Relation between Higher Order Comoments and Dependence Structure of Equity Portfolio*

Mario Cerrato[1], John Crosby[2], Minjoo Kim[1], and Yang Zhao[1]

[1]Adam Smith Business School, University of Glasgow

[2]Centre for Economics and Financial Studies, University of Glasgow

First version: 15 July 2013, this version: 25 September 2015

Abstract

We study a relation between higher order comoments and dependence structure of equity portfolio in the US and UK by relying on a simple portfolio approach where equity portfolios are sorted on the higher order comoments. We find that beta and coskewness are positively related with a copula correlation, whereas cokurtosis is negatively related with it. We also find that beta positively associates with an asymmetric tail dependence whilst coskewness negatively associates with it. Furthermore, two extreme equity portfolios sorted on the higher order comments are closely correlated and their dependence structure is strongly time-varying and nonlinear. Back-testing results of value-at-risk and expected shortfall demonstrate the importance of modeling a dynamic and asymmetric dependence in the risk management.

Key words: Higher order comoments, dependence structure, hyperbolic generalized skewed t copula, generalized autoregressive score, risk management.

JEL codes: C53, G17

*We would like to thank Andrew Patton, Paul Embrechts, Alexander McNeil, Drew Creal and Wolfgang Härdle for sharing their copula codes, and conference participants at 2015 Quant Europe in London, ICBI Global Derivatives conference in Amsterdam, and workshop participants at the University of Glasgow for helpful discussions and comments. All errors remain our own. Mario Cerrato, Email: Mario.Cerrato@glasgow.ac.uk. John Crosby, Email: johnc2205@yahoo.com. Minjoo Kim, Email: Minjoo.Kim@glasgow.ac.uk. Yang Zhao, Email: y.zhao.1@research.gla.ac.uk.

1 Introduction

The Fama-French factor model (Fama and French, 1993) is a monumental turning point in the modern asset pricing literature. Recently, Christoffersen and Langlois (2013) study how an extreme dependence structure associates with the Fama-French factors and address its role in broad area of finance. They also emphasize the importance of copula modeling for the extreme dependence structure. On the other hand, there is a group of researchers supporting the importance of higher order comoments in asset pricing (Harvey and Siddique, 2000; Dittmar, 2002; Bakshi et al., 2003; Ang et al., 2006; Guidolin and Timmermann, 2008; Conrad et al., 2013). Although those are less popular than the Fama-French factors among practitioners, those have been rigorously developed from theoretical perspectives. Hence, it is academically interesting to study how the extreme dependence structure is related with the higher order comoments and address their implications in finance.

A few papers address a relation between the higher order comoments and the tail dependence of equity portfolio. They show that it has a close relationship with, not only beta, but also coskewness. For example, Garcia and Tsafack (2011) show that a strong dependence in lower returns creates a large negative coskewness in their international bond and equity market portfolio analysis. Chabi-Yo et al. (2014) also show that a strong lower tail dependence creates a large negative coskewness. In addition they show that beta is monotonically increasing with respect to the lower tail dependence. From these studies, we are able to draw an inference that the tail dependence is a key driver to create the higher order comoments of the equity portfolio. Thus our first research question is how the higher order comoments associate with the dependence structure of the equity portfolio. We approach our research question by relying on a simple portfolio approach. Specifically, we sort equities into portfolios based on the size of the higher order comoment, i.e., from low beta (coskewness, cokurtosis) to high beta (coskewness, cokurtosis), and test patterns of a copula correlation or an asymmetric dependence across the characteristic-sorted portfolios.

We find that there are statistically significant patterns between the higher order comoments and the dependence structure of the equity portfolio. First, beta and coskewness are positively related with the copula correlation whilst cokurtosis is negatively related with it. Second, we find the asymmetry that the lower tail dependence is stronger than the upper tail dependence for all portfolios. Third, beta is positively related with the asymmetric tail dependence, whereas coskewness is negatively related with it.

Our second research question is what economic implication is contained by the relation between higher order comoments and the dependence structure of the equity portfolio. We find its implication from risk management perspectives. We often long and short two extreme portfolios to hedge their risk. Thus the higher order comoment risk can be also hedged by buying and selling two extreme beta (coskewness, cokurtosis) portfolios, i.e. **B**uying **M**inus **S**elling (BMS) portfolio. However, if our inference is correct in the first research question, the higher order comoments are unable to be key inputs for the risk management of extreme events. Rather, a key driver is the tail dependence which creates the higher order comoments. To investigate our second research question, we apply backtesting tools to alternative models: dynamic copula models, multivariate GARCH model and univariate model. The dynamic copula models fully incorporate the dependence structure of two extreme portfolios whilst the multivariate GARCH model takes into account only the second order comoment. The univariate model considers neither the tail dependence nor the second order comoment.

The backtesting results strongly support the importance of modeling the time-varying and asymmetric dependence of the BMS portfolio. First, we find that the dependence structure of the BMS portfolio is strong, time-varying and asymmetric for all characteristic-sorted portfolios. Second, both the multivariate GARCH model and the univariate model significantly underforecast value-at-risk (VaR) and expected shortfall (ES). Third, the dynamic copula models show not only robust coverage ability but also statistical accuracy for VaR and ES.

Besides two important research questions, we develop a generalized dynamic asymmetric copula. Our proposed model takes into account two important characteristics of equity portfolios; a time-varying dependence and an asymmetric tail dependence. First, we employ a generalized hyperbolic skewed t distribution (see Demarta and McNeil, 2005) to capture the asymmetric dependence structure. Second, the time-varying copula correlation is implied by the generalized autoregressive score (Creal et al., 2013). Hence, our proposed model can cover for the most types of the dependence structure revealed by the equity portfolios. We apply our copula to estimating the dependence structure in our analysis.

Our study makes three contributions. First, we provide comprehensive analysis on the relation between higher order comoments and the dependence structure of the equity portfolio. We find the striking evidence that the higher order comoments are closely related with the dependence structure of the equity portfolio in the US and UK. Second, we demonstrate the importance of modeling the time-varying and asymmetric dependence of the BMS portfolio in the risk management of extreme events. The backtesting results show that the ignorance of dependence asymmetry and dynamics is costly in the risk management. Third, we propose the generalized dynamic asymmetric copula by combining the generalized hyperbolic skewed t distribution and the generalized autoregressive score. Our proposed copula performs well in estimating the dependence structure of the BMS portfolio and forecasting both VaR and ES.

The remainder of this paper is organized as follows. In Section 2, we detail the way we employ for the portfolio construction and the dynamic asymmetric copula we propose. The data used in the paper and the descriptive statistics are in Section 3. In Section 4, we focus on the analysis of the relation between the higher order comoments and the dependence structure. In Section 5, we analyze the role of the dependence structure of the BMS portfolio in the forecasting based risk management application. Finally, conclusions are given in Section 6.

2 Methodology

In this section, we detail a way we employ for the portfolio construction and models we use in this paper.

2.1 Portfolio Construction

A return on an asset is defined as the first difference of the log price, $r_t = \log P_t - \log P_{t-1}$. We construct portfolios sorted on beta, coskewness and cokurtosis, respectively. Following the definition of Bakshi et al. (2003) and Conrad et al. (2013), we define the market beta, coskewness and cokurtosis by

$$BETA_{i,t} = \frac{\mathbb{E}\left[\left(r_{i,t} - \mathbb{E}\left[r_{i,t}\right]\right)\left(r_{m,t} - \mathbb{E}\left[r_{m,t}\right]\right)\right]}{Var\left(r_{m,t}\right)}, \tag{1}$$

$$COSK_{i,t} = \frac{\mathbb{E}\left[\left(r_{i,t} - \mathbb{E}\left[r_{i,t}\right]\right)\left(r_{m,t} - \mathbb{E}\left[r_{m,t}\right]\right)^2\right]}{\sqrt{Var\left(r_{i,t}\right)}Var\left(r_{m,t}\right)}, \tag{2}$$

$$COKT_{i,t} = \frac{\mathbb{E}\left[\left(r_{i,t} - \mathbb{E}\left[r_{i,t}\right]\right)\left(r_{m,t} - \mathbb{E}\left[r_{m,t}\right]\right)^3\right]}{Var\left(r_{i,t}\right)Var\left(r_{m,t}\right)}. \tag{3}$$

All stocks are sorted on each characteristic above and divided into five groups based on the 20th, 40th, 60th and 80th percentiles. We estimate beta, coskewness and cokurtosis each year using all the daily data within the year. Then, we annually rebalance portfolios, value weighted based on the capitalization of each stock.[1] We denote by BETA1 (COSK1, COKT1) the portfolio formed by stocks with the lowest beta (respectively, coskewness, cokurtosis), and BETA5 (COSK5, COKT5) denotes the portfolio formed by stocks with the highest beta (coskewness, cokurtosis).

2.2 Modeling Marginal Density

We allow each portfolio return series to have time-varying conditional mean ($\mu_{i,t}$) and variance ($\sigma_{i,t}^2$), and we also assume that the standardized returns $z_{i,t} = \left(r_{i,t} - \mu_{i,t}\right)/\sigma_{i,t}$ are identically distributed. We fit an AR model to the conditional mean

$$r_{i,t} = c_i + \sum_{k=1}^{p} \phi_{i,k} r_{i,t-k} + \varepsilon_{i,t}, \text{ where } \varepsilon_{i,t} = \sigma_{i,t} z_{i,t} \tag{4}$$

and an asymmetric GARCH model, namely GJR-GARCH(1,1,1) (see Glosten et al., 1993), to the conditional variance

$$\sigma_{i,t}^2 = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2 + \gamma_i \varepsilon_{i,t-1}^2 I_{i,t-1} \tag{5}$$

where $I_{i,t-1} = 1$ if $\epsilon_{i,t-1} < 0$, and $I_{i,t-1} = 0$ if $\epsilon_{i,t-1} \geq 0$.

Let $z_{i,t}$ be a random variable with a continuous distribution F_i. For the parametric model, we

[1] We compute the market capitalization of each company (stock price multiplied by the number of shares outstanding) and then use it to assign weights.

assume that $z_{i,t}$ follows the skewed Student's t distribution of Hansen (1994):

$$z_{i,t} \sim F_{skew-t,i}\left(\eta_i, \lambda_i\right), \; u_{i,t} = F_{skew-t,i}\left(z_{i,t}; \eta_i, \lambda_i\right) \tag{6}$$

where $F_{skew-t,i}$ denotes the cumulative distribution function, η_i denotes the degrees of freedom, λ_i the skewness parameter, and $u_{i,t}$ the probability integral transformation. Hence, we can easily compute the probability given the estimates of parameters; $\hat{\mu}_{i,t}$, $\hat{\sigma}_{i,t}$, $\hat{\eta}_i$ and $\hat{\lambda}_i$. For the nonparametric model, we use the empirical distribution function to obtain the estimate of F_i:

$$\hat{F}_i\left(z\right) \equiv \frac{1}{T+1}\sum_{t=1}^{T} 1\left\{\hat{z}_{i,t} \leq z\right\}, \; \hat{u}_{i,t} = \hat{F}_i\left(\hat{z}_{i,t}\right). \tag{7}$$

We estimate all parameters in (5) – (6) using the maximum likelihood estimation. Then we generate each marginal density parametrically or nonparametrically for the purpose of copula construction.

2.3 Generalized Hyperbolic Skewed t Copulas

In this section, we provide a brief introduction to the generalized hyperbolic skewed t (GHST) distribution which we employ to capture asymmetric extreme dependence structure between equity portfolios in our study. It belongs to the class of multivariate normal variance mixtures and has the stochastic representation

$$\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\gamma} W + \sqrt{W}\mathbf{Z} \tag{8}$$

for a d-dimensional parameter vector $\boldsymbol{\gamma}$. Further, W is a scalar valued random variable following an inverse gamma distribution $W \sim IG\left(\nu/2, \nu/2\right)$ and $\mathbf{Z}$ is a d- dimensional random vector following a normal distribution $\mathbf{Z} \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}\right)$ and is independent of W (see Demarta and McNeil, 2005).

