International Review of Economics and Finance 46 (2016) 55–77

Contents lists available at ScienceDirect

International Review of Economics and Finance

journal homepage: www.elsevier.com/locate/iref

Extreme risk spillover effects in world gold markets and the global financial crisis

Gang-Jin Wang [a,b,c,*], Chi Xie [a,b], Zhi-Qiang Jiang [d,c], H. Eugene Stanley [c]

[a] Business School, Hunan University, Changsha 410082, China
[b] Center of Finance and Investment Management, Hunan University, Changsha 410082, China
[c] Center for Polymer Studies and Department of Physics, Boston University, Boston MA02215, USA
[d] School of Business, East China University of Science and Technology, Shanghai 200237, China

ARTICLE INFO

Article history:
Received 29 January 2016
Received in revised form
23 August 2016
Accepted 23 August 2016
Available online 26 August 2016

JEL classification:
G15
C12
C58

Keywords:
Gold
Risk spillover
Financial crisis
VaR

ABSTRACT

Using the approach of Granger causality in risk, we investigate extreme risk spillover effects among four major world gold markets (London, New York, Tokyo and Shanghai) before and after the recent global financial crisis. We find (i) that there are strong extreme risk spillover effects between London and New York, and London and Shanghai, (ii) that most of the extreme risk spillovers to Tokyo and Shanghai are from New York rather than from London, but London leads New York in risk spillovers, (iii) that extreme risk spillover effects from Tokyo and Shanghai to New York are limited, but those to London play an important role, and (iv) that extreme risk spillover effects between Tokyo and Shanghai are weak or negligible. We also find that extreme risk is more quickly transmitted in the post-crisis era than in the pre-crisis era, an effect that is related to the safe-haven or risk-hedging property or the speculative value of gold.

1. Introduction

Because gold is a special commodity that combines the attributes of a general commodity, a currency, and a financial product, it simultaneously has commercial, monetary, and financial functions. Although the demonetization of gold in 1976 weakened its monetary function, the financial function of gold has increased because gold is sought as an investment or a tool for hedging when uncertainty in a financial system increases beyond a certain point. In particular, the 2008 financial crisis and the European sovereign debt crisis triggered by the US subprime crisis in 2007 caused traders to acquire gold as a safe haven or hedge against other financial instruments (Baur and Lucey, 2010, Baur and McDermott, 2010, Joy, 2011, Reboredo, 2013a, Reboredo, 2013b).

Although gold is an asset that is traded world-wide, the main centers are the London Over-the-Counter (OTC) market, the New York Commodity Exchange (COMEX) market, the Tokyo Commodity Exchange (TOCOM) market, and the Shanghai Futures Exchange (SHFE) market (Lucey et al., 2014). According to Murray (2011) and Lucey et al. (2012), these four gold markets in 2011 carry 98% of the world-wide gold trading volume, the London gold market is the biggest gold trading center with 86.75% of the trading volume (around 90% of which are spot transactions), and the rest three gold markets and others share the rest of the volume (i.e., New York 9.89%, Shanghai 1.38%, and Tokyo 0.98%). In view of the worldwide feature of

* Corresponding author at: Business School, Hunan University, Changsha 410082, China.
E-mail addresses: wanggangjin@hnu.edu.cn (G.-J. Wang), xiechi@hnu.edu.cn (C. Xie), zqjiang@ecust.edu.cn (Z.-Q. Jiang), hes@bu.edu (H.E. Stanley).

http://dx.doi.org/10.1016/j.iref.2016.08.004
1059-0560/

gold trading, a natural question is: what are correlations between gold prices and information spillover effects among different gold markets? Little research pays attention to this question. Despite this topic has emerged on the current work (see a review by O'Connor et al., 2015), most studies focus on correlations or information spillover effects in mean and volatility between two gold markets and ignore their spillover effects in extreme risk. In response to this lack, our goal here is to investigate extreme risk spillover effects among the four major world gold markets mentioned above.

Over the past 15 years gold prices have been volatile. Beginning in 2002 they increased steadily, reached their record high of 1895USD per troy ounce on 5 September 2011, and have declined since that date. High price volatility is characteristic of precious metal markets, and gold traders seeking to minimize risk often diversify their investment portfolios internationally, but world-wide market integration introduces contagion into the system and risk is transmitted between different gold markets when there are market shocks (e.g., financial crises, political crises, or regional wars). Being able to understand and quantify extreme risk (including downside risk and upside risk) spillover effects between different gold markets is thus essential to gold market participants and policy-makers. For investors, understanding the mechanism of extreme risk spillover effects across different gold markets is crucial in managing asset risk and constructing asset portfolios. When investors face extreme adverse market fluctuations and co-movements in an uncertain economic situation, for example, the knowledge of which gold markets are sources of extreme risk spillovers of their holding assets is hugely beneficial in evaluating the possible risk of the holding gold assets and helps investors reconstruct their diversified portfolios and improve their hedging strategies. When policy-makers assess the stability of gold market and monitor its risk, for example, the information of how the extreme risk spillover effects across different gold markets will respond to a financial crisis and market crash would help them formulate more effective policies for influencing their markets.

In our study we use the Granger causality in risk approach proposed by Hong et al. (2009) to examine extreme risk spillovers across the four gold markets. The Granger causality in risk uses the value-at-risk (VaR) measurement introduced by the J.P. Morgan and widely used in both the financial industry and academia to quantify extreme risk in financial markets. VaR is the maximum loss of a given asset portfolio over a specific time horizon with a pre-specified confidence level. Although volatility is a traditional measure of risk and its spillover effects are widely investigated in the literature (e.g., Baklaci et al., 2016; Batten and Lucey, 2010; Baur, 2012; Cheung and Ng, 1996; Hong, 2001), VaR is chosen rather than volatility because it overcomes two drawbacks when using volatility, i.e., (i) the risk of loss measured by volatility is usually underestimated and symmetrical, and (ii) volatility cannot model extreme risk. When the actual loss of a given asset portfolio exceeds the VaR at the given confidence level (e.g., 95%), we say the risk happens at the fixed confidence level. According to the Granger causality in risk, one market is said to Granger cause another market in risk if the past risk profile of the first market assists in predicting the future risk profile of the second market with a greater accuracy than if the second market were to predict its future risk profile using only its own past risk profile. Thus the Granger causality in risk can help market participants and regulators predict and monitor risk in gold markets.

The empirical data we use in our study are the daily closing (fixing) prices of the London, New York, Tokyo, and Shanghai gold markets during the period 30 October 2002–30 October 2015. To analyze the influence of the recent global financial crisis, we divide the this time period into two subperiods, i.e., before and after the financial crisis.[1] We then use the variance–covariance method to estimate the VaRs of these four gold markets at different periods based on the ARMA-(T)GARCH-GED (generalized error distribution) model that captures the "stylized facts" of financial time series, including autocorrelation, volatility clustering, "leverage effect," and fat tails. Finally, we analyze extreme risk spillover effects among the four gold markets by employing the Granger causality in risk.

Our accomplishments here are three-fold and groundbreaking.

(i) We investigate extreme risk spillover effects in world gold markets and examine both the downside and upside risk spillover effects at the 99% and 95% confidence levels. To overcome the non-synchronous trading effect in the gold markets, we develop a modified statistic of the one-way Granger causality in risk that can be extended to the study of risk spillover in other international financial markets.

(ii) We find that extreme risk spillover effects between London and New York and between London and Shanghai are strong and significant, and that there are strong feedback effects between these two pairs of gold markets. The extreme risk spillover from New York to Shanghai is usually stronger than that from Shanghai to New York. Most of the extreme risk spillovers to Tokyo and Shanghai come from New York rather than from London, but the level of one-way Granger causality in risk between London and New York indicates that London leads New York in risk spillovers. Extreme risk spillover effects between Tokyo and Shanghai are weak or negligible, suggesting that the interaction between them is inconspicuous.

(iii) We investigate how extreme risk spillover effects in gold markets before the recent global financial crisis differ from those after the crisis. This is a new contribution to the literature of financial crisis. Empirically we find that extreme risk is more quickly transmitted between gold markets in the post-crisis era than in the pre-crisis era. We attribute this finding to the safe-haven or risk-hedging role or the speculative value of gold during crisis periods.

[1] In our study, we consider the recent global financial crisis including the US subprime crisis, the 2008 financial crisis, and the European sovereign debt crisis.

Overall, our findings provide important references and suggest potential applications of great value to academic researchers, market participants, and market regulators.

We organize our paper as follows. In Section 2 we review the literature. In Section 3 we describe econometric methodologies. In Section 4 we present the empirical data and results. In Section 5 we draw our conclusions.

2. Literature review

The research reported here draws on two topics described in the literature. The first is a focus on correlations or information spillover effects between different gold markets. The early work on this topic began with that reported by Laulajainen (1990), which presents an investigation of the daily prices of the New York, London, and Hong Kong gold markets from 1 October 1987 to 30 September 1988 and uses the level-VAR mode to determine which of the three markets is dominate. They find that the New York gold market had more influence over the other two than the other two had on New York. Dhillon et al. (1997) present an empirical study of the volatility and information flows between the New York and Tokyo gold markets during the period from July 1987 through May 1992. They find that the intraday volatility of the New York gold market is greater than that of the Tokyo gold market, which indicates greater information flows in the New York gold market. Following the idea of Dhillon et al. (1997), Xu and Fung (2005) extend the study by employing a bivariate autoregressive moving average-generalized autoregressive conditional heteroscedasticity (ARMA-GARCH) model to examine cross-market linkages between the New York and Tokyo gold markets during the period 1994–2001. The authors find that the information flows sourced in the New York gold market transmit to the Tokyo gold market in terms of returns, once again implying that the New York market is the world leader in global gold markets. The opposite outcome is reported by Lin et al. (2008), which is an analysis of the dynamic correlations between the New York and Tokyo gold markets from January 1991 to July 2006. Using a bivariate GARCH model, they find that the Tokyo gold market appears to lead the New York gold market. Unlike the previous work that focused on the linkages between the New York and Tokyo gold markets, Kumar and Pandey (2011) use the Baba–Engle–Kraft–Kroner (BEKK) GARCH model to study the cross-market linkages between the New York and Indian (i.e., the Multi Commodity Exchange, MCE) gold markets from 5 May 2005 to 7 April 2008 in terms of return and volatility spillovers. They find that the volatility spillover effect of the New York gold market on the Indian gold market is more than its effect on New York, indicating that the Indian market is a satellite market that assimilates information from the New York market.

More recently, Lucey et al. (2013) utilize the information sharing model to investigate the origin of gold prices and focus on the daily prices of two major gold markets (New York and London) from January 1986 to July 2012. They find that the London gold market may be dominant in accordance with the price information shares but the dominant market is dynamic with no clear link to macroeconomic or political events. Chang et al. (2013) use an augmenting level-VAR model to study the dynamic correlations among the daily prices in five gold markets (New York, London, Tokyo, Hong Kong, and Taiwan) from 2007 to 2010. They find that (i) relationships between the New York and London gold markets and those among the Japanese, Hong Kong, and Taiwan gold markets are bi-directional; (ii) the London gold market is impacted by the other four gold markets; and (iii) the New York gold market is the leading center of gold information and thus is a straightforward influence on the other four gold markets. An empirical study reported by Fuangkasem et al. (2014) examines relationships among the New York, Tokyo, and Indian gold markets using sample data of 5-minute trading prices from April 2011 to August 2011. Using the vector error correction model (VECM) and the information share model, the authors find that these three gold markets are co-integrated and that the New York gold mark dominates the other two markets. Lucey et al. (2014) analyze the integration of the New York, London, Tokyo, and Shanghai gold markets using the daily prices set from 9 January 2008 to 9 October 2013 and the spillover index method. The authors conclude that (i) the Shanghai gold market is an isolated market; (ii) the New York and London gold markets are the strongest integrated pair of gold markets; and (iii) the Tokyo gold market is clearly influenced by the New York and London gold markets and has no significant and consistent effect on other markets. Baklaci et al. (2016) detect the cross-market volatility linkages among the New York, Tokyo, Shanghai, India, Turkey, and Taiwan gold markets from 2 September 2008 to 20 December 2012. Using the VECH-MGARCH model they find that (i) remarkable volatility spillover effects exist among Shanghai, Indian, and Taiwan gold markets; (ii) there is a long-run volatility linkage between the New York gold market and all the other markets except Shanghai; and (iii) the Shanghai gold market is relatively isolated. Recent work reported by Hauptfleisch et al. (2016) employs intraday data during the period 1 January 1997 to 30 November 2014 to uncover which of the New York and London gold markets determines gold prices. They find that both markets have influence but that, on average, the New York market takes the lead. Many sophisticated models have been employed to produce well-documented descriptions of correlations or information spillover effects in global gold markets, but to the best of our knowledge the Granger causality in risk has not been used to study the spillover effects in world gold markets from the perspective of risk.

The second topic in the literature concerns measuring the information spillover effects in financial markets. C.W.J. Granger systematically examined the causality and spillover effects among financial time series, including the Granger causality in mean (Granger, 1969), the Granger causality in variance (or volatility) (Granger et al., 1986), and the general Granger causality (Granger, 1980). Based on the idea of the general Granger causality, Hong et al. (2009) propose the Granger causality in risk using the VaR measure to examine the extreme risk (downside and upside risks) spillover effects in financial markets. Therefore prior to the most recent literature the measurements of the information spillover effects can be categorized as either (i) mean spillover, (ii) volatility spillover, or (iii) risk spillover. These have been used to describe the relationship and co-movement between different financial markets, and they play an important role in understanding information flow among markets. The existing literature has utilized numerous analytical methods of studying the Granger causality and information spillover effects, and these methods fall into two groups.

The first group of methods uses linear regression, including VAR, VAR-GARCH, VECM, and VECM-GARCH, to examine the Granger causality in mean or volatility and is widely used to study spillover effects in gold markets (see, e.g., Baklaci et al., 2016; Chang et al., 2013; Fuangkasem et al., 2014; Kumar and Pandey, 2011). However this linear regression approach (i) considers only the influence of a limited lag and the linear correlation and fails to detect the complex causality between highly heterogeneous financial time series, and (ii) its accuracy can be strongly affected by the "stylized facts" of financial time series, including autocorrelation, non-stationarity, multicollinearity, and heteroscedasticity.

The second group of methods uses a cross-correlation function (CCF) that overcomes the shortcomings of the traditional regression models. The CCF was first proposed by Haugh (1976), extended by McLeod and Li (1983), Koch and Yang (1986), Cheung and Ng (1996), Hong (1996), Hong (2001), Hong et al. (2009), and is robust to distributional assumptions. A two-stage method is used to compute the CCF for the test of Ganger causality. The first stage employs the univariate financial time series models, e.g., AR and ARMA-(T)GARCH, to uncover the stylized features of financial time series. The second stage uses the resulting residual series or their variations (e.g., conditional means and variances) to generate the CCF statistic in order to test the null hypothesis that there is no Granger causality between two financial time series. The Granger causality in risk proposed by Hong et al. (2009) is based on an improved CCF and is widely and successfully applied in different financial markets to quantify their extreme risk spillover effects (see, e.g., de Araújo and Garcia, 2013; Balboa et al., 2015; Du and He, 2015; Fan et al., 2008; Hong et al., 2004; Hwang and Kim, 2015; Lee and Lee, 2009; Liu et al., 2008; Pan and Zhang, 2007; Zhou, 2013).[2] For example, Hong et al. (2004) describe the extreme risk spillover effects between Chinese stock markets and international stock markets. Pan and Zhang (2007) use the GARCH-GED model to estimate the conditional downside and upside VaRs of WTI and Daqing crude oil prices and find extreme risk spillover effects between these two oil markets. Fan et al. (2008) examine the VaRs of WTI and Brent oil prices and their risk spillover effects. Liu et al. (2008) describe the spillover effects in risk between Chinese copper futures and spot markets. Other applications include extreme risk spillover effects between different stock markets (de Araújo and Garcia, 2013; Hwang and Kim, 2015), stock markets and foreign exchange markets (Lee and Lee, 2009), international real estate investment trust markets (Zhou, 2013), and oil and stock markets (Du and He, 2015). All this previous empirical research indicates that the Granger causality in risk can accurately describe and measure extreme risk spillover effects between different markets, and this motives us to utilize it in our study.

3. Methodology

3.1. VaR estimation

Given a specific time period and a confidence level of $(1-\alpha)$ in which $\alpha \in (0, 1)$, VaR is the maximum loss of a given asset portfolio with a probability α. Statistically, VaR is the α-quantile of the conditional probability distribution of returns of asset portfolio. Following Pan and Zhang (2007), Fan et al. (2008), and Liu et al. (2008), we employ the left α-quantile of returns of gold prices to quantify the downside VaR, i.e., the loss of sales revenue for gold providers (or producers) due to the sharp fall in gold prices. We employ the right α-quantile to measure the upside VaR, i.e., the increased expense for gold consumers due to the extreme rise in gold prices. Mathematically, given the gold returns R_t, the downside and upside VaRs at the confidence level of $(1-\alpha)$ are respectively

$$\Pr(R_t < -V_t(\text{down})|\Phi_{t-1}) = \alpha \tag{1}$$

and

$$\Pr(R_t > V_t(\text{up})|\Phi_{t-1}) = \alpha, \tag{2}$$

where $\Phi_{t-1} = \{R_{t-1}, R_{t-2}, \ldots, R_1\}$ is the information set available at $t-1$.

