Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

Extremal Dependence Tests for Contagion

CAMA Working Paper 40/2015
November 2015

Revision - CAMA Working Paper 38/2014

Renée Fry-McKibbin
Crawford School of Public Policy, ANU and
Centre for Applied Macroeconomic Analysis (CAMA), ANU

Cody Yu-Ling Hsiao
School of Business, Macau University of Science and Technology and
Centre for Applied Macroeconomic Analysis (CAMA), ANU

Abstract

A new test for financial market contagion based on changes in extremal dependence defined as co-kurtosis and co-volatility is developed to identify the propagation mechanism of shocks across international financial markets. The proposed approach captures changes in various aspects of the asset return relationships such as cross-market mean and skewness (co-kurtosis) as well as cross-market volatilities (co-volatility). Monte Carlo experiments show that the tests perform well except for when crisis periods are short in duration. Small crisis sample critical values are calculated for use in this case. In an empirical application involving the global financial crisis of 2008-09, the results show that significant contagion effects are widespread from the US banking sector to global equity markets and banking sectors through either the co-kurtosis or the co-volatility channels, reinforcing that higher order moments matter during crises.

Keywords

Co-skewness, Co-kurtosis, Co-volatility, Contagion testing, Extremal dependence, Financial crisis, Lagrange multiplier tests.

JEL Classification

C12, F30, G11, G21

Address for correspondence:

(E) cama.admin@anu.edu.au

Extremal Dependence Tests for Contagion[*]

Renée Fry-McKibbin[#,†] and Cody Yu-Ling Hsiao[^,#]

#Centre for Applied Macroeconomic Analysis (CAMA)

†Crawford School of Public Policy, Australian National University

^School of Business, Macau University of Science and Technology

October 6, 2015

Abstract

A new test for financial market contagion based on changes in extremal dependence defined as co-kurtosis and co-volatility is developed to identify the propagation mechanism of shocks across international financial markets. The proposed approach captures changes in various aspects of the asset return relationships such as cross-market mean and skewness (co-kurtosis) as well as cross-market volatilities (co-volatility). Monte Carlo experiments show that the tests perform well except for when crisis periods are short in duration. Small crisis sample critical values are calculated for use in this case. In an empirical application involving the global financial crisis of 2008-09, the results show that significant contagion effects are widespread from the US banking sector to global equity markets and banking sectors through either the co-kurtosis or the co-volatility channels, reinforcing that higher order moments matter during crises.

Key words: Co-skewness, Co-kurtosis, Co-volatility, Contagion testing, Extremal dependence, Financial crisis, Lagrange multiplier tests.
JEL Classification: C12, F30, G11, G21

1 Introduction

This paper derives a new class of contagion tests for application to financial crises defined broadly by a significant change in extremal dependence between two markets in a non-crisis and a crisis period. Financial crisis episodes are characterized by asset returns falling in value and volatility increasing compared to non-crisis periods. Under the mean-variance framework, this is consistent with investors realizing a high level of excess returns in a non-crisis period in exchange for a high level of risk in a crisis period (Sharpe, 1964; Lintner, 1965; Black, 1972). However, it is evident that return distributions are often asymmetric and fat-tailed suggesting the limitations of the Gaussian distribution for modelling asset returns (Kraus and Litzenberger, 1976; Friend and Westerfield, 1980; Engle et al., 1990; Harvey and Siddique, 2000).

In addition to the changeable properties of the distribution of univariate asset returns, market dependence measured by the co-moments of asset returns is subject to change through the contagious effects of financial crises. Co-moment changes are already the basis of different types of tests for contagion including the correlation test of Forbes and Rigobon (2002) and the co-skewness test of Fry et al. (2010).[1] The asymmetries and fat tails of return distributions mean that linear co-movement may not fully capture market dependence since small and large returns are equally weighted in its calculation (Embrechts et al., 2003), and is a complete measure only if investors display mean-variance preferences. The contagion measures derived in this paper based on extremal dependence changes capture more co-movement than the linear dependence measures. Extremal dependence is measured by co-kurtosis (the relationship between the asset return in market i and return skewness in market j) and co-volatility (the relationship between the return volatility of markets i and j).[2] The extremal dependence measures are similar to the exceedance and co-exceedance tests for contagion developed in papers such as Favero and Giavazzi (2002), Bae et al. (2003), Pesaran and Pick (2007), Diebold and Yilmaz (2009) and Gropp et al. (2010). The non-parametric tests of Rodriguez (2007), Busetti and Harvey (2011) and Garcia and Tsafack (2011) also illustrate the role of the tail dependence coefficient in a range of copula models.

---

[1] In the simplest case contagion is defined as a significant increase in correlation (linear dependence) between two markets during a crisis period compared to a non-crisis period after controlling for market fundamentals (Forbes and Rigobon, 2002).

[2] It should be highlighted that like the correlation and co-skewness contagion tests existing in the literature, the extremal dependence tests are all adjusted for heteroskedasticity by scaling the statistics by a non-linear function of the percentage change in volatility of the source market returns across the regimes.

Asymmetry in the returns distribution measured by skewness generated by the risk preferences of investors is subject to change in different regimes, particularly during a crisis period (Black, 1976; Shleifer and Vishny, 1997; Bekaert and Wu, 2000; Fry et al. 2010; Conrad et al., 2013). Asset returns can also yield leptokurtic behavior with kurtosis rising during crisis periods. The relatively lower kurtosis commonly displayed in non-crisis periods is documented theoretically by Brunnermeier and Pedersen (2008). They find that speculators invest in securities with a positive average return and negative skewness, giving rise to low values of kurtosis. However, extreme events result in speculators investing in securities with a negative average return and negative skewness, thus increasing kurtosis risk but providing a good hedge during the crisis period.[3] Brunnermeier and Pedersen (2008) and Fry et al. (2010) imply different signs of higher order co-moments in the two regimes. It is possible that either sign could eventuate in a crisis period as investors optimize under incomplete information (Vaugirard, 2007; Gorton and Metrick, 2012) and liquidity constraints (Allen and Gale, 2000; Cifuentes et al., 2005), particularly when the source crisis country is also considered a safe haven country (Vayanos, 2004).

To illustrate the new approach the tests are applied to equity markets and banking sectors during the global financial crisis of 2008-09. Table 1 illustrates the co-moment statistics of the equity returns data used in the empirical application for the non-crisis and crisis periods. Each co-moment captures different features of the joint asset return relationship including: i) linear dependence; ii) asymmetric dependence; and iii) extremal dependence. The table shows that in line with the theories touched on earlier, the three types of market dependence clearly change in terms of magnitudes and quite often signs. These statistics emphasize the importance of modelling asymmetric and tail risks.

The rest of the paper proceeds as follows. Section 2 summarizes the investors risk return trade-offs between the expected excess return, and higher order moments and co-moments. Section 3 describes the properties of a bivariate generalized exponential family of distributions with asymmetry and leptokurtosis, which provides the framework for developing tests of co-kurtosis and co-volatility which are at the heart of the tests of contagion based on changes in extremal dependence presented Section 4. Section 5 analyses the finite sample properties of the new tests and compares them to the original co-skewness test previously

---

[3]Brunnermeier and Pedersen (2008) show that the funding constraint influences not only the price level but also the skewness of the price distribution. In extreme events the security price is below the market fundamental price, resulting in negative returns. Holding the security leads to losses as speculators face funding constraints, inducing negative skewness in the price distribution.

derived in Fry et al. (2010). It also presents small sample critical value adjustments for periods when crises durations are short. An illustrative example is presented in Section 6, with the results showing that the extremal dependence contagion tests capture more market co-movements than the asymmetric dependence tests during crises. Section 7 concludes.

2 Higher Order Co-moments and Risky Assets

In standard theory of portfolio choice investors achieve optimal asset allocation by maximizing expected utility subject to the variance of the portfolio in the mean-variance framework (Sharpe, 1964; Lintner, 1965; Black, 1972). To take asymmetries and tail risks into account, the traditional CAPM model can be augmented with higher order moments and co-moments (Fry et al. 2010 and Martellini and Ziemann, 2010). Appendix A.1 contains an expression for the excess returns over the risk free rate R_f for a risky asset R_1 under the scenario of higher order co-moments with respect to asset R_2. The expression is derived from a CAPM model with correlation, co-skewness, co-kurtosis and co-volatility and is

$$
\begin{aligned}
E(R_1) - R_f &= \theta_1 E\left[(R_1 - \mu_1)^2\right] + \theta_2 E\left[(R_2 - \mu_2)^2\right] + \theta_3 E\left[(R_1 - \mu_1)(R_2 - \mu_2)\right] \\
&\quad + \theta_4 E\left[(R_1 - \mu_1)^3\right] + \theta_5 E\left[(R_2 - \mu_2)^3\right] + \theta_6 E\left[(R_1 - \mu_1)^2 (R_2 - \mu_2)\right] \\
&\quad + \theta_7 E\left[(R_1 - \mu_1)(R_2 - \mu_2)^2\right] + \theta_8 E\left[(R_1 - \mu_1)^4\right] + \theta_9 E\left[(R_2 - \mu_2)^4\right] \\
&\quad + \theta_{10} E\left[(R_1 - \mu_1)^3 (R_2 - \mu_2)\right] + \theta_{11} E\left[(R_1 - \mu_1)(R_2 - \mu_2)^3\right] \\
&\quad + \theta_{12} E\left[(R_1 - \mu_1)^2 (R_2 - \mu_2)^2\right].
\end{aligned}
\tag{1}
$$

The parameters in (1) are

$$
\begin{array}{lll}
\theta_1 = \alpha_1^2\left(\frac{\partial E[U(W)]}{\partial \sigma_p^2}\right), & \theta_5 = \alpha_2^3\left(\frac{\partial E[U(W)]}{\partial s_P^3}\right), & \theta_9 = \alpha_2^4\left(\frac{\partial E[U(W)]}{\partial k_P^4}\right), \\
\theta_2 = \alpha_2^2\left(\frac{\partial E[U(W)]}{\partial \sigma_p^2}\right), & \theta_6 = 3\alpha_1^2\alpha_2\left(\frac{\partial E[U(W)]}{\partial s_P^3}\right), & \theta_{10} = 4\alpha_1^3\alpha_2\left(\frac{\partial E[U(W)]}{\partial k_P^4}\right), \\
\theta_3 = 2\alpha_1\alpha_2\left(\frac{\partial E[U(W)]}{\partial \sigma_p^2}\right), & \theta_7 = 3\alpha_1\alpha_2^2\left(\frac{\partial E[U(W)]}{\partial s_P^3}\right), & \theta_{11} = 4\alpha_1\alpha_2^3\left(\frac{\partial E[U(W)]}{\partial k_P^4}\right), \\
\theta_4 = \alpha_1^3\left(\frac{\partial E[U(W)]}{\partial s_P^3}\right), & \theta_8 = \alpha_1^4\left(\frac{\partial E[U(W)]}{\partial k_P^4}\right), & \theta_{12} = 6\alpha_1^2\alpha_2^2\left(\frac{\partial E[U(W)]}{\partial k_P^4}\right),
\end{array}
\tag{2}
$$

where $E[U(W)]$ is the expected utility of the return on investment, and the variance, skew-

ness and kurtosis of the end of period returns can be decomposed respectively into

$$
\begin{aligned}
\sigma_p^2 &= E\left[(R_p - \mu_p)^2\right] \\
&= E\left[\left(\sum_{i=1}^{N=2}\alpha_i\left(R_i - \mu_i\right)\right)^2\right] \\
&= \alpha_1^2 E\left[(R_1 - \mu_1)^2\right] + \alpha_2^2 E\left[(R_2 - \mu_2)^2\right] + 2\alpha_1\alpha_2 E\left[(R_1 - \mu_1)(R_2 - \mu_2)\right],
\end{aligned}
\tag{3}
$$

$$
\begin{aligned}
s_P^3 &= E\left[\left(\sum_{i=1}^{N=2}\alpha_i\left(R_i - \mu_i\right)\right)^3\right] \\
&= \alpha_1^3 E\left[(R_1 - \mu_1)^3\right] + \alpha_2^3 E\left[(R_2 - \mu_2)^3\right] + \\
&\quad 3\alpha_1^2\alpha_2 E\left[(R_1 - \mu_1)^2(R_2 - \mu_2)\right] + 3\alpha_1\alpha_2^2 E\left[(R_1 - \mu_1)(R_2 - \mu_2)^2\right],
\end{aligned}
\tag{4}
$$

$$
\begin{aligned}
k_P^4 &= E\left[\left(\sum_{i=1}^{N=2}\alpha_i\left(R_i - \mu_i\right)\right)^4\right] \\
&= \alpha_1^4 E\left[(R_1 - \mu_1)^4\right] + \alpha_2^4 E\left[(R_2 - \mu_2)^4\right] + 4\alpha_1^3\alpha_2 E\left[(R_1 - \mu_1)^3(R_2 - \mu_2)\right] \\
&\quad + 4\alpha_1\alpha_2^3 E\left[(R_1 - \mu_1)(R_2 - \mu_2)^3\right] + 6\alpha_1^2\alpha_2^2 E\left[(R_1 - \mu_1)^2(R_2 - \mu_2)^2\right].
\end{aligned}
\tag{5}
$$

Equation (1) decomposes the expected excess return for the risky asset 1 in terms of risk prices and risk quantities. The risk prices (the θ_i in (1)) are expressed in terms of the various risk aversion measures arising from the utility function of the investor ($\left(\frac{\partial E[U(W)]}{\partial \sigma_p^2}\right)$, $\left(\frac{\partial E[U(W)]}{\partial s_P^3}\right)$ and $\left(\frac{\partial E[U(W)]}{\partial k_P^4}\right)$) and the shares of the asset in the portfolio (α_i). The risk quantities contain the second moment terms of the variance and covariance, the third moment terms of skewness ($E\left[(R_1 - \mu_1)^3\right]$ and $E\left[(R_2 - \mu_2)^3\right]$) and co-skewness ($E\left[(R_1 - \mu_1)^2(R_2 - \mu_2)\right]$ and $E\left[(R_1 - \mu_1)(R_2 - \mu_2)^2\right]$), as well as the fourth moment terms of kurtosis ($E\left[(R_1 - \mu_1)^4\right]$ and $E\left[(R_2 - \mu_2)^4\right]$), co-kurtosis ($E\left[(R_1 - \mu_1)^3(R_2 - \mu_2)\right]$ and $E\left[(R_1 - \mu_1)(R_2 - \mu_2)^3\right]$) and co-volatility ($E\left[(R_1 - \mu_1)^2(R_2 - \mu_2)^2\right]$).

To illustrate the properties of the expected excess return for a risky asset, the risk-return trade-off surfaces between the expected excess return, variance, skewness and kurtosis, are presented in Figure 1. The figure is generated by simulating R_1 in (1) in terms of second,

third and fourth order moments and co-moments. The parameters values of the simulation are loosely based on the data in Section 6 and are

$$\theta_1 = 0.5, \theta_2 = 0.7, \theta_3 = 2.0, \theta_4 = \theta_5 = \theta_6 = \theta_7 = -1.5, \quad (6)$$
$$\theta_8 = \theta_9 = 4, \theta_{10} = \theta_{11} = \theta_{12} = 1.5,$$

with the values of co-moments given by

$$E\left[(R_1 - \mu_1)(R_2 - \mu_2)\right] = 0.80, E\left[(R_1 - \mu_1)^2 (R_2 - \mu_2)\right] = 0.00, \quad (7)$$
$$E\left[(R_1 - \mu_1)(R_2 - \mu_2)^2\right] = 0.00, E\left[(R_1 - \mu_1)(R_2 - \mu_2)^3\right] = 1.47,$$
$$E\left[(R_1 - \mu_1)^3 (R_2 - \mu_2)\right] = 1.46, E\left[(R_1 - \mu_1)^2 (R_2 - \mu_2)^2\right] = 2.20.$$

The parameters $\theta_4, \ldots, \theta_7$ are restricted to have a negative sign due to the assumption that a risk-averse investor has a utility function with decreasing absolute risk aversion, while the other parameters $\theta_8, \ldots, \theta_{12}$ are set up to have a positive sign due to the utility function exhibiting decreasing absolute prudence (Scott and Horvath, 1980). The skewness and kurtosis value ranges are

$$E[(R_1 - \mu_1)]^3 = E[(R_2 - \mu_2)]^3 = [-1, 1] \quad (8)$$
$$E[(R_1 - \mu_1)]^4 = E[(R_2 - \mu_2)]^4 = [4, 9].$$

Panel A of Figure 1 presents the mean-skewness-kurtosis surface for a case where there is no volatility ($E\left[(R_1 - \mu_1)^2\right] = 0$ and $E\left[(R_2 - \mu_2)^2\right] = 0$) and panel B for a case where there is volatility ($E\left[(R_1 - \mu_1)^2\right] = 2$ and $E\left[(R_2 - \mu_2)^2\right] = 2$). Given any level of volatility, there is a positive relationship between the expected excess return and kurtosis while the relationship between skewness and kurtosis is negative. Comparing panels A and B, an investor needs to be compensated for higher risk (volatility) with higher expected excess returns. Given any level of kurtosis, there is a negative trade-off between the expected excess return and skewness. These findings suggest that an investor requires a higher expected excess return for taking more volatility and kurtosis risks; whereas, an investor also realizes a lower expected excess return for the benefit of positive skewness.[4]

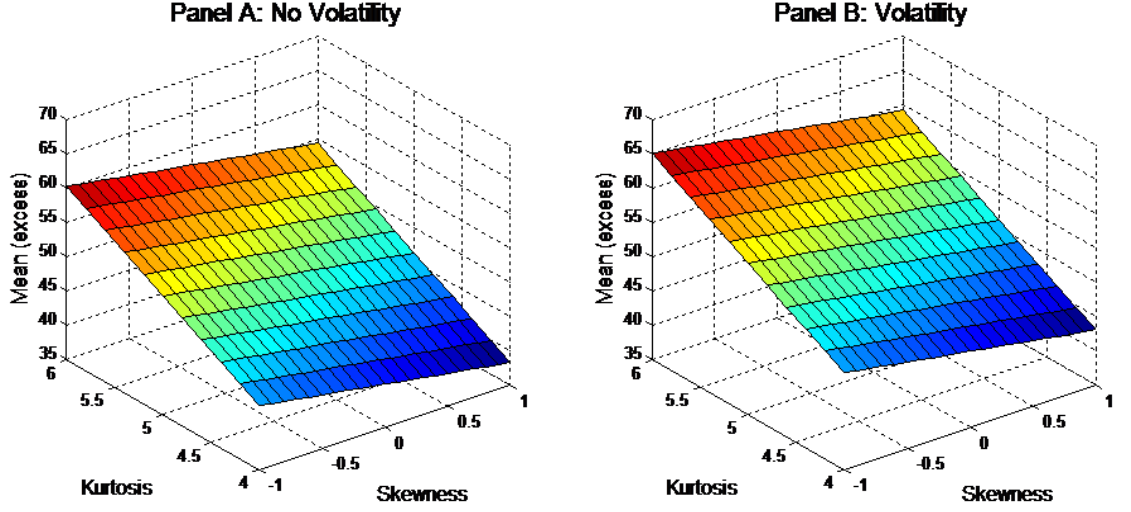[4]Fang and Lai (1997) and Hwang and Satchell (1999) show that the expected excess return for the risky

Figure 1: Mean, volatility, skewness and kurtosis trade-offs. Panel A presents a mean-skewness-kurtosis surface for the case of no volatility $(E\left[(R_1 - \mu_1)^2\right] = 0$ and $E\left[(R_2 - \mu_2)^2\right] = 0)$. Panel B shows a mean-skewness-kurtosis surface for the case of volatility $(E\left[(R_1 - \mu_1)^2\right] = 2$ and $E\left[(R_2 - \mu_2)^2\right] = 2)$. Equations (6) to (8) summarize the parameters of the moments and co-moments chosen for equation (1) which generates these surfaces.

