bar{z}_i z_j}{\sqrt{(A_{\rho,ii} + (\delta_2 + \delta_3) \bar{\sigma}_{ii}^2 \bar{z}_i^2)(A_{\rho,ij} + (\delta_2 + \delta_3) \bar{\sigma}_{jj}^2 \bar{z}_j^2)}}, \quad \text{if } z_i > 0, z_j \leq 0,
$$

where $A_{\rho,ij} = (1 - \delta_2) \bar{q}_{ij} - \delta_3 \bar{n}_{ij}$, with $\bar{q}_{ij}$ and $\bar{n}_{ij}$ the unconditional covariances of $u_t$ and $n_t$ respectively.

From equations (13) to (16), we deduce the following Proposition 3 that 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

$$
S_{ijk}^c(z) = \sum_{r=1}^{n} \omega_{ir}(z) \omega_{jr}(z) \omega_{kr}(z) s_{r}^{Z}(z),
$$

$$
K_{ijkl}^c(z) = \sum_{r=1}^{n} \omega_{ir}(z) \omega_{jr}(z) \omega_{kr}(z) \omega_{ls}(z) k_{r}^{Z}(z) + \sum_{r=1}^{n} \sum_{s \neq r} \psi_{rs}(z),
$$

where $s_{r}^{Z}(z)$ and $k_{r}^{Z}(z)$ are defined in equations (22) and (23) 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

$$
s_{ijkl}^c(z) = \frac{S_{ijk}^c(z)}{\sigma_i(z) \sigma_j(z) \sigma_k(z)},
$$

$$
k_{ijkl}^c(z) = \frac{K_{ijkl}^c(z)}{\sigma_i(z) \sigma_j(z) \sigma_k(z) \sigma_l(z)}.
$$

---

12We remind that the elements of the Choleski decomposition of the covariance matrix are denoted $\Sigma^{1/2}(z) = (\omega_{ij}(z))_{i,j=1,\ldots,n}$. 
In the empirical section, we shall focus on the response of unexpected returns’ moments to shocks, since they combine the effect of shocks on the covariance matrix and on the conditional distribution. The response of the covariance matrix summarizes the joint effect of the shocks on the variances and the correlation. Accordingly, the response of the third and fourth moments summarizes the joint effect of the shocks on the covariance matrix and the individual skewness and kurtosis.

4 Data and estimation results

4.1 Data

We consider in our empirical investigation the four largest international stock markets, i.e. the United States, Japan, the United Kingdom, and Germany. For all countries, we use the reference market index over the period from January 1973 to December 2004, for a total of 8352 daily observations.\textsuperscript{13} The return series $r_t$ are defined as continuously compounded returns in US dollar.

Table 1 displays several sample statistics on market returns (Panel A). Concerning higher moments, we notice a rather large dispersion in the magnitude of skewness and kurtosis across markets. All markets have a significantly negative skewness, meaning that crashes occur more often than booms. In addition, the high level of kurtosis found for all markets is not consistent with the normality assumption: it ranges between 8 and 36, while normality would give a kurtosis of 3.

It may be argued that the observed skewness and kurtosis are affected by the October 1987 crash (Panel B). Indeed, if we remove the return corresponding to the day of the crash, we obtain for the US market a skewness of $-0.19$ (instead of $-1.43$) and a kurtosis of 7 (instead of 36.6). For the Japanese market, the skewness becomes positive once we remove this observation. For Germany, the shape of the distribution is barely affected when we remove the observation of the 1987 crash.

\textsuperscript{13}Indices are the S&P500, the Nikkei 225, the FTSE-100, and the DAX 30. For the FTSE-100, we spliced the series with the FTSE-all Shares before the reference index got established. To capture the lag between opening and closing of the markets, we introduce a lag between the US and Japanese markets.
Since the 1987 crash may have a dramatic effect on the shape of the distribution and on the dynamics of the higher moments, we removed this observation from our sample.\footnote{We also investigated the consequences of keeping this observation in the estimation sample. The main effect was to increase the standard error of parameter estimates and to increase the level of unconditional kurtosis.}

We then test the serial correlation in returns and in squared returns, using the Ljung-Box test and the Lee and King (1993) test respectively. The latter allows testing correlation in squared returns even in presence of serial correlation in returns. As it appears clearly, daily returns display both serial correlation and heteroscedasticity.

### 4.2 Univariate model with constant higher moments

Table 2 reports the estimates of the asymmetric DCC model when innovations are drawn from a Sk-$t$ distribution with constant shape parameters. The purpose of this table is to illustrate the ability of the Sk-$t$ distribution to capture the shape of the return distribution, as compared to the normal distribution. As expected, the conditional mean equation displays a small, yet significant, autoregressive parameter $\rho$, while the conditional volatility equations shows a strongly significant asymmetry parameter $\gamma$, suggesting that in those countries bad news has a much larger effect than good news on volatility. This result confirms the well-known ‘leverage effect’ for our data.

Considering the shape of the conditional distribution, several points are worth emphasizing. First, the degree-of-freedom parameter ranges between 6 and 12, indicating that distribution’s tails are much fatter than those of the normal distribution. In addition, the dispersion of this parameter clearly suggests that assuming the same degree of freedom for all markets would be overly restrictive. Second, the asymmetric parameter is significantly below 1 for the UK and Germany, while it is insignificantly different from 1 for the US and Japan.

On the basis of the Likelihood-Ratio statistics, we conclude for each market that...
the normal distribution is overwhelmingly rejected against the Sk-t distribution with constant shape parameters. We also test the null hypothesis that the fourth moment exists, i.e. \( H_0 : \nu_i = 4 \), versus \( H_A : \nu_i > 4 \). The null hypothesis is strongly rejected in all cases, suggesting that the third and fourth moments actually exist in our sample.\(^{15}\) Therefore, even if for some specific dates (such as the 1987 crash), the conditional skewness and kurtosis may appear infinite, this does not imply that the unconditional skewness and kurtosis are in fact infinite. For this reason, we estimated the model with time-varying shape parameters imposing \( \nu = 4 \) as the minimum value for the degree of freedom.

We also tested the validity of the model using the robust conditional moment test procedure proposed by Wooldridge (1990, 1991). This test aims at detecting if the model fails to capture some particular relations observed in the data. Since our model is designed to capture the dynamics of the first four moments, our moment conditions have to provide a good description of these four moments. As this test is rather well known, we relegate to Appendix 2 the description of its working and of the various moment conditions. Unreported results reveal that the conditional mean is correctly adjusted. In contrast, the Sk-t distribution with constant shape parameters fails to capture at least some aspects of the dynamics of higher moments in all markets. In all cases, the failure concerns the ability of the model to reproduce the asymmetric response of these moments to past shocks. These results confirm that a model with time-varying shape parameters is required to capture this asymmetric pattern.

### 4.3 Multivariate model with time-varying higher moments

In this section, we consider the estimation of the multivariate model with Sk-t distribution and time-varying higher moments. Evidence that univariate higher moments are time varying has been provided earlier by Hansen (1994), Harvey and Siddique (1999) as well as Jondeau and Rockinger (2003). Here, we extend this

\(^{15}\)The null is embedded in an open interval, hence the Chernoff problem does not arise in this context.
analysis to the higher moments and co-moments of the joint distribution and show that they are actually time varying.

We focus on the estimation of all the pairs involving the US market in combination with one of the other markets.\textsuperscript{16} We estimate a version of the asymmetric DCC model, proposed by Cappiello, Engle, and Sheppard (2003) with a Sk-t distribution and time-varying shape parameters as modeled by equations (17) and (18). To save space, robust conditional moment tests a la Wooldridge (1990, 1991) are reported in Appendix 2 (Table A). Overall, only 8 moment conditions are rejected out of 164, suggesting that our specification provides a good description of the data. Four of them are related to the difficulty of the model to capture the asymmetry of the conditional volatility.