There are three ways of estimating VaR, (i) historical simulation, (ii) variance–covariance, (iii) and Monte Carlo simulation. Following a popular VaR estimation method—RiskMetrics™ proposed by Morgan (1996)—we use the variance–covariance method and the GARCH-type model to calculate the VaR of gold returns, but we modify RiskMetrics™, which assumes that asset returns follow a Gaussian distribution, and employ the ARMA(p,q)-(T)GARCH(r,s)-GED model to estimate the conditional mean and volatility of gold returns.[3] The proposed model can both capture the "styled facts" of gold returns,

[2] Note that in the working-paper stage of Hong et al. (2009) the application papers initiated the use of Granger causality in risk.

[3] Other approaches (e.g., simulation and quantile regression) can also be used to estimate VaRs. For example, Engle and Manganelli (2004) propose a widely used VaR estimation method, i.e., the conditional autoregressive value at risk (CAViaR) model, which is based on an autoregressive process and the quantile regression. Kuester et al. (2006) compare the predictive performance of many existing VaR methods and find (i) that a hybrid approach of a fat-tailed (e.g., a skewed-t distribution and the mixed GED) AR-GARCH model and the extreme value theory-based method has the best performance, followed by an improved historical simulation method, and (ii) that none of the CAViaR models perform well. They also find that only conditionally heteroskedastic models obtain acceptable predictions, and this supports our decision to estimate VaRs of gold returns with the ARMA-(T)GARCH-GED model.

e.g., autocorrelation and volatility clustering, and detect the "leverage effect" and fat tails. This approach is in line with previous investigations of the volatility of gold returns. For example, Batten and Lucey (2010) and Hammoudeh et al. (2010) use the ARMA-GARCH model to examine volatility in the gold futures market, and Baur (2012) uses the TGARCH model to describe the asymmetric volatility in the gold market.

The gold returns R_t can be modeled by an ARMA(p,q)-GARCH (r,s)-GED as

$$R_t = \mu_t + \varepsilon_t = c + \sum_{j=1}^{p} \phi_j R_{t-j} + \sum_{j=1}^{q} \varphi_j \varepsilon_{t-j} + \varepsilon_t, \tag{3}$$

$$\varepsilon_t = \eta_t \sqrt{h_t}, \tag{4}$$

$$h_t = \alpha_0 + \sum_{j=1}^{r} \beta_j h_{t-j} + \sum_{j=1}^{s} \alpha_j \varepsilon_{t-j}^2, \tag{5}$$

$$\eta_t \sim iid\ \mathrm{GED}(v), \tag{6}$$

where Eqs. (3) and (5) are the conditional mean and variance equations, respectively, μ_t is the conditional mean, c and ε_t are the constant and innovation terms respectively, ϕ_j and φ_j are autoregressive (AR) and moving average (MA) coefficients of order j, respectively, h_t is the conditional variance of the innovation ε_t with certain restrictions $\left(r \geq 0,\ s > 0,\ \alpha_0 > 0,\ \alpha_j \geq 0\ (j = 1, 2, \ldots, s),\ \beta_j \geq 0\ (j = 1, 2, \ldots, r), \text{ and } \sum_{j=1}^{s} \alpha_j + \sum_{j=1}^{r} \beta_j < 1\right)$, and η_t is the standardized residual that follows the generalized error distribution (GED). According to Nelson (1991), the GED captures the fat tails of the standardized residuals of returns because its tails are fatter than the normal distribution and thinner than the uniform distribution. The probability density function of the GED is defined

$$f(\eta) = \frac{v \exp\left[-0.5|\eta/\lambda|^v\right]}{\lambda 2^{[(1+v)/v]} \Gamma(1/v)}, \quad -\infty < \eta < \infty,\ 0 < v \leq \infty, \tag{7}$$

where $\Gamma(\cdot)$ is the gamma function, $\lambda \equiv \left[2^{-2/v}\Gamma(1/v)/\Gamma(3/v)\right]^{1/2}$, and v is the degree-of-freedom of the GED, which is also the tail-thickness parameter. When $v = 2$ the GED reduces to the standard normal distribution. When $v < 2$ the tail of the GED is thicker than in the standard normal distribution, and when $v > 2$ it is thinner.

Because Baur (2012) reports that there is significant asymmetric volatility in gold markets, we use the threshold GARCH (TGARCH) model to capture the "leverage effect" of asset returns. The "leverage effect" of gold returns indicates that the current volatility caused by previous positive and negative gold return shocks is asymmetric. Following Glosten et al. (1993) and Zakoian (1994), the conditional variance equation of the TGARCH(r,s) model is defined as

$$h_t = \alpha_0 + \sum_{j=1}^{r} \beta_j h_{t-j} + \gamma d_{t-1} \varepsilon_{t-1}^2 + \sum_{j=1}^{s} \alpha_j \varepsilon_{t-j}^2, \tag{8}$$

where $d_{t-1} = 1$ if $\varepsilon_{t-1} < 0$ and $d_{t-1} = 0$ otherwise, $d_{t-1} = 1$ and $d_{t-1} = 0$ are the effects of negative shocks and positive shocks (bad news and good news), respectively, at time $t - 1$, and γ quantifies the difference between the effects of previous positive shocks and negative shocks on the current volatility of gold. Whenever $\gamma \neq 0$, there are asymmetric impacts of lagged positive shocks and lagged negative shocks on the volatility. When $\gamma > 0$ the influence of previous negative shocks on the current volatility is larger than the previous positive shocks, and when $\gamma < 0$ it is smaller.

Using the ARMA-(T)GARCH-GED model and the variance–covariance approach, the downside VaR and upside VaR for gold returns are respectively estimated as

$$V_t(\mathrm{down}) = -\mu_t - z_\alpha \sqrt{h_t} \tag{9}$$

and

$$V_t(\mathrm{up}) = \mu_t + z_{1-\alpha} \sqrt{h_t}, \tag{10}$$

where z_α is the left α-quantile of the GED for the standardized residuals, i.e., $F(z_\alpha) = \alpha$, and $z_{1-\alpha} = -z_\alpha$.

To evaluate the adequacy of the VaR model for estimating extreme risks of four gold markets, we employ five backtesting measures: (i) the number of failure days (violations), (ii) the failure rate, (iii) the likelihood ratio (LR) test of unconditional coverage, (iv) of independent coverage, and (v) of conditional coverage. For a detailed introduction to these backtesting techniques, see Appendix A.

3.2. *Granger causality in risk*

We employ the Granger causality in risk proposed by Hong et al. (2009) to investigate the extreme risk spillover effects among the four gold markets. The Granger causality in risk is an extension of the general Granger causality proposed by Granger (1980) that tests whether the occurrence of past risks in one market can aid in forecasting the occurrence of future risks in another market. We first use the risk indicator introduced by Hong et al. (2009) and as an example take the downside VaR, defined as

$$Z_{mt} = \mathbf{1}(R_{mt} < -V_{mt}), \quad m = 1, 2, \tag{11}$$

where R_{mt} and V_{mt} are the return series and the corresponding downside VaR estimates of gold market m, and $\mathbf{1}(\cdot)$ is the indictor function found in Eq. (A.1), i.e., the risk indicator takes a value of one when the actual loss exceeds the downside VaR estimate, otherwise it takes a value of zero.

Let $\{R_{1t}\}$ and $\{R_{2t}\}$ be the return series of gold markets 1 and 2, respectively. According to Hong et al. (2009), to test the one-way downside risk spillover effects from gold market 2 to gold market 1,[4] the null hypothesis and its alternative hypothesis of Granger causality in risk are

$$H^0: E(Z_{1t}|\Phi_{1(t-1)}) = E(Z_{1t}|\Phi_{t-1}), \tag{12}$$

and

$$H^1: E(Z_{1t}|\Phi_{1(t-1)}) \neq E(Z_{1t}|\Phi_{t-1}), \tag{13}$$

where $\Phi_{t-1} = \{\Phi_{1(t-1)}, \Phi_{2(t-1)}\}$, $\Phi_{1(t-1)} = \{R_{1(t-1)}, R_{1(t-2)}, \ldots, R_{11}\}$, and $\Phi_{2(t-1)} = \{R_{2(t-1)}, R_{2(t-2)}, \ldots, R_{21}\}$.

If $\{\hat{Z}_{1t}\}$ and $\{\hat{Z}_{2t}\}$ are two estimated series of risk indictors of gold markets 1 and 2, the sample cross-covariance function is defined

$$\hat{C}(j) \equiv \begin{cases} T^{-1} \sum_{t=1+j}^{T} \left(\hat{Z}_{1t} - \hat{\alpha}_1\right)\left(\hat{Z}_{2(t-j)} - \hat{\alpha}_2\right), & 0 \le j \le T-1, \\ T^{-1} \sum_{t=1-j}^{T} \left(\hat{Z}_{1(t+j)} - \hat{\alpha}_1\right)\left(\hat{Z}_{2t} - \hat{\alpha}_2\right), & 1-T \le j < 0, \end{cases} \tag{14}$$

where $\hat{\alpha}_m$ is the sample mean of $\left\{\hat{Z}_{mt}\right\}$, m=1,2, and T is the total number of observations in the return series. Then the sample cross-correlation function (CCF) between $\left\{\hat{Z}_{1t}\right\}$ and $\left\{\hat{Z}_{2t}\right\}$ is defined

$$\hat{\rho}(j) \equiv \frac{\hat{C}(j)}{\hat{S}_1\hat{S}_2}, \tag{15}$$

where $\hat{S}_m^2$ is the sample variance of $\left\{\hat{Z}_{mt}\right\}$.

To test the one-way Granger causality in risk from gold market 2 to gold market 1, we use the kernel-based statistic proposed by Hong (2001) and Hong et al. (2009), i.e.,

$$Q_1(M) = \left[T \sum_{j=1}^{T-1} k^2(j/M)\hat{\rho}^2(j) - C_{1T}(M)\right] / \left[D_{1T}(M)\right]^{1/2}, \tag{16}$$

where the centering and standardization constants are defined as

$$C_{1T} = \sum_{j=1}^{T-1} (1 - j/T)k^2(j/M) \tag{17}$$

and

$$D_{1T} = 2\sum_{j=1}^{T-1} (1 - j/T)(1 - (j+1)/T)k^4(j/M). \tag{18}$$

Here $k(\cdot)$ is a kernel function that assigns weights to various lags. In our case, the Daniell kernel is chosen because according to Hong (2001) and Hong et al. (2009) the simulation results indicate that its performance is optimal, i.e., $k(x) = \sin(\pi x)/(\pi x)$.[5] M is the largest effective lag truncation order, which indicates how many lags are used to analyze the risk spillover effects between two gold markets. Lags are taken into consideration because in real-world markets gold investors need time to understand and respond to past information, thus the risk spillover effects have time-lags.

Hong et al. (2009) also propose a two-way Granger causality in risk to test risk spillover effects (including instantaneous risk spillover effects) between two markets, where the statistic is defined

[4] The test for upside risk spillover effects is similar to that of downside risk spillover effects and is not presented due to space limitations.
[5] Note that when x=0 the Daniell kernel $k(x) = 1$ because $\lim_{x\to 0}k(x) = \lim_{x\to 0}\frac{\sin(\pi x)}{\pi x} = 1$.

Table 1
The trading hours of the London, New York, Tokyo, and Shanghai gold markets.

Market	Hours (Greenwich Mean Time)	Hours (Local Time)	Hours (New York Time)
Tokyo	00:00 am–06:15 am	09:00 am–03:15 pm	07:00 pm–01:15 am
Shanghai	01:00 am–03:30 am	09:00 am–11:30 am	08:00 pm–10:30 pm
	05:30 am–07:30 am	01:30 pm–03:30 pm*	00:30 am–02:30 am
London	Morning fix: 10:30 am	Morning fix: 10:30 am	Morning fix: 05:30 am
	Afternoon fix: 03:00 pm	Afternoon fix: 03:00 pm	Afternoon fix: 10:00 am
New York	01:20 pm–06:30 pm	08:20 am–01:30 pm	08:20 am–01:30 pm

Notes: This table shows the gold trading hours of the London Over-the-Counter (OTC) market, the New York Commodity Exchange (COMEX) market, the Tokyo commodity exchange (TOCOM) market, and the Shanghai Gold Exchange (SGE) market. This table only presents the day session for the Tokyo and Shanghai gold markets. Their night sessions start from 04:30 pm to 04:00 am and from 09:00 pm to 02:30 am (Local Time), respectively. For a regular business day in these two markets, one clearing period corresponds to the previous business day's night session (e.g., Shanghai gold market from 09:00 pm) plus today's day session (until 03:30 pm).
Source: Ntim et al. (2015) and authors' elaborations from the websites of these four gold markets. For details, see the following four websites: (i) http://www.lbma.org.uk/pricing-and-statistics (the London OTC market), (ii) http://www.cmegroup.com/trading-hours.html#metals (the New York COMEX market), (iii) http://www.tocom.or.jp/guide/youkou/gold/index.html (the TOCOM market), and (iv) http://www.en.sge.com.cn/rules-regulations/rules/sgerules/523893.shtml (the SGE market).
* For the Shanghai Futures Exchange (SHFE) market, its afternoon session starts from 01:30 pm to 03:00 pm (Local time).

$$Q_2(M) = \left[T \sum_{|j|=1}^{T-1} k^2(j/M)\hat{\rho}^2(j) - C_{2T}(M) \right] / \left[D_{2T}(M) \right]^{1/2}, \tag{19}$$

and the centering and standardization constants are

$$C_{2T} = \sum_{|j|=1}^{T-1} (1 - |j|/T) k^2(j/M) \tag{20}$$

and

$$D_{2T} = 2\left[1 + \hat{\rho}^4(0) \right] \sum_{|j|=1}^{T-1} (1 - |j|/T)(1 - (|j| + 1)/T) k^4(j/M). \tag{21}$$

Here $Q_2(M)$ takes into consideration all possible instantaneous cross-correlations between $\left\{ \hat{Z}_{1t} \right\}$ and $\left\{ \hat{Z}_{2t} \right\}$.

Note that non-synchronous trading is common in world financial markets, i.e., national financial markets (e.g., gold markets) are in different time zones and have different opening and closing times (Malliaris and Urrutia, 1992). In our case, the non-synchronous trading effect can be a problem when testing one-way Granger causality in risk from an Asian gold market (e.g., Shanghai) to an American gold market (e.g., New York) because the opening and closing times of the Shanghai market are prior to those of the New York market (see Table 1). For example, if an important world event (e.g., the discovery of a new source of gold) occurs in China or some other Asian county and is announced while the Shanghai gold market on a given trading day (e.g., Wednesday) is active, the Wednesday closing price of New York gold market will reflect this information, i.e., the closing price of the Shanghai gold market on day t will influence the closing price of the New York gold market on the same calendar day t. The one-way Granger causality in risk based on Eq. (16) ignores this non-synchronous trading effect. To this end, we follow Cheung and Ng (1996) and Lu et al. (2014) and modify the statistic of one-way Granger causality in risk by taking into consideration the CCF with a lag order of zero, i.e.,

$$Q_3(M) = \left[T \sum_{j=0}^{T-1} k^2(j/M)\hat{\rho}^2(j) - C_{3T}(M) \right] / \left[D_{3T}(M) \right]^{1/2}, \tag{22}$$

where the centering and standardization constants are defined as

$$C_{3T} = \sum_{j=0}^{T-1} (1 - j/T) k^2(j/M) \tag{23}$$

and

$$D_{3T} = 2 \sum_{j=0}^{T-1} (1 - j/T)(1 - (j + 1)/T) k^4(j/M). \tag{24}$$

Economically speaking, using a CCF with a lag order of zero in Eq. (22) allows us to take into consideration the instantaneous risk spillover effects from the Shanghai gold market to the New York gold market. In other words, the statistic

$Q_3(M)$ allows us to consider the influence of the price information of the Shanghai gold market on day t on the price information of the New York gold market on the same calendar day t. Thus using the statistic $Q_3(M)$ we can fix the non-synchronous trading effect.

According to Hong et al. (2009), under the null hypothesis, $Q_i(M)$ (i = 1, 2, 3) obeys an asymptotically standard normal distribution, i.e., $Q_i(M) \to N(0, 1)$. If the value of $Q_i(M)$ is greater than the critical value of the right tail of the standard normal distribution, the null hypothesis is rejected, which suggests that there is one-way or two-way Ganger causality in risk from market 2 to market 1 or between them.