3   Co-kurtosis and Co-volatility Statistics

To develop the statistics of higher order co-moments, a non-normal multivariate returns distribution is specified in order to model the distribution with asymmetry and leptokurtosis. Building on Cobb et al. (1983), Lye and Martin (1993) and Fry et al. (2010), the bivariate generalized exponential family of distributions with first to fourth order moments and co-moments is used for developing the co-moment statistics. Defining $r = \{r_1, r_2\}$ as a 2-dimensional vector of random variables, the distribution is expressed as

$$
\begin{aligned}
f(r) &= \exp(h - \eta) && (9) \\
&= \exp\left(\sum_{i=1}^{12} \theta_i g_i(r) - \eta\right) \\
&= \exp\left(\theta_1 r_1^2 + \theta_2 r_2^2 + \theta_3 r_1 r_2 + \theta_4 r_1 r_2^2 + \theta_5 r_1^2 r_2 + \theta_6 r_1^2 r_2^2 \right. \\
&\quad \left. + \theta_7 r_1^1 r_2^3 + \theta_8 r_1^3 r_2^1 + \theta_9 r_1^3 + \theta_{10} r_2^3 + \theta_{11} r_1^4 + \theta_{12} r_2^4 - \eta\right),
\end{aligned}
$$

asset is associated not only with volatility but also with skewness and kurtosis. There is also evidence that a four-moment CAPM is able to price the cross-moments of returns better than the traditional CAPM (Dittmar, 2002; Fry-McKibbin et al., 2014).

where h is a function of parameters θ_i and $g_i(r)$. The choices for $g_i(r)$ are polynomials and cross-products in the elements of r which are assumed to follow an independent bivariate normal distribution, and η is a normalizing constant denoted as

$$
\begin{aligned}
\eta = \ & \ln \int \int \exp \big(\theta_1 r_1^2 + \theta_2 r_2^2 + \theta_3 r_1 r_2 + \theta_4 r_1 r_2^2 + \theta_5 r_1^2 r_2 + \theta_6 r_1^2 r_2^2 \\
& + \theta_7 r_1^1 r_2^3 + \theta_8 r_1^3 r_2^1 + \theta_9 r_1^3 + \theta_{10} r_2^3 + \theta_{11} r_1^4 + \theta_{12} r_2^4 \big) \, dr_1 dr_2.
\end{aligned}
\tag{10}
$$

In terms of equation (9), the parameters θ_1 and θ_2 control the variances of assets 1 and 2 respectively. The parameter θ_3 controls the degree of linear dependence (correlation) in the relationship between assets 1 and 2 (r_1 and r_2). The parameters θ_4 and θ_5 measure the asymmetry of the probability distribution of the two assets, capturing dependence links between the first moment of asset 1 (r_1) and the second moment of asset 2 (r_2^2) and between the second moment of asset 1 (r_1^2) and the first moment of asset 2 (r_2). These parameters represent the co-skewness coefficients. The coefficients of co-volatility and co-kurtosis (the fourth co-moments) measure the extent to which observations tend to have relatively large frequencies around the centre and in the tails of the joint distribution. These parameters (θ_6, θ_7 and θ_8) capture the interaction between the second moment of assets 1 and 2 (r_1^2 and r_2^2), between the first moment and third moment of assets 1 and 2 (r_1^1 and r_2^3) and between the third moment and the first moment of assets 1 and 2 (r_1^3 and r_2^1). The parameter θ_6 denotes the co-volatility coefficient and the parameters θ_7 and θ_8 denote the co-kurtosis coefficients. The parameters θ_9 and θ_{10} as well as θ_{11} and θ_{12} control the skewness (r_1^3 and r_2^3) and kurtosis (r_1^4 and r_2^4) for assets 1 and 2 respectively.

The Lagrange multiplier test is adopted to develop the statistics of co-kurtosis and co-volatility as the bivariate generalized exponential family of the distribution in equation (9) is nested in the bivariate normal distribution by setting the restrictions $\theta_4 = \ldots = \theta_{12} = 0$.

Consider a sample of size T from the bivariate generalized exponential family of the distribution with a finite number K of unknown parameters $\theta = (\theta_1, ..., \theta_K)'$ summarizing the moments of a log likelihood function $\ln L_t(\theta) = h - \eta$ in equation (9) where $h = \sum_{i=1}^{K} \theta_i g_i(r)$ and η is the normalizing constant respectively. The hypothesis to be tested is specified as

$$
H_0 : \theta_4 = \cdots = \theta_p = 0; p \leq K.
\tag{11}
$$

Let $\widehat{\theta}$ be the maximum likelihood estimator of θ. The Lagrange multiplier test statistic is

$$LM = q\left(\widehat{\theta}\right)' I\left(\widehat{\theta}\right)^{-1} q\left(\widehat{\theta}\right),\tag{12}$$

which is asymptotically distributed as the chi-squared with p degrees of freedom

$$LM \xrightarrow{d} \chi_p^2.\tag{13}$$

Here, $q\left(\widehat{\theta}\right)$ is the score function evaluated at $\widehat{\theta}$ given by

$$q\left(\widehat{\theta}\right) = \left(\frac{\partial \ln L(\theta)}{\partial \theta}\right)\Big|_{\theta = \widehat{\theta}},\tag{14}$$

and $I\left(\widehat{\theta}\right)$ is the asymptotic information matrix evaluated at $\widehat{\theta}$, that is

$$I\left(\widehat{\theta}\right) = T\left(E\left[\frac{\partial h}{\partial \theta}\frac{\partial h}{\partial \theta'}\right]\Big|_{\theta = \widehat{\theta}} - E\left[\frac{\partial h}{\partial \theta}\right]\Big|_{\theta = \widehat{\theta}} E\left[\frac{\partial h}{\partial \theta'}\right]\Big|_{\theta = \widehat{\theta}}\right).\tag{15}$$

The proof of the asymptotic information matrix is shown in Appendix A.2.

Consider the restricted model which is the bivariate generalized normal distribution with co-volatility,

$$
\begin{aligned}
f(r_{1,t}, r_{2,t}) \;=\; & \exp\left[-\frac{1}{2}\left(\frac{1}{1-\rho^2}\right)\left(\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^2 + \left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)\right)\right.\\
& \left. +\theta_6\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^2\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^2 - \eta\right],
\end{aligned}\tag{16}
$$

where

$$
\begin{aligned}
\eta \;=\; & \ln \iint \exp\left[\left(\frac{1}{1-\rho^2}\right)\left(\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^2 + \left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)\right)\right.\\
& \left. +\theta_6\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^2\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^2\right] dr_1 dr_2.
\end{aligned}\tag{17}
$$

8

A test of the restriction for the hypothesis of normality is given by

$$H_0 : \theta_6 = 0, \tag{18}$$

which constitutes a test of co-volatility.

If the expression of the fourth order co-moment term, $\theta_6 \left(\frac{r_{1,t} - \mu_1}{\sigma_1} \right)^2 \left(\frac{r_{2,t} - \mu_2}{\sigma_2} \right)^2$, in equations (16) and (17) is replaced with the expression $\theta_7 \left(\frac{r_{1,t} - \mu_1}{\sigma_1} \right)^1 \left(\frac{r_{2,t} - \mu_2}{\sigma_2} \right)^3$ or $\theta_8 \left(\frac{r_{1,t} - \mu_1}{\sigma_1} \right)^3 \left(\frac{r_{2,t} - \mu_2}{\sigma_2} \right)^1$, these are a test for normality. The hypothesis is set up as

$$H_0 : \theta_p = 0, \ p = 7, 8,$$

which constitutes a test of the first form of co-kurtosis ($\theta_7 = 0$) and a test of the second form of co-kurtosis ($\theta_8 = 0$) respectively. Under the null hypothesis, the maximum likelihood estimators of the unknown parameters under the restricted model in equation (16) are

$$\widehat{\mu}_i = \frac{1}{T} \sum_{t=1}^{T} r_{i,t}; \widehat{\sigma}_i^2 = \frac{1}{T} \sum_{t=1}^{T} (r_{i,t} - \widehat{\mu}_i)^2; \widehat{\rho} = \frac{1}{T} \sum_{t=1}^{T} \left(\frac{r_{1,t} - \widehat{\mu}_1}{\widehat{\sigma}_1} \right) \left(\frac{r_{2,t} - \widehat{\mu}_2}{\widehat{\sigma}_2} \right); i = 1, 2. \tag{19}$$

The Lagrange multiplier statistics for co-volatility (LM_1) and for co-kurtosis (LM_2 and LM_3) used to test for extremal dependence are

$$
\begin{aligned}
LM_1 &= \frac{1}{T \left(4\widehat{\rho}^4 + 16\widehat{\rho}^2 + 4 \right)} \left[\sum_{t=1}^{T} \left(\frac{r_{1,t} - \widehat{\mu}_1}{\widehat{\sigma}_1} \right)^2 \left(\frac{r_{2,t} - \widehat{\mu}_2}{\widehat{\sigma}_2} \right)^2 - T \left(1 + 2\widehat{\rho}^2 \right) \right]^2 \\
LM_2 &= \frac{1}{T \left(18\widehat{\rho}^2 + 6 \right)} \left[\sum_{t=1}^{T} \left(\frac{r_{1,t} - \widehat{\mu}_1}{\widehat{\sigma}_1} \right)^1 \left(\frac{r_{2,t} - \widehat{\mu}_2}{\widehat{\sigma}_2} \right)^3 - T \left(3\widehat{\rho} \right) \right]^2 \\
LM_3 &= \frac{1}{T \left(18\widehat{\rho}^2 + 6 \right)} \left[\sum_{t=1}^{T} \left(\frac{r_{1,t} - \widehat{\mu}_1}{\widehat{\sigma}_1} \right)^3 \left(\frac{r_{2,t} - \widehat{\mu}_2}{\widehat{\sigma}_2} \right)^1 - T \left(3\widehat{\rho} \right) \right]^2.
\end{aligned}
\tag{20}
$$

The derivations for the test statistics of co-volatility and co-kurtosis are shown in Appendix A.3. Under the null hypothesis in equation (11), LM_1, LM_2 and LM_3 are distributed asymptotically as χ_1^2.

4 Higher Order Co-moment Contagion Tests

The focus of this section is to develop the extremal dependence contagion tests based on a change in the non-linear dependence between assets. In deriving the contagion tests, the following notation is used. Let x and y denote the non-crisis and crisis periods, respectively. T_x and T_y are the sample sizes of the non-crisis and crisis periods respectively. $T = T_x + T_y$ is the sample size of the full period. Then, the sample correlation coefficient during the non-crisis period (low-volatility) is ρ_x and during the crisis period (high-volatility) is ρ_y. Let i denote the source crisis asset market and j denote the recipient market of contagion. $\widehat{\mu}_{xi}$, $\widehat{\mu}_{xj}$, $\widehat{\mu}_{yi}$ and $\widehat{\mu}_{yj}$ are the sample means of the asset returns for i and j during the non-crisis and crisis periods and $\widehat{\sigma}_{xi}$, $\widehat{\sigma}_{xj}$, $\widehat{\sigma}_{yi}$ and $\widehat{\sigma}_{yj}$ are the sample standard deviations of the asset returns for i and j during the non-crisis and crisis periods. Before developing the extremal dependence contagion tests, the co-skewness contagion test of Fry et al. (2010) which is the precursor to the extremal dependence contagion test is briefly summarized.

4.1 Asymmetric Dependence Tests

The aim of the asymmetric dependence tests of contagion is to identify whether there is a statistically significant change in co-skewness between the non-crisis and crisis period. The first form (CS_{12}) is to test for contagion where the shocks transmit from a change in the returns of a source market i to the volatility of asset returns of a recipient market j. The second form (CS_{21}) is to test for contagion where the shocks spread from the volatility of asset returns of a source market i to the asset returns of a recipient market j. The tests are

$$CS_{12}(i \rightarrow j; r_i^1, r_j^2) = \left(\frac{\widehat{\psi}_y\left(r_i^1, r_j^2\right) - \widehat{\psi}_x\left(r_i^1, r_j^2\right)}{\sqrt{\frac{4\widehat{v}_{y|x_i}^2 + 2}{T_y} + \frac{4\widehat{\rho}_x^2 + 2}{T_x}}} \right)^2 \tag{21}$$

$$CS_{21}(i \rightarrow j; r_i^2, r_j^1) = \left(\frac{\widehat{\psi}_y\left(r_i^2, r_j^1\right) - \widehat{\psi}_x\left(r_i^2, r_j^1\right)}{\sqrt{\frac{4\widehat{v}_{y|x_i}^2 + 2}{T_y} + \frac{4\widehat{\rho}_x^2 + 2}{T_x}}} \right)^2 , \tag{22}$$

where

$$\widehat{\psi}_y(r_i^m, r_j^n) = \frac{1}{T_y} \sum_{t=1}^{T_y} \left(\frac{y_{i,t} - \widehat{\mu}_{yi}}{\widehat{\sigma}_{yi}} \right)^m \left(\frac{y_{j,t} - \widehat{\mu}_{yj}}{\widehat{\sigma}_{yj}} \right)^n$$

$$\widehat{\psi}_x(r_i^m, r_j^n) = \frac{1}{T_x} \sum_{t=1}^{T_x} \left(\frac{x_{i,t} - \widehat{\mu}_{xi}}{\widehat{\sigma}_{xi}} \right)^m \left(\frac{x_{j,t} - \widehat{\mu}_{xj}}{\widehat{\sigma}_{xj}} \right)^n,$$

and

$$\widehat{v}_{y|x_i} = \frac{\widehat{\rho}_y}{\sqrt{1 + \left(\frac{s_{y,i}^2 - s_{x,i}^2}{s_{x,i}^2} \right) \left(1 - \widehat{\rho}_y^2 \right)}}, \tag{23}$$

represents the adjusted correlation coefficient proposed by Forbes and Rigobon (2002). Forbes and Rigobon (2002) argue that estimation of cross-market correlation coefficients is biased due to heteroscedasticity in market returns. The adjusted correlation coefficient is scaled by a non-linear function of the percentage change in volatility of the source market returns $\left(\left(s_{y,i}^2 - s_{x,i}^2 \right) / s_{x,i}^2 \right)$ to solve this problem.