The density function of multivariate GHST distribution is given by

$$\mathbf{f}_{skt}\left(\mathbf{z}; \gamma, \nu, \Sigma\right) = \frac{2^{\frac{2-(\nu+d)}{2}} K_{\frac{\nu+d}{2}}\left(\sqrt{\left(\nu + \mathbf{z}^{*\prime}\Sigma^{-1}\mathbf{z}^{*}\right)\gamma^{\prime}\Sigma^{-1}\gamma}\right) e^{\mathbf{z}^{*\prime}\Sigma^{-1}\gamma}}{\Gamma\left(\frac{\nu}{2}\right)\left(\pi\nu\right)^{\frac{d}{2}}\left|\Sigma\right|^{\frac{1}{2}}\left(\nu + \mathbf{z}^{*\prime}\Sigma^{-1}\mathbf{z}^{*}\right)^{\frac{-\nu+d}{2}}\left(1+\frac{1}{\nu}\mathbf{z}^{*\prime}\Sigma^{-1}\mathbf{z}^{*}\right)^{\frac{-\nu+d}{2}}} \tag{9}$$

where K_λ, ν and γ denote the modified Bessel function of the third kind, the degree of freedom and skewed parameter vector, respectively. The density of multivariate converges to the conventional symmetric t density when γ tends to $\mathbf{0}$. For the parametric case, we define the shocks $z^*_{i,t} = F^{-1}_{skt,i}(u_{i,t}) = F^{-1}_{skt,i}(F_{skew-t,i}(z_{i,t}))$ where $F^{-1}_{skt,i}(u_{i,t})$ denotes the inverse cumulative distribution function of the univariate GHST distribution and it is not known in closed form but can be well approximated via simulation. $F_{skew-t,i}$ denotes the cumulative distribution function of skewed t distribution in Hansen (1994). Note that we use $z^*_{i,t}$ not the standardized return $z_{i,t}$. For the nonparametric case, we use the EDF to obtain the estimate of $u_{i,t}$. A more detailed discussion can be found in Christoffersen et al. (2012).

The probability density function of the GHST copula defined from above multivariate GHST density of Eq. (9) is given by

$$c_{skt}(\mathbf{z};\gamma,\nu,\Sigma) = \frac{2^{\frac{(\nu-2)(d-1)}{2}} K_{\frac{\nu+d}{2}}\left(\sqrt{\left(\nu+\mathbf{z}^{*\prime}\Sigma_t^{-1}\mathbf{z}^*\right)\gamma'\Sigma_t^{-1}\gamma}\right) e^{\mathbf{z}^{*\prime}\Sigma_t^{-1}\gamma}}{\Gamma\left(\frac{\nu}{2}\right)|\Sigma|^{\frac{1}{2}}\left(\nu+\mathbf{z}^{*\prime}\Sigma_t^{-1}\mathbf{z}^*\right)^{\frac{-\nu+d}{2}}\left(1+\frac{1}{\nu}\mathbf{z}^{*\prime}\Sigma_t^{-1}\mathbf{z}^*\right)^{\frac{-\nu+d}{2}}} \times \prod_{i=1}^{d} \frac{\left(\sqrt{\left(\nu+(z_i^*)^2\right)\gamma_i^2}\right)^{-\frac{\nu+1}{2}}\left(1+\frac{1}{\nu}(z_i^*)^2\right)^{\frac{\nu+1}{2}}}{K_{\frac{\nu+1}{2}}\left(\sqrt{\left(\nu+(z_i^*)^2\right)\gamma_i^2}\right) e^{z_i^*\gamma_i}} \tag{10}$$

where Σ_t is the time-varying covariance matrix. Specifically, $\Sigma_t = \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t$, where $\mathbf{D}_t$ is an identity matrix in copula modeling and $\mathbf{R}_t$ is the time-varying correlation matrix. Note that Christoffersen et al. (2012) applied the GHST copula by constraining all the margins to have the same asymmetry parameter. Different from their model, our model consider a more generalized case by allowing the copula to have the different asymmetry parameters across margins. Although our model can be used for high-dimensional copula modeling, in this paper, only the bivariate case is considered as modeling the dependence and market risk of the BMS portfolio is our main task.

2.4 Generalized Autoregressive Score Model

We estimate the dynamic copula model based on the Generalized Autoregressive Score (GAS) model of Creal et al. (2013). We assume that a correlation parameter δ_t is dynamic and updated as function of its own lagged value. For example, the copula correlation is a scalar for the bivariate case and can be obtained from

$$\mathbf{R}_t = \begin{bmatrix} 1 & \delta_t \\ \delta_t & 1 \end{bmatrix}. \tag{11}$$

To make sure that it always lies in a pre-determined range, e.g. $\delta_t \in (-1, 1)$, the GAS model utilizes a strictly increasing transformation. Following Patton (2012), the transformed correlation parameter is denoted by g_t:

$$g_t = h(\delta_t) \Leftrightarrow \delta_t = h^{-1}(g_t), \tag{12}$$

where $\delta_t = (1 - e^{-g_t}) / (1 + e^{-g_t})$. Further, the updated transformed parameter g_{t+1} is a function of a constant $\bar{\omega}$, the lagged transformed parameter g_t, and the standardized score of the copula log-likelihood $Q_t^{-1/2}\mathbf{s}_t$:

$$g_{t+1} = \bar{\omega} + \eta Q_t^{-1/2}\mathbf{s}_t + \varphi g_t, \tag{13}$$

where $\mathbf{s}_t \equiv \partial \log c(u_{i,t}, u_{j,t}; \delta_t) / \partial \delta_t$ and $Q_t \equiv \mathbb{E}_{t-1}[\mathbf{s}_t \mathbf{s}_t']$.

Since the GAS model is an observation driven model, we can estimate the parameters by maximum likelihood estimation

$$\widehat{\delta}_t = \underset{\delta_t}{\operatorname{argmax}} \sum_{t=1}^{n} \log c(u_{i,t}, u_{j,t}; \delta_t). \tag{14}$$

The dynamic copulas are parametrically estimated using maximum likelihood estimation. When the marginal distributions are estimated using the skewed t distribution, the resulting joint distribution is fully parametric. When the marginal distribution is estimated by the empirical distribution function, then the resulting joint distribution is semiparametric. More details can be found in the Appendix A.1 and A.2.

2.5 Monotonicity Test

We test a monotonic pattern between the higher order comoments and the dependence structure of the equity portfolio using a monotonicity test proposed by Patton and Timmermann (2010). It tests whether there is a significantly increasing or decreasing pattern of average dependence measure such as a copula correlation or an asymmetric tail dependence when moving from the portfolio of low higher order comoment (P1) to the one with high higher order comoment (P5).

There are two types of monotonicity tests. One is "MR" test and the other is "UP (Down)" test. The MR test statistic tests for a monotonically increasing dependence. The UP (Down) test is less restrictive and simple. It tests for a generally increasing (decreasing) pattern without requiring the monotonicity of the average dependence measure. Let d_i denote an average dependence measure for a portfolio i. Then the MR test requires that $d_1 < d_2 < \cdots < d_5$ for a monotonically increasing pattern. It formulates the null hypothesis against the alternative one as

$$H_0 : \Delta \leq 0 \qquad \text{against} \qquad H_1 : \min_{i=1,\ldots,4} \Delta_i > 0, \tag{15}$$

wehre Δ is a vector of differences in adjacent average dependences, $(d_2 - d_1, d_3 - d_2, d_4 - d_3, d_5 - d_4)$, and Δ_i is the ith element of Δ. The Up test formulates the null hypothesis of a flat pattern against the alternative hypothesis as

$$H_0 : \Delta = 0 \qquad \text{against} \qquad H_1 : \sum_{i=1}^{4} |\Delta_i| \, 1 \{\Delta_i > 0\} > 0. \tag{16}$$

The Down test follows in an analogous way.

The choice of test statistics for MR, Up and Down are

$$\text{MR: } J_T = \min_{i=1,\ldots,4} \Delta_i, \tag{17}$$

$$\text{Up: } J_T^+ = \sum_{i=1}^{4} |\Delta_i| \, 1 \{\Delta_i < 0\}, \tag{18}$$

$$\text{Down: } J_T^- = \sum_{i=1}^{4} |\Delta_i| \, 1 \{\Delta_i > 0\}. \tag{19}$$

Each test statistic does not have a standard limiting distribution under the null hypothesis, but critical values or p-value can be obtained using a bootstrap approach.

3 Data Sources and Sample Construction

Stock prices are obtained from Datastream. Daily returns of the 500 stocks listed in the S&P 500 and those of the 100 stocks listed in FTSE 100 are used to construct portfolios. Our data, spanning the period of global financial crisis of 2007-2009 and European sovereign debt crisis of 2010-2011, go from January 4, 2000 to December 31, 2012, resulting in 3,268 daily observations for each stock in US and 3,283 daily observations for each stock in UK.

Given the one-year estimation period, we estimate beta, coskewness and cokurtosis using daily data (250 days) for each stock.[2] We rank securities by the estimates of beta (coskewness, cokurtosis) and form into five portfolios, lowest (1st) – highest (5th). Then we calculate daily returns for each portfolio within the estimation period.[3] In this way, we construct fifteen different portfolios for each market. The fifteen portfolios consist of one for each of the three characteristics (beta, coskewness and cokurtosis), divided into five portfolios. We annually rebalance all the portfolios and calculate 12-month daily returns.[4] We skip over presenting descriptive statistics for all portfolios since those have been already reported by many literature.[5]

4 Higher Order Comoments and Dependence Structure

In this section, we investigate a relation between the higher order comoments and the dependence structure of the equity portfolio using a simple portfolio approach. We employ two measures to describe the dependence structure. First, we measure a general dependence between the characteristic-sorted equity portfolio and the market by a copula correlation. Second, we measure the magnitude of asymmetric tail dependence by differencing lower tail dependence and upper tail dependences. We estimate them using our proposed dynamic asymmetric copula model in which the GHST copula takes into account the asymmetric nature of the dependence structure and GAS embodies the time-varying nature of the dependence structure.

4.1 Copula Correlation

Figure 1 plots an average copula correlations for equity portfolios sorted on the higher order comoments. We find the increasing patterns of average copula correlations when moving from the low beta portfolio (BETA1) to the high beta portfolio (BETA5). In particular, the UK stock market shows the monotonically increasing pattern. We also find that the average copula correlations generally increase when moving from the low coskewness portfolio (COSK1) to the high coskewness portfolio

[2]Since we estimate a factor beta on the daily return, we use a short sample period. We also consider alternative longer estimation periods (3 and 5 years) and find consistent results with a one year estimation period.

[3]We also calculate daily reruns for the next 12-months, which are forward looking portfolio returns, and find similar forecasting results. Since we are interested in how higher-order comoments are related with the extreme dependence structure, we prefer portfolio returns calculated within the estimation period to forward looking returns.

[4]We also consider monthly rebalancing of portfolios and find results consistent with annual rebalancing.

[5]The descriptive statistics are available upon request from author.

(COSK5). In contrast, there is the decreasing pattern in the UK stock market when moving from the low cokurtosis portfolio (COKT1) to the high cokurtosis portfolio (COKT5). However, we find no notable pattern for cokurtosis in the US stock market.

