Title
Return distributions and applications

Author
Kim, Young Do

Publication Date
2007

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Return Distributions and Applications

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Economics

by

Young Do Kim

Committee in charge:

Professor Allan G. Timmermann, Chair
Professor James D. Hamilton
Professor Ivana Komunjer
Professor Bruce E. Lehmann
Professor Rossen I. Valkanov

2007

The dissertation of Young Do Kim is approved, and it is acceptable in quality and form for publication on microfilm:

____

____

____

____

____
Chair

University of California, San Diego

2007

To Se-Hyun, Hailey, and my parents.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Firstly I would like to acknowledge Professor Allan G. Timmermann for his support as the chair of my committee and a teacher. Through multiple drafts and many long nights, his guidance has proved to be invaluable. His enthusiasm for academic life is both inspiring and contagious. I sincerely desire to be as good of an advisor and researcher as he is.

In addition to my advisor, I would like to thank my committee members James Hamilton, Ivana Komunjer, Bruce Lehmann, and Rossen Valkanov. Their helpful and significant comments have considerably improved my dissertation. UCSD has been a great place for a young researcher to come and study economics.

Hasung Jang at Korea University and Joosung Jun at Ewha Womans University in Seoul deserve my deep gratitude for showing what the life of an academic researcher is like, and for helping me to pursue it. Joosung is in particular has offered much support and advice over the last a few years.

I thank Se-Hyun Cho who has played numerous roles in my life as a lovely wife, a mother of our daughter, Hailey (Shi Yeon), and also as a graduate student in the Sociology department at UCSD. It is hard for me to imagine how I would have done everything without her. I also would like to thank my parents for supporting and encouraging me in my studies for many years. I would like to dedicate this dissertation to Se-Hyun, Hailey, and my parents.

Multiple generations of UCSD graduate students have made my time here both more enjoyable and productive, My thanks go to Tolga Cenesizoglu, Munechika Katayama, Deniz Kebabci, Kwang Hwan Kim, Yong-Gook Jung, Dong Jin Lee, Jeff Lin, Tatsuyoshi Okimoto, Young Joon Park, Jose Gonzalo Rangel, Karmen Suen, as well as many others.

VITA

11 November, 1973	Born, Busan, Republic of Korea
1999	B.A., Department of Economics, Korea University
1999-2000	Research Assistant, Korea University
1999-2000	Teaching Assistant, Korea University
1999	B.A., Department of Economics, Korea University
2001	M.S., Finance in Business School Korea University
2002-2003	Research Assistant for BK21 project Ewha Womans University
2002-2007	Teaching Assistant University of California, San Diego.
2003	M.A., Department of Economics University of California, San Diego.
2005, 2006	Instructor, Department of Economics University of California, San Diego.
2007	Ph.D., Department of Economics University of California, San Diego.

FIELDS OF STUDY

Major Field: Economics

Studies in Financial Economics
Professors Bruce E. Lehmann, Allan G. Timmermann

Studies in Econometrics
Professors James, D. Hamilton, Ivana Komunjer, Allan G. Timmermann

ABSTRACT OF THE DISSERTATION

Return Distributions and Applications

by

Young Do Kim
Doctor of Philosophy in Economics
University of California, San Diego, 2007
Professor Allan G. Timmermann, Chair

The work presented in this dissertation was motivated by the observation that return distributions are not-normally distributed. Under this circumstance, some economic conclusions based on normal and elliptical distributions could be altered. The three chapters of this dissertation investigate various economic problems using copula functions.

Chapter one studies the conditional time-varying dependence structure in international stock markets. By comparing the conventionally used linear dependence measure and various alternatives, this paper shows that some differences exist in the time path of dependence. Also, the 'correlation breakdown' phenomenon clearly shown in the linear measure is not obvious when the copula model is applied.

Chapter two examines the impact of changes in the joint distribution of asset returns on investors' portfolio holdings under a CRRA utility function. Using simulated returns with moments set to match actual data, I use linear projections to explore how much of the variation in portfolio weights can be explained by different moments of the return distribution. Simple linear decision rules suggest that expected returns can explain from 70% to more than 90% of portfolio holdings. When higher-order moments are added to the decision rule, I find that volatility and skewness are significant and add up to 10% to the explanatory power of the

linear projection, while kurtosis is insignificant in many cases. These results suggest a simple and robust procedure for portfolio choice. This choice is based on linear projections of portfolio weight on the first few (conditional) moments of the return distribution. In a series of out-of-sample forecasting experiments, I find that more information about risk factors may lead people to invest more aggressively. However, this might ruin the performance of the investment, perhaps due to a forecasting error.

Chapter three extends the complete conditional coverage test to multivariate cases to address whether different dependence structures are important for evaluating interval forecasts. In an application to international stock returns, I find that a GARCH-t model for the margins passes specification tests in most cases regardless of dependence structure, while the GARCH-Normal does not. However, there is little evidence that any specific dependence structure dominates others.

I

Dependence Structure in International Financial Markets: Evidence from Asian Stock Markets

I.1 Introduction

Dependence is one of the most important concepts to academic economists and practitioners, because for multivariate distribution it matters to portfolio choice and risk management. Although it is important, only the linear dependence measure has been widely used because of its simplicity. However, such a measure is only exhaustive under some very restrictive conditions, for example, elliptical distributions. If such conditions are not satisfied, the linear measure may fail to reveal true dependence among variables. The copula may be one possible way to overcome shortcomings of the linear measure. The main goal of this paper is to assess the dependence structure of international financial markets, especially in Asian stock markets, using various dependence measures that can capture more general measures of dependence than linear correlation.

Since 1990, technological development facilitated global portfolio management. As a result, there has been growing attention directed toward 'Emerging Markets.' Asian equity markets, which form one major part of the Emerging Markets, also became one of the fastest growing financial markets. However, in late 1997, the Asian market experienced a financial crisis. The crisis first started in Thailand, then spread rapidly, causing turbulence in other East Asian financial markets, such as Indonesia, South Korea, Malaysia, and the Philippines. There also were other regional financial crises during the 1990s, such as the Mexican crisis in 1994, the Russian default crisis in 1998, and the Brazil devaluation in 1999. This series of regional financial crises increased attention to the dependence structure between markets. People paid attention to dependence, because it is closely related to one of the most fundamental concepts in finance, diversification.

There is recent research in international finance and econometrics regarding dependence. Longin and Solnik (2001) find that international stock markets are more correlated in bear markets, using extreme correlation with a copula model. Ang and Chen (2002) and Ang et al. (2006) find that stock return correlation increases during market downturns, and that downside correlations are related to standard size and value factors, as well as momentum. However, recent applied math and econometrics research, such as Embrechts et al. (2002) and Patton (2004), analyze the importance of an alternative measure of risk and apply this alternative to asset allocation by comparing portfolio performances. As analyzed in Embrechts et al. (2002), the conventional dependence measure has a weakness. Rodriguez (2003) adopts the copula model with Markov switching parameters and finds evidence of changing dependence structures during periods of financial turmoil, and increased tail dependence and asymmetry in times of high volatility. If there are differences in dependence measured by the conventional measure and alternatives, then many conclusions based on the conventional measure should be re-considered and re-investigated. Therefore, it is important to take a look at whether such a difference exists.

As mentioned, this paper is devoted to Asian financial markets. Asian markets have grown quickly; therefore, there has been growing practical and academic attention to these regional markets. Also, these markets have experienced huge crises around 1997 and 1998. These events can give financial economists some interesting features regarding market co-movements and dependence structure, and so on. There are a couple of aims of this paper. The first one is an investigation of the differences between the linear measure and the alternative copula model. The basic model is related to Patton (2001), who models copulas with conditional time-varying parameters for the first time. Also, I utilize the dynamic conditional correlation multivariate GARCH model as a benchmark model of the linear measure. Second, I investigate the difference in movements between the general dependence and the tail dependence. Each measure has its own properties and interpretations. Finally, I investigate an asymmetric dependence effect of negative and positive shocks.

I find that there exists differences between the linear and the copula-based alternative measure. Specifically, when I closely examine the crisis period, copula models do not in general provide an increasing dependence result for the crisis period. Therefore this finding does not support the correlation breakdown phenomenon, which is widely supported by the linear measure. This paper also shows that there are different movements in general dependence parameters and tail dependence parameters. Finally, I find that the asymmetric dependence effect is not clear in this framework.

This paper is organized as follows. The next section provides an overview of the copula as a measure of dependence and a comparison of different estimation methods, with a brief simulation. In section 3, I provide a Dynamic Conditional Correlation Multivariate GARCH model as a benchmark and a time-varying model of estimation to apply to the Asian stock markets. Section 4 presents the empirical results and section 5 concludes.

I.2 Dependence Measures and Copula

I.2.A Correlation

In much of the economic literature, especially in the financial economic literature, correlation is one of the most important notions. (Linear) correlation is most frequently used in practice as a dependence measure between two or more random variables.[1] It is not going too far to say that almost all finance theories, such as asset pricing theory and portfolio theory, are developed based on linear correlation. However correlation is only one particular measure of stochastic dependence among many variables and the correlation coefficient as a measure of dependence should be used with caution. I summarize the basic properties of linear correlation below:

Definition I.2.1. *Let X and Y be a vector of random variables with nonzero finite variances. Then the linear correlation coefficient for X and Y is*

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} \tag{I.1}$$

where $Cov(X,Y) = E(XY) - E(X)E(Y)$, the covariance of X and Y

This linear correlation is a popular measure of dependence, since it has many advantages. First, it is easy to calculate. Second, it is also easy to manipulate under linear operations. Finally, it is invariant under strictly increasing *linear* transformations. One other crucial feature of correlation is it is a natural measure of dependence of spherical and elliptical distributions,[2] such as the multivariate normal and the multivariate t-distribution. However, empirical research shows that the distributions of the real world are hardly in this class of distribution. Usually, financial data are fat-tailed, leptokurtic and skewed; therefore, the correlation coefficient might be inadequate in these cases. Therefore as already noticed, correlation coefficients should be used with caution. Thus, it is worthwhile

[1] Embrechts et al. (2003) provide more dependence concepts. Other dependence concepts are presented such as perfect dependence, concordance, Kendall's tau and Spearman's rho, tail dependence as well as linear correlation.

[2] For a more formal definition of spherical and elliptical distributions, see Embrechts et al. (2002).

to re-examine conclusions in financial economic literature using the correlation coefficient, if it does not provide correct information about the true dependence.

I.2.B Copula

What is a copula?[3] In order to answer this question we need to know about the joint distribution function. Basically, the dependence between the real-valued random variables $X_1, ..., X_n$ is completely described by their joint distribution function

$$F(x_1, ..., x_n) = \Pr[X_1 \leq x_1, ..., X_n \leq x_n] \tag{I.2}$$

Their joint distribution function can be separated into two parts. One is the part where the dependence structure is described and the other is the part where the marginal distribution is described.

A copula is a special multivariate joint distribution. We can construct the copula by transforming the random variables $X_1, ..., X_n$ to standard uniform marginal distributions. For simplicity, suppose each of random variables $X_1, ..., X_n$ has a continuous marginal distribution(CDF), $F_1, ..., F_n$. Then, this transformation can be achieved by the following transformation $T : R^n \to R^n, (x_1, ..., x_n) \mapsto (F_1(x_1), ...F_n(x_n))$. The joint distribution function C of $(F_1(x_1), ..., F_n(x_n))$ is the so-called copula of the random variables $X_1, ..., X_n$. It follows that

$$\begin{aligned} F(x_1, ..., x_n) &= \Pr[F_1(X_1) \leq F_1(x_1), ..., F_n(X_n) \leq F_n(x_n)] \\ &= C(F_1(x_1), ..., F_n(x_n)) \end{aligned} \tag{I.3}$$

The following is the formal definition of copula

Definition I.2.2. *An n-dimensional copula is a function* $C : [0,1]^n \to [0,1]$ *such that 1. C is grounded and n-increasing. 2. C has margins* C_k*,* $k = 1, 2, ..., n$*, which satisfy* $C_k(u) = u$ *for all* u *in* $[0,1]$.

