



Recent evolution in Financial Econometrics: Modeling high frequency data, joint dependence and testing forecasting performance

田, 帅如

(Degree)

博士 (経済学)

(Date of Degree)

2017-03-25

(Date of Publication)

2018-03-01

(Resource Type)

doctoral thesis

(Report Number)

甲第6828号

(URL)

<https://hdl.handle.net/20.500.14094/D1006828>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



博士論文

Recent evolution in financial econometrics: modeling high frequency data, joint
dependence and testing forecast performance

金融計量経済学における進展：高頻度データと相互依存関係のモデリング、
及び予測力のテスト

平成28年12月

神戸大学経済学研究科

経済学専攻

指導教員 羽森 茂之

田 帥如

Contents

1. Introduction.....	2
2. Modeling interest rate volatility: A Realized GARCH approach.....	3
2.1 Introduction.....	3
2.2 Methodology.....	4
2.3 Data.....	10
2.4 Empirical results.....	11
2.5 ARMA-RGARCH model.....	16
2.6 Conclusion.....	18
Appendix.....	19
Table and Figure.....	22
3. Realizing Moments.....	38
3.1 Introduction.....	38
3.2 Estimation method of Realizing moments.....	39
3.3 Data.....	41
3.4 Joint behavior between return and realizing moments	41
3.5 Studying return predictability using realizing moments.....	42
3.6 Studying information transmission mechanism using realizing moments.....	43
3.7 Conclusion.....	44
Table and Figure.....	45
4. Improving density forecasting: the role of multivariate conditional dependence....	67
4.1 Introduction.....	67
4.2 Model specification.....	70
4.3 Density forecasting and evaluation.....	74
4.4 Empirical results.....	75
4.5 Conclusion.....	79
Table and Figure.....	80
Reference.....	88

Chapter 1. Introduction.

Over the past three decades, research interest of modeling financial returns has been growing rapidly. Many literatures in financial econometrics have been focused on volatility modeling by fitting parametric models with distributional assumptions (such as GARCH and Stochastic Volatility type models). While as discussed in Anderson (2001), these models suffer from model misspecification, and none is strictly correct. It is also widely recognized that these models suffer from “curse of dimensionality”, and cannot be applied in large dimension cases. In addition, information driven by intraday data is completely ignored in these models.

This paper contributes to two key topics addressed in recent literatures which concerned these problems: (1) modeling high-frequency data and (2) capture comprehensive features of the joint distribution of financial returns.

In chapter 2, we estimate the daily volatility of short term interest rate using a Realized GARCH approach. We find the proposed method fit the data better and provide more accurate volatility forecast by extracting additional information from realized measures. In addition, we propose using the ARMA-Realized GARCH model to capture the volatility clustering and the mean reversion effects of interest rate behavior. We find the ARMA-RGARCH model fits the data better than the simple RGARCH model dose, but it does not provide superior volatility forecasts.

In chapter3, we propose a new type of moments which can describe current features of daily return. The so called “realizing moments” can be estimated using high-frequency intraday data. We show some usages of this new type of moments, such as investigating joint behavior of return and its moments, return predictability, and information transmission mechanism. Many insightful results are found. Future research can use this new type of moments to study many other problems in financial economics.

In chapter 4, we investigate whether the density forecast forecasts of stock returns can be improved by taking account of conditional dependence. The Regular Vine type Copula-GARCH models are applied to accurately capture comprehensive features of the joint distribution of stock returns. Density forecasts are evaluated by adopting the correct specification test based on the probability integral transform. Presented results show that density forecast can be improved significantly after incorporating conditional dependence.

Chapter 2

Modeling interest rate volatility: A Realized GARCH approach

2.1 Introduction

Short-term interest rates are widely recognized as key economic variables. They are used frequently in financial econometrics models because they play an important role in evaluating almost all securities and macroeconomic variables. However, many popular models fail to capture the key features of interest rates and do not fit the data well. A milestone in terms of interest rate models was the development of the generalized regime-switching (GRS) model, proposed by Gray (1996). Conventional GARCH-type and diffusion models failed to handle certain interest rate events, such as explosive volatility, which would cause serious problems in certain applications. He believed this failure may be due to time variations in the parameter values. The GRS model nests many interest rate models as special cases, and allows the parameter values to vary with regime changes. Thus, the model should provide a solution to the problem. He found that the GRS model outperforms conventional single-regime GARCH-type models in out-of-sample forecasting. Unfortunately, the GRS model still does not fit the data well, since almost all the reported parameters of the most generalized version are nonsignificant.

Another conceivable reason for the failure to model short-term interest rates adequately is that the σ -field is not sufficiently informative. When modeling interest rates using conventional GARCH-type models, the only data used are the daily closing prices. All data during trading hours are ignored. As high-frequency data has become more available, recent literatures have introduced a number of more efficient nonparametric estimators of integrated volatility (see Barndorff-Nielsen and Shephard, 2004; Barndorff-Nielsen et al., 2008; and Hansen and Horel, 2009). Examples of models that incorporate these realized measures include the multiplicative error model (MEM) of Engle and Gallo (2006), and the HEAVY model of Shephard and Sheppard (2010). These models incorporate multiple latent variables of daily volatility. Then, within the framework of stochastic volatility models, Takahashi et al. (2009) propose a joint model for return and realized measures. Shirota et al. (2014) introduce the realized stochastic volatility (RSV) model, which incorporates leverage and long memory. A Heston model studies the joint behavior of return and volatility, and shows decisively different extremal behavior to the conventional GARCH and SV models proposed by Ehlert et al. (2015).

In this study, we model the short-term interest rate in the euro-yen market using the Realized GARCH (RGARCH) framework proposed by Hansen et al. (2012). Section 2 introduces the RGARCH framework, including the log-linear specification, the estimation method, the conditional distributions, the robust QLIKE loss function, and the value-at-risk-based loss function. Here, we also present the model confidence set (MCS) procedure we use to evaluate the volatility forecasting performance. Section 3 provides a brief description of our data and the realized measures. Section 4 reports the in-sample empirical results and evaluates the rolling-window volatility forecasting performance. Then, in Section 5, we introduce an ARMA-Realized GARCH (ARMA-RGARCH) framework. Here, we present our empirical results, and compare them to those of the simple RGARCH model. Lastly, Section 6 concludes the paper.

2.2. Methodology

2.2.1. Model specification

We adopt the RGARCH model with its log-linear specification, as proposed by Hansen et al. (2012). In Section 5, we propose an extension of the model to capture the well-known mean reversion and volatility effects in short-term interest rate behavior. The RGARCH(p,q) model is specified as follows:

$$r_t = \sqrt{h_t} z_t \quad (2.1)$$

$$\log h_t = \omega + \sum_{i=1}^p \beta_i \log h_{t-i} + \sum_{j=1}^q \gamma_j \log x_{t-j} \quad (2.2)$$

$$\log x_t = \xi + \varphi \log h_t + \tau(z_t) + u_t, \quad (2.3)$$

where $z_t \sim I.I.D(0,1)$, $u_t \sim I.I.D(0, \sigma_u^2)$, r_t is a zero-mean series, and h_t and x_t denote the conditional variance and realized measure, respectively. Then, $\tau(z_t)$ is the leverage function constructed from Hermite polynomials. Here, we adopt the simple quadratic form: $\tau(z_t) = \eta_1 z_t + \eta_2 (z_t^2 - 1)$. This choice is proper for two reasons. First, it satisfies $E[\tau(z_t)] = 0$ for any standard distribution with $E[z_t] = 0$ and $Var[z_t] = 1$. Second, it is proportional to the news impact curve discussed in Engle and Ng (1993). Thus, it can capture the asymmetric effect of

price shocks on daily volatility. The news impact curve is defined as $v(z) = E(\log h_t | z_{t-1} = z) - E(\log h_t)$.

Equations (1) and (2) construct a GARCH-X model, with the restriction that the coefficient of the squared return is zero. Equation (3) is called the measurement equation, because x_t is a measure of h_t . The measurement equation completes the model, using the leverage function $\tau(z_t)$ to provide a simple way to investigate the joint dependence of r_t and x_t . Numerous studies on market microstructures argue that returns are dependent on trading intensity and liquidity indicators, such as volume, order flow, and the bid–ask spread (see Amihud (2002), Admati and Pfleiderer (1988), Brennan and Subrahmanyam (1996), and Hasbrouck and Seppi (2001)). Because the intraday volatility is linked to the volume, order flow, and trading intensity directly, it should be dependent on the return. Therefore, the measurement equation is important from a theoretical point of view. Unlike other models that incorporate realized measures in the variance equation, such as the MEM and HEAVY models, the realized measure is not treated as an exogenous variable. When the realized measure is a consistent estimator of integrated volatility, it should be viewed as the conditional variance plus an innovation term. The conditional variance is adapted to a much richer σ -field, $F_t = \sigma(r_t, r_{t-1}, \dots, x_t, x_{t-1}, \dots)$. Conventional GARCH and SV-type models only use daily closing prices. In comparison, the additional information included in the realized measure on intraday volatility is expected to promote the fit to the data and the forecasting accuracy of conditional volatility. Moreover, the logarithmic conditional variance can be shown to follow an ARMA process:

$$\log h_t = \omega + \sum_{i=1}^{p \vee q} (\beta_i + \phi \gamma_i) \log h_{t-i} + \sum_{j=1}^q \gamma_j [\tau(z_{t-j}) + u_{t-j} + \xi]. \quad (2.4)$$

The logarithmic conditional variance, h_t , is driven by both the innovation of return and the realized measure. The leverage effect is already indirectly embedded in the variance equation. The persistence parameter, π , is given by $\pi = \sum_i (\beta_i + \phi \gamma_i)$.

2.2.2. Estimation method

Following Hansen et al. (2012), we summarize the estimation method in this section. The model is estimated using the quasi-maximum likelihood (QML) method. The likelihood function of the Gaussian specification is given by

$$l(r, x; \theta) = -\frac{1}{2} \sum_{t=1}^n [\log h_t + \frac{r_t^2}{h_t} + \log(\sigma_u^2) + \frac{u_t^2}{\sigma_u^2}]. \quad (2.5)$$

Since conventional GARCH-type models do not model realized measures, it is meaningless to compare the joint log likelihood to those of the conventional GARCH models. However, we can derive the partial likelihood of the RGARCH model and compare that with the log likelihood of the GARCH models. The joint conditional density of the return and realized measures is

$$f(r_t, x_t | F_{t-1}) = f(r_t | F_{t-1}) f(x_t | r_t, F_{t-1}). \quad (2.6)$$

Then, the logarithmic form is

$$\log f(r_t, x_t | F_{t-1}) = \log f(r_t | F_{t-1}) + \log f(x_t | r_t, F_{t-1}). \quad (2.7)$$

Thus, the joint log likelihood of z_t and u_t under the Gaussian specification can be split as follows:

$$l(r, x; \theta) = -\frac{1}{2} \sum_{t=1}^n [\log(2\pi) + \log h_t + \frac{r_t^2}{h_t}] - \frac{1}{2} \sum_{t=1}^n [\log(2\pi) + \log(\sigma_u^2) + \frac{u_t^2}{\sigma_u^2}]. \quad (2.8)$$

Then, the partial likelihood of the model is defined as

$$l(r; \theta) = -\frac{1}{2} \sum_{t=1}^n [\log(2\pi) + \log h_t + \frac{r_t^2}{h_t}]. \quad (2.9)$$

Hansen et al. (2012) obtained that $\sqrt{n}(\hat{\theta}_n - \theta) \rightarrow N(0, I_\theta^{-1} J_\theta I_\theta^{-1})$ where expressions for I_θ and J_θ are given in the appendix.

An alternative estimation method for GARCH-type models is proposed by Ossandon and Bahamonde (2011). It is possible to have a novel state space representation and an efficient approach based on the Extended Kalman Filter (EKF). Since the structure of the RGARCH model is quite complicated, this is left as a topic for further research.

We also change the assumption of a standard normal distribution on z_t to a more realistic distribution. Here, we adopt the distribution proposed by Fernandez and Steel (1998) that allows for skewness in any symmetric and continuous distribution by changing the scale of the density function:

$$f(z|\varsigma) = \frac{2}{\varsigma + \varsigma^{-1}} [f(\varsigma z)H(-z) + f(\varsigma^{-1}z)H(z)], \quad (2.10)$$

where ς is the shape parameter and $H(\cdot)$ is the Heaviside function. The distribution is symmetric when ς is equal to 1. The mean and variance are defined as

$$\mu_z = M_1(\varsigma - \varsigma^{-1}) \quad (2.11)$$

$$\sigma_z^2 = (M_2 - M_1^2)(\varsigma^2 + \varsigma^{-2}) + 2M_1^2 - M_2, \quad (2.12)$$

respectively, where

$$M_k = 2 \int_0^\infty z^k f(z) dz. \quad (2.13)$$

Here, M_k is the k -th moment on the positive real line. The distribution is called a standardized skewed distribution if $\mu_z = 0$ and $\sigma_z^2 = 1$. We adopt the standardized skewed t distribution. Here, the density function is given by

$$f(z_t|\varsigma, \nu) = \begin{cases} \frac{2\sigma}{\varsigma + \varsigma^{-1}} g[\varsigma(sz_t + \mu)|\nu] & \text{if } z_t < -\frac{\mu}{\sigma} \\ \frac{2\sigma}{\varsigma + \varsigma^{-1}} g[\varsigma^{-1}(sz_t + \mu)|\nu] & \text{if } z_t > -\frac{\mu}{\sigma} \end{cases}, \quad (2.14)$$

where z_t follows a $SKST(0,1,\varsigma,\nu)$ distribution, $g(\cdot)$ is the density function of the standard t distribution, and μ and σ^2 are the mean and variance of a non-standard skewed Student t distribution, respectively. The distribution is symmetric when $\varsigma = 1$.

2.2.3. Benchmark models

We compare the RGARCH model to the benchmark GARCH and EGARCH models. A zero-mean series can be specified by a GARCH model as

$$r_t = \sqrt{h_t} z_t \quad (2.15)$$

$$h_t = \omega + \sum_{i=1}^p \alpha_i r_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}, \quad (2.16)$$

and the EGARCH model of Nelson (1991) is defined as

$$r_t = \sqrt{h_t} z_t \quad (2.17)$$

$$\log h_t = \omega + \sum_{i=1}^p \beta_i \log h_{t-i} + \sum_{j=1}^q \alpha_j [\theta z_{t-j} + \gamma(|z_{t-j}| - E|z_{t-j}|)]. \quad (2.18)$$

The log likelihood of the GARCH and EGARCH models under a Gaussian specification is given by

$$l(r; \theta) = -\frac{1}{2} \sum_{t=1}^n [\log(2\pi) + \log h_t + \frac{r_t^2}{h_t}]. \quad (2.19)$$

Note that the log likelihood of the benchmarks takes the same form as the partial likelihood of the RGARCH model. Therefore, we can compare the fitness of these models.

2.2.4. Forecasting performance evaluation

A reliable model should fit the data well, but should also generate an accurate forecast. We compare the forecasting performance by constructing loss functions. The first loss function we construct is the robust loss function QLIKE, as discussed in Patton (2011). This is called a “direct” evaluation because it compares the volatility forecasts to their realizations directly. As realizations of volatility are unobservable, a volatility proxy is required to construct the loss function. The QLIKE loss function is given by

$$QLIKE: L(\hat{\sigma}^2, h) = \log h + \frac{\hat{\sigma}^2}{h}, \quad (2.20)$$

where $\hat{\sigma}^2$ is a proxy for the true conditional variance and h is a volatility forecast. The squared return is a widely used proxy, but is also known as being rather noisy. Patton (2011) suggests using the realized volatility and intra-daily range as unbiased estimators. These are also more efficient than the squared return if the log price follows a Brownian motion. Although these less

noisy proxies should lead to less distortion, the degree of distortion is still large in many cases. Nagata and Oya (2012) point out that realized volatility does not satisfy the unbiasedness condition owing to market microstructure noise. They show that bias in the volatility proxy can cause misspecified rankings of competing models. As the realized kernel is designed to deal with microstructure noise, it can clearly also be used as an unbiased estimator of conditional variance. Since high frequency prices are only available during trading periods, the realized kernel and realized volatility are directly related to the open-to-close return. Thus, for the close-to-close return, the squared return is the only unbiased proxy. However, the squared open-to-close return is not an unbiased estimator for daily (close-to-close) volatility owing to the existence of the overnight return (see the appendix for the proof). In this study, we use the squared return, realized volatility, and the realized kernel to construct QLIKE loss function.

The second loss function we construct is a value-at-risk (VaR)-based loss function, as proposed by González-Rivera et al. (2004). The value-at-risk is estimated from

$$VaR_t^\alpha = \mu_t + \Phi_t^{-1}(\alpha)\sigma_t, \quad (2.21)$$

where Φ_t is the cumulative distribution function, and μ_t and σ_t are the conditional mean and variance, respectively. Then the asymmetric VaR loss function is given by

$$L(y_t, VaR_t^\alpha) = (\alpha - d_t^\alpha)(y_t - VaR_t^\alpha), \quad (2.22)$$

where $d_t^\alpha = 1(y_t < VaR_t^\alpha)$ is an indicator variable. This is an asymmetric loss function since it penalizes VaR violations more heavily, with weight $(1 - \alpha)$. We call this an “indirect” evaluation because of its risk measurement point of view. Smaller values are preferred for both loss functions.

However, it is difficult to say whether a model with the smallest loss values significantly outperforms alternative models. Thus, we adopt the model confidence set (MCS) procedure proposed by Hansen et al. (2011) to test the equal predictive ability (EPA) hypothesis and to obtain the superior set of models (SSM) under each loss function. The estimation and conditional volatility evaluation are all performed in R(R Core Team (2015)).

2.3. Data

2.3.1. Data and the rolling-over method

For our empirical application, we use high-frequency data of three-month euro-yen futures from the Tokyo Financial Exchange (TFX). Since derivative contracts have a finite life, we construct a continuous time series using the rolling-over method and “volume” criteria, as discussed in Carchano and Pardo (2009). Our full sample comprises 980 observations from January 2006 to December 2009. We estimate both close-to-close return and open-to-close return series. To evaluate the forecasting performance, we use rolling-window method, with a window size of 200 observations, and obtain one-step-ahead out-of-sample forecasts.

2.3.2. Realized measures

Recent literature has introduced a number of realized measures of volatility, such as realized variance, realized kernel, bipower variation, intraday range, and other nonparametric estimators. In this study, we approximate the integrated volatility using two estimators. The first is the five-minute realized variance estimator of Andersen et al. (2001), defined as the summation of the squared intraday returns:

$$RV_t = \sum_{i=1}^n r_{i,t}^2 . \quad (2.23)$$

The second estimator is the realized kernel of Barndorff-Nielsen et al. (2008), given by

$$K(X) = \sum_{h=-H}^H k\left(\frac{h}{H+1}\right) \gamma_h , \quad (2.24)$$

$$\gamma_h = \sum_{j=|h|+1}^n x_j x_{j-|h|} , \quad (2.25)$$

where $k(x)$ is a kernel weight function and H is the bandwidth. Here, we adopt the Parzen kernel, which guarantees non-negativity and satisfies the smoothness conditions, $k'(0) = k'(1) = 0$. The Parzen kernel function is defined as follows:

$$k(x) = \begin{cases} 1 - 6x^2 + 6x^3 & 0 \leq x < 1/2 \\ 2(1-x)^3 & 1/2 \leq x < 1 \\ 0 & x > 1 \end{cases} . \quad (2.26)$$

Figure 1 represents the series of returns and realized measures.

2.4. Empirical results

2.4.1. In-sample empirical results with a normal distribution

In this section, we present our empirical results for the close-to-close return and the open-to-close return with a standard normal distribution, and compare them to those of the benchmarks. For simplicity, RG, G, and EG denote the RGARCH, GARCH, and EGARCH models, respectively.¹ Table 1 reports the main empirical results for the close-to-close return (the complete tables are available online).

Several points are worth noting. First, the joint likelihood of the return and volatility indicate that the simplicity of an RG(1,1) model is proper. The AIC and SBIC also suggest that the RG(1,1) model should be adopted. As a result, we estimate the G(1,1) and EG(1,1) models for comparison. In terms of the partial log likelihood, the RG models clearly outperform the benchmarks, although the RGARCH models and the QML estimation methods do not maximize the partial log likelihood. The additional information contained in the intraday volatility improves the fit to the daily return.

Second, the RG persistence parameter, π , is about 0.91. This indicates strong persistence, but is also the covariance stationarity of the short-term interest rate. Then, the value of γ_2 is almost 0 and is nonsignificant, which indicates that the current volatility is strongly affected only by the latest realized measures. In the case of the GARCH and EGARCH models, the conditional volatility has extremely long memory owing to the high value of β . Furthermore, these models are almost unaffected by recent news, as the values of α in each case are very small and nonsignificant. In contrast, the value of β in the RGARCH models is much smaller, indicating a much shorter memory in the model's volatility process. On the other hand, a larger proportion of the conditional volatility is due to recent news, since γ_1 in the RG models is about 0.29 and is

¹ Following packages are used to obtain the empirical results: rugarch (Ghalanos (2014)), copula (Hofert et al.(2015)), texmex (Southworth and Heffernan(2013)), MCS (Catania and Bernardi(2015)).

highly significant. The positive values of parameters η_1 and η_2 of the leverage function confirm the leverage effect of the interest rate. Furthermore, the significance of these parameters implies the existence of joint return and intraday volatility behavior, because the innovation term of the return is embedded in the measurement equation. Therefore, it is proper to consider that trading intensity and liquidity indicators impact the return.

Then, Θ is the extremal index, which describes the extremal behavior of a stochastic process, as discussed in De Haan et al. (1989), Bhattacharya (2008), and Laurini et al. (2012). The extremal index is estimated by 5000 simulations, with fixed parameters, using the estimation method proposed by Ferro and Segers (2003). Here, $\Theta \in (0,1)$, and lower values indicate higher clusters of extreme values for a given confidence level. In this study, the confidence level is set to 95%. The value of Θ is approximately 0.91 for the GARCH and EGARCH models, and approximately 0.72 for the RGARCH models. This means the GARCH and EGARCH processes assume almost no clustering of extreme values, while the RGARCH processes assume some clustering.

Figure 2 shows the news impact curve defined in Section 2. It is clear that positive price shocks have a larger impact on volatility than do negative shocks. As a result, interest rates rise slowly and fall quickly.

Table 2 reports the empirical results for the open-to-close return. As in the case of the close-to-close return, the values of γ_2 suggest that the conditional variance is tied only to the latest realized volatility. The persistence parameter, π , still suggests strong persistence and covariance stationarity. Then, the extremal index shows that the EGARCH process captures almost the same level of volatility clustering as the RGARCH models do. On the other hand, extreme open-to-close return values are a bit more clustered than are extreme close-to-close return values. Other parameters suggest almost the same results. Although the partial log likelihood of the RGARCH models is smaller than that of the EGARCH model, this is not a problem, because the RGARCH models do not maximize the partial log likelihood. However, in terms of the joint log likelihoods and information criteria, it is not proper to adopt the simplicity of the RG(1,1) model. This is discussed in the next section.

2.4.2. In-sample empirical results with a standard skewed Student t distribution

In this section, we change the assumption of a standard normal distribution to a more generalized distribution, called a standard skewed Student t distribution, or SSKT(0,1). Table 3 presents the estimates for the close-to-close return under the new assumption. The coefficients report almost the same results as in the case of the normal distribution assumption. The skew parameter is approximately 1.03, which means the conditional distribution of the return is almost symmetric. Note that since a positive return means a fall in the interest rate, we refer to this as negative. Thus, negative news has a slightly larger impact on volatility than does positive news. The extremal indices are around 0.85, indicating that under the assumption of a skewed t distribution, extreme returns do not cluster as frequently as they do under a normal distribution. As γ_2 is again 0 and nonsignificant, it is proper to select the best model between RG(1,1) and RG(2,1). The likelihood ratio statistics between these two models for the realized kernel and realized volatility are 8.18 and 8.22, respectively, which are both greater than $6.63 (= \chi^2(1)_{0.01})$. The AIC and SBIC values indicate that the RG(2,1) model fits the data better than does the RG(1,1) model. Therefore, an RG(2,1) model should be proper.

Table 4 reports the empirical results for the open-to-close return under the assumption of a skewed t distribution. Three points deserve to be mentioned. First, the skew parameters are equal to 1, indicating the same impact of positive and negative news on volatility. Second, the extremal indices are larger than those of the normal distribution, which indicates relatively slight volatility clustering. Finally, the likelihood ratio statistics for the RG(1,1) and RG(2,1) models for the realized measures are 4.76 and 4.24, respectively, which are greater than $3.84 (= \chi^2(1)_{0.05})$, but less than $6.63 (= \chi^2(1)_{0.01})$. The AIC and SBIC values suggest different results for model selection.

In a correctly specified model, robust and non-robust standard errors should be in agreement, see Hansen et al.(2012). The non-robust and robust standard errors are diagonal elements of the inverse Fisher information matrix I_θ^{-1} and $I_\theta^{-1}J_\theta I_\theta^{-1}$. Tables 5 to 8 present these standard errors for both return series and the distribution assumptions. The results are as follows. For the close-to-close return, the distances between I_θ^{-1} and $I_\theta^{-1}J_\theta I_\theta^{-1}$ for the GARCH and EGARCH models are quite large, indicating serious misspecifications. However, the RGARCH models are far less misspecified. On the other hand, the RGARCH models with the skewed t distribution show a

much better fit than they do with the normal distribution. The RG(2,1) model with a skewed t distribution clearly performs better than the RG(1,1) model, suggesting the same result as the likelihood ratio test and information criteria. For the open-to-close return, the EGARCH model does not perform worse than the GARCH model, but still cannot compete with the RGARCH models. However, it is quite difficult to distinguish between the normal and skewed t distribution, as well as between the RG(1,1) and RG(2,1) models. Thus, we evaluate the forecasting performance for all these models.

2.4.3. Forecasting performance evaluation

As mentioned in Sections 2 and 3, we evaluate the forecasting performance by constructing QLIKE and asymmetric VaR loss functions. Then, we test the equal predictive ability (EPA) hypothesis using the model confidence set (MCS) procedure. The out-of-sample conditional volatility is estimated by the rolling-window method, with a window size of 200 observations. Each point in rolling-window estimation is checked to avoid sub optimization. The random seed of MCS sampling would affect the p-values to a small extent, but not change the rankings. Following the results in Section 4.1 and 4.2, for the close-to-close return, we evaluate six models, including GARCH, EGARCH, and RG(1,1) with a normal distribution and RG(2,1) with a skewed t distribution, using both realized measures. For the open-to-close return, we evaluate 10 models, including the benchmark models, and RG(1,1) and RG(2,1) with both the normal and skewed t distributions, and using both realized measures. To evaluate the QLIKE loss function, the squared return, realized volatility, and realized kernel are all used as volatility proxies.

2.4.3.1. QLIKE evaluation

Table 9 reports the MCS evaluation for the QLIKE loss function. The confidence level of the EPA hypothesis is 10%. Owing to microstructure noise, the realized volatility is not an unbiased estimator of the open-to-close volatility (Nagata and Oya, 2012), but the squared return and realized kernel are. In addition, the squared return is quite noisy (Patton, 2011), while the realized kernel and realized volatility are more efficient. Although biased proxies may lead to misspecified rankings, we still report these results. Models with ‘ELI’ in the column of MCS p-values were eliminated under the given confidence level using the elimination rule of Hansen et al. (2011).

For the close-to-close returns, using the squared return as a volatility proxy, the RG(1,1) model with a normal distribution and the realized kernel as the realized measure performs best. However, the null hypothesis that the GARCH and RG(1,1) models, using realized volatility, have equal predictive ability cannot be rejected. Since the MCS p-values are close to 1, it is difficult to distinguish between the predictive ability of the two RG models, but they clearly outperform the GARCH model. Using the realized kernel and realized volatility as volatility proxies, no other model has a predictive ability equal to that of the best performing model. The RG(1,1) model with a normal distribution and the realized kernel performs best under the evaluation of all volatility proxies. Therefore, it is proper to consider that using the realized kernel as a realized measure provides better volatility forecasts than when using the realized volatility. For the open-to-close return, it is quite difficult to obtain agreement. The results depend on the volatility proxy used. However, as mentioned above, the realized kernel is unbiased and the most efficient volatility proxy, we adopt the result that the RG(2,1) model with the skewed t distribution and the realized kernel as the realized measure has the best predictive ability.