[INSERT FIGURE 1 ABOUT HERE]

We formally test the increasing or decreasing pattern of copula correlation in Table 1 using the monotonicity test. Panel A reports test results for the US stock market. Since the MR statistics are not rejected for all portfolios, there are no significant monotonic patterns. We however find some significant patterns under less restrictive conditions. The UP statistics are rejected for beta and coskewness. Thus there is the significant increasing pattern when moving from BETA1 (COSK1) to BETA5 (COSK5). Panel B reports test results for the UK stock market. The MR statistic is rejected only for beta. There is thus the significant monotonically increasing pattern when moving from BETA1 to BETA5. We also find significant patterns for coskewness and cokurtosis under less restrictive conditions. The UP statistic is rejected for coskewness and the DOWN statistic is rejected for cokurtosis. Thus there is the significant increasing (decreasing) pattern when moving from COSK1 (COKT1) to COSK5 (COKT5). Overall, the statistical evidences are consistent with the previous descriptive evidences.

[INSERT TABLE 1 ABOUT HERE]

4.2 Asymmetric Tail Dependence

Figure 2 plots the average difference between lower tail dependence (LTD) and upper tail dependence (UTD) for equity portfolios sorted on higher order comoments. We calculate both tail dependence coefficients by the parametric approach of McNeil et al. (2005). See Appendix A.3. for details. First, we find the asymmetry that the average LTD is stronger than the average UTD for all portfolios. Second, beta is positively related with the asymmetric tail dependence. There is the increasing pattern of the average tail difference when moving from BETA1 to BETA5. Hence, the more sensitive the portfolio is to market changes, the more sensitively investors tend to react to the market downturn. This tendency might create the stronger dependence in the lower tail. Third, in contrast, coskewness is negatively related with the asymmetric tail dependence. There is the decreasing pattern when moving from COSK1 to COSK5. Investors would prefer a positive coskewness since it represents a higher probability of extreme positive returns in the portfolio over market returns. Hence, when the portfolio returns are positively coskewed over market returns, investors tend to react less sensitively to market changes.

[INSERT FIGURE 2 ABOUT HERE]

We also formally test the monotonic increasing or decreasing pattern of the average difference in Table 2. The MR statistics are rejected for all portfolios in both stock markets. We find that the Up statistic is not rejected for beta in the UK stock market. Thus there is the significant increasing pattern of the asymmetry for the beta portfolio. We also find that the Down statistic is not rejected for coskewness in the US stock market. Thus there is the significant decreasing pattern of the asymmetry for the coskewness portfolio. Overall, the statistical test results are consistent with the descriptive evidences.

[INSERT TABLE 2 ABOUT HERE]

5 Risk Management Application

We find that there are the significant relations between the higher order comoments and the dependence structure of the equity portfolio in section 4. The evidences imply that the higher order comoments would not be key inputs for the portfolio risk management of extreme events since there are unexplained information for the tail dependence by the higher order comoments. Thus we investigate its implication from the risk management perspectives in this section.

We often long and short two extreme portfolios to hedge its risk. Following this simple strategy, we buy the highest beta (coskewness, cokurtosis) portfolio and sell the lowest beta (coskewness, cokurtosis) one to construct a BMS portfolio,

$$r_{bms,t} = r_{h,t} - r_{l,t}, \tag{20}$$

where $r_{h,t}$ and $r_{l,t}$ denote returns from the highest beta (coskewness, cokurtosis) portfolio and the lowest beta (coskewness, cokurtosis) one, respectively. Note that Table 3 defines several BMS portfolios.

[INSERT TABLE 3 ABOUT HERE]

As demonstrated in the section 4, the higher order comoments are closely related to the extreme dependence structure. Hence, we can expect that two extreme portfolios of the BMS portfolio are able to create a strong extreme dependence structure which should be taken into account in the risk modeling.

5.1 Diagnosis of BMS Portfolio

We first investigate the characteristics of BMS portfolio returns. We look at not only univariate characteristics but also multivariate ones. We get a clue to the modeling of the BMS portfolio returns for VaR and ES from this diagnosis.

5.1.1 Marginal Distribution

Before modeling the joint distribution of portfolio returns, it is necessary to select a suitable model for the marginal return distribution, because the misspecification of the univariate model can lead to biased copula parameter estimates. To allow for autocorrelation, heteroskedasticity and asymmetry, we use the models introduced in Section 2.2.

We estimate model parameters using maximum-likelihood estimation (MLE). The results of AR and GARCH estimations are presented in Table 4. For each portfolio return series, the variance persistence implied by the model is close to 1. For all the series, leverage effect parameters γ are significantly positive implying that a negative return on the series increases volatility more than a positive return with the same magnitude.

The obvious skewness and high kurtosis of returns lead us to consider the skewed Student's t distribution of Hansen (1994) for modeling residuals. To evaluate the goodness-of-fit for the skewed Student's t distribution, the Kolmogorov-Smirnov (KS) and Cramer-von Mises (CvM) tests are implemented and the p-values from these two tests are reported in Table 4.[6] Our results suggest that

[6]The p-values are obtained based on the algorithm suggested in Patton (2012)

the skewed Student's t distribution is suitable for modeling residuals. Thus, in general, the diagnosis provides evidences that our marginal distribution models are well-specified and therefore, we can reliably use the combination of AR, GARCH and skewed Student's t distribution, allied to copulas to model the dependence structure.

[INSERT TABLE 4 ABOUT HERE]

5.1.2 Time Varying Dependence

There is the considerable evidence that the conditional mean, volatility and covariance of financial time series are time-varying. This, possibly, suggests the reasonable inference that the dependence structure may also change over time.

We now consider three tests for time-varying dependence. The first one is a naïve test for a break in rank correlation at specified points in the sample, see Patton (2006). A noticeable limitation of this test is that the break point of dependence structure (e.g. a specified date) must be known a priori. The second test for time-varying dependence allows for a break in the rank correlation coefficient at some prior unspecified date, see Andrews (1993). The third test is the ARCH LM test for time-varying volatility, see Engle (1982). The critical values for the test statistic can be obtained by using a iid bootstrap algorithm, see Patton (2012). The results of the above tests for time-varying dependence are summarized in Table 5. Suppose that there is no priori date for the timing of break, we first consider naïve tests for the break at three chosen points in our sample, at $t^*/T \in \{0.15, 0.50, 0.85\}$, which corresponds to the dates 10-Dec-2001, 03-Jul-2006, and 17-Jan-2011. Then we consider another test in Andrews (1993) for a dependence break of unknown timing. As can be seen from Table 5, for almost all the equity portfolios, the p-value is significant at the 5% significance level showing clear evidence against a constant rank correlation with a one-time break. To detect whether the dependence structures between the high and low portfolios significantly changed during the global financial crisis of 2007-2009 and the European sovereign debt crisis of 2010-2011, we use 15-Sep-2008 (the collapse of Lehman Brothers) and 01-Jan-2010 (EU sovereign debt crisis) as two break points. We find that the dependence between BETA1 and BETA5 significantly changed around those dates, as all the p-values are fairly small. For other portfolio pairs, time homogeneity of the dependence structure is rejected by at least one test.

[INSERT TABLE 5 ABOUT HERE]

Overall, we find the evidence against the time homogeneity of the dependence structure between the standardized residuals of portfolios. This result shows that the standard portfolio diversification and risk management techniques based on constant correlations (or dependence) are inadequate, especially during financial crisis. There are considerable evidences that the time-varying nature is most closely related to the performance of risk forecast.

5.1.3 Asymmetric Tail Dependence

Standard models fail to take into account a noteworthy feature during financial crisis that asset returns often become more highly correlated (in magnitude). To test for the presence of this feature, we use threshold correlations. We find that the lower threshold correlations are always greater than the upper threshold correlation indicating that portfolios are more correlated when both of them perform poorly. To find out whether this asymmetry is statistically significant, we perform the symmetry tests of Hong et al. (2007). Table 6 reports the test results and shows that, as measured by threshold correlation, half of the portfolios are significantly asymmetric: BMS(Beta,US/UK) and BMS(Cokt,UK).

[INSERT TABLE 6 ABOUT HERE]

Although the threshold correlation offers some insights, it is still based on (linear) correlation and, therefore, does not take into account nonlinear information. To capture nonlinear dependence, we consider a copula-based tail dependence. Compared with (linear) correlation, the key advantage of copulas is that they are a "pure measure" of dependence, which cannot be affected by the marginal distributions (see Nelsen, 2006). Table 7 reports the coefficients of lower tail dependence (LTD) and upper tail dependence (UTD) and the difference between them. The coefficients are estimated using McNeil et al. (2005) with the Student's t copula. For example, the LTD coefficient for BETA1 and BETA5 in the US equity market is 0.171 and the UTD coefficient is 0.018. Then we find the significant difference between the UTD and LTD coefficients. In the UK equity market, we also find the evidence of asymmetric dependence in that all the portfolio pairs exhibit greater correlation during market downturns than market upturns.

[INSERT TABLE 7 ABOUT HERE]

This finding about the asymmetric dependence between the high beta (coskewness, cokurtosis) portfolio and the low beta (coskewness, cokurtosis) portfolio is possibly associated with the fact that investors have more uncertainty about the economy, and therefore pessimism and panic spread from one place to another more quickly during market downturns. Another possible explanation is the impact of liquidity risk. Some "uncorrelated" liquid assets suddenly become illiquid during market downturns, and, therefore, even a small trading volume can lead to huge co-movements.

5.2 Backtesting of Value-at-Risk

The diagnosis shows that forecasting models based on a constant dependence or a symmetric dependence are inadequate, especially during the financial crisis. Thus the time varying and asymmetric dependence structure between two extreme portfolios provides us a strong motivation to introduce a dynamic asymmetric copula model for forecasting portfolio VaR and ES defined by

$$VaR_{bms,t}(\alpha) = \inf\left\{x \mid \mathbb{P}\left(r_{bms,t} \le x | \mathcal{F}_{t-1}\right) \le 1-\alpha\right\}, \tag{21}$$

$$ES_{bms,t}(\alpha) = \mathbb{E}\left[r_{bms,t} \mid r_{bms,t} < VaR_{bms,t}(\alpha)\right], \tag{22}$$

where $\mathcal{F}_{t-1}$ represents the information set available at $t-1$. In our study, α is assumed to be either 0.95 or 0.99, and we report results focusing on 0.99 which is the most widely used value for market risk management. Once the dynamic copula parameters have been estimated, Monte Carlo simulation is used to generate 5000 values of $r_{h,t}^{(s)}$ and $r_{l,t}^{(s)}$ and, hence, of $r_{bms,t}^{(s)}$. From the empirical distribution of $r_{bms,t}^{(s)}$, the desired quantile VaR and ES are estimated. See Appendix A.4. for detailed algorithm.

We compare four forecasting models. First, we employ the GHST copula which takes into account the asymmetric tail dependence in the model. The time varying dependence is implied by the GAS model. Second, we employ Student's t (ST) copula combined with the GAS model. The Student's t cannot specify the asymmetric tail dependence in the model. Hence, we can test the effect of the asymmetric tail dependence on the VaR and ES forecasts by comparing ST with GHST. Third, we employ the simulation based multivariate GARCH model, namely dynamic conditional correlation (DCC; Engle, 2002).[7] This model takes into account a linear correlation between two extreme port-

[7]We also consider other multivariate GARCH models such as BEKK-GARCH or CCC-GARCH. We find that DCC-GARCH provides more stable estimation and forecasting results than others. Hence, we report forecasting results by

folios but ignores their tail dependence. Thus we can test the effect of the (tail) dependence on the forecast by comparing DCC with the dynamic copula models. Fourth, we employ the filtered historical simulation (FHS; Barone-Adesi, et al., 2002). FHS is the most popular and successful simulation based univariate VaR model. However, it does not explicitly take into account the dependence structure between two extreme portfolios. Thus we can test the effect of the correlation on the forecast by comparing FHS with the dynamic copula models or DCC.