4. Empirical data and results

In this section, we first introduce the empirical data of the London, New York, Tokyo, and Shanghai gold markets and make a primary analysis on the gold returns. We then conduct the following procedure for examining extreme risk spillover effects across the four gold markets: (i) we employ the ARMA-(T)GARCH-GED model and the variance–covariance approach to estimate (downside and upside) VaRs for four gold markets at different periods, (ii) we use backtesting techniques to evaluate the accuracy and reliability of VaRs, and (iii) we utilize the statistics (including $Q_1(M)$, $Q_2(M)$, and $Q_3(M)$) of the Granger causality in risk to examine extreme risk spillover effects across these four gold markets.

4.1. Data and primary analysis

As noted in Section 1, gold is mainly traded in the London OTC market, the New York COMEX market, the TOCOM market, and the SHFE market in terms of trading volume. The London OTC market is the world's largest gold spot market and the New York COMEX market is the world's largest gold futures market, and together they dominate worldwide gold price discovery (Hauptfleisch et al., 2016; Lucey et al., 2013). Following Lucey et al. (2013, 2014) and Hauptfleisch et al. (2016), we employ the gold spot prices of the London OTC market and the gold futures prices of the New York COMEX market and the TOCOM marekt as the empirical data for the London, New York, and Tokyo gold markets, respectively. The Shanghai gold market has two exchange markets: (i) the Shanghai Gold Exchange (SGE) market for gold spot trading and (ii) the Shanghai Futures Exchange (SHFE) market for gold futures trading. We use the spot price rather than the futures price as the empirical data for the Shanghai gold market because (i) China's gold futures trading began on 9 January 2008 in the SHFE market and thus there is no trading data prior to the global financial crisis for the gold futures market, and (ii) although trading volumes of the London and New York gold markets are both larger than those of the Shanghai gold market, the SGE market has the world's largest physical gold trading volume and thus is a special and interesting representative among world gold markets.[6] Thus we use data from these four gold markets, specifically the daily closing (fixing) gold prices from the London OTC market, the New York COMEX market, the TOCOM market, and the SGE market during the period from 30 October 2002 to 30 October 2015. In the London gold market we select the afternoon 3:00 p.m. fixing prices (local time). According to the London Bullion Market Association (LBMA), the standard bar used for settlement and delivery in the London OTC market is the London Good Delivery gold bar that weighs between 350 troy ounces and 430 troy ounces and has a minimum fineness of 99.5%.[7] In the New York and Tokyo gold futures markets, we select the closing prices of the near-month contract on a continuous rolling basis. The standard for the delivery of gold futures contracts in the New York COMEX market is 100 troy ounces of gold with a minimum purity of 99.5%. Each standard gold futures contract in the TOCOM market represents 1 kg of deliverable grade gold with a minimum fineness of 99.99%. In the Shanghai gold market, we select the closing spot prices of Gold Au99.95, which is the most actively traded gold product in the SGE market and is the underlying asset of the gold futures traded in the SHFE market. The beginning date of the sample is determined by the SGE market because it opened on 30 October 2002. All empirical data were obtained from Thomson Reuters Eikon.

Fig. 1 shows the gold price trends for London, New York, Tokyo, and Shanghai during the 30 October 2002 to 30 October 2015 period. Although the overall price trends are similar in the four markets, in recent years the price trends in the Tokyo gold market differ from those in the other three. As noted above in the methodology section, we focus on the daily returns of gold markets, which is defined as $R_{mt} = 100 \times (\ln P_{mt} - \ln P_{m(t-1)})$, where P_{mt} and $P_{m(t-1)}$ are the closing (fixing) prices of the gold market m on days t and $t-1$, respectively. During the entire 13-year period there are 3372 observations of each return series.

During this 13-year period, global gold markets were affected by the US subprime crisis, the 2008 financial crisis, and the European sovereign debt crisis. To compare the extreme risk spillover effects in the four gold markets before and after the financial crisis, the returns of each gold market are divided into (i) subperiod I (the pre-crisis era) from 31 October 2002 to 29 June 2007 and (ii) subperiod II (the post-crisis era) from 2 July 2007 to 30 October 2015.[8] Because the four gold markets

[6] To check the validity of our results, in Appendix B we conduct a robustness test of the extreme risk spillover effects using the Shanghai gold futures data. The testing data span the period from 9 January 2008 to 30 October 2015, and the testing results are consistent with our central findings.

[7] See http://www.lbma.org.uk/market-tools

[8] We choose the date 2 July 2007 as the break point of the recent global financial crisis for two reasons. (i) In the economics and finance literature many scholars designate the beginning of the collapse of the US subprime mortgage market in July and August of 2007 as the starting of the recent global financial crisis. For example, in the study by Acharya et al. (2009) of the causes of the 2007–2009 financial crisis, they use July 2007 as the starting point.

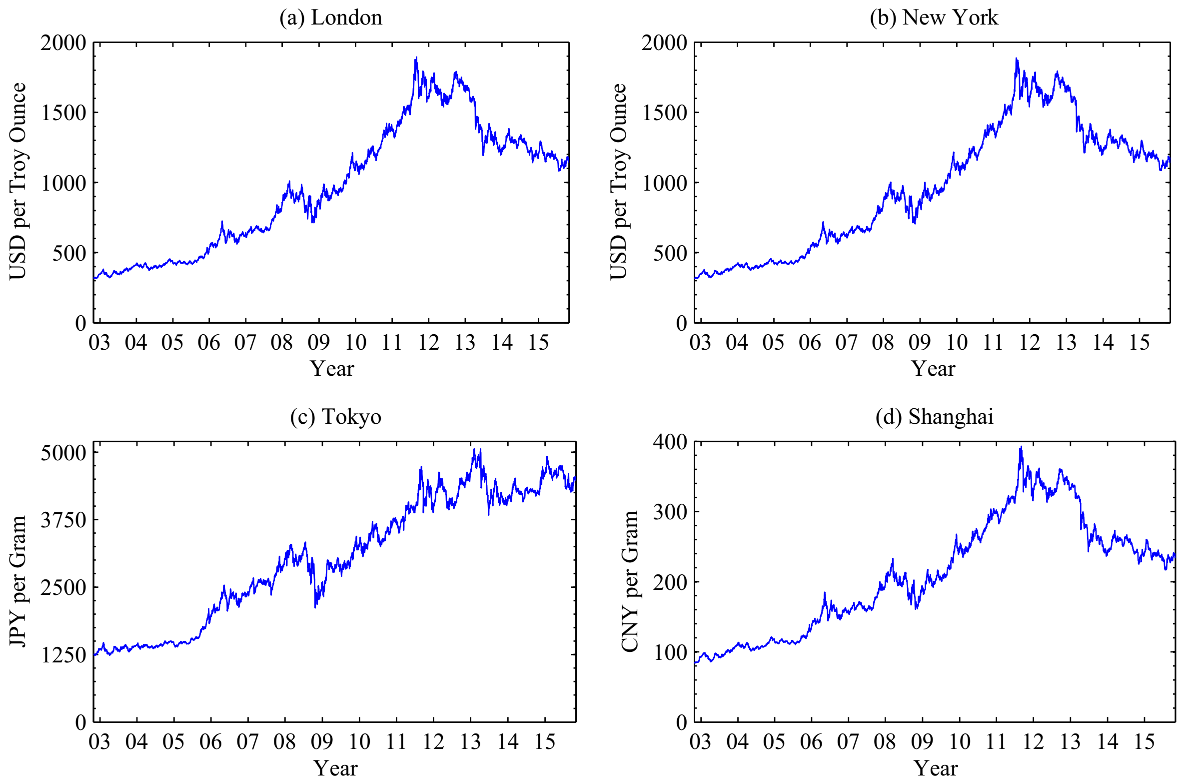

Fig. 1. Daily closing (fixing) prices of the London, New York, Tokyo, and Shanghai gold markets during the entire period from 30 October 2002 to 30 October 2015.

are located in different time zones, their trading hours are different. Table 1 shows the trading hours of the four gold markets in Greenwich Mean Time, local time, and New York time. The second column (Greenwich Mean Time) of Table 1 shows that the Tokyo and Shanghai gold markets trade within or around the same time interval, so the non-synchronous trading effect between them can be ignored. The last column (New York Time) shows that these two Asian gold markets close before the London and New York gold markets open for the trading day, and thus the non-synchronous trading effect should be taken into consideration when studying the one-way Granger causality in risk from Asian gold markets to the London and New York gold markets. Because the afternoon fixing time of the London gold market is three and one-half hours ahead of the closing price time of the New York gold market, we also take the non-synchronous trading effect into account when examining the extreme risk spillover effects (i.e., the one-way Granger causality in risk) from the London gold market to the New York gold market. That is to say, to fix the non-synchronous trading effect we use the statistic $Q_3(M)$ to investigate extreme risk spillover effects from Tokyo or Shanghai to London or New York and from London to New York.

Table 2 provides the gold return statistics for London (LD), New York (NY), Tokyo (TK), and Shanghai (SH) for different periods. The mean return values in subperiod II are smaller than in subperiod I but the standard deviations are larger, indicating that in the post-crisis era gold markets have low return and high volatility. All skewness values are smaller than 0 and kurtosis values greater than 3, suggesting that each return series at different periods follows a leptokurtic distribution with a fat left tail and not a Gaussian distribution. The Jarque-Bera statistic for each returns also rejects the non-hypothesis of Gaussian distribution. In addition, the skewness, kurtosis, and Jarque-Bera values for each return series in the post-crisis era are larger than those in the pre-crisis era, suggesting that large extreme returns have a

(*footnote continued*)
Duchin et al. (2010) and Mishkin (2011) point out that the beginning of the recent global financial crisis is usually set at the collapse of the US subprime mortgage market in July and August of 2007. Garcia-Appendini and Montoriol-Garriga (2013) examine the effect of the global financial crisis on between-firm liquidity provision using July 2007 as the starting point. Carvalho et al. (2015) investigate the spread of bank stress to nonfinancial companies during the financial crisis and use the beginning of July 2007 as the starting point. Choudhry et al. (2015) study correlations between gold and stock markets before and after the global financial crisis and divide the sample period into two subperiods, the pre-financial crisis (January 2000 to June 2007) and the financial crisis period (from July 2007 to March 2014). Hwang and Kim (2015) examine extreme risk spillover effects in financial markets before and after the recent financial crisis and use July 2007 as the beginning of the crisis. (ii) As a robustness test, following Lin et al. (2008) and Zhu et al. (2014), we use the Chow test (Chow, 1960) to check the null hypothesis that there is no structural change in the downside (upside) VaRs of each gold market beginning in 2 July 2007. The *F*-statistic of the Chow test for the downside (upside) VaRs of each gold market rejects the null hypothesis at the 1% significance level. The results of the Chow test are available from the authors upon request.

Table 2
Descriptive statistics of returns of the London (LD), New York (NY), Tokyo (TK), and Shanghai (SH) gold markets for different periods.

Market	Mean	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	LB-Q(20)	ARCH(20)	ADF
Panel A: The entire period: 31 October 2002 to 30 October 2015								
LD	0.0381	1.1823	−0.4159	8.0095	3623.093***	32.9583*	270.3034***	−57.5846***
NY	0.0380	1.1978	−0.4162	8.0774	3719.403***	29.7971*	194.2411***	−58.2964***
TK	0.0377	1.2551	−0.7454	12.5013	1299.591***	64.7255***	306.8119***	−61.8246***
SH	0.0305	1.0862	−0.4507	10.7002	8444.887***	31.2643*	276.0174***	−60.9256***
Panel B: Subperiod I (Pre-crisis era): 31 October 2002 to 29 June 2007								
LD	0.0724	1.0537	−0.4276	6.2339	623.3556***	31.0362*	128.0980***	−35.4317***
NY	0.0730	1.0834	−0.7215	6.1953	684.7775***	29.6461*	121.1929***	−38.0347***
TK	0.0658	1.0485	−0.7862	6.2720	734.1593***	43.5419***	125.1761***	−38.3353***
SH	0.0631	0.9525	−0.5672	9.5340	2450.023***	41.4063***	243.5145***	−38.6321***
Panel C: Subperiod II (Post-crisis era): 2 July 2007 to 30 October 2015								
LD	0.0155	1.2594	−0.3896	8.2922	2614.909***	23.6027	167.4113***	−45.2231***
NY	0.0151	1.2671	−0.2719	8.5266	1221.681***	28.4530*	124.9275***	−44.5556***
TK	0.0192	1.3739	−0.6955	13.0119	2426.250***	46.9003***	182.9114***	−48.2958***
SH	0.0090	1.1655	−0.3831	10.5796	8663.482***	30.3696*	146.3646***	−47.1822***

Notes: The Jarque–Bera statistic tests for the null hypothesis of Gaussian distribution. LB-Q(20) is the Ljung–Box *Q*-test statistic of the sample returns for up to the 20th order serial autocorrelation. ARCH(20) is the Engle's ARCH Lagrange Multiplier (LM) test for residual heteroscedasticity with 20 lags. The ADF statistic denotes the Augmented Dickey–Fuller test for a unit root. The null hypothesis of ADF test is a unit root in the sample returns.

* Rejection of the null hypothesis at 10% significance level.

*** Rejection of the null hypothesis at 1% significance level.

higher probability of occurring in the four gold markets during a period of financial turmoil. The Ljung–Box *Q*-test statistic shows that, with the exception of London in the post-crisis era, serial autocorrelations exist in each set of returns. Thus using the ARMA model we eliminate the serial autocorrelation effect in the returns. The Engle ARCH Lagrange multiplier (LM) test shows that there is significant volatility clustering in each set of returns, and this supports our decision to filter gold returns with a GARCH-type model. The ADF unit root test shows that each set of returns is stationary, implying that it can be further utilized for modeling without spurious regression. The overall differences in the descriptive statistics during the three periods indicate that the behavior of world gold markets differs between before and after the financial crisis, and this may affect their extreme risk spillover effects.

4.2. Estimates for ARMA-(T)GARCH-GED models

First we use an ARMA(p,q) model to filter out the serial autocorrelation in gold returns at different periods, and use a trial and error method to fix the optimal orders of the lag parameters p and q of the ARMA(p,q) model.[9] The resulting ARMA(p,q) model allows (i) the value of Akaike information criterion (AIC) of the model to be at a minimum, (ii) the model coefficients to be significant, and (iii) any serial autocorrelation existing in the residual series of the model to be eliminated. We then employ a (T)GARCH(r,s)-GED model to capture the volatility clustering and fat tails of the gold returns. Similar to the ARMA (p,q) model, the orders of the lag parameters r and s in the (T)GARCH(r,s) are fixed by trial and error and follow certain constraints, i.e., (i) the AIC value of the model should be minimal, (ii) the model coefficients should be significant and positive, and (iii) volatility clustering in the standardized residual series of the model should be eliminated. Finally we obtain the (T)GARCH(1,1)-GED model that confirms the statement made by Brooks (2008) that a GARCH-type model with a lag of one can capture the volatility clustering of financial asset returns. Tables 3 and 4 present the maximum likelihood estimation results from ARMA-(T)GARCH-GED models for the four gold markets over the entire period (Table 3) and during the subperiods (Table 4).

Table 3 shows how ARMA(3,3)-, ARMA(2,2)-, ARMA(5,5)-, and ARMA (0,2)-TGARCH(1,1)-GED models are used to filter out the serial autocorrelation, volatility clustering, "leverage effect," and fat tails of the returns from the London, New York, Tokyo, and Shanghai gold markets during the entire period. The coefficient β_1 of the lagged conditional variance h_{t-1} for each market is significant and positive, and suggests that the current volatility of gold returns is strongly affected by its previous volatility. The coefficient γ_1 for each market is significant and negative, which means (i) that the "leverage effect" is significant in each market during the entire period, and (ii) that previous positive shocks (breaking good news) more strongly influence current volatility than previous negative shocks (breaking bad news), i.e., the volatility of gold returns is more strongly affected by positive shocks (breaking good news), and this is consistent with the finding presented by Baur (2012) that gold takes on a safe-haven role during volatile periods. The degree-of-freedom value ν of the GED for each market is

[9] Note that there is no need to use an ARMA(p,q) model to filter out the serial autocorrelation in gold returns if its Ljung–Box Q-test statistic accepts the null hypothesis of no serial autocorrelation in the sample time series. For example, there is no serial autocorrelation in returns of London in the post-crisis period (see Table 2), which is the reason that the ARMA(0,0) model is shown in subperiod II for London (see Table 4).

Table 3
Maximum likelihood estimation results of ARMA-TGARCH-GED models for the London (LD), New York (NY), Tokyo (TK), and Shanghai (SH) gold markets during the entire period from 31 October 2002 to 30 October 2015.