To test for a significant change in co-skewness between the non-crisis and crisis period, the null and alternative hypotheses are

$$H_0 \; : \; \psi_y(r_i^m, r_j^n) = \psi_x(r_i^m, r_j^n)$$
$$H_1 \; : \; \psi_y(r_i^m, r_j^n) \neq \psi_x(r_i^m, r_j^n).$$

Under the null hypothesis of no contagion, tests of contagion based on changes in co-skewness are asymptotically distributed as

$$CS_{12}, CS_{21}(i \to j) \xrightarrow{d} \chi_1^2.$$

4.2 Extremal Dependence Tests

Three types of extremal dependence tests are specified depending on whether asset returns in the source market are expressed in terms of simple returns or cubed returns in computing co-kurtosis, and in terms of return volatility in computing co-volatility. The first type of statistic CK_{13} is to detect the shocks emanating from the asset returns of a source market i to the cubed returns (akin to skewness) of an asset in a recipient market j. The second type

of statistic CK_{31} is to measure the shocks transmitting from the cubed returns of a source market i to the returns of recipient market j given as

$$CK_{13}(i \rightarrow j; r_i^1, r_j^3) = \left(\frac{\widehat{\xi}_y\left(r_i^1, r_j^3\right) - \widehat{\xi}_x\left(r_i^1, r_j^3\right)}{\sqrt{\frac{18\widehat{v}_{y|x_i}^2 + 6}{T_y} + \frac{18\widehat{\rho}_x^2 + 6}{T_x}}} \right)^2 \tag{24}$$

$$CK_{31}(i \rightarrow j; r_i^3, r_j^1) = \left(\frac{\widehat{\xi}_y\left(r_i^3, r_j^1\right) - \widehat{\xi}_x\left(r_i^3, r_j^1\right)}{\sqrt{\frac{18\widehat{v}_{y|x_i}^2 + 6}{T_y} + \frac{18\widehat{\rho}_x^2 + 6}{T_x}}} \right)^2, \tag{25}$$

where

$$\widehat{\xi}_y(r_i^m, r_j^n) = \frac{1}{T_y}\sum_{t=1}^{T_y}\left(\frac{y_{i,t} - \widehat{\mu}_{yi}}{\widehat{\sigma}_{yi}}\right)^m\left(\frac{y_{j,t} - \widehat{\mu}_{yj}}{\widehat{\sigma}_{yj}}\right)^n - \left(3\widehat{v}_{y|x_i}\right)$$

$$\widehat{\xi}_x(r_i^m, r_j^n) = \frac{1}{T_x}\sum_{t=1}^{T_x}\left(\frac{x_{i,t} - \widehat{\mu}_{xi}}{\widehat{\sigma}_{xi}}\right)^m\left(\frac{x_{j,t} - \widehat{\mu}_{xj}}{\widehat{\sigma}_{xj}}\right)^n - \left(3\widehat{\rho}_x\right),$$

and

$$\widehat{v}_{y|x_i} = \frac{\widehat{\rho}_y}{\sqrt{1 + \left(\frac{s_{y,i}^2 - s_{x,i}^2}{s_{x,i}^2}\right)\left(1 - \widehat{\rho}_y^2\right)}}.$$

The third type of statistic CV_{22} is to detect the shocks transmitting from the volatility of asset returns in a source market i to the volatility of asset returns in a recipient market j. The statistic of co-volatility can be represented as

$$CV_{22}(i \rightarrow j; r_i^2, r_j^2) = \left(\frac{\widehat{\xi}_y\left(r_i^2, r_j^2\right) - \widehat{\xi}_x\left(r_i^2, r_j^2\right)}{\sqrt{\frac{4\widehat{v}_{y|x_i}^4 + 16\widehat{v}_{y|x_i}^2 + 4}{T_y} + \frac{4\widehat{\rho}_x^4 + 16\widehat{\rho}_x^2 + 4}{T_x}}} \right)^2, \tag{26}$$

where

$$\widehat{\xi}_y(r_i^2, r_j^2) = \frac{1}{T_y}\sum_{t=1}^{T_y}\left(\frac{y_{i,t} - \widehat{\mu}_{yi}}{\widehat{\sigma}_{yi}}\right)^2\left(\frac{y_{j,t} - \widehat{\mu}_{yj}}{\widehat{\sigma}_{yj}}\right)^2 - \left(1 + 2\widehat{v}_{y|x_i}^2\right)$$

$$\widehat{\xi}_x(r_i^2, r_j^2) = \frac{1}{T_x}\sum_{t=1}^{T_x}\left(\frac{x_{i,t} - \widehat{\mu}_{xi}}{\widehat{\sigma}_{xi}}\right)^2\left(\frac{x_{j,t} - \widehat{\mu}_{xj}}{\widehat{\sigma}_{xj}}\right)^2 - \left(1 + 2\widehat{\rho}_x^2\right).$$

To test that there is a significant change in co-kurtosis or co-volatility between the non-crisis and crisis period, the null and alternative hypotheses are

$$H_0 \quad : \quad \xi_y(r_i^m, r_j^n) = \xi_x(r_i^m, r_j^n)$$
$$H_1 \quad : \quad \xi_y(r_i^m, r_j^n) \neq \xi_x(r_i^m, r_j^n).$$

Under the null hypothesis of no contagion, tests of contagion based on changes in co-kurtosis or co-volatility are asymptotically distributed as

$$CK_{13}, CK_{31}, CV_{22}(i \rightarrow j) \xrightarrow{d} \chi_1^2.$$

5 Sample Properties

This section examines the sample properties in terms of size and power of the new contagion tests derived in this paper using a range of Monte Carlo experiments. It also presents a comparison to the original co-skewness contagion test. Specifically, the three tests considered are i) the CS_{12} test in (21); ii) the CK_{13} test in (24); and iii) the CV_{22} test in (26). The size and power properties of the tests follow in Sections 5.2 and 5.3, before outlining the critical values appropriate for short crisis periods in Section 5.4.

5.1 Monte Carlo Design

The data generating process (DGP) follows the bivariate normal distribution with higher order moments and co-moments used to capture the linkages between asset markets 1 and 2 with linear dependence $(\theta_{x,3})$, asymmetric dependence $(\theta_{x,4}$ and $\theta_{x,5})$ and extremal dependence $(\theta_{x,6}, \theta_{x,7}$ and $\theta_{x,8})$ during the non-crisis period (x) and is

$$
\begin{aligned}
f(x_1, x_2) \quad = \quad & \exp\left(\theta_{x,1}x_1^2 + \theta_{x,2}x_2^2 + \theta_{x,3}x_1x_2 + \theta_{x,4}x_1x_2^2 + \theta_{x,5}x_1^2x_2 - \theta_{x,6}x_1^2x_2^2 \right. \quad (27) \\
& \left. -\theta_{x,7}x_1x_2^3 - \theta_{x,8}x_1^3x_2 + \theta_{x,9}x_1^3 + \theta_{x,10}x_2^3 + \theta_{x,11}x_1^4 + \theta_{x,12}x_2^4 - \eta_x\right),
\end{aligned}
$$

with $\theta_{x,1} = \theta_{x,2} = -0.5/\left(1 - \rho_x^2\right)$ and $\theta_{x,3} = \rho_x/\left(1 - \rho_x^2\right)$. During the crisis period (y)

$$
\begin{aligned}
f(y_1, y_2) \;=\; & \exp\left(\theta_{y,1}y_1^2 + \theta_{y,2}y_2^2 + \theta_{y,3}y_1y_2 + \theta_{y,4}y_1y_2^2 + \theta_{y,5}y_1^2y_2 - \theta_{y,6}y_1^2y_2^2 \right. \\
& \left. -\theta_{y,7}y_1y_2^3 - \theta_{y,8}y_1^3y_2 + \theta_{y,9}y_1^3 + \theta_{y,10}y_2^3 + \theta_{y,11}y_1^4 + \theta_{y,12}y_2^4 - \eta_y\right),
\end{aligned}
\tag{28}
$$

with $\theta_{y,1} = \theta_{y,2} = -0.5/\left(1 - \rho_y^2\right)$ and $\theta_{y,3} = \rho_y/\left(1 - \rho_y^2\right)$. η_x and η_y are the normalizing constants such that $\int\int f(x_1, x_2)dx_1dx_2 = 1$ and $\int\int f(y_1, y_2)dy_1dy_2 = 1$. The non-crisis $\{x_{1,t}, x_{2,t}\}$ and the crisis data $\{y_{1,t}, y_{2,t}\}$ for asset markets 1 and 2 are generated from the bivariate normal distribution with higher order moments and co-moments in (27) and (28) respectively, given their cumulative distribution functions and using the inverse-transform method. For the purposes of the Monte Carlo experiments the correlation parameters ρ_x and ρ_y are set equal to zero so that $\theta_{x,3} = \theta_{y,3} = 0$ and $\theta_{x,1} = \theta_{x,2} = \theta_{y,1} = \theta_{y,2} = -0.5$.

To evaluate the size properties of the contagion tests, the parameters in the Monte Carlo experiments conducted under the null hypothesis of no contagion are set to

$$
\theta_{x,i} = \theta_{y,i} = 0, \quad \forall i = 4, ..., 12,
\tag{29}
$$

in (27) and (28). To allow for varying crisis period sample sizes, five experiments are conducted, with the non-crisis period being set to 500 days (T_x) and the crisis period varying from 100 to 500 days $(T_y = 100, 200, 300, 400, 500)$.

To investigate the power properties of the contagion tests, three Monte Carlo experiments are conducted under the alternative hypothesis of contagion

$$
\begin{aligned}
&\text{Experiment I} &&: \theta_{y,6} > 0, \\
&\text{Experiment II} &&: \theta_{y,4} > 0,\ \theta_{y,6} = 0.5, \\
&\text{Experiment III} &&: \theta_{y,7} > 0,\ \theta_{y,6} = 0.5,
\end{aligned}
\tag{30}
$$

in (27) and (28). Each experiment investigates a particular transmission channel of contagion between asset markets 1 and 2: through the co-volatility channel in experiment I; through the co-skewness channel in experiment II; and through the co-kurtosis channel in experiment III. The strength of contagion is controlled by the values of the contagion parameters $\theta_{y,4}$, $\theta_{y,6}$ and $\theta_{y,7}$. Table 2 summarizes the restrictions on the parameters in (27) and (28) for

both the size and power tests. Note that experiment II is designed to allow a channel of contagion through co-skewness; however, to ensure boundedness, the co-volatility parameter needs to be fixed and is set to $\theta_{y,6} = 0.5$ in experiment II. This setting is also applied to the design for experiment III.

5.2 Size

Table 3 gives the results of the size experiments for the contagion tests, with attention to increasing the duration of the crisis period T_y from 100 to 500 days. The duration of the non-crisis period is set to $T_x = 500$ days. The number of replications is $500,000$ for all simulations. To compute the size values for each test statistic, a series of random samples of non-crisis and crisis data are generated using equations (27) and (28) with the restriction of no contagion and the assumption of no skewness and kurtosis in either asset market in (29). The size for each test is based on the 5% asymptotic χ_1^2 distribution critical value.

The results show that given a relatively large crisis sample period, both the asymmetric dependence tests (CS_{12}) and extremal dependence tests of contagion (CK_{13} and CV_{22}) exhibit the correct size in most cases, with the simulated size being close to the nominal size of 5%. However, if the crisis period sample size is on the small side ($T_y < 100$), the CK_{13} and CV_{22} tests are slightly undersized with values of 4.3% and 4.0%, respectively.

5.3 Power

Figure 2 shows the power functions of the four contagion statistics across experiments I to III with the sample sizes of the non-crisis and crisis period being 500 days ($T_x = T_y = 500$). The DGP in each experiment contains a particular transmission channel of contagion through the co-volatility channel in experiment I, the co-skewness and co-volatility channels in experiment II, and the co-kurtosis and co-volatility channels in experiment III. The power functions are determined by increasing the intensity level of contagion. The summary of the experiments are contained in (30) and in Table 2.

In experiment I shown in Figure 2, the power functions for each test are simulated based on the alternative hypothesis of contagion through the co-volatility channel for increasing values of $\theta_{y,6}$. Given a contagion strength value of zero, all tests yield a probability of 5% for finding contagion as the power functions are size adjusted ($\theta_{x,6} = \theta_{y,6} = 0$). The CV_{22}
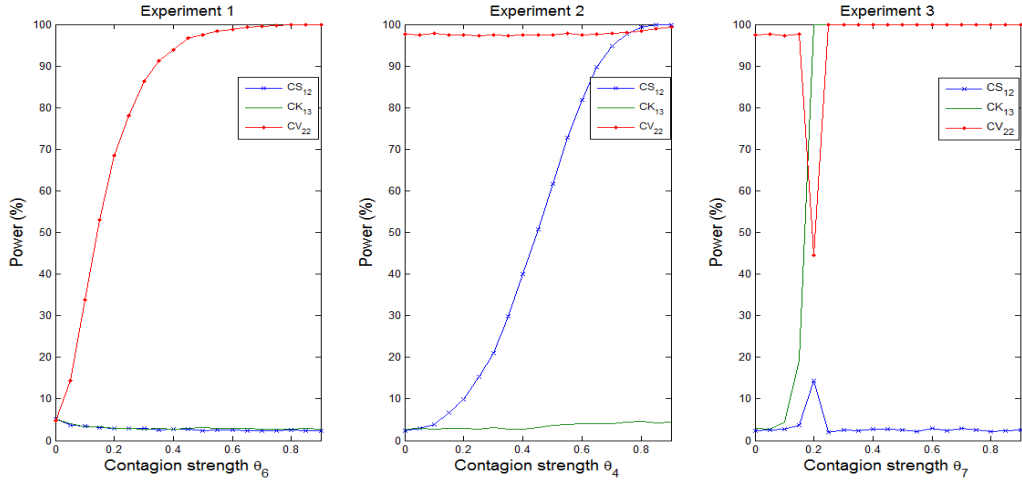
Figure 2: Simulated power functions of contagion statistics (CS_{12}, CK_{13} and CV_{22}) for various experiments. The sample sizes of non-crisis and crisis periods are set at $T_x = T_y = 500$. The number of replications is set at $100,000$.

test is the most powerful among the four tests over the range of $\theta_{y,6}$, approaching power of 100% as the contagion strength exceeds 0.6. As expected, the CS_{12} and CK_{13} tests exhibit no power in detecting contagion as the DGP in experiment I only contains the transmission channel of contagion through the co-volatility channel and does not include changes in the co-skewness or co-kurtosis channels.

In experiment II of Figure 2, the power functions for each test are simulated based on the alternative hypothesis of contagion through the co-skewness and co-volatility channels where $\theta_{y,4} > 0$ and $\theta_{y,6} = 0.5$. The co-volatility contagion parameter is kept constant at a value of $\theta_{y,6} = 0.5$ to ensure boundedness. The CS_{12} test shows a monotonically increasing power function over the range of contagion strength controlled by $\theta_{y,4}$ from 0 to 0.9. The CK_{13} and CV_{22} tests exhibit constant power for the experiment, with as expected, low power of 5% for the co-kurtosis change test and a high power of 95% for the co-volatility change test.

The power functions for each test in experiment III are simulated based on the alternative hypothesis of contagion through the co-kurtosis and co-volatility channels where $\theta_{y,7} > 0$ and $\theta_{y,6} = 0.5$. The CK_{13} test has a monotonically increasing power function and its power quickly approaches 100% as contagion strength $\theta_{y,7}$ reaches 0.2. As for the CS_{12} test, there is low constant power of 5% for most values of $\theta_{y,7}$ as expected, but its power jumps to 15% when $\theta_{y,7} = 0.2$. Similar to experiment II, the CV_{22} test yields a probability of finding
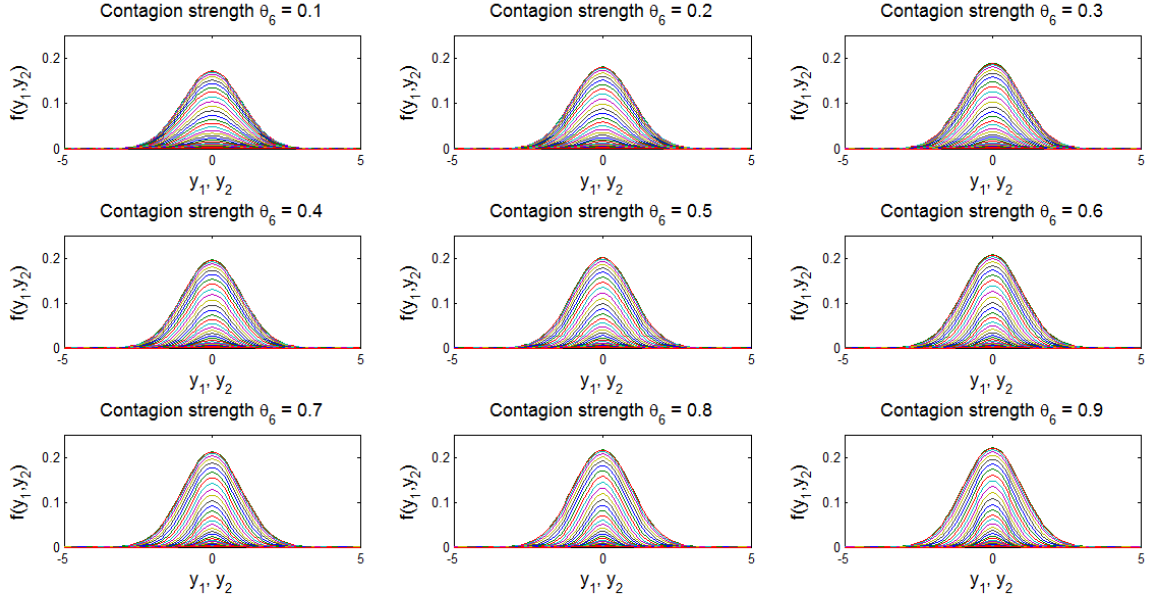
16

Figure 3: Simulated bivariate normal distributions with values of the co-volatility parameter θ_6 ranging from 0.1 to 0.9. The distributions are simulated using (28) and experiment I of (30) with the parameters summarized in Table 2.

contagion over most of the range of values of $\theta_{y,7}$ between 95% to 100%. However, when the co-kurtosis contagion strength value is $\theta_{y,7} = 0.2$, the power of the CV_{22} test falls to 45%. This is the same point at which the power function for the co-skewness test momentarily jumps.

To investigate the anomaly surrounding the point when $\theta_{y,7} = 0.2$ in Experiment III, the bivariate normal distribution with a range of parameter values for co-volatility, co-skewness and co-kurtosis for Experiments I to III is simulated and shown in Figures 3 to 5. Figure 3 reveals that the simulated distributions exhibit higher peaks around the mean, as the co-volatility parameter increases while Figure 4 shows that the simulated distributions exhibit longer and fatter tails as the co-skewness parameter increases. Figure 5 shows that the simulated distribution is symmetric in the case when the co-kurtosis parameter is set to 0.1. However, as the co-kurtosis parameter increases, the simulated distribution switches from a unimodal to bimodal distribution, explaining the results illustrated in the power functions for experiment III.