We apply standard backtesting tools to evaluating the coverage ability and the statistical accuracy of the VaR models. The coverage ability is evaluated by the empirical coverage probability (ECP) and Basel Penalty Zone (BPZ; Basel Committee on Banking and Supervision, 1996). The statistical accuracy is evaluated by the conditional coverage test (CC test; Christoffersen, 1998) and the dynamic quantile test (DQ test; Engle and Manganelli, 2004). See Appendix A.5. for details.

Backtesting of ES is not a straightforward task because it fails to satisfy elicitability (see Gneiting, 2011). Thus we simply evaluate the ES forecast based on a loss function which enables researchers to rank the models and specify a utility function reflecting the concern of the risk manager. We define two loss functions:

$$\text{Absolute error} :=\left| r_{bms,s}-\widehat{ES}_{bms,s}\left(\alpha\right)\right| 1\left\{ r_{bms,s}<\widehat{VaR}_{bms,s}\left(\alpha\right)\right\} , \tag{23}$$

$$\text{Squared error} :=\left(r_{bms,s}-\widehat{ES}_{bms,s}\left(\alpha\right)\right)^{2} 1\left\{ r_{bms,s}<\widehat{VaR}_{bms,s}\left(\alpha\right)\right\} , \tag{24}$$

for $s=1,\ldots,N$. In order to rank the models, we compute the mean absolute error (MAE) and the mean squared error (MSE). This evaluation is in line with the framework proposed by Lopez (1999) for the VaR evaluation. The smaller value indicates more accurate forecast.

In order to evaluate VaR and ES forecasts, we use a rolling window instead of the full sample period and set a window size at 250 (one trading year) for all the data sets.[8] All the models are recursively re-estimated throughout the out-of-sample period and the time-varying correlation coefficients of

DCC-GARCH in our paper.

[8]The reason we use a moving window of 250 days instead of other window length or expending window is because a moving window of 250 days is the standard estimation period by the Basel accord. In practice the selection of an optimal sample size is a nontrivial issue. As the window size increases, estimation and forecasting precision generally improves. On the other hand it also raises uncertainty about the latent market regimes caused by a sequence of rare or extreme shocks hitting the market in which case it would be more desirable to select the shorter and homogeneous sample rather than longer and heterogeneous ones.

copulas are implied by the GAS model. For the UK portfolios, we estimate the VaR and ES models using 250 business days over the period 4 Jan. 2000 - 15 Dec. 2000, and compute the one-day-ahead forecast of the 99 percent VaR and ES for 18 Dec. 2000. We conduct rolling forecast by moving forward a day at a time and end with the forecast for 31 Dec. 2012. This generates 3,033 out-of-sample daily forecasts. Next we repeat the same process for the US portfolios. It starts with the forecast for 18 Dec. 2000 and ends with the forecast for 31 Dec. 2012. This generates 3,018 out-of-sample daily forecasts.

We evaluate the coverage ability by ECP and BPZ as follows: We calculate ECP for each portfolio and then report bias and Root Mean Square Error (RMSE). Bias is the average deviation of ECP from the nominal coverage probability (1% in our case). The smaller the bias is, the more accurate the VaR forecast is. RMSE is the average of the squared deviation. It shows the dispersion of ECP from the nominal coverage probability. It makes up for the defect of bias due to the offset of positive and negative deviations. Financial regulators would prefer a VaR model with, simultaneously, a small bias and small RMSE. BPZ describes the coverage ability of the VaR model through the test of failure rate. It counts the number of failures over the previous 250 business days.

5.2.1 Coverage Ability

Table 8 presents the evaluation results of the coverage ability. Panel A shows the ECPs of the VaR models. We find that the GHST copula produces the smallest bias (-0.02%) with the smallest RMSE (0.11%). It means that the ECPs of all portfolios are more concentrated around the nominal one than other models. The ST copula also produces the smallest bias but RMSE (0.13%) is slightly larger than GHST. We thus can draw inference that the asymmetric tail dependence slightly contributes to the VaR forecast from this marginal difference. On the other hand, DCC produces a huge bias (0.28%) implying that it significantly underforecasts VaR. We can easily understand that this huge underforecast is mostly generated by the lack of tail dependence in DCC. We also find that FHS largely underforecast VaR. Interestingly, FHS considers neither correlation nor tail dependence but its bias (0.31%) is similar with DCC. This tells us that the correlation (the second comoment) is not the key input in the extreme event forecast.

Panel B shows the BPZ of the VaR models. We find that all models achieve 12 Green zone using the framework of Basel committee. This result is however not surprising as the “traffic lights” backtest is not as rigorous as other statistical tests such as CC test and DQ test.

[INSERT TABLE 8 ABOUT HERE]

Consequently, the evaluation results of the coverage ability clearly show the importance of modeling the time varying asymmetric dependence structure in the extreme event forecast. In particular, we are able to confirm that the modeling of linear correlation does not improve the forecasting accuracy of extreme event at all.

5.2.2 Statistical Accuracy

We evaluate the statistical accuracy by the CC test and the DQ test as follows: We calculate both statistics for each portfolio and test them at the 5% significance level. Then we report the number of rejected portfolios.

Table 9 represents the results of statistical tests. Panel A reports the CC test results. The GHST and ST copulas are rejected for 2 portfolios and 1 portfolio, respectively. Thus there is not the significant difference between the asymmetric copula and the symmetric copula. It thus repeatedly verifies that the asymmetric tail dependence slightly contributes to the VaR forecast. DCC and FHS are rejected for 3 and 2 portfolios, respectively. Panel B reports the results of the DQ test. The results are mostly consistent with the CC test. The GHST and ST copulas are rejected for 2 portfolios whilst DCC and FHS are rejected for 3 portfolios. Although there is not the striking difference of rejection frequency, overall the dynamic copula models are less rejected than DCC and FHS.

[INSERT TABLE 9 ABOUT HERE]

5.3 Backtesting of Expected Shortfall

Table 10 presents MAE and MSE for the ES forecasts. Panel A reports the MAE results. The dynamic copula models provide the most accurate forecasts (lowest MAEs) in 10 out of 12 portfolios. Also, The GHST copula generates the lower average MAE in general comparing with the ST copula, as it takes into account the asymmetric dependence between portfolios. As a robustness check, the MSE results reported in Paenl B also confirm this conclusion. The dynamic copula models have better performance than both DCC and FHS in almost all cases.

[INSERT TABLE 10 ABOUT HERE]

In sum, we find that two extreme portfolios sorted on the higher order comoments are closely correlated each other over their whole distribution structure. Their dependence structure is time-varying and nonlinear. We perform the forecasting exercise for the extreme event and have the following implications from the backtesting results. Firstly, the dynamic modeling of the tail dependence is more effective than the dynamic modeling of the linear correlation for the accurate forecast of the extreme event. Especially, the lack of the tail dependence in the forecast model generates the huge underforecast of the extreme event. Secondly, there is little difference between the asymmetric dependence and the symmetric dependence in our portfolios. Both provide almost equivalent performances. Thirdly, the linear correlation does not help it to improve the extreme event forecast at all. There is little difference between DCC and FHS. Overall, the evaluation results strongly support the importance of modeling time-varying and asymmetric dependence in the market risk management. Note that we also examine the forecasting performance of all the candidate models at 95% and 97.5% significance level. The consistent results confirm our conclusion and suggest that data mining are unlikely explanations.[9]

6 Conclusion

The higher order comoments are occupies an important position in asset pricing with the Fama-French factors. There is strong evidence of nonlinear dependence across factors (Christoffersen and Langlois, 2013), which is the key input for equity portfolio selection and risk management. We therefore empirically study a relation between the higher order comoments and the dependence structure of the equity portfolio in the US and UK. We also investigate the role of dependence structure in the risk management of extreme events using equity portfolios sorted on the higher order comoments.

There are three notable findings in this paper.

First, our analysis shows that there are clear patterns in the relation between the higher order comoments and the dependence structure of the equity portfolio. Our simple portfolio approach provides the significant evidences of increasing or decreasing patterns in the copula correlation and

[9] All the robustness checks are available on request from the authors.

the asymmetric tail dependence. First, beta and coskewness are positively related with the copula correlation, whereas cokurtosis is negatively correlated with it. Second, beta is positively related with the asymmetric tail dependence whilst coskewness is negatively related with it.

Second, we find that two extreme equity portfolios sorted on higher order comoments are closely correlated. In particular, the dependence structure of the BMS portfolio (high beta (coskewness, cokurtosis) minus low beta (coskewness, cokurtosis)) is strongly time-varying and nonlinear. The backtesting results show that the dependence structure is the key input for the robust risk management of extreme events. The forecasting model employing a linear correlation significantly underforecasts value-at-risk and expected shortfall whilst the dynamic copula models forecast them very accurately.

Third, we use the new copula model to investigate the economic and statistical importance of modeling the time-varying and asymmetric dependence between equity portfolios sorted on the higher order comoments. We combine a hyperbolic generalized skewed t distribution with the generalized autoregressive score to capture both dynamics and asymmetries. Using a forecasting based risk management exercise, we demonstrate economic and statistical gains from modeling dynamic and asymmetric dependence. Our proposed copula achieves stronger coverage ability and better statistical accuracy.

Overall, we conclude that the higher order comoments are closely related with the dependence structure of the equity portfolio. The dependence structure is time-varying and nonlinear, which is the key input for the risk management of extreme events. The forecasting based risk management exercise demonstrates the importance of modeling the dynamic and asymmetric dependence using copulas.

Appendix

A.1. Estimation of Parametric Copula Model

The log-likelihood of a fully parametric copula model for conditional distribution of $\mathbf{z}_t$ takes the form:

$$L(\boldsymbol{\theta}) = \prod_{t=1}^{T} \mathbf{f}(\mathbf{z}_t|\mathcal{F}_{t-1};\boldsymbol{\theta}) = \prod_{t=1}^{T} \left[c_t(u_{1,t},...,u_{d,t}|\mathcal{F}_{t-1};\boldsymbol{\theta}_C) \prod_{i=1}^{N} f_{i,t}(z_{i,t}|\mathcal{F}_{t-1};\theta_i) \right] \tag{A.1}$$

with log-likelihood

$$\begin{aligned} \sum_{t=1}^{T} \log \mathbf{f}(\mathbf{z}_t|\mathcal{F}_{t-1};\boldsymbol{\theta}) = & \sum_{t=1}^{T} \sum_{i=1}^{d} \log f_{i,t}(z_{i,t}|\mathcal{F}_{t-1};\theta_i) \\ & + \sum_{t=1}^{T} \log c_t\left(F_{1,t}(z_{1,t}|\mathcal{F}_{t-1};\theta_1), \ldots, F_{d,t}(z_{d,t}|\mathcal{F}_{t-1};\theta_d)|\mathcal{F}_{t-1};\boldsymbol{\theta}_C\right) \end{aligned} \tag{A.2}$$

where $\boldsymbol{\theta}$ denotes the parameter vector for the full model parameters, θ_i denotes the parameters for the ith marginals, θ_C denotes the parameters of copula model and $\mathcal{F}_{t-1}$ denotes the information set at time $t-1$. Following the two-stage maximum likelihood estimation (also known as the Inference method for marginals) of Joe and Xu (1996), we first estimate the parameters of marginal models using maximum likelihood:

$$\hat{\theta}_{\mathbf{i}} = \underset{\theta_i}{\operatorname{argmax}} \sum_{t=1}^{T} \log f_{i,t}(z_{i,t}|\mathcal{F}_{t-1};\theta_i),\ i = 1,...,N, \tag{A.3}$$

and then using the estimations in the first stage, we calculate $F_{i,t}$ and estimate the copula parameters via maximum likelihood:

$$\hat{\theta}_{\mathbf{C}} = \underset{\theta_C}{\operatorname{argmax}} \sum_{t=1}^{T} \log c_t\left(F_{1,t}(z_{1,t}|\mathcal{F}_{t-1};\theta_1),...,F_{d,t}(z_{d,t}|\mathcal{F}_{t-1};\theta_d)|\mathcal{F}_{t-1};\theta_C\right) \tag{A.4}$$

A.2. Estimation of Semiparametric Copula Model

In the semiparametric estimation (also known as Canonical Maximum Likelihood Estimation), the univariate marginals are estimated nonparametrically using the empirical distribution function (EDF) and the copula model is again parametrically estimated via maximum likelihood.

$$\hat{F}_i(z) \equiv \frac{1}{T+1}\sum_{t=1}^{T} 1\left\{\hat{z}_{i,t} \le z\right\} \tag{A.5}$$

$$\hat{u}_{i,t} \equiv \hat{F}_i(z) \sim Unif\ (0,1)\,,\ i = 1,2,...,N \tag{A.6}$$

$$\hat{\theta}_{\mathbf{C}} = \arg\max_{\theta_{\mathbf{C}}}\sum_{t=1}^{T}\log c_t\left(\hat{u}_{1,t},...,\hat{u}_{i,t}|\mathcal{F}_{t-1};\theta_{\mathbf{C}}\right) \tag{A.7}$$

where $z_{i,t}$ are the standardized residuals of the marginal model and $\hat{F}_i$ is different from the standard empirical CDF by the scalar $1/(n+1)$ (in order to ensure that the transformed data cannot be on the boundary of the unit interval $[0,1]$).

