

The Impact of Shocks on Higher Moments

Technical Appendix

Eric Jondeau^a and Michael Rockinger^b

November 2008

^aSwiss Finance Institute and University of Lausanne. Faculty of Business and Economics, CH 1012 Lausanne, Switzerland. E-mail: eric.jondeau@unil.ch.

^bCorresponding author. Swiss Finance Institute, University of Lausanne, and CEPR. University of Lausanne. Faculty of Business and Economics, CH 1012 Lausanne, Switzerland. E-mail: michael.rockinger@unil.ch.

1 Proofs of the propositions

Proof of Proposition 1 (1) Expressions (19) and (20) correspond to the response of the degree of freedom and asymmetry parameter to a shock z . Using equations (16) and (17), we deduce the expressions for the conditional distribution's shape parameters $\nu_i^Z(z)$ and $\xi_i^Z(z)$ in equations (19) and (20) respectively.

(2) 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 Sk- t distribution.

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

Proof of Proposition 2 (1) The expression for the covariance news impact curve in the ABEKK model is based on the observation that $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are correlated if there is a shock $z = (z_1, z_2)'$ at date t , through $\varepsilon(z) = \bar{\Sigma}^{1/2}z$.

(2) The expression for the covariance matrix follows from the definition of a covariance as a product between standard deviations and correlation. The expression for the correlation matrix can be found in Cappiello, Engle, and Sheppard (2006, Appendix A.2). Again, since normalized unexpected returns u_t are defined as $u_t = \bar{D}^{-1}\varepsilon_t$, if there is a shock $z = (z_1, z_2)'$ at date t , this yields $u(z) = \bar{D}^{-1}\bar{\Sigma}^{1/2}z$.

Proof of Proposition 3 (1) The expressions for the third and fourth central co-moments directly follow from the expressions given in equations (14) and (15).

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

2 Optimal number of lags for expected returns

To capture the possible serial correlation in returns we assume for returns an Auto-Regressive process of order p . Under such a specification the conditional mean of returns may be written as

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

where μ is an $n \times 1$ vector and φ_i are $n \times n$ diagonal matrices, $i = 1, \dots, p$. To determine the optimal number of lags p , we minimize the Schwarz (1978) Bayesian Information Criterion (BIC) given by

$$BIC(p) = T \log(RSS(p)/T) + p \log(T),$$

where T is the number of observations and $RSS(p)$ is the residual sum of squares estimated with p lags.

Table 1 reports, for each country and lags ranging from 0 to 5, the R^2 of the regression, the standard error of residuals ($RSS(p)/T$) and the $BIC(p)$ statistics. The optimal number of lags is found to be 0 in Japan and Germany and 1 for the US and the UK.

3 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 (equations (1) to (9)). As expected, the conditional mean equation displays only little serial correlation. Only the first lagged return gives a small, yet significant, autoregressive parameter φ_1 . The conditional volatility equations display a strongly significant asymmetry parameter ψ , 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 ranges between 6 and 11, indicating that distribution’s tails are much fatter than those of the normal distribution. The dispersion of this parameter clearly suggests that assuming the same degree of freedom for all markets would be overly restrictive.¹ Second, the asymmetry 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 does not exist, 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 of the innovation distribution actually exist in our sample. Therefore, even if for some specific dates (such as the October 1987 crash), the skewness and kurtosis may take very large values, this does not prevent the sample skewness and kurtosis from being in fact finite. For this reason, we estimated the model with time-varying shape parameters imposing $\underline{\nu} = 4$ as the minimum value for the degree of freedom.

¹A Wald test for the null hypothesis that the four degrees of freedom are equal to each other can be based on the delta method, the statistics being distributed under the null as a $\chi^2(3)$. The test statistics is equal to 12.1 and the null is rejected with a p-value of 0.7%.

The test indicates that the conditional mean is correctly adjusted. In contrast, the Sk- t distribution with constant shape parameters fails to capture 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 Moment conditions for specification tests

In this section, we describe how we implemented the robust conditional moment test procedure proposed by Wooldridge (1990, 1991). We report the corresponding results in **Tables 3 and 4**. We first define some generalized residuals $v_{i,t}$ 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 $v_t^1 = \varepsilon_t$, $v_t^2 = \varepsilon_t^2 - \sigma_t^2$, $v_t^3 = z_t^3 - sk_t^Z$, and $v_t^4 = z_t^4 - ku_t^Z$, for the conditional mean, variance, skewness and kurtosis respectively (we omitted the index of the country to lighten notations). We then define various variables $x_{j,t-1}^k$ that may help predicting the generalized residuals v_t^k . For the conditional mean, we set