Equivalently, an n-copula is a function $C : [0,1]^n \to [0,1]$ with the following properties.

---
[3]For more formal definition and concepts, see Nelsen (1999).

1. For every $\mathbf{u}$ in $[0,1]^n$, $C(\mathbf{u}) = 0$ if at least one coordinate of $\mathbf{u}$ is 0, and $C(\mathbf{u}) = u_k$ if all coordinates of $\mathbf{u}$ are equal to 1 except $u_{k.}$
2. For all $(a_1, ..., a_n), (b_1, ..., b_n) \in [0,1]^n$ with $a_i \leq b_i$ for all i, $V_C([\mathbf{a}, \mathbf{b}]) \geq 0$, where $V_C([\mathbf{a}, \mathbf{b}])$ is the C-volume of $[\mathbf{a}, \mathbf{b}]$[4]

As mentioned above, the joint distribution can be separated into two parts: the dependence structure part and the marginal distribution part. This separating idea has led to the concept of a copula. The following theorem is known as Sklar's Theorem. This is the most important and fundamental theorem in the theory and application of copula.

Theorem I.2.3. *Let F be an n-dimensional distribution function with margins $F_1, ..., F_n$. Then there exists an n-copula C such that for all $\mathbf{x}$ in R^n,*

$$F(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n)) \tag{I.4}$$

If $F_1, ..., F_n$ are all continuous, then C is unique. Conversely, if C is an n-copula and $F_1, ..., F_n$ are distribution functions, then the function F defined above is an n-dimensional distribution function with margins $F_1, ..., F_n$.

From this theorem, we can see that for continuous multivariate distribution functions, the univariate margins and the multivariate dependence structure can be separated, and the dependence structure can be represented by a copula. Although the copula has not achieved popularity in the economic literature yet, there are many kinds of copulas that have been studied because of their nice properties.[5] In this paper I use some popular copula models, including the one-parameter family and the two-parameter family, which are used often in literature.[6] Next, I will introduce some properties of different copula models used in this paper.

[4] For the definition, see Embrechts et al. (2003).

[5] Joe (1997) summarizes many useful parametric families of copulas together with their properties. he includes 12 one-parametric, 10 two-parametric copulas, and some multivariate copulas as an extension of a (one-parameter) bivariate copula family.

[6] It is possible to use different name for the same copula.

I.2.B.a The Elliptical copula

Since the elliptical copula shares many of the properties of the multivariate normal or t distribution, it is useful to use for general comparison. There are, however, some drawbacks: elliptical copulas do not have closed form expressions and are restricted to radial symmetry. This copula family has been derived from certain families of multivariate distribution functions, so elliptical copulas are simply the distribution functions of componentwise transformed elliptically distributed random vectors. I present two main elliptical copulas belonging to the elliptical copula family and their properties.

Gaussian(Normal) copula If the univariate margins $F_1, \cdots, F_n$ are Gaussians, the random vector $X = (X_1, \cdots, X_n)$ is multivariate normal, and the dependence structure among the margins is described by the following copula function

$$C^N(u_1, \cdots, u_n; \Sigma) = \Phi_\Sigma(\Phi^{-1}(u_1), \cdots, \Phi^{-1}(u_n)) \tag{I.5}$$

where Φ_Σ denotes the joint distribution function of the n-variate standard normal distribution function with linear correlation. matrix Σ, and Φ^{-1} is the inverse of the standard univariate Gaussian. This copula form is called Gaussian copula. When $n = 2$, the copula expression can be written as

$$C^N(u, v; \rho) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left\{\frac{-(r^2 - 2\rho rs + s^2)}{2(1-\rho^2)}\right\} drds \tag{I.6}$$

where ρ is simply the usual linear correlation coefficient of the corresponding bivariate normal distribution, so $\rho \in (-1, 1)$. Remember that for nondegenerate Gaussian margins with finite variances, ρ is just the usual linear correlation matrix. But for nondegenerate non-Gaussian margins with finite variances, ρ does not correspond to the linear correlation matrix exactly [7].

t copula If the univariate margins $F_1, \cdots, F_n$ are Student $t-$distributions, the random vector $X = (X_1, \cdots, X_n)$ is multivariate t, and the dependence structure

[7] For more theoretical proof, see Embrechts et al. (2003).

among the margins is described by the following copula function

$$C^t(u_1, \cdots, u_n; \Sigma, \nu) = T_{\Sigma,\nu}(t_\nu^{-1}(u_1), \cdots, t_\nu^{-1}(u_n)) \tag{I.7}$$

where $T_{\Sigma,\nu}$ denotes the joint distribution function of the n-variate standard Student's t distribution function with linear correlation. matrix Σ and the degree of freedom ν, and t_ν^{-1} is the inverse of the Student's t distribution function. This copula form is called Student's t copula. When $n = 2$, the copula expression can be written as

$$C^t(u, v; \rho, \nu) = \int_{-\infty}^{t_\nu^{-1}(u)} \int_{-\infty}^{t_\nu^{-1}(v)} \frac{1}{2\pi\sqrt{(1-\rho^2)}} \left\{1 + \frac{r^2 - 2\rho rs + s^2}{2(1-\rho^2)}\right\}^{-(\nu+2)/2} drds \tag{I.8}$$

where ρ is simply the usual linear correlation coefficient of the corresponding bivariate t_ν-distribution if $\nu > 2$[8]. Like the Gaussian case, for nondegenerate Student's t margins with finite variances, ρ is just the usual linear correlation matrix. But for nondegenerate non-Student's t margins with finite variances, ρ does not exactly correspond to the linear correlation matrix.

I.2.B.b The Archimedean copula

In contrast to the elliptical copula, the Archimedean copulas have closed form expressions, allowing for a great variety of different dependence structures. Also, it is relatively easy to construct and many parametric families of copulas belong to this class. The general definition of Archimedean copula is the following:

Definition I.2.4. *Let $\varphi^{[-1]}$ be the pseudo-inverse function of φ. An Archimedean copula can be written in the following way*

$$C(U_1, ..., U_n) = \varphi^{[-1]}[\varphi(U_1) + ... + \varphi(U_n)] \tag{I.9}$$

for all $0 \le U_1, ..., U_n \le 1$ and where φ is a continuous function, $\varphi : [0, 1] \to [0, \infty]$, termed a generator, of the copula satisfying:

[8] When $\nu \le 2$, we cannot define the covariance matrix of Student's t distribution. In this case we just interpret ρ as being the shape parameter of the distribution.

1. $\varphi(1) = 0$

2. $\varphi(t)$ *is a strictly decreasing function,* $\varphi'(t) < 0$ *for all* $t \in (0, 1)$

3. $\varphi(t)$ *is a convex function,* $\varphi''(t) \geq 0$ *for all* $t \in (0, 1)$[9]

This type of copula has nice properties to capture dependence in the distribution tail. I summarize the concept of tail dependence and then I present the most-used Archimedean copulas and their properties. Tail dependence[10] is a concept related to the amount of dependence in the upper-right-quadrant tail or lower-left-quadrant tail of a multivariate distribution. This is very useful for investigating the dependence of extreme values. In many financial applications, especially in the financial contagion literature, a main concern is extreme events rather than normal events. Therefore tail dependence gives more insight into analyzing a high volatility period or some crises. An intuitive explanation of tail dependence is a probability measure of extreme events, for instance, markets crashing together or booming together.

Let (X, Y) be a vector of continuous random variables with marginal distribution functions F and G. The coefficient of upper tail dependence of (X, Y) is

$$\lim_{u \to 1} \Pr\{Y > G^{-1}(u) | X > F^{-1}(u)\} = \lambda_U \tag{I.10}$$

provided that the limit $\lambda_U \in [0, 1]$ exists. Then, X and Y have asymptotic upper tail dependence if $\lambda_U \in (0, 1]$, and asymptotic upper tail independence if $\lambda_U = 0$. Since $\Pr\{Y > G^{-1}(u) | X > F^{-1}(u)\}$ can be written as

$$\frac{1 - \Pr\{X \leq F^{-1}(u)\} - \Pr\{Y \leq G^{-1}(u)\} + \Pr\{X \leq F^{-1}(u), Y \leq G^{-1}(u)\}}{1 - \Pr\{X \leq F^{-1}(u)\}} \tag{I.11}$$

the alternative definition is the following

Definition I.2.5. *If a bivariate copula* C *is such that*

$$\lambda_U = \lim_{u \to 1} \frac{(1 - 2u + C(u, u))}{1 - u} \tag{I.12}$$

[9] *For some background on bivariate Archimedean copulas see Genest and Rivest (1993). Also nice references on this family copula in general is Nelsen (1999).*

[10] This is based on Embrechts et al. (2003). Also see Joe (1997).

exists, then C has upper tail dependence if $\lambda_U \in (0,1]$, and upper tail independence if $\lambda_U = 0$.

The concept of lower tail dependence can be defined in a similar way. The coefficient of lower tail dependence of (X,Y) is

$$\lim_{u\to 0} \Pr\{Y < G^{-1}(u)|X < F^{-1}(u)\} = \lambda_L \tag{I.13}$$

provided that the limit $\lambda_L \in [0,1]$ exists. Then X and Y has asymptotic lower tail dependence if $\lambda_U \in (0,1]$, and asymptotic lower tail independence if $\lambda_L = 0$. Since $\Pr\{Y < G^{-1}(u)|X < F^{-1}(u)\} = \Pr\{Y \leq G^{-1}(u)|X \leq F^{-1}(u)\}$ can be written as

$$\frac{\Pr\{X \leq F^{-1}(u), Y \leq G^{-1}(u)\}}{\Pr\{X \leq F^{-1}(u)\}} \tag{I.14}$$

The alternative definition is the following:

Definition I.2.6. *If a bivariate copula C is such that*

$$\lim_{u\to 0} \frac{C(u,u)}{u} = \lambda_L \tag{I.15}$$

exists, then C has lower tail dependence if $\lambda_L \in (0,1]$, and lower tail independence if $\lambda_L = 0$.

In the Gaussian copula, we can prove $\lambda_U = 0$ for $\rho < 1$ using a standard result in the statistic theory[11]. Given the radial symmetry property of the Gaussian distribution, the lower tail dependence also is null confirming the tail independence in the Gaussian copula. In Student's t copula, the tail dependence is $\lambda_U = 2 - 2t_{\nu+1}\left[\sqrt{\nu+1}\cdot\frac{\sqrt{1-\rho}}{\sqrt{1+\rho}}\right] = \lambda_U$, which is increasing in ρ and decreasing in ν. As the number of degrees of freedom goes to infinity, λ_U tends to 0 for $\rho < 1$.