A.3. Computation of Asymmetric Dependence

A primary goal of our paper is to investigate how the characteristic-sorted portfolio returns covary and whether their dependence structures are asymmetric. Consequently, we consider three different dependence structures: The threshold correlation; the quantile dependence; and the tail dependence.

Following Longin and Solnik (2001) and Ang and Chen (2002), the threshold correlation for probability level *p* is given by

$$\rho^{-} = Corr\left(r_{h,t}, r_{l,t}|r_{h,t} \le r_h(p)\ \text{ and } r_{l,t} \le r_l(p)\right) \text{ if } p \le 0.5 \tag{A.8}$$

$$\rho^{+} = Corr\left(r_{h,t}, r_{l,t}|r_{h,t} > r_h(p)\ \text{ and } r_{l,t} > r_l(p)\right) \text{ if } p > 0.5 \tag{A.9}$$

where $r(p)$ denotes the corresponding empirical percentile for asset returns $r_{h,t}$ and $r_{l,t}$. In words, we compute the correlation between two assets conditional on both of them being less (respectively, greater) than their *p*th percentile value when $p \le 0.5$ (respectively, $p > 0.5$). To examine whether this asymmetry is statistically significant, we consider a model-free test proposed by Hong et al. (2007). If the null hypothesis of $\rho^{+} = \rho^{-}$ is rejected, then there exists a linear asymmetric correlation between $r_{h,t}$ and $r_{l,t}$.

The quantile dependence provides a more precise measure of dependence structure than the threshold correlation, as it contains more detailed information. In addition, from the risk management perspective, tails are more important than the centre. Following Patton (2012), the quantile dependence can be defined as

$$\lambda^q = \begin{cases} \mathbb{P}\left\{u_{h,t} \leq q | u_{l,t} \leq q\right\} = \frac{C(q,q)}{q} & \text{if } 0 < q \leq 0.5 \\ \mathbb{P}\left\{u_{h,t} > q | u_{l,t} > q\right\} = \frac{1-2q+C(q,q)}{1-q} & \text{if } 0.5 < q \leq 1 \end{cases} \tag{A.10}$$

and nonparametrically estimated by

$$\hat{\lambda}^q = \begin{cases} \frac{1}{Tq} \sum_{t=1}^{T} 1\left\{\hat{u}_{h,t} \leq q, \hat{u}_{l,t} \leq q\right\} & \text{if } 0 < q \leq 0.5 \\ \frac{1}{T(1-q)} \sum_{t=1}^{T} 1\left\{\hat{u}_{h,t} > q, \hat{u}_{l,t} > q\right\} & \text{if } 0.5 < q < 1. \end{cases}, \tag{A.11}$$

where C denotes the corresponding copula function.

The tail dependence coefficient (TDC) is a measure of the degree of dependence in the tail of a bivariate distribution (see McNeil et al., 2005; Frahm et al., 2005; Joe et al., 2010, among others). Let z_h and z_l be random variables with continuous distribution functions F_h and F_l. Then the coefficients of upper and lower tail dependence of z_h and z_l are

$$\lambda^L = \lim_{q \to 0^+} \frac{\mathbb{P}\left\{z_h \leq F_h^{-1}(q), z_l \leq F_l^{-1}(q)\right\}}{\mathbb{P}\left\{z_l \leq F_l^{-1}(q)\right\}} = \lim_{q \to 0^+} \frac{C(q,q)}{q} \tag{A.12}$$

$$\lambda^U = \lim_{q \to 1^-} \frac{\mathbb{P}\left\{z_h > F_h^{-1}(q), z_l > F_l^{-1}(q)\right\}}{\mathbb{P}\left\{z_l > F_l^{-1}(q)\right\}} = \lim_{q \to 1^-} \frac{1-2q+C(q,q)}{1-q} \tag{A.13}$$

The coefficients can be easily calculated when the copula C has a closed form. The copula C has upper tail dependence if $\lambda^U \in (0,1]$ and no upper tail dependence if $\lambda^U = 0$. A similar conclusion holds for the lower tail dependence. If the copulas are symmetric, then $\lambda^L = \lambda^U$, otherwise, $\lambda^L \neq \lambda^U$ (see Joe, 1997). McNeil et al. (2005) state that the copula of the bivariate t distribution is asymptotically dependent in both the upper and lower tail. We use the Student's t copula to estimate the tail dependence coefficient between portfolios.

A.4. Algorithm for Forecasting VaR and ES

[Step 1] Determine the in sample and out-of-sample period for VaR and ES forecasting.

[Step 2] We predict conditional mean and conditional volatility from the prespecified time series model on rolling window and do one step ahead forecasting for each margins.

[Step 3] Estimate the density model to get the probabilities for each forecasted margin. We consider both parametric (univariate skewed t) and nonparametric (EDF) estimation on sliding window.

[Step 4] Estimate the parameters for full parametric and semiparametric copulas using using maximum likelihood estimation.

[Step 5] Using the estimated parameters in [Step 4] as initial values, we estimate time-varying dependence parameters for asymmetric (GH skewed t) copulas based on the GAS framework.

[Step 6] With the estimated time-varying copula parameters in hand, we can apply Monte Carlo simulation to generate N samples of shocks and then portfolio returns.

[Step 7] Based on the empirical α-quantile of forecasted portfolio return, it is straightforward to forecast corresponding VaR.

[Step 8] Given the N simulated portfolio returns, we can also calculate α-quantile ES.

[Step 9] Use the realized portfolio returns to backtest VaR and ES forecasts.

A.5. Backtesting VaR and ES

We first define the failure of VaR as the event that a realized return r_s is not covered by the predicted VaR. We identify it by the indicator function taking the value unity in the case of failure:

$$I_s = 1\left\{r_s < \widehat{VaR}_s\left(\alpha|\mathcal{F}_{s-1}\right)\right\},\ s = 1,\dots,N, \tag{A.14}$$

where $\widehat{VaR}_s\left(\alpha|\mathcal{F}_{s-1}\right)$ is the VaR forecast based on the information set at $s-1$, denoted by $\mathcal{F}_{s-1}$, with a nominal coverage probability α. Henceforth, we abbreviate the notation $\widehat{VaR}_s\left(\alpha|\mathcal{F}_{s-1}\right)$ to $\widehat{Var}_s\left(\alpha\right)$.

Empirical Coverage Probability (ECP) is calculated by the sample average of I_s, $\hat{\alpha} = N^{-1}\sum_{s=1}^{N} I_s$ which is a consistent estimator of the coverage probability. The VaR model for which ECP is closest to its nominal coverage probability is preferred. BPZ is suggested by Basel Committee on Banking and Supervision (1996). It describes the strength of the VaR model through the test of failure rate. It records the number of failures of the 99 percent VaR in the previous 250 business days. One may expect, on average, 2.5 failures out of the previous 250 VaR forecasts given the correct forecasting model. The Basel Committee rules that up to four failures are acceptable for banks and defines the range as a "Green" zone. If the failures are five or more, the banks fall into a "Yellow" (5–9) or "Red" (10+) zone. The VaR model for which BPZ is in the "Green" zone is preferred.

Accurate VaR forecasts should satisfy the condition that the conditional expectation of the failure is the nominal coverage probability:

$$\mathbb{E}\left[I_s|\mathcal{F}_{s-1}\right] = \alpha. \tag{A.15}$$

Christoffersen (1998) shows that it is equivalent to testing if $I_s|\mathcal{F}_{s-1}$ follows an i.i.d. Bernoulli distribution with parameter α:

$$H_0 : I_s|\mathcal{F}_{s-1} \sim i.i.d.\ Bernoulli\left(\alpha\right). \tag{A.16}$$

The CC test uses the LR statistic which follows the chi-squared distribution with two degrees-of-freedom under the null hypothesis, Eq. (A.16). The DQ test is a general extension of the CC test allowing for more time-dependent information of $\{I_s\}_{s=1}^{N}$. The out-of-sample DQ test is given by

$$DQ = \frac{\left(\tilde{\mathbf{I}}'\mathbf{Z}\right)\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\left(\mathbf{Z}'\tilde{\mathbf{I}}\right)}{\alpha\left(1-\alpha\right)} \overset{d}{\sim} \chi^2_{p+2}, \tag{A.17}$$

where $\tilde{\mathbf{I}} = \left(\tilde{I}_{p+1}, \tilde{I}_{p+2}, \ldots, \tilde{I}_N\right)'$, $\tilde{I}_s = I_s - \alpha$, $\mathbf{Z} = \left(\mathbf{z}_{p+1}, \ldots, \mathbf{z}_N\right)'$ and $\mathbf{z}_s = \left(1, \tilde{I}_{s-1}, \ldots, \tilde{I}_{s-p}, \widehat{VaR}_s\left(\alpha\right)\right)'$. We use the first four lags for our evaluation, i.e., $\mathbf{z}_s = \left(1, \tilde{I}_{s-1}, \ldots, \tilde{I}_{s-4}, \widehat{VaR}_s\left(\alpha\right)\right)'$.

Backtesting of ES is not a straightforward task because it fails to satisfy elicitability (see Gneiting, 2011). We consider a backtesting for the ES forecast given the sample of N ES forecasts. We simply evaluate the ES forecast based on a loss function which enables researchers to rank the models and specify a utility function reflecting the concern of the risk manager. We define two loss functions:

$$\text{Absolute error} := \left|r_s - \widehat{ES}_s\left(\alpha\right)\right| I_s,\ \ \text{Squared error} := \left(r_s - \widehat{ES}_s\left(\alpha\right)\right)^2 I_s, \tag{A.18}$$

where $I_s = 1\left\{r_s < \widehat{VaR}_s\left(\alpha\right)\right\}$. In order to rank the models, we compute the mean absolute error (MAE) and the mean squared error (MSE):

$$MAE = \frac{1}{N}\sum_{s=1}^{N}\left|r_s - \widehat{ES}_s\left(\alpha\right)\right| I_s,\ \ MSE = \frac{1}{N}\sum_{s=1}^{N}\left(r_s - \widehat{ES}_s\left(\alpha\right)\right)^2 I_s. \tag{A.19}$$

This evaluation is in line with the framework proposed by Lopez (1999) for the VaR evaluation. The smaller value indicates more accurate forecast.

References

Andrews, D.W., 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821-856.