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

For the conditional variance, we set

$$\begin{aligned} x_{1,t-1}^2 &= \varepsilon_{t-1}^2 - \sigma_{t-1}^2, & x_{2,t-1}^2 &= \varepsilon_{t-2}^2 - \sigma_{t-2}^2, \\ x_{3,t-1}^2 &= \varepsilon_{t-1} \mathbf{1}_{\{\varepsilon_{t-1} \leq 0\}}, & x_{4,t-1}^2 &= \varepsilon_{t-1} \mathbf{1}_{\{\varepsilon_{t-1} > 0\}}. \end{aligned}$$

For the conditional skewness, we set

$$\begin{aligned} x_{1,t-1}^3 &= z_{t-1}^3 - sk_{t-1}^Z, & x_{2,t-1}^3 &= z_{t-2}^3 - sk_{t-2}^Z, \\ x_{3,t-1}^3 &= z_{t-1} \mathbf{1}_{\{z_{t-1} \leq 0\}}, & x_{4,t-1}^3 &= z_{t-1} \mathbf{1}_{\{z_{t-1} > 0\}}. \end{aligned}$$

For the conditional kurtosis, we set

$$\begin{aligned} x_{1,t-1}^4 &= z_{t-1}^4 - ku_{t-1}^Z, & x_{2,t-1}^4 &= z_{t-2}^4 - ku_{t-2}^Z, \\ x_{3,t-1}^4 &= z_{t-1} \mathbf{1}_{\{z_{t-1} \leq 0\}}, & x_{4,t-1}^4 &= z_{t-1} \mathbf{1}_{\{z_{t-1} > 0\}}. \end{aligned}$$

For the multivariate generalized residuals, we use the same moment conditions as in Kroner and Ng (1998) or Cappiello, Engle, and Sheppard (2006) 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,

$$\begin{aligned} v_t^5 &= u_{i,t}u_{j,t} - \rho_{ij,t}, & v_t^6 &= u_{i,t}^2u_{j,t} - sk_{ij,t}^\varepsilon, \\ v_t^7 &= u_{i,t}u_{j,t}^2 - sk_{ijj,t}^\varepsilon, & v_t^8 &= u_{i,t}^2u_{j,t}^2 - ku_{ijj,t}^\varepsilon, \\ v_t^9 &= u_{i,t}^3u_{j,t} - ku_{iii,t}^\varepsilon, & v_t^{10} &= u_{i,t}u_{j,t}^3 - ku_{ijj,t}^\varepsilon, \end{aligned}$$

where $u_{i,t} = \varepsilon_{i,t}/\sigma_{i,t}$. To keep notations as simple as possible, introduce $D_{i,j}^{\#_1, \#_2} = 1_{\{u_{i,t-1} \#_1 0\}} 1_{\{u_{j,t-1} \#_2 0\}}$, where the symbols $\#_1$ and $\#_2$ represent elements in the set $\{\leq, >\}$. For the conditional correlation $\rho_{ij,t}$, we use as instruments

$$\begin{aligned} x_{1,t-1}^5 &= u_{i,t-1}u_{j,t-1} - \rho_{ij,t-1}, & x_{2,t-1}^5 &= u_{i,t-2}u_{j,t-2} - \rho_{ij,t-2}, \\ x_{3,t-1}^5 &= u_{i,t-1}u_{j,t-1}D_{i,j}^{\leq, \leq}, & x_{4,t-1}^5 &= u_{i,t-1}u_{j,t-1}D_{i,j}^{\leq, >}, \\ x_{5,t-1}^5 &= u_{i,t-1}u_{j,t-1}D_{i,j}^{>, \leq}, & x_{6,t-1}^5 &= u_{i,t-1}u_{j,t-1}D_{i,j}^{>, >}. \end{aligned}$$

For the conditional co-skewness $sk_{ij,t}^\varepsilon$, the variables are

$$\begin{aligned} x_{1,t-1}^6 &= u_{i,t-1}^2u_{j,t-1} - sk_{ij,t-1}^\varepsilon, & x_{2,t-1}^6 &= u_{i,t-2}^2u_{j,t-2} - sk_{ij,t-2}^\varepsilon, \\ x_{3,t-1}^6 &= u_{i,t-1}^2u_{j,t-1}D_{i,j}^{\leq, \leq}, & x_{4,t-1}^6 &= u_{i,t-1}^2u_{j,t-1}D_{i,j}^{\leq, >}, \\ x_{5,t-1}^6 &= u_{i,t-1}^2u_{j,t-1}D_{i,j}^{>, \leq}, & x_{6,t-1}^6 &= u_{i,t-1}^2u_{j,t-1}D_{i,j}^{>, >}. \end{aligned}$$