Frank Let $\varphi(t) = -\ln\frac{e^{-\delta t}-1}{e^{-\delta}-1}$, where $\delta \in R\backslash\{0\}$. When $n = 2$,this gives the Frank Copula form as the following

$$C^F(u,v;\delta) = -\frac{1}{\delta}\ln\left(1 + \frac{(e^{-\delta u}-1)(e^{-\delta v}-1)}{e^{-\delta}-1}\right), \delta \in R\backslash\{0\} \tag{I.16}$$

[11] For a proof see Embrechts et al. (2003)

The Frank copula is a strict[12] Archimedean copula and characterized by upper and lower tail independence. In this copula $\delta \to 0$ implies the independence copula, i.e. $C^F(u,v;\delta) = uv$. $\delta \to \infty$ and $\delta \to -\infty$ implies the upper Frèchet-Hoeffding bound and the lower Frèchet-Hoeffding bound, respectively.[13] is attained. But lower and upper tail dependence is zero($\lambda_U = 0$ and $\lambda_L = 0$). In general, the basic feature of the Frank copula is symmetric and it assigns zero probability to events that are deep in the both tails.

Gumbel Let $\varphi(t) = (-\ln t)^\delta$, where $\delta \in [1,\infty)$. When $n = 2$, this gives the Gumbel Copula form[14] as following

$$C^G(u,v;\delta) = \exp\left\{-\left[(-\ln u)^\delta + (-\ln v)^\delta\right]^{\frac{1}{\delta}}\right\}, \delta \in [1,\infty) \tag{I.17}$$

The Gumbel copula is a strict Archimedean copula and characterized by upper tail dependence and lower tail independence. In this copula $\delta = 1$ implies the independence copula, i.e. $C^G(u,v;\delta) = uv$, and as $\delta \to \infty$, the upper Frèchet-Hoeffding bound is attained. Lower tail dependence is zero($\lambda_L = 0$) and upper tail dependence is $2 - 2^{1/\delta}(\lambda_U = 2 - 2^{1/\delta})$. In general, the basic feature of the Gumbel copula is asymmetry and it assigns more probability mass to events in the upper right tail.

Clayton (or generalized Cook and Johnson) Let $\varphi(t) = \frac{t^{-\delta}-1}{\delta}$, where $\delta \in [-1,\infty)\backslash\{0\}$. When $n = 2$, this gives the Clayton Copula form as following

$$C^C(u,v;\delta) = \max\left\{\left(u^{-\delta} + v^{-\delta} - 1\right)^{-\frac{1}{\delta}}, 0\right\}, \delta \in [-1,\infty)\backslash\{0\} \tag{I.18}$$

The Clayton copula is characterized by upper tail independence and lower tail dependence. In this copula $\delta \to 0$ implies the independence copula, i.e. $C^C(u,v;\delta) =$

[12]If $\varphi(0) = \infty$, the φ is called as a strict generator and in this case, $\varphi^{[-1]} = \varphi^{-1}$ and $C(U_1,...,U_n) = \varphi^{-1}[\varphi(U_1) + ... + \varphi(U_n)]$ is said to be a strict Archimedean copula.

[13]The upper Frèchet-Hoeffding bound is $Min(u,v)$. This can be interpreted as perfect positive dependence between random variables. Also the lower Frèchet-Hoeffding bound is $min(u+v-1,1)$ which can be interpreted as perfect negative relationship.

[14]Longin and Solnik (2001) investigate the extreme correlation of international equity markets assuming each marginal distribution as extreme value distribution(Generalized Pareto Distribution) and find asymmetric correlation in right and left tail dependence. But the multivariate(more exactly bivariate) dependence function they use is nothing but one of copulas, the Gumbel copula.

uv. As $\delta \to \infty$, the upper Frèchet-Hoeffding bound is attained, and As $\delta = -1$, the lower Frèchet-Hoeffding bound is attained. Lower tail dependence is $2^{-1/\delta}(\lambda_L = 2^{-1/\delta})$ and upper tail dependence is zero($\lambda_U = 0$). In general, the basic feature of Clayton copula is asymmetry and it assigns more probability mass to events in the lower left tail.

I.2.B.c Other copula families

There are, of course, many other types of copulas. One well-known way to construct copulas is using algebraic methods[15], which use an algebraic relationship between the joint distribution function and its margins. Here I introduce Plackett's copula.

Plackett In Plackett's copula, a measure of dependence is the cross product ratio, or odds ratio. Therefore this copula is restricted to the bivariate case. Using the probability transforms $u = F(x)$, $v = G(y)$, we can write the dependence measure as

$$\theta = \frac{C(u,v)[1-u-v+C(u,v)]}{[u-C(u,v)][v-C(u,v)]} \tag{I.19}$$

and solve for $C(u,v)$. Then, by the boundary condition for $C(u,0)$ and $C(u,1)$, we can get the following form

$$C^P(u,v;\theta) = \frac{1}{2\eta}\left(1+\eta(u+v)-\sqrt{(1+\eta(u+v))^2-4\theta\eta uv}\right), \theta \in [0,\infty)\backslash\{1\} \tag{I.20}$$

where, $\eta = \theta - 1$. In the Plackett copula, $\theta \to 1$ implies the independence copula and the limits as θ goes to 0 and to ∞ are the lower Frèchet-Hoeffding bound and the upper Frèchet-Hoeffding bound. Like the Gaussian and the Frank copula, the tail dependence of this copula is 0. by the definition, the upper tail dependence is

$$\lim_{u\to 1}\frac{1-2u+C(u,u)}{1-u} = \lim \frac{1-2u+\frac{1}{2\eta}\left(1+2u\eta-\sqrt{(1+2u\eta)^2-4\theta\eta u^2}\right)}{1-u} \tag{I.21}$$

[15] The other ways are the inversion methods, the geometric methods. For more theoretical explanation, see Nelsen (1999)

Solving this expression using L'Hospital's rule, $\lambda_U = 0$. By the symmetry property, the lower tail dependence also is 0.

Joe-Clayton This is one of two-parameter families and it is constructed by taking a particular Laplace transformation of a Clayton copula. Then, the functional form is

$$C^{JC}(u, v; \theta, \theta) = 1 - \left(1 - [(1 - \bar{u}^{\theta})^{-\delta} - (1 - \bar{v}^{\theta})^{-\delta} - 1]^{-1/\delta}\right)^{1/\theta}, \theta \geq 1, \delta > 0 \tag{I.22}$$

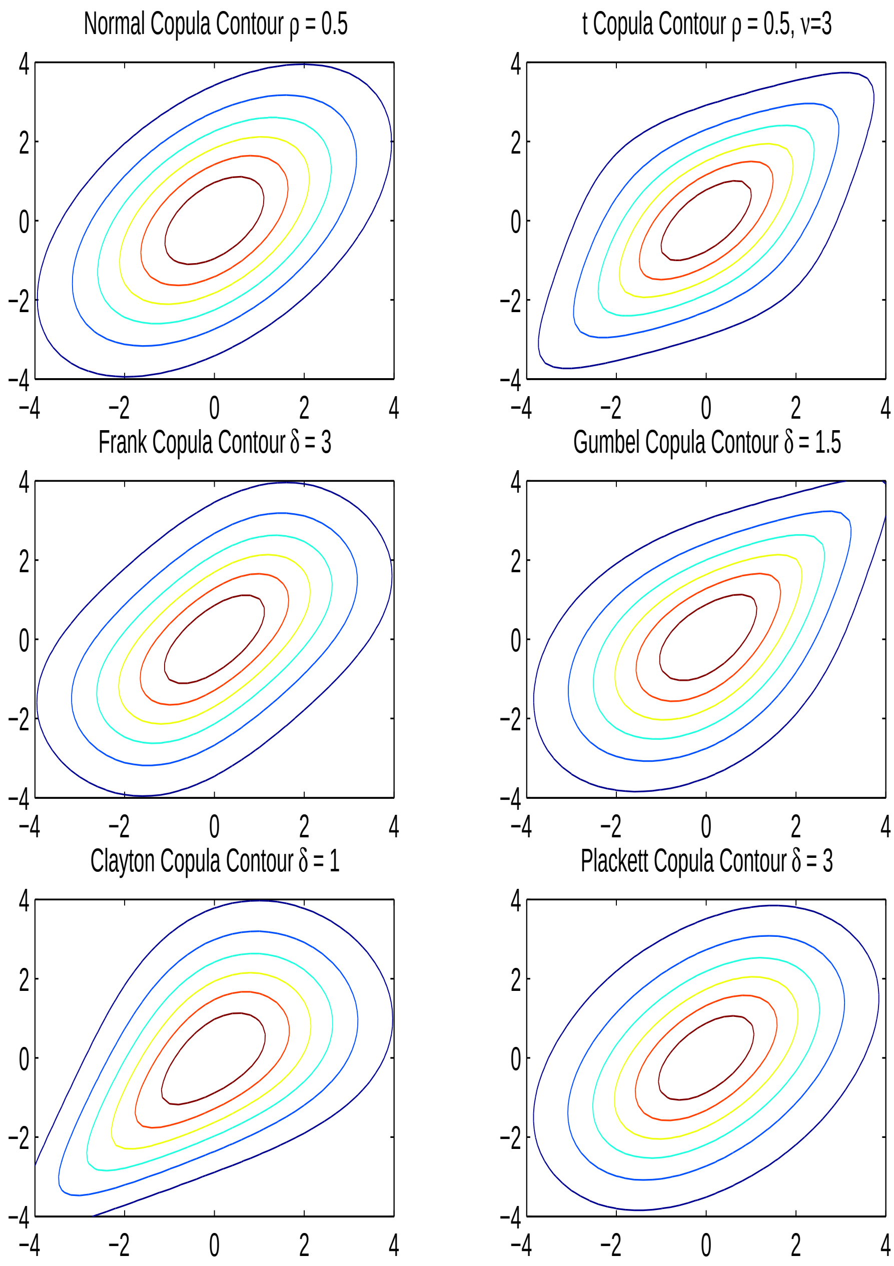
The Clayton copula is obtained when $\theta = 1$, and the Joe copula[16] is obtained as $\delta \to 0$. The upper Frèchet-Hoeffding bound obtains as $\theta \to \infty$ or $\delta \to \infty$. The lower tail dependence parameter is $2^{-1/\delta}$, independent of θ, and the upper tail dependence parameter is $2 - 2^{1/\theta}$, independent of δ.

Table I.1 summarizes the characteristics of each copula. Also, Figure I.1 shows the contour plots of different copulas. Basically the Elliptical copula family, the Frank and the Plackett are symmetric, and the Gumbel and the Clayton show a high tendency to happen together in upper-right and lower-left corner, respectively. However this tendency depends on each margin. Figure I.2 shows how different margins affects the whole joint distribution. Depending on which margins are used, we can see the different shape of the Gumbel copula contour plot. This is evidence of the copula's flexibility. In next section, I will discuss the model selection in copula application.

I.2.C Model Specification

Using the copula seems to raise an issue about which type of copula specification is right one among the many alternatives. This question is quite natural. Although many studies, such as Baig and Goldfajn (1999), use well-known distributions such as the multivariate normal distribution for modeling financial

[16]This copula is characterized by the upper tail dependence. For more information about this copula, see Joe (1997)



Note: This graph shows the contour plot of different copulas. The top two are the Elliptical copula and the middle and lower-left are the Archimedean copulas. The lower-right is the Plackett copula. Each copula has the normal margins, $N(0,2)$, for both axes.