Ang, A., Chen, J., 2002. Asymmetric correlations of equity portfolios. *Journal of Financial Economics* 63, 443-494.

Ang, A., Chen, J., Xing, Y., 2006. Downside risk. *Review of Financial Studies* 19, 1191-1239.

Bakshi, G., Kapadia, N., Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies* 16, 101-143.

Barone-Adesi, G., Giannopoulos, K., Vosper, L., 2002. Backtesting derivative portfolios with filtered historical simulation (FHS). *European Financial Management* 8, 31-58.

Basel Committee., 1996. Supervisory framework for the use of "backtesting" in conjunction with the internal models approach to market risk capital requirements. *Basel Committee on Bank Supervision.*

Chabi-Yo, F., Ruenzi, S., Weigert, F., 2014. Crash sensitivity and the cross-section of expected stock returns. *Unpublished Working Paper, University of Mannheim.*

Chen, X., Fan, Y., 2006. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics* 135, 125-154.

Christoffersen, P., 1998. Evaluating interval forecasts. *International Economic Review* 39, 841-862.

Christoffersen, P., Errunza, V., Jacobs, K., Langlois, H., 2012. Is the potential for international diversification disappearing? A dynamic copula approach. *Review of Financial Studies* 25, 3711-3751.

Christoffersen, P., Langlois, H., 2013. The joint dynamics of equity market factors. *Journal of Financial and Quantitative Analysis* 48, 1371-1404.

Chung, Y.P., Johnson, H., Schill, M.J., 2006. Asset pricing when returns are nonnormal: Fama-French factors versus higher-order systematic comoments. *Journal of Business* 79, 923-940.

Conrad, J., Dittmar, R.F., Ghysels, E., 2013. Ex ante skewness and expected stock returns. *Journal of Finance* 68, 85-124.

Creal, D., Koopman, S.J., Lucas, A., 2013. Generalized autoregressive score models with applications. *Journal of Applied Econometrics* 28, 777-795.

Demarta, S., McNeil, A.J., 2005. The t Copula and related copulas. *International Statistical Review* 73, 111-129.

Dittmar, R.F., 2002. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *Journal of Finance* 57, 369-403.

Engle, R., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* 50, 987-1007.

Engle, R., 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* 20, 339-350.

Engle, R., Manganelli, S., 2004. CAViaR: Conditional autoregressive Value at Risk by regression quantiles. *Journal of Business and Economic Statistics* 22, 367-381.

Fama, E. F., French, K. R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3-56.

Frahm, G., Junker, M., Schmidt, R., 2005. Estimating the tail-dependence coefficient: properties and pitfalls. *Insurance: Mathematics and Economics* 37, 80-100.

Garcia, R., Tsafack, G., 2011. Dependence structure and extreme comovements in international equity and bond markets. *Journal of Banking and Finance* 35, 1954-1970.

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

Gneiting, T., 2011. Making and evaluating point forecasts. *Journal of the American Statistical Association* 106, 746-762.

Guidolin, M., Timmermann, A., 2008. International asset allocation under regime switching, skew, and kurtosis preferences. *Review of Financial Studies* 21, 889-935.

Hansen, B.E., 1994. Autoregressive conditional density estimation. *International Economic Review* 35, 705-730.

Harvey, C.R., Siddique. A., 2000. Conditional skewness in asset pricing tests. *Journal of Finance* 55, 1263-1295.

Hong, Y., Tu, J., Zhou, G., 2007. Asymmetries in stock returns: Statistical tests and economic evaluation. *Review of Financial Studies* 20, 1547-1581.

Joe, H., 1997. *Multivariate Models and Dependence Concepts.* Chapmann & Hall, London.

Joe, H., Li, H., Nikoloulopoulos, A.K., 2010. Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101, 252-270.

Joe, H., Xu, J.J., 1996. The estimation method of inference functions for margins for multivariate models. *Technical Report no. 166, Department of Statistics, University of British Columbia.*

Longin, F., Solnik, B., 1995. Is the correlation in international equity returns constant: 1960–1990?. *Journal of International Money and Finance* 14, 3-26.

Longin, F., Solnik, B., 2001. Extreme correlation of international equity markets. *Journal of Finance* 56, 649-676.

Lopez, J.A., 1999. Methods for evaluating Value-at-Risk estimates. *Federal Reserve Bank of New York, Economic Policy Review* 2, 3-17.

McNeil, A.J., Frey, R., Embrechts, P., 2005. *Quantitative Risk Management: Concepts, Techniques, and Tools.* Princeton University Press.

Nelsen, R.B., 2006. *An Introduction to Copulas.* Springer.

Patton, A.J., 2006. Modelling asymmetric exchange rate dependence. *International Economic Review* 47, 527-556.

Patton, A.J., 2012. Copula methods for forecasting multivariate time series. *Handbook of Economic Forecasting*, 2, 899-960.

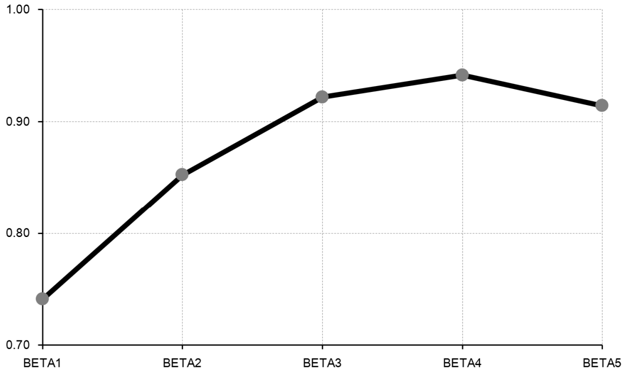
Patton, A.J., Timmermann, A., 2010. Monotonicity in asset returns: New tests with applications to the term structure, the CAPM, and portfolio sorts. *Journal of Financial Economics* 98, 605-625.

Sklar, M., 1959. *Fonctions de Répartition à n Dimensions et Leurs Marges*, Université Paris 8.

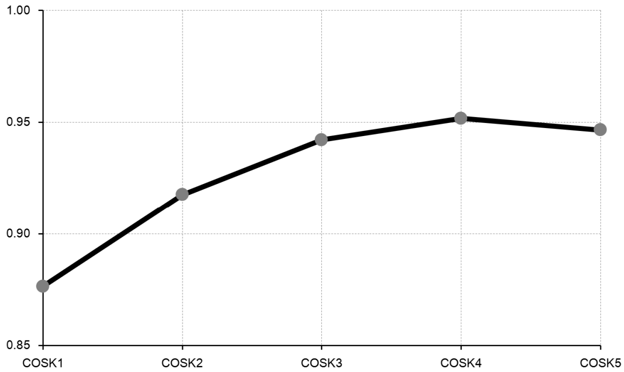
Figure 1: Relation between Higher Order Comoments and Copula Correlation

This figure plots the average copula correlations for equity portfolios sorted on higher order comoments. BETA1 (COSK1, COKT1) denotes the lowest beta (coskewness, cokurtosis) portfolio whilst BETA5 (COSK5, COKT5) denotes the highest beta (coskewness, cokurtosis) portfolio. We estimate the correlation of GHST copula where the time-varying correlations are implied by the GAS model.

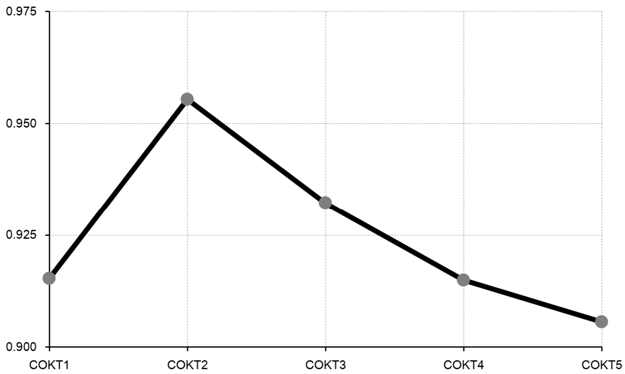
Panel A: US Stock Market



(a) BETA

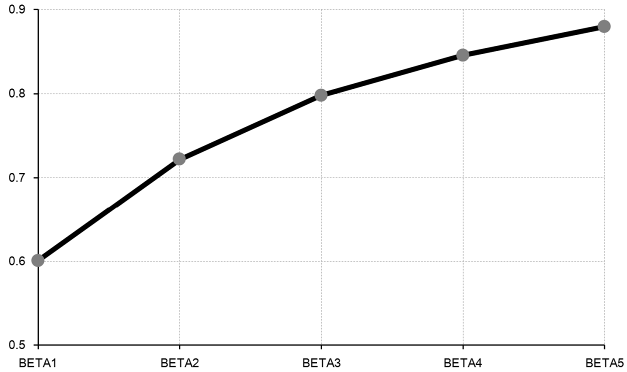


(b) Coskewness

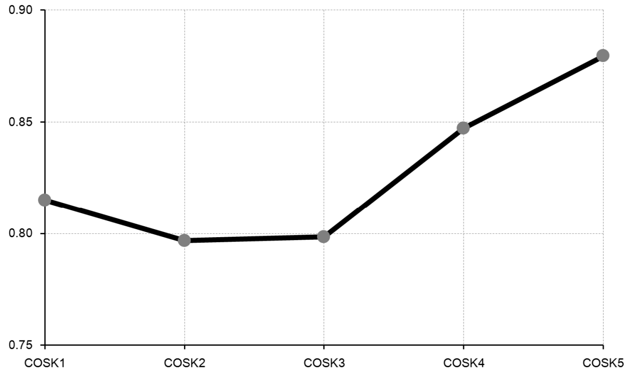


(c) Cokurtosis

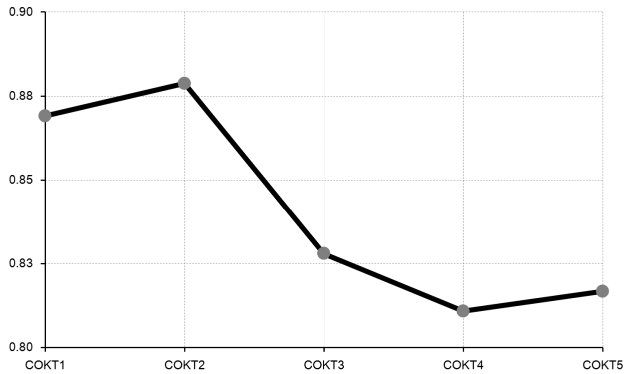
Panel B: UK Stock Market



(a) BETA



(b) Coskewness

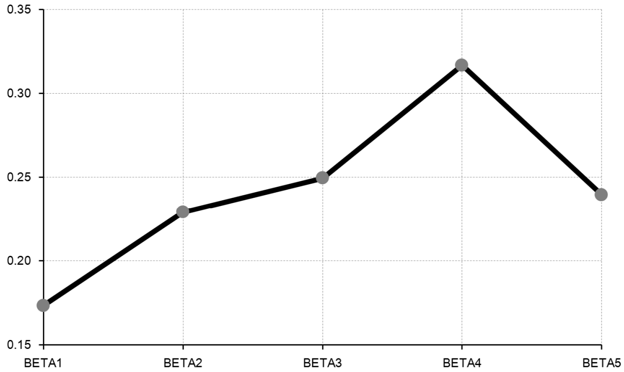


(c) Cokurtosis

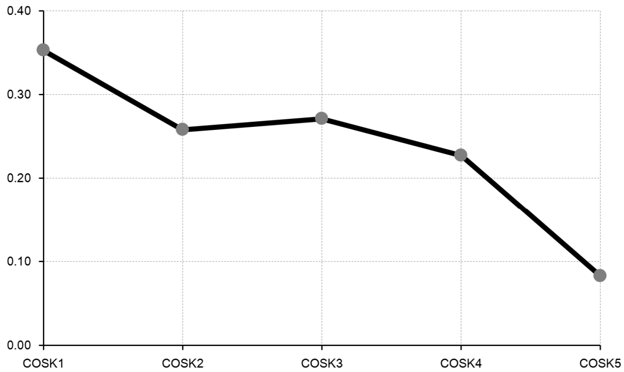
Figure 2: Relation between Higher Order Comoments and Asymmetric Tail Dependence

This figure plots the average difference between lower tail dependence (LTD) and upper tail dependence (UTD) for equity portfolios sorted on higher order comoments, $DIFF = LTD - UTD$. The average difference measures the extent of asymmetry for the tail dependence. The tail dependence coefficients are calculated by the parametric approach in McNeil et al. (2005). We estimate the GHST copula where the time-varying correlations are implied by the GAS model.