Overall, the simulation experiments reveal that the extremal dependence tests (CK_{13} and CV) are the most powerful, followed by the asymmetric dependence test (CS_{12}). The

17

Figure 4: Simulated bivariate normal distributions with values of the co-skewness parameter θ_4 ranging from 0.1 to 0.9. The distributions are simulated using (28) and experiment II of (30) with the parameters summarized in Table 2.
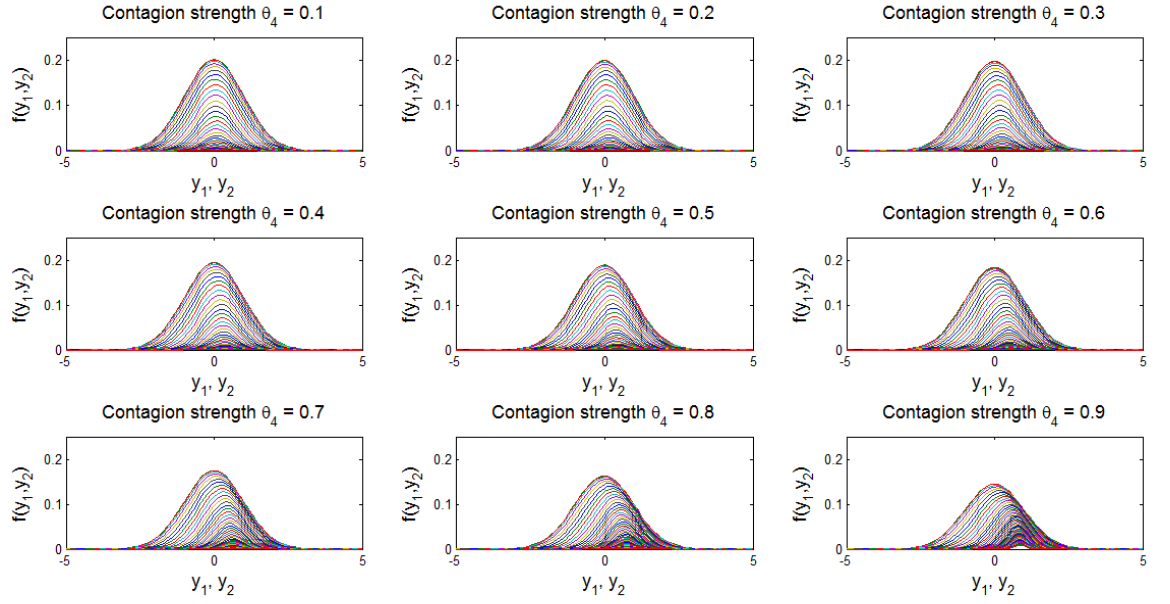

Figure 5: Simulated bivariate normal distributions with values of the co-kurtosis parameter θ_7 ranging from 0.1 to 0.9. The distributions are simulated using (28) and experiment III of (30) with the parameters summarized in Table 2.
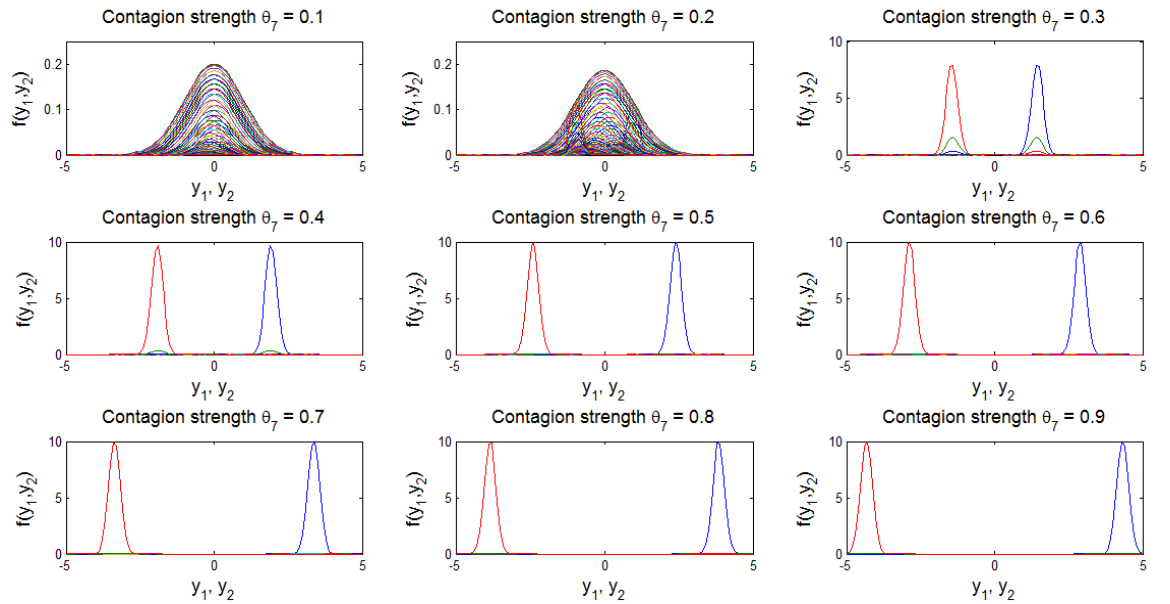
speed of approaching 100% power for the CK_{13} and CV_{22} tests is faster than that for the CS_{12} test.

5.4 Adjustment for Short Duration Crisis Periods

If a crisis period is short, standard χ_1^2 critical values will not suffice in testing for contagion as the test statistic is undersized. The duration of the last four financial crises are long, ranging from 187 days for the Argentinean crisis to more than 390 days for the US 2008-2009 financial crisis. However, prior to that most crises were of much shorter duration. For example, the Russian crisis lasted for 99 days, while the Brazilian crisis was 36 days.[5] To provide guidance to the correct critical values to use for small crisis durations, the Monte Carlo experiments for the CS_{12}, CK_{13} and CV_{22} contagion tests are run for cases where the crisis sample period is allowed to vary from 15 to 200 days. Table 4 presents the relevant critical values.

6 Empirical Application

The financial crisis of 2008-09 originated in the US interbank market and spread across the financial system, exposing vulnerabilities elsewhere in global banking and equity markets. This section applies the asymmetric and extremal tests of contagion to banking and equity returns of eight countries from Asia, Europe, Latin America and North America to the crisis episode to illustrate the workings of the tests.[6]

The data are composed of daily equity returns in US dollars expressed in percent for each market i calculated as

$$Z_{i,t} = 100 \left(ln(P_{i,t}) - ln(P_{i,t-1}) \right),$$

where $P_{i,t}$ is the equity index in market i at time t and $Z_{i,t}$ is the percentage return. The sample period starts on April 1, 2005 and ends on August 31, 2009. The non-crisis period is defined from April 1, 2005 to February 29, 2008 ($T_x = 760$ observations), and the crisis

---

[5]See Fry-McKibbin, Hsiao and Tang (2013) for a discussion on crisis duration.

[6]The purpose of the data set is illustrative only and covers Hong Kong, Korea, France, Germany, the UK, Chile, Mexico and the US. The banking data is daily banking equity price returns.

period from March 3, 2008 to August 31, 2009 ($T_y = 391$).[7,8] The crisis source is assumed to be the US banking sector. All returns are plotted in Figure 6.[9] The figure illustrates that the volatility of equity returns changes dramatically in equity markets and banking sectors globally during the financial crisis of 2008-09.

As outlined earlier, Table 1 contains the co-moment statistics of the dataset, while for completeness, Table 5 also report the moment statistics. Several features characteristic of crises are highlighted in this table and are compared to a period in which there is no crisis. Table 5 reports descriptive statistics of the own-moments of the mean, standard deviation, skewness and kurtosis for each return series in both equity markets and banking sectors for each period. The first and second moments in the table show that average daily returns of the equity indices decrease and volatility increases across the board in the crisis period as expected.[10] These statistics are consistent with a risk-averse investor realizing a higher level of excess returns in the non-crisis period in conjunction with a higher level of risk across the two regimes (Sharpe, 1964; Lintner, 1965; Black, 1972). Inspection of the third and fourth moments of the returns shows that it is not only the first and second moments that change. Non-normality with asymmetry and fat tails is a major characteristic of returns in the equity

---

[7]The starting date of the crisis coincides with the Fed rescue of Bear Stearns on March 3, 2008. The end of the crisis is August 31, 2009 as economic activity in the US improved. This paper does not consider the period of the European sovereign debt crisis, and jitters in international banking markets such as the issues with Northern Rock as early as 2007 are acknowledged.

[8]The data are filtered in the same way as in Forbes and Rigobon (2002). That is, a VAR model is first specified and estimated to control for market fundamentals (country-specific and cross market relationships that always exist) and also to handle serial correlation in the data set. The model specification is given by

$$\begin{aligned} Z_t &= \phi(L)Z_t + R_t \\ Z_t &= \{x_t, y_t\}'. \end{aligned} \tag{31}$$

where Z_t is a transposed vector of returns across a set of equity markets and banking sectors during the non-crisis (x_t) and crisis (y_t) periods; $\phi(L)$ is a vector of lags and R_t is a vector of the residual terms. Two-day rolling average returns are used to deal with time zone issues. $L = 5$ based on the criteria of the sequential modified log-likelihood ratio test statistic and the Akaike information Criterion. The residuals estimated from the VAR (R_t) are used in computing the contagion statistics.

[9]The equity indices are collected from Datastream. The mneumonics are: Hong Kong - Hang Seng price index (HNGKNGI); Korea - Korea Se Composite price index (KORCOMP); Chile - General price index (IGPAGEN); Mexico - Mexico Ipc Bolsa price index (MXIPC35); France - CAC 40 price index (FRCAC40); Germany - MDAX Frankfurt price index (MDAXIDX); the UK - FTSE100 price index (FTSE100); and the US - Dow Jones Industrials (DJINDUS). The banking equity indices are collected from Bloomberg. The mneumonics are: Hong Kong - FTSE China A 600 Banks (XA81); Korea - Korea Banking Index (KOSPBANK); Chile - MSCI Chile Banks (MXCL0BK); Mexico - MSCI Mexico Banks (MXMX0BK); Germany - MSCI Germany Banks (MXDE0BK); the UK - FTSE 350 Banking Index (F3BANK); and the US - PHLX KBW Bank Sector Index (BKX).

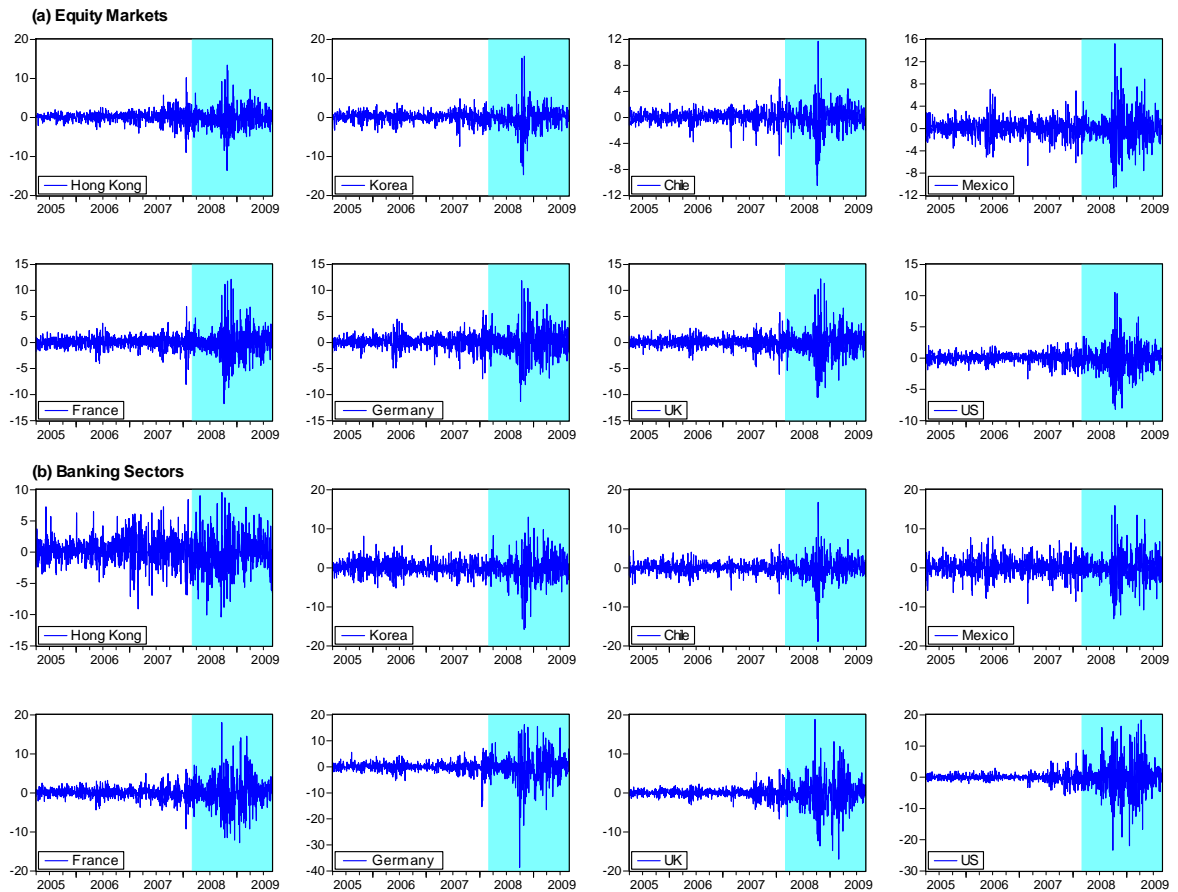[10]Volatility in the banking sectors display much higher values than that in the equity markets.

Figure 6: Daily percentage returns of equity markets and equity returns of banking sectors for eight countries (Hong Kong, Korea, Chile, Mexico, France, Germany, the UK and the US) during January 1, 2005 to August 31, 2009. The shaded areas refer to the global crisis starting from March 3, 2008 to August 31, 2009.

markets and banking sectors. The third moment of skewness is generally negative in the non-crisis period but switches sign to be positive in the crisis period for most markets. Asset returns also yield leptokurtic behavior and kurtosis rises during the crisis period, illustrating that the return distributions are far from normal.

Table 1 reports statistics of the higher order co-moments of the equity and banking returns with the US banking returns during the non-crisis and crisis periods. Correlation between the US banking sector and all other returns increase apart from the correlation of the US banking sector with the US equity market indicating foremost the nature of the crisis as being one of a banking issue, rather than of the aggregate equity markets in the US.

Turning to the asymmetric dependence statistics, an increase in the magnitude of the value of co-skewness$_{12}$ indicates that the market volatilities for the countries in the sample are more strongly related to the returns in the US banking sector than previously. Negative co-skewness$_{12}$ values suggest that an asset invested in the US banking sector should achieve a relatively higher return when other markets are less volatile. Positive co-skewness$_{12}$ is desirable from the point of view of the US investor as higher returns during the volatile period provide a good hedge. That is, an asset pays higher returns to the US banking sector as other equity markets become more volatile. The logic is reversed for the co-skewness$_{21}$ case. The co-skewness$_{12}$ statistics in the equity markets show a fall in value in most cases, while for the banking sector all coefficients are higher. A similar property is presented in Table 1 where co-skewness$_{21}$ rises or moves towards positive values in almost all markets during the crisis period.

Table 1 presents statistics on the extremal (fourth order) co-moments of the returns for the two periods. As the table shows, co-kurtosis$_{13}$ and co-kurtosis$_{31}$ are positive during both periods for all countries apart from Hong Kong and Korea where both versions are negative during the non-crisis period for the banking sector and for Korea for the equity sector. Co-kurtosis$_{13}$ ranges from -0.401 (Hong Kong) to 9.386 (US) during the non-crisis period and from 0.061 (Hong Kong) to 5.563 (US) during the crisis period. Co-kurtosis increases in the crisis period in most cases, indicating that the joint distribution of returns has a sharper peak and longer and fatter tails during the crisis period. The higher values of co-kurtosis$_{13}$ in the crisis period implies returns in the US banking sector are low with a positive skewness of returns in the other markets, thus increasing contagion risk in the crisis period.

Similarly, co-kurtosis$_{31}$ is higher in the crisis period, suggesting that the distribution of

returns in the US banking sector exhibits negative skewness as the returns in the other markets are low during the crisis period, again increasing risk. The co-volatility$_{22}$ statistics, which measures the correlation between volatilities, are all positive in both periods. Interestingly, the co-volatility relationship does not change as systematically as the other statistics. Sometimes the statistic takes a higher value than in the non-crisis period and other times it does not. Most interesting is that the co-volatility statistics for the US are lower for both the equity and banking sectors in the crisis period.

Contagion through Asymmetric Dependence   Table 6 presents the empirical results for the asymmetric dependence tests of contagion from US banks during the crisis of 2008-09. Inspection of this table reveals that significant evidence of contagion is found within banking sectors through the co-skewness channel, but less is found between the US banking sector and the equity markets. Panel A of Table 6 shows no significant cross-border contagious linkages from the US banking sector to equity market volatility apart from that of Chile through the first co-skewness channel (CS_{12}), while the equity markets in Hong Kong and Korea are exposed to contagion from the volatility of the US banking sector through the second channel (CS_{21}). The majority of the banking sectors, apart from Mexico's are affected by at least one of the asymmetric dependence contagion channels during the financial crisis. The significant contagion effects are from US banking to the volatility of the banking sectors in Korea, France, Germany, the UK and Chile and also from the volatility of the US banking sector to the returns of the banking sectors in Hong Kong, Korea and the UK. The US banking sector is of systemic importance through the asymmetric dependence channel.[11]

Contagion through Extremal Dependence   The results of the extremal dependence tests presented in Table 7 show significant contagion effects from the US banking sector to the equity (Panel A) and banking markets (Panel B) in the four regions during the crisis of 2008-09 through the three extremal dependence measures, indicating again that the US banking sector is of systemic importance. Among the regions, Europe is most exposed with all banking and equity sectors experiencing dramatic increases in extremal dependence (i.e. co-kurtosis and co-volatility). The smallest value of the test statistic for the equity markets in Europe is 55.82 for the linkages between the volatility of the US banking sector and the

---

[11]Gropp and Moerman (2004) define the term "systemic importance" in terms of the banking sector which tends to have a net contagious influence on other banks.

volatility of the German equity market (CV_{22}), and is 79.54 in the banking sector for the links between the US banks and the skewness of the German banking sector (CK_{13}). These values are much higher than the critical value at the 1% significance level.

The evidence of contagion to Asia and Latin America is strong through the two forms of the co-kurtosis channels. The channels operating are all cases of the CK_{13} channel and for all cases except for the Hong Kong banking sector for the CK_{31} channel. There is evidence of the co-volatility CV_{22} contagion channel in 50% of cases, with no evidence of contagion to the volatility of the Hong Kong and Chilean markets through this channel. As for North America, the US equity market is affected by the banking sector to its skewness (CK_{13}) and to the volatility channel (CV_{22}).

Comparing Tables 6 and 7, more contagion is detected through extremal dependence than through asymmetric dependence during the financial crisis of 2008-09. The finding suggests that in extreme events, the extremal dependence links are more important in capturing markets linkages than the asymmetric dependence linkages.

7   Conclusions

This paper proposes a class of new tests of financial contagion based on changes in extremal dependence measured by co-kurtosis and co-volatility during a financial crisis. The extremal dependence structure is non-linear and is subject to change in different regimes. The changes can be attributed to the risk return trade-off effect between the expected excess returns and higher order moments. In deriving the new tests for contagion, a bivariate generalized exponential distribution is specified to allow for higher order moments and co-moments. The Lagrange multiplier test is adopted to develop the statistics of co-kurtosis and co-volatility, which provides the framework for developing the contagion tests.