Figure I.1 Contour Plot of Copulas

Table I.1 The Characteristics of Copulas

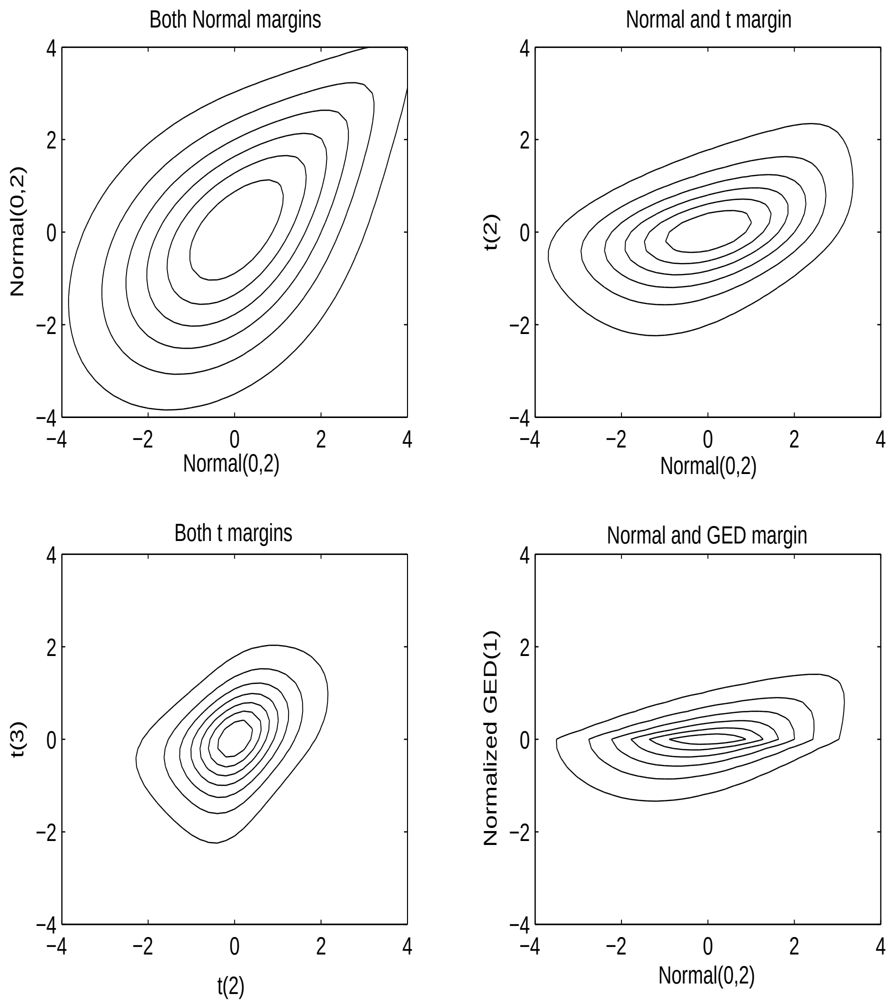
	upper tail dep.	lower tail dep.	C_U^*	C_I^*	C_L^*
Normal $C^N(u,v;\rho)$	0	0	$\rho=1$	$\rho=0$	$\rho=-1$
t $C^t(u,v;\rho,\nu)$	$2\bar{t}^{**}_{\nu+1}\left(\sqrt{\nu+1}\frac{\sqrt{1-\rho}}{\sqrt{1+\rho}}\right)$	$2\bar{t}^{**}_{\nu+1}\left(\sqrt{\nu+1}\frac{\sqrt{1-\rho}}{\sqrt{1+\rho}}\right)$	$\rho=1$	$\rho=0$	$\rho=-1$
Frank $C^F(u,v;\delta)$	0	0	$\delta\to\infty$	$\delta\to 0$	$\delta\to-\infty$
Gumbel $C^G(u,v;\delta)$	0	$2-2^{1/\delta}$	$\delta\to\infty$	$\delta=1$	-
Clayton $C^C(u,v;\delta)$	$-2^{-1/\delta}$	0	$\delta\to\infty$	$\delta\to 0$	$\delta=-1$
Plackett $C^P(u,v;\theta)$	0	0	$\theta\to\infty$	$\theta\to 1$	$\theta=0$

Note: C_U, C_I, and C_L are used for the upper Frèchet-Hoeffding bound, independence and the lower Frèchet-Hoeffding bound copulas, respectively.
$\bar{t}_{\nu+1}$ denotes the tail of a univariate t-distribution.

markets, there is skepticism about using those distributions, because these may not be correctly specified. In the copula literature, some authors, such as Embrechts et al. (2002), are paying attention to this problem. The copula is nothing but a joint distribution which has some nice properties, such as dividing the whole distribution into two parts: the marginal distribution and dependence structure. Therefore, one issue in the copula literature is choosing the 'right' or 'best' one that provides the best fit with the data. In addition, since some copulas have completely different dependence concepts, it is important to check whether some copula models are appropriate to fit the data dependence structure. This procedure usually can be done by goodness-of-fit tests statistically. Fermanian and Scaillet (2003) suggest an informal test through nonparametric estimation using a kernel method. Breymann et al. (2003) applied classical statistical tests such as χ^2 and the Anderson-Darling tests of a parametric copula specification to finance data. Also, Chen et al. (2004) develop two simple goodness-of-fit tests for dependence models. The first test is consistent but is recommended for testing the dependence structure between a small number of assets because it requires the estimation of a multivariate density function. Since I investigate the dependence structure of bivariate cases in this paper, I will provide the first test result in Section I.4[17].

Another issue here is how to specify each marginal distribution and how

[17] See Genest and Rivest (1993) for Archimedean type copulas test, Fermanian et al. (2002) for the limiting law of copula empirical processes, Fermanian (2003) for distribution free goodness-of-fitness test for *i.i.d* case, etc.



Note: This graph shows how different margins in the same copula affect the contour shape. Here the Gumbel copula with different margins shows a quite different shape. The upper-left uses both normal margins. The upper-right uses the Normal and the student's t margin. The lower-left uses both the student's t margins with different degree of freedom parameter. The lower-right uses the Normal and the normalized Generalized Error Distribution.

Figure I.2 Contour Plot of Gumbel Copula

to estimate the parameters in copulas. The copula is nothing but the joint distribution, having a marginal distribution part and a dependence part. Therefore, if we assume that the true copula belongs to a parametric family $\mathcal{C} = \{C_\theta, \theta \in \Theta\}$, then consistent and asymptotically normally distributed estimates of the parameter θ can be obtained through Maximum Likelihood Estimation (MLE) methods. There are mainly two estimation methods to be considered in this procedure: a parametric method and a semi-parametric method[18]. Parametric methods also can be done using two different procedures: a one-step and a two-step estimation estimation.

The one-step parametric method involves estimating the parameters of margins and copula together. Given an assumption of each marginal distribution and a parametric form of copula, the density of the copula function is

$$f(x_1;\alpha_1,\cdots,x_n;\alpha_n;\delta) = c(F_1(x_1;\alpha_1),\cdots,F_n(x_n;\alpha_n);\delta)\cdot\prod_{i=1}^{n} f_i(x_i;\alpha_i) \qquad \text{(I.23)}$$

where f_i is the density of the marginal distribution F_i, and α_i and δ are the parameters for each marginal distribution and for copula, respectively.

Suppose that we have data on n objects over T time periods which we are interested in and the parameter vector is $\theta = (\alpha_1,\cdots,\alpha_n;\delta)$. The log-likelihood function is the following:

$$l(\theta) = \sum_{t=1}^{T}\log(c(F_1(x_{1,t};\alpha_1),\cdots,F_n(x_{n,t};\alpha_n);\delta)) + \sum_{t=1}^{T}\sum_{i=1}^{n}\log(f_i(x_{i,t};\alpha_i)) \qquad \text{(I.24)}$$

The ML estimator $\hat{\theta}$ maximize equation (I.24) and it verifies the asymptotic normality.Durrleman et al. (2000)

$$\sqrt{T}(\widehat{\theta}_{ML} - \theta_0) \rightarrow N(0, I^{-1}(\theta_0)) \qquad \text{(I.25)}$$

where $I(\theta_0)$ is the Fisher Information matrix.

The two-step approach is sometimes called the Inference Functions for Margins method (IFM). In the two-step approach, the marginal distribution parameters are estimated in a first step, and in the second step, we optimize the copula likelihood for the copula parameter only. This procedure is less efficient than the one step procedure. However, when there are large numbers of parameters, the one-step approach result can be numerically misleading. The asymptotic normality of the two-step estimators is verified in Joe and Xu (1996).

$$\sqrt{T}(\widehat{\theta}_{IFM} - \theta_0) \rightarrow N(0, V^{-1}(\theta_0)) \qquad \text{(I.26)}$$

[18] for more formal detail about these two methods, see Genest et al. (1995)

with $V(\theta_0)$ the Godambe Information matrix. If we define the score function in the following way $g(\theta) = (\partial_{\alpha_1} l^1, \cdots, \partial_{\alpha_n} l^n, \partial_\delta l^c)$, the Godambe Information matrix is

$$V(\theta_0) = D^{-1} M (D^{-1})' \tag{I.27}$$

where $D = E[\partial g(\theta)'/\partial\theta]$ and $M = E[g(\theta)' g(\theta)]$. Note that the two step(IFM) method could be viewed as a special case of the GMM with an identity weight matrix.

The semiparametric estimation procedure, sometimes called the Canonical Maximum Likelihood (CML) method, is similar to the two-step procedure in terms of two-step estimation. But it differs from the parametric method because margins are left unspecified. This procedure reduces the risk on the right specification of all margins, although it suffers from a loss of efficiency. The estimation procedure of the semiparametric estimation is performed in two steps. In the first step, the dataset $(x_{1,t}, \cdots, x_{n,t}), t = 1, \cdots T$ is transformed into uniform variates $(\widehat{u}_{1.t}, \cdots, \widehat{u}_{n.t})$ using the rescaled empirical distributions, $\widehat{F}_n(\cdot)$, defined as follows:

$$\widehat{F}_n(\cdot) = \frac{1}{T+1} \sum_{t=1}^{T} 1_{\{X_{n,t} \leq \cdot\}} \tag{I.28}$$

where $1_{\{X_{n,t} \leq \cdot\}}$ is the indicator function. In the second step, the copula parameters are estimated as follows:

$$\widehat{\theta} = \arg\max \sum_{t=1}^{T} \log(c(\widehat{u}_{1.t}, \cdots, \widehat{u}_{n.t}); \theta)) \tag{I.29}$$

The proposed estimator is shown to be consistent, asymptotically normal and fully efficient in independence case[19], and the natural estimator of its asymptotic variance is proved to be consistent. If the marginal distributions are correctly specified, the two-step parametric and the semiparametric approaches give the same results.

[19] Genest and Rivest (1993) and Genest et al. (1995) show this asymptotic result for independent and identically distributed ($i.i.d$) data. But this is not hold for time-dependent data. Chen and Fan (2004) show the condition for asymptotically normal and consistent in the class of β-mixing case.

I.2.D Monte Carlo Results

In the parametric method, the specification of each margin is very important. Misspecified margins remain potential problem on the estimation of the copula parameter. Table I.2 shows the impact of misspecified margin on the copula parameter using Monte Carlo simulation. The design of the Monte Carlo simulation is the following. In the first simulation, the true model is a bivariate Normal copula with two AR-GARCH-Normal processes. The copula parameter, ρ, is equal to 0.3 and each AR-GARCH process is the following:

$$\begin{aligned} Y_{1t} &= 0.05 + 0.01Y_{1,t-1} + e_{1t} \\ e_{1t} &= \sqrt{h_{1t}} \cdot v_{1t}, \quad v_{1t} \sim i.i.d.\ N(0,1) \\ h_{1t} &= 0.10 + 0.50h_{1,t-1} + 0.30e^2_{1,t-1} \end{aligned} \tag{I.30}$$

$$\begin{aligned} Y_{2t} &= 0.08 + 0.02Y_{2,t-1} + e_{2t} \\ e_{2t} &= \sqrt{h_{2t}} \cdot v_{2t}, \quad v_{2t} \sim i.i.d.\ N(0,1) \\ h_{2t} &= 0.02 + 0.45h_{2,t-1} + 0.35e^2_{1,t-1} \end{aligned} \tag{I.31}$$

The sample size n equals to 100, 500 and 1000. The pseudo model for the margins is Gaussian $N(\mu, \sigma)$, that is, I estimate using a Gaussian instead of an AR-GARCH process for each margin. The estimation procedures used here are the one-step parametric, the two-step parametric, and the semiparametric method. This Monte Carlo simulation is done 1000 times.