2.4.3.2. *Value-at-Risk evaluation*

The VaR evaluation is a widely used for conditional volatility assessment. A VaR estimate is said to be valid if it satisfies the unconditional convergence condition of Kupiec (1995) and the independence and conditional convergence of Christoffersen (1998). We also report the results of the VaR independence duration test proposed by Christoffersen and Pelletier (2004). Table 10 presents the VaR backtesting results. For the close-to-close return, all models generate valid VaR forecasts, except the EGARCH model. For the open-to-close return, the RG model with the normal distribution and the realized kernel performs best. Other models fail to reject several hypotheses. Since the exceedance ratio of the RG model with the skewed t distribution is 0 for 0.5% VaR, we cannot calculate the likelihood ratio statistics for the VaR backtesting. Overall, the VaR forecasts generated by the RG models are reliable.

However, instead of validity, it is of more interest to investigate the predictive accuracy of the VaR estimate. Thus, we construct the asymmetric VaR loss function mentioned in Section 2 and test the EPA hypothesis. Table 11 reports the MCS evaluation for the VaR loss function. For the close-to-close return, it is clear that the RG(1,1) model with the normal distribution dominates

the alternatives for both the 0.5% and 1% VaR estimates. The predictive accuracy of estimation using the realized kernel is a little better than that of the realized volatility. For the open-to-close return, the RG models with the normal distribution have almost equal predictive ability, and dominate the other models.

Considering all our in-sample empirical results, the QLIKE loss function, and the VaR estimate, the RGARCH models clearly dominate the benchmarks. Although the RGARCH models with the skewed t distribution fit the data better for the in-sample close-to-close return, they cannot provide superior volatility and VaR forecasts. Clearly, the forecast performance is affected more by the specifications of the models than by the chosen realized measures. The RG model with a normal distribution and the realized kernel as a realized measure performs best in the QLIKE evaluation, using all volatility proxies. The model also performs slightly better than the model that uses the realized volatility in the VaR evaluation. Thus, it is proper to consider that the realized kernel is a better realized measure than is the realized volatility. On the other hand, it is quite difficult to select the best performing model for the open-to-close return. Consequently, we extend the mean process of RG(1,1) with the normal distribution and the realized kernel for the close-to-close return in Section 5.

2.5. ARMA-RGARCH model

Gray (1996) suggests that models of short-term interest rates should capture two well-known empirical attributes, namely mean reversion and leptokurtosis. Engle (1982) shows that leptokurtosis in an unconditional distribution may be caused by conditional heteroskedasticity. Conditional heteroskedasticity is already modeled by the volatility process and leverage function. The RGARCH model assumes the conditional mean $E[r_t | F_{t-1}] = 0$. To capture the level behavior of short-term interest rates, a more general specification of the mean process is proper. A simple way to model the mean reversion effect is to impose an ARMA process on the mean process. In this section, we generalize the mean process to an ARMA process to model both the mean reversion and the volatility clustering effects simultaneously. The ARMA-RGARCH model is specified as follows:

$$r_t = \mu + \sum_{i=1}^p \alpha_i r_{t-i} + \sum_{j=1}^q \kappa_j \varepsilon_{t-j} + \varepsilon_t$$

$$\varepsilon_t = \sqrt{h_t} z_t$$

$$\log h_t = \omega + \sum_{m=1}^r \beta_m \log h_{t-m} + \sum_{n=1}^s \gamma_n \log x_{t-n}$$

$$\log x_t = \xi + \varphi \log h_t + \tau(z_t) + u_t .$$

The RGARCH model is nested within the ARMA-RGARCH model by imposing $\mu = \alpha_i = \beta_j = 0$. Here, α_1 should be positive to model volatility clustering, and κ_1 should be negative if the mean reversion effect exists. For simplicity, we estimate an ARMA(2,2)-RGARCH(1,1) model with a standard normal distribution and using the realized kernel as the realized measure, as we have already discussed the volatility process in Section 4. The results are as follows:

$$r_t = -0.000003 + 1.621r_{t-1} - 0.995r_{t-2} - 1.62\varepsilon_{t-1} + 1.004\varepsilon_{t-2} + \varepsilon_t$$

(6e-6) (0.003) (0.004) (3e-5) (3e-6)

$$\log h_t = -1.523 + 0.627 \log h_{t-1} + 0.282 \log x_{t-1}$$

(0.273) (0.012) (0.008)

$$\log x_t = -0.212 + 1.018 \log h_{t-1} + 0.072 z_t + 0.02(z_t^2 - 1) + u_t$$

(0.964) (0.061) (0.038) (0.003)

$$l(r, x | \theta) = 6068.111 \quad l(r | \theta) = 6975.64 \quad \pi = 0.91 \quad \Theta = 0.72 .$$

Here, the numbers in parentheses are robust standard errors. Note the following points. First, the likelihood ratio statistics of the ARMA-RGARCH model and of its benchmark RG(1,1) model are 21.102, which is larger than $\chi_{0.01}^2(5) = 15.09$. This suggests a significant increase in fitting the data. This finding is also evident in the increase of the partial likelihood, after imposing an ARMA process. Second, the coefficients α_1 and κ_1 are positive and negative, respectively, and also highly significant, showing the existence of strong volatility clustering and mean reversion effects. Other properties are similar to the simple RGARCH model. Moreover, since we assumed independence between z_t and u_t , $\{\hat{z}_t, \hat{u}_t\}$ should be independently distributed if the model is well specified. Genest and Rémillard (2004) studied tests of independence and randomness based on a decomposition of an empirical copula process. Kojadinovic and Yan (2011) proposed a

generalization of the decomposition to test serial independence in a continuous multivariate time series. Since the statistics of serial dependence are not distribution free, they studied the consistency using the bootstrap methodology. We investigate the serial independence between z_t and u_t of the RGARCH and ARMA-RGARCH models, assuming a realized kernel and standard normal distribution on the innovation process, using the empirical copula method. Table 12 reports the p-values of the global Cramer–Von Mises test, the combined tests from the Mobius decomposition with Fisher’s rule, and Tippet’s rule. The table also evaluates the forecasting performance of the ARMA-RGARCH and simple RGARCH models. The results are as follows. The p-values for the multivariate serial independence test strongly suggest serial independence between z_t and u_t for both models. As a result, neither the ARMA-RGARCH model nor the simple RGARCH model are misspecified. The simple RGARCH model performs better in the VaR estimate. However, the evaluation of the QLIKE loss function does not suggest the same result. Using a squared return as a volatility proxy, the MCS p-values cannot reject the EPA null hypothesis. Then, using the realized kernel and realized volatility as proxies, the ARMA-RGARCH model performs better than the benchmarks. Figure 3 represents the series of parameters and the conditional standard errors using the rolling-window method for the two models. The scales of y-axis are 70 times of standard errors for beta and gamma and 10 times for eta1 and eta2. It is clear that the patterns of parameters for the two models are almost the same. The variation of parameters is quite small over time. For the ARMA-RGARCH model, parameters are more volatile than simple RGARCH model. Overall, although modeling the volatility clustering and mean reversion effects improve the fit to the data, we cannot say that the forecasting accuracy has improved.

2.6. Conclusion

This study proposes an RGARCH approach to model short-term interest rates. The important empirical results include the following. First, the RGARCH model is an effective tool to model and forecast the volatility of an interest rate. The more informative σ -field improves the fit to the data and generates more accurate volatility forecasts and value-at-risk estimates. Second, generalizing the assumption of the conditional distribution does not improve the forecasting accuracy. The asymptotic properties and the QML estimation method perform well enough that the RGARCH model with a standard normal distribution can outperform conventional GARCH-

type models. Finally, the proposed ARMA-RGARCH model can capture the volatility clustering and mean reversion effects. The generalization of the mean process fits the data better than the simple RGARCH model does, but does not improve the forecasting accuracy of daily volatility. Neither the RGARCH nor the ARMA-RGARCH models are misspecified. Finally, although we only investigate one market, our results should be a general indicator for other markets.

Appendix A: Asymptotic distribution

Following Hansen et al (2012), we write the leverage function as $\tau(z_t) = \eta_1 a_1(z_1) + \dots + \eta_t a_t(z_t)$, and denote the parameters in the model by $\theta = (\lambda', \psi', \sigma_u^2)'$, where $\lambda = (\omega, \beta_1, \dots, \beta_p, \gamma_1, \dots, \gamma_q)'$ and $\psi = (\xi, \phi, \eta)'$. To simplify the notation we denote $v_t = (1, \log h_{t-1}, \dots, \log h_{t-p}, \log x_{t-1}, \log x_{t-q})'$ and $q_t = (1, \log h_t, a_t)'$. Thus, equation (2) and (3) can be expressed as $\log h_t = \lambda' v_t$ and $\log x_t = \psi' q_t + u_t$, respectively. The score, $\frac{\partial l}{\partial \theta} = \sum_{t=1}^n \frac{\partial l_t}{\partial \theta}$, is given by

$$\frac{\partial l_t}{\partial \theta} = -\frac{1}{2} \left\{ \left(1 - z_t^2 + \frac{2u_t}{\sigma_u^2} \dot{u}_t \right) \dot{h}_t, -\frac{2u_t}{\sigma_u^2} q_t, \frac{\sigma_u^2 - u_t^2}{\sigma_u^4} \right\}'$$

where $\dot{h}_t = \frac{\partial \log h_t}{\partial \lambda}$, $\dot{u}_t = \frac{\partial u_t}{\partial \log h_t} = -\phi + \frac{1}{2} z_t \eta' \dot{a}_t$ with $\dot{a}_t = \frac{\partial a(z_t)}{\partial z_t}$. The second derivative,

$\frac{\partial^2 l}{\partial \theta \partial \theta'} = \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \theta \partial \theta'}$, is given by

$$\frac{\partial^2 l_t}{\partial \theta \partial \theta'} = \begin{pmatrix} -\frac{1}{2} \left\{ z_t^2 + \frac{2(\dot{u}_t^2 + u_t \ddot{u}_t)}{\sigma_u^2} \right\} \dot{h}_t \dot{h}_t' - \frac{1}{2} \left\{ 1 - z_t^2 + \frac{2u_t \dot{u}_t}{\sigma_u^2} \right\} \ddot{h}_t & \bullet & \bullet \\ \frac{\dot{u}_t}{\sigma_u^2} q_t \dot{h}_t' + \frac{u_t}{\sigma_u^2} v_t \dot{h}_t' & -\frac{1}{\sigma_u^2} q_t q_t' & \bullet \\ \frac{u_t \dot{u}_t}{\sigma_u^4} \dot{h}_t' & \frac{u_t}{\sigma_u^4} q_t' & \frac{1}{2} \frac{\sigma_u^2 - 2u_t^2}{\sigma_u^6} \end{pmatrix}$$

where $\ddot{h}_t = \frac{\partial^2 \log h_t}{\partial \lambda \partial \lambda'}$, $v_t = (0, 1, -\frac{1}{2} z_t \dot{a}_t)$, $\ddot{u}_t = -\frac{1}{4} \eta' \{z_t \dot{a}_t + z_t^2 \ddot{a}_t\}$ with $\ddot{a}_t = \frac{\partial^2 a(z_t)}{\partial z_t^2}$. Suppose

$\{(r_t, x_t, \log h_t)\}$ is stationary and ergodic, then $\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial l_t}{\partial \theta} \xrightarrow{d} N(0, J_\theta)$ and $-\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \theta \partial \theta'} \xrightarrow{p} I_\theta$ with

$$J_\theta = \begin{pmatrix} \frac{1}{4} E(1 - z_t^2 + \frac{2u_t}{\sigma_u^2} \dot{u}_t)^2 E(\dot{h}_t \dot{h}_t') & \bullet & \bullet \\ -\frac{1}{\sigma_u^2} E(\dot{u}_t q_t \dot{h}_t') & \frac{1}{\sigma_u^2} E(q_t q_t') & \bullet \\ \frac{-E(u_t^3) E(\dot{u}_t)}{2\sigma_u^6} E(\dot{h}_t') & \frac{E(u_t^3)}{2\sigma_u^6} E(q_t') & \frac{E(u_t^2 / \sigma_u^2 - 1)^2}{4\sigma_u^4} \end{pmatrix}$$

and

$$I_\theta = \begin{pmatrix} \left\{ \frac{1}{2} + \frac{E(\dot{u}_t^2)}{\sigma_u^2} \right\} E(\dot{h}_t \dot{h}_t') & \bullet & 0 \\ -\frac{1}{\sigma_u^2} E\{(\dot{u}_t q_t + u_t v_t) \dot{h}_t'\} & \frac{1}{\sigma_u^2} E(q_t q_t') & 0 \\ 0 & 0 & \frac{1}{2\sigma_u^4} \end{pmatrix}$$

are finite.

Appendix B: Proof

Here, we prove that the squared open-to-close return is not an unbiased estimator of the true daily conditional variance. Denote r_{cc} , r_{oc} , and r_{ov} as the close-to-close, open-to-close, and overnight return, respectively. It is clear that $r_{cc,t} = r_{oc,t} + r_{ov,t}$,

$E[r_{cc,t} | F_{t-1}] = E[r_{oc,t} | F_{t-1}] = E[r_{ov,t} | F_{t-1}] = 0$, and $Var[r_{cc,t} | F_{t-1}] = E[r_{cc,t}^2 | F_{t-1}] = \sigma_t^2$, where σ_t^2 is the true conditional variance. Thus,

$$Var[r_{oc,t} | F_{t-1}] = Var[r_{cc,t} - r_{ov,t} | F_{t-1}]$$

$$\begin{aligned}
&= E[(r_{cc,t} - r_{ov,t})^2 | F_{t-1}] - E[r_{cc,t} - r_{ov,t} | F_{t-1}]^2 \\
&= E[r_{cc,t}^2 - 2r_{cc,t}r_{ov,t} + r_{ov,t}^2 | F_{t-1}] + \{E[r_{cc,t} | F_{t-1}] - E[r_{ov,t} | F_{t-1}]\}^2 \\
&= \sigma_t^2 - 2\text{cov}[r_{cc,t}r_{ov,t} | F_{t-1}] + \text{Var}[r_{ov,t} | F_{t-1}].
\end{aligned}$$

The squared open-to-close return is the unbiased estimator of the true conditional variance if and only if $\text{Var}[r_{ov,t} | F_{t-1}] = 2\text{cov}[r_{cc,t}r_{ov,t} | F_{t-1}]$, which is not an ordinary condition. For the same reason, the realized volatility and realized kernel are also not unbiased estimators of the daily volatility of the close-to-close return.

Table 1 Empirical results for close-to-close returns with standard normal distribution

			Realized Kernel				Realized Volatility			
	G(1,1)	EG(1,1)	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)
α	0.02	-0.05								
	(0.99)	(1.17)								
β_1	0.95	0.98	0.62	0.46	0.62	0.46	0.61	0.42	0.60	0.42
	(21.55)	(40.97)	(43.41)	(16.14)	(59.82)	(26.82)	(33.02)	(13.09)	(38.75)	(18.60)
β_2				0.16		0.16		0.18		0.18
				(2.48)		(1.90)		(2.35)		(1.91)
γ_1		0.06	0.28	0.29	0.28	0.29	0.28	0.30	0.29	0.30
		(1.38)	(18.84)	(5.28)	(57.63)	(4.71)	(13.07)	(3.89)	(21.88)	(3.84)
γ_2					0	1e-6			1e-6	0
					(5e-6)	(1.5e-5)			(1.4e-5)	(0)
η_1			0.07	0.07	0.07	0.07	0.08	0.08	0.08	0.08
			(1.85)	(1.83)	(1.85)	(1.83)	(2.17)	(2.17)	(2.17)	(2.17)
η_2			0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
			(3.79)	(3.77)	(3.76)	(3.74)	(3.89)	(3.91)	(3.88)	(3.88)
$l(r,x)$			6057.56	6057.74	6055.97	6057.74	6157.71	6158.17	6155.82	6158.17
$l(r)$	6931.84	6943.76	6963.15	6962.77	6962.84	6962.77	6957.35	6957.33	6957.00	6957.33
Θ	0.91	0.91	0.72	0.70	0.72	0.70	0.74	0.72	0.74	0.72
π	0.97	0.98	0.91	0.92	0.91	0.92	0.91	0.92	0.91	0.92
AIC	-14.145	-14.163	-12.346	-12.344	-12.341	-12.342	-12.550	-12.549	-12.545	-12.547
SBIC	-14.140	-14.143	-12.306	-12.299	-12.296	-12.292	-12.511	-12.504	-12.500	-12.497

Note: G, EG, and RG represent the GARCH, EGARCH, and RGARCH models, respectively. The realized kernel and realized volatility are the realized measures used in the estimations. The

values in parentheses are t values calculated using a robust standard error, and $l(r, x)$ and $l(r)$ denote the joint log likelihood and partial likelihood, respectively.

Table 2 Empirical results for the open-to-close return with standard normal distribution

	Realized Kernel					Realized Volatility				
	G(1,1)	EG(1,1)	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)
α	0.03	-5e-4								
	(3.9e-3)	(0.02)								
β_1	0.96	0.97	0.65	0.36	0.65	0.36	0.60	0.37	0.60	0.37
	(0.12)	(1272.86)	(1546.78)	(280.29)	(2215.93)	(441.43)	(1083.51)	(126.37)	(550.56)	(120.65)
β_2				0.28		0.28		0.23		0.23
				(75.44)		(81.29)		(8.42)		(8.21)
γ_1		0.20	0.32	0.35	0.32	0.35	0.40	0.42	0.40	0.42
		(17.18)	(179.97)	(123.63)	(58.25)	(105.55)	(443.75)	(130.23)	(534.07)	(124.90)
γ_2					0	0			0	0
					(2e-6)	(0)			(0)	(0)
η_1			-0.05	-0.05	-0.05	-0.05	-0.03	-0.03	-0.03	-0.03
			(1.98)	(1.97)	(1.98)	(1.97)	(1.58)	(1.59)	(1.59)	(1.59)
η_2			0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
			(5.05)	(4.93)	(5.05)	(4.97)	(5.65)	(5.53)	(5.62)	(5.51)
$l(r,x)$			6518.69	6520.37	6516.87	6520.37	6638.14	6639.12	6635.91	6639.12
$l(r)$	7339.18	7366.21	7356.23	7357.65	7355.84	7357.65	7359.31	7359.82	7358.87	7359.82
Θ	0.77	0.67	0.69	0.67	0.70	0.67	0.69	0.67	0.69	0.66
π	0.99	0.97	0.92	0.93	0.92	0.93	0.91	0.92	0.91	0.92
AIC	-14.972	-15.025	-13.287	-13.289	-13.281	-13.286	-13.531	-13.531	-13.524	-13.529

SBIC	-14.957	-15.005							
			-13.247	-13.244	-13.236	-13.237	-13.491	-13.486	-13.479

Note: G, EG, and RG represent the GARCH, EGARCH, and RGARCH models, respectively. The realized kernel and realized volatility are the realized measures used in the estimations. The values in parentheses are t values calculated using a robust standard error, and $l(r, x)$ and $l(r)$ denote the joint log likelihood and partial likelihood, respectively.

Table 3 Empirical results for the close-to-close return with skewed student t distribution

	Realized Kernel				Realized Volatility			
	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)
β_1	0.63	0.36	0.64	0.36	0.63	0.37	0.64	0.37
	(233.62)	(32.51)	(363.12)	(35.94)	(289.92)	(20.18)	(356.39)	(24.32)
β_2		0.28		0.28		0.26		0.26
		(6.69)		(6.88)		(3.95)		(4.37)
γ_1	0.36	0.37	0.34	0.36	0.36	0.37	0.35	0.37
	(105.11)	(7.80)	(118.74)	(9.29)	(101.41)	(6.64)	(173.63)	(8.52)
γ_2			0	0			0	0
			(0)	(0)			(0)	(0)
η_1	0.09	0.13	0.12	0.13	0.12	0.15	0.14	0.15
	(1.03)	(1.53)	(1.41)	(1.53)	(0.86)	(1.96)	(1.86)	(1.96)
η_2	0.03	0.06	0.06	0.06	0.04	0.07	0.06	0.07
	(0.68)	(1.42)	(1.14)	(1.42)	(0.49)	(2.03)	(1.77)	(2.04)
skew	1.03	1.03	1.03	1.03	1.03	1.02	1.03	1.02
$l(r, x)$	6318.43	6322.52	6318.55	6322.52	6421.82	6425.93	6422.15	6425.93
Θ	0.88	0.84	0.87	0.85	0.87	0.85	0.89	0.85
AIC	-12.874	-12.881	-12.873	-12.879	-13.085	-13.092	-13.084	-13.090

SBIC	-12.824	-12.826	-12.818	-12.819	-13.035	-13.037	-13.029	-13.030
------	---------	---------	---------	---------	---------	---------	---------	---------

Note: The realized kernel and realized volatility are the realized measures used in the estimations. The values in parentheses are t values calculated using a robust standard error, and $l(r, x)$ denotes the joint log likelihood.

Table 4 Empirical results for the open-to-close return with skewed student t distribution

	Realized Kernel				Realized Volatility			
	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)
$\beta 1$	0.64	0.32	0.64	0.32	0.62	0.32	0.62	0.32
	(1271.43)	(301.56)	(1934.73)	(485.13)	(1387.56)	(140.40)	(342.98)	(161.54)
$\beta 2$		0.31		0.31		0.29		0.29
		(116.61)		(133.06)		(24.86)		(30.86)
$\gamma 1$	0.32	0.35	0.32	0.35	0.37	0.40	0.37	0.40
	(38.65)	(163.59)	(9.19)	(172.01)	(24.39)	(228.15)	(8.81)	(257.88)
$\gamma 2$			0	1e-6			0	2e-4
			(0)	(2.8e-5)			(0)	(5.6e-5)
$\eta 1$	-0.05	-0.05	-0.05	-0.05	-0.03	-0.03	-0.04	-0.04
	(1.96)	(1.93)	(1.94)	(1.93)	(1.55)	(1.54)	(1.31)	(1.54)
$\eta 2$	0.10	0.11	0.11	0.11	0.11	0.11	0.11	0.11
	(4.52)	(3.69)	(4.01)	(3.67)	(3.91)	(3.44)	(1.72)	(3.40)
skew	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$l(r, x)$	6586.31	6588.69	6584.86	6588.69	6702.15	6704.27	6700.36	6704.27
Θ	0.85	0.82	0.85	0.84	0.83	0.83	0.83	0.80
AIC	-13.421	-13.424	-13.416	-13.422	-13.657	-13.660	-13.652	-13.658
SBIC	-13.371	-13.369	-13.361	-13.362	-13.608	-13.605	-13.597	-13.598

Note: The realized kernel and realized volatility are the realized measures used in the estimations. The values in parentheses are t values calculated using a robust standard error, and $l(r, x)$ denotes the joint log likelihood.

Table 5 Robust and non-robust standard errors for the close-to-close return with normal distribution

			Realized Kernel				Realized Volatility			
	G(1,1)	EG(1,1)	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)
I_{θ}^{-1}										
α	2.5e-3	0.01								
$\beta 1$	3.3e-3	9e-6	4.4e-3	0.10	4e-3	7.2e-3	6e-3	0.01	6e-3	0.01
$\beta 2$										
$\gamma 1$				0.04		0.04		0.04		0.04
$\gamma 2$										
$\eta 1$		1.1e-3	6e-3	0.03	7.7e-3	0.03	0.1	0.03	0.01	0.03
$\eta 2$										
					0.03	0.03			0.03	0.03
			0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
			4e-3	4.1e-3						
					4.1e-3	4.1e-3	3.6e-3	3.7e-3	3.7e-3	3.7e-3
$I_{\theta}^{-1} J_{\theta} I_{\theta}^{-1}$										
α	0.02	0.04								
$\beta 1$	0.04	2.7e-5								
$\beta 2$			0.01	0.03	0.01	0.02	0.02	0.03	0.02	0.02
$\gamma 1$										
$\gamma 2$				0.07		0.09		0.08		0.09
$\eta 1$		3.4e-3								
$\eta 2$			0.02	0.06	4.9e-3	0.06	0.02	0.08	0.01	0.08
					0.06	0.07			0.06	0.06
			0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
			5.4e-3	5.4e-3	5.5e-3	5.5e-3	5.2e-3	5.2e-3	5.2e-3	5.2e-3

Note: I_{θ}^{-1} and $I_{\theta}^{-1} J_{\theta} I_{\theta}^{-1}$ denote the non-robust and robust standard errors.

Table 6 Robust and non-robust standard errors for the close-to-close return with skewed t distribution

	Realized Kernel				Realized Volatility			
	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)
I_{θ}^{-1}								
$\beta 1$	2e-3	0.02	1.8e-3	0.02	1.9e-3	0.02	1.7e-3	0.02
$\beta 2$		0.04		0.04		0.05		0.05
$\gamma 1$	3.5e-3	0.03	4.5e-3	0.03	3.4e-3	0.03	4.1e-3	0.03
$\gamma 2$			0.03	0.03			0.04	0.03
$\eta 1$	0.04	0.05	0.05	0.05	0.06	0.05	0.05	0.05
$\eta 2$	0.02	0.03	0.03	0.03	0.04	0.03	0.03	0.03
$I_{\theta}^{-1} J_{\theta} I_{\theta}^{-1}$								
$\beta 1$	2.7e-3	0.01	1.8e-3	0.01	2.2e-3	0.02	1.8e-3	0.02
$\beta 2$		0.04		0.04		0.07		0.06
$\gamma 1$	3.4e-3	0.05	2.9e-3	0.04	3.5e-3	0.06	2e-3	0.04
$\gamma 2$			0.04	0.03			0.04	0.04
$\eta 1$	0.09	0.08	0.09	0.08	0.14	0.08	0.08	0.08
$\eta 2$	0.05	0.04	0.05	0.04	0.09	0.03	0.03	0.03

Note: I_{θ}^{-1} and $I_{\theta}^{-1} J_{\theta} I_{\theta}^{-1}$ denote the non-robust and robust standard errors.

Table 7 Robust and non-robust standard errors for the open-to-close return with normal distribution

	Realized Kernel				Realized Volatility					
	G(1,1)	EG(1,1)	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)
I_{θ}^{-1}										

α	4.6e-3	0.02								
β_1	4.5e-3	4.5e-4	3.5e-4	1.3e-3	3.3e-4	1.3e-3	3.3e-4	2.5e-3	1e-3	3e-3
β_2				0.01		0.01		0.02		0.03
γ_1		4e-3	2e-3	0.01	7.5e-4	0.01	2e-3	3.7e-3	2e-3	4e-3
γ_2					0.03	0.02			0.03	0.04
η_1			0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
η_2			0.01	0.01						
					0.01	0.01	0.01	0.01	0.01	0.01
$I_{\theta}^{-1} J_{\theta} I_{\theta}^{-1}$										
α	0.06	0.03								
β_1	0.07	7.6e-4	4e-4	1.3e-3	3e-4	8e-4	5.6e-4	2.9e-3	1e-3	3e-3
β_2				3.7e-3		3.5e-3		0.03		0.03
γ_1		0.01	2e-3	2.8e-3	5.5e-3	3.3e-3	1e-3	3.2e-3	7.5e-4	3.4e-3
γ_2					0.04	0.04			0.04	0.04
η_1			0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
η_2			0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Note: I_{θ}^{-1} and $I_{\theta}^{-1} J_{\theta} I_{\theta}^{-1}$ denote the non-robust and robust standard errors.

Table 8 Robust and non-robust standard errors for the open-to-close return with skewed t distribution

	Realized Kernel				Realized Volatility			
	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)	RG(1,1)	RG(2,1)	RG(1,2)	RG(2,2)
I_{θ}^{-1}								
β_1	5e-4	2e-3	3e-4	2e-3	5e-4	3.5e-3	1e-3	4e-3
β_2		0.01		0.01		0.02		0.02

$\gamma 1$	0.01	0.01	0.04	0.01	0.01	4e-3	0.02	4e-3
$\gamma 2$			0.02	0.03			0.04	0.03
$\eta 1$	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
$\eta 2$	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02
$I_{\theta}^{-1} J_{\theta} I_{\theta}^{-1}$								
$\beta 1$	5e-4	1e-3	3e-4	6.5e-4	4e-4	2e-3	2e-3	2e-3
$\beta 2$		3e-3		2.3e-3		0.01		0.01
$\gamma 1$	0.01	2e-3	0.03	2e-3	0.02	2e-3	0.04	1.5e-3
$\gamma 2$			0.02	0.03			0.08	0.03
$\eta 1$	0.02	0.03	0.03	0.03	0.02	0.02	0.03	0.02
$\eta 2$	0.02	0.03	0.03	0.03	0.03	0.03	0.06	0.03

Note: I_{θ}^{-1} and $I_{\theta}^{-1} J_{\theta} I_{\theta}^{-1}$ denote the non-robust and robust standard errors.