For the conditional co-kurtosis $ku_{ijk,t}^\varepsilon$, we use

$$\begin{aligned} x_{1,t-1}^7 &= u_{i,t-1}^2u_{j,t-1}u_{k,t-1} - ku_{ijk,t-1}^\varepsilon, & x_{2,t-1}^7 &= u_{i,t-2}^2u_{j,t-2}u_{k,t-2} - ku_{ijk,t-2}^\varepsilon, \\ x_{3,t-1}^7 &= u_{i,t-1}^2u_{j,t-1}u_{k,t-1}D_{i,j}^{\leq, \leq}, & x_{4,t-1}^7 &= u_{i,t-1}^2u_{j,t-1}u_{k,t-1}D_{i,j}^{\leq, >}, \\ x_{5,t-1}^7 &= u_{i,t-1}^2u_{j,t-1}u_{k,t-1}D_{i,j}^{>, \leq}, & x_{6,t-1}^7 &= u_{i,t-1}^2u_{j,t-1}u_{k,t-1}D_{i,j}^{>, >}. \end{aligned}$$

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

$$C_j^k = \frac{\left(\frac{1}{T} \sum_{t=1}^T v_t^k \lambda_{j,t-1}^k\right)^2}{\frac{1}{T} \sum_{t=1}^T (v_t^k)^2 (\lambda_{j,t-1}^k)^2}.$$

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 $\chi^2(1)$.

In Table 3, we consider the ability of the model to correctly fit the first four individual moments (mean, variance, skewness, and kurtosis), and in Table 4, we focus on the ability of the model to capture the main features of the correlation, co-skewness, and co-kurtosis dynamics. As the tables reveal, almost all moment conditions are satisfied. There are four rejections concerning the asymmetry in the volatility dynamics, while three other rejections concern the asymmetry in the conditional correlation. It clearly appears from Table 4 that all moment conditions are satisfied by the model concerning the dynamics of the co-skewness and the co-kurtosis matrices.

5 Effect of eliminating the October 1987 crash

As explained in the text, we also estimated the model when the observation of the October 1987 crash is removed from the sample. The reason is that this observation

may have an effect on the shape of the distribution and on the dynamics of the higher moments. For this purpose, we removed a single return, the one of October 19th, 1987.

Table 5 reports the estimates of the ADCC model when innovations are drawn from a Sk- t distribution with time-varying shape parameters. If we compare this table with Table 2 in the text, we observe that the main consequence of including the October 1987 crash in the estimation is to increase the standard error of some parameter estimates in the dynamics of the degree of freedom and the asymmetry parameter. It also slightly affects the dynamics of the degree of freedom for the German market. We may conclude that our model performs well even if the data contains important outliers.

6 Estimation with inverse ordering in the markets

The baseline case reported in the text is based on the following idea: the dataset is constructed in such a way that at date t the US market closes first, then the Japanese market, then the European markets. For this reason, in all our bivariate models, we put the US return first. Then, the underlying assumption behind the Cholesky decomposition of the covariance matrix is that the US market at date t is not affected contemporaneously by the return on the other markets. On the opposite, the US market is allowed to affect the other markets contemporaneously.

In order to investigate the consequences of our choice of a Cholesky decomposition of the covariance matrix, based on the assumption that US news affect other markets, we consider in this section the case of the inverse ordering in the system. We construct the dataset in such a way that at date t the Japanese market closes first, then the European markets, then the US market. As a consequence, in all our bivariate models, we put the other market first and then the US market. Consequently, the Cholesky decomposition implies that the other market at date t cannot be affected by the US market contemporaneously.

Table 6 reports the estimates of the ADCC model when the Cholesky decomposition of the covariance matrix is based on this inverse ordering. As it appears clearly, there is no significant change in the parameter estimates relative to Table 2 in the main text. These results suggest that the ordering of the variables is consistent with the actual causal links between market returns.

7 Estimation with spectral decomposition of the covariance matrix

An alternative way to construct the “square root” of the covariance matrix is to use a spectral decomposition, such that $\Sigma_t = V_t \Lambda_t V_t'$, where V_t is the $n \times n$ matrix of eigenvectors and Λ_t is the diagonal matrix of eigenvalues. In this case, the “square root” is simply given by $\Sigma_t^{1/2} = V_t \Lambda_t^{1/2}$. The spectral decomposition is clearly different from the Cholesky decomposition. In particular, it does not assume any particular ordering of the variables in the system. In our context, we argue that the Cholesky decomposition is probably more appropriate because there is a natural ordering of the markets, given their non-synchronicity. Yet, in this section, we provide and discuss the estimates obtained with the spectral decomposition.

Table 7 reports the estimates of the ADCC model with the spectral decomposition of the covariance matrix. It reveals that the parameters pertaining to the conditional mean vector and covariance matrix are barely affected by the change in the decomposition of the covariance matrix.