Parameter	LD	NY	TK	SH
Mean equation				
c	0.0379*	0.0379*	0.0368***	0.0305***
	(0.0203)	(0.0202)	(0.0136)	(0.0187)
ϕ_1	−0.4294***	0.5296***	0.3403***	–
	(0.0043)	(0.0251)	(0.0570)	
ϕ_2	−0.4761***	−0.9531***	0.2761***	–
	(0.02332)	(0.0242)	(0.0646)	
ϕ_3	−0.9612***	–	−0.3626***	–
	(0.0429)		(0.0600)	
φ_1	0.4260***	−0.5425***	−0.3938***	−0.0481***
	(0.0412)	(0.0283)	(0.0598)	(0.0172)
φ_2	0.4681***	0.9397***	−0.2323***	0.0492***
	(0.0226)	(0.0276)	(0.0723)	(0.0172)
φ_3	0.9645***	–	0.3595***	–
	(0.0410)		(0.0680)	
Variance equation				
α_0	0.0107***	0.0121***	0.0127***	0.0103***
	(0.0035)	(0.0039)	(0.0035)	(0.0031)
α_1	0.0455***	0.0497***	0.0814***	0.0737***
	(0.0096)	(0.0103)	(0.0129)	(0.0126)
β_1	0.9568***	0.9537***	0.9290***	0.9354***
	(0.0076)	(0.0070)	(0.0090)	(0.0085)
γ_1	−0.0188*	−0.0216**	−0.0322**	−0.0302**
	(0.0101)	(0.0107)	(0.0139)	(0.0133)
υ	1.1344***	1.1814***	1.1366***	1.0300***
	(0.0300)	(0.034)	(0.0317)	(0.0283)
Diagnostic				
Log(L)	−4905.861	−5002.698	−4871.120	−4441.410
AIC	2.9127	2.9702	2.8921	2.6373
LB-Q(20)	16.5145	9.2988	12.3677	16.9034
	[0.6842]	[0.9792]	[0.9028]	[0.6592]
ARCH(20)	10.5913	23.1055	20.8305	18.1378
	[0.9561]	[0.2836]	[0.4072]	[0.5783]

Notes: The estimated models for the London, New York, Tokyo, and Shanghai gold markets during the entire period are ARMA(3,3)-, ARMA(2,2)-, ARMA (5,5)-, and ARMA (0,2)-TGARCH(1,1)-GED respectively. The numbers in parentheses are Std. Errors of the estimates. The coefficients (Std. Errors) of ϕ_4, ϕ_5, φ_4, and φ_5 of ARMA(5,5) for the Tokyo gold market in the entire period are −0.24403 (0.0629), 0.8195 (0.0474), 0.1944 (0.0711), and −0.8200 (0.0525) at 1% significant level. υ is the degree-of-freedom (the tail-thickness parameter) of the GED. Log(L) is the logarithm maximum likelihood function value. The numbers in square brackets are p-values of the statistics. LB-Q(20) and ARCH(20) are the Ljung–Box Q-test and the Engle's ARCH test statistics of the standardized residuals with 20 lags respectively.

* Significance at 10% level.

** Significance at 5% level.

*** Significance at 1% level.

significant and smaller than 2, which again confirms that gold returns are leptokurtic and fat-tailed.

Fig. 2 shows the conditional variance trends in the London, New York, Tokyo, and Shanghai gold markets during the entire period. The overall trend for the four gold markets is similar, but their volatility levels differ. The volatility is more severe in the Asian markets, i.e., the volatility level (amplitude) of the Tokyo gold market is the largest, followed by the Shanghai gold market. The conditional variance curves in the London and New York gold markets seem to have the same volatility trend and nearly overlap. In addition, during the 2008 financial crisis and the European sovereign debt crisis the conditional variances in each gold market exhibited a huge peak, rapidly increasing and then falling dramatically. The largest value of conditional variance for the Tokyo gold market, for example, was 10 times the average volatility level. This finding implies that (i) during the recent financial crisis the risk was extreme in the four gold markets and (ii) the extreme risk spillover effects in the markets in subperiod I differed from those in subperiod II.

Table 4 presents the data showing that the estimation models for the London, New York, Tokyo, and Shanghai gold markets are ARMA(3,2)-, ARMA(2,2)-, ARMA(5,5)-, and ARMA(2,2)-TGARCH(1,1)-GED in subperiod I, and ARMA(0,0)-, ARMA (3,3)-, ARMA(3,1)- and ARMA(3,3)-GARCH(1,1)-GED in subperiod II. The outcomes for these two subperiods differ because the volatility "leverage effect" does not have a significant effect in subperiod II. In addition, the estimated GED degree-of-freedom for each set of returns in subperiod II is smaller than in subperiod I, suggesting that in the post-crisis era the gold return tails are thicker than in the pre-crisis era, which indicates that in the post-crisis era the risk in the gold return tails is more extreme.

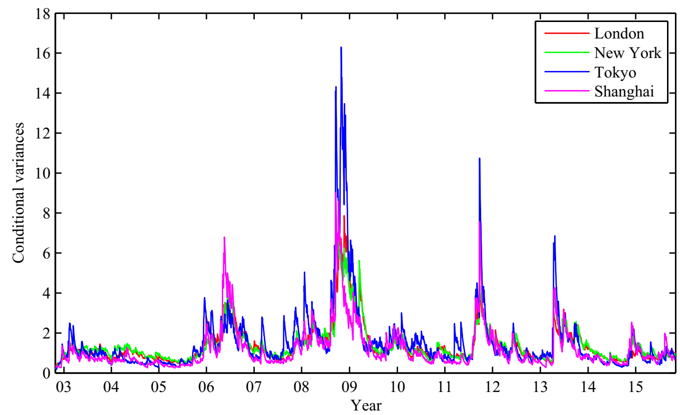

Fig. 2. Conditional variances of returns for the London, New York, Tokyo, and Shanghai gold markets during the entire period from 31 October 2002 to 30 October 2015.

Table 4
Maximum likelihood estimation results of ARMA-(T)GARCH-GED models for the London (LD), New York (NY), Tokyo (TK), and Shanghai (SH) gold markets in subperiods I (from 31 October 2002 to 29 June 2007) and II (from 2 July 2007 to 30 October 2015).

Parameter	Subperiod I				Subperiod II			
	LD	NY	TK	SH	LD	NY	TK	SH
Mean equation								
c	0.0596**	0.0594*	0.0587***	0.0525*	–	0.0261	0.0235**	0.0175
	(0.0634)	(0.0317)	(0.0206)	(0.0282)		(0.0270)	(0.0106)	(0.0242)
ϕ_1	0.7629***	0.9601***	0.1574***	0.8134***	–	−0.3319***	0.9184***	−0.9927***
	(0.0963)	(0.0968)	(0.0483)	(0.0674)		(0.1014)	(0.0221)	(0.3033)
ϕ_2	−0.8338***	−0.7785***	0.5920***	−0.7554***	–	−0.5348***	0.1053***	−1.2209***
	(0.0880)	(0.093)	(0.0396)	(0.0667)		(0.0541)	(0.0291)	(0.1290)
ϕ_3	0.0935***	–	−0.5629***	–	–	−0.8612***	−0.0449**	−0.5587*
	(0.0312)		(0.0263)			(0.1010)	(0.0216)	(0.2964)
φ_1	−0.7422***	−1.0086***	−0.1793***	−0.8815***	–	0.3483***	−0.9928***	0.9483***
	(0.0934)	(0.0838)	(0.0402)	(0.0539)		(0.0935)	(0.0054)	(0.3129)
φ_2	0.7772***	0.8408***	−0.5987***	0.8523***	–	0.5171***	–	1.2168***
	(0.0918)	(0.0804)	(0.0332)	(0.0532)		(0.0504)		(0.1313)
φ_3	–	–	0. 6250***	–	–	0.8824***	–	0.5264*
			(0.0184)			(0.0929)		(0.3110)
Variance equation								
α_0	0.0058	0.0058	0.0146***	0.0055**	0.0148**	0.0133**	0.0190***	0.01781***
	(0.0041)	(0.0041)	(0.0055)	(0.0030)	(0.0053)	(0.0056)	(0.0063)	(0.0058)
α_1	0.0506***	0.0627***	0.1147***	0.0919***	0.0368***	0.0425***	0.0598***	0.0592***
	(0.0152)	(0.0184)	(0.0269)	(0.0197)	(0.0084)	(0.0117)	(0.0105)	(0.0109)
β_1	0.9664***	0.9652***	0.9100***	0.9413***	0.9536***	0.9529***	0.9297***	0.9284***
	(0.0109)	(0.0104)	(0.0188)	(0.0127)	(0.0099)	(0.0089)	(0.0114)	(0.0123)
γ_1	−0.0411***	−0.0575***	−0.0683**	−0.0706***	–	–	–	–
	(0.0169)	(0.0178)	(0.0279)	(0.0213)				
ν	1.2194***	1.3469***	1.2861***	1.1447***	1.0939***	1.1411***	1.0833***	1.05395***
	(0.0680)	(0.0741)	(0.0685)	(0.0633)	(0.0341)	(0.0411)	(0.0368)	(0.0347)
Diagnostic								
Log(L)	−1656.874	−1714.330	−1570.016	−1420.315	−3245.674	−3284.165	−3300.123	−3025.776
AIC	2.7537	2.8489	2.6098	2.3617	3.0020	3.0385	3.0523	2.7989
LB-Q(20)	17.4429	13.8368	11.9212	24.5217	15.6136	10.2478	20.5349	12.4961
	[0.6241]	[0.8387]	[0.9188]	[0.2203]	[0.7403]	[0.9635]	[0.4249]	[0.8979]
ARCH(20)	26.9147	17.0093	12.1113	22.6940	5.6772	21.0809	17.3771	13.0350
	[0.1377]	[0.6524]	[0.9122]	[0.3041]	[0.9993]	[0.3924]	[0.6284]	[0.8759]

Notes: The estimated models for the London, New York, Tokyo, and Shanghai gold markets are ARMA(3,2)-, ARMA(2,2)-, ARMA(5,5)-, and ARMA(2,2)-TGARCH(1,1)-GED in subperiod I, and ARMA(0,0)-, ARMA(3,3)-, ARMA(3,1)-, and ARMA(3,3)-GARCH(1,1)-GED in subperiod II. The numbers in parentheses are Std. Errors of the estimates. The coefficients (Std. Errors) of ϕ_4, ϕ_5, φ_4, and φ_5 of ARMA(5,5) for the Tokyo gold market in subperiod I are −0.2348 (0.0387), 0.8731 (0.0461), 0.1974 (0.033), and −0.9235 (0.0381) at 1% significant level. ν is the degree-of-freedom (the tail-thickness parameter) of the GED. Log(L) is the logarithm maximum likelihood function value. The numbers in square brackets are p-values of the statistics. LB-Q(20) and ARCH(20) are the Ljung–Box Q-test and the Engle's ARCH test statistics of the standardized residuals with 20 lags respectively.

* Significance at 10% level.

** Significance at 5% level.

*** Significance at 1% level.

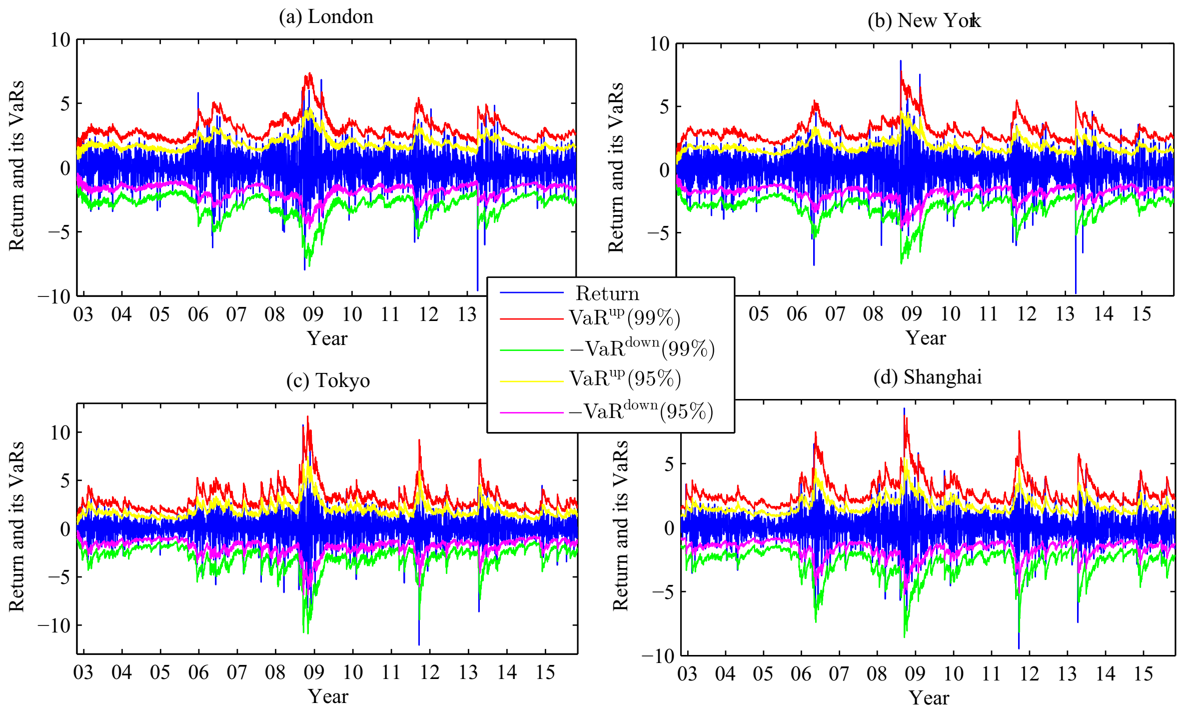

Fig. 3. Daily returns of the London, New York, Tokyo, and Shanghai gold markets during the entire period from 31 October 2002 to 30 October 2015 and their upside and downside VaRs at confidence levels of 99% and 95%.

4.3. Estimates for VaR models

We use Eqs. (8) and (9) and for different periods estimate the downside and upside VaRs in the London, New York, Tokyo, and Shanghai gold markets at confidence levels of 99% and 95%. For example, Fig. 3 shows the daily returns in the London, New York, Tokyo and Shanghai markets during the entire period and their upside and downside VaRs at confidence levels of 99% and 95%. To evaluate the accuracy and reliability of the VaR estimation, we use the backtesting techniques described in Appendix A. Table 5 shows the number of failure days (violations), failure rate results, LR_{uc}, LR_{ind}, and LR_{cc}, of the downside and upside VaRs for the four gold markets at the 99% and 95% confidence levels for different periods.

During the entire period, the downside VaRs for each gold market at the 99% and 95% confidence levels exhibit LR_{uc} and LR_{cc} values that are smaller than the corresponding critical values, and their LR_{ind} values—with the exception of Tokyo—are also smaller. This indicates that the downside VaR models for the four gold markets are acceptable. The LR_{ind} values for the upside VaRs in each gold market at the 99% and 95% confidence levels are less than the corresponding critical values, indicating that their VaR exceptions are independently distributed. In contrast, the LR_{uc} and LR_{cc} values of the upside VaRs for London and New York at the 99% confidence level and for Tokyo at the 95% confidence level are greater than the corresponding critical values, indicating that the upside VaR models at the corresponding levels for these markets are unacceptable. Thus we should be wary of the VaRs for these three markets at the corresponding confidence levels. As noted in Section 4.2, the GARCH-type market models before the financial crises differ from those after, possibly because GARCH-type models of the entire period do not take into full account the “styled facts” (e.g., the volatility clustering and fat tails) in gold returns.

Panel B of Table 5 shows that for the upside and downside VaRs, their LR_{uc}, LR_{ind}, and LR_{cc} values for each gold market at the 99% and 95% confidence levels in subperiod I are smaller than the corresponding critical values. This indicates that both the downside and upside VaR models are acceptable for the four gold markets in the pre-crisis period. From the LR_{uc}, LR_{ind} and LR_{cc} values in subperiod II shown in panel C of Table 5, we see that the downside and upside VaR models at the 99% and 95% confidence levels are significant for the four gold markets. Thus both the downside and upside VaRs at the 99% and 95% confidence levels for each gold market in the two subperiods can be used for further study.

4.4. Results for extreme risk spillover effects

Using the estimated downside and upside VaRs at the 99% and 95% confidence levels for the London, New York, Tokyo, and Shanghai gold markets, we compute the statistic values (including $Q_1(M)$, $Q_2(M)$, and $Q_3(M)$) and the corresponding p-values of two-way and one-way Granger causalities in risk. Following Hong (2001) and Hong et al. (2009), we investigate the extreme risk spillover effects with the largest effective lag truncation orders 5, 10, and 20, which are approximately equal to one trading

Table 5
Backtesting of VaR estimations for the London (LD), New York (NY), Tokyo (TK), and Shanghai (SH) gold markets at the 99% and 95% confidence levels for different periods.