Table 9 MCS evaluation for the QLIKE loss function

	Close-close return				Open-close return		
	QLIKE	p-MCS	Rank		QLIKE	p-MCS	Rank
Volatility proxy: r^2							
G	-16.036	0.23	3	G	-16.500	ELI	10
EG	-15.748	ELI	4	EG	-16.877	1.00	1
RG11normRK	-16.081	1.00	1	RG11normRK	-16.838	0.34	3
RG11normRV	-16.057	0.99	2	RG21normRK	-16.839	0.42	2
RG21skstRK	-15.243	ELI	6	RG11normRV	-16.837	0.18	4
RG21skstRV	-15.262	ELI	5	RG21normRV	-16.836	0.14	5
				RG11skstRK	-16.748	ELI	7
				RG21skstRK	-16.662	ELI	8
				RG11skstRV	-16.782	ELI	6
				RG21skstRV	-16.572	ELI	9
Volatility proxy: Realized Kernel							
G	-15.680	ELI	3	G	-11.998	ELI	10

EG	-15.103	ELI	6	EG	-15.152	ELI	9
RG11normRK	-15.866	1.00	1	RG11normRK	-15.289	ELI	8
RG11normRV	-15.856	ELI	2	RG21normRK	-15.312	ELI	6
RG21skstRK	-15.176	ELI	5	RG11normRV	-15.294	ELI	7
RG21skstRV	-15.197	ELI	4	RG21normRV	-15.316	ELI	5
				RG11skstRK	-15.505	ELI	3
				RG21skstRK	-15.720	1.00	1
				RG11skstRV	-15.496	ELI	4
				RG21skstRV	-15.712	0.244	2
Volatility proxy: Realized Volatility							
G	-15.527	ELI	3	G	-11.176	ELI	10
EG	-14.883	ELI	6	EG	-14.836	ELI	9
RG11normRK	-15.758	1.00	1	RG11normRK	-14.989	ELI	8
RG11normRV	-15.752	ELI	2	RG21normRK	-15.018	ELI	6
RG21skstRK	-15.149	ELI	5	RG11normRV	-14.996	ELI	7
RG21skstRV	-15.169	ELI	4	RG21normRV	-15.024	ELI	5
				RG11skstRK	-15.257	ELI	3
				RG21skstRK	-15.533	0.30	2
				RG11skstRV	-15.240	ELI	4
				RG21skstRV	-15.540	1.00	1

Note: r^2 , realized kernel and realized volatility are the volatility proxies used in constructing QLIKE loss functions. The values of p-MCS are the p-values of the model confidence set procedure. Models with ‘ELI’ in the column of p-MCS were eliminated using the MCS elimination rule under the given 10% confidence level.

Table10 VaR backtesting

	VaR0.5%				VaR1%			
	Ratio	UC	CC	INDDR	Ratio	UC	CC	INDDR
close-close return								
G	0.0064	0.59	0.84	0.64	0.009	0.77	0.9	0.89
EG	0.0179	0	0	0.08	0.0231	0	0.01	0.22
RG11normRK	0.0064	0.59	0.84	0.8	0.0064	0.28	0.54	0.8
RG11normRV	0.0064	0.59	0.84	0.8	0.0064	0.28	0.54	0.8
RG21skstRK	0.0038	0.64	0.88	0.54	0.0051	0.13	0.32	0.67
RG21skstRV	0.0038	0.64	0.88	0.54	0.0051	0.13	0.32	0.67

open-close return								
G	0.0308	0	0	0.08	0.0436	0	0	0.94
EG	0.0051	0.96	0.98	0.09	0.009	0.77	0.9	0.97
RG11normRK	0.0064	0.59	0.84	0.1	0.0115	0.67	0.82	0.24
RG21normRK	0.0064	0.59	0.84	0.15	0.0115	0.67	0.82	0.24
RG11normRV	0.0103	0.07	0.17	0.18	0.0167	0.09	0.1	0.46
RG21normRV	0.0064	0.59	0.84	0.1	0.0141	0.28	0.19	0.73
RG11skstRK	0				0.0051	0.13	0.32	0.54
RG21skstRK	0				0.0038	0.05	0.14	0.98
RG11skstRV	0				0.0051	0.13	0.32	0.54
RG21skstRV	0				0.0038	0.05	0.14	0.98

Note: Ratio, UC, CC, and INDDR represent the value-at-risk exceedance ratio, p-values for unconditional convergence, conditional convergence, and duration-based independence hypothesis, respectively.

Table 11 MCS evaluation for the VaR loss function

Close-close return				Open-close return			
	VaRLoss	p-MCS	Rank		VaRLoss	p-MCS	Rank
VaR 0.5%							
G	5.632	0.34	5	G	3.519	ELI	10
EG	6.642	ELI	6	EG	2.462	0.20	5
RG11normRK	4.980	1.00	1	RG11normRK	2.322	1.00	3
RG11normRV	5.053	1.00	2	RG21normRK	2.357	0.99	4
RG21skstRK	5.574	0.82	3	RG11normRV	2.256	1.00	1
RG21skstRV	5.624	0.37	4	RG21normRV	2.296	1.00	2
				RG11skstRK	2.555	ELI	7
				RG21skstRK	2.904	ELI	8
				RG11skstRV	2.547	ELI	6
				RG21skstRV	3.058	ELI	9
VaR 1%							
G	8.094	0.36	3	G	5.383	ELI	10
EG	9.133	ELI	6	EG	4.175	1.00	5
RG11normRK	7.776	1.00	1	RG11normRK	4.116	1.00	1
RG11normRV	7.870	0.99	2	RG21normRK	4.129	1.00	2
RG21skstRK	8.433	ELI	4	RG11normRV	4.130	1.00	3

RG21skstRV	8.467	ELI	5	RG21normRV	4.147	1.00	4
				RG11skstRK	4.408	ELI	7
				RG21skstRK	4.709	ELI	8
				RG11skstRV	4.302	0.22	6
				RG21skstRV	4.801	ELI	9

Note: VaRLoss and p-MCS represent the values of the asymmetric VaR-based loss function $\times 10^6$ and p-values for MCS procedure, respectively. Models with ‘ELI’ in the column of p-MCS were eliminated using the MCS elimination rule under the given 10% confidence level.

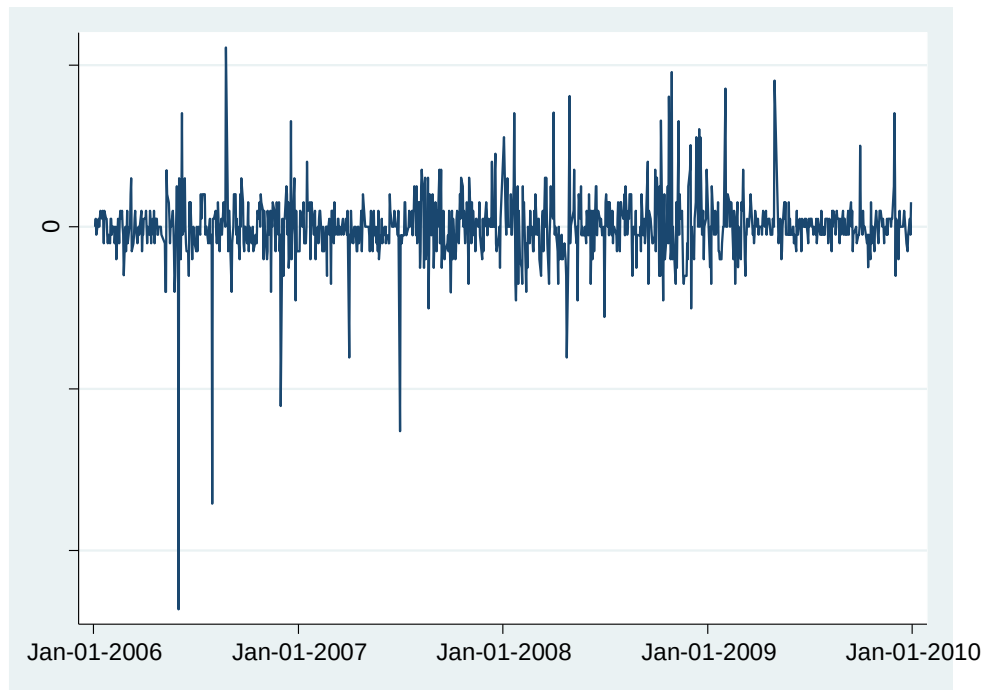
Table 12 Comparison of ARMA-RGARCH and RGARCH

	VaRLoss*10 ⁶		QLIKE			Serial independence		
	0.5%	1%	R2	RK	RV	CVM	Fisher	Tippet
RG	4.98	7.78	-16.0623	-15.8306	-15.7201	0.29	0.41	0.56
	(1)	(1)	(1)	(5e-3)	(0.03)			
ARMA- RG	5.11	7.95	-16.0570	-15.8323	-15.7214	0.36	0.72	0.71
	(7e-4)	(0)	(0.35)	(1)	(1)			

Note: the values in parentheses are p-values for the MCS procedure; CVM, Fisher, Tippet represent the Cramer–Von Mises test, Fisher’s rule, and Tippet’s rule, respectively.

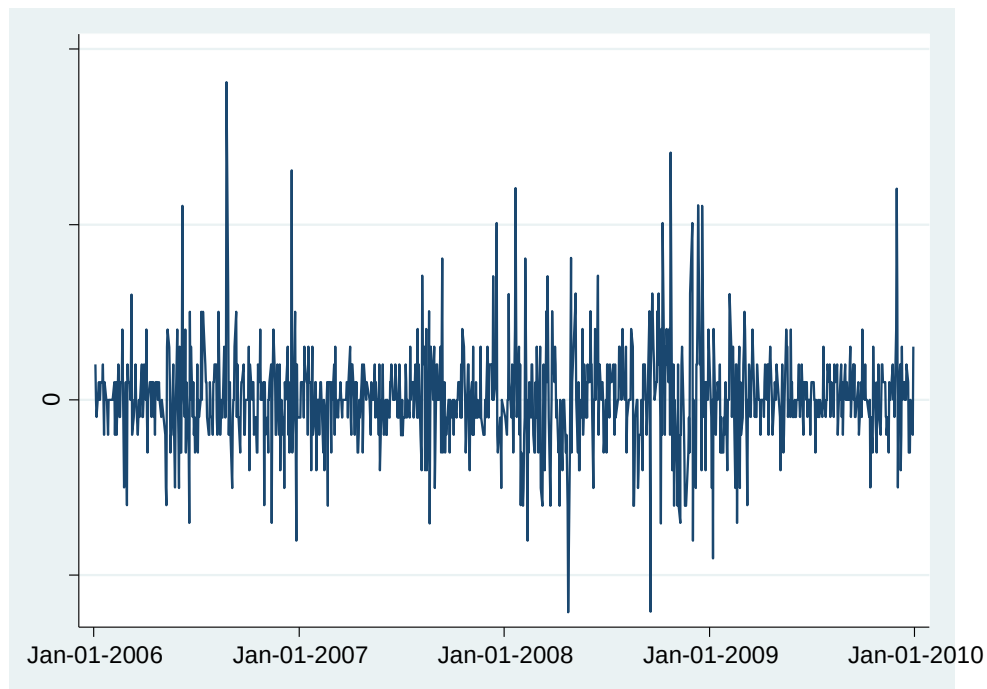
Figure 1 Time series of returns and realized measures

Close-to-close return

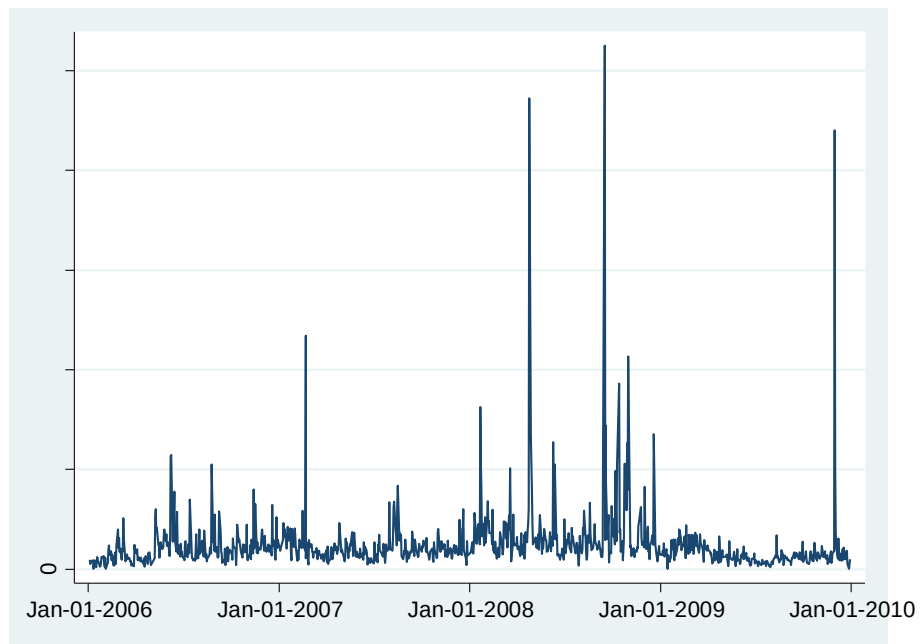


Open-to-close

return



Realized kernel



Realized volatility

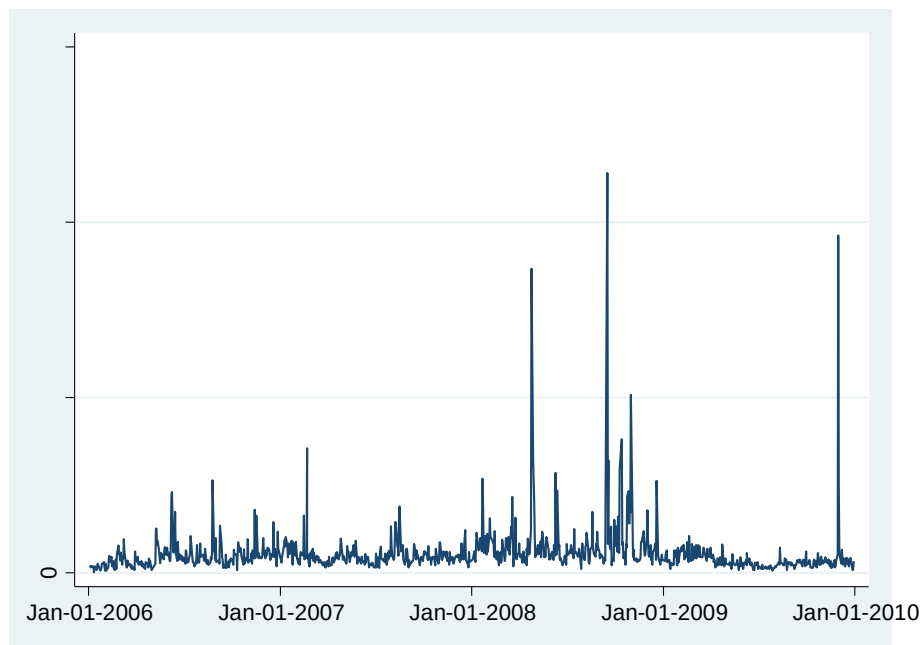


Figure 2 News impact curve

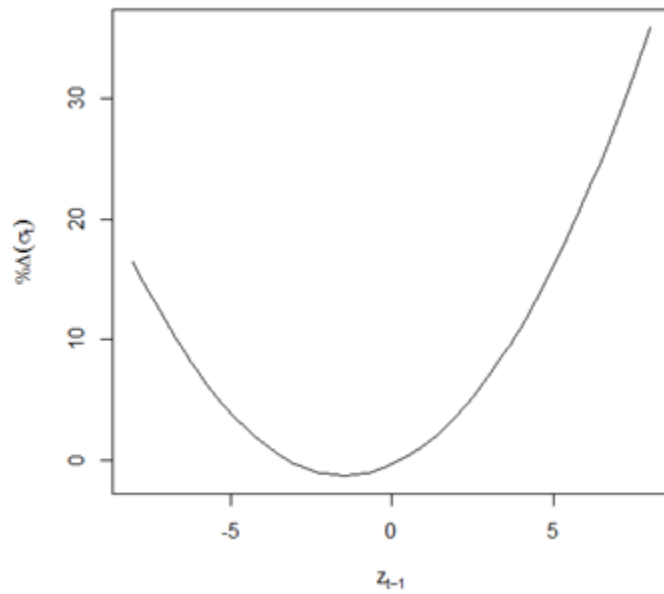
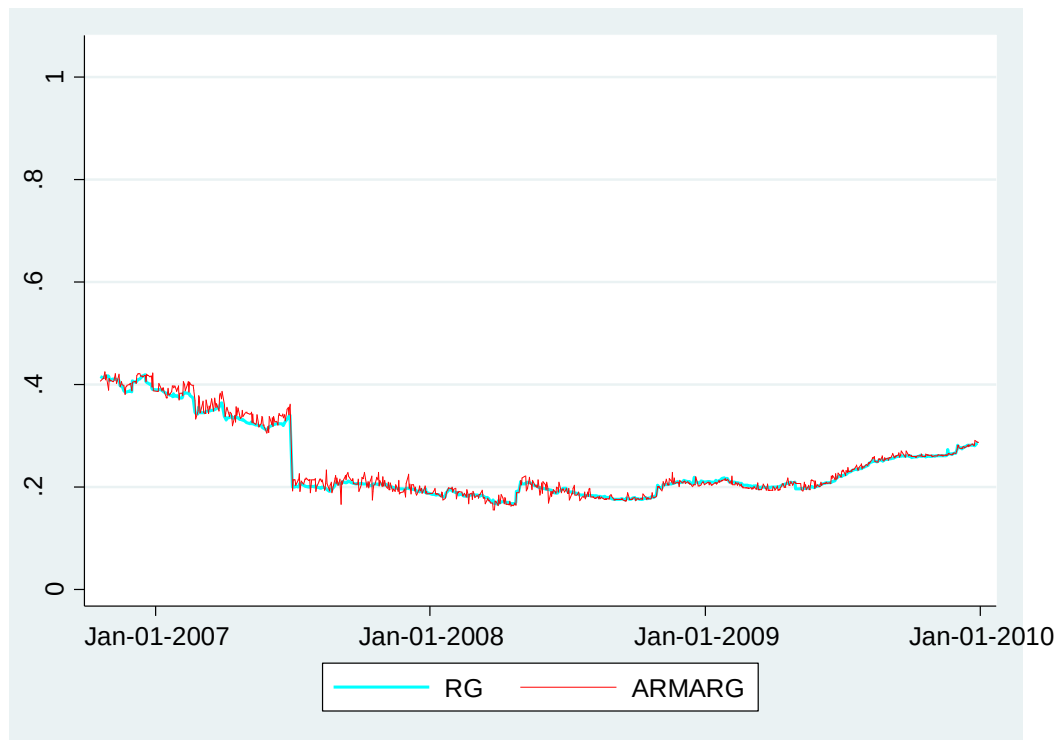
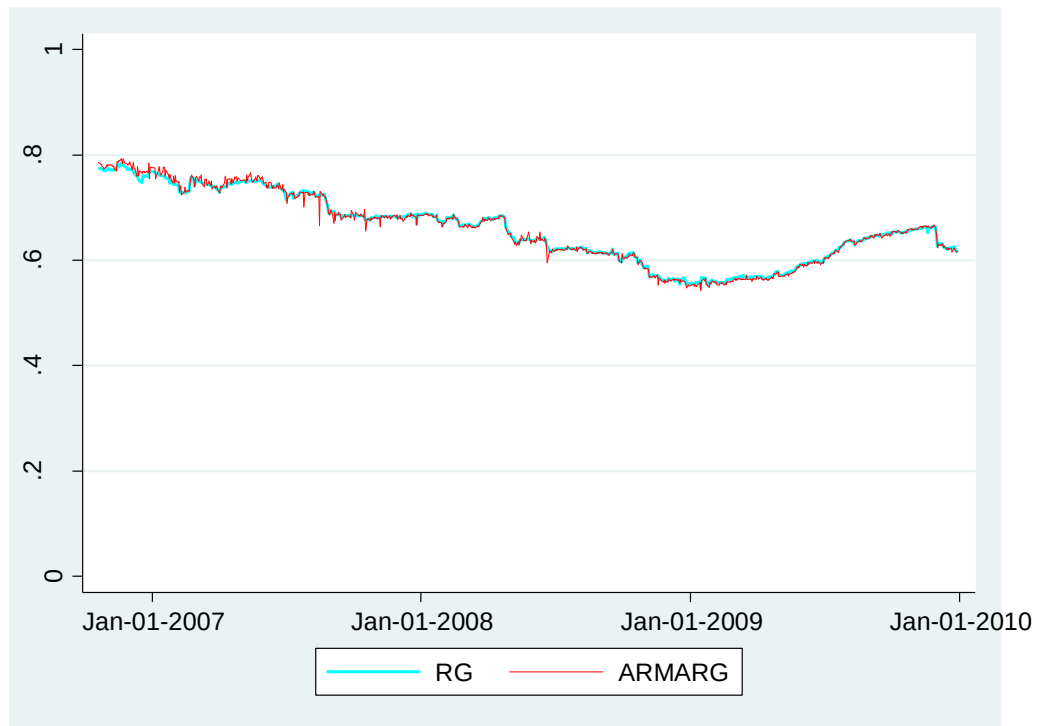


Figure 3 Parameters and conditional standard errors

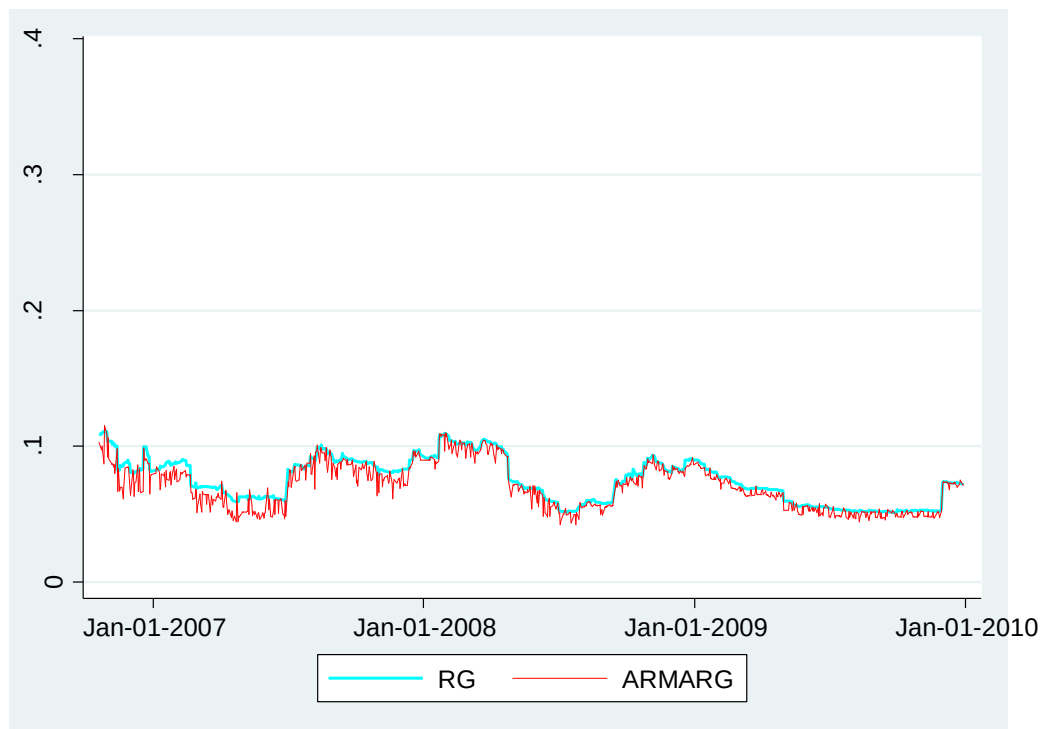
Gamma



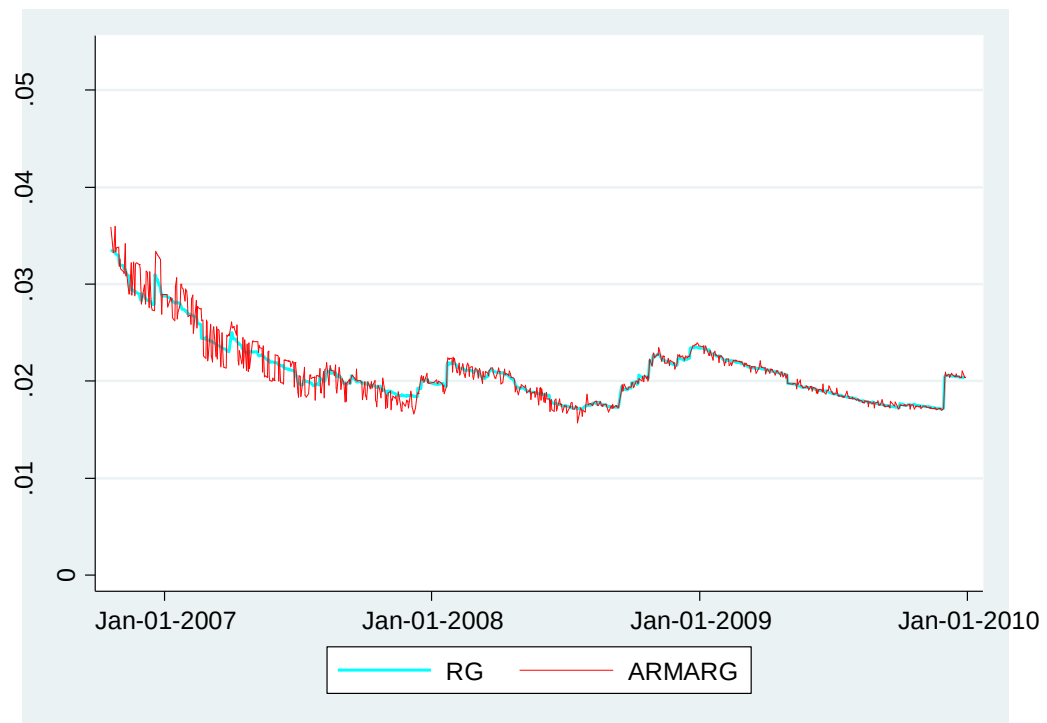
Beta



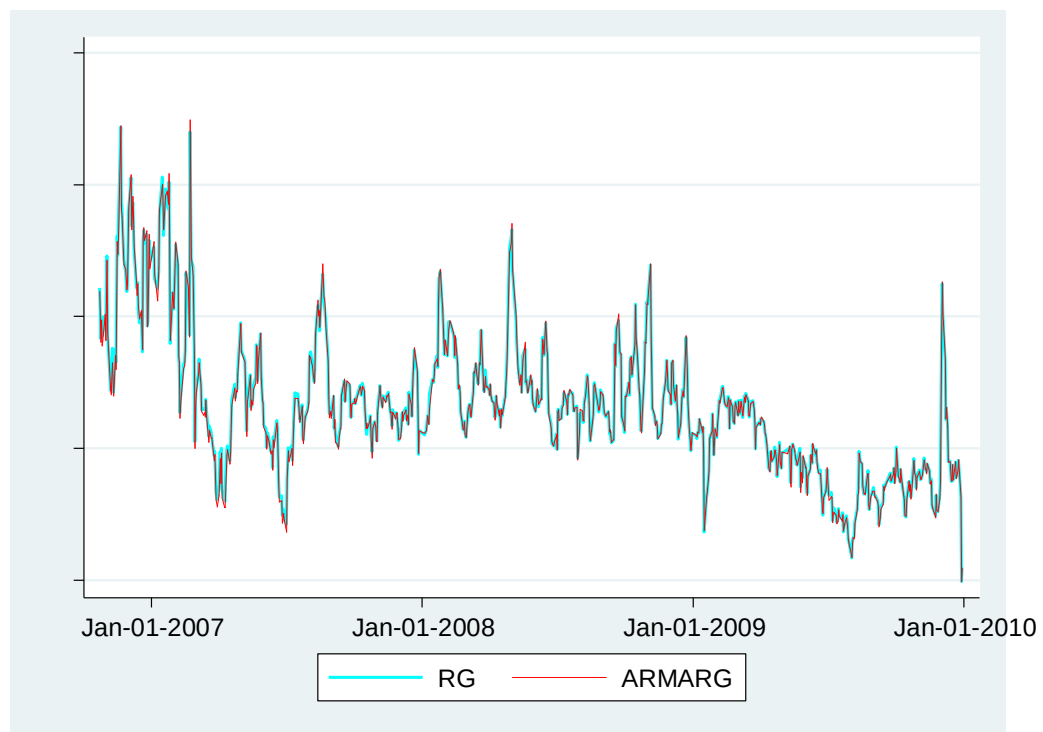
Eta1



Eta2



Conditional standard error



Chapter 3

Realizing moments

3.1. Introduction

Many financial problems such as asset pricing, risk management and portfolio allocation require distributional characteristics of asset returns. The most critical feature of return distribution is its second moment, which has been studied by using various methods over the past three decades. Such approaches include fitting a parametric model (such as GARCH and SV type models) with a distributional assumption, extracting implied volatility from option price using specific option pricing models, and calculating realized volatility which is the sum of intraday squared returns. As discussed in Andersen et al. (2001), parametric volatility models and option pricing models suffer from model misspecification. Regardless of model misspecification, volatility estimates obtained from the parametric volatility models and option pricing models are conditioned on the past information, but financial practitioners prefer indicators which contain information up-to-date. On the other hand, Andersen et al. (2001, 2003) argue that under suitable condition, realized volatility is the unbiased and consistent estimator of quadratic variation. This type of volatility can be considered as the volatility during a certain period, however, volatility at the moment is of more interest.