The first component of the spectral decomposition can be interpreted as the US shocks. As a consequence, the responses of its higher moments to shocks are very similar to the ones obtained with the Cholesky decomposition. The second component can be interpreted as an average of the shocks affecting the two markets. Therefore, it may imply some changes in the responses of its higher moments to shocks, as compared to the Cholesky decomposition. The table shows that the degree of freedom and the asymmetry parameter are barely affected. In all cases, the changes in the parameter estimates are not significantly different from zero.

8 Estimation of the complete system

A last issue we address is the joint estimation of all the parameters of the complete system, for the four markets. As we argue in the text, the main motivation of the bivariate estimation is that the subsequent analysis of the NIS relies on the bivariate models. We show in this section that estimating all the parameters simultaneously is possible and produces essentially the same estimates as in the bivariate models (see Table 2 in the text). We retain here the Cholesky decomposition as covariance-matrix decomposition. In terms of modeling assumptions, the main difference between both approaches is that here the parameters ($\delta_1, \delta_2, \delta_3$) are the same for all the correlations.

Table 8 reports the estimates of the ADCC model when innovations are drawn

from a Sk- t distribution with time-varying shape parameters for the four markets simultaneously (equations (1) to (9)). The parameter estimates for the US, Japanese, and UK markets are essentially unaffected when all the parameters are simultaneously estimated. The only sizeable change is in the parameters pertaining to the degree of freedom and the asymmetry parameter of the German market. We also notice a significant decrease in the standard error of several parameter estimates.

9 Computation of the *GIR*

This section describes the algorithm used to compute the Generalized Impulse Response function (*GIR*). We follow an approach similar to that developed by Koop, Pesaran, and Potter (1996).

1. Simulate a history ω_{t-1} . Since the conditional distribution of z_t varies over time, we draw at each date t in a different joint distribution, whose characteristics are given by the past realization of z_{t-1} . We initialize the simulation at date $\tau = 0$ using unconditional moments. Given the unconditional skewness and kurtosis, we compute the implied degree of freedom and asymmetry parameter that gives us the conditional distribution for date $\tau = 1$ to simulate z_1 . Then, using the unconditional covariance matrix, we compute the unexpected returns ε_1 and using the unconditional mean, we finally obtain r_1 . Then, we iterate. Given z_1 , we estimate the degree of freedom and asymmetry parameter for date $\tau = 2$ and therefore simulate z_2 . Given ε_1 , we estimate the covariance matrix for date $\tau = 2$ and therefore we deduce unexpected returns ε_2 and r_2 . And so on. Following this approach, we simulate 1000 observations to obtain the state of the history ω_{t-1} just before the shock experiment.²
2. For a given horizon H , simulate J random samples of H values of the innovation process $\{z_{t+h}^j\}_{h=1, \dots, H}^{j=1, \dots, J}$. At each date $t+h$, the conditional distribution of z_{t+h} is $g(z_{t+h}|\eta_{t+h})$, where η_{t+h} is the vector of shape parameters at date $t+h$ whose dynamics is given by equations (16) and (17) depending on past innovations.
3. Define a shock event at date t as a vector $v_t = (v_{1,t}, v_{2,t})'$ that corresponds to the experiment we are interested in. For instance, $v_t = (1, 0)'$ would correspond to a shock of 1% on the US market. Then construct J samples of

²To increase the speed of convergence of the statistics resulting from the simulation, we use the technique of antithetic variates.

$(H + 1)$ realizations of $\{r_{t+h}^j\}_{h=0,\dots,H}^{j=1,\dots,J}$ based on the history ω_{t-1} and the innovations $\{z_{t+h}^j\}_{h=1,\dots,H}^{j=1,\dots,J}$ and the assumed shock v_t as innovation at date t . Similarly, estimate the $(H + 1)$ realizations of the covariance matrix and the higher moments. They are obtained by iterating on the nonlinear model. These are the samples with shock.

4. Construct J samples of $(H + 1)$ realizations of $\{\tilde{r}_{t+h}^j\}_{h=0,\dots,H}^{j=1,\dots,J}$ based on the history ω_{t-1} and the innovations $\{z_{t+h}^j\}_{h=1,\dots,H}^{j=1,\dots,J}$ and a shock $\nu_t = (0, 0)'$ as innovation at date t . Similarly, estimate the $(H + 1)$ realizations of the covariance matrix and higher moments. These are the samples without shock.
5. Compute the mean for each individual component over the J samples as

$$r_{t+h}^N = \frac{1}{J} \sum_{j=1}^J r_{t+h}^j, \quad h = 1, \dots, H$$

$$\tilde{r}_{t+h}^N = \frac{1}{J} \sum_{j=1}^J \tilde{r}_{t+h}^j, \quad h = 1, \dots, H$$

These sample means provide estimates of the conditional expectations $E[r_{t+h}|\omega_{t-1}]$ and $E[r_{t+h}|z_t, \omega_{t-1}]$, $h = 1, \dots, H$, respectively.³ The difference $\tilde{r}_{t+h}^N - r_{t+h}^N$ gives a Monte-Carlo estimate of the $GIR(h, v_t, \omega_{t-1})$. Similarly, compute the Monte-Carlo estimate of the GIR for the covariance matrix and higher moments.