Conf. level	Market	Downside VaR					Upside VaR				
		Violations	Failure rate	LR_{uc}	LR_{ind}	LR_{cc}	Violations	Failure rate	LR_{uc}	LR_{ind}	LR_{cc}
Panel A: The entire period: 31 October 2002 to 30 October 2015											
99%	LD	38	0.0113	0.5297*	0.8665*	1.3962*	15	0.0044	13.2324	0.1341*	13.3665
	NY	47	0.0139	4.7142*	1.8532*	6.5674*	17	0.0050	10.2278	0.1723*	10.4001
	TK	48	0.0142	5.4083*	1.3867*	6.7950*	25	0.0074	2.4968*	0.3736*	2.8704*
	SH	36	0.0107	0.1537*	3.5148*	3.6685*	22	0.0065	4.6839*	0.2890*	4.9729*
95%	LD	179	0.0531	0.6690*	0.6801*	1.3491*	161	0.0477	0.3611*	0.0137*	0.3748*
	NY	172	0.0510	0.0739*	0.0063*	0.0802*	162	0.0480	0.2713*	4.1502*	4.4215*
	TK	181	0.0537	0.9462*	7.9656	8.9118*	122	0.0362	14.9098	0.0433*	14.9531
	SH	157	0.0466	0.8518*	1.0232*	1.8750*	154	0.0457	1.3598*	1.2268*	2.5866*
Panel B: Subperiod I (Pre-crisis era): 31 October 2002 to 29 June 2007											
99%	LD	14	0.0116	0.2997*	2.0630*	2.3627*	6	0.0050	3.7731*	0.0600*	3.8331*
	NY	18	0.0149	2.5668*	0.5455*	3.1123*	6	0.0050	3.7731*	0.0600*	3.8331*
	TK	16	0.0133	1.1792*	1.6048*	2.7840*	8	0.0066	1.5665*	0.1068*	1.6733*
	SH	15	0.0124	0.6719*	0.3778*	1.0497*	7	0.0058	2.5256*	0.0817*	2.6073*
95%	LD	69	0.0572	1.2651*	3.7272*	4.9923*	52	0.0431	1.2588*	0.9323*	2.1911*
	NY	66	0.0547	0.5510*	0.9445*	1.4955*	45	0.0373	4.4632*	3.4892*	7.9524*
	TK	67	0.0555	0.7576*	2.6423*	3.3999*	54	0.0447	0.7170*	0.0837*	0.8007*
	SH	65	0.0539	0.3765*	3.1141*	3.4906*	58	0.0481	0.0935*	3.1603*	3.2538*
Panel C: Subperiod II (Post-crisis era): 2 July 2007 to 30 October 2015											
99%	LD	23	0.0106	0.0846*	0.4942*	0.5788*	10	0.0046	7.9040	0.0929*	7.9969*
	NY	26	0.0120	0.8336*	0.9892*	1.8228*	12	0.0055	5.1720*	0.1338*	5.3058*
	TK	29	0.0134	2.2850*	0.7878*	3.0728*	17	0.0079	1.0848*	0.2692*	1.3540*
	SH	22	0.0102	0.0060*	5.5099*	5.5159*	15	0.0069	2.3058*	0.2094*	2.5152*
95%	LD	113	0.0522	0.2211*	0.1606*	0.3817*	108	0.0499	0.0004*	0.0320*	0.0324*
	NY	113	0.0522	0.2211*	0.1606*	0.3817*	100	0.0462	0.6704*	0.7062*	1.3766*
	TK	115	0.0531	0.4412*	2.3603*	2.8015*	81	0.0374	7.8541	0.4273*	8.2814*
	SH	102	0.0471	0.3809*	0.1408*	0.5217*	91	0.0420	3.0353*	2.3470*	5.3823*

Notes: A violation, i.e., a failure day, is defined as occurring when the gold return R_t is greater (smaller) than the upside (negative downside) VaR on day t. The failure rate is defined as the ratio between the number of violations (failure days) and the total number of observations. LR_{uc}, LR_{ind}, and LR_{cc} represent statistics of likelihood ratio (LR) tests of unconditional coverage (uc), independent (ind) coverage, and conditional coverage (cc), respectively. $LR_{uc} \sim \chi^2(1)$, $LR_{ind} \sim \chi^2(1)$, and $LR_{cc} \sim \chi^2(2)$. The chi-square critical values of the LR_{uc} (or LR_{ind}) and LR_{cc} statistics at 1% significance level are 6.64 and 9.21, respectively.
* VaR model passes the corresponding test at 1% significance level. For a detailed introduction of these backtesting techniques, see Appendix A.

week, two trading weeks, and one trading month, respectively.[10] Tables 6–11 present the testing results for extreme risk spillover effects between six pairs gold markers at different periods. To see and examine all the testing results together, in Table 12 we summarize extreme risk spillover effects across the London, New York, Tokyo, and Shanghai gold markets at different periods.[11] Because some VaR models are not accurate when applied to the entire period, we will focus our analysis on the two subperiods.

Table 6 reports the testing results for the extreme risk spillover effects between the London and New York gold markets. In each period and at both downside and upside risks at the 99% and 95% confidence levels we find significant two-way and one-way Granger causalities in risk between London and New York, indicating the presence of extreme risk spillover effects between the two markets. In most cases the one-way Granger causality in risk values from London to New York are greater than those from New York to London, indicating a more extreme risk spillover effect from London to New York than from New York to London. The London Bullion Market Association (LBMA) has stated that "London is home to the international prices for gold…"[12] This adds support to the judgment that London leads New York in extreme risk spillovers.

Table 7 shows the testing results between London and Tokyo. For the downside and upside risks at the 95% confidence level in each period, and for both two-way and one-way Granger causalities in risk, we see extreme risk spillover effects between London and Tokyo. There are exceptions at the 99% confidence level between the two subperiods, i.e., in subperiod I we find (i) no downside (upside) risk spillover effects from London (Tokyo) to Tokyo (London), (ii) no two-way Granger causality in risk between London and Tokyo, and (iii) weak upside risk spillover effects from London to Tokyo because the statistic is significant only for $M = 10$ and 20 and its values are small. These findings suggest that there are no extreme risk spillover effects between London and Tokyo at the 1% risk level before the financial crisis. Although the one-way and two-

[10] Other lag orders like 15, 25, and 30 are also examined but not presented in the paper due to space limitations, and their results are consistent with the outcomes using the lag orders of 5, 10, and 20. For a detailed introduction of the lag order, refer to Hong (2001) and Hong et al. (2009).

[11] Note that Table 12 only shows whether there are extreme risk spillover effects between six pairs of gold markets. For the level of statistics of Granger causality in risk, refer to Tables 6–11.

[12] See http://www.lbma.org.uk/pricing-and-statistics

Table 6
Testing for extreme risk spillover effects between the London (LD) and New York (NY) gold markets for different periods.

Conf. level	Spillover direction	Downside risk spillover			Upside risk spillover		
		$M=5$	$M=10$	$M=20$	$M=5$	$M=10$	$M=20$
Panel A: The entire period: 31 October 2002 to 30 October 2015							
99%	LD ⇔ NY	37.080 (0.000)[a]	28.084 (0.000)[a]	19.903 (0.000)[a]	15.670 (0.000)[a]	11.488 (0.000)[a]	7.711 (0.000)[a]
	LD ⇒ NY	168.328 (0.000)[a]	126.923 (0.000)[a]	92.952 (0.000)[a]	53.475 (0.000)[a]	39.428 (0.000)[a]	27.714 (0.000)[a]
	LD ⇐ NY	61.029 (0.000)[a]	42.912 (0.000)[a]	29.207 (0.000)[a]	27.280 (0.000)[a]	19.607 (0.000)[a]	13.971 (0.000)[a]
95%	LD ⇔ NY	17.818 (0.000)[a]	13.935 (0.000)[a]	9.756 (0.000)[a]	30.641 (0.000)[a]	23.835 (0.000)[a]	18.965 (0.000)[a]
	LD ⇒ NY	319.628 (0.000)[a]	241.035 (0.000)[a]	176.442 (0.000)[a]	178.988 (0.000)[a]	134.288 (0.000)[a]	98.807 (0.000)[a]
	LD ⇐ NY	27.576 (0.000)[a]	19.541 (0.000)[a]	13.094 (0.000)[a]	51.441 (0.000)[a]	37.776 (0.000)[a]	28.265 (0.000)[a]
Panel B: Subperiod I (Pre-crisis era): 31 October 2002 to 29 June 2007							
99%	LD ⇔ NY	12.314 (0.000)[a]	9.031 (0.000)[a]	6.340 (0.000)[a]	9.327 (0.000)[a]	5.938 (0.000)[a]	2.678 (0.004)[a]
	LD ⇒ NY	33.250 (0.000)[a]	25.192 (0.000)[a]	19.076 (0.000)[a]	13.921 (0.000)[a]	9.704 (0.000)[a]	6.011 (0.000)[a]
	LD ⇐ NY	20.849 (0.000)[a]	14.118 (0.000)[a]	8.881 (0.000)[a]	16.992 (0.000)[a]	11.149 (0.000)[a]	6.540 (0.000)[a]
95%	LD ⇔ NY	11.068 (0.000)[a]	8.170 (0.000)[a]	5.155 (0.000)[a]	8.592 (0.000)[a]	6.098 (0.000)[a]	4.099 (0.000)[a]
	LD ⇒ NY	78.060 (0.000)[a]	58.396 (0.000)[a]	42.153 (0.000)[a]	20.940 (0.000)[a]	14.992 (0.000)[a]	10.411 (0.000)[a]
	LD ⇐ NY	13.662 (0.000)[a]	9.898 (0.000)[a]	6.470 (0.000)[a]	15.192 (0.000)[a]	10.982 (0.000)[a]	7.781 (0.000)[a]
Panel C: Subperiod II (Post-crisis era): 2 July 2007 to 30 October 2015							
99%	LD ⇔ NY	57.307 (0.000)[a]	42.853 (0.000)[a]	30.170 (0.000)[a]	4.102 (0.000)[a]	3.105 (0.000)[a]	2.566 (0.005)[a]
	LD ⇒ NY	105.178 (0.000)[a]	63.232 (0.000)[a]	57.042 (0.000)[a]	74.512 (0.000)[a]	55.205 (0.000)[a]	39.244 (0.000)[a]
	LD ⇐ NY	94.915 (0.000)[a]	78.807 (0.000)[a]	45.982 (0.000)[a]	8.469 (0.000)[a]	6.902 (0.000)[a]	6.507 (0.000)[a]
95%	LD ⇔ NY	6.509 (0.000)[a]	4.429 (0.000)[a]	2.539 (0.006)[a]	12.13 (0.000)[a]	9.672 (0.000)[a]	8.188 (0.000)[a]
	LD ⇒ NY	264.050 (0.000)[a]	198.333 (0.000)[a]	144.717 (0.000)[a]	144.543 (0.000)[a]	108.075 (0.000)[a]	78.228 (0.000)[a]
	LD ⇐ NY	12.565 (0.000)[a]	8.378 (0.000)[a]	5.321 (0.000)[a]	20.698 (0.000)[a]	16.137 (0.000)[a]	13.945 (0.000)[a]

Notes: "⇔" denotes two-way Granger causality in risk between the two gold markets; "⇒" ("⇐") denotes one-way Granger causality in risk from the former (the latter) to the latter (the former); and the numbers in parentheses are the corresponding p-values.

[a] Significant extreme risk spillover effect exists in the corresponding test at 1% level.

Table 7
Testing for extreme risk spillover effects between the London (LD) and Tokyo (TK) gold markets for different periods.

Conf. level	Spillover direction	Downside risk spillover			Upside risk spillover		
		$M=5$	$M=10$	$M=20$	$M=5$	$M=10$	$M=20$
Panel A: The entire period: 31 October 2002 to 30 October 2015							
99%	LD ⇔ TK	3.768 (0.000)[a]	2.892 (0.002)[a]	1.514 (0.065)[b]	**0.878 (0.190)**	**0.478 (0.316)**	**0.180 (0.429)**
	LD ⇒ TK	6.998 (0.000)[a]	4.879 (0.000)[a]	2.804 (0.003)[a]	3.017 (0.001)[a]	1.364 (0.086)[b]	**0.394 (0.347)**
	LD ⇐ TK	49.765 (0.000)[a]	37.514 (0.000)[a]	27.210 (0.000)[a]	14.321 (0.000)[a]	11.129 (0.000)[a]	8.428 (0.000)[a]
95%	LD ⇔ TK	61.457 (0.000)[a]	47.530 (0.000)[a]	34.660 (0.000)[a]	34.466 (0.000)[a]	25.843 (0.000)[a]	19.079 (0.000)[a]
	LD ⇒ TK	99.414 (0.000)[a]	70.876 (0.000)[a]	49.404 (0.000)[a]	57.802 (0.000)[a]	40.956 (0.000)[a]	28.988 (0.000)[a]
	LD ⇐ TK	86.598 (0.000)[a]	66.156 (0.000)[a]	48.776 (0.000)[a]	35.321 (0.000)[a]	26.202 (0.000)[a]	19.247 (0.000)[a]
Panel B: Subperiod I (Pre-crisis era): 31 October 2002 to 29 June 2007							
99%	LD ⇔ TK	**1.275 (0.101)**	**0.871 (0.192)**	**−0.373 (0.645)**	**−1.327 (0.908)**	**1.146 (0.126)**	**0.487 (0.313)**
	LD ⇒ TK	**−0.991 (0.839)**	**−1.404(0.920)**	**−2.000 (0.977)**	**−0.380 (0.648)**	3.899 (0.000)[a]	3.403 (0.000)[a]
	LD ⇐ TK	21.662 (0.000)[a]	16.867 (0.000)[a]	11.965 (0.000)[a]	**−1.366 (0.914)**	**−1.858 (0.968)**	**−2.493 (0.994)**
95%	LD ⇔ TK	24.559 (0.000)[a]	19.263 (0.000)[a]	13.721 (0.000)[a]	6.745 (0.000)[a]	4.625 (0.000)[a]	4.090 (0.000)[a]
	LD ⇒ TK	40.452 (0.000)[a]	29.274 (0.000)[a]	20.098 (0.000)[a]	11.175 (0.000)[a]	7.695 (0.000)[a]	5.947 (0.000)[a]
	LD ⇐ TK	27.760 (0.000)[a]	21.262 (0.000)[a]	15.559 (0.000)[a]	25.186 (0.000)[a]	18.513 (0.000)[a]	14.024 (0.000)[a]
Panel C: Subperiod II (Post-crisis era): 2 July 2007 to 30 October 2015							
99%	LD ⇔ TK	7.702 (0.000)[a]	5.655 (0.000)[a]	3.716 (0.000)[a]	**−1.677 (0.953)**	**−1.033 (0.849)**	**−1.030 (0.849)**
	LD ⇒ TK	13.503 (0.000)[a]	9.398 (0.000)[a]	6.066 (0.000)[a]	**−1.170 (0.879)**	**−1.661 (0.952)**	**−2.331 (0.990)**
	LD ⇐ TK	34.270 (0.000)[a]	25.666 (0.000)[a]	18.713 (0.000)[a]	21.623 (0.000)[a]	17.230 (0.000)[a]	13.324 (0.000)[a]
95%	LD ⇔ TK	30.693 (0.000)[a]	23.309 (0.000)[a]	16.771 (0.000)[a]	22.918 (0.000)[a]	17.843 (0.000)[a]	12.708 (0.000)[a]
	LD ⇒ TK	51.098 (0.000)[a]	35.971 (0.000)[a]	24.942 (0.000)[a]	38.427 (0.000)[a]	27.294 (0.000)[a]	18.593 (0.000)[a]
	LD ⇐ TK	50.460 (0.000)[a]	38.245 (0.000)[a]	27.974 (0.000)[a]	24.380 (0.000)[a]	18.910 (0.000)[a]	13.959 (0.000)[a]

Notes: "⇔" denotes two-way Granger causality in risk between the two gold markets; "⇒" ("⇐") denotes one-way Granger causality in risk from the former (the latter) to the latter (the former); and the numbers in parentheses are the corresponding p-values. Statistics without significant extreme risk spillover effects are highlighted in bold.

[a] Significant extreme risk spillover effect exists in the corresponding test at 1% level.

[b] Significant extreme risk spillover effect exists in the corresponding test at 10% level.

Table 8
Testing for extreme risk spillover effects between the London (LD) and Shanghai (SH) gold markets for different periods.