The second simulation reflects a similar setting except that the true model is a Clayton copula ($\theta = 1.5$) with two Gaussian margins ($\mu = 0$, and $\sigma = 2$), and the pseudo model for the margins is a student-t distribution. Table I.2 summarizes this experimental result. According to this result, misspecified margins may lead to negative bias in small samples, and this does not disappear even in large samples. Although there is a tendency of upward bias in Panel B, using the nonparametric estimation for each margin leads to smaller bias. Also, its Mean Squared Error (MSE) is smaller than any other misspecified cases, and in large samples, MSE is close to the correct specified case. Therefore, when we are not sure about the

Table I.2 Effect of Misspecified Margins

Panel A	$\rho = 0.3$	Incorrect Margins			Correct Margins	
		One Step	Two Step	Semipara	One Step	Two Step
n=100	Bias	-0.0229	-0.0206	0.0032	0.0003	-0.0117
	MSE	0.0099	0.0099	0.0099	0.0090	0.0085
n=500	Bias	-0.0271	-0.0266	-0.0145	-0.0047	-0.0070
	MSE	0.0026	0.0025	0.0020	0.0017	0.0017
n=1000	Bias	-0.0278	-0.0276	-0.0178	-0.0060	-0.0071
	MSE	0.0017	0.0017	0.0012	0.0009	0.0009

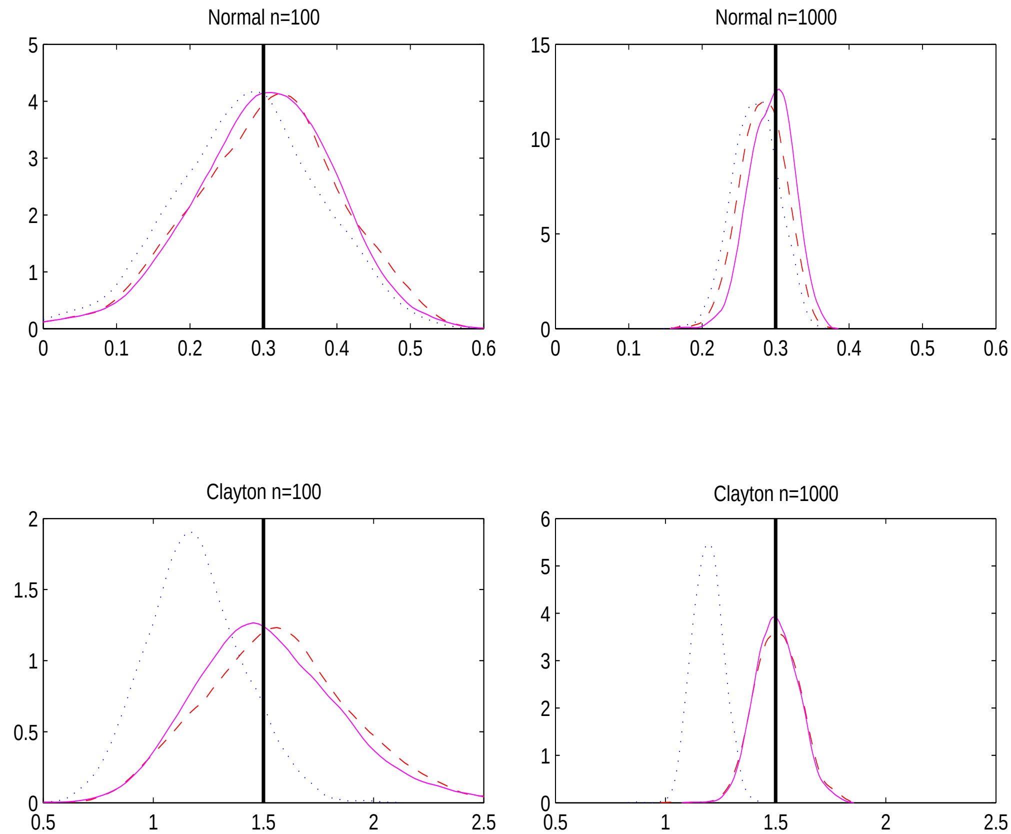
Panel B	$\theta = 1.5$	Incorrect Margins			Correct Margins	
		One Step	Two Step	Semipara	One Step	Two Step
n=100	Bias	-0.3120	-0.2439	0.0684	0.0220	0.0156
	MSE	0.1444	0.1195	0.1188	0.1047	0.1025
n=500	Bias	-0.3107	-0.2486	0.0123	0.0053	0.0029
	MSE	0.1054	0.0726	0.0193	0.0180	0.0178
n=1000	Bias	-0.3074	-0.2457	0.0093	0.0061	0.0049
	MSE	0.0991	0.0660	0.0107	0.0094	0.0094

correct marginal distributions, the semiparametric method offers a lot of gains and little loss in estimation result. Figure I.3 shows the density plot of this simulation. This also confirms that the incorrect margin assumption leads to bias which does not disappear even in large samples.

I.3 Estimation Models

The main purpose of this paper is to estimate and assess the dependence structure (parameter) of Asian markets using a copula model. Asian markets experienced a quite interesting rise and fall. Asian markets grew in the early 1990s quickly, and suddenly plummeted together. Figure I.4 shows how the Asian market indices changed during 1990s[20], and for the purpose of comparison later, I separate the sample into Thailand and the other countries. Each index is normalized to 100 on April 6th, 1990. As shown in Figure I.4, Thailand started to drop first around early 1997 and then other markets crashed together, nearly 50% in most cases. The

[20]These indices are from Data Stream index



Note: In each graph, dotted line is for incorrect margin assumption with one step estimation, dashed line is for semiparametric estimation, and solid line is for correct margin assumption with one-step estimation.

Figure I.3 Simulation Density Plot

Hong Kong index dropped from nearly 600 to less than 300. This phenomenon is named "Asian Flu." Asian index except Japan also shows a sudden drop, implying that this is not a one-country specific problem but a regional phenomenon. But can this co-movement be explained by close relationships among these indices (or markets)? There are a number of papers regarding this phenomenon. Specifically, some literature reports that there is a significant change in the correlation between the markets during the crisis period using the linear correlation coefficient.[21] But this conclusion could be from misleading evidence. As discussed in previous sec-

[21] One of literature about this problem is contagion literature. In order to detect the financial contagion between the countries, many papers try to capture "Correlation Breakdown", which means that a significant increase in cross-market relationship after a shock to one or a group of countries. King and Wadhwani (1990) and Baig and Goldfajn (1999) find that there is favorable evidence of correlation change between quiet and turbulent periods. But recent papers, such as Boyer et al. (1999) and Forbes and Rigobon (2002), show the opposite conclusion, after adjusting for upward bias in correlation coefficient.

tion, the linear correlation coefficient is only one kind of dependence measure. In addition if the joint distribution among variables is not a spherical or elliptical distribution, this conclusion could fail to capture what really happened in their dependence during this period. Therefore, this paper investigates the change in dependence structure of Asian markets using different dependence measures.

I.3.A Data Description

The data used here are seven Asian Stock Market indices: Indonesia, Hong Kong, S. Korea, Malaysia, Philippines, Thailand, and Taiwan. I use weekly indices from the Data stream country index from April 6th, 1990 to September 24th, 2004, so there are total 756 data points.[22] In this paper, I establish a bivariate model[23] between each country and Thailand, from which, people believe, the Asian financial crisis started.[24]

Table I.3 shows descriptive statistics of each series. As usual, I take the log-difference of each series and multiply by 100 to calculate each series return. There is a significant skewness, although some are positive and some are negative, and a significant kurtosis in each series. The most volatile market in terms of standard deviation is the Taiwan market and the next are Thailand, S. Korea and so on. As expected, every series strongly rejects the Jarque-Bera test, implying non-normality of the series. This is one reason why using multivariate normal distribution could be questionable.

I.3.B Model and Estimation Method

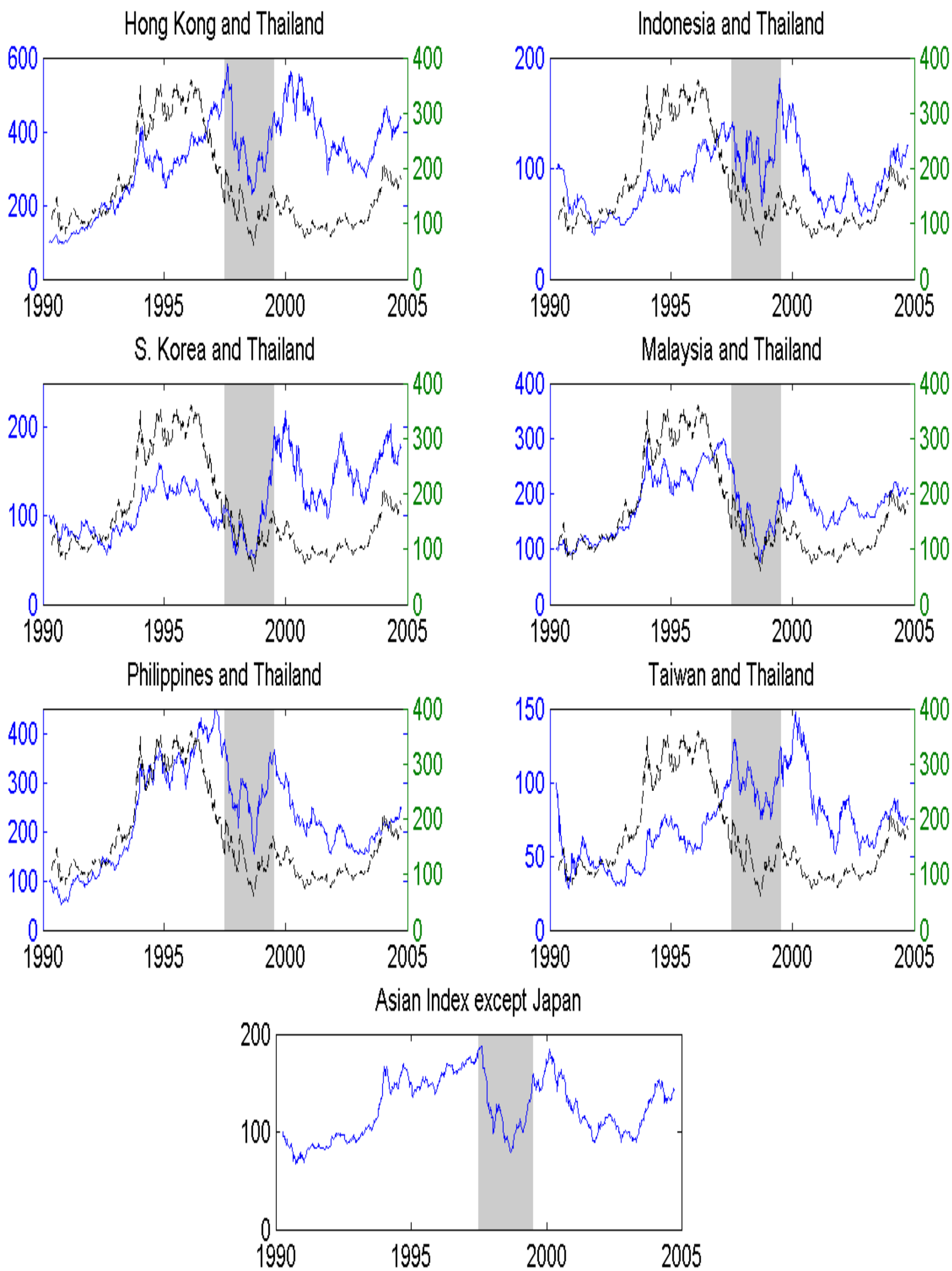
I.3.B.a Benchmark model

Asian markets seem to have high co-movement during the 1990s at the index level. Also, each market shows similar volatility patterns in Figure I.5. Some

[22]Most of literature use the main stock market index in their empirical test. But the correlation between the main stock market index and the Data stream country index is so high, therefore I assume that the coming result would be similar. Also because of data availability of some countries, this data set starts from April 6th, 1990

[23]Of course it is possible to extend to multivariate model.