**Dynamics of the covariance matrix.** As indicated in Section 2, it is of great interest to disentangle the asymmetry coming from the dynamic of the covariance matrix and the asymmetry coming from the distribution of innovations.\textsuperscript{17} As Table 3 shows, the correlation dynamics is strongly persistent ($\delta_1$ is close to 0.99) and parameter $\delta_2$ is significantly positive, indicating that when the two markets are simultaneously affected by shocks of the same sign, the subsequent correlation increases more than when the markets are affected by shocks of opposite signs. This result is consistent with a number of papers which document that correlation increases after a common (negative or positive) shock (Cappiello, Engle, and Sheppard, 2003, Ang and Chen, 2002). We also found that the asymmetry parameter $\delta_3$ in the DCC model is insignificant in our estimations. This finding appears at first glance at odds with the evidence reported in Cappiello, Engle, and Sheppard (2003). Their significant estimate of $\delta_3$ 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. But it is not due to the increase in the

\textsuperscript{16}Notice that, due to computational burden and for expositional purpose, we shall focus on bivariate models only.

\textsuperscript{17}This problem is similar to that developed by Engle (1982) who shows that return’s non-normality partly comes from innovation’s non-normality and from volatility clustering.
conditional correlation, but rather to the persistence in higher moments: a given co-crash induces a more negative skewness, so that the probability of occurrence of another co-crash increases.\textsuperscript{18}

In Figure 3, we display the conditional correlation for the three pairs of markets under study. For the US-Japan pair, the conditional correlation ranges between 0 and 0.4 and it steadily increased since 1993. For the US-UK pair, the figure indicates that the correlation between the two markets slightly increased between 1973 and 2004, from 0.2 to 0.3. The German market is more and more correlated with the US market. The correlation began to rise at the end of the nineties, with the development of the Internet bubble and eventually dropped with the bubble burst.

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. It means that independently of its sign, a large shock generates a subsequent fatter-tailed distribution. We also notice that $|c_{i,1}^-| < |c_{i,1}^+|$, suggesting that a positive shock increases more the distribution’s tails than a negative shock does. Turning to the skewness parameter $\xi_{i,t}$, we observe that the two parameters $d_{i,1}^-$ and $d_{i,1}^+$ are positive, meaning that a negative shock increases the probability to obtain another negative shock in the subsequent period. A positive shock also tends to increase the probability of another positive shock 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. As a consequence, the estimated model incorporates two different features of asymmetry, that complement each other: first, the lagged unexpected return $\varepsilon_{t-1}$ has a different effect on the subsequent volatility depending on its sign. Second, the lagged innovation $z_{t-1}$ affects the asymmetry of the subsequent conditional distribution. These two

\textsuperscript{18}Longin and Solnik (2001) document that the correlation between two markets instantaneously increases when one of the two markets crashes, so that the probability that the other market crashes at the same time increases. Our estimate indicates that the increase in correlation is the same after a joint crash or a joint boom of the same magnitude.
features are complementary. Indeed, 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 variations
of the asymmetry and tail-fatness parameters, 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. Eventually, the two phenomena affect
the shape of the unconditional distribution of returns: a bad news is followed by
a more dispersed and more negatively skewed distribution, so that a subsequent
negative shock is more likely.

Deducing the behavior of higher moments by contemplating parameter estimates
is a difficult task, because skewness and kurtosis are jointly related to shape param-
eters, and this in a highly non-linear way. In Figure 4, we display the dynamics
of the conditional skewness and kurtosis for the US market return. We observe a
similar evolution for the other markets. As it appears in the figure, conditional
skewness and kurtosis take rather reasonable values and vary quite a lot through
time. The typical range for skewness is $[-0.2, 0.2]$ with some positive peaks in 1982
and 1991. The range for kurtosis is $[4, 8]$. Over the last ten years of the sample,
changes in the dynamics of the skewness and kurtosis tend to be less erratic than
over the beginning of the sample. Trends also appear to be more pronounced. In
particular, we observe an increase in the skewness of the US return between 1994
and 1998 (from $-0.2$ to $0.3$) and then a significant drop (to $-0.2$ in 2002). This
trend may reflect the dynamics of the markets during the development and burst of
the Internet bubble.$^{19}$

$^{19}$The estimation of the model is done under the restriction that the degree of freedom is above
$\nu = 4$, in order to ensure that the estimated kurtosis at date $t$ is finite. This restriction implies an
upper bound for kurtosis around 63 (in the absence of asymmetry). On the opposite side, there is
no lower bound on kurtosis. When $c_{-1}$ and $c_{+1}$ are negative, the maximum value for the degree
of freedom is given by $c_0$. Therefore, the minimum value for kurtosis is $[\nu + \exp(c_0/(1 - c_2))]$. In
the figure, the dynamics of the higher moments are smoothed over 4 weeks using a simple moving
average.
5 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, as well as 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.

5.1 Response of individual higher moments to shocks

Figure 5 displays the NIC of 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 large shocks tend to cluster. As already mentioned, the response of the US skewness indicates a significant asymmetry, since skewness increases more after a large positive shock than after a large negative shock. This suggests that large positive shocks are more likely to cluster than large negative shocks. This pattern is typical of Case 3 described above.

Regarding the NIC of conditional kurtosis, we observe a U-shape pattern for the Japanese market, suggesting 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. For the US, we observe a similar result, but the increase in the kurtosis in case of a positive shock is much more pronounced than for a negative shock.

These patterns are confirmed by simply estimating the value of the skewness and kurtosis of returns after a large shock. We computed over our sample the skewness and kurtosis only for realizations that followed a large (negative or positive) return (with a standard deviation of ± one standard deviation). We obtained that the US skewness is equal to −0.1 after a large negative return but 0.8 after a large positive return. In addition, the kurtosis is equal to 6.1 and 8.9 after a large negative and a large positive return respectively. Interestingly, we do not observe such an asymmetry for the Japanese market.
5.2 Response of the covariance matrix to shocks

Figure 6 displays the NIS of the various components of the covariance matrix between US and Japanese daily returns.\textsuperscript{20} We first notice that the US variance is not affected by a shock on the Japanese market. Since we use a Choleski decomposition of the covariance matrix, the US unexpected return is not affected, by construction, by a shock on the Japanese market. The NIS of the US variance thus provides the same amount of information as the standard NIC of the variance in a univariate model. On the other side, 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 in case of a negative Japanese shock, the response of the Japanese variance is much more pronounced if there is also a negative shock on the US market. Vice versa, in case of a positive Japanese shock, the response of the Japanese variance is more pronounced if there is also a positive shock on the US market.

Inspection of the correlation surface reveals that, at the daily frequency, shocks of the same sign on the two markets (positive or negative) imply an increase in the subsequent correlation, while shocks of opposite signs are followed by a decrease in subsequent correlation. Now, if we consider the effect on the covariance, we observe that it increases after large shocks, whatever their 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 assume a 3 standard-deviation shock for the two markets, then the subsequent covariance is equal to 0.83 if the two shocks are negative, 0.54 if the two shocks are positive, but only 0.36 if the two shocks are of opposite side. We observe similar patterns for all other pairs of market returns.

5.3 Response of the higher moments and co-moments to shocks

We now describe how changes in the correlation matrix and in the conditional distribution combine to affect the higher moments of the joint distribution of unexpected

\textsuperscript{20}Here, $z_{1,t-1}$ denotes the shock on the US market return, while $z_{2,t-1}$ denotes the shock on the Japanese market return.
returns. **Figures 7 and 8** display for the US-Japan pair the NIS of third and fourth central moments.

We begin with the individual third central moments $S_{iii}(z)$ which are closely related to the standard measure of skewness. The response of the US third central moment $S_{111}(z)$ was already discussed in Section 5.1. For other countries, the shape of the third moment is affected by the sign and the size of the US shock, because the covariance between the two markets intervenes in the computation of $S_{222}(z)$. In particular, for a given shock on the Japanese return, the third moment is higher, the larger the shock on the US market. This effect suggests that in case of a simultaneous crash on the two markets, the probability of another large negative event is higher.