Monte Carlo simulations show that the statistics for contagion provide a good approximation of the finite sample distribution. However, where the crisis period is relatively short the statistics are undersized. Small sample critical values are computed for the tests for use in this situation. Compared to the most closely related tests in the literature, the power functions of the new tests are consistently better in that they detect contagion earlier and more definitively when contagion exists. The Monte Carlo simulations also reveal that in the case of strong co-kurtosis, the distribution of returns is bimodal. This has some impact on

the power functions when contagion is low, but does not appear to have adverse implications when the strength of contagion is high.

The new approach is illustrated with an application to test for financial contagion in equity markets and banking sectors during the financial crisis of 2008-09. The results of the tests show significant contagion effects from the US banking sector to global equity and banking sector markets through extremal dependence. The extremal dependence tests capture more market co-movements than the asymmetric dependence tests during extreme events.

References

[1] Allen, F. and Gale, D. (2000), "Financial Contagion", *Journal of Political Economy*, 108, 1-33.

[2] Bae, K.H., Karolyi, G.A., and Stulz, R.M. (2003), "A New Approach to Measuring Financial Contagion," *Review of Financial Studies*, 16, 717-763.

[3] Bekaert, G., and Wu, G. (2000), "Asymmetric Volatility and Risk in Equity Markets," *The Review of Financial Studies*, 13, 1, 4-42.

[4] Black, F. (1972), "Capital Market Equilibrium with Restricted Borrowing," *Journal of Business*, 45, 3, 444-54.

[5] Black, F. (1976), "Studies of Stock Price Volatility Changes,"Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section, 177-181.

[6] Brunnermeier, M.K., and Pedersen, L.H. (2008), "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22, 6, 2201-38.

[7] Busetti, F., and Harvey, A.C. (2011), "When is a Copula Constant? A Test for Changing Relationships", *Journal of Financial Econometrics*, 9, 106-131.

[8] Cifuentes, R., Ferrucci, G. and Shin, H. (2005), "Liquidity Risk and Contagion", *Journal of the European Economic Association*, 3(2-3), 556-66.

[9] Cobb, L., Koppstein, P., and Chen, N.H. (1983), "Estimation and Moment Recursion Relations for Multimodal Distributions of the Exponential Family," *Journal of the American Statistical Association*, 78, 124-30.

[10] Conrad, J., Dittmar, R.F., and Ghysels, E. (2013), "Ex Ante Skewness and Expected Stock Returns," Journal of Finance, 68(1), 85-124.

[11] Diebold, F. X., and Yilmaz, K. (2009), "Measuring Financial Asset Return and Volatility Spillovers, with Application to Global Equity Markets", *The Economic Journal*, 119, 158-171.

[12] Dittmar, R.F. (2002), "Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross-section of Equity Returns," *The Journal of Finance*, 57, 1, 369-403.

[13] Embrechts, P., Lindskog, F., and McNeil, A. (2003), "Modelling Dependence with Copulas and Applications to Risk Management," Handbook of Heavy Tailed Distributions in Finance, Chapter 8, 329-84.

[14] Engle, R.F., Ng, V.K., and Rothschild, M. (1990), "Asset Pricing with a Factor-ARCH Covariance Structure: Empirical Estimates for Treasury Bills," *Journal of Econometrics,* 45, 213-237.

[15] Fang, H., and Lai, T.Y. (1997), "Co-Kurtosis and Capital Asset Pricing," *The Financial Review*, 32, 2, 293-307.

[16] Favero, C.A. and Giavazzi, F. (2002), "Is the International Propagation of Financial Shocks Non-linear? Evidence From the ERM", *Journal of International Economics*, 57 (1), 231-46.

[17] Forbes, K., and Rigobon, R. (2002), "No Contagion, Only Interdependence: Measuring Stock Market Co-movements," *The Journal of Finance*, 57, 5, 2223-61.

[18] Friend, I., and Westerfield, R. (1980), "Co-skewness and Capital Asset Pricing,", *The Journal of Finance,* 35, 897-913.

[19] Fry-McKibbin, R.A., and Hsiao, C.Y., (2015), "Extremal Dependence Tests for Contagion", CAMA Working Paper, forthcoming.

[20] Fry-McKibbin, R.A., Hsiao, C.Y-L, and Tang, C. (2013), "Contagion and Global Financial Crises: Lessons from Nine Crises Episodes", Open Economies Review, 2013, 1-50.

[21] Fry, R.A., Martin, V.L., and Tang, C. (2010), "A New Class of Tests of Contagion with Applications," *Journal of Business and Economic Statistics*, 28, 3, 423-437.

[22] Fry-McKibbin, R.A., Martin, V.L., and Tang, C. (2014), "Financial Contagion and Asset Pricing,", *Journal of Banking and Finance*, 47, 296-308.

[23] Garcia, R. and G. Tsafack (2011). Dependence Structure and Extreme Comovements in International Equity and Bond Markets. *Journal of Banking and Finance* 35(8), 1954-70.

[24] "Gorton, G. and Metrick, A. (2012), "Securitized Banking and the Run on Repo," *Journal of Financial Economics*, 104, 425-51.

[25] Gropp, R., Duca, M.L., and Vesala J. (2010), "Cross-border Bank Contagion in Europe," *International Journal of Central Banking*, 5, 1, 97-139.

[26] Gropp, R., and Moerman, G. (2004), "Measurement of Contagion in Banks' Equity Prices," *Journal of International Money and Finance*, 23, 405-59.

[27] Harvey C.R., and Siddique, A. (2000), "Conditional Skewness in Asset Pricing Tests," *The Journal of Finance*, 55, 3, 1263-95.

[28] Hwang, S., and Satchell, S.E. (1999), "Modelling Emerging Market Risk Premia Using Higher Moments," *International Journal of Finance and Economics*, 4, 271-296.

[29] Kraus, A., and Litzenberger, R. (1976), "Skewness Preferences and the Valuation of Risk Assets,," *Journal of Finance*, 31, 1085-1100.

[30] Lintner, J. (1965), "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics*, 47, 1, 13-37.

[31] Lye, J.N., and Martin, V.L. (1993), "Robust Estimation, Non-Normalities and Generalized Exponential Distributions," *Journal of the American Statistical Association*, 88, 253-59.

[32] Martellini, L., and Ziemann, V. (2010), "Improved Estimates of Higher-order Comoments and Implications for Portfolio Selection," *Review of Financial Studies*, 23(4), 1467-1502.

[33] Pesaran, H. and Pick, A. (2007), "Econometric Issues in the Analysis of Contagion", *Journal of Economic Dynamics and Control*, 31, 1245-1277.

[34] Rodriguez, J.C. (2007), "Measuring Financial Contagion: A Copula Approach", *Journal of Empirical Finance*, 14, 401-423.

[35] Scott, R.C., and Horvath, P.A. (1980), "On the Direction of Preference for Moments of Higher Order than the Variance," *The Journal of Finance*, 35, 4, 915-19.

[36] Sharpe, W.F. (1964), "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *The Journal of Finance*, 19, 3, 425-42.

[37] Shleifer, A., and Vishny, R.W. (1997), "The Limits of Arbitrage," *The Journal of Finance*, 52, 1, 35-55.

[38] Vayanos, D. (2004), "Flight to Quality, Flight to Liquidity, and the Pricing of Risk", NBER Working Paper #10327.

[39] Vaugirard, V. (2007), "Informational Contagion of Bank Runs in a Third-Generation Crisis Model", *Journal of International Money and Finance,* 26, 403-429.

A   Appendix: Derivation of the Test Statistics

A.1   Portfolio Choice

This section outlines a model of portfolio choice accounting for higher order moments and co-moments. Consider the standard utility function of an investor allocating their portfolio across N risky assets to maximize their end of period wealth W

$$\underset{\alpha_0, \alpha_1, \ldots, \alpha_N}{Max} \quad E\left[U\left(W\right)\right], \tag{32}$$

given the budget constraint

$$\alpha_0 + \sum_{i=1}^{N} \alpha_i = 1.$$

The fraction of wealth allocated to the risk-free asset is α_0, and the fraction of wealth allocated to the ith risky asset is α_i. The end of period wealth is

$$W = \alpha_0\left(1 + R_f\right) + \sum_{i=1}^{N} \alpha_i\left(1 + R_i\right), \tag{33}$$

where R_f and R_i are the rates of return on the risk-free and risky assets respectively.

Assume the distribution of returns of the investor's portfolio of risky assets to be asymmetric and fat-tailed. In this case, the investor's utility function is constructed to include higher order moments since the quadratic mean and variance do not completely determine the distribution. Scott and Horvath (1980) show that investor preferences for higher moments are essential for portfolio selection. The expected utility of the return on investment W of the investor in this case is given by an infinite-order Taylor series expansion around expected wealth:

$$E\left[U\left(W\right)\right] = \sum_{q=0}^{\infty} \frac{U^{(q)}\left(\overline{W}\right)}{q!} E\left[\left(W - \overline{W}\right)^q\right], \tag{34}$$

where the expected value of the end of period wealth is $\overline{W} = E\left(W\right)$ and the q-th derivative of U is $U^{(q)}$.

Equation (34) can be decomposed into the investor's risk preferences $U^{(q)}\left(\overline{W}\right)$ and the q moments of the distribution $E\left[\left(W - \overline{W}\right)^q\right]$. Scott and Horvath (1980) show that under the assumption of positive marginal utility of wealth, a utility function exhibits decreasing absolute risk aversion at all wealth levels with strict consistency for moment preferences of $U^{(q)}\left(\overline{W}\right) > 0$ if q is odd and $U^{(q)}\left(\overline{W}\right) < 0$ if q is even.

Consider equation (34) for the case of the moments $q = 1$ to $q = 4$. The expected utility

function can be written as

$$
\begin{aligned}
E\left[U\left(W\right)\right] \quad = \quad & U\left(\overline{W}\right) + U'\left(\overline{W}\right) E\left[\left(W - \overline{W}\right)\right] + \frac{1}{2}U''\left(\overline{W}\right) E\left[\left(W - \overline{W}\right)^2\right] + \\
& \frac{1}{3!}U'''\left(\overline{W}\right) E\left[\left(W - \overline{W}\right)^3\right] + \frac{1}{4!}U^{(4)}\left(\overline{W}\right) E\left[\left(W - \overline{W}\right)^4\right] + \\
& o(W).
\end{aligned}
\tag{35}
$$

where $o(W)$ is the Taylor remainder. The expected return, variance, skewness and kurtosis for the end of period return, R_p, are denoted by

$$
\begin{aligned}
\mu_p \quad &= \quad E\left[R_p\right] \\
\sigma_p^2 \quad &= \quad E\left[\left(R_p - \mu_p\right)^2\right] = E\left[\left(W - \overline{W}\right)^2\right] \\
s_p^3 \quad &= \quad E\left[\left(R_p - \mu_p\right)^3\right] = E\left[\left(W - \overline{W}\right)^3\right] \\
k_p^4 \quad &= \quad E\left[\left(R_p - \mu_p\right)^4\right] = E\left[\left(W - \overline{W}\right)^4\right],
\end{aligned}
\tag{36}
$$

where μ_p is the expected return on the portfolio defined as

$$
\mu_p = E\left[\alpha_0 R_f\right] + E\left[\sum_{i=1}^{N}\alpha_i R_i\right] = \sum_{i=0}^{N}\alpha_i \mu_i,
\tag{37}
$$

where α_i is the weight of the ith asset in the portfolio with $\alpha_i \geq 0$ and $\sum_{i=0}^{N}\alpha_i = 1$.

The investor's equilibrium condition is solved by taking the first order conditions of the Lagrange of the utility function of wealth subject to the budget constraint where

$$
L = E\left[U\left(W\right)\right] + \lambda\left(1 - \alpha_0 + \sum_{i=1}^{N}\alpha_i\right).
\tag{38}
$$

In solving this problem by taking the two partial derivatives of the Lagrange function, two pairs of simultaneous linear equations are written as

$$
\frac{\partial L}{\partial \alpha_0} = \frac{\partial U\left(\overline{W}\right)}{\partial \overline{W}}R_f - \lambda = 0,
\tag{39}
$$

and

$$
\begin{aligned}
\frac{\partial L}{\partial \alpha_i} \quad = \quad & \frac{\partial U\left(\overline{W}\right)}{\partial \overline{W}}E\left(R_i\right) + \frac{\partial\left(\frac{1}{2}U''\left(\overline{W}\right)\right)}{\partial \alpha_i}\sigma_p^2 \\
& + \frac{\partial\left(\frac{1}{3!}U'''\left(\overline{W}\right)\right)}{\partial \alpha_i}s_p^3 + \frac{\partial\left(\frac{1}{4!}U^{(4)}\left(\overline{W}\right)\right)}{\partial \alpha_i}k_p^4 - \lambda = 0.
\end{aligned}
\tag{40}
$$

The expected excess return on each risky asset over the risk free rate is given by

$$E(R_i) - R_f = \left(\frac{\partial E\left[U\left(W\right)\right]}{\partial \sigma_p^2}\right)\sigma_p^2 + \left(\frac{\partial E\left[U\left(W\right)\right]}{\partial s_p^3}\right)s_P^3 + \left(\frac{\partial E\left[U\left(W\right)\right]}{\partial k_p^4}\right)k_P^4, \quad (41)$$

where

$$\left(\frac{\partial E\left[U\left(W\right)\right]}{\partial \sigma_p^2}\right) = -\frac{\frac{\partial\left(\frac{1}{2}U''\left(\overline{W}\right)\right)}{\partial \alpha_i}}{\frac{\partial U\left(\overline{W}\right)}{\partial \overline{W}}} \quad (42)$$

$$\left(\frac{\partial E\left[U\left(W\right)\right]}{\partial s_p^3}\right) = -\frac{\frac{\partial\left(\frac{1}{3!}U'''\left(\overline{W}\right)\right)}{\partial \alpha_i}}{\frac{\partial U\left(\overline{W}\right)}{\partial \overline{W}}}$$

$$\left(\frac{\partial E\left[U\left(W\right)\right]}{\partial k_p^4}\right) = -\frac{\frac{\partial\left(\frac{1}{4!}U^{(4)}\left(\overline{W}\right)\right)}{\partial \alpha_i}}{\frac{\partial U\left(\overline{W}\right)}{\partial \overline{W}}}.$$

Equation (41) shows that the expected excess return on each risky asset contains two elements: i) the portfolio risk premium involving higher order moments of volatility (σ_p^2), skewness (s_p^3) and kurtosis (k_p^4); and ii) measures of the investor's risk preferences for volatility $\left(\frac{\partial E[U(W)]}{\partial \sigma_p^2}\right)$, skewness $\left(\frac{\partial E[U(W)]}{\partial s_p^3}\right)$ and kurtosis $\left(\frac{\partial E[U(W)]}{\partial k_p^4}\right)$.

If an investor invests in two risky assets, $N = 2$, then the variance, skewness and kurtosis of the end of period returns can be decomposed into

$$\begin{aligned}
\sigma_p^2 &= E\left[\left(R_p - \mu_p\right)^2\right] \quad (43)\\
&= E\left[\left(\sum_{i=1}^{N=2}\alpha_i\left(R_i - \mu_i\right)\right)^2\right]\\
&= \alpha_1^2 E\left[\left(R_1 - \mu_1\right)^2\right] + \alpha_2^2 E\left[\left(R_2 - \mu_2\right)^2\right] + 2\alpha_1\alpha_2 E\left[\left(R_1 - \mu_1\right)\left(R_2 - \mu_2\right)\right],
\end{aligned}$$

$$\begin{aligned}
s_P^3 &= E\left[\left(\sum_{i=1}^{N=2}\alpha_i\left(R_i - \mu_i\right)\right)^3\right] \quad (44)\\
&= \alpha_1^3 E\left[\left(R_1 - \mu_1\right)^3\right] + \alpha_2^3 E\left[\left(R_2 - \mu_2\right)^3\right] +\\
&\quad 3\alpha_1^2\alpha_2 E\left[\left(R_1 - \mu_1\right)^2\left(R_2 - \mu_2\right)\right] + 3\alpha_1\alpha_2^2 E\left[\left(R_1 - \mu_1\right)\left(R_2 - \mu_2\right)^2\right],
\end{aligned}$$

$$k_P^4 = E\left[\left(\sum_{i=1}^{N=2}\alpha_i\left(R_i-\mu_i\right)\right)^4\right] \tag{45}$$

$$= \alpha_1^4 E\left[\left(R_1-\mu_1\right)^4\right] + \alpha_2^4 E\left[\left(R_2-\mu_2\right)^4\right] + 4\alpha_1^3\alpha_2 E\left[\left(R_1-\mu_1\right)^3\left(R_2-\mu_2\right)\right]$$
$$+ 4\alpha_1\alpha_2^3 E\left[\left(R_1-\mu_1\right)\left(R_2-\mu_2\right)^3\right] + 6\alpha_1^2\alpha_2^2 E\left[\left(R_1-\mu_1\right)^2\left(R_2-\mu_2\right)^2\right],$$

respectively.