6. Estimate the 50% confidence interval of the GIR using the first and third quartiles of the empirical distribution.

References

- [1] Schwarz, G. (1978), Estimating the Dimension of a Model. *Annals of Statistics*, 6, 461–464.

³Simple means sometimes provide erratic estimates of the conditional expectations given the nonlinearity of the model and the occurrence of some outliers. We experimented with trimmed means and medians and obtained essentially the same results. The GIR displayed in the Figures 8 to 11 are computed with trimmed means with 1% of the observations removed on the left and right hand sides of the distribution.

	lag p	R^2	$RSS(p)/T$	$BIC(p)$
US	0	0.0000	0.9659	-0.0347
	1	0.0024	0.9637	-0.0359
	2	0.0025	0.9638	-0.0347
	3	0.0042	0.9624	-0.0351
	4	0.0045	0.9623	-0.0341
	5	0.0046	0.9624	-0.0329
Japan	0	0.0000	1.8550	0.6179
	1	0.0005	1.8545	0.6187
	2	0.0008	1.8543	0.6197
	3	0.0009	1.8546	0.6209
	4	0.0009	1.8550	0.6222
	5	0.0010	1.8552	0.6234
UK	0	0.0000	1.3586	0.3064
	1	0.0151	1.3383	0.2925
	2	0.0154	1.3383	0.2936
	3	0.0156	1.3383	0.2946
	4	0.0168	1.3369	0.2947
	5	0.0169	1.3372	0.2960
Germany	0	0.0000	1.6968	0.5288
	1	0.0000	1.6969	0.5299
	2	0.0007	1.6960	0.5304
	3	0.0009	1.6959	0.5315
	4	0.0013	1.6957	0.5324
	5	0.0015	1.6957	0.5335

Table 1: This table reports statistics on the optimal number of lags for the dynamics of the expected returns. R^2 is the R-square of the regression, RSS/T is the standard error of residuals, BIC is the Bayesian information criterion.

	US		Japan		UK		Germany	
	Param.	Std err.	Param.	Std err.	Param.	Std err.	Param.	Std err.
Conditional mean								
μ	0.025	(0.008)	0.030	(0.011)	0.025	(0.010)	0.032	(0.007)
φ_1	0.057	(0.011)	0.024	(0.011)	0.116	(0.011)	0.014	(0.049)
Conditional variance								
ω	0.009	(0.002)	0.023	(0.004)	0.022	(0.004)	0.022	(0.004)
α	0.024	(0.005)	0.040	(0.006)	0.050	(0.007)	0.045	(0.006)
ψ	0.060	(0.009)	0.085	(0.011)	0.045	(0.009)	0.047	(0.009)
β	0.937	(0.006)	0.908	(0.007)	0.909	(0.009)	0.918	(0.005)
Shape parameters								
ν	7.313	(0.561)	6.236	(0.418)	10.355	(1.007)	8.988	(0.567)
ξ	0.986	(0.014)	1.001	(0.015)	0.937	(0.015)	0.953	(0.012)
lnL	10627.0		13126.9		12048.1		12979.0	
LR(norm.)	399.0		593.2		267.1		370.1	
$t(\nu = 4)$	5.905		5.351		6.313		8.792	
$t(\xi = 1)$	0.964		0.093		4.272		4.043	

Table 2: This table reports parameter estimates and specification tests for the univariate models with Sk- t distribution and constant shape parameters. All figures in parenthesis represent standard errors. lnL is the log-likelihood of the sample. LR_{norm} is the LR statistics for the test of the null hypothesis that the distribution is normal. 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.