In addition, investigating higher moments such as skewness and kurtosis attract increasing attention, and such studies (Jondeau and Rickinger (2003), Conrad et al. (2013), Kang and Lee (2016)) find strong relationship between higher moments and returns. These studies also belongs to the framework of conditional, implied and realized moments.

In this paper, we propose a new type of moments which can be estimated from intraday data. The so called “realizing moments” differ from conditional, implied and realized moments, since these moments contain information up-to-date and describes the current return distribution at daily frequency.

Generally, we only have the point observation of return at daily, weekly, monthly and longer frequencies, but we cannot observe the probability density of return. Finance literatures consider that all information would be absorbed and reflected by market price. However, it is hardly to consider that a point observation of return can describe these information well. Financial returns

are rather noisy, especially at daily frequency. It is more acceptable to consider that information would be well described by the probability density of return. In addition, the characteristics of a return distribution can be described well by its moments. Therefore, we extract the moments of the return distribution and use these moment to study some problems widely concerned in finance. In this paper, we investigate the joint behavior of return and its moments, the predictability of return, and information transmission mechanism using realizing moments, and provide many insightful results. This new type of moments can be used for many purpose in further finance research.

This paper is organized as follows: Section 2 describe the estimation method of realizing moments. Section 3 briefly describes our data. Section 4 provides insightful results of joint behavior between return and realizing moments. Section 5 investigates the predictability of return using realizing moments. Section 6 studies the information transmission mechanism using this new type of moments based on the framework of spillover index proposed by Diebold and Yilmaz (2012). Section 7 concludes.

3.2. Estimation method of Realizing moments

To estimate the moments of daily return from intraday data, we build our framework on the Hallam and Olmo (2014) to first approximate the probability density of current return at daily frequency. The estimation method relies on the theory of self-affine process.

3.2.1 Self-Affine process

A self-affine process performs distributional scaling behavior, which suggests that the distribution of the process at different time scales are identical after an appropriate transformation.

Definition 1: A stochastic process $\{X(t)\}$ satisfies

$$\{X(ct_1), \dots, X(ct_k)\} \stackrel{d}{=} \{c^H X(t_1), \dots, c^H X(t_k)\} \quad (3.1)$$

for some $H > 0$ and $c, k, t_1, \dots, t_k \geq 0$, is called self-affine. H is the self-affinity index and describes the relationship between distributions of $\{X(t)\}$ at different time scales.

3.2.2 Estimation of Realizing moments

Many empirical studies have confirmed the existence of distributional scaling behavior in a wide range of assets (see Calvet and Fisher (2002), Matia et al. (2003), Calvet and Fisher (2004), Di Matteo et al. (2005), Di Matteo (2007), Onali and Goddard (2009)). To approximate the probability density of return using intraday data and estimate the corresponding realizing moments, we assume

Assumption 1: The stochastic logarithmic price process $\{X(t)\}$ is self-affine and has stationary increments $\Delta X(t) = X(t + \Delta) - X(t)$, where $\Delta X(t)$ is the return process.

Let r_D denotes daily return and r_i denotes intraday return. Under the assumption of self-affinity, the return process satisfies the distribution scaling behavior and we have the relationship

$$f(r_D) = f(c^H r_i) \quad (3.2)$$

Equation 2.2 says the probability density of daily return r_D is identical to the probability density of intraday return r_i rescaled by the factor c^H , where c is called the prefactor and equals to the relative length of two sampling intervals. For example, for a market with 9 hours trading, if we collect 5 minute intraday return, then c would be 108. H is the self-affinity index and can be estimated using various. Here, we apply the Detrended Moving Average (DMA) method.

The DMA estimator can be obtained in following way. For a discrete time series $x(t)$, $t = 1, \dots, T$, select a range of window sizes n , $n_{\min} \leq n \leq n_{\max}$, and filter the original series $x(t)$ using a standard moving average with each n .

$$x_n^{MA}(t) = \frac{1}{n} \sum_{k=0}^{n-1} x(t-k) \quad (3.3)$$

For each MA filtered series $\{x_n^{MA}(t)\}$, we calculate the value of σ_n^2 , where

$$\sigma_n^2 = \frac{1}{T-n} \sum_{i=n}^T [x(i) - x_n^{MA}(i)]^2 \quad (3.4)$$

Under the assumption of self-affinity, we have the relationship $\sigma_n \propto n^H$, and we can obtain the estimates of the self-affinity index H by running a linear regression of the logarithm of σ_n^2 on the logarithm of n ,

$$\log \sigma_n = \alpha + H \log n \quad (3.5)$$

Then we can rescale intraday returns by $c^{\hat{H}}$. The distribution of these rescaled intraday returns is identical to the daily return distribution. Consequently, we can obtain the moments of daily return by calculating sample moments of rescaled intraday returns.

$$E[r_D] = \frac{1}{T} \sum_{i=1}^T r_R$$

$$V[r_D] = \frac{1}{T-1} \sum_{i=1}^T (r_R - \bar{r}_R)^2$$

$$Skew[r_D] = \frac{\frac{1}{T} \sum_{i=1}^T (r_R - \bar{r}_R)^3}{V[r_D]^{3/2}}$$

$$Kurt[r_D] = \frac{\frac{1}{T} \sum_{i=1}^T (r_R - \bar{r}_R)^4}{[\frac{1}{T} \sum_{i=1}^T (r_R - \bar{r}_R)^2]^2}$$

3.3. Data

Our data consists of 7 frequently traded commodities in Tokyo commodity exchange, including gold, silver, platinum, gasoline, crude oil, kerosene and rubber. Since derivative contracts have a finite life, we construct our time series using the rolling-over method and “volume” criteria, as discussed in Carchano and Pardo (2009). The sample period is from Jan 2005 to Dec 2011, including 1707 observations. The constructed price series are represented in figure 1.

3.4. Joint behavior between return and realizing moments

Figure 2 represent the time series of returns and estimates of realizing moments in each market. It is obvious that return, (realizing) mean and skewness fluctuate around zero, and return has much larger fluctuation than its mean.

We also study the joint behavior of returns and their moments. Figure 3 shows the joint histogram between return and realizing moments in crude oil market. For other markets, these figures of joint histogram are provided in appendix and they suggest similar results.

Figure 3a represents the relationship between return and mean. If market is perfectly efficient, return should be coincide with its mean, but this is not the case in crude oil (and other commodity) market, suggesting daily return is rather noisy. In general, return is positively dependent with the mean, and has a relative larger probability to coincide with the mean in the tail and center part. That is to say, even under the condition of imperfectly efficient market, return would coincide with the mean of its probability density in two conditions: (1) an extreme event occurred and (2) nothing happened.

From figure 3b, we can find the return is almost independent with its variance, suggesting that market participants are risk neutral in crude oil market. Figure 3c shows that return are positively correlated with skewness, with explicit tail dependence. This is very intuitive: if the return distribution is right skewed, a positive return would be archived and vice versa; on the other hand, an extreme large/small return would happen if its distribution is extremely right/left skewed. Figure 3d indicates that the cluster level of probability density of return rise when the absolute return is large, and the kurtosis is relative low when absolute return is small.

Figure 3e investigates the relationship between mean and variance of daily return density. We can see a clear V shape joint histogram. Generally, the variance and absolute mean of return density are almost positively correlated. The larger the absolute mean, the larger the price variation. While the absolute mean is small (around 0.5), the variance is always small. That is to say, when there is nothing large happens in the market, the price fluctuate less. If the mean (and corresponding return) is extremely large or small, the price variation is also very large. Figure 3f represent the relationship of realizing mean and skewness. Basically, we can see that mean and skewness are positively dependent. From figure 3g, 3h and 3i, we can say that mean is independent of kurtosis, and variance is independent of skewness and kurtosis.

Finally, figure 3j suggest that there is a V shape relationship between skewness and kurtosis. If the return distribution is highly skewed (which means the absolute return is extremely large), it is also highly clustered. If the return distribution is almost not skewed, it is also not clustered.

3.5. Studying return predictability using realizing moments

The predictability of asset return is a key issue in financial economics. In this paper, we focus on investigating whether daily return is predictable based on its past information.

All information should be described by the probability density of return. Therefore, we study the return predictability by investigating whether return is dependent of its past realizing moments. In this paper, we apply the independence test using the empirical copula method proposed by Genest and Remillard (2004). The null hypothesis is that two series are independent. We test the lags up to 50.

Figure 4 represents p-values of the independence test of return and its moments in each market. Generally, the null hypothesis of independence is rarely rejected in the long term. But in short term, sometimes the null hypothesis is rejected. These results suggest that past information of return would help to predict only in short term (generally 2 or 3 days, and no more than one week).

3.6. Studying information transmission mechanism using realizing moments.

Studying cross market information transmission mechanism is a typical issue in financial economics. As the realizing moments describe information contained in return density, we can use this new type of moments to study cross market information transmission mechanism. We build our framework on the spillover index proposed by Diebold and Yilmaz (2012), and study volatility, skewness and kurtosis spillover cross gold, silver, platinum, gasoline, crude oil, kerosene and rubber markets.

Table 1, 2 and 3 represents the total volatility, skewness and kurtosis spillover index of 1-, 3-, and 10- steps. It is obvious that volatility shocks almost transmit nothing to other markets, however, a considerable proportion of information driven by skewness and kurtosis transmit to other markets and help to forecast the skewness and kurtosis in other markets.

In addition, we can observe that the gold, silver and platinum markets transmit much more information to each other than other markets. Same phenomenon happens in gasoline, crude oil and kerosene markets. Therefore, we can conclude that information mostly transmit between markets belong to same class.

Figure 5, 6 and 7 represent the cross market dynamic total volatility, skewness and kurtosis spillover. It is clear that during the global financial crisis and European debt crisis, there is a significant increase in transmission of volatility shock. Moreover, effect of volatility shock spillover generally reach the peak in 3 to 5 days, and decay fast. For skewness shock and kurtosis shock, the spillover level reach the peak level immediately, and almost do not decay in 10 days.

3.7. Conclusion

We propose a new type of moments, which can describe the current probability density of daily return. The so called “realizing moments” contain information up-to-date. This new type of moments can be estimated using high-frequency intraday data by assuming the self-affinity of return process.

We also show various use of realizing moments, including studying the joint behavior of return and its moments, predictability of return, and cross market information transmission mechanism. Many insightful results are provided. Future research can investigate many financial problems using this new type of moments.

Tabel 1 Volatility spillover index

	Gold	Silver	Platinum	Gasoline	CrudeOil	Kerosene	Rubber	Directional from others
1-step ahead								
Gold	99.7	0	0	0.3	0	0	0	0
Silver	0	96.5	0	0.6	0	2.9	0	4
Platinum	0	0.1	99	0	0.8	0	0	1
Gasoline	0.3	0.6	0	99.1	0	0	0	1
CrudeOil	0	0	0.8	0	99	0.1	0	1
Kerosene	0	2.9	0	0	0.1	97	0	3
Rubber	0	0	0	0	0	0	99.9	0
Directional to others	0	4	1	1	1	3	0	10
Directional including own	100	100	100	100	100	100	100	1.4%
3-step ahead								
Gold	99.5	0.2	0	0.3	0	0	0	1
Silver	0	95.1	1.3	0.6	0.1	2.9	0	5
Platinum	0	0.2	98.7	0.1	0.9	0	0	1
Gasoline	0.3	0.6	0.1	99	0	0	0	1
CrudeOil	0	0.1	0.8	0	99	0.1	0	1
Kerosene	0	2.9	0.1	0	0.1	96.9	0	3
Rubber	0	0	0.1	0	0	0	99.8	0
Directional to others	0	4	2	1	1	3	0	12
Directional including own	100	99	100	100	100	100	100	1.7%
10-step ahead								
Gold	99.5	0.2	0	0.3	0	0	0	1
Silver	0	95.1	1.3	0.6	0.1	2.9	0	5
Platinum	0	0.2	98.7	0.1	0.9	0	0	1
Gasoline	0.3	0.6	0.1	99	0	0	0	1
CrudeOil	0	0.1	0.8	0	99	0.1	0	1
Kerosene	0	2.9	0.1	0	0.1	96.9	0	3
Rubber	0	0	0.1	0	0	0	99.8	0

Directional others	to	0	4	2	1	1	3	0	12
Directional including own		100	99	100	100	100	100	100	1.7%

Tabel 2 Skewness spillover index

		Gold	Silver	Platinum	Gasoline	CrudeOil	Kerosene	Rubber	Directional from others
1-step ahead									
Gold		39.3	20.2	23.3	2	2.8	2	10.5	61
Silver		22.2	43.2	20.2	1.7	2.3	1.7	8.7	57
Platinum		22.3	17.6	37.6	3.1	3.5	2.3	13.7	62
Gasoline		2.5	1.9	3.9	48	21.5	16.7	5.4	52
CrudeOil		3.1	2.3	4.2	19.9	44.3	16.2	10.1	56
Kerosene		2.6	2.1	3.1	17.7	18.6	50.8	5.1	49
Rubber		11.7	8.9	16	5	10	4.4	44	56
Directional others	to	64	53	71	49	59	43	53	393
Directional including own		104	96	108	97	103	94	97	56.1%
3-step ahead									
Gold		39.1	20.1	23.2	2.1	2.8	2.1	10.6	61
Silver		22.2	43.2	20.2	1.7	2.3	1.8	8.7	57
Platinum		22.3	17.6	37.5	3.1	3.5	2.3	13.6	62
Gasoline		2.5	2.2	4	47.7	21.6	16.6	5.6	52
CrudeOil		3.1	2.4	4.2	20	44.1	16.3	10	56
Kerosene		2.6	2	3	17.4	18.6	51	5.3	49
Rubber		11.7	8.9	15.9	5	10.2	4.4	43.8	56
Directional others	to	103	96	108	97	103	95	98	394
Directional including own		103	96	108	97	103	95	98	56.2%
10-step ahead									

Gold		39.1	20.1	23.2	2.1	2.8	2.1	10.6	61
Silver		22.2	43.2	20.2	1.7	2.3	1.8	8.7	57
Platinum		22.3	17.6	37.5	3.1	3.5	2.3	13.6	62
Gasoline		2.5	2.2	4	47.7	21.6	16.6	5.6	52
CrudeOil		3.1	2.4	4.2	20	44.1	16.3	10	56
Kerosene		2.6	2	3	17.4	18.6	51	5.3	49
Rubber		11.7	8.9	15.9	5	10.2	4.4	43.8	56
Directional	to	103	96	108	97	103	95	98	394
others									
Directional		103	96	108	97	103	95	98	56.2%
including own									

Table 3 Kurtosis spillover index

		Gold	Silver	Platinum	Gasoline	CrudeOil	Kerosene	Rubber	Directional from others
1-step ahead									
Gold		53.5	15	22	0.4	1.1	0.2	7.9	46
Silver		16.7	59.7	13.8	1.3	1	0.5	6.9	40
Platinum		20.9	11.8	50.9	1.1	1.4	0.7	13.3	49
Gasoline		0.3	1	1	44.6	20.4	29.1	3.5	55
CrudeOil		1	0.9	1.4	23.5	51.2	19.2	2.8	49
Kerosene		0.2	0.4	0.7	30.9	17.8	47.3	2.8	53
Rubber		8.6	6.7	15.2	4.6	3.2	3.4	58.3	42
Directional	to	48	36	54	62	45	53	37	335
others									
Directional		101	95	105	106	96	101	95	47.8%
including own									
3-step ahead									
Gold		53.1	15	22.3	0.4	1.1	0.3	7.9	47
Silver		16.6	59.2	14	1.6	1.1	0.6	6.9	41
Platinum		20.7	11.8	50.4	1.5	1.6	0.9	13.2	50
Gasoline		0.3	1	1	44.6	20.4	29.1	3.5	55
CrudeOil		1.1	1	1.3	24.2	49.9	19.8	2.7	50

Kerosene		0.3	0.4	0.7	30.8	17.7	47.4	2.8	53
Rubber		8.9	6.8	15.2	4.8	3.6	3.5	57.2	43
Directional	to	48	36	55	63	45	54	37	338
others									
Directional		101	95	105	108	95	102	94	48.3%
including own									
10-step ahead									
Gold		53.1	15	22.3	0.4	1.1	0.3	7.9	47
Silver		16.6	59.2	14	1.6	1.1	0.6	6.9	41
Platinum		20.7	11.8	50.4	1.5	1.6	0.9	13.2	50
Gasoline		0.3	1	1	44.6	20.4	29.1	3.5	55
CrudeOil		1.1	1	1.3	24.2	49.9	19.8	2.7	50
Kerosene		0.3	0.4	0.7	30.8	17.7	47.4	2.8	53
Rubber		8.9	6.8	15.2	4.8	3.6	3.5	57.2	43
Directional	to	48	36	55	63	45	54	37	338
others									
Directional		101	95	105	108	95	102	94	48.3%
including own									

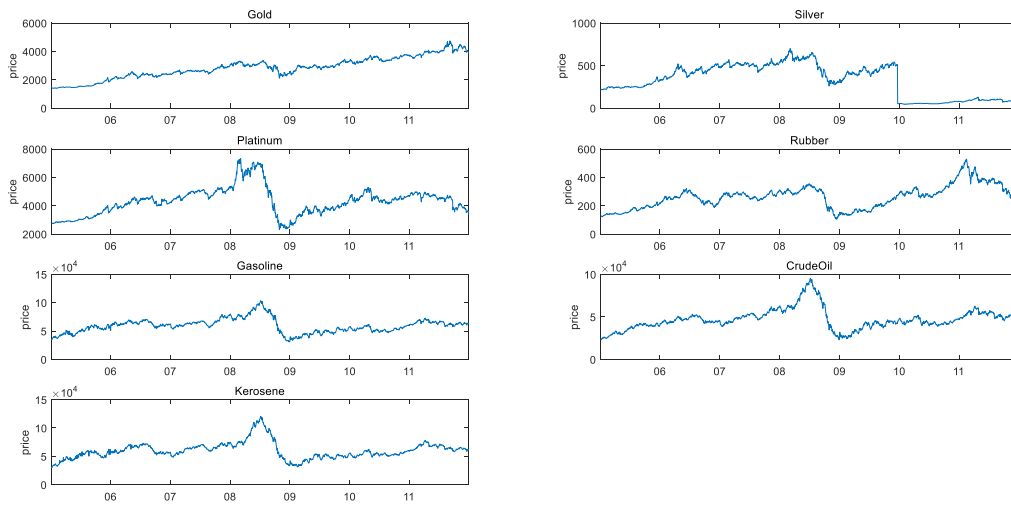
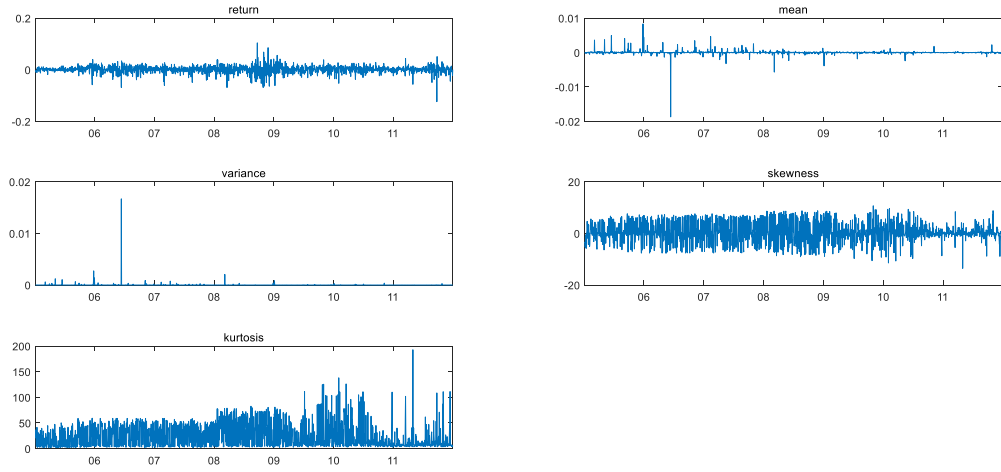
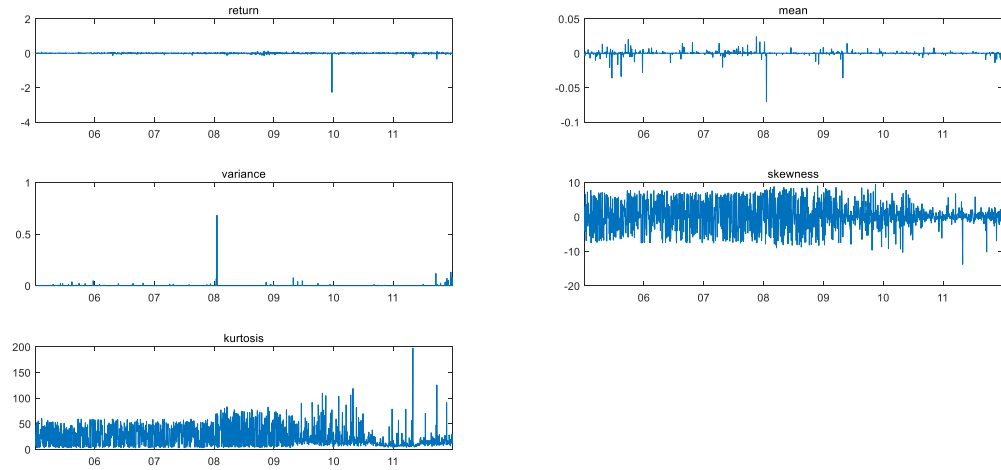