Substituting equations (43) to (45) into equation (41) gives the expected excess return for asset i in terms of the second, third and fourth order moments and co-moments. The co-moments include the correlation, co-skewness, co-kurtosis and co-volatility as shown below:

$$\begin{aligned}E(R_i)-R_f =\ & \theta_1 E\left[\left(R_1-\mu_1\right)^2\right] + \theta_2 E\left[\left(R_2-\mu_2\right)^2\right] + \theta_3 E\left[\left(R_1-\mu_1\right)\left(R_2-\mu_2\right)\right]\\
& +\theta_4 E\left[\left(R_1-\mu_1\right)^3\right] + \theta_5 E\left[\left(R_2-\mu_2\right)^3\right] + \theta_6 E\left[\left(R_1-\mu_1\right)^2\left(R_2-\mu_2\right)\right]\\
& +\theta_7 E\left[\left(R_1-\mu_1\right)\left(R_2-\mu_2\right)^2\right] + \theta_8 E\left[\left(R_1-\mu_1\right)^4\right] + \theta_9 E\left[\left(R_2-\mu_2\right)^4\right]\\
& +\theta_{10} E\left[\left(R_1-\mu_1\right)^3\left(R_2-\mu_2\right)\right] + \theta_{11} E\left[\left(R_1-\mu_1\right)\left(R_2-\mu_2\right)^3\right]\\
& +\theta_{12} E\left[\left(R_1-\mu_1\right)^2\left(R_2-\mu_2\right)^2\right],\end{aligned} \tag{46}$$

where

$$\begin{aligned}
&\theta_1 = \alpha_1^2\left(\frac{\partial E[U(W)]}{\partial\sigma_p^2}\right), &&\theta_5 = \alpha_2^3\left(\frac{\partial E[U(W)]}{\partial s_P^3}\right), &&\theta_9 = \alpha_2^4\left(\frac{\partial E[U(W)]}{\partial k_P^4}\right),\\
&\theta_2 = \alpha_2^2\left(\frac{\partial E[U(W)]}{\partial\sigma_p^2}\right), &&\theta_6 = 3\alpha_1^2\alpha_2\left(\frac{\partial E[U(W)]}{\partial s_P^3}\right), &&\theta_{10} = 4\alpha_1^3\alpha_2\left(\frac{\partial E[U(W)]}{\partial k_P^4}\right),\\
&\theta_3 = 2\alpha_1\alpha_2\left(\frac{\partial E[U(W)]}{\partial\sigma_p^2}\right), &&\theta_7 = 3\alpha_1\alpha_2^2\left(\frac{\partial E[U(W)]}{\partial s_P^3}\right), &&\theta_{11} = 4\alpha_1\alpha_2^3\left(\frac{\partial E[U(W)]}{\partial k_P^4}\right),\\
&\theta_4 = \alpha_1^3\left(\frac{\partial E[U(W)]}{\partial s_P^3}\right), &&\theta_8 = \alpha_1^4\left(\frac{\partial E[U(W)]}{\partial k_P^4}\right), &&\theta_{12} = 6\alpha_1^2\alpha_2^2\left(\frac{\partial E[U(W)]}{\partial k_P^4}\right).
\end{aligned} \tag{47}$$

Equations (46) and (47) are valid for all i.

A.2 Proof of the Asymptotic Information Matrix

The log of the likelihood function at time t in equation (9) is

$$\ln L_t(\theta) = h - \eta. \tag{48}$$

The first and second derivatives are given respectively by

$$\frac{\partial\ln L_t(\theta)}{\partial\theta} = \frac{\partial h}{\partial\theta} - \frac{\partial\eta}{\partial\theta}, \tag{49}$$

and

$$\frac{\partial^2\ln L_t(\theta)}{\partial\theta\partial\theta'} = \frac{\partial^2 h}{\partial\theta\partial\theta'} - \frac{\partial^2\eta}{\partial\theta\partial\theta'}. \tag{50}$$

The information matrix at time t is

$$
\begin{aligned}
I_t(\theta) &= -E\left[\frac{\partial^2 \ln L_t(\theta)}{\partial\theta\partial\theta'}\right] \\
&= \frac{\partial^2\eta}{\partial\theta\partial\theta'} - E\left[\frac{\partial^2 h}{\partial\theta\partial\theta'}\right].
\end{aligned}
\tag{51}
$$

Differentiating a first and second time gives

$$
\frac{\partial\eta}{\partial\theta} = \frac{\int\left(\frac{\partial h}{\partial\theta}\right)\exp(h)\,dr}{\int\exp(h)\,dr} = E\left[\frac{\partial h}{\partial\theta}\right].
\tag{52}
$$

$$
\begin{aligned}
\frac{\partial^2\eta}{\partial\theta\partial\theta'} &= \frac{\left(\int\left(\frac{\partial^2 h}{\partial\theta\partial\theta'}\right)\exp(h)\,dr + \int\left(\frac{\partial h}{\partial\theta}\right)\left(\frac{\partial h}{\partial\theta'}\right)\exp(h)\,dr\right)\left(\int\exp(h)\,dr\right)}{\left(\int\exp(h)\,dr\right)^2} \\
&\quad - \frac{\left(\int\left(\frac{\partial h}{\partial\theta}\right)\exp(h)\,dr\right)\left(\int\left(\frac{\partial h}{\partial\theta}\right)\exp(h)\,dr\right)}{\left(\int\exp(h)\,dr\right)^2} \\
&= E\left[\frac{\partial^2 h}{\partial\theta\partial\theta'}\right] + E\left[\frac{\partial h}{\partial\theta}\frac{\partial h}{\partial\theta'}\right] - E\left[\frac{\partial h}{\partial\theta}\right]E\left[\frac{\partial h}{\partial\theta'}\right].
\end{aligned}
\tag{53}
$$

Substituting equation (53) into the information matrix at time t in equation (51) yields

$$
\begin{aligned}
I_t(\theta) &= E\left[\frac{\partial^2 h}{\partial\theta\partial\theta'}\right] + E\left[\frac{\partial h}{\partial\theta}\frac{\partial h}{\partial\theta'}\right] - E\left[\frac{\partial h}{\partial\theta}\right]E\left[\frac{\partial h}{\partial\theta'}\right] - E\left[\frac{\partial^2 h}{\partial\theta\partial\theta'}\right] \\
&= E\left[\frac{\partial h}{\partial\theta}\frac{\partial h}{\partial\theta'}\right] - E\left[\frac{\partial h}{\partial\theta}\right]E\left[\frac{\partial h}{\partial\theta'}\right].
\end{aligned}
\tag{54}
$$

Finally, $I\left(\widehat{\theta}\right)$ is the asymptotic information matrix evaluated at $\widehat{\theta}$, that is

$$
\begin{aligned}
I\left(\widehat{\theta}\right) &= \sum_t I_t\left(\widehat{\theta}\right) \\
&= T\left(E\left[\frac{\partial h}{\partial\theta}\frac{\partial h}{\partial\theta'}\right]\bigg|_{\theta=\widehat{\theta}} - E\left[\frac{\partial h}{\partial\theta}\right]\bigg|_{\theta=\widehat{\theta}}E\left[\frac{\partial h}{\partial\theta'}\right]\bigg|_{\theta=\widehat{\theta}}\right).
\end{aligned}
\tag{55}
$$

A.3 Derivation of the Test Statistics

A.3.1 Statistics of Co-volatility

Consider the following bivariate generalized normal distribution with co-volatility

$$
\begin{aligned}
f(r_{1,t}, r_{2,t}) &= \exp\left[-\frac{1}{2}\left(\frac{1}{1-\rho^2}\right)\left(\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^2 + \left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)\right)\right. \\
&\qquad \left. +\theta_6\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^2\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^2 - \eta\right] \\
&= \exp\left[h - \eta\right],
\end{aligned} \tag{56}
$$

where

$$
\begin{aligned}
\eta &= \ln\iint \exp\left[-\frac{1}{2}\left(\frac{1}{1-\rho^2}\right)\left(\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^2 + \left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)\right)\right. \\
&\qquad \left. +\theta_6\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^2\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^2\right] dr_1 dr_2 \\
&= \ln\iint \exp\left[h\right] dr_1 dr_2,
\end{aligned} \tag{57}
$$

and

$$
\begin{aligned}
h &= -\frac{1}{2}\left(\frac{1}{1-\rho^2}\right)\left(\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^2 + \left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)\right) \\
&\qquad +\theta_6\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^2\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^2.
\end{aligned} \tag{58}
$$

A test of the restriction for normality is

$$
H_0 : \theta_6 = 0. \tag{59}
$$

Under the null hypothesis, the maximum likelihood estimators of the unknown parameters are

$$
\widehat{\mu}_i = \frac{1}{T}\sum_{t=1}^{T} r_{i,t}; \widehat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T}(r_{i,t} - \widehat{\mu}_i)^2; \widehat{\rho} = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_{1,t}-\widehat{\mu}_1}{\widehat{\sigma}_1}\right)\left(\frac{r_{2,t}-\widehat{\mu}_2}{\widehat{\sigma}_2}\right); i = 1, 2. \tag{60}
$$

Let the parameters of equation (56) be

$$
\theta = \left\{\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho, \theta_6\right\}. \tag{61}
$$

By taking the log function in equation (56), the log likelihood function at time t is given by

$$
\ln L_t(\theta) = -\frac{1}{2}\left(\frac{1}{1-\rho^2}\right)\left(\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^2 + \left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)\right)
$$
$$
+\theta_6\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^2\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^2 - \eta \tag{62}
$$
$$
= h - \eta,
$$

where h and η are given by equations (58) and (57).

Taking the expectations for the first and second derivatives of the log likelihood function $(\ln L_t(\theta))$ in equation (58) under the null hypothesis of the bivariate normality $(\theta_6 = 0)$ is

$$
E\left(\frac{\partial h}{\partial \mu_1}\right) = 0
$$
$$
E\left(\frac{\partial h}{\partial \mu_2}\right) = 0
$$
$$
E\left(\frac{\partial h}{\partial \sigma_1^2}\right) = \frac{1}{2}\frac{1}{\sigma_1^2}
$$
$$
E\left(\frac{\partial h}{\partial \sigma_2^2}\right) = \frac{1}{2}\frac{1}{\sigma_2^2}
$$
$$
E\left(\frac{\partial h}{\partial \rho}\right) = \frac{-\rho}{1-\rho^2}
$$
$$
E\left(\frac{\partial h}{\partial \theta_6}\right) = 1+2\rho^2,
$$

$$
E\left[\left(\frac{\partial h}{\partial \mu_1}\right)^2\right] = \frac{1}{1-\rho^2}\frac{1}{\sigma_1^2}
$$
$$
E\left[\left(\frac{\partial h}{\partial \mu_2}\right)^2\right] = \frac{1}{1-\rho^2}\frac{1}{\sigma_2^2}
$$
$$
E\left[\left(\frac{\partial h}{\partial \sigma_1^2}\right)^2\right] = -\frac{1}{4}\frac{1}{1-\rho^2}\left(\frac{1}{\sigma_1^2}\right)^2\left(2\rho^2-3\right)
$$
$$
E\left[\left(\frac{\partial h}{\partial \sigma_2^2}\right)^2\right] = -\frac{1}{4}\frac{1}{1-\rho^2}\left(\frac{1}{\sigma_2^2}\right)^2\left(2\rho^2-3\right)
$$
$$
E\left[\left(\frac{\partial h}{\partial \rho}\right)^2\right] = \left(\frac{1}{1-\rho^2}\right)^4\left(2\rho^6-3\rho^4+1\right)
$$
$$
E\left[\left(\frac{\partial h}{\partial \theta_6}\right)^2\right] = 9+72\rho^2+24\rho^4,
$$

$$E\left[\left(\frac{\partial h}{\partial \mu_1}\right)\left(\frac{\partial h}{\partial \mu_2}\right)\right] = -\frac{1}{\sigma_1}\frac{1}{\sigma_2}\left(\frac{\rho}{1-\rho^2}\right)$$

$$E\left[\left(\frac{\partial h}{\partial \mu_1}\right)\left(\frac{\partial h}{\partial \sigma_1^2}\right)\right] = 0$$

$$E\left[\left(\frac{\partial h}{\partial \mu_1}\right)\left(\frac{\partial h}{\partial \sigma_2^2}\right)\right] = 0$$

$$E\left[\left(\frac{\partial h}{\partial \mu_1}\right)\left(\frac{\partial h}{\partial \rho}\right)\right] = 0$$

$$E\left[\left(\frac{\partial h}{\partial \mu_1}\right)\left(\frac{\partial h}{\partial \theta_6}\right)\right] = 0,$$

$$E\left[\left(\frac{\partial h}{\partial \mu_2}\right)\left(\frac{\partial h}{\partial \sigma_1^2}\right)\right] = 0$$

$$E\left[\left(\frac{\partial h}{\partial \mu_2}\right)\left(\frac{\partial h}{\partial \sigma_2^2}\right)\right] = 0$$

$$E\left[\left(\frac{\partial h}{\partial \mu_2}\right)\left(\frac{\partial h}{\partial \rho}\right)\right] = 0$$

$$E\left[\left(\frac{\partial h}{\partial \mu_2}\right)\left(\frac{\partial h}{\partial \theta_6}\right)\right] = 0,$$

$$E\left[\left(\frac{\partial h}{\partial \sigma_1^2}\right)\left(\frac{\partial h}{\partial \sigma_2^2}\right)\right] = -\frac{1}{4}\frac{1}{\sigma_1^2}\frac{1}{\sigma_2^2}(2\rho^2-1)\left(\frac{1}{1-\rho^2}\right)$$

$$E\left[\left(\frac{\partial h}{\partial \sigma_1^2}\right)\left(\frac{\partial h}{\partial \rho}\right)\right] = -\rho\frac{1}{\sigma_1^2}\left(\frac{1}{1-\rho^2}\right)$$

$$E\left[\left(\frac{\partial h}{\partial \sigma_1^2}\right)\left(\frac{\partial h}{\partial \theta_6}\right)\right] = \frac{3}{2}\frac{1}{\sigma_1^2}\left(2\rho^2+1\right),$$

$$E\left[\left(\frac{\partial h}{\partial \sigma_2^2}\right)\left(\frac{\partial h}{\partial \rho}\right)\right] = -\rho\frac{1}{\sigma_2^2}\left(\frac{1}{1-\rho^2}\right)$$

$$E\left[\left(\frac{\partial h}{\partial \sigma_2^2}\right)\left(\frac{\partial h}{\partial \theta_6}\right)\right] = \frac{3}{2}\frac{1}{\sigma_2^2}(2\rho^2+1)$$

$$E\left[\left(\frac{\partial h}{\partial \rho}\right)\left(\frac{\partial h}{\partial \theta_6}\right)\right] = 3(-2\rho^3+\rho)\left(\frac{1}{1-\rho^2}\right),$$

where

$$\mu_i = E[r_{i,t}]$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^2\right] = 1$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)\left(\frac{r_{j,t} - \mu_j}{\sigma_j}\right)\right] = \rho, i \neq j$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^1\left(\frac{r_{j,t} - \mu_j}{\sigma_j}\right)^2\right] = 0$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^2\left(\frac{r_{j,t} - \mu_j}{\sigma_j}\right)^3\right] = 0$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^1\left(\frac{r_{j,t} - \mu_j}{\sigma_j}\right)^4\right] = 0$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^3\left(\frac{r_{j,t} - \mu_j}{\sigma_j}\right)^4\right] = 0$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^3\right] = 0$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^4\right] = 3$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^2\left(\frac{r_{j,t} - \mu_j}{\sigma_j}\right)^2\right] = 1 + 2\rho^2$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^1\left(\frac{r_{j,t} - \mu_j}{\sigma_j}\right)^3\right] = 3\rho$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^2\left(\frac{r_{j,t} - \mu_j}{\sigma_j}\right)^4\right] = 3 + 12\rho^2$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^3\left(\frac{r_{j,t} - \mu_j}{\sigma_j}\right)^3\right] = 9\rho + 6\rho^3$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^4\left(\frac{r_{j,t} - \mu_j}{\sigma_j}\right)^4\right] = 9 + 72\rho^2 + 24\rho^4.$$

The elements of the information matrix at time t, evaluated under the null are

$$I_{1,1,t} = E\left[\frac{\partial h}{\partial \mu_1}\frac{\partial h}{\partial \mu_1}\right] - E\left[\frac{\partial h}{\partial \mu_1}\right]E\left[\frac{\partial h}{\partial \mu_1}\right]$$

$$= \left(\frac{1}{1 - \rho^2}\right)\left(\frac{1}{\sigma_1^2}\right)$$

$$I_{1,2,t} = E\left[\frac{\partial h}{\partial \mu_1}\frac{\partial h}{\partial \mu_2}\right] - E\left[\frac{\partial h}{\partial \mu_1}\right]E\left[\frac{\partial h}{\partial \mu_2}\right]$$

$$= \left(\frac{1}{1-\rho^2}\right)\left(\frac{-\rho}{\sigma_1\sigma_2}\right)$$

$$I_{1,3,t} = E\left[\frac{\partial h}{\partial \mu_1}\frac{\partial h}{\partial \sigma_1^2}\right] - E\left[\frac{\partial h}{\partial \mu_1}\right]E\left[\frac{\partial h}{\partial \sigma_1^2}\right]$$

$$= 0$$

$$I_{1,4,t} = E\left[\frac{\partial h}{\partial \mu_1}\frac{\partial h}{\partial \sigma_2^2}\right] - E\left[\frac{\partial h}{\partial \mu_1}\right]E\left[\frac{\partial h}{\partial \sigma_2^2}\right]$$

$$= 0$$

$$I_{1,5,t} = E\left[\frac{\partial h}{\partial \mu_1}\frac{\partial h}{\partial \rho}\right] - E\left[\frac{\partial h}{\partial \mu_1}\right]E\left[\frac{\partial h}{\partial \rho}\right]$$

$$= 0$$

$$I_{1,6,t} = E\left[\frac{\partial h}{\partial \mu_1}\frac{\partial h}{\partial \theta_6}\right] - E\left[\frac{\partial h}{\partial \mu_1}\right]E\left[\frac{\partial h}{\partial \theta_6}\right]$$

$$= 0$$

$$I_{2,2,t} = E\left[\frac{\partial h}{\partial \mu_2}\frac{\partial h}{\partial \mu_2}\right] - E\left[\frac{\partial h}{\partial \mu_2}\right]E\left[\frac{\partial h}{\partial \mu_2}\right]$$

$$= \left(\frac{1}{1-\rho^2}\right)\left(\frac{1}{\sigma_2^2}\right)$$

$$I_{2,3,t} = E\left[\frac{\partial h}{\partial \mu_2}\frac{\partial h}{\partial \sigma_1^2}\right] - E\left[\frac{\partial h}{\partial \mu_2}\right]E\left[\frac{\partial h}{\partial \sigma_1^2}\right]$$

$$= 0$$

$$I_{2,4,t} = E\left[\frac{\partial h}{\partial \mu_2}\frac{\partial h}{\partial \sigma_2^2}\right] - E\left[\frac{\partial h}{\partial \mu_2}\right]E\left[\frac{\partial h}{\partial \sigma_2^2}\right]$$

$$= 0$$

$$I_{2,5,t} = E\left[\frac{\partial h}{\partial \mu_2}\frac{\partial h}{\partial \rho}\right] - E\left[\frac{\partial h}{\partial \mu_2}\right]E\left[\frac{\partial h}{\partial \rho}\right]$$