We now turn to cross third central moments of the form $S_{ij}(z)$. It indicates if return in a given country $j$ provides a good hedge against adverse volatility changes in country $i$. A positive value implies in fact that if the volatility in country $i$ goes up, the return in country $j$ also goes up, thereby providing a hedge. Figure 7 shows that moments $S_{112}(z)$ and $S_{122}(z)$ are higher, the higher the shocks on the two markets: after two large negative shocks ($-$3 standard deviation for the two markets), the subsequent moment $S_{112}(z)$ is equal to $-0.8$, while it is as high as 0.9 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 NIS of the fourth central moments are depicted in **Figure 8**. The individual fourth moments $K_{iiii}(z)$ 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 fat-tailedness of the Japanese return is also strongly affected by the US shock: the response of $K_{2222}(z)$ to two large negative shocks is 132, while the response to a negative Japanese shock and a positive US shock is only 47. Similarly, the response to two large positive shocks is 90, while the response to a positive Japanese shock and a negative US

\[\text{returns. **Figures 7 and 8** display for the US-Japan pair the NIS of third and fourth central moments.}

We begin with the individual third central moments $S_{iii}(z)$ which are closely related to the standard measure of skewness. The response of the US third central moment $S_{111}(z)$ was already discussed in Section 5.1. For other countries, the shape of the third moment is affected by the sign and the size of the US shock, because the covariance between the two markets intervenes in the computation of $S_{222}(z)$. In particular, for a given shock on the Japanese return, the third moment is higher, the larger the shock on the US market. This effect suggests that in case of a simultaneous crash on the two markets, the probability of another large negative event is higher.

We now turn to cross third central moments of the form $S_{ij}(z)$. It indicates if return in a given country $j$ provides a good hedge against adverse volatility changes in country $i$. A positive value implies in fact that if the volatility in country $i$ goes up, the return in country $j$ also goes up, thereby providing a hedge. Figure 7 shows that moments $S_{112}(z)$ and $S_{122}(z)$ are higher, the higher the shocks on the two markets: after two large negative shocks ($-$3 standard deviation for the two markets), the subsequent moment $S_{112}(z)$ is equal to $-0.8$, while it is as high as 0.9 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 NIS of the fourth central moments are depicted in **Figure 8**. The individual fourth moments $K_{iiii}(z)$ 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 fat-tailedness of the Japanese return is also strongly affected by the US shock: the response of $K_{2222}(z)$ to two large negative shocks is 132, while the response to a negative Japanese shock and a positive US shock is only 47. Similarly, the response to two large positive shocks is 90, while the response to a positive Japanese shock and a negative US shock is only 47.

---

\[21\text{Since we consider bivariate model in this work, terms such as } S_{ijk}(z) \text{ do not appear. Similarly, terms such as } K_{iijk}(z) \text{ and } K_{ijkl}(z) \text{ do not appear.}\]
shock is only 53. Therefore, the US shock tends to reinforce the dominant effect of the domestic shock.

The cross fourth moment $\mathcal{K}^z_{1122}$ can be interpreted as a measure of the strength between the two variances. A large value indicates that the volatilities in both markets move together. Both markets therefore provide a bad hedge against high volatility in the other market. Contemplating 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 it is during periods of high volatility in both markets that diversification provides the worst hedge. Such phenomenon can be related to the evidence provided by Ang and Bekaert (2002).

Last, we turn to cross fourth moments of the form $\mathcal{K}^z_{iiij}(z)$. 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. Market $i$ would therefore be a bad hedge against a fall in market $j$. Concerning moments $\mathcal{K}^z_{1112}(z)$ and $\mathcal{K}^z_{1222}(z)$, we obtain the following pattern: first, they increase after a (positive or negative) US shock; but after a negative US shock, they decrease with the Japanese shock; and after a positive US shock, they increase with the Japanese shock. This result suggests that the fourth-order dependence between the two markets is positive reflecting the positive correlation between the two markets.

6 News impact surfaces for the optimal portfolio

The evidence reported so far indicates that international markets are not likely to be good hedges to each other in case of a large drop in one market. The main reason is that correlation between market returns increases during turbulent periods. Therefore, the benefits of diversification are likely to be smaller than expected (Ang and Bekaert, 2002). Up to now, the consequences of joint shocks on the subsequent distribution of the optimal portfolio return have not been analyzed. This analysis is possible in our set-up, since we can deduce the behavior of the higher moments of the portfolio return from the one of individual assets.

We consider the following asset allocation problem. The investor allocates her
portfolio at date $t$ for date $t+1$ between two risky assets with return $(r_{1,t+1}, r_{2,t+1})'$ and the risk free asset with return $r_{f,t}$. At date $t-1$, all the moments of the joint distribution are at their unconditional level. At date $t$, there is a shock $z = (z_1, z_2)'$. Given this shock, the investor forecasts the vector of expected returns and the covariance matrix of market returns for date $t+1$ using the model described above. She uses the mean-variance criterion to allocate her wealth. Formally, optimal portfolio weights $\alpha_t = (\alpha_{1,t}, \alpha_{2,t})'$ at time $t$ are obtained by maximizing the expression $\mu_{p,t+1} - \gamma \sigma_{p,t+1}^2$, where $\mu_{p,t+1}$ and $\sigma_{p,t+1}^2$ are the expected return and variance of the portfolio. The parameter $\gamma$ is the coefficient of risk aversion.

For a given portfolio weight vector $\alpha_t$, the expected return, variance, skewness, and kurtosis of the portfolio are easily defined as:

$$
\mu_{p,t+1} = r_{f,t} + \alpha_t' (\mu_{t+1} - r_{f,t}e),
$$

$$
\sigma_{p,t+1}^2 = \alpha_t' \Sigma_{t+1} \alpha_t,
$$

$$
sk_{p,t+1} = \alpha_t' S_{t+1}^e (\alpha_t \otimes \alpha_t) / \sigma_{p,t+1}^3,
$$

$$
ku_{p,t+1} = \alpha_t' K_{t+1}^e (\alpha_t \otimes \alpha_t \otimes \alpha_t) / \sigma_{p,t+1}^4,
$$

where $e = (1,1)'$.

We now consider a set of shocks $z = (z_1, z_2)'$ at date $t$ and evaluate their effect on the optimal portfolio weights and the moments of the optimal portfolio return. Using the results of Section 3, we obtain the following expressions.

**Proposition 4** When short selling and borrowing are allowed, the NIS of the opti-

---

22 Adopting another criterion taking higher moments into account would not substantially alter the shape of the following NIS even if it may affect the value of the portfolio weights or moments.
mal portfolio weights and the moments of the portfolio return are given by

\[
\begin{align*}
\alpha(z) &= \frac{1}{\gamma} \Sigma(z)^{-1} (\hat{\mu}(z) - r_f e), \\
\mu_p(z) &= r_f + \alpha(z)' (\hat{\mu}(z) - r_f e), \\
\sigma_p^2(z) &= \alpha(z)' \Sigma(z) \alpha(z), \\
skp_p(z) &= \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_i(z) \alpha_j(z) \alpha_k(z) S_{ijk}^\varepsilon(z) / \sigma_p^4(z), \\
kup_p(z) &= \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{l=1}^{n} \alpha_i(z) \alpha_j(z) \alpha_k(z) \alpha_l(z) K_{ijkl}^\varepsilon(z) / \sigma_p^4(z),
\end{align*}
\]

where the vector of expected returns is \( \hat{\mu}(z) = (I_2 - \varphi)^{-1} \mu + \varphi \Sigma^{1/2} z, \) with \( \Sigma \) the unconditional covariance matrix of \( \varepsilon_t, \) and \( I_2 \) is the \( 2 \times 2 \) identity matrix.