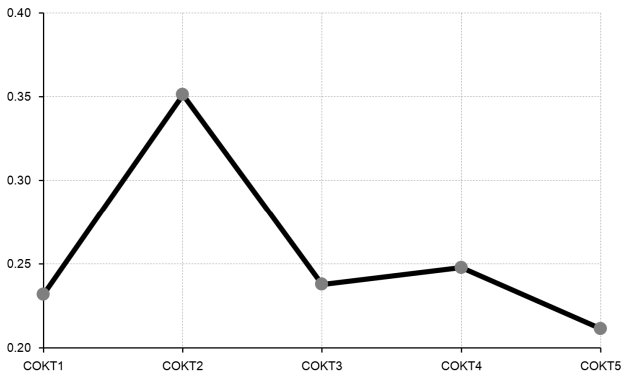
Panel A: US Stock Market



(a) BETA

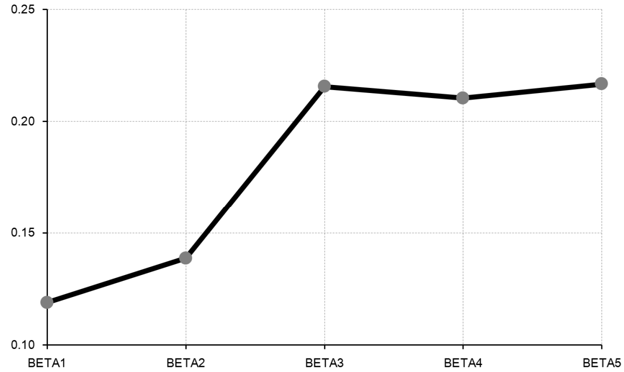


(b) Coskewness

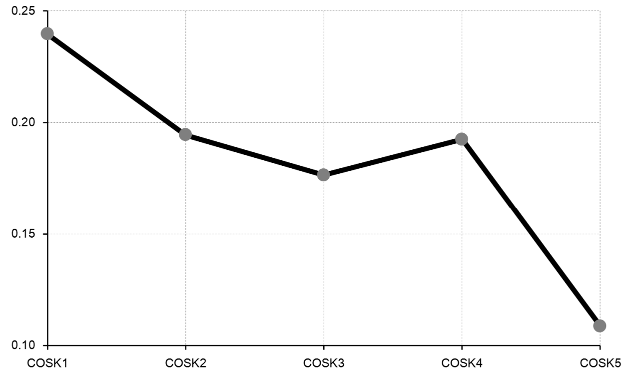


(c) Cokurtosis

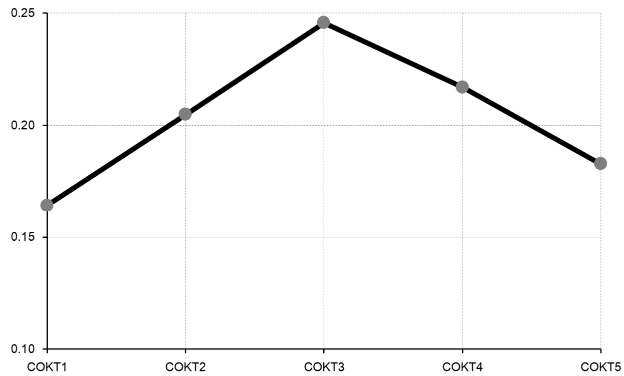
Panel B: UK Stock Market



(a) BETA



(b) Coskewness



(c) Cokurtosis

Table 1: Relation between Higher Order Comoments and Copula Correlation

This table presents average copula correlations for equity portfolios sorted on higher order comoments. We estimate the correlation of GHST copula where the time-varying correlations are implied by the GAS model. The column P1 denotes a portfolio with the lowest higher order comoment whilst the column P5 denotes one with the highest higher order comment. Columns 'MR', 'Up' and 'Down' report p-values for tests of correlation monotonicity (Patton and Timmermann, 2010). '–' indicates a case that both 'Up' test and 'Down' test are rejected, i.e., inconclusive.

Portfolio	P1	P2	P3	P4	P5	MR	Up	Down
Panel A: US Stock Market								
BETA	0.741	0.852	0.922	0.942	0.914	1.000	0.000	0.013
Coskewness	0.876	0.918	0.942	0.952	0.947	0.972	0.000	0.230
Cokurtosis	0.915	0.955	0.932	0.915	0.906	1.000	–	–
Panel B: UK Stock Market								
BETA	0.600	0.722	0.798	0.846	0.880	0.000	0.000	0.910
Coskewness	0.815	0.797	0.798	0.847	0.880	0.905	0.000	0.323
Cokurtosis	0.869	0.879	0.828	0.811	0.817	0.938	0.101	0.000

Table 2: Relation between Higher Order Comoments and Asymmetry of Tail Dependence

This table presents the average difference of lower tail dependence (LTD) and upper tail dependence (UTD) for equity portfolios sorted on higher order comoments, $DIFF = LTD - UTD$. The average difference measures the extent of asymmetry for the tail dependence. The tail dependences are calculated by the parametric approach in McNeil et al. (2005). We estimate the GHST copula where the time-varying correlations are implied by the GAS model. The column P1 denotes a portfolio with the lowest higher order comoment whilst the column P5 denotes one with the highest higher order comment. Columns 'MR', 'Up' and 'Down' report p-values for tests of correlation monotonicity (Patton and Timmermann, 2010). '–' indicates a case that both 'Up' test and 'Down' test are rejected, i.e., inconclusive.

Portfolio	P1	P2	P3	P4	P5	MR	Up	Down
Panel A: US Stock Market								
BETA	0.173	0.229	0.250	0.317	0.240	1.000	–	–
Coskewness	0.353	0.258	0.271	0.227	0.083	1.000	0.026	0.000
Cokurtosis	0.232	0.351	0.238	0.248	0.212	1.000	–	–
Panel B: UK Stock Market								
BETA	0.119	0.139	0.216	0.210	0.217	0.777	0.000	0.488
Coskewness	0.240	0.195	0.176	0.193	0.109	1.000	–	–
Cokurtosis	0.164	0.205	0.246	0.217	0.183	1.000	–	–

Table 3: Definitions of BMS Portfolios

This table describes the 12 BMS portfolios that we constructed for the purpose of empirical analysis in our study. Portfolios are sorted by market beta, coskewness and cokurtosis. All the portfolios are annually rebalanced.

Portfolio	Description
Panel A: US Stock Market	
$BMS\,(Beta, L/S; US)$	Long (short) BETA5 and short (long) BETA1
$BMS\,(Cosk, L/S; US)$	Long (short) COSK5 and short (long) COSK1
$BMS\,(Cokt, L/S; US)$	Long (short) COKT5 and short (long) COKT1
Panel B: UK Stock Market	
$BMS\,(Beta, L/S; UK)$	Long (short) BETA5 and short (long) BETA1
$BMS\,(Cosk, L/S; UK)$	Long (short) COSK5 and short (long) COSK1
$BMS\,(Cokt, L/S; UK)$	Long (short) COKT5 and short (long) COKT1

Table 4: Parameter Estimates and Goodness of Fit Test for Univariate Modeling

This table reports parameter estimates from AR and GJR-GARCH models for conditional mean and conditional variance of portfolio returns. We estimate all parameters using the sample from January 4, 2000 to December 31, 2012, which correspond to a sample of 3,268 observations for US market and a sample of 3,283 observations for UK market. We use * to indicate the significance of estimate at the 5% significance level. We also report the p-values of two goodness-of-fit tests for the skewed Student's t distribution. We employ Kolmogorov-Smirnov (KS) and Cramer-von Mises (CvM) tests.

Panel A: US Stock Market

Porftolio	BETA1	BETA5	COSK1	COSK5	COKT1	COKT5
ϕ_0	0.025	-0.021	0.004	0.006	-0.011	0.029
ϕ_1	-0.064*	–	–	-0.091*	–	-0.085*
ϕ_2	–	–	–	-0.085*	–	-0.061*
ω	0.012*	0.048*	0.027*	0.015*	0.021*	0.016*
α	0.018	0.001*	0.0235*	0.000	0.023*	0.008
γ	0.122*	0.133*	0.104*	0.163*	0.126*	0.179*
β	0.901*	0.923*	0.903*	0.916*	0.903*	0.896*
ν	11.507	12.805	9.404	12.331	24.329	5.637
λ	-0.111	-0.072	-0.246	0.006	-0.099	-0.099
KS	0.61	0.17	0.43	0.11	0.14	0.97
CvM	0.33	0.10	0.42	0.10	0.16	1.00

Panel B: UK Stock Market

Porftolio	BETA1	BETA5	COSK1	COSK5	COKT1	COKT5
ϕ_0	0.039*	-0.023	0.018	-0.001	0.009	0.008
ϕ_1	–	–	–	-0.045*	–	-0.060*
ϕ_2	–	–	–	–	–	–
ω	0.014*	0.036*	0.021*	0.019*	0.014*	0.028*
α	0.012	0.004	0.031*	0.006	0.016	0.047*
γ	0.111*	0.130*	0.092*	0.135*	0.103*	0.142*
β	0.913*	0.924*	0.911*	0.919*	0.925*	0.867*
ν	9.593	27.910	7.384	33.361	94.667	7.475
λ	-0.076	-0.054	-0.168	0.025	-0.062	-0.076
KS	0.77	0.96	0.53	0.53	0.35	0.36
CvM	0.87	0.92	0.30	0.40	0.45	0.27

Table 5: Tests for Time-varying Dependence between High and Low Portfolios

We report the p-value from tests for time-varying rank correlation between the high portfolio (e.g. BETA5) and the low portfolio (e.g. BETA1). Having no a priori dates to consider for the timing of a break, we consider naive tests for breaks at three chosen points in sample period, at $t^*/T \in \{0.15, 0.50, 0.85\}$, which corresponds to the dates 10-Dec-2001, 03-Jul-2006, 17-Jan-2011. The 'Any' column reports the results of test for dependence break of unknown timing proposed by Andrews (1993). To detect whether the dependence structures between characteristic-sorted portfolios significantly changed after the US and EU crisis broke out, we use 15-Sep-2008 (the collapse of Lehman Brothers) and 01-Jan-2010 (EU sovereign debt crisis) as two break points and the 'Crisis' panel reports the results for this test. The 'AR' panel presents the results from the ARCH LM test for time-varying volatility proposed by Engle (1982). Under the null hypothesis of a constant conditional copula, we test autocorrelation in a measure of dependence (see Patton, 2012).