$$= 0$$

$$
\begin{aligned}
I_{2,6,t} &= E\left[\frac{\partial h}{\partial \mu_2}\frac{\partial h}{\partial \theta_6}\right] - E\left[\frac{\partial h}{\partial \mu_2}\right]E\left[\frac{\partial h}{\partial \theta_6}\right] \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
I_{3,3,t} &= E\left[\frac{\partial h}{\partial \sigma_1^2}\frac{\partial h}{\partial \sigma_1^2}\right] - E\left[\frac{\partial h}{\partial \sigma_1^2}\right]E\left[\frac{\partial h}{\partial \sigma_1^2}\right] \\
&= \left(\frac{1}{1-\rho^2}\right)\left(\frac{-\rho^2+2}{4\sigma_1^4}\right)
\end{aligned}
$$

$$
\begin{aligned}
I_{3,4,t} &= E\left[\frac{\partial h}{\partial \sigma_1^2}\frac{\partial h}{\partial \sigma_2^2}\right] - E\left[\frac{\partial h}{\partial \sigma_1^2}\right]E\left[\frac{\partial h}{\partial \sigma_2^2}\right] \\
&= \left(\frac{1}{1-\rho^2}\right)\left(\frac{-\rho^2}{4\sigma_1^2\sigma_2^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
I_{3,5,t} &= E\left[\frac{\partial h}{\partial \sigma_1^2}\frac{\partial h}{\partial \rho}\right] - E\left[\frac{\partial h}{\partial \sigma_1^2}\right]E\left[\frac{\partial h}{\partial \rho}\right] \\
&= \left(\frac{1}{1-\rho^2}\right)\left(\frac{-\rho}{2\sigma_1^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
I_{3,6,t} &= E\left[\frac{\partial h}{\partial \sigma_1^2}\frac{\partial h}{\partial \theta_6}\right] - E\left[\frac{\partial h}{\partial \sigma_1^2}\right]E\left[\frac{\partial h}{\partial \theta_6}\right] \\
&= \frac{(2\rho^2+1)}{\sigma_1^2}
\end{aligned}
$$

$$
\begin{aligned}
I_{4,4,t} &= E\left[\frac{\partial h}{\partial \sigma_2^2}\frac{\partial h}{\partial \sigma_2^2}\right] - E\left[\frac{\partial h}{\partial \sigma_2^2}\right]E\left[\frac{\partial h}{\partial \sigma_2^2}\right] \\
&= \left(\frac{1}{1-\rho^2}\right)\left(\frac{-\rho^2+2}{4\sigma_2^4}\right)
\end{aligned}
$$

$$
\begin{aligned}
I_{4,5,t} &= E\left[\frac{\partial h}{\partial \sigma_2^2}\frac{\partial h}{\partial \rho}\right] - E\left[\frac{\partial h}{\partial \sigma_2^2}\right]E\left[\frac{\partial h}{\partial \rho}\right] \\
&= \left(\frac{1}{1-\rho^2}\right)\left(\frac{-\rho}{2\sigma_2^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
I_{4,6,t} &= E\left[\frac{\partial h}{\partial \sigma_2^2}\frac{\partial h}{\partial \theta_6}\right] - E\left[\frac{\partial h}{\partial \sigma_2^2}\right]E\left[\frac{\partial h}{\partial \theta_6}\right] \\
&= \frac{(2\rho^2+1)}{\sigma_2^2}
\end{aligned}
$$

$$I_{5,5,t} = E\left[\frac{\partial h}{\partial \rho}\frac{\partial h}{\partial \rho}\right] - E\left[\frac{\partial h}{\partial \rho}\right]E\left[\frac{\partial h}{\partial \rho}\right]$$

$$= \left(\frac{1}{1-\rho^2}\right)\left(\frac{\rho^2+1}{1-\rho^2}\right)$$

$$I_{5,6,t} = E\left[\frac{\partial h}{\partial \rho}\frac{\partial h}{\partial \theta_6}\right] - E\left[\frac{\partial h}{\partial \rho}\right]E\left[\frac{\partial h}{\partial \theta_6}\right]$$

$$= 4\rho$$

$$I_{6,6,t} = E\left[\frac{\partial h}{\partial \theta_6}\frac{\partial h}{\partial \theta_6}\right] - E\left[\frac{\partial h}{\partial \theta_6}\right]E\left[\frac{\partial h}{\partial \theta_6}\right]$$

$$= (8 + 68\rho^2 + 20\rho^4).$$

The information matrix under the null hypothesis of the bivariate normality $(\theta_6 = 0)$ is

$$I\left(\widehat{\theta}\right) = T\left(E\left[\frac{\partial h}{\partial \theta}\frac{\partial h}{\partial \theta'}\right]\Big|_{\theta_6=0} - E\left[\frac{\partial h}{\partial \theta}\right]\Big|_{\theta_6=0}E\left[\frac{\partial h}{\partial \theta'}\right]\Big|_{\theta_6=0}\right) \tag{63}$$

$$= \left(\frac{T}{1-\widehat{\rho}^2}\right) \times$$

$$\begin{bmatrix} \frac{1}{\widehat{\sigma}_1^2} & \frac{-\widehat{\rho}}{\widehat{\sigma}_1\widehat{\sigma}_2} & 0 & 0 & 0 & 0 \\ \frac{-\widehat{\rho}}{\widehat{\sigma}_1\widehat{\sigma}_2} & \frac{1}{\widehat{\sigma}_2^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{-\widehat{\rho}^2+2}{4\widehat{\sigma}_1^4} & \frac{-\widehat{\rho}^2}{4\widehat{\sigma}_1^2\widehat{\sigma}_2^2} & \frac{-\widehat{\rho}}{2\widehat{\sigma}_1^2} & \frac{(2\widehat{\rho}^2+1)(1-\widehat{\rho}^2)}{\widehat{\sigma}_1^2} \\ 0 & 0 & \frac{-\widehat{\rho}^2}{4\widehat{\sigma}_1^2\widehat{\sigma}_2^2} & \frac{-\widehat{\rho}^2+2}{4\widehat{\sigma}_2^4} & \frac{-\widehat{\rho}}{2\widehat{\sigma}_2^2} & \frac{(2\widehat{\rho}^2+1)(1-\widehat{\rho}^2)}{\widehat{\sigma}_2^2} \\ 0 & 0 & \frac{-\widehat{\rho}}{2\widehat{\sigma}_1^2} & \frac{-\widehat{\rho}}{2\widehat{\sigma}_2^2} & \frac{\widehat{\rho}^2+1}{1-\widehat{\rho}^2} & 4\widehat{\rho}(1-\widehat{\rho}^2) \\ 0 & 0 & \frac{(2\widehat{\rho}^2+1)(1-\widehat{\rho}^2)}{\widehat{\sigma}_1^2} & \frac{(2\widehat{\rho}^2+1)(1-\widehat{\rho}^2)}{\widehat{\sigma}_2^2} & 4\widehat{\rho}(1-\widehat{\rho}^2) & (8+68\widehat{\rho}^2+20\widehat{\rho}^4)(1-\widehat{\rho}^2) \end{bmatrix},$$

so that

$$I\left(\widehat{\theta}\right)^{-1} = \left(\frac{1-\widehat{\rho}^2}{T}\right) \times \tag{64}$$

$$\begin{bmatrix} \frac{1}{\widehat{\sigma}_1^2} & \frac{-\widehat{\rho}}{\widehat{\sigma}_1\widehat{\sigma}_2} & 0 & 0 & 0 & 0 \\ \frac{-\widehat{\rho}}{\widehat{\sigma}_1\widehat{\sigma}_2} & \frac{1}{\widehat{\sigma}_2^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{-\widehat{\rho}^2+2}{4\widehat{\sigma}_1^4} & \frac{-\widehat{\rho}^2}{4\widehat{\sigma}_1^2\widehat{\sigma}_2^2} & \frac{-\widehat{\rho}}{2\widehat{\sigma}_1^2} & \frac{(2\widehat{\rho}^2+1)(1-\widehat{\rho}^2)}{\widehat{\sigma}_1^2} \\ 0 & 0 & \frac{-\widehat{\rho}^2}{4\widehat{\sigma}_1^2\widehat{\sigma}_2^2} & \frac{-\widehat{\rho}^2+2}{4\widehat{\sigma}_2^4} & \frac{-\widehat{\rho}}{2\widehat{\sigma}_2^2} & \frac{(2\widehat{\rho}^2+1)(1-\widehat{\rho}^2)}{\widehat{\sigma}_2^2} \\ 0 & 0 & \frac{-\widehat{\rho}}{2\widehat{\sigma}_1^2} & \frac{-\widehat{\rho}}{2\widehat{\sigma}_2^2} & \frac{\widehat{\rho}^2+1}{1-\widehat{\rho}^2} & 4\widehat{\rho}(1-\widehat{\rho}^2) \\ 0 & 0 & \frac{(2\widehat{\rho}^2+1)(1-\widehat{\rho}^2)}{\widehat{\sigma}_1^2} & \frac{(2\widehat{\rho}^2+1)(1-\widehat{\rho}^2)}{\widehat{\sigma}_2^2} & 4\widehat{\rho}(1-\widehat{\rho}^2) & (8+68\widehat{\rho}^2+20\widehat{\rho}^4)(1-\widehat{\rho}^2) \end{bmatrix}^{-1}.$$

Evaluating the gradient for θ_6 under the null hypothesis gives

$$
\begin{aligned}
\left.\frac{\partial \ln L_t(\theta)}{\partial \theta}\right|_{\theta_6 = 0} &= \sum_{t=1}^{T}\left(\frac{\partial h}{\partial \theta_6}\right) - T\left(\frac{\partial \eta}{\partial \theta_6}\right) \\
&= \sum_{t=1}^{T}\left(\frac{r_{1,t} - \mu_1}{\sigma_1}\right)^2\left(\frac{r_{2,t} - \mu_2}{\sigma_2}\right)^2 - T\left(1 + 2\rho^2\right).
\end{aligned}
\tag{65}
$$

By taking the first derivatives of the log likelihood function, the score function under H_0 is given as

$$
\begin{aligned}
q\left(\widehat{\theta}\right) &= \left.\frac{\partial \ln L_t(\theta)}{\partial \theta}\right|_{\theta_6 = 0} \\
&= \left[\begin{array}{cccccc} 0 & 0 & 0 & 0 & 0 & \sum_{t=1}^{T}\left(\frac{r_{1,t}-\widehat{\mu}_1}{\widehat{\sigma}_1}\right)^2\left(\frac{r_{2,t}-\widehat{\mu}_2}{\widehat{\sigma}_2}\right)^2 - T\left(1 + 2\widehat{\rho}^2\right)\end{array}\right]'.
\end{aligned}
\tag{66}
$$

Substituting equations (64) and (66) into the Lagrange multiplier statistic in equation (15) gives

$$
\begin{aligned}
LM_1 &= q\left(\widehat{\theta}\right)' I\left(\widehat{\theta}\right)^{-1} q\left(\widehat{\theta}\right) \\
&= \frac{1}{T\left(4\widehat{\rho}^4 + 16\widehat{\rho}^2 + 4\right)}\left[\sum_{t=1}^{T}\left(\frac{r_{1,t} - \widehat{\mu}_1}{\widehat{\sigma}_1}\right)^2\left(\frac{r_{2,t} - \widehat{\mu}_2}{\widehat{\sigma}_2}\right)^2 - T\left(1 + 2\widehat{\rho}^2\right)\right]^2 \\
&= \left(\frac{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_{1,t}-\widehat{\mu}_1}{\widehat{\sigma}_1}\right)^2\left(\frac{r_{2,t}-\widehat{\mu}_2}{\widehat{\sigma}_2}\right)^2 - \left(1 + 2\widehat{\rho}^2\right)}{\sqrt{\frac{\left(4\widehat{\rho}^4 + 16\widehat{\rho}^2 + 4\right)}{T}}}\right)^2.
\end{aligned}
\tag{67}
$$

A.3.2   A Statistic of Co-kurtosis

Consider the following bivariate generalized normal distribution with co-kurtosis given as

$$
\begin{aligned}
f(r_{1,t}, r_{2,t}) &= \exp\left[-\frac{1}{2}\left(\frac{1}{1 - \rho^2}\right)\left(\left(\frac{r_{1,t} - \mu_1}{\sigma_1}\right)^2 + \left(\frac{r_{2,t} - \mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{r_{1,t} - \mu_1}{\sigma_1}\right)\left(\frac{r_{2,t} - \mu_2}{\sigma_2}\right)\right)\right. \\
&\qquad \left. + \theta_7\left(\frac{r_{1,t} - \mu_1}{\sigma_1}\right)^1\left(\frac{r_{2,t} - \mu_2}{\sigma_2}\right)^3 - \eta\right] \\
&= \exp\left[h - \eta\right],
\end{aligned}
\tag{68}
$$

where

$$
\begin{aligned}
\eta &= \ln \iint \exp \left[-\frac{1}{2} \left(\frac{1}{1-\rho^2} \right) \left(\left(\frac{r_{1,t} - \mu_1}{\sigma_1} \right)^2 + \left(\frac{r_{2,t} - \mu_2}{\sigma_2} \right)^2 - 2\rho \left(\frac{r_{1,t} - \mu_1}{\sigma_1} \right) \left(\frac{r_{2,t} - \mu_2}{\sigma_2} \right) \right) \right. \\
&\qquad \left. + \theta_7 \left(\frac{r_{1,t} - \mu_1}{\sigma_1} \right)^1 \left(\frac{r_{2,t} - \mu_2}{\sigma_2} \right)^3 \right] dr_1 dr_2 \qquad (69) \\
&= \ln \iint \exp \left[h \right] dr_1 dr_2,
\end{aligned}
$$

and

$$
\begin{aligned}
h &= -\frac{1}{2} \left(\frac{1}{1-\rho^2} \right) \left(\left(\frac{r_{1,t} - \mu_1}{\sigma_1} \right)^2 + \left(\frac{r_{2,t} - \mu_2}{\sigma_2} \right)^2 - 2\rho \left(\frac{r_{1,t} - \mu_1}{\sigma_1} \right) \left(\frac{r_{2,t} - \mu_2}{\sigma_2} \right) \right) \\
&\qquad + \theta_7 \left(\frac{r_{1,t} - \mu_1}{\sigma_1} \right)^1 \left(\frac{r_{2,t} - \mu_2}{\sigma_2} \right)^3. \qquad (70)
\end{aligned}
$$

A test of the restriction for normality is

$$
H_0 : \theta_7 = 0.
$$

Let the parameters of equation (70) be

$$
\theta = \left\{ \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho, \theta_7 \right\}. \qquad (71)
$$

By taking the log function of equation (68), the log likelihood function at time t is given by

$$
\begin{aligned}
\ln L_t(\theta) &= -\frac{1}{2} \left(\frac{1}{1-\rho^2} \right) \left(\left(\frac{r_{1,t} - \mu_1}{\sigma_1} \right)^2 + \left(\frac{r_{2,t} - \mu_2}{\sigma_2} \right)^2 - 2\rho \left(\frac{r_{1,t} - \mu_1}{\sigma_1} \right) \left(\frac{r_{2,t} - \mu_2}{\sigma_2} \right) \right) \\
&\qquad + \theta_7 \left(\frac{r_{1,t} - \mu_1}{\sigma_1} \right)^1 \left(\frac{r_{2,t} - \mu_2}{\sigma_2} \right)^3 - \eta \qquad (72) \\
&= h - \eta,
\end{aligned}
$$

where h and η are given by equations (70) and (69).

Taking the expectations for the first and second derivatives of the log likelihood function

in equation (72) under the null hypothesis of the bivariate normality ($\theta_7 = 0$) is

$$E\left(\frac{\partial h}{\partial \theta_7}\right) = 3\rho$$

$$E\left[\left(\frac{\partial h}{\partial \theta_7}\right)^2\right] = 15 + 90\rho^2$$

$$E\left[\left(\frac{\partial h}{\partial \mu_1}\right)\left(\frac{\partial h}{\partial \theta_7}\right)\right] = 0$$

$$E\left[\left(\frac{\partial h}{\partial \mu_2}\right)\left(\frac{\partial h}{\partial \theta_7}\right)\right] = 0$$

$$E\left[\left(\frac{\partial h}{\partial \sigma_1^2}\right)\left(\frac{\partial h}{\partial \theta_7}\right)\right] = \frac{3\rho}{\sigma_1^2}$$

$$E\left[\left(\frac{\partial h}{\partial \sigma_2^2}\right)\left(\frac{\partial h}{\partial \theta_7}\right)\right] = \frac{6\rho}{\sigma_2^2}$$

$$E\left[\left(\frac{\partial h}{\partial \rho}\right)\left(\frac{\partial h}{\partial \theta_7}\right)\right] = 3(1 - 2\rho^2)\left(\frac{1}{1 - \rho^2}\right),$$

where

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^6\right] = 15$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^1\left(\frac{r_{j,t} - \mu_j}{\sigma_j}\right)^5\right] = 15\rho, i \neq j$$

$$E\left[\left(\frac{r_{i,t} - \mu_i}{\sigma_i}\right)^2\left(\frac{r_{j,t} - \mu_j}{\sigma_j}\right)^6\right] = 15 + 90\rho^2, i \neq j.$$

The elements of the information matrix at observation t, evaluated under the null are

$$\begin{aligned}
I_{1,6,t} &= E\left[\frac{\partial h}{\partial \mu_1}\frac{\partial h}{\partial \theta_7}\right] - E\left[\frac{\partial h}{\partial \mu_1}\right]E\left[\frac{\partial h}{\partial \theta_7}\right] \\
&= 0
\end{aligned}$$

$$\begin{aligned}
I_{2,6,t} &= E\left[\frac{\partial h}{\partial \mu_2}\frac{\partial h}{\partial \theta_7}\right] - E\left[\frac{\partial h}{\partial \mu_2}\right]E\left[\frac{\partial h}{\partial \theta_7}\right] \\
&= 0
\end{aligned}$$

$$\begin{aligned}
I_{3,6,t} &= E\left[\frac{\partial h}{\partial \sigma_1^2}\frac{\partial h}{\partial \theta_7}\right] - E\left[\frac{\partial h}{\partial \sigma_1^2}\right]E\left[\frac{\partial h}{\partial \theta_7}\right] \\
&= \left(\frac{3\rho}{2\sigma_1^2}\right)
\end{aligned}$$

$$
\begin{aligned}
I_{4,6,t} &= E\left[\frac{\partial h}{\partial \sigma_2^2}\frac{\partial h}{\partial \theta_7}\right] - E\left[\frac{\partial h}{\partial \sigma_2^2}\right]E\left[\frac{\partial h}{\partial \theta_7}\right] \\
&= \left(\frac{9\rho}{2\sigma_2^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
I_{5,6,t} &= E\left[\frac{\partial h}{\partial \rho}\frac{\partial h}{\partial \theta_7}\right] - E\left[\frac{\partial h}{\partial \rho}\right]E\left[\frac{\partial h}{\partial \theta_7}\right] \\
&= 3
\end{aligned}
$$

$$
\begin{aligned}
I_{6,6,t} &= E\left[\frac{\partial h}{\partial \theta_7}\frac{\partial h}{\partial \theta_7}\right] - E\left[\frac{\partial h}{\partial \theta_7}\right]E\left[\frac{\partial h}{\partial \theta_7}\right] \\
&= (15 + 81\rho^2).
\end{aligned}
$$