Moment condition	US		Japan		UK		Germany	
	Stat.	p-val.	Stat.	p-val.	Stat.	p-val.	Stat.	p-val.
Conditional mean								
1	0.163	(0.686)	0.182	(0.670)	9.757	(0.002)	2.632	(0.105)
2	0.711	(0.399)	1.575	(0.209)	0.428	(0.513)	2.619	(0.106)
Conditional variance								
1	0.930	(0.335)	0.696	(0.404)	2.811	(0.094)	0.044	(0.834)
2	2.691	(0.101)	0.200	(0.655)	5.918	(0.015)	2.289	(0.130)
3	0.455	(0.500)	0.770	(0.380)	1.694	(0.193)	7.645	(0.006)
4	8.434	(0.004)	1.143	(0.285)	2.130	(0.145)	5.808	(0.016)
Conditional skewness								
1	2.505	(0.114)	1.173	(0.279)	0.248	(0.619)	1.451	(0.228)
2	0.860	(0.354)	1.178	(0.278)	0.017	(0.897)	0.263	(0.608)
3	2.241	(0.134)	0.496	(0.481)	1.471	(0.225)	1.478	(0.224)
4	2.485	(0.115)	0.785	(0.376)	1.615	(0.204)	0.835	(0.361)
Conditional kurtosis								
1	1.293	(0.256)	1.205	(0.272)	0.956	(0.328)	1.067	(0.302)
2	1.394	(0.238)	1.000	(0.317)	0.884	(0.347)	0.577	(0.448)
3	0.011	(0.916)	0.761	(0.383)	0.923	(0.337)	1.136	(0.287)
4	2.894	(0.089)	0.485	(0.486)	0.829	(0.363)	1.146	(0.284)
Conditional correlation								
1	—		1.200	(0.273)	0.847	(0.358)	1.365	(0.243)
2	—		0.958	(0.328)	0.620	(0.431)	0.569	(0.451)
3	—		1.288	(0.256)	1.065	(0.302)	2.505	(0.114)
4	—		0.419	(0.517)	0.624	(0.430)	1.778	(0.182)
5	—		0.843	(0.359)	0.432	(0.511)	5.757	(0.016)
6	—		0.755	(0.385)	10.228	(0.001)	18.115	(0.000)

Table 3: This table reports robust moment condition tests for the ADCC model with Sk- t distribution and time-varying shape parameters. In this table, we report the diagnostic tests for the univariate moments and conditional correlations.

Moment condition	US - Japan		US - UK		US - Germany	
	Stat.	p-val.	Stat.	p-val.	Stat.	p-val.
Conditional skewness(1,1,2)						
1	1.090	(0.297)	0.967	(0.325)	1.909	(0.167)
2	0.982	(0.322)	0.693	(0.405)	0.809	(0.368)
3	1.030	(0.310)	0.986	(0.321)	1.672	(0.196)
4	0.912	(0.340)	1.170	(0.279)	0.828	(0.363)
5	0.015	(0.901)	0.046	(0.830)	0.200	(0.655)
6	1.441	(0.230)	2.249	(0.134)	2.209	(0.137)
Conditional skewness(1,2,2)						
1	1.010	(0.315)	0.745	(0.388)	0.115	(0.735)
2	1.005	(0.316)	1.788	(0.181)	0.037	(0.848)
3	1.031	(0.310)	0.707	(0.401)	0.820	(0.365)
4	1.228	(0.268)	0.787	(0.375)	0.992	(0.319)
5	1.153	(0.283)	0.192	(0.662)	0.351	(0.554)
6	0.234	(0.629)	2.676	(0.102)	3.168	(0.075)
Conditional kurtosis(1,1,1,2)						
1	1.079	(0.299)	0.799	(0.372)	1.679	(0.195)
2	1.036	(0.309)	1.143	(0.285)	0.090	(0.765)
3	1.090	(0.297)	0.735	(0.391)	1.774	(0.183)
4	1.039	(0.308)	2.018	(0.155)	0.426	(0.514)
5	0.669	(0.414)	0.377	(0.539)	0.044	(0.835)
6	0.745	(0.388)	0.478	(0.489)	0.109	(0.741)
Conditional kurtosis(1,2,2,2)						
1	1.002	(0.317)	0.303	(0.582)	0.685	(0.408)
2	1.001	(0.317)	1.982	(0.159)	1.476	(0.224)
3	1.003	(0.317)	0.301	(0.583)	0.642	(0.423)
4	0.951	(0.330)	0.037	(0.848)	1.475	(0.225)
5	0.729	(0.393)	0.052	(0.820)	0.069	(0.793)
6	0.794	(0.373)	0.731	(0.393)	0.230	(0.632)
Conditional kurtosis(1,1,2,2)						
1	1.015	(0.314)	0.867	(0.352)	1.133	(0.287)
2	1.014	(0.314)	1.634	(0.201)	0.264	(0.607)
3	1.016	(0.313)	0.866	(0.352)	1.131	(0.288)
4	1.034	(0.309)	1.299	(0.255)	0.592	(0.442)
5	1.124	(0.289)	0.286	(0.593)	0.799	(0.371)
6	0.975	(0.324)	0.635	(0.426)	0.004	(0.948)

Table 4: This table reports robust moment condition tests for the ADCC model with Sk- t distribution and time-varying shape parameters. In this table, we report the same diagnostic tests as in Table 3 but for higher co-moment conditions.