We now compare with this graphical tool two alternative investment strategies: the first one is the static strategy, based on the sample mean and variance of market returns. The optimal allocation is therefore independent on recent shocks. The second strategy is the dynamic mean-variance strategy, based on the best forecast at date \( t \) of the expected returns and variances for date \( t+1. \) This allocation depends on the realization of the recent shocks. Thus, for each pair of shocks, we compute the optimal weights and the moments of the portfolio return. The comparison between the two strategies is made with the US-Japan pair. We also assume \( \gamma = 5. \)

We first consider the optimal allocation when short selling and borrowing are allowed. In this context, the optimal static allocation is given by portfolio weights \((0.206; 0.100; 0.693)'\) for the US, the Japanese, and the US T-bill markets respectively. Figure 9 displays the various characteristics of the optimal dynamic portfolio. The point in the figure corresponding to \((z_1, z_2) = (0, 0)\) is the optimal static allocation. Given the dependence of current returns on past shocks, portfolio weights for a given market are obviously increasing with the value of the shock on this market. Since short selling and borrowing are allowed, the investor typically sells assets with negative expected return and buys assets with positive expected return. As a result, the expected return of the optimal portfolio increases after large (negative as well as positive) past shocks. This also implies that the portfolio variance increases after large past shocks, since a large shock (of either sign) implies
a large subsequent variance of individual market returns. However, while individual market volatilities are more affected by bad news than by good news, we obtain the opposite result for the optimal portfolio volatility. The reason is that, after large negative shocks, the portfolio is massively invested in the risk free asset, while the magnitude of the short-selling of the risky assets remains moderate.

Regarding the higher moments of the portfolio return, we notice that the skewness is negative for small shocks (it is equal to −0.6 for the static portfolio), but it increases significantly for large shocks of either sign. The reason is the following: market returns with a negative (positive) shock are expected to have a subsequent negative (positive) skewness and simultaneously a negative (positive) weight. As a consequence, the portfolio skewness increases in either case. Last, the kurtosis of the portfolio is rather small, except after a large positive shock on the US market, so that the weight of this asset is large. As compared to the static strategy, the dynamic strategy also results in a larger portfolio skewness for most combinations of shocks, given the ability to short sell assets with negative expected returns. On the opposite, the kurtosis of the dynamic portfolio significantly exceeds the kurtosis of the static portfolio only when the weight of the US market is very large (i.e., after a very large positive shock on the US return).

The figure also shows the NIS of the Sharpe ratio of the optimal dynamic strategy. As expected, it is always larger than the Sharpe ratio of the optimal static strategy, since the dynamic investor is able to use more information on future returns. The difference is small only when the two past shocks are large and positive.

In the last experiment, we consider an all-equity portfolio, excluding the risk free asset from the opportunity set and the possibility of short-selling. The optimal static allocation is given by portfolio weights (0.69; 0.31)′ for the US and Japanese markets respectively. As shown by Figure 10, the characteristics of the optimal dynamic portfolio are rather different from the previous case: the investor can no longer benefit from bear markets by short selling her position in the assets with negative expected return. We obtain a similar result for the portfolio skewness, which turns out to be negative when both expected returns are negative. On the
opposite, the portfolio variance and kurtosis do not exceed the variance and kurtosis of the individual assets.

The last plot of the figure shows that the Sharpe ratio of the dynamic portfolio is higher when the shocks on the two markets are of opposite sign. The reason is that in this case the investor invests all her wealth in the asset with a positive expected return, which also has a positive skewness. On the opposite the optimal static portfolio is diversified, so that it includes the asset with a negative expected return and consequently with a negative skewness.

It is worth emphasizing that these properties of the optimal dynamic portfolios are not specific to the US and Japanese markets. Indeed for the other pairs of markets, we obtained essentially the same patterns. This result may be explained by the fact that the estimates of the multivariate models involving the US market are very close, suggesting that the interactions between the US market and other international markets are similar from one market to the other. Also, given the weight of the US market in international portfolios, the main characteristics of the US return in terms of higher moments (in particular, the large increase in the skewness and kurtosis of the innovation’s distribution after a positive shock) can be expected to be found in the optimal portfolio return. This is basically what we obtain for the all-equity portfolio.

7 Conclusion

In this paper, we propose two methodological contributions. The first contribution is that we model asset returns in a multivariate setting within which we allow for feedback effects of shocks not only on the covariance matrix but also on the higher moments (co-skewness and co-kurtosis) of the joint distribution. We achieve this by extending the DCC model of Engle (2002) to the case of innovations drawn from a Sk-$t$ distribution. The second contribution is the development of a graphical tool that extends the concept of news impact curve (NIC). In a univariate setting, this leads us to the NIC of skewness and kurtosis. In a multivariate setting, we obtain the news impact surfaces (NIS) of the various co-moments, thus allowing a better
characterization of the joint distribution of returns.

In the empirical section, we demonstrate the working of our methodology, and using daily index returns of four developed markets, we establish some stylized features of feedback effects. We find that subsequent to a shock of a given sign, another shock of similar sign is likely to take place. Unlike the volatility asymmetry documented by Glosten, Jagannathan, and Runkle (1993) whereby volatility is larger after a negative shock than after a positive shock of the same magnitude, the effects of shocks on subsequent return’s skewness is more complex. Even though large negative (positive) events tend to be followed by another negative (positive) event, the conditional skewness tends to be higher after a positive shock than after a negative shock of the same magnitude. A positive return will tend to trigger a larger positive return than if the return had been negative of the same magnitude, in which case we would of course expect a negative return. In a multivariate conditional setting, we establish some stylized features. In particular, we document that past foreign events have little impact on the current skewness or kurtosis beyond that information contained in domestic past events.

There exists a strong interest in the finance literature to test theoretical models by considering their predictions in terms of response of future moments to current shocks. For instance, the econometric model of Harvey and Siddique (1999) showed that there is time variation in skewness and kurtosis. This finding can be related to several theoretical work that showed that higher moments should be affected by past returns (Veronesi, 1999, Cao, Coval, and Hirshleifer, 2002, Hong and Stein, 2003, Chen, Hong, and Stein, 2001). Our methodology should be useful to allow for a better discrimination of which theoretical model provides best predictions.

In the multivariate setting, the study of feedback effects has been relatively scarce since most investigations focused on the contemporaneous dependency of markets. While earlier contributions focused on correlations, a more recent literature investigates the dependency further out in the tails of distributions. For instance in the literature on contagion of financial markets, Longin and Solnik (2001) or Poon, Rockinger, and Tawn (2004) show that the dependency in the left tail of
distributions is stronger than in the right tail. These contributions are, however, not of a dynamic nature. The recent dynamic study by Bekaert, Harvey, and Ng (2005) emphasizes correlation rather than higher moments. Our model and proposed graphical tool should prove useful to gain a better understanding of the validity of theoretical contagion models.

Our analysis of the response of the optimal portfolio to shocks on individual assets shows how the moments of the portfolio return are affected by past shocks. Our finding suggests that the characteristics of the optimal portfolio are actually very sensitive to past shocks. Therefore, it may be of great importance for the investor to adopt an allocation strategy that incorporates the effect of shocks on the distribution of portfolio returns. This result seems at odds with some empirical evidence reported for instance by Harvey et al. (2002) and Das and Uppal (2004) that the mean-variance criterion approximates correctly the expected utility except in case of large departure from normality or for higher levered portfolios. It should be noticed however that these contributions assume that the distribution of the opportunity set is constant through time and therefore do not consider the possibility that the current distribution actually depends on past shocks. In a regime-switching framework, Ang and Bekaert (2002) as well as Guidolin and Timmermann (2003) demonstrate the importance of taking higher moments into account for asset allocation purposes. We leave the investigation of the optimal allocation strategy in the case of a fully time-varying distribution of returns for further research.
8 Appendices

8.1 Appendix 1: Proofs of the propositions

For all propositions, we consider that the unconditional distribution applies at date \( t - 1 \), just before the shock. At date \( t \), a shock occurs. Then, we evaluate the various moments at date \( t + 1 \) conditionally on this shock.