Figure 1 Price series

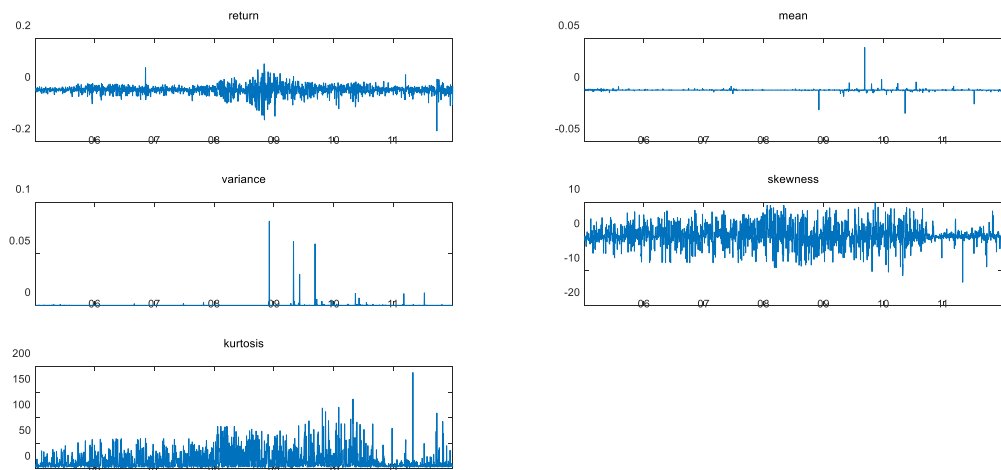
Gold



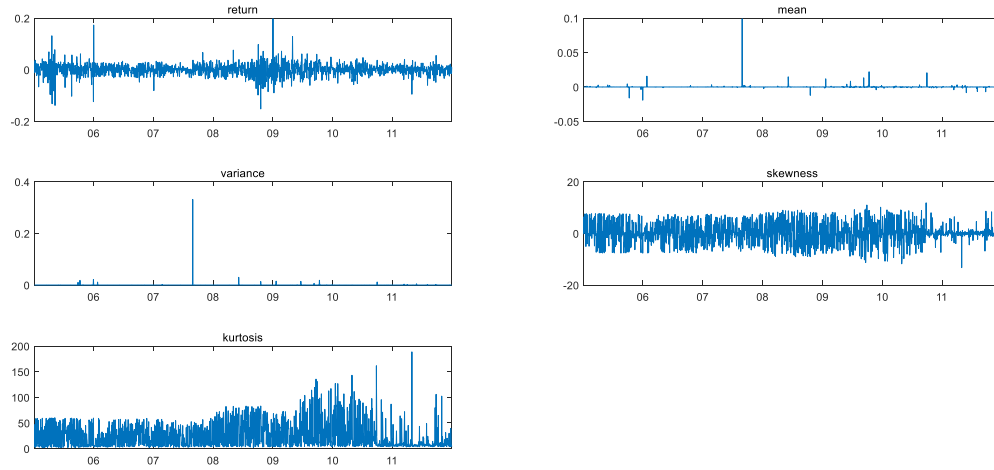
Silver



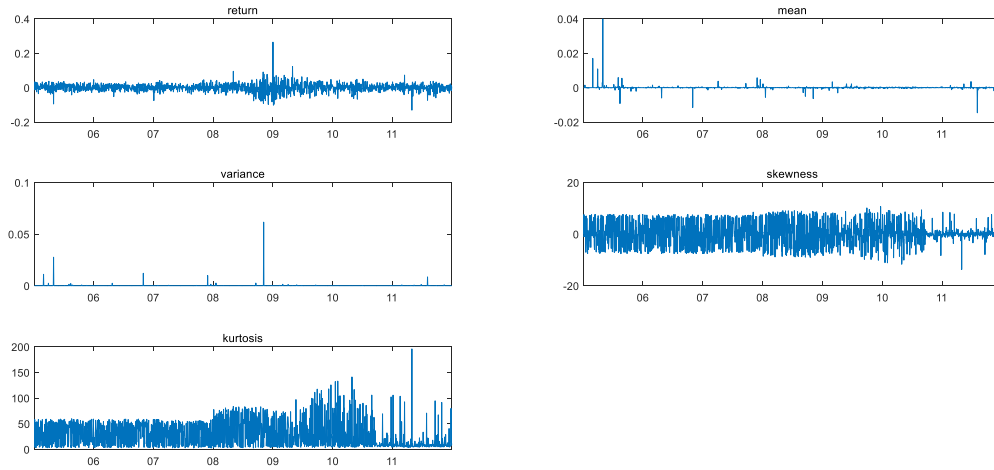
Platinum



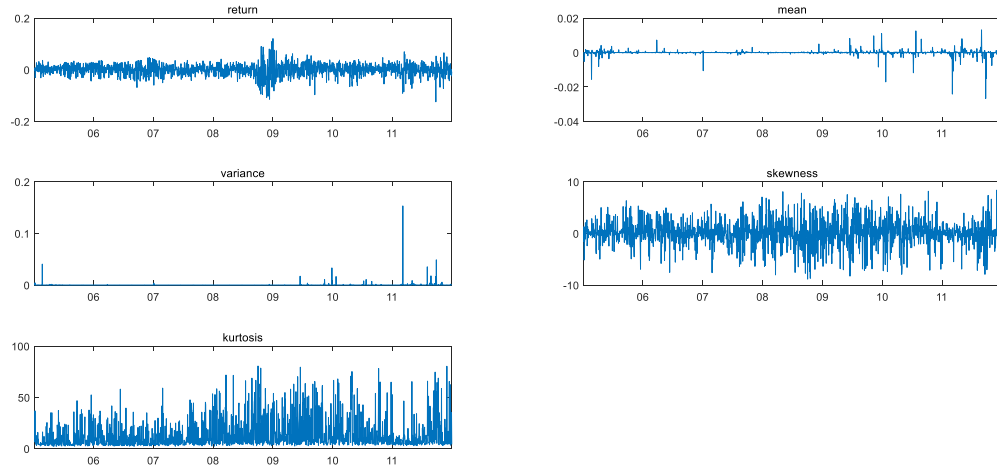
Gasoline



CrudeOil



Rubber



Kerosene

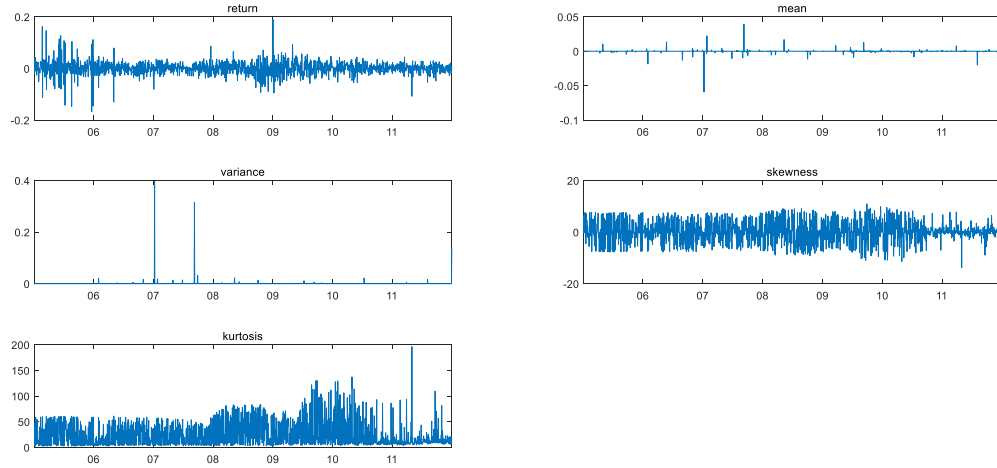
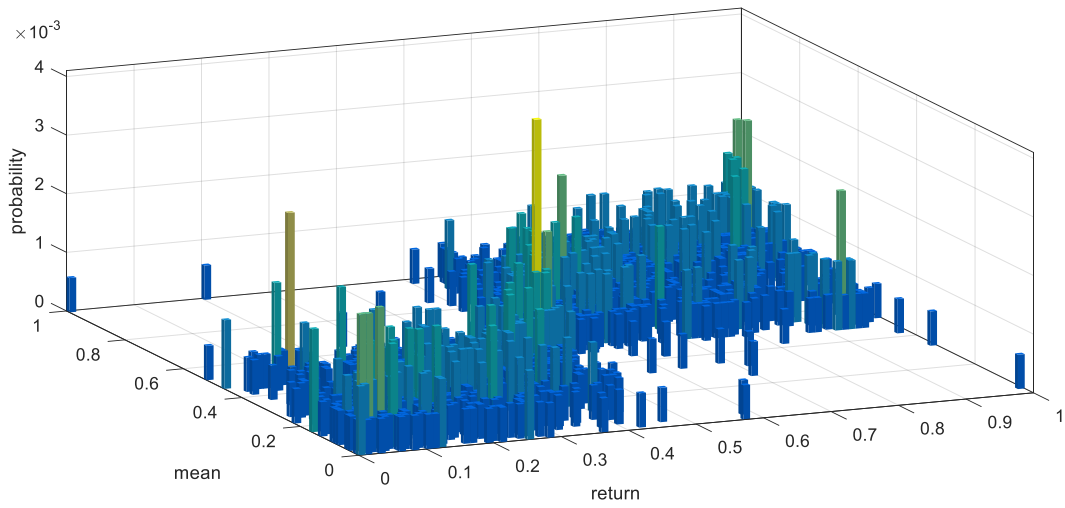
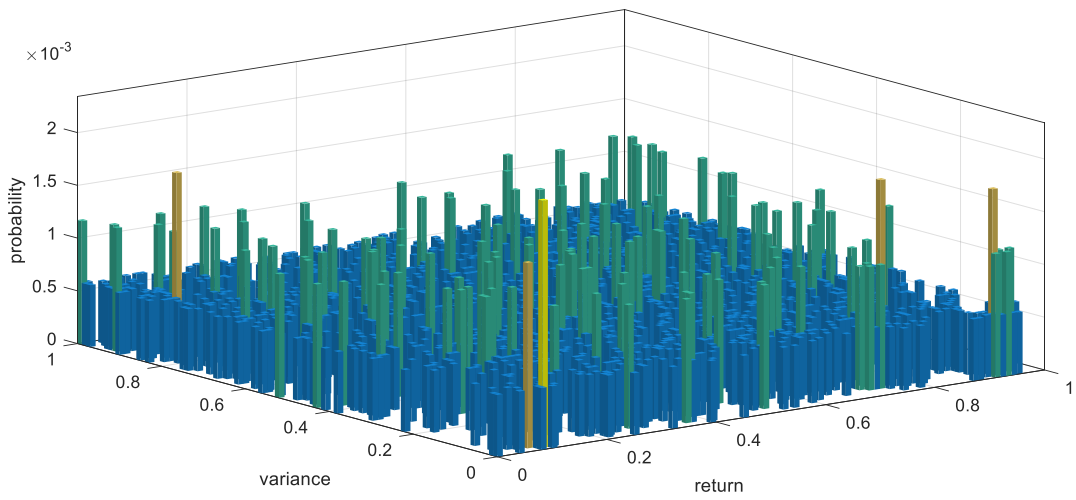


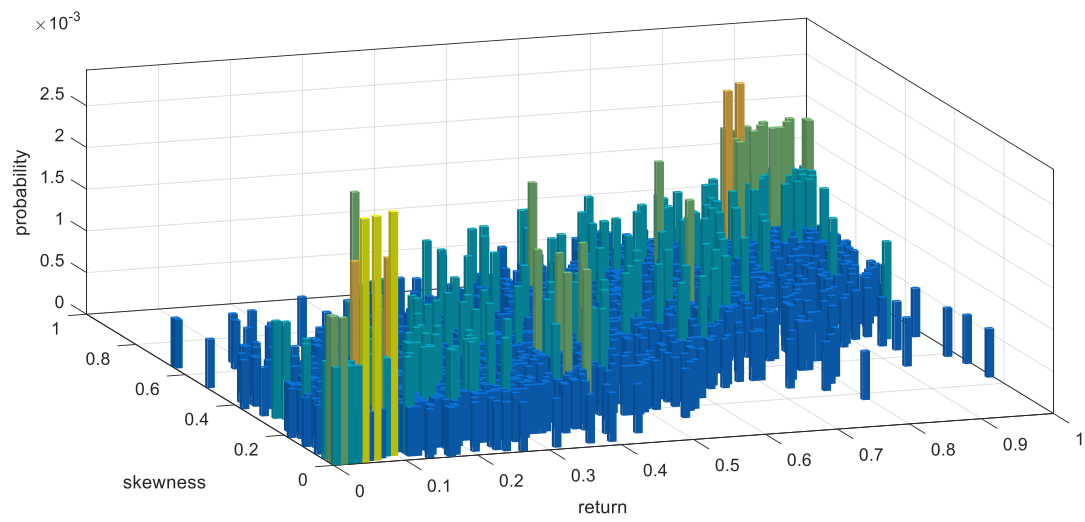
Figure 2 return and estimates of realizing moments



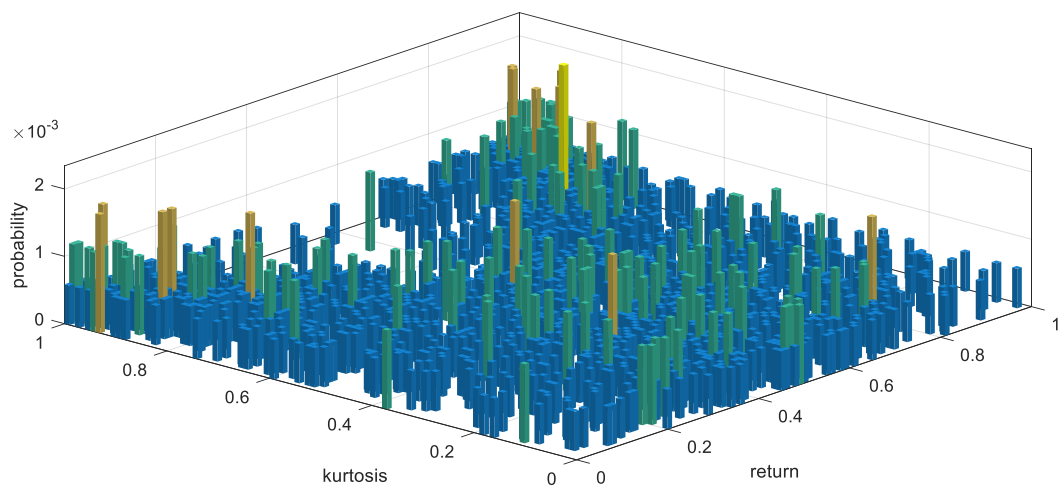
(a) Joint histogram of return and its mean



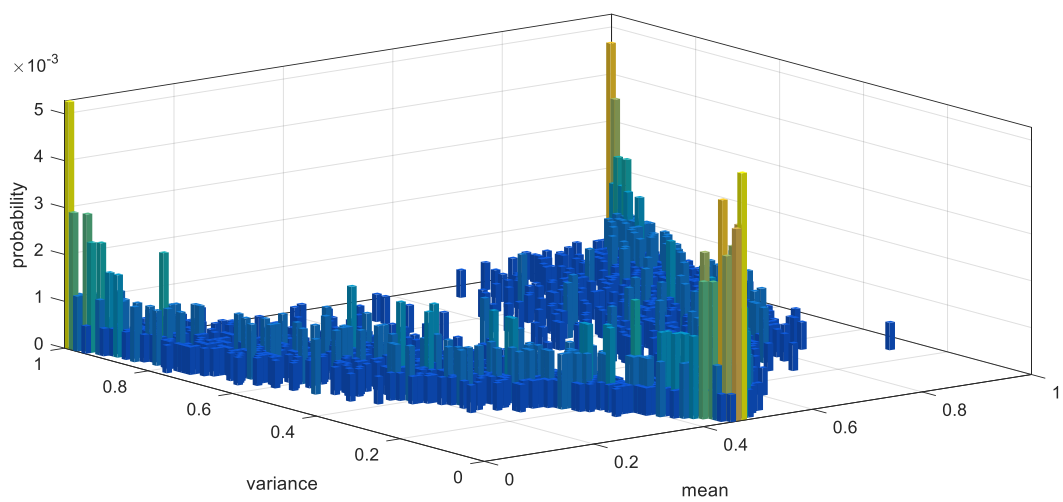
(b) Joint histogram of return and its variance



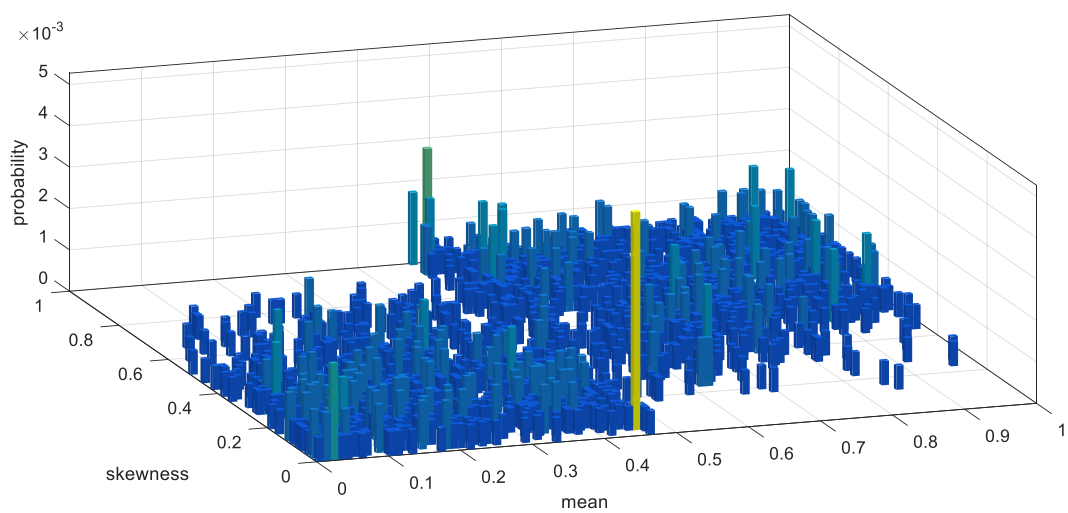
(c) Joint histogram of return and skewness



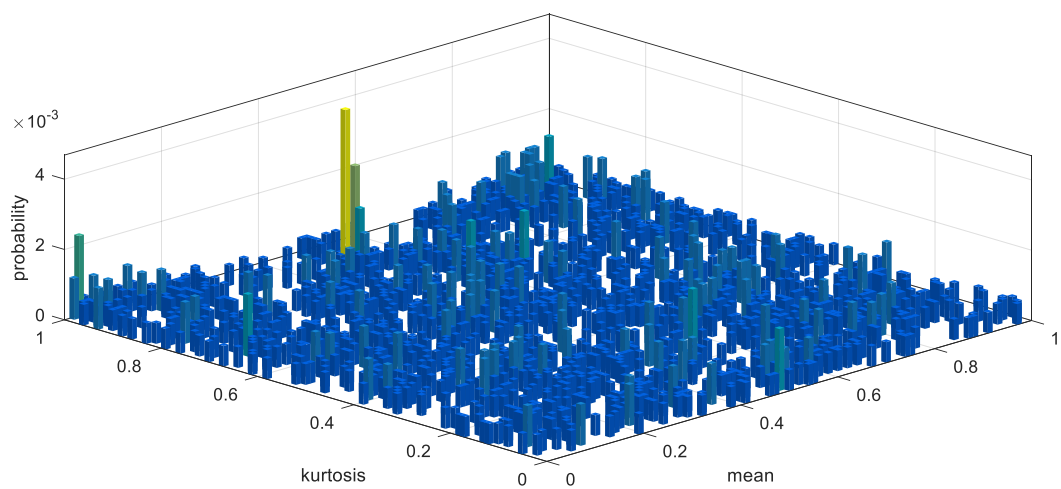
(d) Joint histogram of return and kurtosis



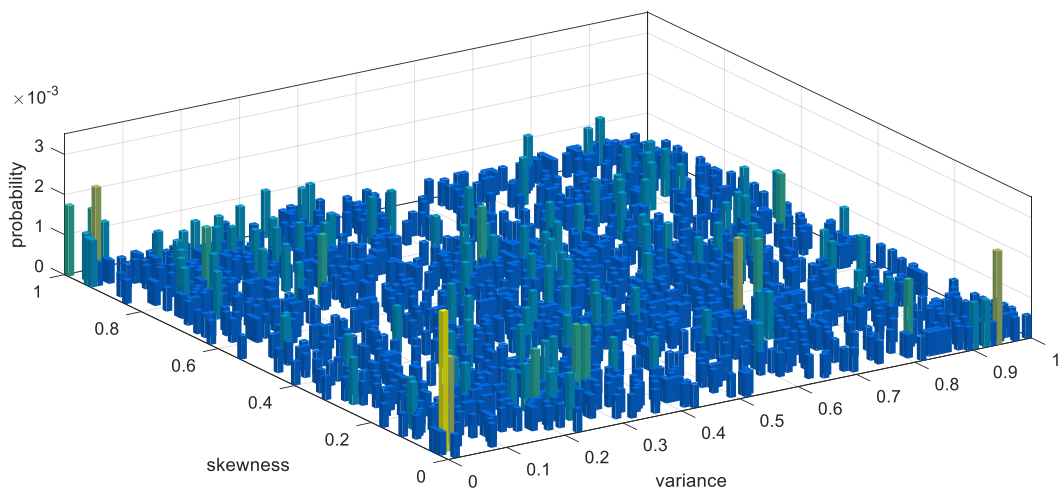
(e) Joint histogram of mean and variance



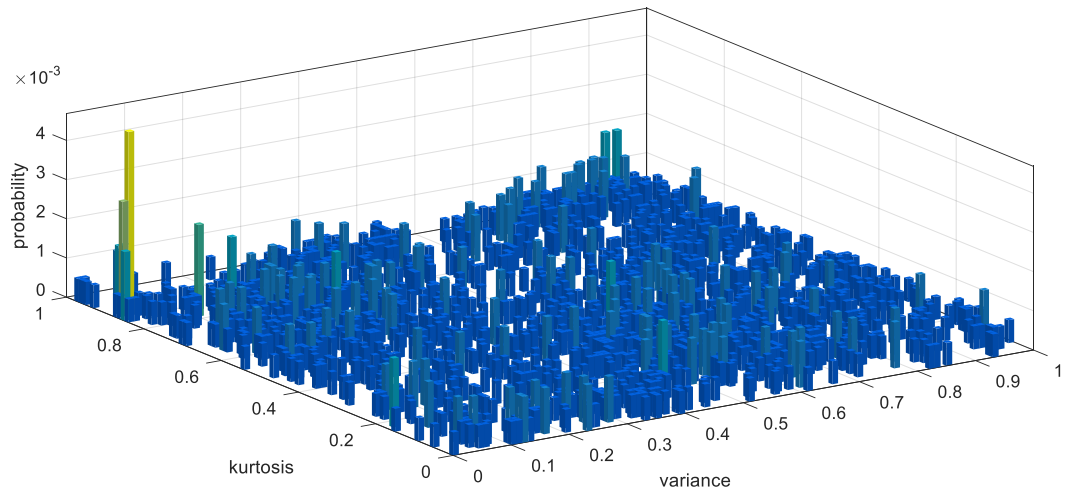
(f) Joint histogram of mean and skewness



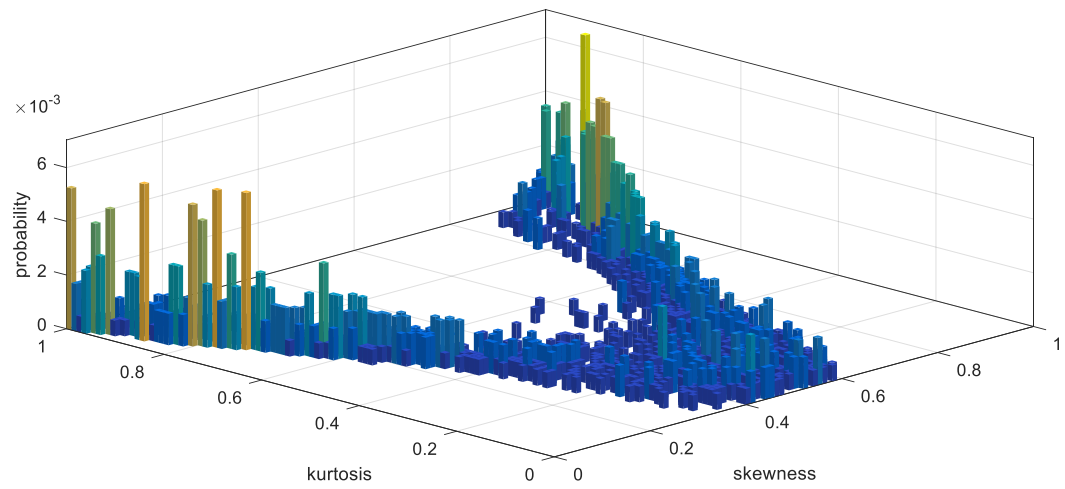
(g) Joint histogram of mean and kurtosis



(h) Joint histogram of variance and skewness

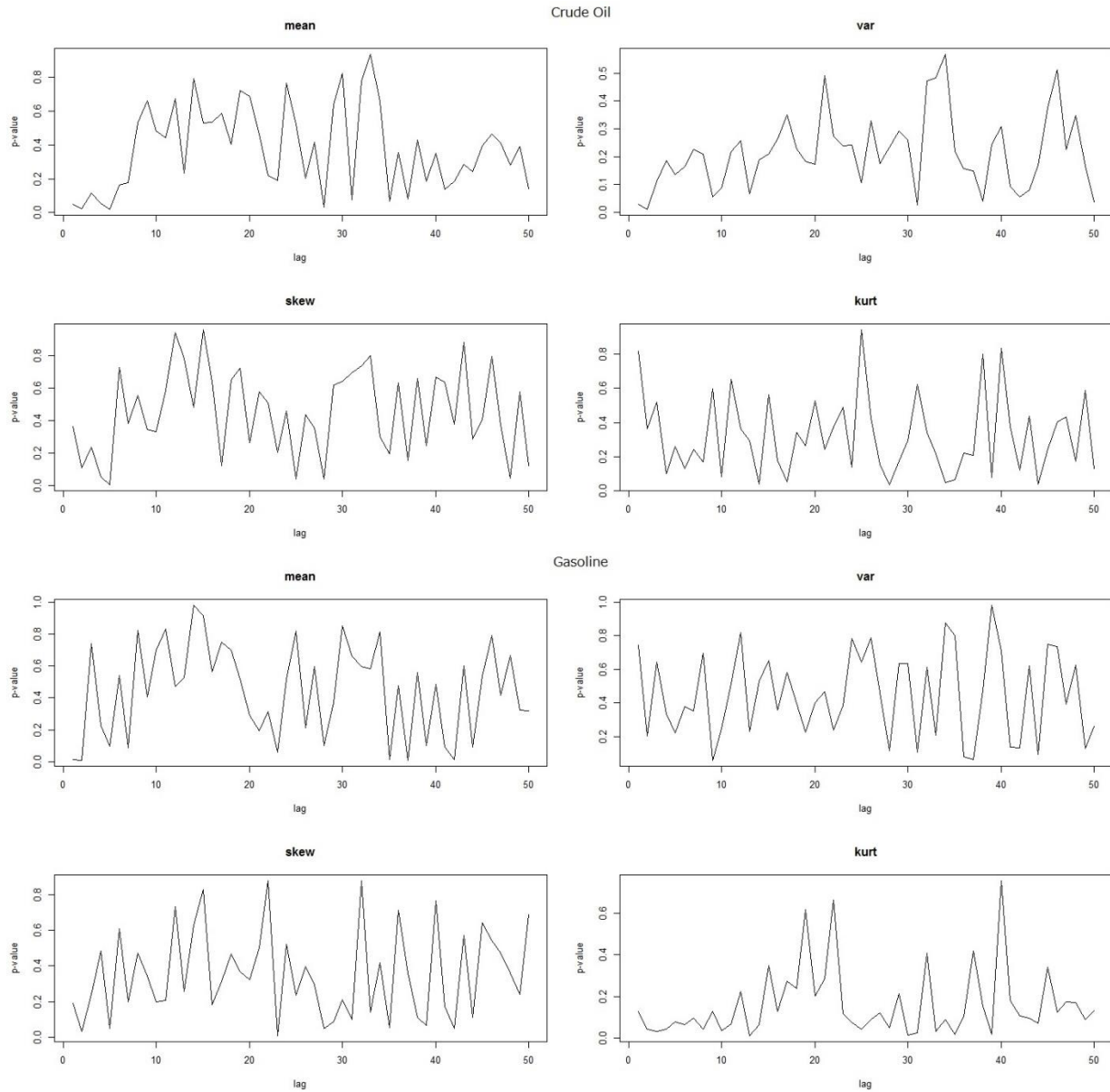


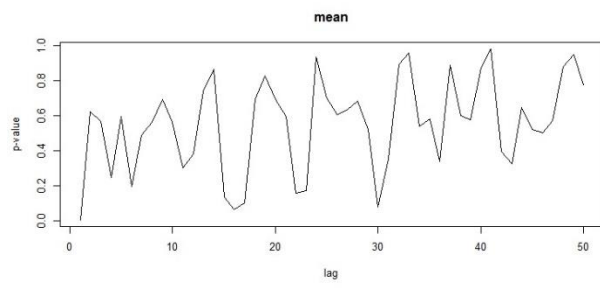
(i) Joint histogram of variance and kurtosis



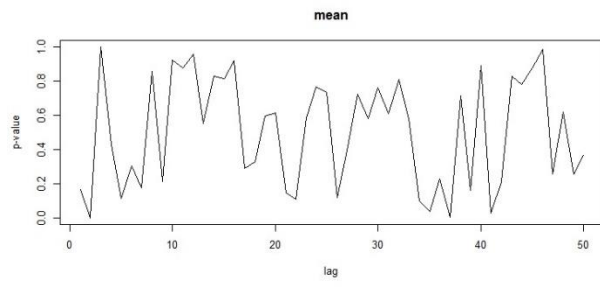
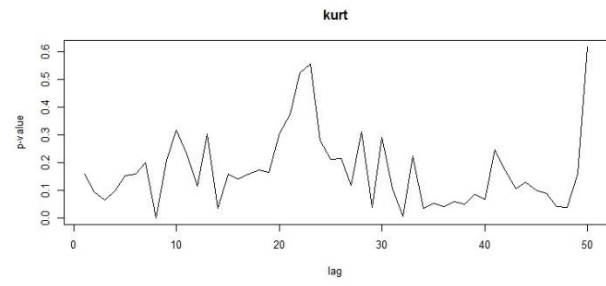
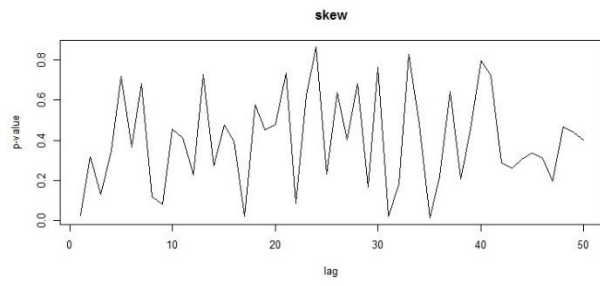
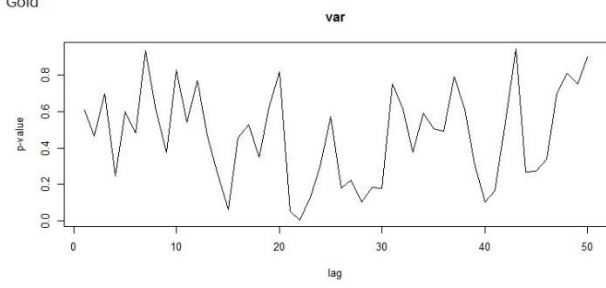
(j) Joint histogram of skewness and kurtosis

Figure 3 Joint histogram of return and realizing moments in crude oil market

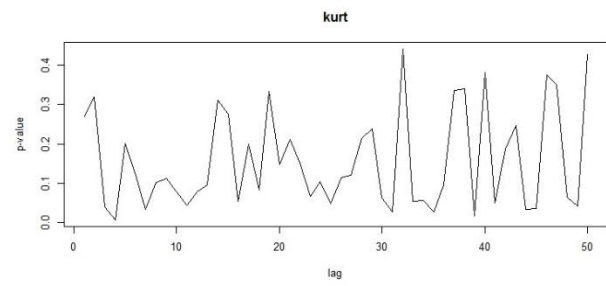
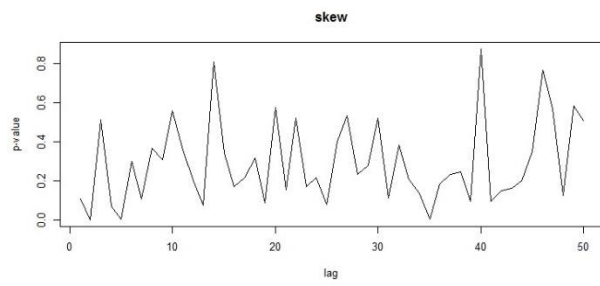
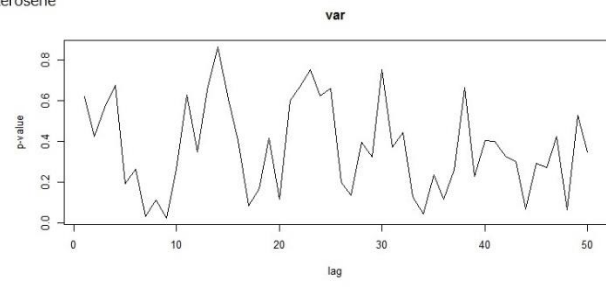