Conf. level	Spillover direction	Downside risk spillover			Upside risk spillover		
		$M=5$	$M=10$	$M=20$	$M=5$	$M=10$	$M=20$
Panel A: The entire period: 31 October 2002 to 30 October 2015							
99%	LD ⇔SH	54.259 (0.000)[a]	40.653 (0.000)[a]	29.169 (0.000)[a]	4.028 (0.000)[a]	4.861 (0.000)[a]	3.915 (0.000)[a]
	LD ⇒ SH	82.601 (0.000)[a]	58.393 (0.000)[a]	40.114 (0.000)[a]	8.126 (0.000)[a]	7.499 (0.000)[a]	4.867 (0.000)[a]
	LD⇐SH	91.022 (0.000)[a]	68.213 (0.000)[a]	50.131 (0.000)[a]	2.791 (0.003)[a]	2.923 (0.002)[a]	3.016 (0.001)[a]
95%	LD ⇔ SH	74.607 (0.000)[a]	57.706 (0.000)[a]	41.701 (0.000)[a]	36.394 (0.000)[a]	27.825 (0.000)[a]	19.777 (0.000)[a]
	LD ⇒ SH	119.652 (0.000)[a]	86.218 (0.000)[a]	60.202 (0.000)[a]	61.000 (0.000)[a]	44.024 (0.000)[a]	30.626 (0.000)[a]
	LD⇐SH	133.538 (0.000)[a]	100.928 (0.000)[a]	73.676 (0.000)[a]	61.908 (0.000)[a]	46.192 (0.000)[a]	33.335 (0.000)[a]
Panel B: Subperiod I (Pre-crisis era): 31 October 2002 to 29 June 2007							
99%	LD ⇔ SH	14.955 (0.000)[a]	10.417 (0.000)[a]	6.541 (0.000)[a]	8.326 (0.000)[a]	9.556 (0.000)[a]	7.638 (0.000)[a]
	LD ⇒ SH	25.860 (0.000)[a]	17.649 (0.000)[a]	11.825 (0.000)[a]	15.284 (0.000)[a]	16.229 (0.000)[a]	12.925 (0.000)[a]
	LD ⇐ SH	8.205 (0.000)[a]	5.522 (0.000)[a]	3.098 (0.000)[a]	11.664 (0.000)[a]	8.348 (0.000)[a]	5.579 (0.000)[a]
95%	LD ⇔ SH	31.822 (0.000)[a]	24.834 (0.000)[a]	18.251 (0.000)[a]	7.257 (0.000)[a]	5.806 (0.000)[a]	4.245 (0.000)[a]
	LD ⇒ SH	51.649 (0.000)[a]	37.670 (0.000)[a]	26.328 (0.000)[a]	13.367 (0.000)[a]	10.409 (0.000)[a]	7.543 (0.000)[a]
	LD ⇐ SH	37.932 (0.000)[a]	28.623 (0.000)[a]	21.275 (0.000)[a]	7.683 (0.000)[a]	5.331 (0.000)[a]	3.675 (0.000)[a]
Panel C: Subperiod II (Post-crisis era): 2 July 2007 to 30 October 2015							
99%	LD ⇔ SH	22.277 (0.000)[a]	16.407 (0.000)[a]	12.259 (0.000)[a]	2.513 (0.006)[a]	4.671 (0.000)[a]	4.046 (0.000)[a]
	LD ⇒ SH	34.818 (0.000)[a]	24.553 (0.000)[a]	16.351 (0.000)[a]	5.646 (0.000)[a]	6.861 (0.000)[a]	4.636 (0.000)[a]
	LD ⇐ SH	96.126 (0.000)[a]	71.8596 (0.000)[a]	53.833 (0.000)[a]	4.848 (0.000)[a]	4.764 (0.000)[a]	4.539 (0.000)[a]
95%	LD ⇔ SH	40.262 (0.000)[a]	30.083 (0.000)[a]	20.984 (0.000)[a]	28.700 (0.000)[a]	23.109 (0.000)[a]	16.920 (0.000)[a]
	LD ⇒ SH	66.751 (0.000)[a]	47.026 (0.000)[a]	32.280 (0.000)[a]	47.759 (0.000)[a]	35.053 (0.000)[a]	24.364 (0.000)[a]
	LD ⇐ SH	70.040 (0.000)[a]	52.380 (0.000)[a]	37.703 (0.000)[a]	52.775 (0.000)[a]	40.470 (0.000)[a]	29.957 (0.000)[a]

Notes: "⇔" denotes two-way Granger causality in risk between the two gold markets; "⇒" ("⇐") denotes one-way Granger causality in risk from the former (the latter) to the latter (the former); and the numbers in parentheses are the corresponding p-values.
[a] Significant extreme risk spillover effect exists in the corresponding test at 1% level.

Table 9
Testing for extreme risk spillover effects between the New York (NY) and Tokyo (TK) gold markets for different periods.

Conf. level	Spillover direction	Downside risk spillover			Upside risk spillover		
		$M=5$	$M=10$	$M=20$	$M=5$	$M=10$	$M=20$
Panel A: The entire period: 31 October 2002 to 30 October 2015							
99%	NY ⇔ TK	95.900 (0.000)[a]	74.502 (0.000)[a]	55.488 (0.000)[a]	173.434 (0.000)[a]	130.807 (0.000)[a]	94.089 (0.000)[a]
	NY ⇒ TK	155.533 (0.000)[a]	109.902 (0.000)[a]	75.890 (0.000)[a]	284.184 (0.000)[a]	200.876 (0.000)[a]	138.671 (0.000)[a]
	NY⇐TK	4.233 (0.000)[a]	6.063 (0.000)[a]	7.247 (0.000)[a]	1.738 (0.041)[b]	**1.1478 (0.126)**	**1.221 (0.111)**
95%	NY⇔TK	168.395 (0.000)[a]	127.804 (0.000)[a]	91.736 (0.000)[a]	109.576 (0.000)[a]	82.802 (0.000)[a]	59.827 (0.000)[a]
	NY ⇒ TK	275.438 (0.000)[a]	196.394 (0.000)[a]	136.449 (0.000)[a]	179.357 (0.000)[a]	127.584 (0.000)[a]	89.281 (0.000)[a]
	NY ⇐ TK	23.984 (0.000)[a]	17.485 (0.000)[a]	12.092 (0.000)[a]	−**0.564 (0.714)**	−**1.010 (0.844)**	−**1.258 (0.8960)**
Panel B: Subperiod I (Pre-crisis era): 31 October 2002 to 29 June 2007							
99%	NY⇔TK	35.923 (0.000)[a]	27.602 (0.000)[a]	19.170 (0.000)[a]	32.752 (0.000)[a]	27.827 (0.000)[a]	22.6801 (0.000)[a]
	NY ⇒ TK	54.914 (0.000)[a]	38.368 (0.000)[a]	25.776 (0.000)[a]	54.814 (0.000)[a]	42.165 (0.000)[a]	31.231 (0.000)[a]
	NY⇐TK	2.624 (0.004)[a]	3.234 (0.000)[a]	2.224 (0.013)[b]	−**1.106 (0.866)**	**0.266**(**0.395**)	1.951 (0.026)[b]
95%	NY⇔TK	49.305 (0.000)[a]	37.062 (0.000)[a]	25.943 (0.000)[a]	36.017 (0.000)[a]	27.077 (0.000)[a]	20.027 (0.000)[a]
	NY ⇒ TK	79.946 (0.000)[a]	56.804 (0.000)[a]	38.957 (0.000)[a]	59.646 (0.000)[a]	41.906 (0.000)[a]	29.656 (0.000)[a]
	NY⇐TK	12.357 (0.000)[a]	8.859 (0.000)[a]	5.858 (0.000)[a]	1.565 (0.059)[c]	**1.244** (**0.107**)	**1.070**(**0.142**)
Panel C: Subperiod II (Post-crisis era): 2 July 2007 to 30 October 2015							
99%	NY⇔TK	59.792 (0.000)[a]	45.946 (0.000)[a]	33.342 (0.000)[a]	91.244 (0.000)[a]	68.732 (0.000)[a]	48.923 (0.000)[a]
	NY ⇒ TK	98.157 (0.000)[a]	69.381 (0.000)[a]	47.535 (0.000)[a]	147.081 (0.000)[a]	103.549 (0.000)[a]	70.847 (0.000)[a]
	NY⇐TK	3.137 (0.000)[a]	3.424 (0.000)[a]	3.390 (0.000)[a]	5.343 (0.000)[a]	4.335 (0.000)[a]	3.286 (0.000)[a]
95%	NY⇔TK	95.622 (0.000)[a]	72.642 (0.000)[a]	51.986 (0.000)[a]	78.235 (0.000)[a]	59.914 (0.000)[a]	43.100 (0.000)[a]
	NY ⇒ TK	154.948 (0.000)[a]	110.067 (0.000)[a]	76.088 (0.000)[a]	123.051 (0.000)[a]	87.513 (0.000)[a]	60.701 (0.000)[a]
	NY⇐TK	10.692 (0.000)[a]	8.236 (0.000)[a]	5.793 (0.000)[a]	**1.222**(**0.111**)	1.713 (0.043)[b]	**1.187**(**0.118**)

Notes: "⇔" denotes two-way Granger causality in risk between the two gold markets; "⇒" ("⇐") denotes one-way Granger causality in risk from the former (the latter) to the latter (the former); and the numbers in parentheses are the corresponding p-values. Statistics without significant extreme risk spillover effects are highlighted in bold.
[a] Significant extreme risk spillover effect exists in the corresponding test at 1% level.
[b] Significant extreme risk spillover effect exists in the corresponding test at 5% level.
[c] Significant extreme risk spillover effect exists in the corresponding test at 10% level.

Table 10
Testing for extreme risk spillover effects between the New York (NY) and Shanghai (SH) gold markets for different periods.

Conf. level	Spillover direction	Downside risk spillover			Upside risk spillover		
		$M=5$	$M=10$	$M=20$	$M=5$	$M=10$	$M=20$
Panel A: The entire period: 31 October 2002 to 30 October 2015							
99%	NY ⇔ SH	330.866 (0.000)[a]	250.678 (0.000)[a]	180.780 (0.000)[a]	48.244 (0.000)[a]	37.391 (0.000)[a]	26.646 (0.000)[a]
	NY ⇒ SH	540.150 (0.000)[a]	383.760 (0.000)[a]	266.525 (0.000)[a]	79.043 (0.000)[a]	56.791 (0.000)[a]	38.846 (0.000)[a]
	NY ⇐ SH	11.109 (0.000)[a]	7.938 (0.000)[a]	5.722 (0.000)[a]	**-0.354 (0.638)**	**0.240** (**0.405**)	**0.289**(**0.386**)
95%	NY ⇔ SH	203.668 (0.000)[a]	153.607 (0.000)[a]	110.022 (0.000)[a]	161.344 (0.000)[a]	122.542 (0.000)[a]	88.277 (0.000)[a]
	NY ⇒ SH	333.067 (0.000)[a]	235.882 (0.000)[a]	163.475 (0.000)[a]	263.886 (0.000)[a]	187.587 (0.000)[a]	129.846 (0.000)[a]
	NY ⇐SH	28.134 (0.000)[a]	20.602 (0.000)[a]	14.293 (0.000)[a]	2.229 (0.013)[b]	1.761 (0.039)[b]	1.511 (0.065)[c]
Panel B: Subperiod I (Pre-crisis era): 31 October 2002 to 29 June 2007							
99%	NY ⇔SH	96.212 (0.000)[a]	72.833 (0.000)[a]	52.264 (0.000)[a]	10.398 (0.000)[a]	13.101 (0.000)[a]	10.738 (0.000)[a]
	NY ⇒ SH	158.190 (0.000)[a]	112.237 (0.000)[a]	77.558 (0.000)[a]	18.741 (0.000)[a]	22.175 (0.000)[a]	18.555 (0.000)[a]
	NY ⇐ SH	**-1.048 (0.853)**	**-0.898 (0.815)**	**-0.686 (0.754)**	**-1.373 (0.915)**	**-1.878 (0.970)**	**-2.572 (0.995)**
95%	NY ⇔ SH	60.206 (0.000)[a]	47.201 (0.000)[a]	35.261 (0.000)[a]	43.179 (0.000)[a]	32.667 (0.000)[a]	22.962 (0.000)[a]
	NY ⇒ SH	98.900 (0.000)[a]	72.304 (0.000)[a]	51.568 (0.000)[a]	72.118 (0.000)[a]	51.409 (0.000)[a]	35.434 (0.000)[a]
	NY ⇐ SH	7.485 (0.000)[a]	5.862 (0.000)[a]	4.688 (0.000)[a]	**-1.169 (0.879)**	**-1.206 (0.886)**	**-1.472 (0.930)**
Panel C: Subperiod II (Post-crisis era): 2 July 2007 to 30 October 2015							
99%	NY ⇔ SH	219.675 (0.000)[a]	166.048 (0.000)[a]	119.219 (0.000)[a]	64.972 (0.000)[a]	49.444 (0.000)[a]	34.559 (0.000)[a]
	NY ⇒ SH	359.315 (0.000)[a]	254.959 (0.000)[a]	176.572 (0.000)[a]	105.829 (0.000)[a]	74.274 (0.000)[a]	50.549 (0.000)[a]
	NY ⇐ SH	12.839 (0.000)[a]	9.129 (0.000)[a]	6.344 (0.000)[a]	−**0.003**(**0.501**)	**1.032**(**0.151**)	**0.223**(**0.412**)
95%	NY ⇔ SH	138.045 (0.000)[a]	103.911 (0.000)[a]	74.391 (0.000)[a]	87.012 (0.000)[a]	66.426 (0.000)[a]	47.490 (0.000)[a]
	NY ⇒ SH	226.331 (0.000)[a]	160.225 (0.000)[a]	111.170 (0.000)[a]	141.979 (0.000)[a]	101.606 (0.000)[a]	70.484 (0.000)[a]
	NY ⇐ SH	15.029 (0.000)[a]	10.709 (0.000)[a]	7.139 (0.000)[a]	3.677 (0.000)[a]	2.690 (0.004)[a]	1.455 (0.073)[c]

Notes: "⇔" denotes two-way Granger causality in risk between the two gold markets; "⇒" ("⇐") denotes one-way Granger causality in risk from the former (the latter) to the latter (the former); and the numbers in parentheses are the corresponding *p*-values. Statistics without significant extreme risk spillover effects are highlighted in bold.

[a] Significant extreme risk spillover effect exists in the corresponding test at 1% level.
[b] Significant extreme risk spillover effect exists in the corresponding test at 5% level.
[c] Significant extreme risk spillover effect exists in the corresponding test at 10% level.

Table 11
Testing for extreme risk spillover effects between the Tokyo (TK) and Shanghai (SH) gold markets for different periods.

Conf. level	Spillover direction	Downside risk spillover			Upside risk spillover		
		$M=5$	$M=10$	$M=20$	$M=5$	$M=10$	$M=20$
Panel A: The entire period: 31 October 2002 to 30 October 2015							
99%	TK ⇔ SH	4.401 (0.000)[a]	5.542 (0.000)[a]	4.800 (0.000)[a]	−**0.515 (0.697)**	−**0.036 (0.514)**	−**0.333 (0.630)**
	TK ⇒ SH	3.701 (0.000)[a]	6.493 (0.000)[a]	6.355 (0.000)[a]	**0.519 (0.302)**	**0.779 (0.218)**	**0.397 (0.346)**
	TK ⇐ SH	1.627 (0.052)[c]	**0.713 (0.238)**	−**0.153 (0.561)**	−**0.884 (0.812)**	−**0.494 (0.689)**	−**0.652 (0.743)**
95%	TK ⇔ SH	5.208 (0.000)[a]	3.970 (0.000)[a]	2.303 (0.011)[b]	1.596 (0.055)[c]	1.680 (0.047)[b]	**1.096 (0.137)**
	TK ⇒ SH	1.895 (0.029)[b]	**1.139 (0.127)**	**0.493 (0.311)**	3.039 (0.001)[a]	2.884 (0.002)[a]	2.006 (0.022)[b]
	TK ⇐ SH	4.115 (0.000)[a]	3.164 (0.000)[a]	1.644 (0.050)[b]	−**0.068 (0.527)**	−**0.052 (0.521)**	−**0.210 (0.583)**
Panel B: Subperiod I (Pre-crisis era): 31 October 2002 to 29 June 2007							
99%	TK ⇔ SH	5.923 (0.000)[a]	5.593 (0.000)[a]	3.247 (0.000)[a]	3.140 (0.000)[a]	5.330 (0.000)[a]	3.569 (0.000)[a]
	TK ⇒ SH	11.228 (0.000)[a]	10.416 (0.000)[a]	7.052 (0.000)[a]	6.193 (0.000)[a]	5.532 (0.000)[a]	2.978 (0.001)[a]
	TK ⇐ SH	−**1.017 (0.845)**	−**1.465 (0.929)**	−**2.006 (0.978)**	−**0.527 (0.701)**	2.995 (0.001)[a]	2.539 (0.006)[a]
95%	TK ⇔ SH	7.756 (0.000)[a]	6.235 (0.000)[a]	4.669 (0.000)[a]	5.556 (0.000)[a]	7.071 (0.000)[a]	5.611 (0.000)[a]
	TK ⇒ SH	3.966 (0.000)[a]	3.538 (0.000)[a]	2.541 (0.006)[a]	8.587 (0.000)[a]	10.482 (0.000)[a]	8.231 (0.000)[a]
	TK ⇐ SH	4.261 (0.000)[a]	2.856 (0.002)[a]	2.148 (0.016)[b]	**0.628 (0.265)**	**0.458 (0.324)**	**0.122 (0.452)**
Panel C: Subperiod II (Post-crisis era): 2 July 2007 to 30 October 2015							
99%	TK ⇔ SH	**0.502 (0.308)**	**1.029 (0.152)**	1.543 (0.061)[c]	2.989 (0.001)[a]	**1.263 (0.103)**	−**0.073 (0.529)**
	TK ⇒ SH	**0.200 (0.421)**	1.474 (0.070)[c]	1.696 (0.045)[b]	−**1.061 (0.856)**	−**1.477 (0.930)**	−**1.382 (0.917)**
	TK ⇐ SH	**0.222 (0.412)**	−**0.161 (0.564)**	**0.421 (0.337)**	2.673 (0.004)[a]	**1.089 (0.138)**	−**0.339 (0.910)**
95%	TK ⇔ SH	−**0.026 (0.510)**	−**0.279 (0.610)**	−**0.624 (0.734)**	**0.919 (0.179)**	**0.490 (0.312)**	**0.854 (0.197)**
	TK ⇒ SH	−**0.378 (0.647)**	−**0.577 (0.718)**	−**0.480 (0.684)**	−**0.108 (0.543)**	−**0.413 (0.660)**	−**0.128 (0.551)**
	TK ⇐ SH	**0.095 (0.462)**	−**0.034 (0.514)**	−**0.591 (0.723)**	**0.964 (0.167)**	**0.704 (0.241)**	**0.467 (0.320)**

Notes: "⇔" denotes two-way Granger causality in risk between the two gold markets; "⇒" ("⇐") denotes one-way Granger causality in risk from the former (the latter) to the latter (the former); and the numbers in parentheses are the corresponding *p*-values. Statistics without significant extreme risk spillover effects are highlighted in bold.