Proof of Proposition 1  
(1) The first expression corresponds to the response of the variance to a shock \( z \). The Glosten, Jagannathan, and Runkle (1993) model (equation (5)) can be rewritten as follows, if volatility at date \( t \) is at its unconditional level:

\[
\sigma_{i,t+1}^2 = \omega_i + \beta_i \tilde{\sigma}_i^2 + \alpha_i \varepsilon_{i,t}^2 + \gamma_i \varepsilon_{i,t}^2 1_{\{\varepsilon_{i,t} \leq 0\}}.
\]

Since \( \varepsilon_{i,t} = \sigma_{i,t} z_{i,t} \), this expression gives

\[
\sigma_{i,t+1}^2 = \omega_i + \beta_i \tilde{\sigma}_i^2 + \left( \alpha_i + \gamma_i 1_{\{z \leq 0\}} \right) \tilde{\sigma}_i^2 z^2.
\]

so that

\[
\sigma_i(z)^2 = \begin{cases} 
A_{\sigma,i} + (\alpha_i + \gamma_i) \tilde{\sigma}_i^2 z^2, & \text{if } z \leq 0, \\
A_{\sigma,i} + \alpha_i \tilde{\sigma}_i^2 z^2, & \text{if } z > 0,
\end{cases}
\]

where \( A_{\sigma,i} = \omega_i + \beta_i \tilde{\sigma}_i^2 \).

(2) Using a similar approach and equations (17) and (18), we deduce the expressions for the conditional distribution’s shape parameters \( \nu_i^2(z) \) and \( \xi_i^2(z) \).

(3) The responses of conditional skewness and kurtosis are directly deduced from the relations (11) and (12) together with the definition of the \( r \)-th moments \( M_{i,r} \) for the skewed \( t \) distribution.

(4) The expressions for the conditional third and fourth moments make use of relations (13) and (14) in the case \( i = j = k = l \).

Proof of Proposition 2  
(1) The expression for the covariance matrix is obvious, since it just defines the covariance terms as function of the variances and correlations.
(2) The expression for the correlation matrix comes from Cappiello, Engle, and Sheppard (2003, Appendix B.2).

Proof of Proposition 3  (1) The expressions for the third and fourth central co-moments directly come from the expressions given in equations (15) and (16).

(2) The expressions for the co-skewness and co-kurtosis make use, once again, of relations (13) and (14) for any $i, j, k,$ and $l$.

Proof of Proposition 4  The expressions in this proposition are given by the definition of the moments of a portfolio return.

8.2 Appendix 2: Moment conditions for specification tests

We first define some generalized residuals $v_{t,t}$ that should have a conditional expectation equal to zero for the various moments we are interested in. Natural definitions of the generalized residuals for a given univariate distribution are $v_1^t = \varepsilon_t$, $v_2^t = \varepsilon_t^2 - \sigma_t^2$, $v_3^t = z_t^3 - sk_t^Z$, and $v_4^t = z_t^4 - ku_t^Z$, for the conditional mean, variance, skewness and kurtosis respectively (we omitted the index of the country to lighten notations). We then define various sets of moment conditions $x_{j,t-1}^k$ that may help to predict the generalized residual $v_k^t$. For the conditional mean, we set

$$x_{1, t-1}^1 = \varepsilon_{t-1}, \quad x_{1, t-1}^1 = \varepsilon_{t-2}.$$  

For the conditional variance, we set

$$x_{1, t-1}^2 = \varepsilon_{t-1}^2 - \sigma_{t-1}^2, \quad x_{2, t-1}^2 = \varepsilon_{t-2}^2 - \sigma_{t-2}^2;$$

$$x_{3, t-1}^2 = \varepsilon_{t-1}1_{\{\varepsilon_{t-1} \leq 0\}}, \quad x_{4, t-1}^2 = \varepsilon_{t-1}1_{\{\varepsilon_{t-1} > 0\}}.$$  

For the conditional skewness, we set

$$x_{1, t-1}^3 = z_{t-1}^3 - sk_{t-1}^Z, \quad x_{2, t-1}^3 = z_{t-2}^3 - sk_{t-2}^Z;$$

$$x_{3, t-1}^3 = z_{t-1}1_{\{z_{t-1} \leq 0\}}, \quad x_{4, t-1}^3 = z_{t-1}1_{\{z_{t-1} > 0\}}.$$  

For the conditional kurtosis we set

$$x_{1, t-1}^4 = z_{t-1}^4 - ku_{t-1}^Z, \quad x_{2, t-1}^4 = z_{t-2}^4 - ku_{t-2}^Z;$$

$$x_{3, t-1}^4 = z_{t-1}1_{\{z_{t-1} \leq 0\}}, \quad x_{4, t-1}^4 = z_{t-1}1_{\{z_{t-1} > 0\}}.$$  

34
For the multivariate generalized residuals reported in Table A, we use the same moment conditions as in Kroner and Ng (1998) or Cappiello, Engle, and Sheppard (2003) for the correlation and similar moment conditions for the co-skewness and co-kurtosis. For a pair of markets \( i \) and \( j \), we adopt the following generalized residuals,

\[
\begin{align*}
  v_i^5 &= u_{i,t}u_{j,t} - \rho_{ij,t}, \\
  v_i^7 &= u_{i,t}u_{j,t}^2 - sk_{ijj,t}, \\
  v_i^9 &= u_{i,t}^3u_{j,t} - ku_{ijj,t}, \\
  v_i^6 &= u_{i,t}^2u_{j,t} - sk_{ijj,t}, \\
  v_i^8 &= u_{i,t}^2u_{j,t}^2 - ku_{ijj,t}, \\
  v_i^{10} &= u_{i,t}^3u_{j,t}^3 - ku_{ijj,t},
\end{align*}
\]

where \( u_{i,t} = \varepsilon_{i,t}/\sigma_{i,t} \). To keep notations as simple as possible, introduce \( D_{i,j}^{\#1,\#2} = 1_{\{u_{i,t-1,#0}\}}1_{\{u_{j,t-1,#20}\}} \), where the symbols \( \#_1 \) and \( \#_2 \) represent elements in the set \( \{\leq, >\} \). For the conditional correlation \( \rho_{ij,t} \), the set of moment conditions that may help to predict the residual is given by

\[
\begin{align*}
  x_{5,1,t-1}^5 &= u_{i,t-1}u_{j,t-1} - \rho_{ij,t-1}, \\
  x_{5,3,t-1}^5 &= u_{i,t-1}u_{j,t-1}D_{ij}^{\leq\leq}, \\
  x_{5,5,t-1}^5 &= u_{i,t-1}u_{j,t-1}D_{ij}^{\leq>}, \\
  x_{6,1,t-1}^6 &= u_{i,t}u_{j,t-1} - sk_{ijj,t-1}, \\
  x_{6,3,t-1}^6 &= u_{i,t}u_{j,t-1}D_{ij}^{\leq<}, \\
  x_{6,5,t-1}^6 &= u_{i,t}u_{j,t-1}D_{ij}^{<>},
\end{align*}
\]

For the conditional co-skewness \( sk_{ijj,t} \), the conditions are

\[
\begin{align*}
  x_{6,1,t-1}^6 &= u_{i,t}u_{j,t-1} - sk_{ijj,t-1}, \\
  x_{6,3,t-1}^6 &= u_{i,t}u_{j,t-1}D_{ij}^{\leq<}, \\
  x_{6,5,t-1}^6 &= u_{i,t}u_{j,t-1}D_{ij}^{<>},
\end{align*}
\]

For the conditional co-kurtosis \( ku_{ijj,k,t} \), the conditions are

\[
\begin{align*}
  x_{7,1,t-1}^7 &= u_{i,t-1}u_{j,t-1}u_{k,t-1} - ku_{ijj,k,t-1}, \\
  x_{7,3,t-1}^7 &= u_{i,t-1}u_{j,t-1}u_{k,t-1}D_{ijj}^{\leq\leq}, \\
  x_{7,5,t-1}^7 &= u_{i,t-1}u_{j,t-1}u_{k,t-1}D_{ijj}^{\leq>},
\end{align*}
\]

Following Wooldridge (1990, 1991), the moment conditions \( x_{j,t-1}^k \) are first regressed on the expected gradient of the generalized residuals \( v_i^k \) with respect to the parameters. We define the residual of this regression as \( \lambda_{j,t-1}^k \). Finally, the test statistic is given by

\[
C_j^k = \frac{\left( \frac{1}{T} \sum_{t=1}^{T} v_i^k \lambda_{j,t-1}^k \right)^2}{\frac{1}{T} \sum_{t=1}^{T} (v_i^k)^2 (\lambda_{j,t-1}^k)^2}.
\]
Under the null hypothesis that the model correctly fits the moment condition $x_{j,t-1}^k$ for the $k$th moment, $TC_{j}^k$ is asymptotically distributed as a $\chi^2(1)$.