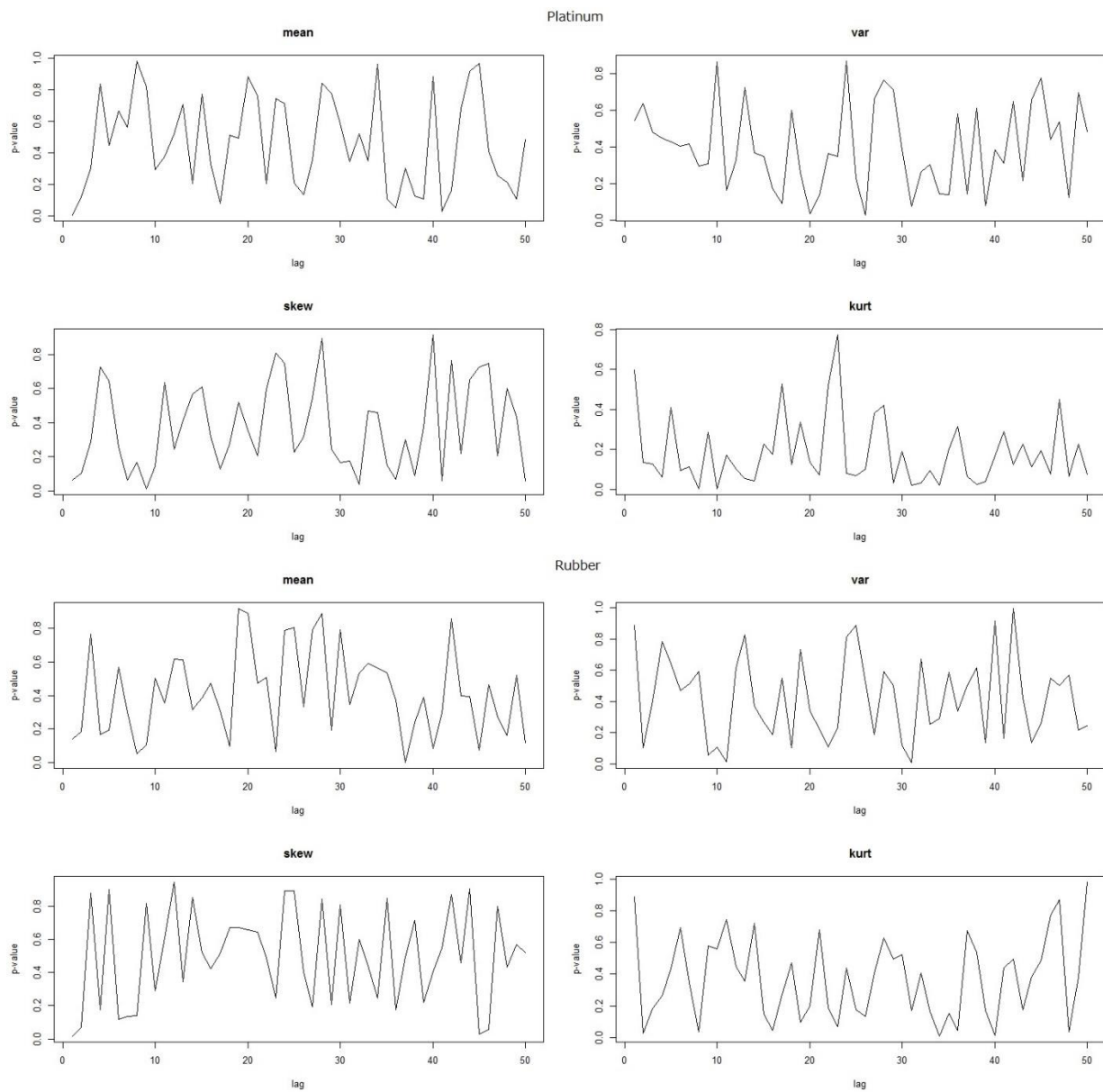


Gold



Kerosene





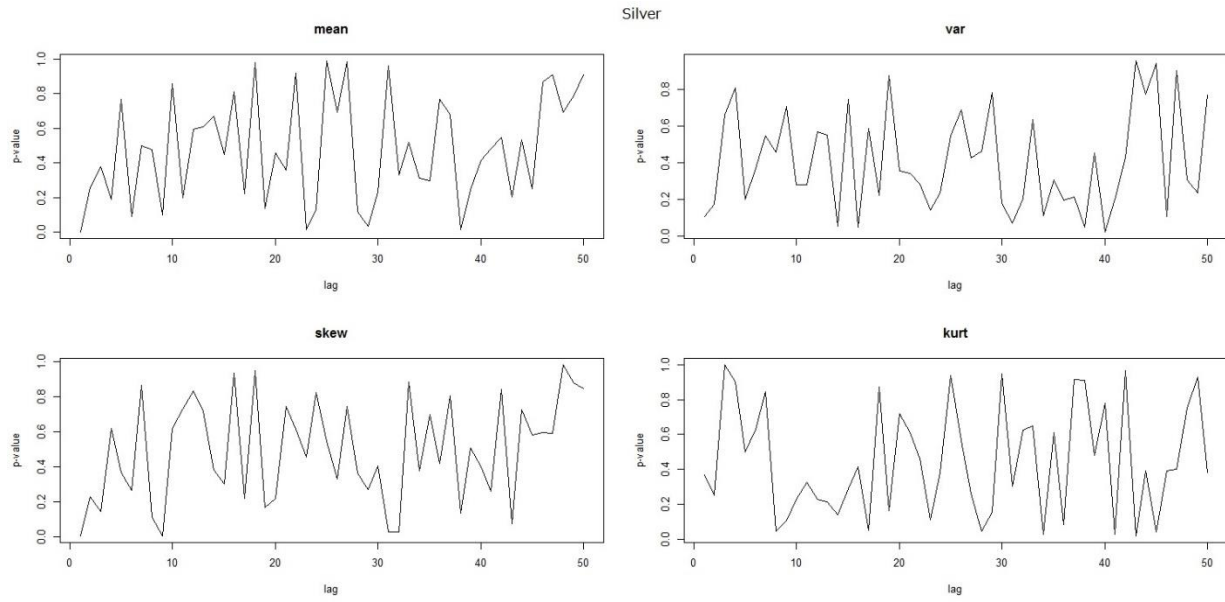


Figure 4 P-values for independence test between return and realizing moments in each market

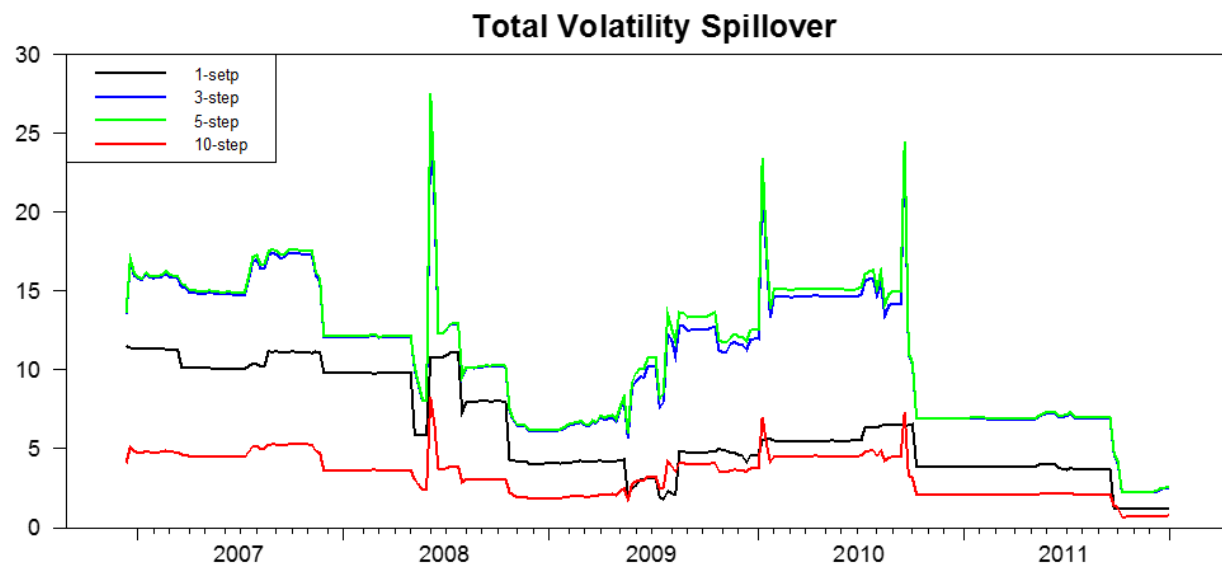


Figure 5 Dynamic total volatility spillover

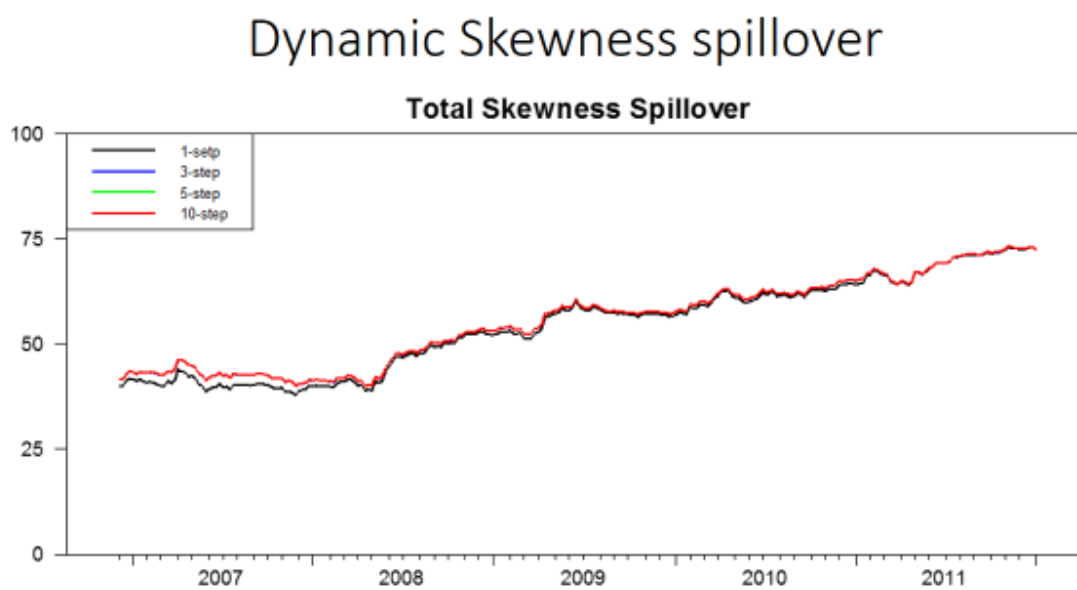


Figure 6 Dynamic total skewness spillover

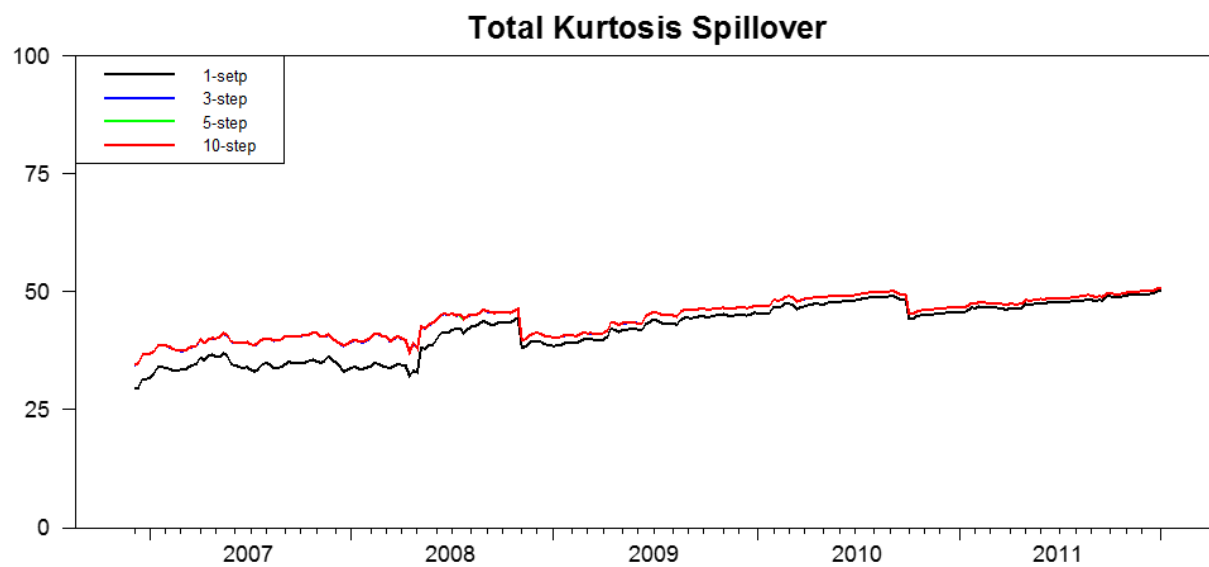
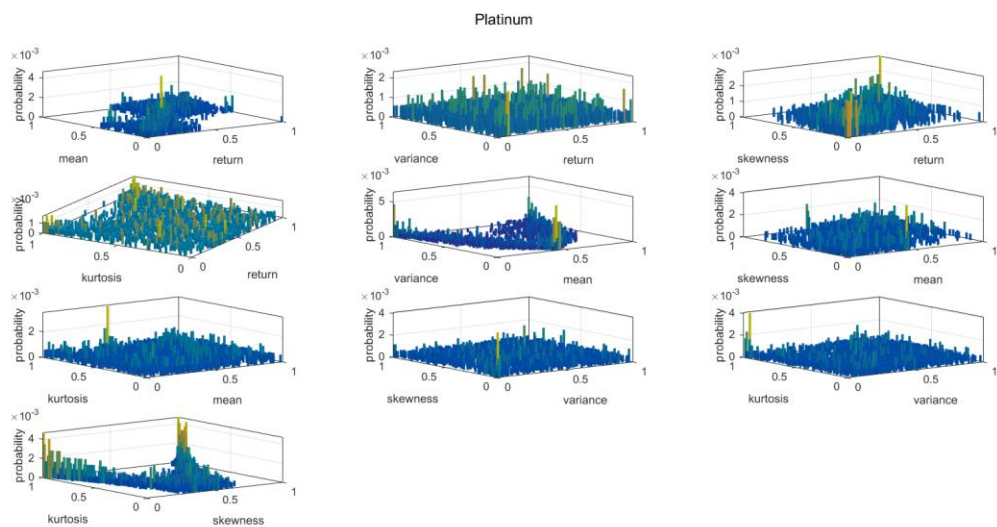
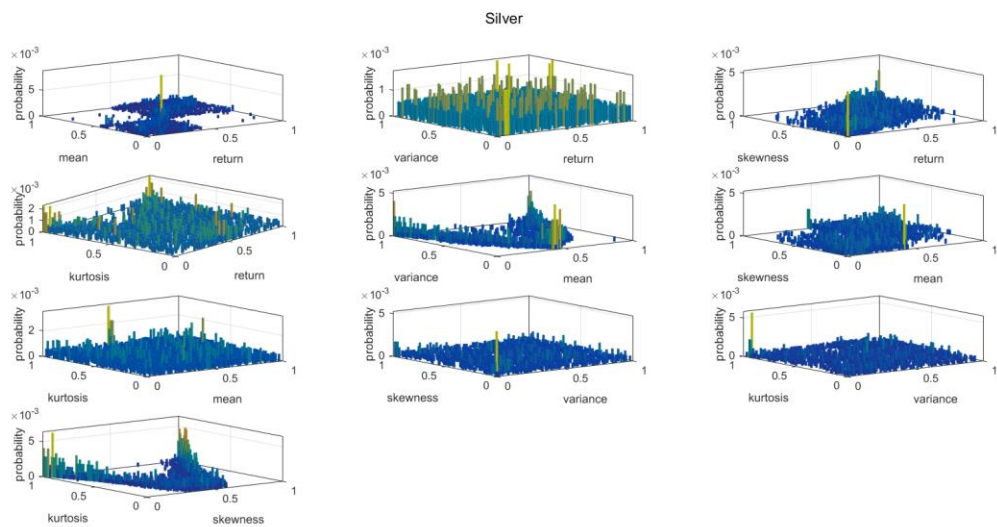
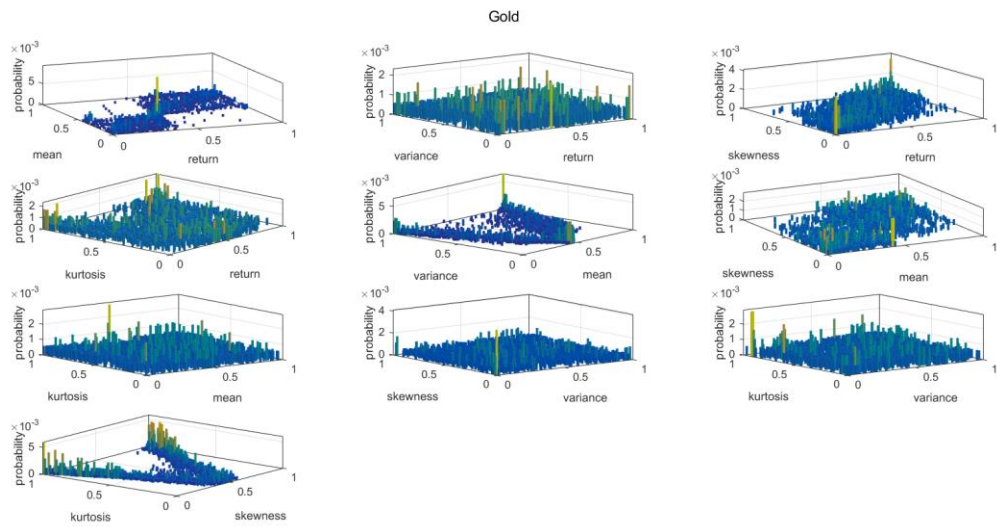
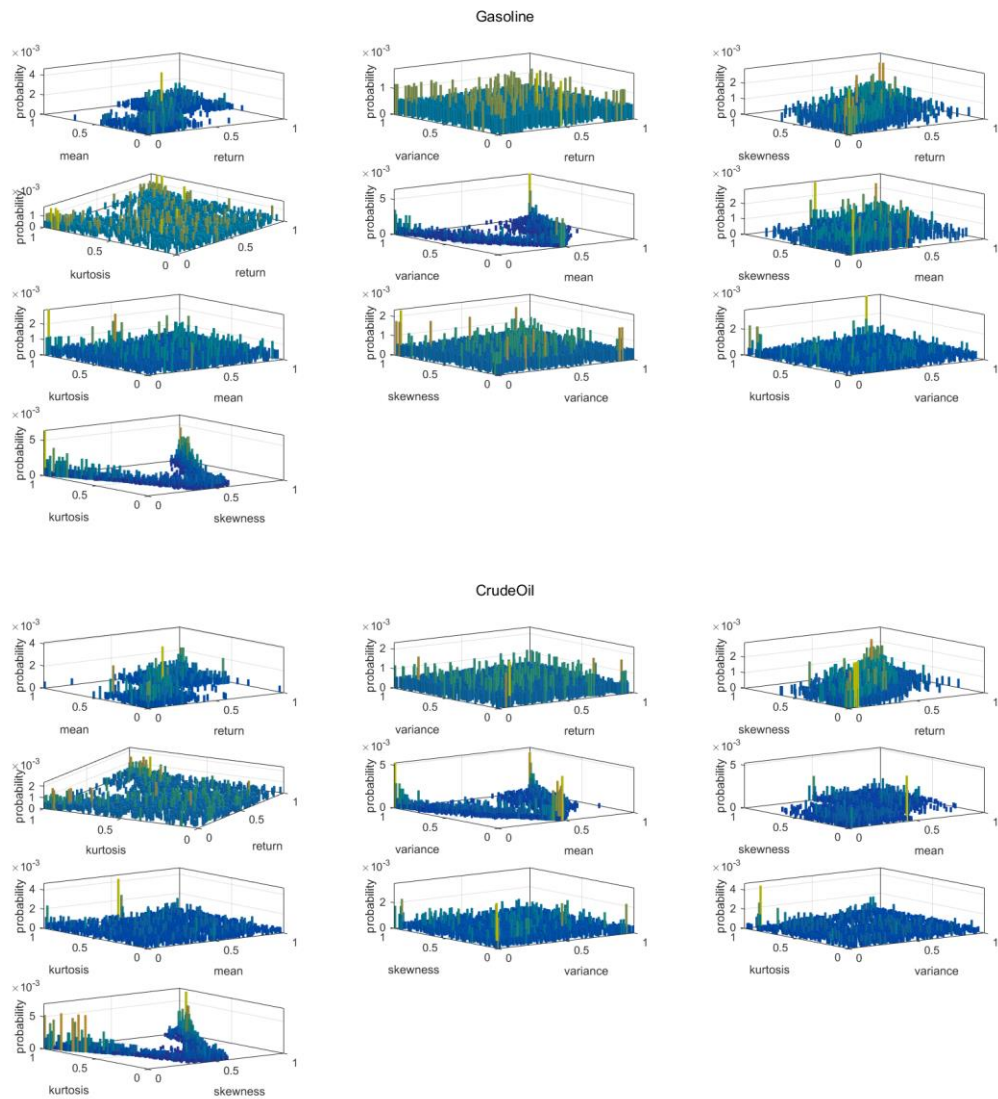
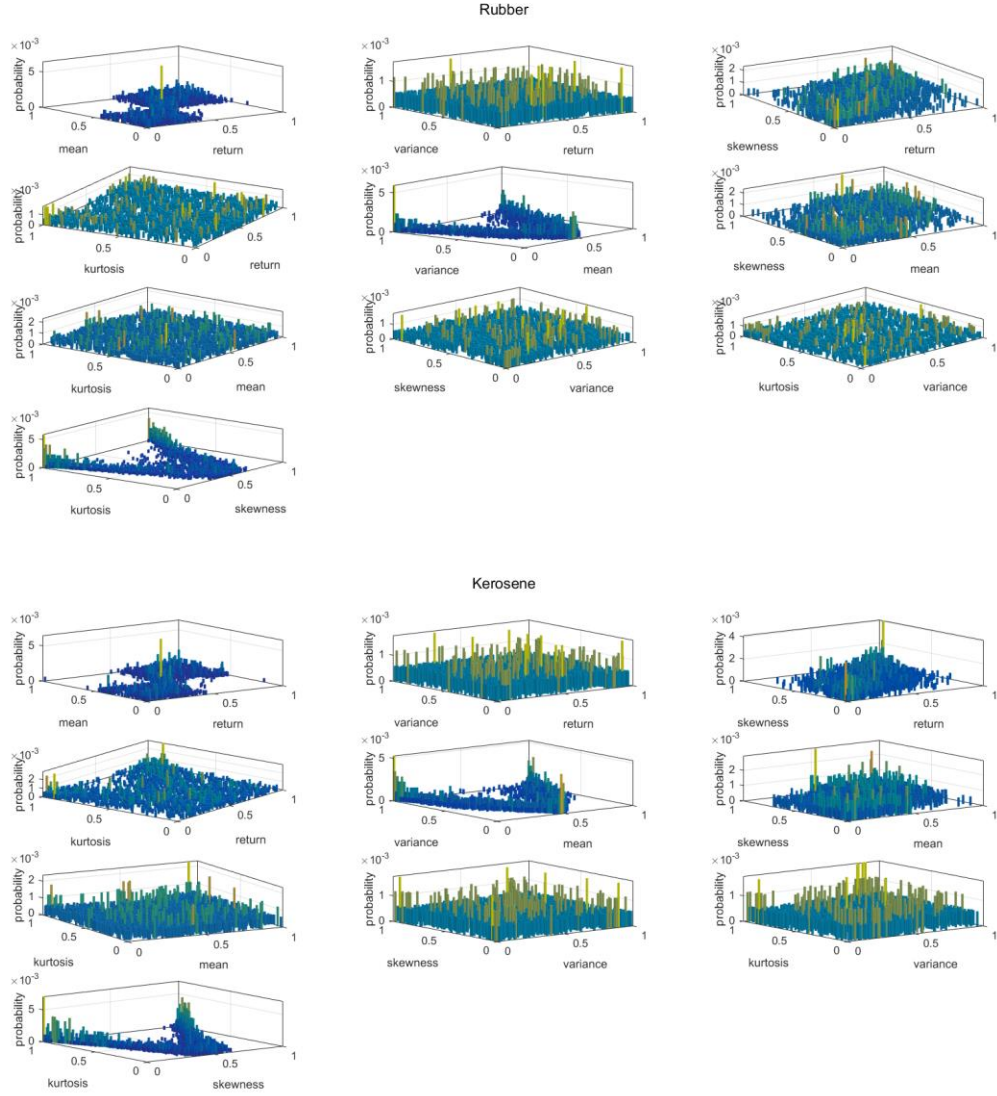


Figure 7 Dynamic total kurtosis spillover







Appendix A. Joint histogram of return and realizing moments in each market

Chapter 4

Improving density forecasting: The role of multivariate conditional dependence

4.1.Introduction

Many financial problems such as risk management, option pricing, and portfolio selection require fully characterized predictive return density. Predictive accuracy is becoming extremely important for investors. One conceivable method of improving forecast accuracy is taking account of conditional dependence, since much more information is available in the predictive densities $F_{i,t+h}(y_{i,t+h} | y_{1,t}, y_{1,t-1}, \dots, y_{i,t}, y_{i,t-1}, \dots, y_{n,t}, y_{n,t-1}, \dots)$ for multivariate models than $F_{i,t+h}(y_{i,t+h} | y_{i,t}, y_{i,t-1}, \dots)$ for univariate models. Various studies have aimed to capture the conditional dependence of financial returns by using multivariate GARCH (MGARCH) models. However, MGARCH models have poor forecast performance, even compared with univariate models. For instance, Efimova and Serletis (2014) investigate energy market volatilities and point out that univariate models provide more accurate forecasts than MGARCH models. Wang and Wu (2012) showed that univariate GARCH models with asymmetric effects outperform MGARCH models in forecasting crack spread volatility in energy markets. Sadorsky (2006) tested forecast performance in the petroleum futures by market using several loss functions and found that the BEKK model does not perform as well as single equation GARCH models. Caporin and McAleer (2014) compared the forecast performance of a range of MGARCH models and concluded that “all models are wrong, but some are useful than others.”

The poor forecast performance of MGARCH models is attributed to their model complexity and restrictions, especially elliptical joint density assumption of innovations, which cannot capture the well-known asymmetric properties of financial returns. These asymmetries include asymmetric volatility clustering (Glosten et al. (1993)), skewness within the distribution of individual stock returns (Ait-Sahalia and Brandt (2001)), and asymmetric dependence (Ang and Chen (2002), Longin and Solnik (2001), Patton (2004)). In sum, MGARCH type models are not well-specified to capture complicated properties of the joint distribution and joint behavior of financial returns.

In this study, we thus examine whether the density forecasts of returns can be improved by incorporating multivariate conditional dependence correctly. To capture the more complex

properties of multivariate distributions, the copula method has attracted substantial recent attention. Copula theory, developed by Sklar (1959), explains that a joint distribution can be decomposed into its marginal distributions and a copula, which can characterize the dependence between variables. The concept of the conditional copula was introduced by Patton (2006), who adapted the copula framework to a time-series case by allowing for time variation in the parameters of marginal distributions. To extend copula applications from bivariate to multivariate cases, Aas et al. (2009) showed that multivariate data that exhibit complex patterns of tail dependence can be modeled by using the pair-copula, which is constructed by using only bivariate copulas for the joint distribution of certain pairs of variables conditional on a specified set of variables. This idea dates back to Joe (1996), and it was also investigated in-depth by Bedford and Cooke (2001, 2002). The identification of conditioning and conditioned variables is facilitated by a sequence of trees. These trees and the corresponding multivariate distributions are called regular vines and the R-Vine distribution, respectively. Although an R-Vine provides a flexible class of multivariate dependence, its flexibility comes along with exponentially increasing complexity in large dimensions. Most empirical studies focus on its two restrictive cases: canonical vines (C-Vines) and drawable vines (D-Vines). An automated model selection and estimation technique for R-Vines was proposed by Dißmann et al. (2013). They suggested determining the dependence structure of each level of the R-Vine tree by first maximizing the empirical Kendall's tau and then selecting the copula generators for each pair of conditioning variables according to selected information criteria. Another truncation method for use with regular vines was proposed by Brechmann et al. (2012). They suggested truncating the R-Vine tree according to the Vuong (1989) test, which can be used to compare non-nested models based on the likelihood ratio, and multivariate goodness-of-fit test, finding that the most important dependencies and tail behavior are captured in the first several levels of the trees.