[a] Significant extreme risk spillover effect exists in the corresponding test at 1% level.
[b] Significant extreme risk spillover effect exists in the corresponding test at 5% level.
[c] Significant extreme risk spillover effect exists in the corresponding test at 10% level.

Table 12
Summary of extreme risk spillover effects across the London (LD), New York (NY), Tokyo (TK), and Shanghai (SH) gold markets for different periods.

Conf. level	Spillover direction	Downside risk spillover: Entire period	Downside risk spillover: Subperiod I	Downside risk spillover: Subperiod II	Upside risk spillover: Entire period	Upside risk spillover: Subperiod I	Upside risk spillover: Subperiod II
Panel A: London and New York							
99%	LD$\Leftrightarrow$NY	*	*	*	*	*	*
	LD$\Rightarrow$NY	*	*	*	*	*	*
	LD$\Leftarrow$NY	*	*	*	*	*	*
95%	LD$\Leftrightarrow$NY	*	*	*	*	*	*
	LD$\Rightarrow$NY	*	*	*	*	*	*
	LD$\Leftarrow$NY	*	*	*	*	*	*
Panel B: London and Tokyo							
99%	LD$\Leftrightarrow$TK	*	×	*	×	×	×
	LD$\Rightarrow$TK	*	×	*	+	+	×
	LD$\Leftarrow$TK	*	*	*	*	×	*
95%	LD$\Leftrightarrow$TK	*	*	*	*	*	*
	LD$\Rightarrow$TK	*	*	*	*	*	*
	LD$\Leftarrow$TK	*	*	*	*	*	*
Panel C: London and Shanghai							
99%	LD$\Leftrightarrow$SH	*	*	*	*	*	*
	LD$\Rightarrow$SH	*	*	*	*	*	*
	LD$\Leftarrow$SH	*	*	*	*	*	*
95%	LD$\Leftrightarrow$SH	*	*	*	*	*	*
	LD$\Rightarrow$SH	*	*	*	*	*	*
	LD$\Leftarrow$SH	*	*	*	*	*	*
Panel D: New York and Tokyo							
99%	NY$\Leftrightarrow$TK	*	*	*	*	*	*
	NY$\Rightarrow$TK	*	*	*	*	*	*
	NY$\Leftarrow$TK	*	*	*	+	+	*
95%	NY$\Leftrightarrow$TK	*	*	*	*	*	*
	NY$\Rightarrow$TK	*	*	*	*	*	*
	NY$\Leftarrow$TK	*	*	*	×	+	+
Panel E: New York and Shanghai							
99%	NY$\Leftrightarrow$SH	*	*	*	*	*	*
	NY$\Rightarrow$SH	*	*	*	*	*	*
	NY$\Leftarrow$SH	*	×	*	×	×	×
95%	NY$\Leftrightarrow$SH	*	*	*	*	*	*
	NY$\Rightarrow$SH	*	*	*	*	*	*
	NY$\Leftarrow$SH	*	*	*	*	×	*
Panel F: Tokyo and Shanghai							
99%	TK$\Leftrightarrow$SH	*	*	+	×	*	+
	TK$\Rightarrow$SH	*	*	+	×	*	×
	TK$\Leftarrow$SH	+	×	×	×	+	+
95%	TK$\Leftrightarrow$SH	*	*	×	+	*	×
	TK$\Rightarrow$SH	+	*	×	*	*	×
	TK$\Leftarrow$SH	*	*	×	×	×	×

Notes: "$\Leftrightarrow$" denotes two-way Granger causality in risk between the two gold markets; and "$\Rightarrow$" ("$\Leftarrow$") denotes one-way Granger causality in risk from the former (the latter) to the latter (the former). The entire period is from 31 October 2002 to 30 October 2015, the subperiod I is from 31 October 2002 to 29 June 2007, and the subperiod II is from 2 July 2007 to 30 October 2015. * denotes a significant extreme risk spillover effect. + denotes a weak extreme risk spillover effect because the statistic is not fully significant for all M. × with a gray background indicates that there is an insignificant extreme risk spillover effect in the corresponding test. This table only shows whether there are extreme risk spillover effects between six pairs of gold markets. For the level of statistics of Granger causality in risk refer to Tables 6–11.

way Granger causalities in upside risk from London to Tokyo and between London and Tokyo are not significant in subperiod II at the 99% confidence level, we do see remarkable extreme risk spillover effects between London and Tokyo. This indicates that the relationship between London and Tokyo becomes tighter in the post-crisis era.

Table 8 presents the testing results between London and Shanghai, which are similar to those between London and New York. All the statistic values for one-way and two-way Granger causalities in risk show significant extreme risk spillover effects between London and Shanghai. This may be because (i) London is the largest gold-trading center and is the gold-pricing center and (ii) China is the largest gold-producing country and among the top two gold-consuming countries in the world. In 2014 China's gold production reached approximately 460 metric tons and its consumption approximately 886 metric tons (GFMS, 2015). Thus significant real-time interactions between the London and Shanghai markets are to be expected. The statistic values during the pre-crisis period indicate that the level of spillover effects from London to Shanghai is stronger than that from Shanghai to London. However, most statistical values during the post-crisis period show that the level of spillover effects from Shanghai to London is stronger than that in the opposite direction, indicating that the market role of Shanghai may have shifted following the crisis.

Table 9 presents the testing results between New York and Tokyo. There are extreme downside risk spillover effects at the 99% and 95% confidence levels between New York and Tokyo, and there are extreme upside risk spillover effects from New York to Tokyo but not from Tokyo to New York. In particular, the upside risk spillover effects from Tokyo to New York in the two subperiods are weak and in the entire period at the 99% confidence level they disappear. We thus conclude (i) that the old saying "Bad news travels quickly, good news stagnates" is true—that negative return shocks (breaking bad news) are more easily transferred between New York and Tokyo than positive return shocks (breaking good news), (ii) that the risk spillover effects from New York to Tokyo are strong, and (iii) that the risk spillover effects from Tokyo to New York are weak.

Table 10 gives the testing results between New York and Shanghai, which are similar to those between New York and Tokyo. For both downside and upside risks at the 99% and 95% confidence levels in each period there is a two-way Granger causality in risk between New York and Shanghai and a one-way Granger causality in risk from New York to Shanghai. Extreme risk spillover effects from Shanghai to New York are weak or insignificant, especially in the pre-crisis era. Even in the post-crisis era, the one-way Granger causality in risk from Shanghai to New York is much less than from New York to Shanghai. In addition, at the same confidence level the statistic values indicate that the downside risk is greater than the upside risk. We thus find (i) that the downside risk between New York and Shanghai is transmitted more rapidly than the upside risk, (ii) that New York is a major source of extreme risk for Shanghai, and (iii) that risk spillover effects from Shanghai to New York are limited.

Table 11 gives the testing results between Tokyo and Shanghai. Although we might expect significant extreme risk spillover effects between these two Asian markets from a regional integration perspective, the statistic values show that the one-way Granger causality in risk from Tokyo to Shanghai, and from Shanghai to Tokyo, and the two-way Granger causality in risk between them are all weak or insignificant. In the post-crisis era extreme risk spillover effects between Tokyo and Shanghai effectively disappear. Although these two gold markets share the same trading time, instantaneous risk spillover effects between them are weak or insignificant. From an investment diversification perspective, the Tokyo and Shanghai gold markets have shown insignificant extreme risk spillover effects, which shows the diversification potential of these two gold assets and thus provides the opportunity of risk diversification benefits for market participants.

Tables 6–10 present a common picture. Both one-way and two-way Granger causalities in risk are stronger in the post-crisis period than in the pre-crisis period.[13] The level of extreme risk spillover effects during a financial crisis is thus stronger than when there is no crisis. On another note, extreme risk is more quickly transmitted between gold markets during periods of crisis because during crisis periods many market participants choose gold as a safe haven or risk-hedging tool or as a speculative vehicle[14] and this leads to shocks in gold pricing that strongly affect price volatility and increase risk.

5. Conclusions

It is important that gold market participants be able to monitor and control extreme risk and the spillover effects of risk. In this study, we have investigated the extreme risk spillover effects among the London, New York, Tokyo, and Shanghai gold markets. In particular, we have used the ARMA-(T)GARCH-GED model and the variance–covariance method to estimate the downside and upside VaRs and have used backtesting techniques to test their reliability. We have used the one-way and two-way Granger causalities in risk to study the downside and upside risk spillover effects between six pairs of gold markets. We have also analyzed how the extreme risk spillover effects in the four gold markets before the recent financial crisis differ from those after. Our findings fall into six categories.

(i) We find significant volatility clustering in the four gold markets. The conditional variances show that the volatility in the two Asian gold markets is higher than in the two Western markets, and that volatility is highest in the Tokyo market. During the 2008 financial crisis and the European sovereign debt crisis, the volatility in each gold market exhibited a peak wave, indicating the presence of extreme risk.

(ii) We find a significant "leverage effect" in the volatility in all four gold markets during the pre-crisis period and during the entire period studied. This effect becomes insignificant in all four markets during the post-crisis period. We find fat tails in the gold returns that are thicker in the post-crisis period than in the pre-crisis period, indicating that risk is more

[13] Note that the interaction between Tokyo and Shanghai shown in Table 11 is an exception to this.

[14] Because for each gold market during the post-crisis period there is a heightened volatility (as in Table 2 and Fig. 3) that may possibly stem from higher speculation.

extreme during a crisis period.

(iii) We find at the 99% and 95% confidence levels one-way and two-way Granger causalities in downside and upside risk between London and New York and between London and Shanghai, indicating the presence of significant extreme risk spillover effects in both cases.

(iv) We find that the level of extreme risk spillover effects is stronger from New York to Shanghai and Tokyo than from Shanghai and Tokyo to New York, indicating that New York is a major source of extreme risk to Shanghai and Tokyo. We also find that the level of spillover risk from London to New York is greater than from New York to London.

(v) We find that the one-way and two-way extreme risk spillover effects between Tokyo and Shanghai are weak or insignificant, and that the extreme risk spillover effects from Tokyo and Shanghai to New York are weak, as are the feedback effects from the two Asian markets to New York.

(vi) We find that the downside risk spillover effects are greater than the upside risk spillover effects, indicating that downside risk is transmitted more quickly than upside risk. Similarly, extreme risk is more quickly transmitted between gold markets in the post-crisis era than in the pre-crisis era.

The goal of this research is to provide the first empirical study of extreme risk spillover effects in global gold markets before and after the recent global financial crisis. Our findings have important applications in risk management for gold market participants and regulators. Most gold market participants choose gold as a safe haven or risk-hedging tool and thus pay particular attention to risks associated with gold pricing and their spillover effects among different gold markets. To avoid possible losses and predict the upcoming risk of holding gold they should be able to determine which gold markets are sources of risk spillover effects on their target gold market. Gold investors in China facing financial and macroeconomic uncertainty, for example, should pay close attention to the strong extreme risk spillovers from New York and London. Market regulators monitoring risk also pay special attention to extreme risk spillover effects. If they understand the mechanism of extreme risk spillovers across different gold markets they can use policy instruments and enhanced international and regional policy coordination to lower the probability that extreme risk will occur in the market under their care. Our study indicates that regulators of the two Asian gold markets should strengthen their cooperation and improve the exchange of information in order to enhance the movements between the two markets, forming a regional alliance and taking the lead in gold pricing.

Our work sheds light on how extreme risk spillover effects in the four gold markets before the recent global financial crisis differed from those after. We assume that extreme risk spillover effects change over time, and thus our future research on gold markets will focus on the dynamic and time-varying extreme risk spillover effects between them.

Acknowledgments

We are grateful to the Editor (Carl R. Chen) and three anonymous reviewers for their insightful suggestions that greatly improved the quality of this paper. This work was supported by the National Natural Science Foundation of China (grant numbers 71501066 and 71373072); the China Scholarship Council (grant number 201506135022); the Specialized Research Fund for the Doctoral Program of Higher Education (grant number 20130161110031); and the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (grant number 71521061).

Appendix A. VaR evaluation: backtesting

We use five backtesting techniques to evaluate the accuracy and reliability of the VaR estimate of extreme risk, (i) the number of violations (failure days), (ii) the failure rate, (iii) the likelihood ratio (LR) test of unconditional coverage, (iv) of independent coverage, and (v) of conditional coverage. Taking the downside VaR as an example, the failure rate is defined as the ratio between the number of failure days (violations) in which gold returns are smaller than negative downside VaR estimates and the total number of observations T. Mathematically, the failure rate f and the number of failure days N are $f=N/T$ and $N=\sum_{t=1}^{T} I_t$, respectively, where the indicator variable I_t is defined

$$I_t = \begin{cases} 1, & R_t < -V_t(\text{down}), \\ 0, & R_t \geq -V_t(\text{down}). \end{cases} \tag{A.1}$$

The null hypothesis of the LR test of unconditional coverage ($\mathrm{LR_{uc}}$) proposed by Kupiec (1995) is that the failure rate f for each trial equals to the specified probability α, where the statistics is defined as:

$$\mathrm{LR_{uc}} = -2\log\left(\frac{(1-\alpha)^{T-N}\alpha^N}{(1-f)^{T-N}f^N}\right) \sim \chi^2(1). \tag{A.2}$$

If $\mathrm{LR_{uc}}$ is greater than the critical value of $\chi^2(1)$ at a confidence level of $(1-\alpha)$, the null hypothesis is rejected with a probability α, indicating that the VaR estimation is unacceptable. Although the $\mathrm{LR_{uc}}$ test rejects both high and low failures, it neglects the time variation in the returns and does not exclude failures when they quickly cluster in a short period of time.

To improve the LR_{uc} test, Christoffersen (1998) proposes an LR test of conditional coverage, which is a combination of the LR tests of unconditional coverage and independence. Based on a binary first-order Markov chain model, the null hypothesis of the LR test of independence (LR_{ind}) is that the VaR exceptions are independently distributed, where the statistic is defined

$$LR_{ind} = -2\log\left(\frac{(1-\pi)^{n_{00}+n_{10}}\pi^{n_{01}+n_{11}}}{(1-\pi_0)^{n_{00}}\pi_0^{n_{01}}(1-\pi_1)^{n_{10}}\pi_1^{n_{11}}}\right) \sim \chi^2(1). \tag{A.3}$$

In Eq. (A.3), n_{ij} ($i,j=0,1$) is the number of failure days in which state j occurs and state i occurs on the previous day, i.e., $\pi_0 = n_{01}/(n_{00}+n_{01})$, $\pi_1 = n_{11}/(n_{10}+n_{11})$, and $\pi = (n_{01}+n_{11})/(n_{00}+n_{01}+n_{10}+n_{11})$. The LR_{ind} test determines whether the probability of a failure on any day is determined by the result on the previous day, i.e., under the null hypothesis of the LR_{ind} test the probability π_0 should be equal to π_1. See Christoffersen (1998) for a more detailed explanation of the variables.

The LR test of conditional coverage (LR_{cc}) is the joint test of unconditional coverage and independence, and the statistic is defined

$$LR_{cc} = LR_{uc} + LR_{ind} \sim \chi^2(2). \tag{A.4}$$

Unlike the LR_{uc} test, the LR_{cc} test can reject the VaR estimation that produces either too many or too few clustered exceptions.

Appendix B. Robustness test using the Shanghai gold futures data

In Section 4, we choose gold spot prices rather than gold futures prices as the empirical data for the Shanghai gold market because of the data constraint of gold futures prices. As noted in Section 4.1, the gold futures trading in China started on 9 January 2008 in the Shanghai Futures Exchange (SHFE) market, and thus there is no data prior to the global financial crisis for the Shanghai gold futures. A large body of literature shows that for most assets price discovery happens primarily in the futures market. Thus a major concern is whether the gold futures data would affect our central findings. To check the validity of our results, in this appendix we conduct the same testing of extreme risk spillover effects using the Shanghai gold futures data.