[24]Rodriguez (2003) does the same thing. He estimates the copula parameter between Thailand and each other country for Asian markets and between Mexico and each other country for Latin-American markets.



Note: Each index is normalized to 100 on April 6th, 1990. These indices are weekly data and from Data Stream index data. Each graph of first three rows include the Thailand index scaled in right axis and the corresponding country's index scaled in left axis. The last row is Asian except Japan index.

Figure I.4 Asian Market Index

Table I.3 Descriptive Statistics Summary

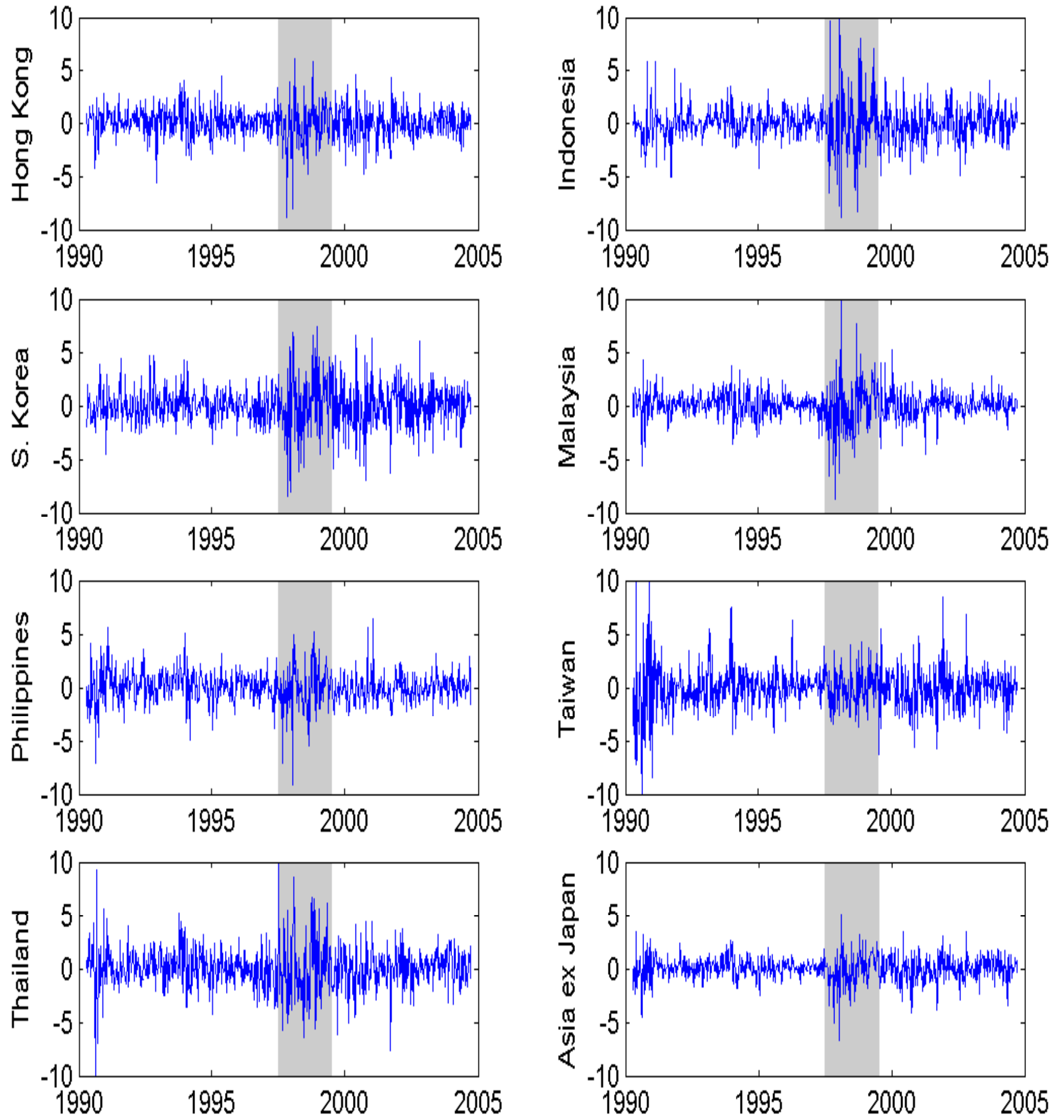
	IND	HK	KOR	MAL	PHI	TWN	THA	ASIA
OBS	755	755	755	755	755	755	755	755
Mean	0.0111	0.0848	0.0329	0.0428	0.0525	-0.0152	0.0345	0.0206
Median	-0.0271	0.1336	-0.0334	0.0605	0.0484	0.0312	0.0182	0.0606
Max	9.8681	6.1068	7.3602	10.0140	6.4019	11.1160	10.9730	5.0429
Min	-8.8741	-8.7794	-8.5012	-8.6603	-9.1786	-10.2990	-11.6760	-6.6433
Std. Dev	1.9052	1.5353	1.9867	1.4886	1.5488	2.0856	2.0796	1.1715
Skewness	0.2187	-0.4835	-0.0449	-0.0740	-0.3004	0.1475	0.0973	-0.5185
Kurtosis	7.1381	5.9980	4.8627	9.3608	6.5603	6.9559	6.5804	5.8655
J-B stat.	539.8	309.1	107.9	1263.6	406.2	490.4	400.5	289.2
J-B p-val.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Note: This table presents descriptive statistics of each country series. The sample period runs from April 6th, 1990 to September 24th, 2004.
Under the null hypothesis, the Jarque-Bera test statistics has a χ^2 distribution with degree of freedom 2.

markets, such as Indonesia, Taiwan, Thailand, were very volatile around 1990, but in general, during the first half of 1990s, each market was relatively tranquil. After that, they became so volatile, around 1997, starting from Thailand. Other markets became volatile after Thailand. People call this phenomenon "Asian Flu." However, one thing noticed here is that after 2 years since 1997 it became less volatile, although there are some differences in degree. Asian index except Japan shows less volatility, but more volatility compared to the first half of 1990s as a whole. Therefore, it is interesting to take a look at the dependence structure changing as time goes on.

I use the copula framework to detect the change in dependence structure. Of course, there is a huge literature in which researchers investigate the dependence structure using the correlation coefficient. In this paper, I use the Dynamic Conditional Correlation Multivariate GARCH model(DCC-MVGARCH) introduced by Engle and Sheppard (2001) and Engle (2002) as the benchmark model of the copula framework. According to correlation coefficient literature, there exists evidence of increases in correlation during the volatile period in sample[25]. More specifically,

[25] Although this result is skeptical in Boyer et al. (1999) and Loretan and English (2000), and etc, most academic literature and non-academic articles tend to accept that.



Note: This graph shows that Asian markets became highly volatile around 1997 and became less volatile but a bit more volatile comparing to the first half of 1990s as a whole.

Figure I.5 Asian Market Return from 1990 and 2004 some recent literature, such as Tse (2000) and Engle (2002), report that conditional correlation between equity returns is not constant. Since I use the conditional copula framework, previously introduced by Patton (2001), it is more attractive to use the conditional correlation coefficient rather than to use the unconditional correlation. In the DCC-MVGARCH model, univariate GARCH models are estimated for each asset series, and then, using the standardized residuals resulting from the first step, a time varying correlation matrix is estimated using a simple specification. This parameterization preserves the simple interpretation of univariate GARCH models, so it is easy to compute correlation estimator. Therefore, I use the DCC-MVGARCH model to parameterize the time varying correlation matrix. Once the univariate volatility models are estimated, the standardized residuals of each univariate series are used to estimate the dynamics of the correlation. The conditional correlation matrix in a DCC multivariate GARCH model is modeled as

$$\begin{gathered} r_t|\mathcal{F}_{t-1} \sim N(0, H_t) \\ H_t \equiv D_t R_t D_t \end{gathered} \tag{I.32}$$

where D_t is the $k \times k$ diagonal matrix of time varying standard deviations from univariate GARCH models with $\sqrt{h_{it}}$ on the ith diagonal, and R_t is the time varying correlation matrix. The proposed dynamic correlation structure is

$$\begin{gathered} Q_t = \left(1 - \sum_{m=1}^{M} \alpha_m - \sum_{n=1}^{N} \beta_n\right) \overline{Q} + \sum_{m=1}^{M} \alpha_m (\epsilon_{t-m}\epsilon'_{t-m}) + \sum_{n=1}^{N} \beta_n Q_{t-n} \\ R_t = Q_t^{*-1} Q_t Q_t^{*-1} \end{gathered} \tag{I.33}$$

where $\overline{Q}$ is the unconditional covariance of the standardized residuals resulting from the first stage estimation, and

$$Q_t^{*-1} = \begin{bmatrix} \sqrt{q_{11}} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{q_{22}} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{q_{kk}} \end{bmatrix}$$

so that Q_t^* is a diagonal matrix composed of the square root of the diagonal elements of Q_t. The typical element of R_t will be of the form $\rho_{ij,t} = q_{ij,t}/\sqrt{q_{ii}q_{jj}}$.

In this paper, the lag of DCC MVGARCH model is DCC(1,1), which means that $m = 1$ and $n = 1$ in equation (I.33) for simplicity. I do not report the parameter estimates for DCC-MVGARCH model in this paper as these are not of direct interest. However I will provide the time-varying conditional correlation plot in order to compare to the estimation result of time varying copula model. The other point that should be mentioned here is that I compare Thailand and each country here. Although it is possible to estimate a number of series together in DCC-MVGARCH model, it is more reasonable to match only a pair of countries for comparison purposes.

I.3.B.b Time-varying Dependence Model

Now, I introduce the copula model to estimate the dependence structure (each type copula parameter) for the Asian Markets during the 1990s. Simply by allowing time-variation in the dependence parameter, I investigate how volatile periods affect the dependence between markets, how persistent the shock effect is, as well as how different the dependence of copula is, compared to the linear conditional correlation represented by the DCC-MVGARCH model. The conditional copula model was first investigated by Patton (2001). Jondeau and Rockinger (2001) also use the conditional copula model. Recently, Dias and Embrechts (2003) also apply this method with high frequency data. For more theoretical justification about the conditional copula, see Patton (2001).