	US		Japan		UK		Germany	
	Param.	Std err.	Param.	Std err.	Param.	Std err.	Param.	Std err.
Conditional mean								
μ	0.028	(0.011)	0.030	(0.011)	0.027	(0.008)	0.036	(0.004)
φ_1	0.061	(0.011)	0.021	(0.014)	0.065	(0.012)	-0.017	(0.018)
Conditional variance								
ω	0.007	(0.002)	0.019	(0.003)	0.021	(0.004)	0.020	(0.004)
α	0.034	(0.006)	0.036	(0.006)	0.048	(0.007)	0.046	(0.006)
ψ	0.051	(0.008)	0.076	(0.009)	0.040	(0.009)	0.037	(0.007)
β	0.939	(0.005)	0.914	(0.007)	0.912	(0.009)	0.917	(0.008)
Conditional correlation								
δ_1	—		0.9881	(0.0045)	0.9947	(0.0034)	0.9837	(0.0047)
δ_2	—		0.0068	(0.0022)	0.0020	(0.0011)	0.0101	(0.0026)
δ_3	—		0.0001	(0.0011)	0.0003	(0.0006)	0.0001	(0.0017)
Conditional degree of freedom								
$c_0(/100)$	0.036	(0.009)	0.037	(0.013)	0.020	(0.014)	0.058	(0.009)
c_1^-	-0.541	(0.255)	-0.211	(0.169)	0.552	(0.197)	-0.067	(0.192)
c_1^+	-2.176	(0.571)	-0.506	(0.153)	-0.333	(0.383)	-0.584	(0.219)
c_2	0.574	(0.103)	-0.308	(0.071)	0.576	(0.337)	0.004	(0.002)
Conditional asymmetry								
d_0	0.973	(0.011)	1.006	(0.016)	0.985	(0.010)	0.988	(0.099)
d_1^-	0.031	(0.005)	0.053	(0.029)	0.026	(0.009)	0.035	(0.120)
d_1^+	0.099	(0.022)	0.043	(0.018)	0.035	(0.014)	0.046	(0.017)
d_2	0.613	(0.092)	0.550	(0.191)	0.841	(0.069)	0.760	(0.076)
lnL	—		23445.2		22390.1		23203.0	

Table 5: This table reports parameter estimates for the ADCC model with Sk- t distribution and time-varying higher moments, when the October 1987 crash is excluded from the sample. 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. lnL is the log-likelihood of the sample.

	US		Japan		UK		Germany	
	Param.	Std err.	Param.	Std err.	Param.	Std err.	Param.	Std err.
Conditional mean								
μ	0.024	(0.007)	0.030	(0.002)	0.030	(0.006)	0.030	(0.007)
φ_1	0.112	(0.010)	0.023	(0.001)	0.023	(0.003)	0.033	(0.012)
Conditional variance								
ω	0.026	(0.003)	0.011	(0.002)	0.011	(0.002)	0.012	(0.002)
α	0.056	(0.005)	0.031	(0.005)	0.031	(0.005)	0.029	(0.005)
ψ	0.050	(0.012)	0.051	(0.006)	0.051	(0.009)	0.055	(0.009)
β	0.906	(0.008)	0.929	(0.003)	0.929	(0.004)	0.927	(0.006)
Conditional correlation								
δ_1	—		0.9934	(0.0011)	0.9934	(0.0010)	0.9918	(0.0032)
δ_2	—		0.0045	(0.0006)	0.0045	(0.0005)	0.0055	(0.0017)
δ_3	—		0.0001	(0.0006)	0.0001	(0.0006)	0.0001	(0.0013)
Conditional degree of freedom								
$c_0(/100)$	0.023	(0.006)	0.032	(0.007)	0.032	(0.007)	0.031	(0.007)
c_1^-	0.585	(0.281)	-0.449	(0.088)	-0.449	(0.172)	-0.484	(0.278)
c_1^+	-0.439	(0.184)	-1.492	(0.500)	-1.492	(0.363)	-1.415	(0.579)
c_2	0.502	(0.081)	0.565	(0.142)	0.565	(0.141)	0.540	(0.116)
Conditional asymmetry								
d_0	0.972	(0.020)	0.961	(0.012)	0.961	(0.011)	0.965	(0.010)
d_1^-	0.026	(0.020)	0.017	(0.005)	0.017	(0.009)	0.028	(0.026)
d_1^+	0.059	(0.030)	0.120	(0.023)	0.120	(0.020)	0.118	(0.025)
d_2	0.769	(0.101)	0.551	(0.127)	0.551	(0.050)	0.615	(0.097)
lnL	—		22391.0		22391.0		23320.7	

Table 6: This table reports parameter estimates for the ADCC model with Sk- t distribution and time-varying higher moments, when the inverse ordering is used for the Cholesky decomposition of the covariance matrix. 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. lnL is the log-likelihood of the sample.