For robustness testing, we use daily closing (fixing) gold prices from the London OTC market, the New York COMEX market, the TOCOM market, and the SHFE market during the period from 9 January 2008 to 30 October 2015. Note that in the Shanghai gold market we use the daily closing gold futures prices of the continuous contract. As in Section 4, we first use the ARMA-(T)GARCH-GED model and the variance–covariance method to calculate the downside and upside VaRs. Then we employ backtesting techniques to evaluate the reliability of estimated VaRs. Finally, using Granger causality in risk we investigate extreme risk spillover effects across the London, New York, Tokyo, and Shanghai gold markets. To save space, Table B1 only shows the testing results of extreme risk spillover effects between the Shanghai gold market and other three markets.[15]

Panel A of Table B1 shows the testing results between the London and Shanghai gold markets. The statistical values of one-way and two-way Granger causalities in risk between London and Shanghai are significant at the 99% and 95% confidence levels, except for the one-way Granger causality in upside risk from Shanghai to London at 99% confidence level. The results confirm our main finding that there are significant extreme risk spillover effects between London and Shanghai.

Panel B of Table B1 reports the testing results between the New York and Shanghai gold markets. For both the upside and downside risks at the 99% and 95% confidence levels, the statistic values of one-way Granger causality in risk from New York to Shanghai and two-way Granger causality in risk between them are significant, but those of one-way Granger causality in risk from Shanghai to New York are insignificant. These results confirm our central findings between New York and Shanghai, i.e., extreme risk spillover effects from New York to Shanghai are significant and those from Shanghai to New York are insignificant.

Panel C of Table B1 presents the testing results between the Tokyo and Shanghai gold markets. Only the statistical values for the downside risk at the 99% confidence level show that there are extreme risk spillover effects between Tokyo and Shanghai. Most statistic values confirm our major finding that extreme risk spillover effects between Tokyo and Shanghai are weak or insignificant.

Overall, the robustness test using the Shanghai gold future data produces results consistent with our central findings, i.e., (i) there are strong extreme risk spillovers between London and Shanghai and from New York to Shanghai, (ii) extreme risk spillover effects from Shanghai to New York are limited (even insignificant), but risk spillover effects from Shanghai to London play a role, and (iii) extreme risk spillover effects between Tokyo and Shanghai are weak or negligible.

[15] Other detailed results can be obtained from the authors upon request.

Table B1
Testing for extreme risk spillover effects using the Shanghai gold futures data from 10 January 2008 to 30 October 2015.

Conf. level	Spillover direction	Downside risk spillover			Upside risk spillover		
		$M=5$	$M=10$	$M=20$	$M=5$	$M=10$	$M=20$
Panel A: London and Shanghai							
99%	LD ⇔ SH	75.271 (0.000)[a]	56.484 (0.000)[a]	40.288 (0.000)[a]	45.472 (0.000)[a]	34.323 (0.000)[a]	24.730 (0.000)[a]
	LD ⇒ SH	124.060 (0.000)[a]	87.895 (0.000)[a]	60.945 (0.000)[a]	75.702 (0.000)[a]	54.185 (0.000)[a]	36.838 (0.000)[a]
	LD ⇐ SH	6.032 (0.000)[a]	3.893 (0.000)[a]	2.293 (0.011)[b]	**−1.219 (0.889)**	**−1.544 (0.939)**	**−0.458 (0.677)**
95%	LD ⇔ SH	47.729 (0.000)[a]	36.701 (0.000)[a]	26.271 (0.000)[a]	21.396 (0.000)[a]	16.940 (0.000)[a]	12.790 (0.000)[a]
	LD ⇒ SH	78.427 (0.000)[a]	56.817 (0.000)[a]	39.991 (0.000)[a]	34.887 (0.000)[a]	26.140 (0.000)[a]	18.551 (0.000)[a]
	LD ⇐ SH	20.485 (0.000)[a]	15.086 (0.000)[a]	10.259 (0.000)[a]	9.827 (0.000)[a]	7.156 (0.000)[a]	5.641 (0.000)[a]
Panel B: New York and Shanghai							
99%	NY ⇔ SH	78.752 (0.000)[a]	63.256 (0.000)[a]	45.985 (0.000)[a]	17.455 (0.000)[a]	12.406 (0.000)[a]	7.660 (0.000)[a]
	NY ⇒ SH	129.633 (0.000)[a]	97.708 (0.000)[a]	68.440 (0.000)[a]	30.116 (0.000)[a]	20.576 (0.000)[a]	13.238 (0.000)[a]
	NY ⇐ SH	**−0.978 (0.836)**	**−0.984 (0.838)**	**−0.795 (0.787)**	**−1.278 (0.899)**	**−1.417 (0.922)**	**−1.838 (0.967)**
95%	NY ⇔ SH	110.604 (0.000)[a]	83.958 (0.000)[a]	60.612 (0.000)[a]	38.089 (0.000)[a]	29.160 (0.000)[a]	21.227 (0.000)[a]
	NY ⇒ SH	181.665 (0.000)[a]	129.002 (0.000)[a]	90.209 (0.000)[a]	63.529 (0.000)[a]	45.987 (0.000)[a]	32.437 (0.000)[a]
	NY ⇐ SH	**1.085 (0.139)**	**0.906 (0.182)**	**0.143 (0.443)**	**−0.417 (0.662)**	**−0.746 (0.773)**	**−0.769 (0.779)**
Panel C: Tokyo and Shanghai							
99%	TK ⇔ SH	30.689 (0.000)[a]	23.011 (0.000)[a]	16.871 (0.000)[a]	**−1.570 (0.942)**	**−2.103 (0.982)**	**−2.136 (0.984)**
	TK ⇒ SH	16.930 (0.000)[a]	11.913 (0.000)[a]	9.416 (0.000)[a]	**−0.984 (0.837)**	**−1.443 (0.925)**	**−1.109 (0.866)**
	TK ⇐ SH	16.456 (0.000)[a]	11.411 (0.000)[a]	7.166 (0.000)[a]	**−1.066 (0.857)**	**−1.407 (0.920)**	**−1.779 (0.962)**
95%	TK ⇔ SH	5.796 (0.000)[a]	3.810 (0.000)[a]	2.073 (0.000)[a]	**0.280 (0.390)**	**0.350 (0.363)**	1.480 (0.069)[c]
	TK ⇒ SH	10.276 (0.000)[a]	6.800 (0.000)[a]	4.286 (0.000)[a]	1.663 (0.048)[b]	**1.182 (0.119)**	**1.204 (0.114)**
	TK ⇐ SH	**−0.319 (0.625)**	**−0.628 (0.735)**	**−1.002 (0.842)**	**−0.637 (0.738)**	**−0.243 (0.596)**	**1.255 (0.105)**

Notes: "⇔" denotes two-way Granger causality in risk between the two gold markets; "⇒" ("⇐") denotes one-way Granger causality in risk from the former (the latter) to the latter (the former); and the numbers in parentheses are the corresponding *p*-values. Statistics without significant extreme risk spillover effects are highlighted in bold.

[a] Significant extreme risk spillover effect exists in the corresponding test at 1% level.
[b] Significant extreme risk spillover effect exists in the corresponding test at 5% level.
[c] Significant extreme risk spillover effect exists in the corresponding test at 10% level.

References

Acharya, V., Philippon, T., Richardson, M., & Roubini, N. (2009). The financial crisis of 2007–2009: Causes and remedies. *Financial Markets, Institutions & Instruments*, *18*, 89–137.

de Araújo, A. D. S., & Garcia, M. T. M. (2013). Risk contagion in the north-western and southern European stock markets. *Journal of Economics and Business*, *69*, 1–34.

Baklaci, H. F., Süer, O., & Yelkenci, T. (2016). Volatility linkages among gold futures in emerging markets. *Emerging Markets Finance and Trade*, *52*, 1–9.

Balboa, M., López-Espinosa, G., & Rubia, A. (2015). Granger causality and systemic risk. *Finance Research Letters*, *15*, 49–58.

Batten, J. A., & Lucey, B. M. (2010). Volatility in the gold futures market. *Applied Economics Letters*, *17*, 187–190.

Baur, D. G. (2012). Asymmetric volatility in the gold market. *The Journal of Alternative Investments*, *14*, 26–38.

Baur, D. G., & Lucey, B. M. (2010). Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *Financial Review*, *45*, 217–229.

Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking & Finance*, *34*, 1886–1898.

Brooks, C. (2008). *Introductory Econometrics for Finance* (2nd ed.). New York: Cambridge University Press.

Carvalho, D., Ferreira, M. A., & Matos, P. (2015). Lending relationships and the effect of bank distress: Evidence from the 2007–2009 financial crisis. *Journal of Financial and Quantitative Analysis*, *50*, 1165–1197.

Chang, C.-L., Della Chang, J.-C., & Huang, Y.-W. (2013). Dynamic price integration in the global gold market. *The North American Journal of Economics and Finance*, *26*, 227–235.

Cheung, Y.-W., & Ng, L. K. (1996). A causality-in-variance test and its application to financial market prices. *Journal of Econometrics*, *72*, 33–48.

Choudhry, T., Hassan, S. S., & Shabi, S. (2015). Relationship between gold and stock markets during the global financial crisis: Evidence from nonlinear causality tests. *International Review of Financial Analysis*, *41*, 247–256.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, *28*, 591–605.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, *39*, 841–862.

Dhillon, U. S., Lasser, D. J., & Watanabe, T. (1997). Volatility, information, and double versus walrasian auction pricing in US and Japanese futures markets. *Journal of Banking & Finance*, *21*, 1045–1061.

Du, L., & He, Y. (2015). Extreme risk spillovers between crude oil and stock markets. *Energy Economics*, *51*, 455–465.

Duchin, R., Ozbas, O., & Sensoy, B. A. (2010). Costly external finance, corporate investment, and the subprime mortgage credit crisis. *Journal of Financial Economics*, *97*, 418–435.

Engle, R. F., & Manganelli, S. (2004). CAViaR: conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, *22*, 367–381.

Fan, Y., Zhang, Y.-J., Tsai, H.-T., & Wei, Y.-M. (2008). Estimating value at risk of crude oil price and its spillover effect using the GED-GARCH approach. *Energy Economics*, *30*, 3156–3171.

Fuangkasem, R., Chunhachinda, P., & Nathaphan, S. (2014). Information transmission among world major gold futures markets: Evidence from high frequency synchronous trading data. *Journal of US-China Public Administration*, *11*, 255–269.

Garcia-Appendini, E., & Montoriol-Garriga, J. (2013). Firms as liquidity providers: Evidence from the 2007–2008 financial crisis. *Journal of Financial Economics*, *109*, 272–291.

GFMS (2015). *GFMS Gold Survey 2015*. London: Thomson Reuters.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*, 1779–1801.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*, 424–438.

Granger, C. W. J. (1980). Testing for causality: a personal viewpoint. *Journal of Economic Dynamics and Control*, *2*, 329–352.
Granger, C. W. J., Robins, P. P., & Engle, R. F. (1986). Wholesale and retail prices: Bivariate time-series modeling with forecastable error variances. In D. A. Belsley, & E. Kuh (Eds.), *Model reliability* (pp. 1–17). Cambridge: MIT Press.
Hammoudeh, S. M., Yuan, Y., McAleer, M., & Thompson, M. A. (2010). Precious metals-exchange rate volatility transmissions and hedging strategies. *International Review of Economics & Finance*, *19*, 633–647.
Haugh, L. D. (1976). Checking the independence of two covariance-stationary time series: A univariate residual cross-correlation approach. *Journal of the American Statistical Association*, *71*, 378–385.
Hauptfleisch, M., Putniņš, T. J., & Lucey, B. M. (2016). Who sets the price of gold? London or New York. *Journal of Futures Markets*, *36*, 564–586.
Hong, Y. (1996). Testing for independence between two covariance stationary time series. *Biometrika*, *83*, 615–625.
Hong, Y. (2001). A test for volatility spillover with application to exchange rates. *Journal of Econometrics*, *103*, 183–224.
Hong, Y., Chen, S., Liu, Y., & Wang, S. (2004). Extreme risk spillover between Chinese stock market and international stock markets. *China Economic Quarterly*, *3*, 703–726.
Hong, Y., Liu, Y., & Wang, S. (2009). Granger causality in risk and detection of extreme risk spillover between financial markets. *Journal of Econometrics*, *150*, 271–287.
Hwang, J., & Kim, J.-Y. (2015). Extreme risk spillover in financial markets: Evidence from the recent financial crisis. *Seoul Journal of Economics*, *28*, 171–198.
Joy, M. (2011). Gold and the US dollar: Hedge or haven? *Finance Research Letters*, *8*, 120–131.
Koch, P. D., & Yang, S. S. (1986). A method for testing the independence of two time series that accounts for a potential pattern in the cross-correlation function. *Journal of the American Statistical Association*, *81*, 533–544.
Kuester, K., Mittnik, S., & Paolella, M. S. (2006). Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, *4*, 53–89.
Kumar, B., & Pandey, A. (2011). International linkages of the Indian commodity futures markets. *Modern Economy*, *2*, 213–227.
Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, *3*, 73–84.
Laulajainen, R. (1990). Gold price round the clock: Technical and fundamental issues. *Resources Policy*, *16*, 143–152.
Lee, J., & Lee, H. (2009). Testing for risk spillover between stock market and foreign exchange market in Korea. *Journal of Economic Research*, *14*, 329–340.
Lin, H.-N., Chiang, S.-M., & Chen, K.-H. (2008). The dynamic relationships between gold futures markets: Evidence from COMEX and TOCOM. *Applied Financial Economics Letters*, *4*, 19–24.
Liu, X., Cheng, S., Wang, S., Hong, Y., & Li, Y. (2008). An empirical study on information spillover effects between the Chinese copper futures market and spot market. *Physica A: Statistical Mechanics and its Applications*, *387*, 899–914.
Lu, F.-b., Hong, Y.-m., Wang, S.-y., Lai, K.-k., & Liu, J. (2014). Time-varying Granger causality tests for applications in global crude oil markets. *Energy Economics*, *42*, 289–298.
Lucey, B. M., Larkin, C. J., & O'Connor, F. A. (2012). London or New York: Where does the gold price come from. *Alchemist*, *68*, 8–9.
Lucey, B. M., Larkin, C., & O'Connor, F. A. (2013). Where and when does the gold price originate London or New York. *Applied Economics Letters*, *20*, 813–817.
Lucey, B. M., Larkin, C., & O'Connor, F. (2014). Gold markets around the world—Who spills over what, to whom, when. *Applied Economics Letters*, *21*, 887–892.
Malliaris, A. G., & Urrutia, J. L. (1992). The international crash of October 1987: Causality tests. *The Journal of Financial and Quantitative Analysis*, *27*, 353–364.
McLeod, A. I., & Li, W. K. (1983). Diagnostic checking ARMA time series models using squared residual autocorrelations. *Journal of Time Series Analysis*, *4*, 269–273.
Mishkin, F. S. (2011). Over the cliff: from the subprime to the global financial crisis. *Journal of Economic Perspectives*, *25*, 49–70.
Morgan, J. P. (1996). *RiskMetrics™ Technical document* (4th ed.). Morgan: Guaranty Trust Company.
Murray, S. (2011). Loco London liquidity survey. *Alchemist*, *63*, 9–10.
Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, *59*, 347–370.
Ntim, C. G., English, J., Nwachukwu, J., & Wang, Y. (2015). On the efficiency of the global gold markets. *International Review of Financial Analysis*, *41*, 218–236.
O'Connor, F. A., Lucey, B. M., Batten, J. A., & Baur, D. G. (2015). The financial economics of gold–A survey. *International Review of Financial Analysis*, *41*, 186–205.
Pan, H.-F., & Zhang, J.-S. (2007). The extreme risk spillover effect between international and domestic oil markets. *Chinese Journal of Management Science*, *15*, 25–30.
Reboredo, J. C. (2013a). Is gold a hedge or safe haven against oil price movements? *Resources Policy*, *38*, 130–137.
Reboredo, J. C. (2013b). Is gold a safe haven or a hedge for the US dollar? Implications for risk management. *Journal of Banking & Finance*, *37*, 2665–2676.
Xu, X. E., & Fung, H.-G. (2005). Cross-market linkages between U.S. and Japanese precious metals futures trading. *Journal of International Financial Markets, Institutions and Money*, *15*, 107–124.
Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, *18*, 931–955.
Zhou, J. (2013). Extreme risk spillover among international REIT markets. *Applied Financial Economics*, *23*, 91–103.
Zhu, H.-M., Li, R., & Li, S. (2014). Modelling dynamic dependence between crude oil prices and Asia-Pacific stock market returns. *International Review of Economics & Finance*, *29*, 208–223.