In order to estimate a bivariate (or multivariate) distribution, the first thing to do is to make an assumption about each univariate marginal distribution. After that, we can estimate the dependence structure using each copula. In this paper, I assume each marginal distribution follows three different processes. The first one is an AR(p)-GARCH(1,1) normal process for simplicity.[26] The AR(p)-GARCH(1,1) normal process is the following form:

[26] This assumption could be reconsidered. Although GARCH process can capture the time varying volatility of each series, each series has different feature which would not be considered in GARCH process. Table I.3 also show that there is a skewness and a high kurtosis.

$$
\begin{aligned}
y_{i,t} = \mu_i + \lambda_1 y_{i,t-1} + \cdots + \lambda_p y_{i,t-p} + \varepsilon_{i,t} \text{ for } i = 1, 2 \\
\varepsilon_{i,t} | I_{t-1} \sim N(0, h_{i,t-1}) \\
h_{i,t} = c_i + \alpha_i h_{i,t-1} + \beta_i \varepsilon_{i,t-1}^2
\end{aligned}
\tag{I.34}
$$

The second specification of each margin is an AR(p)-GARCH(1,1)-t process introduced by Bollerslev (1987). General financial return data shows fat-tailed and leptokurtic properties. Therefore, it is more reasonable to assume that each margin process is an AR(p)-GARCH(1,1)-t process rather than just an AR(p)-GARCH(1,1) normal process. The AR(p)-GARCH(1,1)-t process follows

$$
\begin{aligned}
y_{i,t} = \mu_i + \lambda_1 y_{i,t-1} + \cdots + \lambda_p y_{i,t-p} + \varepsilon_{i,t} \text{ for } i = 1, 2 \\
\sqrt{\frac{\nu}{h_{i,t}(\nu - 2)}} \varepsilon_{i,t} | I_{t-1} \sim t(\nu) \\
h_{i,t} = c_i + \alpha_i h_{i,t-1} + \beta_i \varepsilon_{i,t-1}^2
\end{aligned}
\tag{I.35}
$$

The last specification of each margin uses non-parametric estimation. Wrong specification of each margin could be a source of a bias in estimation, as seen in section I.2.C. The empirical distributions are calculated as the equation (I.28).

After estimating each series process using three alternatives, the next thing to do is to estimate the copula parameter. This is of course a two step estimation procedure. Although estimating all parameters together would give more efficient results, it is extremely hard to estimate all parameters at one time using maximum likelihood estimation with many parameters. In the simple bivariate case with a simple GARCH(1,1) normal process, for example, we have 9 parameters: 4 parameters from each GARCH(1,1) process and 1 dependence parameter from the copula. Further, when we allow a time-varying dependence parameter, then the number of parameters increases again. The dependency parameter(s) of the copula may be modeled as a convoluted expression of the parameters. Therefore, this two step estimation is convenient.

Estimating a copula parameter can be done by maximum likelihood estimation. Since the copula function is the same as the joint distribution function, it is easy to calculate the likelihood function. Let $u = F(x;\theta_x)$ and $v = G(y;\theta_y)$, where θ_x and θ_y are the vectors of parameters of each marginal distribution. Given $C(u,v;\delta) = C(F(x),G(y);\delta)$, the copula density is

$$c(u,v;\delta) = \frac{\partial^2 C(u,v;\delta)}{\partial u \partial v} \tag{I.36}$$

Then the joint density of an observation (x_t, y_t) is

$$\begin{aligned} c(x,y;\delta) &= \frac{\partial^2 C(u,v;\delta)}{\partial u \partial v} \cdot \frac{\partial u}{\partial x} \cdot \frac{\partial v}{\partial y} \\ &= c(u,v;\delta) \cdot f(x) \cdot g(y) \end{aligned} \tag{I.37}$$

The likelihood of each observation is the same as the joint density of an observation (x_t, y_t), so that the log-likelihood of a sample, as a consequence, becomes

$$\begin{aligned} L(x_t, y_t; \delta, \theta_x, \theta_y) &= \sum_{t=1}^{T} \ln(c(u_t, v_t, ; \delta) \cdot f(x_t) \cdot g(y_t)) \\ &= \sum_{t=1}^{T} \ln(c(F(x_t;\theta_x), G(y_t;\theta_y); \delta) \cdot f(x_t;\theta_x) \cdot g(y_t;\theta_x)) \end{aligned} \tag{I.38}$$

One thing noticed here is that the density estimation of each margin, $f(x_t;\theta_x)$ and $g(y_t;\theta_x)$) does not matter in estimating the copula parameter with a two-step estimation method. Since equation (I.38) can be written as

$$\begin{aligned} L(x_t, y_t; \delta, \theta_x, \theta_y) &= \sum_{t=1}^{T} \ln(c(u_t, v_t, ; \delta) \cdot f(x_t) \cdot g(y_t)) \\ &= \sum_{t=1}^{T} \{\ln c(F(x_t;\theta_x), G(y_t;\theta_y); \delta) + \ln f(x_t;\theta_x) + \ln g(y_t;\theta_x)\} \end{aligned} \tag{I.39}$$

and in fact each margin is estimated in the first step, thus margins are only constant in the second step. In order to estimate the copula parameter, therefore, we only need to maximize $\sum_{t=1}^{T} \ln c(F(x_t;\theta_x), G(y_t;\theta_y);\delta)$. But in this paper I use equation (I.38) in order to compare the likelihood value of each specification.

The following expressions are for each copula density.[27]

[27] Joe-Clayton copula density is long and complicated, so I do not present it here. See Andrew Patton's homepage(http://fmg.lse.ac.uk/ patton/).

1. Normal copula density

$$c_N(u,v;\delta) = \frac{1}{\sqrt{1-\rho^2}} \exp\left\{ \frac{\Phi^{-1}(u)^2 + \Phi^{-1}(v)^2 - 2\rho\Phi^{-1}(u)\Phi^{-1}(v)}{2(1-\rho^2)} + \frac{\Phi^{-1}(u)+\Phi^{-1}(v)}{2} \right\} \tag{I.40}$$

2. Frank copula density

$$c_F(u,v;\delta) = \delta\eta \frac{e^{-\delta(u+v)}}{[\eta - (1-e^{-\delta u})(1-e^{-\delta v})]^2} \tag{I.41}$$

where $\eta = 1 - e^{-\delta}$.

3. Gumbel copula density

$$c_G(u,v;\delta) = C_G(u,v;\delta)(uv)^{-1} \frac{(\tilde{u}\tilde{v})^{\delta-1}}{(\tilde{u}^\delta + \tilde{v}^\delta)^{2-1/\delta}} [(\tilde{u}^\delta + \tilde{v}^\delta)^{1/\delta} + \delta - 1] \tag{I.42}$$

where $\tilde{u} = -\ln u$ and $\tilde{v} = -\ln v$.

4. Clayton copula density

$$c_C(u,v;\delta) = (1+\delta)[uv]^{-\delta-1}(u^{-\delta} + v^{-\delta} - 1)^{-2-1/\delta} \tag{I.43}$$

Given these functional forms, first I estimate the copula parameter assuming it is constant. Second, I estimate the parameter(ρ or δ), assumed as time-varying, with the following form.

$$\begin{aligned} \delta_t &= \eta_0 + \eta_1\delta_{t-1} + f(u^*_{t-1}) + g(v^*_{t-1}) \\ f(u^*_t) &= \omega_1 u^*_t + \lambda_1 u^{*-}_t \\ g(v^*_t) &= \omega_2 v^*_t + \lambda_2 v^{*-}_t \end{aligned} \tag{I.44}$$

where $u^*_t = u_t - 0.5$, $v^*_t = v_t - 0.5$, $u_t = F(x_t;\theta_x)$, $v_t = G(y_t;\theta_y)$, $u^{*-}_t \equiv min[u^*_t, 0]$, and $v^{*-}_t \equiv min[v^*_t, 0]$.

Basically, this time-varying modeling is a modification of Patton (2001). But there are some different points, compared to the model used in Patton (2001) and others. The first thing I want to point out is the form of $f(u^*_t)$ and $g(v^*_t)$. I use two different evolution equations as equation (I.44). In this model δ_t evolves over time by the force of u_t^* and v_t^*. This may be a reasonable assumption, since u_t^* and v_t^* are the factors which constitute the copula distribution.[28] Unlike previous papers, I assume that each evolution force, u_t^* and v_t^* may have different effects on the dependence structure depending on its sign. In order to express different effects, I include u_t^{*-} and v_t^{*-} in the evolution equation. Since many recent papers find an asymmetric effect of positive and negative tails, it might be an improvement to model the asymmetric effect of u_t^* and v_t^*. Therefore, when $u_t^* \leq 0$, the effect on δ_t is $\omega_1 + \lambda_1$, and when $u_t^* > 0$, the effect on δ_t is just ω_1. For v_t^*, the same thing happens. This asymmetric information effect is the so-called 'leverage effect.' This kind of effect should be considered if time-variation also has an asymmetric response. In GARCH model extensions, this idea is reflected in the Threshold ARCH (TARCH) model, introduced in Glosten et al. (1993). The idea of TARCH models is to divide the distribution of the innovations into disjoint intervals and then approximate a piecewise linear function for the conditional standard deviation. If there are only two intervals, the division is normally at zero, i.e., the influence of positive and negative innovations on the volatility is differentiated. I apply this time-varying evolution equation to each copula to capture the leverage effect on the dependence structure in asset return. Also, separating $g(u_t^*)$ and $g(v_t^*)$ would give a clear picture of how u_t and v_t affect the dependence structure.

Finally, I want to mention the transformation of the evolution equation. Unlike Patton (2001) and others, I do not use the transformation of the evolution equation. This is because for almost all copula models, the parameter space is not the whole real number line, but is restricted to a certain range. For example, the normal copula dependence parameter should be in a range between -1 to 1, and the Gumbel copula dependence parameter should be in a range between 1 to ∞. Therefore, the transformation of the function is one way to solve the problem in case parameters go over a certain range. Patton (2001) and others use, for instance, $\delta_t = \Lambda(\eta_0 + \eta_1 \delta_{t-1} + f(u_{t-1}^*) + g(v_{t-1}^*))$, where Λ is any function where

[28] Here I use u_t^* and v_t^* instead of u_t and v_t. u_t and v_t are uniformly distributed in [0,1], therefore u_t^* and v_t^* are uniformly distributed in [-0.5,0.5] mapping from R. I found that using u_t^* and v_t^*gives more likelihood rather than using u_t and v_t in evolution equation.

the range is restricted to be within a certain interval, such as a logistic function. But if we can estimate this MLE procedure without the transformation, it would be better for interpreting the marginal effect of shocks and the persistence on the dependence structure directly. In this paper, I estimate equation (I.44) without any transformation, along with a numerical trick.

I.4 Estimation Results

I.4.A Simple Correlation

Before talking about the copula estimation, it is worthwhile to take a look at the linear correlation and the Spearman correlation of the data series. Table I.4 shows the linear correlation and the Spearman correlation among the data series. The linear correlation is one of many possibilities to evaluate dependence. As mentioned, to use it we must assume that the data in the pairs come from Normal distributions and the data are at least in the category of equal interval data. If these two conditions are not met, another possibility is to use the Spearman (Rank) Correlation Coefficient. Specifically the copula dependence parameter is easily converted to this rank correlation with closed form. The linear correlation coefficients range from less than 0.15 to more than 0.43 with different pairs. According to Table I.4, Thailand, where we believe the crisis was triggered, has relatively high correlations with other countries, showing values of more than 0.35 in most cases. The Spearman correlations range from about 0.13 to 0.40. This rank correlation is a little bit less than the linear correlation as a whole. The mnemonics are HK for Hong Kong, IND for Indonesia, KOR for South Korea, MAL for Malaysia, PHI for Philippines, TWN for Taiwan, and THA for Thailand.