Owing to the large number of previous studies of copulas and regular vines, we can thus specify flexible multivariate distributions to study how multivariate dependence contributes to density forecasting. Our main contributions are threefold. First, we propose a semi-parametric one-step-ahead out-of-sample density forecast generation method from regular vine models and test density forecast performance. Although R-Vine is a recent introduction for capturing the abundant properties of the multivariate distribution, few studies have thus far tested the forecast

performance of this type of model. Our study therefore contributes to test the density forecast performance of regular vine models.

Second, we focus on the forecast of the whole return density instead of moments or quantile forecasts. Although most financial applications are simple multivariate problems with volatility forecasts (Laurent et al. (2012)), many recent studies have shown that conditional skewness and kurtosis also play an important role in financial forecasting (Leon et al. (2005), Virbickaite et al. (2016), Jondeau and Rockinger (2003), Bali et al. (2008), Harvey and Siddique (2000)). The whole density forecast is much more informative than simple moments (volatility, skewness, and kurtosis) or quantile forecasts, and thus considerably more useful for investors.

Third, we evaluate the density forecasts based on the probability integral transform (PIT). Various studies have compared forecast performance based on loss functions, whereas whether the density forecasts generated from superior models are accurate remains unclear. The idea of using the PIT can be traced to Diebold et al. (1998). These authors proved that if a density forecast is accurate, the PIT should be i.i.d. uniformly distributed. Therefore, we can test whether the predictive densities are accurate. DGT(Diebold, Gunther and Tay)-type tests focus on evaluating the absolute performance of a model's predictive ability instead of the relative performance of competing models. This test is independent of users' loss functions. Since Diebold et al. (1998), a number of DGT-type tests based on the PIT have been proposed (Hong and Li (2005), Corradi and Swanson (2006), Gonzalez-Rivera and Yoldas (2012), Gonzalez-Rivera and Sun (2015), Bai (2003), Rossi and Sekhposyan (2016)). However, few works investigate the forecast performance of financial returns by using this type of test owing to its relative complexity. Our study bridges this gap in forecasting and evaluation applications of financial return densities based on the PIT.

In this paper, we run a 16-variable application, which is a large study compared with recent studies. Our results show that a large proportion of density forecasts are inaccurate in univariate cases, while their degree of accuracy improves markedly after taking conditional dependence into consideration. We also test each part of the density forecasts, including the left/right tail, left/right half, and center. We find that the misspecification of univariate models can mainly be attributed to the left tail and that this part improved most. This result suggests that regular vine-type models perform much better in risk management (e.g., (conditional) value-at-risk

estimation), which is the problem of the quantile forecasting of the left tail. Further, these results are robust during the global financial crisis and European debt crisis. However, although the density forecast improved after taking conditional dependence into account, a small proportion of density forecasts remained inaccurate when testing each part of the predictive densities, indicating that current R-Vine-type models and estimation methods are not perfect. In summary, this type of model may capture some statistical but not actual dependence.

The remainder of this paper is organized as follows. Section 2 briefly reviews the specification of regular vines and presents the estimation method. The methods for generating and evaluating the predictive density are discussed in Section 3. Section 4 presents the empirical results and Section 5 concludes.

4.2. Model specification

In this section, we briefly review the specification of the regular vine copula and estimation method of the R-Vine models. Vine models or pair-copula constructions (PCCs) can provide extreme flexibility when modeling joint distributions. For each pair of variables, a copula can be selected according to particular criteria. Here, not all variables necessarily share the same theoretical joint density (e.g., hierarchical Archimedean copula and multivariate copula). Furthermore, the dependence structure is estimated according to empirical dependence, which is realistic. The R-Vine model or PCC is defined next.

4.2.1. Regular vine copula

The notion of a regular vine distribution was introduced by Bedford and Cooke (2001, 2002) and described in more detail by Kurowicka and Cooke (2006). It involves the specification of a sequence of trees where each edge corresponds to a bivariate copula, a so-called pair-copula. These pair-copulas then constitute the building blocks of the joint regular vine distribution. A regular vine is defined as follows:

Definition 2.1 R-Vine

$\mathcal{V} = (T_1, \dots, T_{n-1})$ is an R-Vine on n elements if

- (i) T_1 is a tree with nodes $N_1 = \{1, \dots, n\}$ and a set of edges denoted E_1 .

(ii) For $i = 2, \dots, n-1$, T_i is a tree with nodes $N_i = E_{i-1}$ and edge set E_i .

(iii) If two edges in tree T_i are to be joined by an edge in tree T_{i+1} , they must share a common node (proximity condition).

To build a statistical model on R-Vine trees with node set $\mathcal{N} = \{N_1, \dots, N_{d-1}\}$ and edge set $\mathcal{E} = \{E_1, \dots, E_{d-1}\}$, each edge $e = e(a), e(b) \mid D(e)$ in E_i is associated with a bivariate copula density $c_{e(a), e(b) \mid D(e)}$. The nodes $e(a)$ and $e(b)$ are called conditioned nodes, while $D(e)$ is the conditioning set. An R-Vine distribution of the random vector $\mathcal{X}$ with a conditional copula density of $(x_{e(a)}, x_{e(b)})$ given the variables $x_{D(e)}$ is specified as $c_{e(a), e(b) \mid D(e)}$ for R-Vine trees with node set $\mathcal{N}$ and edge set $\mathcal{E}$.

Definition 2.2 R-Vine Copula

$(\mathcal{F}, \mathcal{V}, \mathcal{B})$ is an R-Vine copula specification if $\mathcal{F} = (F_1, \dots, F_n)$ is a vector of the continuous invertible distribution function, $\mathcal{V}$ is an n -dimensional R-Vine, and $\mathcal{B} = \{B_e \mid i = 1, \dots, n-1; e \in E_i\}$ is a set of copula with B_e being a bivariate copula, a so-called pair-copula.

As noted in the Introduction, most studies focus on two special cases of regular vines, namely C-Vines and D-Vines. In particular, an R-Vine is called a C-Vine if each tree T_i has a unique node with $d-i$ edges to other nodes and a D-Vine if each node in T_1 has no more than two edges to other nodes. In this study, we examine how the multivariate dependence provided by an R-Vine and its truncated versions (C-Vines and D-Vines) contributes to density forecasting.

According to Dißmann et al. (2013), the R-Vine density is given by

$$f_{1..n}(x) = \prod_{k=1}^n f_k(x_k) \prod_{i=1}^{n-1} \prod_{e \in E_i} c_{Ce(a), Ce(b) \mid D(e)}(F_{Ce(a) \mid D(e)}(x_{Ce(a)} \mid x_{D(e)}), F_{Ce(b) \mid D(e)}(x_{Ce(b)} \mid x_{D(e)})),$$

where $x = (x_1, \dots, x_n)$, $e = \{a, b\}$, and $x_{D(e)}$ stands for the variables in $D(e)$. f_i denotes the density of F_i for $i = 1, \dots, n$.

4.2.2. Estimation method

Given that the number of different possible R-Vines increases explosively as the dimension rises, estimating all potential R-Vines and selecting the best one is challenging. A sequential and heuristic estimation method for R-Vines was proposed by Dißmann et al. (2013). They suggested that fitting an R-Vine to a given data set copula specification requires the following separate tasks: (i) selecting the R-Vine structure, (ii) choosing a bivariate copula family for each pair selected in (i), and (iii) estimating the corresponding parameters for each copula.

They proposed selecting the tree by using the maximum spanning tree algorithm of Prim (1957), which maximizes the sum of the absolute empirical Kendall's tau. Their algorithm is summarized as follows:

For an n -dimensional random vector,

(1) Calculate the empirical Kendall's tau $\hat{\tau}_{j,k}$ for all possible variable pairs $\{j, k\}, 1 \leq j \leq k \leq n$.

(2) Select the spanning tree that maximizes the sum of the absolute empirical Kendall's tau

$$\max \sum_{e=\{j,k\} \text{ in spanning tree}} |\hat{\tau}_{j,k}|.$$

(3) For each edge $\{j, k\}$ in the selected spanning tree, select a copula and estimate the corresponding parameters. Then, transform $\hat{F}_{j|k}(x_{lj} | x_{lk})$ and $\hat{F}_{k|j}(x_{lk} | x_{lj})$, $l = 1, \dots, N$, using the fitted copula $\hat{C}_{jk}$.

(4) **For** $i = 2, \dots, n-1$ **do** (Iteration over the trees).

(5) Calculate the empirical Kendall's tau $\hat{\tau}_{j,k|D}$ for all conditional variable pairs $\{j, k | D\}$ that can be part of tree T_i , i.e., all edges fulfilling the proximity condition.

(6) Among these edges, select the spanning tree that maximizes the sum of the absolute empirical Kendall's tau,

$$\max \sum_{e=\{j,k|D\} \text{ in spanning tree}} |\hat{\tau}_{j,k}|.$$

(7) For each edge $\{j, k | D\}$ in the selected spanning tree, select a conditional copula and estimate the corresponding parameters. Then, transform $\hat{F}_{jk \cup D}(x_{lj} | x_{lk}, x_{lD})$ and

$\hat{F}_{k|j \cup D}(x_{lk} | x_{lj}, x_{lD}), l = 1, \dots, N$, using the fitted copula $\hat{C}_{jk|D}$.

(8) End for

In addition to the steps described above, a copula family for each pair of variables is necessary. In the empirical studies of this paper, we take the following copula families into consideration:

- (a) Gaussian (symmetric, no tail dependence).
- (b) Student-t (symmetric, tail dependence).
- (c) Frank (symmetric, no tail dependence).
- (d) Gumbel (asymmetric, upper tail dependence), survival Gumbel (180 degree), and rotated Gumbel (90 and 270 degrees).
- (e) Clayton (asymmetric, lower tail dependence), survival Clayton (180 degree), and rotated Clayton (90 and 270 degrees).
- (f) Joe (asymmetric, upper tail dependence), survival Joe (180 degree), and rotated Joe (90 and 270 degrees).
- (g) Independence copula

Given these options, we select the best fitting copula for each pair of variables according to the AIC, which has been found to be a reliable criterion for selecting bivariate copula (Manner, 2007; Brechmann, 2010). We also consider a truncation method that applies a preliminary independence test based on Kendall's tau, as described in Genest and Favre (2007), before copula selection. If the test indicates independence, the independence copula is chosen. The given significance level is 5%.

4.2.3. Marginal specification

We specify the margins by using an AR(1)-EGARCH(1,1) model with general error distribution. The model is specified as follows:

$$r_t = \mu + \phi r_{t-1} + \sqrt{h_t} z_t$$

$$\log h_t = \omega + \beta \log h_{t-1} + \theta z_{t-1} + \gamma(|z_{t-1}| - E|z_{t-1}|).$$

To study the role of multivariate conditional dependence in forecasting return distributions, we also generate the predictive densities from the univariate model for comparison purposes.

4.3. Density forecasting and evaluation

Density forecasting has attracted growing scholarly attention since a fully specified forecast of the return distribution plays an important role in financial applications. The evaluation of density forecasts is difficult since the true density cannot be observed. Indeed, although many studies have evaluated the density forecasts of competing models by comparing the loss functions based on predictive likelihood, whether the predictive density generated by the superior model is indeed accurate remains unclear.

Diebold et al. (1998) showed that the predictive density can be evaluated by using the PIT, which is the conditional C.D.F. $F_{t+h|t}$ evaluated at the realized values y_{t+h} :

$$z_{t+h} = \int_{-\infty}^{y_{t+h}} f_{t+h}(y_{t+h} | \Omega_t) dy = F_{t+h}(y_{t+h} | \Omega_t)$$

These authors proved that if the predictive density is accurate, then the series of the PITs is i.i.d. uniformly distributed: $\{z_t\}_{t=1}^T \sim i.i.d U(0,1)$. Therefore, the problem of testing whether the predictive density is accurate can be converted into testing the autocorrelation and uniformity of the PITs.

Here, we generate one-step-ahead out-of-sample density forecasts by using regular vine models and calculate the PITs according to the following steps:

- (1) Estimate the univariate models, extract the residuals $(\hat{\varepsilon}_1', \dots, \hat{\varepsilon}_d')$, and transform into uniform data $(\hat{u}_1', \dots, \hat{u}_d')$ through the empirical C.D.F. of each margin.
- (2) Estimate the R-Vine models to obtain the joint distribution $\hat{F}(u_1, \dots, u_d)$.
- (3) Simulate pseudo observations $(u_1, \dots, u_d)$ from the joint density. This step takes conditional dependence into consideration for each margin.
- (4) Inverse the simulated pseudo observations $(u_1, \dots, u_d)$ to the empirical C.D.F.s of the shocks obtained in step (1), which is $F_{\varepsilon_1}^{-1}(u_1), \dots, F_{\varepsilon_d}^{-1}(u_d)$. Then, we can derive the simulated densities of shocks $F_1(\varepsilon_1), \dots, F_d(\varepsilon_d)$ ². This step takes account of the historical information of each margin, which means the simulated densities conditioned on past information.

² For a continuous distribution function F_i , and an uniform variable u , $F_i^{-1}(u) \sim F_i$.

(5) Convert the simulated densities of the error terms $F_1(\varepsilon_1), \dots, F_d(\varepsilon_d)$ into return densities $F_1(r_1), \dots, F_d(r_d)$ by using the estimated parameters $\hat{\theta}_i$ of each margin i obtained in step (1).

(6) Calculate the PITs.

The predictive densities are simulated by using the rolling window method with a window size of 300 observations. Each density is simulated 100 thousand times. We also simulate the predictive densities and calculate the PITs for the univariate models by way of a comparison.

To test whether predictive density is accurate, we adopt the correct specification test proposed by Rossi and Sekhposyan (2016). This test evaluates the actual forecasting ability of models at the estimated parameter values in finite samples, which is more empirically useful than previous DGT-type tests. Given the parametric model and estimation method specified by the researcher, whether the predictive densities are correctly specified can be tested. Accordingly, this test does not require an asymptotic correction for parameter estimation errors. Even in dynamic misspecifications, limiting distributions that are nuisance parameter free for one-step-ahead density forecasts are obtained. Hence, the null hypothesis of the correct specification of the density forecast can still hold even if the forecasting model is dynamically misspecified. We thus focus on testing whether the conditional distribution is equal to the true density:

$$H_0 : F_{i,t+h}(y_{i,t+h} | y_{i,t}, y_{i,t-1}, \dots) = F_{i,t+h}(y_{i,t+h}) \text{ for univariate model and}$$

$$H_0 : F_{i,t+h}(y_{i,t+h} | y_{1,t}, y_{1,t-1}, \dots, y_{i,t}, y_{i,t-1}, \dots, y_{n,t}, y_{n,t-1}, \dots) = F_{i,t+h}(y_{i,t+h}) \text{ for Vine models,}$$

where $i \in [1, n]$ denote each margin i .

The test statistics are given by

$$\kappa_p = \sup_{u \in [0,1]} \Psi_p(u)^2 \Rightarrow \sup_{u \in [0,1]} \Psi(u)^2$$

where $\Psi_p(u) \equiv P^{1/2} \sum_{t=R}^T \xi_{t+h}(u)$ and $\xi_{t+h}(u) \equiv (1\{F_{t+h}(r_{t+h} | \Omega_{t-R+1}^t) \leq u\} - u)$. Critical values are available in Rossi and Sekhposyan (2016).

4.4. Empirical results

4.4.1. Data

Our dataset consists of the stock market indices from 16 countries: Italy, Germany, Japan, Australia, China, Singapore, the United Kingdom, the United States, Brazil, Chile, Russia, South Africa, Mexico, Poland, Turkey, and Hong Kong. All data are MSCI total return indices presented in US dollars. The sample period runs from January 3, 2005 to November 30, 2015, including 2814 observations. The data are summarized in Table 1.

4.4.2. In-sample results

We consider six models, namely the R-Vine, C-Vine, and D-Vine models as well as their truncated versions, by applying an independence test before choosing the copulas. Figure 1 presents the first level of the estimated dependence structure of the R-Vine, C-Vine, and D-Vine models. The structure of the trees remains regardless of whether an independence test is performed.

The structure of the R-Vine tree reflects the expected relationship among the shocks of the stock indices. Regional countries are grouped together, such as Asian advanced economies (Australia, Singapore, Japan, Hong Kong), American countries (the United States, Brazil, Mexico), and European countries (the United Kingdom, Germany, Poland, Russia, Turkey, Italy), owing to the integration of regional economies and financial markets. This finding suggests that the maximum spanning tree method of determining the tree structure performs well.

The C-Vine tree indicates that the United Kingdom has the strongest aggregated dependence with all the other countries. One conceivable reason for the center selection of the United Kingdom rather than, say, the United States or Germany is that our data consist of six European countries. Since the maximum spanning tree method maximizes the empirical Kendall's tau, it is logical that the United Kingdom would be selected. Even in the R-Vine and D-Vine trees, the United Kingdom is located in important (i.e., centered) locations.

Table 2 presents the log-likelihood, AIC, and BIC values of the six models. This table shows that the R-Vine model best fits the data, since an R-Vine provides the most flexible dependence structure. The C-Vines and D-Vines are truncated versions, which are especially restricted in large dimensions. Another truncation method (i.e., that which incorporates the independence copula) performs less well. We observe that the R-VineInd model provides a lower log-likelihood than the C-Vine model, while the C-VineInd model performs worse than the D-Vine

model. The AIC suggests the same results as the log-likelihood. However, according to the BIC, the R-VineInd and D-VineInd models outperform the R-Vine and D-Vine models, respectively. The difference between the C-Vine and C-VineInd model also reduces. Thus, there is a tradeoff between better fit and the number of parameters. Although some pairs of variables are tested to be independent at the given significance level, modeling their dependence can better fit the data.

Table 3 presents the results of the Vuong test. The Vuong test statistics can also be corrected for the number of parameters according to the penalty terms of the AIC and BIC. Without correcting for the number of parameters, we find that the R-Vine model outperforms the other models significantly. Even with the AIC correction, the R-Vine model is still the best. However, with the BIC correction, the R-VineInd model significantly outperforms all the other models, while the difference between C-Vine and C-VineInd becomes insignificant. Overall, the most flexible R-Vine/R-VineInd models outperform the C-Vine and D-Vine models significantly according to all correction criteria.

4.4.3. Out-of-sample results

4.4.3.1. Testing the whole predictive densities

In this section, we represent results of testing density forecast performance based on the PIT. The one-step-ahead out-of-sample predictive densities are simulated by using the rolling window method with a window size of 300 observations. We calculate and test the corresponding PITs according to steps stated in section 3.

Table 4 presents the results of the density forecast evaluation for all the predictive densities over the sample period. Test statistics in blue and red advocate the rejection of the null hypothesis of the correct specification at the 10% and 5% significance levels, respectively. Based on the test statistics of uniformity mentioned in Section 3, we find that six of the 16 density forecasts provided by the univariate model (China, the United Kingdom, the United States, Chile, South Africa, Hong Kong) are inaccurate under the given 5% or 10% significance level. Further, by incorporating multivariate dependence, the null hypothesis of accurate density forecasts cannot be rejected in all cases. Although the C-Vine and D-Vine models do not fit the data as well as the R-Vine model does, they do also provide accurate density forecasts. Truncate Vine models using

independent copula would also not affect predictive ability. P-values of the Ljung–Box test suggest that none of the PITs is autocorrelated.

4.4.3.2 Testing each part of the predictive densities

The previous section showed that all the predictive densities are accurate, whereas we are more interested in whether each part of the predictive densities is accurate. Here, we focus on the following parts: (1) left tail $[0, 0.25]$; (2) right tail $[0.75, 1]$; (3) left half $[0, 0.5]$; (4) right half $[0.5, 1]$; and (5) center $[0.25, 0.75]$.

Table 5 provides these test statistics for each part of the predictive densities. The left tails are most likely to be misspecified and improved most after taking conditional dependence into account. Specifically, the null hypothesis of correct specification is rejected in 10 of 16 of the densities at the 5% or 10% significance level for the univariate model. After taking account of conditional dependence, this number reduces to two of the 16 for the RVineInd/DVine/DVineInd models and one of the 16 for the RVine/CVine/CVineInd models. Consequently, Vine models should perform much better in risk management, such as VaR or CVaR prediction. For the other parts of the predictive densities, two to six of the 16 densities are misspecified for the univariate model and most of them become accurate after using Vine models. The only exception is the right tail of the United States, which becomes inaccurate at the 10% significance level for three Vine models after taking conditional dependence into account. Overall, consistent with the results of evaluating all the predictive densities, each part of the predictive densities also improves considerably.

4.4.3.3 Robustness check during the crisis period

The previous results show that Vine models are useful during the whole sample period. However, whether these models also perform well in turbulent periods remains unclear. Therefore, we test forecast performance during the global financial crisis and European debt crisis period, from August 2007 to June 2012. Table 6 presents these test statistics during these crisis periods. In panel A, three of the 16 densities are inaccurate for the univariate model and they all become accurate for Vine models. In panel B, half of the left tails of the predictive densities are inaccurate for the univariate models. Five of them become accurate after using Vine models and three remain inaccurate. However, all the predictive density of Italy and the left tail

of the United Kingdom become inaccurate after taking account of conditional dependence, indicating that Vine models may capture some statistical but not real dependence. Overall, we can conclude that density forecasts can be improved by taking account of conditional dependence during the crisis period, but this improvement is not as large as that across the whole sample period.

4.5. Conclusion

In this study, we investigate whether density forecasts can be improved by incorporating multivariate conditional dependence correctly. MGARCH models have poor forecast performance since they cannot specify the joint distribution well. We try to capture the joint density by using Vine-type models, which can provide much more flexibility when capturing the asymmetric dependence of financial returns. A semi-parametric density forecast simulation method from Vine models is proposed. The predictive densities are evaluated by using a DGT-type test based on the PIT. We also test each part of the predictive densities and check the robustness during the global financial crisis and European debt crisis periods.

The presented results show that the multivariate distribution can be described well by R-Vine models. In particular, the in-sample results suggest that the R-Vine/R-VineInd models outperforms the C-Vine and D-Vine models since the R-Vine model provides the most flexible and reasonable dependence structure. Furthermore, by evaluating the forecast performance, we find that the predictive densities improve markedly after incorporating conditional dependence. By testing each part of the predictive densities, the left tails are found to be most likely to be misspecified and improved after taking conditional dependence into account, suggesting that Vine models perform well in risk management. The robustness check shows that density forecasts can also be improved during the crisis period. Although the R-Vine model fits the data better than its special cases such as C-Vine and D-Vine models, forecast performance is almost the same. Truncating Vine models by using independence copula would also not affect predictive ability. Finally, the forecast performance of Vine-type models is not perfect. Existing Vine models and their current estimation method may thus capture some statistical but not real dependence, which may lead to the misspecification of predictive densities.

Table 1 Descriptive statistics

	IT	GE	JP	AU	CN	SI	UK	US	BR	CH	RU	SA	MX	PO	TU	HK
Mean*10 ⁴	-0.54	2.42	0.85	2.44	2.11	2.85	1.45	2.81	1.59	4.16	0.81	2.50	3.19	0.41	1.51	3.01
Stdev*100	1.82	1.63	1.39	1.69	1.51	1.34	1.46	1.25	2.28	1.80	2.46	1.91	1.70	2.09	2.35	1.34
Maximum*10	1.25	1.16	1.15	0.88	1.03	0.86	1.22	1.10	1.66	1.40	2.40	1.24	1.52	1.42	1.62	1.04
Minimum*-10	1.09	0.96	0.95	1.60	1.42	0.98	1.04	0.95	1.83	1.28	2.56	1.36	1.09	1.34	1.47	1.24
Skewness	-0.01	-0.04	-0.19	-0.80	-0.77	-0.23	-0.12	-0.35	-0.29	-0.02	-0.44	-0.26	-0.07	-0.23	-0.28	-0.19
Kurtosis	8.31	8.94	8.46	11.51	13.41	9.11	12.55	14.34	10.69	10.15	18.18	7.35	10.50	7.70	7.06	11.27

Note: IT, GE, JP, AU, CN, SI, UK, US, BR, CH, RU, SA, MX, PO, TU and HK represent Italy, Germany, Japan, Australia, China, Singapore, United Kingdom, United States, Brazil, Chile, Russia, South Africa, Mexico, Poland, Turkey and Hong Kong, respectively.

Table 2 Likelihood based criteria

	RVine	RVineInd	CVine	CVineInd	DVine	DVineInd
loglikelihood	20042.61	19884.76	19988.85	19841.09	19877.37	19783.75
AIC	-39685.22	-39475.53	-39573.71	-39342.17	-39344.74	-39223.49
BIC	-38496.74	-38602	-38373.35	-38331.97	-38126.56	-38201.41

Note: RVine, CVine and DVine denote Regular Vine, Canonical Vine and Drawable Vine model, respectively. 'Ind' denotes models incorporate independent copula.

Table 3 Vuong Test

	RVine	RVineInd	CVine	CVineInd	DVine
No correction					
RVineInd	8.41 (0)				
CVine	1.98 (0.05)	-3.34 (0)			
CVineInd	5.95 (0)	1.27 (0.21)	6.82 (0)		
DVine	5.4 (0)	0.23 (0.82)	3.14 (0)	-0.94 (0.35)	
DVineInd	7.88 (0)	3.17 (0)	5.39 (0)	1.43 (0.15)	6.51 (0)

AIC correction					
RVineInd	5.59				
	(0)				
CVine	2.05	-1.58			
	(0.04)	(0.11)			
CVineInd	5.07	1.93	5.35		
	(0)	(0.05)	(0)		
DVine	5.56	2.02	3.22	-0.03	
	(0)	(0.04)	(0)	(0.97)	
DVineInd	7.03	3.95	4.61	1.48	4.21
	(0)	(0)	(0)	(0.14)	(0)
BIC correction					
RVineInd	-2.81				
	(0.01)				
CVine	2.27	3.67			
	(0.02)	(0)			
CVineInd	2.43	3.91	0.96		
	(0.01)	(0)	(0.34)		
DVine	6.05	7.33	3.47	2.65	
	(0)	(0)	(0)	(0.01)	
DVineInd	4.5	6.28	2.26	1.63	-2.6
	(0)	(0)	(0.02)	(0.1)	(0.01)

Note: The values in parentheses are p-values. Positive values of statistics favor the model in the column. P-values larger than 0.1 indicate that the difference between the two models is not significant.

Table 4 Density forecast performance evaluation for the whole predictive densities

	IT	GE	JP	AU	CN	SI	UK	US	BR	CH	RU	SA	MX	PO	TU	HK
Panel A. Test statistics																
Rvine	0.81	0.86	0.59	0.81	0.77	0.95	0.79	0.88	0.87	0.60	0.93	0.62	0.67	0.67	0.60	0.80
RvineInd	0.84	0.95	0.56	0.80	0.78	0.99	0.86	0.89	0.79	0.59	0.93	0.61	0.64	0.62	0.60	0.75
Cvine	0.84	0.87	0.63	0.81	0.74	0.95	0.84	0.92	0.84	0.59	0.93	0.64	0.72	0.64	0.60	0.79
CvineInd	0.82	0.87	0.61	0.73	0.81	1.01	0.84	0.84	0.84	0.58	0.92	0.58	0.69	0.64	0.60	0.78
Dvine	0.83	0.95	0.60	0.80	0.78	1.00	0.86	0.86	0.79	0.59	0.94	0.62	0.64	0.64	0.61	0.81
DvineInd	0.81	0.93	0.61	0.82	0.77	0.98	0.78	0.87	0.81	0.58	0.91	0.61	0.70	0.61	0.65	0.80
Uni	0.94	0.86	0.97	1.02	1.58	1.13	1.28	1.22	0.84	1.37	1.12	1.75	0.97	0.57	1.10	1.26
Panel B. P-values for Ljung-Box test																
Rvine	0.88	0.84	0.55	0.87	0.3	0.26	0.55	0.17	0.81	0.48	0.56	0.22	0.12	0.39	0.35	0.5
RvineInd	0.87	0.83	0.55	0.87	0.32	0.29	0.57	0.17	0.82	0.48	0.57	0.23	0.12	0.38	0.37	0.52
Cvine	0.87	0.83	0.58	0.88	0.32	0.28	0.56	0.16	0.8	0.48	0.55	0.23	0.12	0.39	0.34	0.54
CvineInd	0.88	0.85	0.57	0.87	0.31	0.28	0.58	0.17	0.83	0.49	0.56	0.21	0.13	0.39	0.35	0.52
Dvine	0.87	0.83	0.58	0.88	0.31	0.29	0.57	0.17	0.81	0.47	0.55	0.22	0.13	0.38	0.36	0.51
DvineInd	0.87	0.83	0.58	0.88	0.3	0.28	0.57	0.17	0.83	0.5	0.56	0.23	0.13	0.41	0.34	0.52
Uni	0.99	0.88	0.61	0.84	0.59	0.22	0.68	0.29	0.93	0.48	0.83	0.45	0.34	0.32	0.52	0.23

Note: Values in blue and red advocate rejection of the null hypothesis of correct specification at 10% and 5% significance level, respectively.