We report in Table A the robust conditional moment test statistics for the multivariate model. In Panel a, we consider the ability of the model to correctly fit the first four individual moments (mean, variance, skewness, and kurtosis), while in Panel b, we focus on the ability of the model to capture the main features of the correlation, co-skewness, and co-kurtosis dynamics. As the table reveals, almost all moment conditions are now satisfied. Three of them concern the asymmetry in the volatility dynamics, while the last two concern the modeling of the conditional skewness of the US return. As it appears clearly from the table, almost all moment conditions are satisfied by the model concerning the dynamics of the correlation. This result indicates that the insignificance in the asymmetry parameter $\delta_3$ effectively reflects the absence of asymmetry in the correlation between daily returns. Concerning the co-skewness and the co-kurtosis matrices, the table reveals that all moment conditions are satisfied by the model.
References


Captions

**Table 1:** This table reports summary statistics on stock-market returns. The Jarque-Bera statistic is denoted by JB. The Ljung-Box statistic 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 statistic is distributed as a $N(0, 1)$. Panel A of the table corresponds to summary statistics over the full sample, including the 1987 crash. Panel B gives the same statistics but without the day of the 1987 crash (19 October 1987). $^a$ and $^b$ indicate that the statistic is significant is significant at the 1% and 5% level respectively.

**Table 2:** This table reports parameter estimates and specification tests for the GJR model with Sk-t distribution and constant higher moments. All figures in parenthesis represent standard errors. The log-likelihood of the sample is denoted lnL. The LR statistic for the test of the null hypothesis that the distribution is normal is denoted LR$_{norm}$. Under the null, it is distributed as a $\chi^2(2)$. The t-stat for the null hypotheses $\nu = 4$ and $\xi = 1$ is denoted by $t(\nu = 4)$ and $t(\xi = 1)$ respectively.

**Table 3:** This table reports parameter estimates for the multivariate DCC model with Sk-t distribution and time-varying higher moments. All figures in parenthesis represent standard errors. All the 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.

**Table A:** This table reports robust moment condition tests for the model with Sk-t distribution and time-varying shape parameters. Panel a reports the diagnostic tests for the individual moment and correlation conditions. Panel b reports the same diagnostic tests for the higher co-moment conditions.

**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 that is distributed along the horizontal axis being negatively skewed ($\xi_1 = -0.5, \xi_2 = 0.5$).
Figure 2: This figure displays various patterns of the news impact curves for the skewness (top) and kurtosis (bottom). We consider three different cases, which are of particular interest for empirical analysis. The first case corresponds to $c_{i,-} = c_{i,+} = -0.5$ and $d_{i,-} = d_{i,+} = 0.5$. The second case corresponds to $c_{i,-} = c_{i,+} = 0.5$ and $d_{i,-} = d_{i,+} = -0.5$. The last case corresponds to $c_{i,-} = -0.25$, $c_{i,+} = -0.75$, $d_{i,-} = 0.25$, and $d_{i,+} = 0.75$.

Figure 3: This figure displays the evolution of the conditional correlation for the three pairs of markets under study, resulting from parameter estimates reported in Table 3.

Figure 4: This figure displays the dynamics of the conditional skewness and kurtosis for the US return, resulting from parameter estimates reported in Table 3. They are smoothed over 4 weeks using a simple moving average.

Figure 5: This figure displays the news impact curves of conditional skewness and kurtosis for the US-Japan returns.

Figure 6: These figures display the news impact surfaces of the covariance matrix for the US-Japan returns.

Figure 7: These figures display the news impact surfaces of the co-skewness matrix for the US-Japan returns.

Figure 8: These figures display the news impact surfaces of the co-kurtosis matrix for the US-Japan returns.

Figure 9: These figures display the news impact surfaces of various measures of the optimal dynamic allocation between the US and Japan returns, when short selling and borrowing are allowed. We assume $\gamma = 5$.

Figure 10: These figures display the news impact surfaces of various measures of the optimal dynamic allocation between the US and Japan returns, for the all-equity portfolio. We assume $\gamma = 5$. 
### Table 1: Summary statistics on market returns

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th>Japan</th>
<th>UK</th>
<th>Germany</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: including the 1987 crash</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Moments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.029</td>
<td>0.023</td>
<td>0.027</td>
<td>0.035</td>
</tr>
<tr>
<td>Std dev.</td>
<td>1.032</td>
<td>1.387</td>
<td>1.184</td>
<td>1.315</td>
</tr>
<tr>
<td>Maximum</td>
<td>8.709</td>
<td>12.571</td>
<td>9.048</td>
<td>9.332</td>
</tr>
<tr>
<td>Skewness</td>
<td>-1.432</td>
<td>-0.043</td>
<td>-0.305</td>
<td>-0.306</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>36.609</td>
<td>10.856</td>
<td>8.870</td>
<td>8.269</td>
</tr>
<tr>
<td>JB test</td>
<td>382532.6</td>
<td>21126.7</td>
<td>11834.0</td>
<td>9649.5</td>
</tr>
<tr>
<td><strong>Panel B: excluding the 1987 crash</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Moments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.030</td>
<td>0.024</td>
<td>0.030</td>
<td>0.036</td>
</tr>
<tr>
<td>Std dev.</td>
<td>0.993</td>
<td>1.369</td>
<td>1.167</td>
<td>1.308</td>
</tr>
<tr>
<td>Maximum</td>
<td>5.573</td>
<td>12.571</td>
<td>9.048</td>
<td>9.332</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.190</td>
<td>0.190</td>
<td>-0.071</td>
<td>-0.270</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>6.991</td>
<td>7.759</td>
<td>6.578</td>
<td>8.103</td>
</tr>
<tr>
<td>JB test</td>
<td>5399.5</td>
<td>7795.4</td>
<td>4356.0</td>
<td>9024.2</td>
</tr>
<tr>
<td><strong>Serial correlation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LB(5)</td>
<td>45.746</td>
<td>7.330</td>
<td>156.492</td>
<td>14.341</td>
</tr>
<tr>
<td>LK(5)</td>
<td>25.246</td>
<td>22.153</td>
<td>33.159</td>
<td>24.427</td>
</tr>
<tr>
<td><strong>Correlation matrix</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US</td>
<td>1</td>
<td>0.224</td>
<td>0.234</td>
<td>0.273</td>
</tr>
<tr>
<td>Japan</td>
<td>0.224</td>
<td>1</td>
<td>0.102</td>
<td>0.131</td>
</tr>
<tr>
<td>UK</td>
<td>0.234</td>
<td>0.229</td>
<td>1</td>
<td>0.452</td>
</tr>
<tr>
<td>Germany</td>
<td>0.273</td>
<td>0.131</td>
<td>0.452</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2: Parameter estimates of the GJR model with Sk-\(t\) distribution and constant shape parameters