The information matrix under the null hypothesis of the bivariate normality ($\theta_7 = 0$) is given as

$$
I\left(\widehat{\theta}\right) = T\left(E\left[\frac{\partial h}{\partial \theta}\frac{\partial h}{\partial \theta'}\right]\Big|_{\theta_7 = 0} - E\left[\frac{\partial h}{\partial \theta}\right]\Big|_{\theta_7 = 0}E\left[\frac{\partial h}{\partial \theta'}\right]\Big|_{\theta_7 = 0}\right) \tag{73}
$$

$$
= \left(\frac{T}{1-\widehat{\rho}^2}\right)
\begin{bmatrix}
\frac{1}{\widehat{\sigma}_1^2} & \frac{-\widehat{\rho}}{\widehat{\sigma}_1\widehat{\sigma}_2} & 0 & 0 & 0 & 0 \\
\frac{-\widehat{\rho}}{\widehat{\sigma}_1\widehat{\sigma}_2} & \frac{1}{\widehat{\sigma}_2^2} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{-\widehat{\rho}^2+2}{4\widehat{\sigma}_1^4} & \frac{-\widehat{\rho}^2}{4\widehat{\sigma}_1^2\widehat{\sigma}_2^2} & \frac{-\widehat{\rho}}{2\widehat{\sigma}_1^2} & \frac{3\widehat{\rho}(1-\widehat{\rho}^2)}{2\widehat{\sigma}_1^2} \\
0 & 0 & \frac{-\widehat{\rho}^2}{4\widehat{\sigma}_1^2\widehat{\sigma}_2^2} & \frac{-\widehat{\rho}^2+2}{4\widehat{\sigma}_2^4} & \frac{-\widehat{\rho}}{2\widehat{\sigma}_2^2} & \frac{9\widehat{\rho}(1-\widehat{\rho}^2)}{2\widehat{\sigma}_2^2} \\
0 & 0 & \frac{-\widehat{\rho}}{2\widehat{\sigma}_1^2} & \frac{-\widehat{\rho}}{2\widehat{\sigma}_2^2} & \frac{\widehat{\rho}^2+1}{1-\widehat{\rho}^2} & 3(1-\widehat{\rho}^2) \\
0 & 0 & \frac{3\widehat{\rho}(1-\widehat{\rho}^2)}{2\widehat{\sigma}_1^2} & \frac{9\widehat{\rho}(1-\widehat{\rho}^2)}{2\widehat{\sigma}_2^2} & 3(1-\widehat{\rho}^2) & (15+81\widehat{\rho}^2)(1-\widehat{\rho}^2)
\end{bmatrix},
$$

so that

$$
I\left(\widehat{\theta}\right)^{-1} = \left(\frac{1-\widehat{\rho}^2}{T}\right)
\begin{bmatrix}
\frac{1}{\widehat{\sigma}_1^2} & \frac{-\widehat{\rho}}{\widehat{\sigma}_1\widehat{\sigma}_2} & 0 & 0 & 0 & 0 \\
\frac{-\widehat{\rho}}{\widehat{\sigma}_1\widehat{\sigma}_2} & \frac{1}{\widehat{\sigma}_2^2} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{-\widehat{\rho}^2+2}{4\widehat{\sigma}_1^4} & \frac{-\widehat{\rho}^2}{4\widehat{\sigma}_1^2\widehat{\sigma}_2^2} & \frac{-\widehat{\rho}}{2\widehat{\sigma}_1^2} & \frac{3\widehat{\rho}(1-\widehat{\rho}^2)}{2\widehat{\sigma}_1^2} \\
0 & 0 & \frac{-\widehat{\rho}^2}{4\widehat{\sigma}_1^2\widehat{\sigma}_2^2} & \frac{-\widehat{\rho}^2+2}{4\widehat{\sigma}_2^4} & \frac{-\widehat{\rho}}{2\widehat{\sigma}_2^2} & \frac{9\widehat{\rho}(1-\widehat{\rho}^2)}{2\widehat{\sigma}_2^2} \\
0 & 0 & \frac{-\widehat{\rho}}{2\widehat{\sigma}_1^2} & \frac{-\widehat{\rho}}{2\widehat{\sigma}_2^2} & \frac{\widehat{\rho}^2+1}{1-\widehat{\rho}^2} & 3(1-\widehat{\rho}^2) \\
0 & 0 & \frac{3\widehat{\rho}(1-\widehat{\rho}^2)}{2\widehat{\sigma}_1^2} & \frac{9\widehat{\rho}(1-\widehat{\rho}^2)}{2\widehat{\sigma}_2^2} & 3(1-\widehat{\rho}^2) & (15+81\widehat{\rho}^2)(1-\widehat{\rho}^2)
\end{bmatrix}.
$$
$$\tag{74}$$

Evaluating the gradient for θ_7 under the null hypothesis gives

$$
\begin{aligned}
\frac{\partial \ln L_t(\theta)}{\partial \theta_7}\Big|_{\theta_7 = 0} &= \sum_{t=1}^{T}\left(\frac{\partial h}{\partial \theta_7}\right) - T\left(\frac{\partial \eta}{\partial \theta_7}\right) \\
&= \sum_{t=1}^{T}\left(\frac{r_{1,t}-\mu_1}{\sigma_1}\right)^1\left(\frac{r_{2,t}-\mu_2}{\sigma_2}\right)^3 - T(3\rho).
\end{aligned}
\tag{75}
$$

43

By taking the first derivatives of the log likelihood function, the score function under H_0 is given as

$$q\left(\widehat{\theta}\right) = \left.\frac{\partial \ln L_t(\theta)}{\partial \theta}\right|_{\theta_7 = 0} \tag{76}$$

$$= \left[\begin{array}{cccccc} 0 & 0 & 0 & 0 & 0 & \sum_{t=1}^{T}\left(\frac{r_{1,t}-\widehat{\mu}_1}{\widehat{\sigma}_1}\right)^1 \left(\frac{r_{2,t}-\widehat{\mu}_2}{\widehat{\sigma}_2}\right)^3 - T\left(3\widehat{\rho}\right) \end{array} \right]'.$$

Substituting equations (74) and (76) into the Lagrange multiplier statistic in equation (15) gives

$$LM = q\left(\widehat{\theta}\right)' I\left(\widehat{\theta}\right)^{-1} q\left(\widehat{\theta}\right) \tag{77}$$

$$= \frac{1}{T\left(18\widehat{\rho}^2 + 6\right)} \left[\sum_{t=1}^{T}\left(\frac{r_{1,t}-\widehat{\mu}_1}{\widehat{\sigma}_1}\right)^1 \left(\frac{r_{2,t}-\widehat{\mu}_2}{\widehat{\sigma}_2}\right)^3 - T\left(3\widehat{\rho}\right) \right]^2$$

$$= \left(\frac{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_{1,t}-\widehat{\mu}_1}{\widehat{\sigma}_1}\right)^1 \left(\frac{r_{2,t}-\widehat{\mu}_2}{\widehat{\sigma}_2}\right)^3 - \left(3\widehat{\rho}\right)}{\sqrt{\frac{\left(18\widehat{\rho}^2 + 6\right)}{T}}} \right)^2.$$

Table 1: Linear, asymmetric and extremal co-moments of returns for eight equity markets with the US banking sector during the non-crisis (NC) and crisis (C) periods. The non-crisis period is from April 1, 2005 to February 29, 2008 and the crisis period is from March 3, 2008 to August 31, 2009.

| | Linear | | Asymmetric | | | |
| | Correlation | | Co-skewness$_{12}$ | | Co-skewness$_{21}$ | |
Country	NC	C	NC	C	NC	C
Asia						
Hong Kong	-0.020	0.166	0.476	0.223	0.345	0.085
Korea	-0.082	0.173	0.193	0.228	-0.202	0.031
Europe						
France	0.266	0.404	0.155	0.012	-0.495	-0.105
Germany	0.230	0.404	0.058	-0.103	-0.417	-0.177
UK	0.262	0.380	0.214	-0.064	-0.397	-0.191
Latin America						
Chile	0.309	0.376	0.045	-0.165	-0.440	-0.066
Mexico	0.492	0.585	-0.023	0.025	-0.158	-0.036
North America						
US	0.799	0.772	-0.360	-0.116	-0.291	-0.122

| | Extremal | | | | | |
	Co-kurtosis$_{13}$		Co-kurtosis$_{31}$		Co-volatility$_{22}$	
Asia						
Hong Kong	1.436	1.392	1.271	1.063	4.853	1.494
Korea	-1.728	1.792	-1.186	0.682	2.202	1.344
Europe						
France	0.868	2.178	0.296	2.457	2.922	2.307
Germany	1.289	1.967	0.699	2.544	2.398	2.263
UK	0.887	2.144	0.840	2.482	2.651	2.494
Latin America						
Chile	2.189	2.467	1.759	2.317	2.966	2.207
Mexico	2.670	3.222	3.677	3.169	2.830	2.818
North America						
US	4.491	4.031	6.283	3.967	5.045	3.549

Notes: The statistics are measured in terms of:-
Correlation: returns of market j and returns of US bank.
Co-skewness$_{12}$: returns of market j and squared returns of US bank.
Co-skewness$_{21}$: squared returns of market j and returns of US bank.
Co-kurtosis$_{13}$: returns of market j and cubed returns of US bank.
Co-kurtosis$_{31}$: cubed returns of market j and returns of US bank.
Co-volatility$_{22}$: squared returns of market j and squared returns
         of US bank.

Table 2:
Summary of the restrictions on the parameters in equations (27) and (28) for the size and power tests.

Restrictions	Size	Power		
		Experiment I	Experiment II	Experiment III
Sample size				
T_x	500	500	500	500
T_y	$[100, 500]$	500	500	500
Parameters in (27)				
ρ_x	0	0	0	0
$\theta_{x,i}, \forall i = 4, ..., 12$	0	0	0	0
Parameters in (28)				
ρ_y	0	0	0	0
$\theta_{y,4}$	0	0	$(0, 0.9]$	0
$\theta_{y,5}$	0	0	0	0
$\theta_{y,6}$	0	$(0, 0.9]$	0.5	0.5
$\theta_{y,7}$	0	0	0	$(0, 0.9]$
$\theta_{y,i}, \forall i = 8, ..., 12$	0	0	0	0

Note: $f(x_1, x_2) = \exp\left(\theta_{x,1}x_1^2 + \theta_{x,2}x_2^2 + \theta_{x,3}x_1x_2 + \theta_{x,4}x_1x_2^2 + \theta_{x,5}x_1^2x_2 - \theta_{x,6}x_1^2x_2^2 \right.$
$\left. -\theta_{x,7}x_1x_2^3 - \theta_{x,8}x_1^3x_2 + \theta_{x,9}x_1^3 + \theta_{x,10}x_2^3 + \theta_{x,11}x_1^4 + \theta_{x,12}x_2^4 - \eta_x\right)$ in (27)
and $f(y_1, y_2) = \exp\left(\theta_{y,1}y_1^2 + \theta_{y,2}y_2^2 + \theta_{y,3}y_1y_2 + \theta_{y,4}y_1y_2^2 + \theta_{y,5}y_1^2y_2 - \theta_{y,6}y_1^2y_2^2 \right.$
$\left. -\theta_{y,7}y_1y_2^3 - \theta_{y,8}y_1^3y_2 + \theta_{y,9}y_1^3 + \theta_{y,10}y_2^3 + \theta_{y,11}y_1^4 + \theta_{y,12}y_2^4 - \eta_y\right)$ in (28)

Table 3:
Simulated sizes of alternative contagion tests based on the different sizes of the crisis sample period T_y. The non-crisis sample period is $T_x = 500$. The size of the CS_{12}, CK_{13}, and CV_{22} tests are based on the 5% asymptotic χ_1^2 critical values with one degree of freedom. The results are based on 500,000 replications.

	Sample size of crisis period (T_y)				
Tests	500	400	300	200	100
CS_{12}	0.049	0.048	0.048	0.049	0.047
CK_{13}	0.049	0.050	0.048	0.047	0.043
CV_{22}	0.049	0.048	0.047	0.046	0.040

Table 4:
Simulated critical values of alternative test statistics of contagion for small crisis period sample sizes T_y. The non-crisis sample period is $T_x = 500$. The results are based on 500,000 replications.

Sig. level/	CS_{12} test			CK_{13} test			CV_{22} test		
T_y	2.5%	5%	10%	2.5%	5%	10%	2.5%	5%	10%
15	3.89	2.92	2.04	2.73	1.93	1.30	2.36	1.82	1.35
30	4.51	3.39	2.35	3.99	2.74	1.79	3.52	2.55	1.80
60	4.81	3.64	2.53	4.75	3.31	2.16	4.31	3.13	2.17
90	4.85	3.70	2.59	4.89	3.49	2.32	4.62	3.38	2.32
150	4.91	3.74	2.63	5.01	3.66	2.47	4.89	3.61	2.49
200	4.97	3.79	2.66	5.04	3.71	2.53	4.94	3.70	2.54

Table 5: Summary moments of equity returns for eight equity markets and banking sectors during the non-crisis (NC) and crisis (C) periods. The non-crisis period is from April 1, 2005 to February 29, 2008 and the crisis period is from March 3, 2008 to August 31, 2009.

Country	Moment							
	Mean		Std.Dev		Skewness		kurtosis	
	NC	C	NC	C	NC	C	NC	C
(A) Equity markets								
Asia								
Hong Kong	0.078	-0.053	0.886	1.951	-0.107	0.218	11.446	6.917
Korea	0.084	-0.091	0.929	1.843	-0.676	-0.082	5.672	11.011
Europe								
France	0.042	-0.084	0.839	1.872	-0.528	0.251	8.641	7.100
Germany	0.081	-0.092	0.938	2.005	-0.526	0.087	6.584	5.505
UK	0.031	-0.097	0.780	1.819	-0.312	0.079	6.715	7.082
Latin America								
Chile	0.078	-0.016	0.713	1.275	-0.470	-0.404	8.319	10.430
Mexico	0.113	-0.063	1.122	1.920	-0.230	0.198	5.453	7.043
North America								
US	0.022	-0.065	0.557	1.498	-0.514	0.166	4.937	6.728
(B) Banking sectors								
Asia								
Hong Kong	0.219	-0.119	1.542	2.206	0.005	-0.008	4.781	4.200
Korea	0.046	0.001	1.289	2.321	0.170	-0.379	4.102	6.522
Europe								
France	-0.003	-0.050	1.016	2.792	-0.400	0.264	6.838	5.446
Germany	0.024	-0.312	1.173	3.584	-1.389	-1.010	15.664	10.701
UK	-0.018	-0.118	0.807	2.651	-0.032	0.051	7.888	6.562
Latin America								
Chile	0.050	-0.016	0.974	1.819	-0.181	-0.425	5.368	12.854
Mexico	0.118	-0.051	1.624	2.518	-0.170	0.097	4.716	5.971
North America								
US	-0.020	-0.141	0.785	3.796	-0.063	0.000	9.386	5.563

Table 6: Test statistics for contagion based on changes in asymmetric dependence during the global financial crisis of 2008-09.

Country	(A) Equity markets		(B) Banking sectors	
	$CS_{12}^{(a)}$	$CS_{21}^{(b)}$	$CS_{12}^{(a)}$	$CS_{21}^{(b)}$
Asia				
Hong Kong	0.03	9.01**	0.14	8.08**
Korea	0.40	3.17*	12.94**	4.06**
Europe				
France	0.18	0.21	7.78**	0.97
Germany	0.30	0.18	5.68**	1.33
UK	2.08	1.19	17.63**	4.13**
Latin America				
Chile	14.91**	1.83	3.42*	2.53
Mexico	0.19	0.64	0.13	0.01
North America				
US	0.04	0.04	n.a.	n.a.

Notes: (a) CS_{12}: Co-skewness contagion test measured in terms of the returns of the US banking sector and squared returns of market j. (b) CS_{21}: Co-skewness contagion test measured in terms of the squared returns of the US banking sector and returns of market j. * denotes significance at the 10% level and ** significance at the 5% level.

Table 7: Test statistics for contagion based on changes in extremal dependence during the global financial crisis of 2008-09.

Country	(A) Equity markets			(B) Banking sectors		
	$CK_{13}^{(a)}$	$CK_{31}^{(b)}$	$CV_{22}^{(c)}$	$CK_{13}^{(a)}$	$CK_{31}^{(b)}$	$CV_{22}^{(c)}$
Asia						
Hong Kong	90.68**	210.36**	0.87	1.42	23.73**	2.42
Korea	137.86**	170.76**	14.66**	11.88**	107.80**	9.83**
Europe						
France	111.21**	155.68**	86.34**	218.07**	116.85**	157.81**
Germany	83.51**	151.50**	55.82**	351.51**	79.54**	167.91**
UK	151.14**	100.17**	88.87**	246.91**	83.32**	128.01**
Latin America						
Chile	25.56**	7.93**	2.62	62.29**	9.71**	1.31
Mexico	46.40**	21.23**	20.45**	140.49**	104.35**	95.01**
North America						
US	28.42**	2.08	5.77**	n.a.	n.a.	n.a.

Notes: (a) CK_{13} : Co-kurtosis contagion test measured in terms of the returns of the US banking sector and cubed returns of market j. (b) CK_{31} : Co-kurtosis contagion test measured in terms of the cubed returns of the US banking sector and returns of market j. (c) CV_{22} : Co-volatility contagion test measured in terms of the squared returns of the US banking sector and squared returns of market j. ** denotes the significance of contagion at the 5% level.