Table 5 Density forecast performance evaluation for each part of the predictive densities

	IT	GE	JP	AU	CN	SI	UK	US	BR	CH	RU	SA	MX	PO	TU	HK
Panel A. Test statistics for left tail [0,0.25]																
Rvine	0.80	0.86	0.49	0.81	0.58	0.95	0.79	0.69	0.69	0.54	0.85	0.62	0.61	0.55	0.60	0.66
RvineInd	0.73	0.95	0.46	0.80	0.55	0.99	0.86	0.69	0.72	0.57	0.80	0.61	0.63	0.57	0.60	0.65
Cvine	0.77	0.87	0.50	0.81	0.55	0.95	0.84	0.65	0.68	0.55	0.86	0.64	0.65	0.55	0.60	0.67
CvineInd	0.82	0.87	0.48	0.73	0.56	1.01	0.84	0.65	0.70	0.55	0.77	0.58	0.64	0.52	0.60	0.66
Dvine	0.77	0.95	0.51	0.80	0.56	1.00	0.86	0.70	0.65	0.59	0.85	0.62	0.64	0.51	0.61	0.67
DvineInd	0.80	0.93	0.48	0.82	0.63	0.98	0.78	0.63	0.69	0.54	0.82	0.61	0.64	0.52	0.62	0.65
Uni	0.94	0.86	0.59	1.02	1.25	1.13	0.87	1.22	0.84	0.97	1.12	0.96	0.68	0.52	1.10	1.26
Panel B. Test statistics for right tail [0.75,1]																
Rvine	0.81	0.64	0.42	0.65	0.49	0.87	0.61	0.88	0.58	0.60	0.44	0.45	0.46	0.42	0.42	0.80
RvineInd	0.84	0.63	0.43	0.65	0.50	0.83	0.66	0.89	0.54	0.59	0.46	0.49	0.46	0.44	0.37	0.75
Cvine	0.84	0.63	0.46	0.62	0.47	0.82	0.66	0.92	0.58	0.59	0.48	0.46	0.47	0.44	0.38	0.79
CvineInd	0.81	0.67	0.42	0.66	0.53	0.86	0.63	0.84	0.56	0.58	0.45	0.45	0.42	0.42	0.40	0.78
Dvine	0.83	0.65	0.41	0.60	0.52	0.85	0.63	0.86	0.55	0.57	0.48	0.48	0.47	0.48	0.40	0.81
DvineInd	0.81	0.61	0.50	0.65	0.47	0.82	0.63	0.87	0.56	0.58	0.48	0.47	0.42	0.48	0.39	0.80
Uni	0.57	0.37	0.79	0.49	0.52	0.59	0.65	0.71	0.23	1.37	0.77	0.55	0.41	0.50	0.52	1.20
Panel C. Test statistics for left half [0,0.5]																
Rvine	0.80	0.86	0.59	0.81	0.73	0.95	0.79	0.69	0.87	0.54	0.93	0.62	0.61	0.55	0.60	0.66
RvineInd	0.73	0.95	0.56	0.80	0.69	0.99	0.86	0.69	0.79	0.57	0.93	0.61	0.63	0.57	0.60	0.65
Cvine	0.77	0.87	0.63	0.81	0.74	0.95	0.84	0.65	0.84	0.55	0.93	0.64	0.65	0.56	0.60	0.67
CvineInd	0.82	0.87	0.61	0.73	0.66	1.01	0.84	0.65	0.84	0.55	0.92	0.58	0.64	0.53	0.60	0.66
Dvine	0.77	0.95	0.60	0.80	0.72	1.00	0.86	0.70	0.79	0.59	0.94	0.62	0.64	0.51	0.61	0.67
DvineInd	0.80	0.93	0.61	0.82	0.71	0.98	0.78	0.63	0.81	0.54	0.91	0.61	0.64	0.59	0.62	0.65
Uni	0.94	0.86	0.59	1.02	1.44	1.13	1.11	1.22	0.84	0.97	1.12	1.60	0.68	0.52	1.10	1.26
Panel D. Test statistics for right half [0.5,1]																
Rvine	0.81	0.64	0.48	0.65	0.77	0.87	0.61	0.88	0.58	0.60	0.63	0.45	0.67	0.67	0.60	0.80
RvineInd	0.84	0.63	0.43	0.65	0.78	0.86	0.66	0.89	0.54	0.59	0.70	0.49	0.64	0.62	0.57	0.75
Cvine	0.84	0.63	0.46	0.62	0.72	0.82	0.66	0.92	0.58	0.59	0.67	0.54	0.72	0.64	0.58	0.79
CvineInd	0.81	0.67	0.46	0.66	0.81	0.86	0.63	0.84	0.56	0.58	0.67	0.48	0.69	0.64	0.57	0.78
Dvine	0.83	0.65	0.45	0.60	0.78	0.85	0.63	0.86	0.56	0.57	0.62	0.48	0.63	0.64	0.59	0.81
DvineInd	0.81	0.61	0.50	0.65	0.77	0.82	0.63	0.87	0.56	0.58	0.61	0.50	0.70	0.61	0.65	0.80
Uni	0.81	0.76	0.97	0.90	1.58	0.59	1.28	0.80	0.50	1.37	0.77	1.75	0.97	0.57	0.73	1.24
Panel E. Test statistics for center [0.25,0.75]																
Rvine	0.61	0.49	0.59	0.61	0.77	0.80	0.51	0.56	0.87	0.53	0.93	0.57	0.67	0.67	0.60	0.61
RvineInd	0.65	0.54	0.56	0.57	0.78	0.86	0.49	0.51	0.79	0.53	0.93	0.48	0.64	0.62	0.57	0.58
Cvine	0.69	0.52	0.63	0.55	0.74	0.82	0.46	0.58	0.84	0.47	0.93	0.58	0.72	0.64	0.58	0.65
CvineInd	0.61	0.63	0.61	0.53	0.81	0.77	0.45	0.51	0.84	0.46	0.92	0.50	0.69	0.64	0.57	0.63
Dvine	0.72	0.65	0.60	0.60	0.78	0.84	0.45	0.51	0.79	0.46	0.94	0.53	0.63	0.64	0.59	0.63
DvineInd	0.71	0.58	0.61	0.56	0.77	0.78	0.47	0.53	0.81	0.43	0.91	0.52	0.70	0.61	0.65	0.62
Uni	0.81	0.78	0.97	0.95	1.58	0.57	1.28	1.09	0.57	0.93	1.11	1.75	0.97	0.57	0.73	1.24

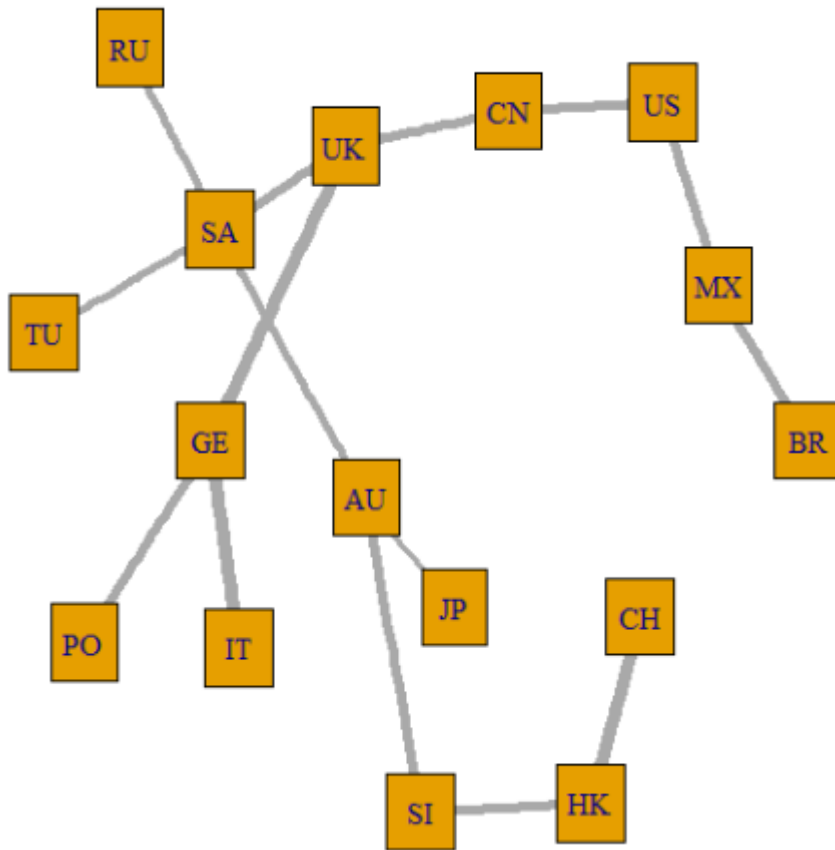
Note: Values in blue and red advocate rejection of the null hypothesis of correct specification at 10% and 5% significance level, respectively.

Table 6 Density forecast performance evaluation during global financial crisis and European debt crisis period (Aug 2007 ~ June 2012)

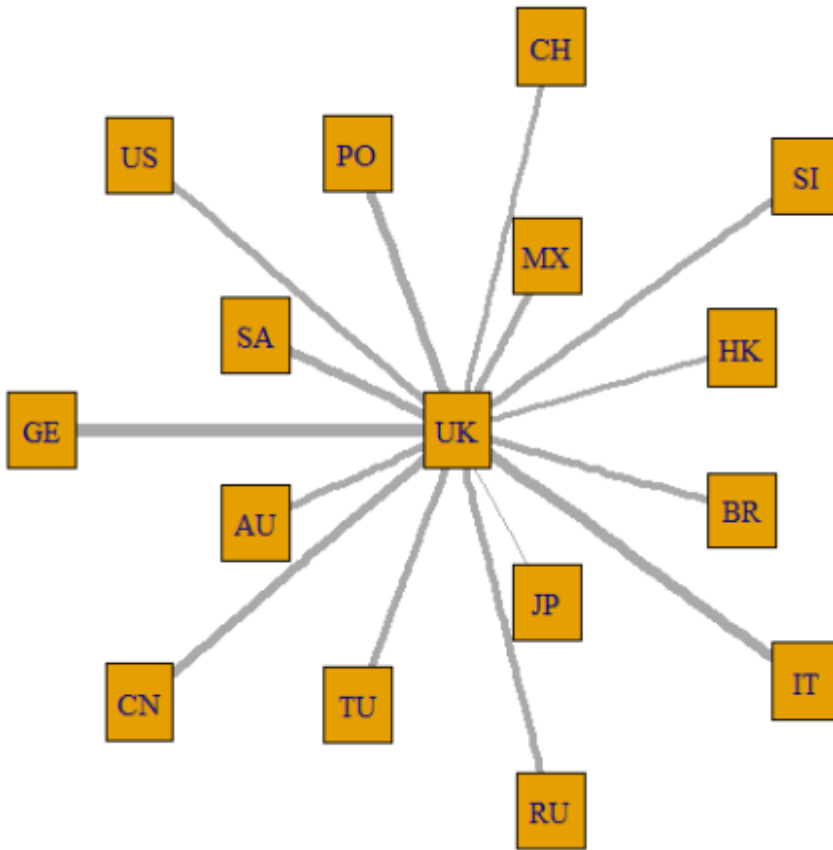
	IT	GE	JP	AU	CN	SI	UK	US	BR	CH	RU	SA	MX	PO	TU	HK
Panel A. Test statistics																
Rvine	1.29	1.14	0.74	0.82	0.84	0.73	1.12	1.01	0.71	0.71	0.69	0.59	0.59	0.83	0.48	0.74
RvineInd	1.36	1.15	0.77	0.82	0.86	0.78	1.14	0.97	0.82	0.70	0.73	0.60	0.60	0.89	0.53	0.66
Cvine	1.33	1.08	0.78	0.78	0.84	0.75	1.13	1.01	0.82	0.71	0.73	0.58	0.61	0.82	0.53	0.70
CvineInd	1.35	1.15	0.76	0.86	0.83	0.76	1.19	0.95	0.82	0.66	0.74	0.62	0.64	0.83	0.50	0.71
Dvine	1.39	1.10	0.74	0.82	0.84	0.77	1.12	0.99	0.79	0.70	0.69	0.59	0.59	0.83	0.56	0.69
DvineInd	1.36	1.14	0.75	0.82	0.87	0.74	1.17	1.01	0.82	0.72	0.73	0.61	0.62	0.84	0.54	0.72
Uni	1.06	0.97	1.00	0.87	1.32	0.93	1.32	1.36	0.74	0.94	0.94	1.04	0.71	0.73	0.69	1.14
Panel B. Test statistics for left tail [0,0.25]																
Rvine	1.25	1.14	0.58	0.75	0.83	0.72	1.12	0.89	0.71	0.65	0.69	0.59	0.59	0.83	0.48	0.67
RvineInd	1.36	1.15	0.55	0.81	0.80	0.72	1.14	0.90	0.82	0.68	0.73	0.60	0.60	0.89	0.53	0.63
Cvine	1.33	1.08	0.62	0.78	0.84	0.69	1.13	0.94	0.82	0.55	0.73	0.58	0.61	0.82	0.53	0.61
CvineInd	1.35	1.15	0.55	0.86	0.83	0.68	1.19	0.95	0.82	0.62	0.74	0.56	0.64	0.83	0.50	0.62
Dvine	1.36	1.10	0.57	0.81	0.84	0.70	1.12	0.93	0.79	0.63	0.69	0.59	0.58	0.83	0.56	0.64
DvineInd	1.36	1.14	0.58	0.78	0.86	0.74	1.17	0.89	0.82	0.65	0.73	0.61	0.62	0.84	0.51	0.61
Uni	1.06	0.97	0.53	0.87	1.18	0.93	0.76	1.36	0.74	0.94	0.94	0.77	0.70	0.55	0.68	1.05

Note: Values in blue and red advocate rejection of the null hypothesis of correct specification at 10% and 5% significance level, respectively.

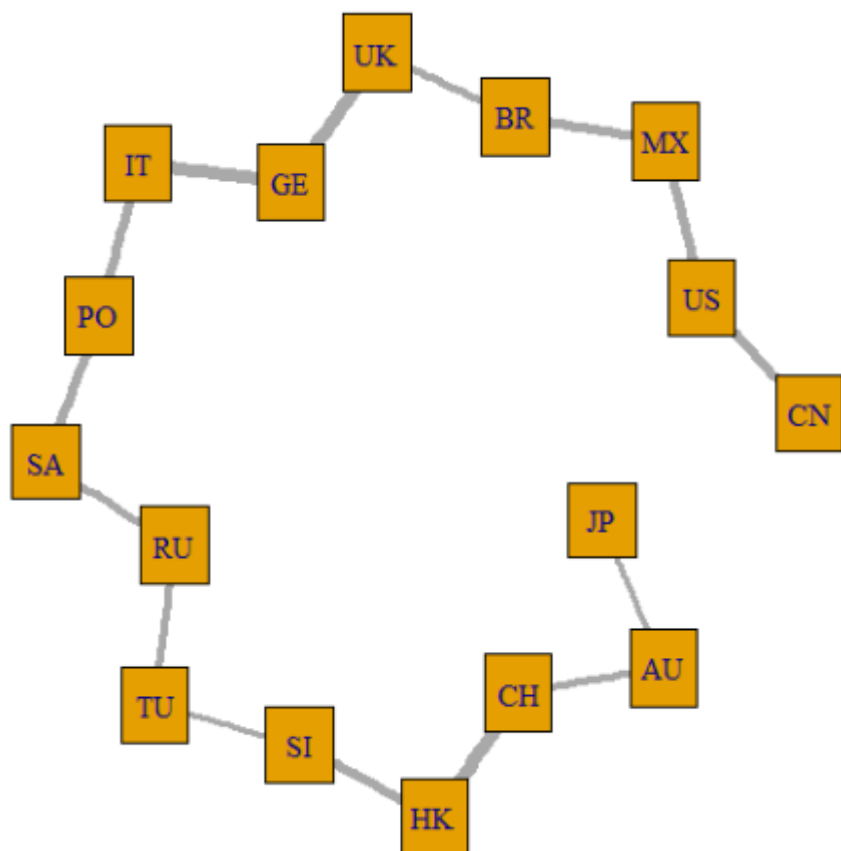
Figure 1 First level of R-Vine, C-Vine and D-Vine trees.



1(a) Regular Vine



1(b) Canonical Vine



1(c) Drawable Vine

Acknowledgement

Firstly, I want to thank Shigeyuki Hamori for his supervision and guidance. From Kobe University I thank everyone who has provided me comments, suggestions and help. This includes but is not limited to, Akio Namba, Xiaojing Cai, Takuji Kinkyo and Naoya Sueishi. Finally, I want to thank all of my family for their love and support.

Reference

- Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics* 44, 182–198.
- Admati, A., Pfleiderer, P., 1988. A theory of intraday patterns: volume and price variability. *Review of Financial Studies* 1, 3-40.
- Ait-Sahalia, Y., Brandt, M.W., 2001. Variable selection for portfolio choice. *Journal of Finance* 56, 1297-1351.
- Amihud, Y. 2002. Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets* 5, 31-56.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P. (2001). The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association*, 96, 42-55.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica*, 71, 579-625.
- Ang, A., Chen, J., 2002. Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 443-494. Bai, J., 2003. Testing parametric conditional distributions of dynamic models. *Review of Economics and Statistics* 85, 531-549.
- Bali, T.G., Mo, H., Tang, Y., 2008. The role of autoregressive conditional skewness and kurtosis in the estimation of conditional VaR. *Journal of Banking & Finance* 32, 269-282.
- Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2008. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica* 76, 1481–1536.

- Barndorff-Nielsen, O.E., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics* 2, 1–48.
- Bedford, T., Cooke, R.M., 2001. Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence* 32, 245–268.
- Bhattacharya, B., 2008. The extremal index and the maximum of a dependent stationary pulse load process observed above a high threshold. *Structural Safety* 30, 34–48.
- Bedford, T., Cooke, R.M., 2002. Vines—a new graphical model for dependent random variables. *Annals of Statistics* 30(4), 1031–1068.
- Brechmann, E.C., 2010. Truncated and simplified regular vines and their applications. Diploma Thesis, Technische University München.
- Brechmann, E.C., Czado, C., Aas, K., 2012. Truncated regular vines in high dimensions with applications to financial data. *Canadian Journal of Statistics* 40(1), 68–85.
- Brennan, M. J., Subrahmanyam, A., 1996. Market microstructure and asset pricing: on the compensation for illiquidity in stock returns. *Journal of Financial Economics* 41, 441–464.
- Calvet, L., Fisher, A. (2002). Multifractality in Asset Returns: Theory and Evidence. *Review of Economics and Statistics*, 84, 381–406.
- Calvet, L., Fisher, A. (2004). How to Forecast Long-run Volatility: Regime Switching and the Estimation of Multifractal Process. *Journal of Financial Econometrics*, 2, 49–83.
- Caporin, M., McAleer, M., 2014. Robust ranking of multivariate GARCH models by problem dimension. *Computational Statistics and Data Analysis* 76, 172–185.
- Carchano, O., Pardo, A., 2009. Rolling over stock index futures contracts. *Journal of Futures Markets* 29(7), 684–694.
- Catania, L., Bernardi, M., 2015. MCS: Model Confidence Set Procedure. R package version 0.1.1. <http://CRAN.R-project.org/package=MCS>.
- Christoffersen, P., 1998. Evaluating interval forecasts. *International Economic Review* 841–862.

- Christoffersen, P., Pelletier, D., 2004. Backtesting value-at-risk: A duration-based approach. *Journal of Financial Econometrics* 2(1), 84–108.
- Conrad, J., Dittmar, R.F., Ghysels, E. (2013). Ex Ante Skewness and Expected Stock Returns. *Journal of Finance*, Vol. 68, 1, 85-124.
- Corradi, V., Swanson, N.R., 2006. Bootstrap conditional distribution tests in the presence of dynamic misspecification. *Journal of Econometrics* 133, 779-806.
- De Haan, L., Resnick, S.I., Rootzén, H., De Vries, C.G., 1989. Extremal behavior of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Process and their Applications* 32, 213–224.
- Di Matteo, T. (2007). Multi-scaling in Finance. *Quantitative Finance*, 8, 21-36.
- Di Matteo, T., Aste, T., Dacorogna, M. (2005). Long-term Memories of Developed and Emerging Markets: Using the Scaling Analysis to Characterize their Stage of Development. *Journal of Banking & Finance*, 29, 827-851.
- Diebold, F.X., Gunther, T.A., Tay, A.S., 1998. Evaluating density forecasts with applications to financial risk management. *International Economic Review* 39(4), 863–883.
- Diebold, F.X., Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28, 57-66.
- Dißmann, J., Brechmann, E.C., Czado, C., Kurowicka, D., 2013. Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics and Data Analysis* 59, 52–69.
- Efimova, O., Serletis, A., 2014. Energy markets volatility modeling using GARCH. *Energy Economics* 43, 264-273.
- Ehlert, A., Fiebig, U.R., Janßen, A., Schlather, M., 2015. Joint extremal behavior of hidden and observable time series with applications to GARCH processes. *Extremes* 18:109-140.
- Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* Vol. 50, No. 4

- Engle, R.F., Gallo, G., 2006. A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics* 131, 3–27.
- Engle, R.F., Ng, V., 1993. Measuring and testing the impact of news on volatility. *Journal of Finance* 48, 1747–1778.
- Fernandez, C., Steel, M.F., 1998. On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association* 93(441), 359–371.
- Ferro, C., Segers, J., 2003. Inference for clusters of extreme values. *Journal of Royal Statistical Society Series B* 65, 545–556.
- Genest, C., Favre, A-C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering* 12(4), 347–368.
- Genest, C., Rémillard, B., 2004. Tests of independence and randomness based on the empirical copula process. *Test* 13(2), 335–369.
- Ghalanos, A., 2014. rugarch: Univariate GARCH models. R package version 1.3-4.
- Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between and expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779-1801.
- González-Rivera, G., Lee, T.H., Mishra, S., 2004. Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. *International Journal of Forecasting* 20(4), 629–645.
- Gonzalez-Rivera, G., Sun, Y., 2015. Generalized autocountours: Evaluation of multivariate density models. *International Journal of Forecasting* 31, 799-814.
- Gonzalez-Rivera, G., Yoldas, E., 2012. Autocontour-based evaluation of multivariate predictive densities. *International Journal of Forecasting* 28, 328-342.
- Gray, S.F., 1996. Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics* 42, 27–62.
- Hallam, M., Olmo, J. (2014). Semiparametric Density Forecasts of Daily Financial Returns from Intraday Data. *Journal of Financial Econometrics*, Vol. 12, 2, 408-432.

- Hansen, P.R., Horel, G., 2009. Quadratic variation by Markov chains. University of Aarhus Department of Economics Research Paper No. 13.
- Hansen, P.R., Lunde, A., Nason, J.M., 2011. The model confidence set. *Econometrica* 79(2), 453–497.
- Hansen, P.R., Huang, Z., Shek, H., 2012. Realized GARCH: A joint model of returns and realized measures of volatility. *Journal of Applied Econometrics* 27, 877–906.
- Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing. *Journal of Finance* 55, 1263-1295.
- Hasbrouck, J., Seppi, D., 2001. Common factors in prices, order flows, and liquidity, *Journal of Financial Economics*, 59, 383-411.
- Hofert, M., Kojadinovic, I., Maechler, M., Yan, J., 2015. copula: Multivariate Dependence with Copulas. R package version 0.999-13. URL <http://CRAN.R-project.org/package=copula>.
- Hong, Y., Li, H., 2005. Nonparametric specification testing for continuous- time models with applications to term structure of interest rates. *Review of Financial Studies* 18(1), 37-84.
- Joe, H., 1996. Families of m-variate distributions with given margins and $m(m - 1)/2$ bivariate dependence parameters. In: Rüschendorf, L., Schweizer, B., Taylor, M.D. (Eds.), *Distributions with Fixed Marginals and Related Topics*. Institute of Mathematical Statistics, Hayward, pp. 120–141.
- Jondeau, E., Rockinger, M., 2003. Conditional volatility, skewness, and skewness: existence, persistence, and comovements. *Journal of Economic Dynamics & Control* 27, 1699-1737.
- Kang, J., Lee, S. (2016) Is the information on the Higher Moments of Underlying Returns Correctly Reflected in Option Prices? *Journal of Futures Markets*, Vol. 36, 8, 722-744.
- Kojadinovic, I., Yan, J., 2011. Tests of serial independence for continuous multivariate time series based on a Möbius decomposition of the independence empirical copula process. *Annals of the Institute of Statistical Mathematics* 63, 347–373.
- Kupiec, P., 1995. Techniques for verifying the accuracy of risk management models. *Journal of Derivatives* 3, 73–84.

- Kurowicka, D., Cooke, R., 2006. Uncertainty Analysis with High Dimensional Dependence Modelling. Wiley, Chichester.
- Laurent, A., Rombouts, J.V.K., Violante, A.F., 2012. On the forecasting accuracy of Multivariate GARCH models. *Journal of Applied Econometrics* 27, 934-955.
- Laurini F , Tawn, J.A., 2012. The extremal index for GARCH(1,1) processes. *Extremes* 15, 511–529.
- Leon, A., Rubio, G., Serna, G., 2005. Autoregressive conditional volatility, skewness and kurtosis. *Quarterly Journal of Economics and Finance* 45, 599-618.
- Longin, F., Solnik, B., 2001. Extreme correlation of international equity markets. *Journal of Finance*, 649-676.
- Manner, H., 2007. Estimation and model selection of copulas with an application to exchange rates. METEOR Research Memorandum 07/056, Maastricht University.
- Matia, K., Ashkenazy, Y., Stanley, H. (2003). Multifractal Properties of Price Fluctuations of Stocks and Commodities. *Europhysics Letters*, 61, 422-428.
- Nagata, S., Oya, K., 2012. Volatility forecast comparison with biased proxy. Discussion paper.
- Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* 59, 347–370.
- Onali, E., Goddard, J. (2009). Unifractality and Multifractality in the Italian Stock Market. *International Review of Financial Analysis*, 18, 154-163.
- Ossandon, S., Bahamonde, N., 2011. On the nonlinear estimation of GARCH models using an Extended Kalman Filter. *Proceedings of World Congress on Engineering 2011 Vol I*.
- Patton, A.J., 2004. On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics* 2, 130-168.
- Patton, A.J, 2006. Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2), 527–556.

- Patton, A.J., 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* 160, 246–256.
- Prim, R.C., 1957. Shortest connection networks and some generalizations. *Bell System Technical Journal* 36, 1389–1401.
- R Core Team, 2015. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <http://www.R-project.org/>, ISBN 3-900051-07-0.
- Rossi, B, Sekhposyan, T., 2016. Alternative tests for correct specification of conditional predictive densities. Submitted for publication. Available at <http://www.tateviksekhposyan.org>.
- Sadorsky, P., 2006. Modeling and forecasting petroleum futures volatility. *Energy Economics* 28, 467-488.
- Shephard, N., Sheppard, K., 2010. Realising the future: Forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics* 25, 197–231.
- Shirota, S., Hizu, T., Omori, Y., 2014. Realized stochastic volatility with leverage and long memory. *Computational Statistics and Data Analysis* 76, 618–641.
- Sklar, A., 1959. Fonctions de Répartition à n Dimensions et Leurs Marges. *Publications de Institut Statistique de Université de Paris* 8, 229–231.
- Southworth, H., Heffernan, J.E., 2013. texmex: Statistical modelling of extreme values. R package version 2.1.
- Takahashi, M., Omori, Y., Watanabe, T., 2009. Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics and Data Analysis* 53, 2404–2406.
- Virbickaite, A., Ausin, M.C., Galeano, P., 2016. A Bayesian non-parametric approach to asymmetric dynamic conditional correlation model with application to portfolio selection. *Computational Statistics and Data Analysis* 100, 814-829.
- Vuong, Q.H., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57(2), 307–333.

Wang, Y., Wu, C., 2012. Forecasting energy market volatility using GARCH models: Can multivariate models beat univariate models? *Energy Economics* 34, 2167-2181.