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conditional mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\mu)</td>
<td>0.025</td>
<td>(0.010)</td>
<td>0.032</td>
<td>(0.011)</td>
<td>0.025</td>
<td>(0.011)</td>
<td>0.033</td>
<td>(0.017)</td>
</tr>
<tr>
<td>(\rho)</td>
<td>0.062</td>
<td>(0.011)</td>
<td>0.028</td>
<td>(0.010)</td>
<td>0.117</td>
<td>(0.012)</td>
<td>0.015</td>
<td>(0.013)</td>
</tr>
<tr>
<td><strong>Conditional variance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\omega)</td>
<td>0.009</td>
<td>(0.002)</td>
<td>0.023</td>
<td>(0.005)</td>
<td>0.020</td>
<td>(0.004)</td>
<td>0.022</td>
<td>(0.005)</td>
</tr>
<tr>
<td>(\alpha)</td>
<td>0.022</td>
<td>(0.004)</td>
<td>0.039</td>
<td>(0.006)</td>
<td>0.050</td>
<td>(0.007)</td>
<td>0.047</td>
<td>(0.007)</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>0.061</td>
<td>(0.011)</td>
<td>0.084</td>
<td>(0.012)</td>
<td>0.043</td>
<td>(0.009)</td>
<td>0.046</td>
<td>(0.011)</td>
</tr>
<tr>
<td>(\beta)</td>
<td>0.939</td>
<td>(0.007)</td>
<td>0.910</td>
<td>(0.009)</td>
<td>0.914</td>
<td>(0.010)</td>
<td>0.916</td>
<td>(0.011)</td>
</tr>
<tr>
<td><strong>Shape parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\nu)</td>
<td>8.921</td>
<td>(0.917)</td>
<td>6.822</td>
<td>(0.508)</td>
<td>11.448</td>
<td>(1.812)</td>
<td>9.830</td>
<td>(1.202)</td>
</tr>
<tr>
<td>(\xi)</td>
<td>0.980</td>
<td>(0.015)</td>
<td>1.003</td>
<td>(0.015)</td>
<td>0.924</td>
<td>(0.016)</td>
<td>0.948</td>
<td>(0.015)</td>
</tr>
<tr>
<td>lnL</td>
<td>10412.8</td>
<td></td>
<td>13009.6</td>
<td></td>
<td>11858.5</td>
<td></td>
<td>12857.4</td>
<td></td>
</tr>
<tr>
<td>LR(_{norm})</td>
<td>328.6</td>
<td></td>
<td>551.4</td>
<td></td>
<td>239.6</td>
<td></td>
<td>342.4</td>
<td></td>
</tr>
<tr>
<td>(t(\nu = 4))</td>
<td>5.366</td>
<td></td>
<td>5.550</td>
<td></td>
<td>4.111</td>
<td></td>
<td>4.849</td>
<td></td>
</tr>
<tr>
<td>(t(\xi = 1))</td>
<td>1.351</td>
<td></td>
<td>0.211</td>
<td></td>
<td>4.861</td>
<td></td>
<td>3.411</td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Parameter estimates of the DCC model with Sk-\(t\) distribution and time-varying shape parameters