	US		Japan		UK		Germany	
	Param.	Std err.	Param.	Std err.	Param.	Std err.	Param.	Std err.
Conditional mean								
μ	0.030	(0.009)	0.029	(0.009)	0.027	(0.004)	0.035	(0.014)
φ_1	0.061	(0.011)	0.037	(0.011)	0.069	(0.011)	-0.016	(0.017)
Conditional variance								
ω	0.009	(0.002)	0.023	(0.004)	0.023	(0.004)	0.021	(0.004)
α	0.034	(0.007)	0.045	(0.008)	0.053	(0.008)	0.053	(0.008)
ψ	0.057	(0.010)	0.079	(0.009)	0.043	(0.008)	0.038	(0.008)
β	0.931	(0.007)	0.907	(0.009)	0.908	(0.008)	0.916	(0.008)
Conditional correlation								
δ_1	—		0.9918	(0.0047)	0.9851	(0.0076)	0.9800	(0.0057)
δ_2	—		0.0028	(0.0024)	0.0051	(0.0020)	0.0127	(0.0028)
δ_3	—		0.0003	(0.0011)	0.0001	(0.0013)	0.0001	(0.0020)
Conditional degree of freedom								
$c_0(/100)$	0.041	(0.014)	-0.012	(0.005)	0.020	(0.024)	0.039	(0.011)
c_1^-	-0.769	(0.415)	0.085	(0.322)	0.603	(0.261)	0.023	(0.006)
c_1^+	-2.325	(0.795)	-0.253	(0.248)	-0.351	(0.909)	-0.834	(0.362)
c_2	0.539	(0.097)	0.830	(0.143)	0.550	(0.523)	0.257	(0.205)
Conditional asymmetry								
d_0	0.968	(0.019)	1.005	(0.015)	0.981	(0.005)	0.988	(0.013)
d_1^-	0.033	(0.017)	0.050	(0.022)	0.030	(0.043)	0.049	(0.023)
d_1^+	0.116	(0.036)	0.039	(0.023)	0.042	(0.013)	0.048	(0.023)
d_2	0.608	(0.115)	0.475	(0.107)	0.805	(0.089)	0.720	(0.103)
lnL	—		23678.8		22416.7		23213.4	

Table 7: This table reports parameter estimates for the ADCC model with Sk- t distribution and time-varying higher moments, when a spectral decomposition of the covariance matrix is used. 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. lnL is the log-likelihood of the sample.

	US		Japan		UK		Germany	
	Param.	Std err.	Param.	Std err.	Param.	Std err.	Param.	Std err.
Conditional mean								
μ	0.027	(0.005)	0.031	(0.008)	0.025	(0.008)	0.029	(0.007)
φ_1	0.076	(0.012)	0.033	(0.003)	0.047	(0.008)	0.002	(0.014)
Conditional variance								
ω	0.010	(0.002)	0.025	(0.003)	0.029	(0.005)	0.026	(0.004)
α	0.033	(0.004)	0.045	(0.008)	0.059	(0.011)	0.029	(0.005)
ψ	0.054	(0.008)	0.087	(0.006)	0.040	(0.014)	0.031	(0.005)
β	0.938	(0.005)	0.910	(0.005)	0.915	(0.008)	0.908	(0.010)
Conditional correlation								
δ_1	0.9913	(0.0010)	—	—	—	—	—	—
δ_2	0.0054	(0.0005)	—	—	—	—	—	—
δ_3	0.0006	(0.0005)	—	—	—	—	—	—
Conditional degree of freedom								
$c_0(/100)$	0.039	(0.004)	0.031	(0.014)	0.021	(0.004)	0.174	(0.019)
c_1^-	-0.588	(0.050)	-0.301	(0.154)	0.444	(0.672)	-0.108	(0.024)
c_1^+	-1.921	(0.289)	-0.530	(0.117)	-0.391	(0.198)	0.170	(0.149)
c_2	0.519	(0.121)	-0.258	(0.398)	0.536	(0.093)	-0.822	(0.059)
Conditional asymmetry								
d_0	0.964	(0.011)	1.010	(0.010)	0.986	(0.010)	0.933	(0.020)
d_1^-	0.026	(0.011)	0.066	(0.009)	0.012	(0.010)	-0.006	(0.001)
d_1^+	0.120	(0.022)	0.046	(0.007)	0.029	(0.014)	0.084	(0.020)
d_2	0.584	(0.113)	0.475	(0.106)	0.898	(0.058)	-0.322	(0.153)
lnL	46702.4		—	—	—	—	—	—

Table 8: This table reports parameter estimates for the ADCC model with Sk- t distribution and time-varying higher moments for the four markets simultaneously. All figures in parenthesis represent standard errors. lnL is the log-likelihood of the sample.