	Panel A: Break				Panel B: Crisis		Panel C: AR(p)		
Portfolio	0.15	0.5	0.85	Any	US	EU	AR(1)	AR(5)	AR(10)
US BETA1&5	0.00	0.00	0.04	0.00	0.00	0.07	0.00	0.00	0.00
US COSK1&5	0.00	0.03	0.82	0.04	0.22	0.38	0.00	0.00	0.00
US COKT1&5	0.02	0.30	0.67	0.25	0.92	0.42	0.18	0.75	0.09
UK BETA1&5	0.00	0.00	0.17	0.00	0.04	0.08	0.00	0.12	0.00
UK COSK1&5	0.59	0.03	0.62	0.03	0.07	0.25	0.01	0.00	0.02
UK COKT1&5	0.98	0.24	0.36	0.24	0.24	0.13	0.00	0.00	0.00

Table 6: Testing Differences of Exceedence Correlations

This table presents the statistics and p-values from a model-free symmetry test proposed by Hong et al. (2007) to examine whether the exceedance correlations between low portfolio (i.e. BETA1) and high portfolio (i.e. BETA5) are asymmetric. We report p-values in $[\cdot]$. The $\mathcal{J}$ statistics for testing the null hypothesis of symmetric correlation, $\rho^{+}(c) = \rho^{-}(c)$, is defined as

$$J_{\rho} = T\left(\hat{\rho}^{+} - \hat{\rho}^{-}\right)' \hat{\Omega}^{-1}\left(\hat{\rho}^{+} - \hat{\rho}^{-}\right)$$

where $\hat{\Omega} = \sum_{l=1-T}^{T-1} k\left(l/p\right)\hat{\gamma}_l$ and k is a kernel function that assigns a suitable weight to each lag of order l, and p is the smoothing parameter or lag truncation order (see Hong et al. (2007) for more details).

Panel A: US Stock Market			Panel B: UK Stock Market		
BETA1&5	COSK1&5	COKT1&5	BETA1&5	COSK1&5	COKT1&5
48.471	40.246	44.363	56.249	38.655	46.367
[0.06]	[0.25]	[0.13]	[0.01]	[0.31]	[0.09]

Table 7: Testing Differences of Tail Dependences

This table reports the coefficients of lower tail dependence (LTD) and upper tail dependence (UTD) and the difference between them for all the portfolios pairs. The estimations are calculated by the parametric approach in McNeil et al. (2005). λ^L and λ^U denote LTD and UTD estimated by t copula. The p-values of testing $\lambda^L = \lambda^U$ are computed by a bootstrapping with 500 replications and reported in $[\cdot]$.

Portfolio	λ^L	λ^U	$\lambda^L - \lambda^U$
US BETA1&5	0.171	0.018	0.153 [0.02]
US COSK1&5	0.200	0.153	0.047 [0.53]
US COKT1&5	0.153	0.103	0.050 [0.13]
UK BETA1&5	0.104	0.018	0.086 [0.00]
UK COSK1&5	0.203	0.062	0.141 [0.01]
UK COKT1&5	0.137	0.052	0.085 [0.21]

Table 8: Backtesting of Value-at-Risk: Empirical Coverage Probability

This table presents ECP in Panel A and BPZ in Panel B for each BMS portfolio and VaR model. Bias summarises the average deviation of 12 portfolios from the nominal coverage probability, 1%, for each VaR model, and RMSE (Root Mean Square Error) summarises the fluctuation of the deviation across 12 portfolios for each VaR model,

$$Bias = \frac{1}{12}\sum_{p=1}^{12}\left(ECP_p - 1\%\right), \ RMSE = \sqrt{\frac{1}{12}\sum_{p=1}^{12}\left(ECP_p - 1\%\right)^2}.$$

BPZ counts the number of failures of the 99 percent VaR in the previous 250 VaR forecasts. Up to four failures, on average, the portfolio falls into the range of a "Green" zone. If the failures are five or more, the portfolio falls into a "Yellow" (5–9) or "Red" (10+) zone. The VaR model of which BPZ is "Green" zone is preferred. For the UK portfolios, we estimate the VaR models using 250 business days over the period 4 Jan. 2000 - 15 Dec. 2000, and compute the one-day-ahead forecast of the 99 percent VaR for 18 Dec. 2000. We conduct rolling forecasting by moving forward a day at a time and end with the forecast for 31 Dec. 2012. This generates 3,033 out-of-sample daily forecasts. Next we repeat the same process for the US portfolios. It starts with the forecast for 18 Dec. 2000 and ends with the forecast for 31 Dec. 2012. This generates 3,018 out-of-sample daily forecasts.

	Panel A: ECP				Panel B: BPZ			
Portfolio	GHST	ST	DCC	FHS	GHST	ST	DCC	FHS
$BMS\,(Beta, L; US)$	0.86%	0.86%	1.46%	1.09%	Green	Green	Green	Green
$BMS\,(Cosk, L; US)$	1.03%	1.06%	1.36%	1.56%	Green	Green	Green	Green
$BMS\,(Cokt, L; US)$	0.86%	0.89%	1.19%	1.36%	Green	Green	Green	Green
$BMS\,(Beta, S; US)$	1.13%	1.19%	1.39%	1.26%	Green	Green	Green	Green
$BMS\,(Cosk, S; US)$	0.93%	0.99%	1.42%	1.23%	Green	Green	Green	Green
$BMS\,(Cokt, S; US)$	1.03%	0.86%	1.29%	1.29%	Green	Green	Green	Green
$BMS\,(Beta, L; UK)$	0.92%	0.86%	1.13%	1.39%	Green	Green	Green	Green
$BMS\,(Cosk, L; UK)$	0.96%	0.92%	1.19%	1.16%	Green	Green	Green	Green
$BMS\,(Cokt, L; UK)$	0.99%	0.99%	1.19%	1.16%	Green	Green	Green	Green
$BMS\,(Beta, S; UK)$	0.89%	0.92%	1.06%	1.52%	Green	Green	Green	Green
$BMS\,(Cosk, S; UK)$	1.22%	1.25%	1.26%	1.33%	Green	Green	Green	Green
$BMS\,(Cokt, S; UK)$	0.92%	0.96%	1.36%	1.33%	Green	Green	Green	Green
Bias (*Green/Yellow/Red*)	-0.02%	-0.02%	0.28%	0.31%	12/0/0	12/0/0	12/0/0	12/0/0
RMSE	0.11%	0.13%	0.30%	0.34%				

Table 9: Backtesting of Value-at-Risk: Statistical Accuracy

This table presents statistical tests for the accuracy of VaR forecasts. Panel A reports the CC results. The CC test uses the LR statistic and it follows the Chi-squared distribution with two degrees-of-freedom under the null hypothesis. Panel B reports the DQ test results. The DQ test uses the Wald statistic and it follows the Chi-squared distribution with 6 degrees-of-freedom under the null hypothesis. For the UK portfolios, we estimate the VaR models using 250 business days over the period 4 Jan. 2000 - 15 Dec. 2000, and compute the one-day-ahead forecast of the 99 percent VaR for 18 Dec. 2000. We conduct rolling forecasting by moving forward a day at a time and end with the forecast for 31 Dec. 2012. This generates 3,033 out-of-sample daily forecasts. Next we repeat the same process for the US portfolios. It starts with the forecast for 18 Dec. 2000 and ends with the forecast for 31 Dec. 2012. This generates 3,018 out-of-sample daily forecasts. * indicates that the VaR model is rejected at the 5% significance level. GHST, ST, DCC and FHS denote "Generalized Hyperbolic Skewed Student's t copula", "Student's t copula", "DCC-GARCH" and "Filtered Historical Simulation", respectively.

	Panel A: The CC Test				Panel B: The DQ Test			
Portfolio	GHST	ST	DCC	FHS	GHST	ST	DCC	FHS
$BMS\,(Beta, L; US)$	3.44	1.06	6.91*	0.99	9.75	5.49	34.03*	16.73*
$BMS\,(Cosk, L; US)$	8.60*	8.33*	6.83*	8.20*	35.42*	33.83*	19.74*	12.25
$BMS\,(Cokt, L; US)$	6.07*	5.52	1.67	3.85	15.19*	14.17*	3.64	6.72
$BMS\,(Beta, S; US)$	4.01	4.22	4.43	2.34	10.84	10.36	6.17	4.55
$BMS\,(Cosk, S; US)$	1.42	1.05	6.11*	1.96	3.18	2.71	7.67	4.14
$BMS\,(Cokt, S; US)$	1.06	1.06	2.78	3.41	1.63	1.63	4.86	4.69
$BMS\,(Beta, L; UK)$	0.71	1.10	1.15	4.31	1.18	1.42	3.64	24.63*
$BMS\,(Cosk, L; UK)$	0.62	0.71	4.18	4.05	11.19	11.84	11.87	10.67
$BMS\,(Cokt, L; UK)$	0.74	0.60	1.88	1.34	2.54	2.54	3.24	3.18
$BMS\,(Beta, S; UK)$	0.87	0.71	0.78	11.46*	5.20	4.71	3.65	27.39*
$BMS\,(Cosk, S; UK)$	2.30	2.78	2.78	3.18	3.85	4.53	9.86	5.71
$BMS\,(Cokt, S; UK)$	0.71	0.62	4.55	3.18	1.18	1.17	25.44*	5.57
Number of Rejection	2	1	3	2	2	2	3	3

Table 10: Backtesing of Expected Shortfall

This table presents the mean absolute error (MAE) in Panel A and the mean squared absolute error (MSE) for each BMS portfolio and ES model. For the UK portfolios, we forecast the ES models using 250 business days over the period 4 Jan. 2000 - 15 Dec. 2000, and compute the one-day-ahead forecast of the 99 percent ES for 18 Dec. 2000. We conduct rolling forecasting by moving forward a day at a time and end with the forecast for 31 Dec. 2012. This generates 3,033 out-of-sample daily forecasts. Next we repeat the same process for the US portfolios. It starts with the forecast for 18 Dec. 2000 and ends with the forecast for 31 Dec. 2012. This generates 3,018 out-of-sample daily forecasts. The average MAE and MSE are reported at the bottom of this table. GHST, ST, DCC and FHS denote "Generalized Hyperbolic Skewed Student's t copula", "Student's t copula", "DCC-GARCH" and "Filtered Historical Simulation", respectively.

	Panel A: MAE				Panel B: MSE			
Portfolio	GHST	ST	DCC	FHS	GHST	ST	DCC	FHS
$BMS\,(Beta, L; US)$	0.0068	0.0073	0.0108	0.0100	0.0168	0.0178	0.0287	0.0237
$BMS\,(Cosk, L; US)$	0.0047	0.0042	0.0030	0.0055	0.0053	0.0054	0.0016	0.0083
$BMS\,(Cokt, L; US)$	0.0084	0.0058	0.0056	0.0081	0.0176	0.0130	0.0169	0.0234
$BMS\,(Beta, S; US)$	0.0054	0.0172	0.0149	0.0112	0.0253	0.0514	0.0497	0.0286
$BMS\,(Cosk, S; US)$	0.0112	0.0029	0.0094	0.0093	0.0282	0.0127	0.0155	0.0184
$BMS\,(Cokt, S; US)$	0.0099	0.0102	0.0060	0.0069	0.0219	0.0237	0.0069	0.0089
$BMS\,(Beta, L; UK)$	0.0070	0.0071	0.0073	0.0127	0.0105	0.0121	0.0200	0.0257
$BMS\,(Cosk, L; UK)$	0.0048	0.0043	0.0054	0.0049	0.0045	0.0033	0.0044	0.0042
$BMS\,(Cokt, L; UK)$	0.0071	0.0072	0.0096	0.0108	0.0148	0.0178	0.0330	0.0550
$BMS\,(Beta, S; UK)$	0.0058	0.0062	0.0099	0.0136	0.0119	0.0122	0.0155	0.0346
$BMS\,(Cosk, S; UK)$	0.0100	0.0101	0.0125	0.0139	0.0214	0.0215	0.0351	0.0394
$BMS\,(Cokt, S; UK)$	0.0046	0.0050	0.0067	0.0064	0.0046	0.0050	0.0068	0.0063
Average	0.0071	0.0073	0.0084	0.0094	0.0152	0.0163	0.0195	0.0230