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th>Japan</th>
<th>UK</th>
<th>Germany</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conditional mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\mu)</td>
<td>0.028 (0.014)</td>
<td>0.030 (0.025)</td>
<td>0.027 (0.017)</td>
<td>0.037 (0.025)</td>
</tr>
<tr>
<td>(\rho)</td>
<td>0.061 (0.017)</td>
<td>0.021 (0.016)</td>
<td>0.110 (0.013)</td>
<td>0.025 (0.015)</td>
</tr>
<tr>
<td><strong>Conditional variance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\omega)</td>
<td>0.007 (0.002)</td>
<td>0.019 (0.005)</td>
<td>0.020 (0.006)</td>
<td>0.036 (0.006)</td>
</tr>
<tr>
<td>(\alpha)</td>
<td>0.034 (0.009)</td>
<td>0.036 (0.005)</td>
<td>0.051 (0.011)</td>
<td>0.051 (0.010)</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>0.051 (0.012)</td>
<td>0.076 (0.012)</td>
<td>0.039 (0.013)</td>
<td>0.046 (0.016)</td>
</tr>
<tr>
<td>(\beta)</td>
<td>0.939 (0.009)</td>
<td>0.914 (0.010)</td>
<td>0.909 (0.014)</td>
<td>0.897 (0.009)</td>
</tr>
<tr>
<td><strong>Conditional correlation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\delta_1)</td>
<td></td>
<td>0.9881 (0.0070)</td>
<td>0.9922 (0.0097)</td>
<td>0.9896 (0.0012)</td>
</tr>
<tr>
<td>(\delta_2)</td>
<td></td>
<td>0.0068 (0.0033)</td>
<td>0.0039 (0.0041)</td>
<td>0.0073 (0.0010)</td>
</tr>
<tr>
<td>(\delta_3)</td>
<td></td>
<td>0.0001 (0.0012)</td>
<td>0.0009 (0.0015)</td>
<td>0.0001 (0.0011)</td>
</tr>
<tr>
<td><strong>Conditional degree of freedom</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c_0) /100</td>
<td>0.036 (0.009)</td>
<td>0.037 (0.018)</td>
<td>0.032 (0.013)</td>
<td>0.037 (0.014)</td>
</tr>
<tr>
<td>(c^+_1)</td>
<td>-0.540 (0.319)</td>
<td>-0.211 (0.184)</td>
<td>0.222 (0.090)</td>
<td>-0.109 (0.059)</td>
</tr>
<tr>
<td>(c^-_1)</td>
<td>-2.176 (0.521)</td>
<td>-0.506 (0.188)</td>
<td>-0.529 (0.509)</td>
<td>-0.750 (0.439)</td>
</tr>
<tr>
<td>(c_2)</td>
<td>0.574 (0.111)</td>
<td>-0.307 (0.468)</td>
<td>0.342 (0.476)</td>
<td>0.273 (0.279)</td>
</tr>
<tr>
<td><strong>Conditional asymmetry</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d_0)</td>
<td>0.973 (0.011)</td>
<td>1.006 (0.019)</td>
<td>0.967 (0.032)</td>
<td>0.981 (0.013)</td>
</tr>
<tr>
<td>(d^-_1)</td>
<td>0.031 (0.013)</td>
<td>0.053 (0.021)</td>
<td>0.024 (0.040)</td>
<td>0.052 (0.022)</td>
</tr>
<tr>
<td>(d^+_1)</td>
<td>0.099 (0.023)</td>
<td>0.043 (0.032)</td>
<td>0.069 (0.049)</td>
<td>0.067 (0.027)</td>
</tr>
<tr>
<td>(d_2)</td>
<td>0.613 (0.100)</td>
<td>0.550 (0.163)</td>
<td>0.680 (0.295)</td>
<td>0.642 (0.106)</td>
</tr>
<tr>
<td>lnL</td>
<td>23445.2</td>
<td>22365.1</td>
<td>23272.9</td>
<td></td>
</tr>
</tbody>
</table>
Table Aa: Robust moment condition tests (individual moment and correlation conditions)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conditional mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.815 (0.367)</td>
<td>0.135 (0.713)</td>
<td>1.358 (0.244)</td>
<td>0.854 (0.355)</td>
</tr>
<tr>
<td>2</td>
<td>0.604 (0.437)</td>
<td>1.571 (0.210)</td>
<td>0.862 (0.353)</td>
<td>2.581 (0.108)</td>
</tr>
<tr>
<td><strong>Conditional variance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.119 (0.730)</td>
<td>0.733 (0.392)</td>
<td>2.994 (0.084)</td>
<td>0.019 (0.892)</td>
</tr>
<tr>
<td>2</td>
<td>3.285 (0.070)</td>
<td>0.172 (0.678)</td>
<td>5.646 (0.018)</td>
<td>1.493 (0.222)</td>
</tr>
<tr>
<td>3</td>
<td>0.360 (0.548)</td>
<td>0.677 (0.411)</td>
<td>1.689 (0.194)</td>
<td>6.477 (0.011)</td>
</tr>
<tr>
<td>4</td>
<td>8.665 (0.003)</td>
<td>1.364 (0.243)</td>
<td>1.672 (0.196)</td>
<td>5.884 (0.015)</td>
</tr>
<tr>
<td><strong>Conditional skewness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.237 (0.135)</td>
<td>0.969 (0.325)</td>
<td>1.399 (0.237)</td>
<td>0.879 (0.348)</td>
</tr>
<tr>
<td>2</td>
<td>2.275 (0.132)</td>
<td>1.098 (0.295)</td>
<td>1.413 (0.235)</td>
<td>0.001 (0.973)</td>
</tr>
<tr>
<td>3</td>
<td>2.767 (0.096)</td>
<td>0.942 (0.332)</td>
<td>2.770 (0.096)</td>
<td>0.919 (0.338)</td>
</tr>
<tr>
<td>4</td>
<td>4.200 (0.040)</td>
<td>0.929 (0.335)</td>
<td>3.850 (0.050)</td>
<td>1.112 (0.292)</td>
</tr>
<tr>
<td><strong>Conditional kurtosis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.044 (0.833)</td>
<td>0.994 (0.319)</td>
<td>1.727 (0.189)</td>
<td>0.228 (0.633)</td>
</tr>
<tr>
<td>2</td>
<td>0.094 (0.760)</td>
<td>0.986 (0.321)</td>
<td>1.134 (0.287)</td>
<td>0.000 (0.991)</td>
</tr>
<tr>
<td>3</td>
<td>0.437 (0.509)</td>
<td>1.010 (0.315)</td>
<td>1.266 (0.261)</td>
<td>0.980 (0.322)</td>
</tr>
<tr>
<td>4</td>
<td>1.784 (0.182)</td>
<td>0.943 (0.332)</td>
<td>1.293 (0.256)</td>
<td>1.171 (0.279)</td>
</tr>
<tr>
<td><strong>Conditional correlation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>–</td>
<td>1.255 (0.263)</td>
<td>0.949 (0.330)</td>
<td>1.429 (0.232)</td>
</tr>
<tr>
<td>2</td>
<td>–</td>
<td>0.997 (0.318)</td>
<td>4.965 (0.026)</td>
<td>6.077 (0.014)</td>
</tr>
<tr>
<td>3</td>
<td>–</td>
<td>1.321 (0.250)</td>
<td>1.069 (0.301)</td>
<td>0.322 (0.570)</td>
</tr>
<tr>
<td>4</td>
<td>–</td>
<td>0.308 (0.579)</td>
<td>3.484 (0.062)</td>
<td>0.418 (0.518)</td>
</tr>
<tr>
<td>5</td>
<td>–</td>
<td>1.090 (0.296)</td>
<td>0.225 (0.636)</td>
<td>0.199 (0.656)</td>
</tr>
<tr>
<td>6</td>
<td>–</td>
<td>0.691 (0.406)</td>
<td>2.671 (0.102)</td>
<td>7.201 (0.007)</td>
</tr>
</tbody>
</table>
## Table Ab: Robust moment condition tests (higher co-moment conditions)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conditional skewness(1,1,2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.083 (0.298)</td>
<td>0.542 (0.462)</td>
<td>1.116 (0.291)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.000 (0.317)</td>
<td>1.112 (0.292)</td>
<td>1.239 (0.266)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.014 (0.314)</td>
<td>0.575 (0.448)</td>
<td>1.192 (0.275)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.857 (0.355)</td>
<td>0.931 (0.335)</td>
<td>1.015 (0.314)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.552 (0.457)</td>
<td>1.776 (0.183)</td>
<td>3.800 (0.051)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.569 (0.210)</td>
<td>1.333 (0.248)</td>
<td>1.102 (0.294)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Conditional skewness(1,2,2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.031 (0.310)</td>
<td>0.110 (0.741)</td>
<td>0.930 (0.335)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.030 (0.310)</td>
<td>0.921 (0.337)</td>
<td>1.979 (0.160)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.055 (0.304)</td>
<td>0.886 (0.347)</td>
<td>0.429 (0.513)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.557 (0.455)</td>
<td>1.248 (0.264)</td>
<td>0.478 (0.490)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.207 (0.272)</td>
<td>1.406 (0.236)</td>
<td>2.468 (0.116)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.217 (0.641)</td>
<td>0.995 (0.319)</td>
<td>1.849 (0.174)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Conditional kurtosis(1,1,1,2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.086 (0.297)</td>
<td>0.737 (0.391)</td>
<td>0.165 (0.685)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.036 (0.309)</td>
<td>0.977 (0.323)</td>
<td>0.927 (0.336)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.098 (0.295)</td>
<td>1.239 (0.266)</td>
<td>0.341 (0.560)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.042 (0.307)</td>
<td>1.067 (0.302)</td>
<td>1.069 (0.301)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.928 (0.335)</td>
<td>0.537 (0.464)</td>
<td>1.429 (0.232)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.720 (0.396)</td>
<td>0.668 (0.414)</td>
<td>0.591 (0.442)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Conditional kurtosis(1,2,2,2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.986 (0.321)</td>
<td>0.229 (0.632)</td>
<td>0.191 (0.662)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.007 (0.316)</td>
<td>0.961 (0.327)</td>
<td>1.199 (0.274)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.986 (0.321)</td>
<td>0.226 (0.634)</td>
<td>0.452 (0.501)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.463 (0.226)</td>
<td>1.396 (0.237)</td>
<td>0.228 (0.633)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.760 (0.383)</td>
<td>1.088 (0.297)</td>
<td>2.587 (0.108)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.880 (0.348)</td>
<td>1.127 (0.289)</td>
<td>1.147 (0.284)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Conditional kurtosis(1,1,2,2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.008 (0.316)</td>
<td>0.473 (0.492)</td>
<td>0.476 (0.490)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.020 (0.312)</td>
<td>1.005 (0.316)</td>
<td>1.021 (0.312)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.010 (0.315)</td>
<td>0.368 (0.544)</td>
<td>0.711 (0.399)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.045 (0.307)</td>
<td>1.023 (0.312)</td>
<td>1.055 (0.304)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.092 (0.296)</td>
<td>0.944 (0.331)</td>
<td>1.192 (0.275)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.014 (0.314)</td>
<td>0.953 (0.329)</td>
<td>0.724 (0.395)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: Contour plot of the bivariate Sk-t distribution for various levels of asymmetry.

- **Innovations: No asymmetry**
  - $z_1$ vs $z_2$
  - $\varepsilon_1$ vs $\varepsilon_2$

- **Innovations: Negative skewness**
  - $z_1$ vs $z_2$
  - $\varepsilon_1$ vs $\varepsilon_2$

- **Innovations: Opposite asymmetry**
  - $z_1$ vs $z_2$
  - $\varepsilon_1$ vs $\varepsilon_2$
Figure 2: News impact curve of conditional higher moments (illustrative case)
Figure 3: Evolution of the conditional correlation between returns

US – Japan

US – UK

US – Germany
Figure 4: Evolution of the conditional higher moments for the US return
Figure 5: NIC of innovations’ higher moments (US-Japan returns)

Skewness $sk^Z_{i,t-1}$

Kurtosis $ku^Z_{i,t-1}$

US

Japan
Figure 6: NIS of unexpected returns’ covariance matrix (US-Japan returns)
Figure 7: NIS of unexpected returns’ co-skewness matrix (US-Japan returns)

The diagram illustrates the co-skewness matrix for unexpected returns in the US and Japan, focusing on third moments. The plots show the skewness matrices for different combinations of return series:

- US third moment $S_{1,1,1}^{e}$
- Japanese third moment $S_{2,2,2}^{e}$
- Third moment $S_{1,1,2}^{e}$
- Third moment $S_{1,2,2}^{e}$

Each plot represents the skewness matrix with axes indicating the levels of expected returns $z_{1,t-1}$ and $z_{2,t-1}$.
Figure 8: NIS of unexpected returns’ co-kurtosis matrix (US-Japan returns)

US fourth moment $K_e^{1,1,1,1}$

Japanese fourth moment $K_e^{2,2,2,2}$

Fourth moment $K_e^{1,1,1,2}$

Fourth moment $K_e^{1,1,2,2}$

Fourth moment $K_e^{1,2,2,2}$
Figure 9: Characteristics of the optimal dynamic allocation between the US and Japanese returns (with short selling or borrowing)
Figure 10: Characteristics of the optimal dynamic allocation between the US and Japanese returns (all-equity portfolio)