Before going on to the estimation result, it is worthwhile to take a look at Table I.5. Table I.5 comes from Joe (1997, pp.146-7), and it gives an indication of the amount of dependence that exists as the δ varies for some copula models. For example, Kendall's τ 0.3 corresponds to 1.86, 1.25 and 0.86 for the Frank,

Table I.4 Correlation Coefficients among Countries

Correlation	HK	IND	KOR	MAL	PHI	TWN	THA	Asia
HK	1.0000							
IND	0.2885	1.0000						
KOR	0.3722	0.2424	1.0000					
MAL	0.4266	0.3471	0.2185	1.0000				
PHI	0.3876	0.4039	0.1797	0.4127	1.0000			
TWN	0.3210	0.1453	0.2628	0.2588	0.2662	1.0000		
THA	0.3988	0.3683	0.3771	0.4306	0.4229	0.2664	1.0000	
ASIA	0.8019	0.4078	0.5720	0.5745	0.4923	0.6409	0.5636	1.0000

Spearman	HK	IND	KOR	MAL	PHI	TWN	THA	ASIA
HK	1.0000							
IND	0.2600	1.0000						
KOR	0.3547	0.1941	1.0000					
MAL	0.3884	0.2935	0.2324	1.0000				
PHI	0.3170	0.3543	0.1633	0.3352	1.0000			
TWN	0.3110	0.1354	0.2728	0.2351	0.2627	1.0000		
THA	0.3598	0.3200	0.3067	0.3936	0.3657	0.2359	1.0000	
ASIA	0.7587	0.3624	0.5650	0.5377	0.4385	0.6164	0.5232	1.0000

Gumbel and Clayton copulas respectively. For all three cases, Spearman's ρ_s is greater than Kendall's τ for the corresponding value of δ.[29] In the sample, the Spearman's $\rho_(s)$ ranges from about 0.14 to about 0.4. 0.4 corresponds to 0.313 of a Normal copula parameter, 2.61 of a Frank copula parameter, and so on.

I.4.B Test and Estimation of the Marginal Model

As seen in section I.2.C, correct estimation of the marginal distribution is important, especially in two-step estimation. First, I select the different lags model for the mean equation. The variance equation remains the GARCH(1,1) for each country. The selection criterion is the Akaike Information Criterion(AIC). Using AIC, I finally choose AR(2) mean equation for Hong Kong, AR(7) for Indonesia, AR(4) for Korea, AR(7) for Malaysia, AR(3) for the Philippines, AR(4) for Taiwan, and AR(2) for Thailand. I do not report this estimation result, since

[29]Spearman's ρ_s and Kendall's τ are one of dependence measures. In brief both measure can be interpreted as concordance of random variables.

Table I.5 Parameter Values corresponding to given Kendall τ and Spearman ρ

τ	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Normal ρ	0.0	0.156	0.309	0.454	0.588	0.707	0.809	0.891	0.951	0.988	1.0
Frank δ	0	0.91	1.86	2.92	4.16	5.74	7.93	11.4	18.2	20.9	∞
Gumbel δ	1	1.11	1.25	1.43	1.67	2.00	2.50	3.33	5.00	10.0	∞
Clayton δ	0	0.22	0.50	0.86	1.33	2.00	3.00	4.67	8.00	18.0	∞
ρ_s	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Normal ρ	0.0	0.105	0.209	0.313	0.416	0.518	0.618	0.717	0.813	0.908	1.0
Frank δ	0	0.60	1.22	1.88	2.61	3.45	4.47	5.82	7.90	12.2	∞
Gumbel δ	1	1.07	1.16	1.26	1.38	1.54	1.75	2.07	2.58	3.73	∞
Clayton δ	0	0.14	0.31	0.51	0.76	1.06	1.51	2.14	3.19	5.56	∞

Note: This table is from Joe (1997).

it is not the main concern of this paper. After that, I test whether the marginal distribution assumption is correct using a variety of techniques. White (1982) proposed one easy way to test model misspecification. To conduct this test, all we have to do is the simple OLS regression of the constant unity on the vector, $(\nabla \ln \widehat{f}_{t|t-1}, [\text{vec}\nabla \ln \widehat{f'}_{t|t-1}(\nabla \ln \widehat{f}_{t-1|t-2}, \ldots, \nabla \ln \widehat{f}_{t-\lambda|t-\lambda-1})]'A')$,$t = 1, \ldots, n$, where $\ln \widehat{f}_{t|t-1}$ is the score vector evaluated at the MLE $\hat{\theta}$, and A is a finite non-stochastic $p \times k^2\lambda$ matrix. Here k is a number of parameters. Then, given assumption of some regularity conditions satisfied, nR^2 or the explained sum of squares of this repression $\overset{A}{\sim} \chi^2_p$, where R^2 is the squared multiple correlation coefficient of the OLS regression.[30]

White (1982) and Hamilton (1996) provided interpretation of this dynamic information matrix test in some special cases. For example, in the AR(p)-GARCH(1,1) normal case, there are total $4+p$ parameters, $1+p$ in mean equation and 3 in variance equation. The (1,1) element of $\nabla \ln \widehat{f'}_{t|t-1} \nabla \ln \widehat{f}_{t-1|t-2}$, which corresponds to a constant term in mean equation, is $\frac{\varepsilon_t}{h_t}\frac{\varepsilon_{t-1}}{h_{t-1}}$. Thus, $E\left(\frac{\varepsilon_t}{h_t}\frac{\varepsilon_{t-1}}{h_{t-1}}\right) = 0$ as a moment condition can be interpreted as a 1st-order autocorrelation test, which is itself interpretable as a dynamic information matrix test. By adding the (1,1) elements of $\nabla \ln \widehat{f'}_{t|t-1} \nabla \ln \widehat{f}_{t-\lambda|t-\lambda-1}$, $\lambda = 2, \ldots, k$, this test leads to tests that

[30] Under appropriate conditions this is asymptotically equivalent to the familiar Lagrange Multiplier(LM) statistic.

are asymptotically equivalent to LM statistics for kth-order autocorrelation. In a similar fashion, the (2+p,2+p) element of $\nabla \ln \widehat{f'}_{t|t-1} \nabla \ln \widehat{f}_{t-1|t-2}$, which corresponds to a constant term in variance equation, is $\dfrac{\varepsilon_t^2 - h_t}{2h_t^2}\dfrac{\varepsilon_{t-1}^2 - h_{t-1}}{2h_{t-1}^2}$. Therefore, $E\Big(\dfrac{\varepsilon_t^2 - h_t}{2h_t^2}\dfrac{\varepsilon_{t-1}^2 - h_{t-1}}{2h_{t-1}^2}\Big) = 0$ as a moment condition corresponds to ARCH(1) test under appropriate conditions. This is asymptotically equivalent to Engel's LM test for ARCH(1). Again by adjoining (2+p,2+p) elements of $\nabla \ln \widehat{f'}_{t|t-1} \nabla \ln \widehat{f}_{t-\lambda|t-\lambda-1}$, $\lambda = 2, \ldots, k$ we can get dynamic misspecification indicators which is asymptotically equivalent to LM tests for kth-order ARCH.

Second, in the AR(p)-GARCH(1,1) t case, we can obtain a similar interpretation using the (1,1) and (2+p,2+p) elements of $\nabla \ln \widehat{f'}_{t|t-1} \nabla \ln \widehat{f}_{t-\lambda|t-\lambda-1}$, $\lambda = 1, 2, \ldots, k$. The (1,1) element when $\lambda = 1$ is $\dfrac{(\nu+1)\varepsilon_t}{h_t(\nu-2)+\varepsilon_t^2}\dfrac{(\nu+1)\varepsilon_{t-1}}{h_{t-1}(\nu-2)+\varepsilon_{t-1}^2}$. Thus, $E\Big(\dfrac{(\nu+1)\varepsilon_t}{h_t(\nu-2)+\varepsilon_t^2}\dfrac{(\nu+1)\varepsilon_{t-1}}{h_{t-1}(\nu-2)+\varepsilon_{t-1}^2}\Big) = 0$ implies $E(\varepsilon_t\varepsilon_{t-1}) = 0$, and this moment condition corresponds to a 1st-order autocorrelation test. Again we can expand this test for kth-order autocorrelation by adding (1,1) elements of $\nabla \ln \widehat{f'}_{t|t-1} \nabla \ln \widehat{f}_{t-\lambda|t-\lambda-1}$, $\lambda = 2, \ldots, k$. Likewise, the (2+p,2+p) element when $\lambda = 1$ is $\dfrac{\nu\varepsilon_t^2 - h_t(\nu-2)}{2h_t(h_t(\nu-2)+\varepsilon_t^2)}\dfrac{\nu\varepsilon_{t-1}^2 - h_{t-1}(\nu-2)}{2h_{t-1}(h_{t-1}(\nu-2)+\varepsilon_{t-1}^2)}$ and again, the moment condition, $E\big(\dfrac{\nu\varepsilon_t^2 - h_t(\nu-2)}{2h_t(h_t(\nu-2)+\varepsilon_t^2)}\dfrac{\nu\varepsilon_{t-1}^2 - h_{t-1}(\nu-2)}{2h_{t-1}(h_{t-1}(\nu-2)+\varepsilon_{t-1}^2)}\big) = 0$ implies $E\big((\nu\varepsilon_t^2 - h_t(\nu-2))(\nu\varepsilon_{t-1}^2 - h_{t-1}(\nu-2))\big) = E\left(\nu^2\Big(\varepsilon_t^2 - h_t\dfrac{(\nu-2)}{\nu}\Big)\Big(\varepsilon_{t-1}^2 - h_{t-1}\dfrac{(\nu-2)}{\nu}\Big)\right) = 0$. Therefore, again, this moment condition leads to the test to detect ARCH(1) effect. The ARCH(k) effect can be tested by adding the lagged terms in the same way.

Table I.6 summarizes the dynamic information matrix test result. Generally speaking, this result tells that the GARCH-t distribution assumption is better than the GARCH-Normal distribution assumption. Since the null hypothesis of this test is no autocorrelation or no ARCH effect on the residuals, accepting the null leads to the conclusion of correctly specified model. In GARCH-t assumption, we cannot reject the null except a few cases, whereas we can reject the null in many cases with the GARCH-Normal assumption.

Another way to test the model misspecification can be done by suggestion of Diebold et al. (1998). In their paper, they suggested examination of, first, a less formal, but more revealing, graphical density estimate of the probability integral transform series and then, second, the correlogram of $(z-\bar{z})$, $(z-\bar{z})^2$, $(z-\bar{z})^3$, and $(z-\bar{z})^4$, where z is the probability integral transform series. Figure I.6 shows the graphical density estimate of the probability integral transform residuals from each country's model selection with 20 bins.[31] Basically, the histograms of the first row, GARCH Normal assumption, are close to uniform shape , but they still display slight peaks in the middle indicating that many of the realizations fall in the middle of density in some cases, such as Indonesia, Malaysia, etc. However, the second row, GARCH-t assumption, has a uniform shape, almost all within the confidential interval with a few exceptions. This result favors the GARCH-t distribution assumption rather than the GARCH Normal, in addition to previous results. Of course, the last row looks like exact uniform distribution, since that is the empirical CDF of each series.

The second step, suggested by Diebold et al. (1998), is an examination of the correlogram of $(z-\bar{z})$, $(z-\bar{z})^2$, $(z-\bar{z})^3$, and $(z-\bar{z})^4$. Each one will reveal dependence operative through the conditional mean, conditional variance, conditional skewness, or conditional kurtosis. Figure I.7 to Figure I.9 show the correlogram of $(z-\bar{z})$, $(z-\bar{z})^2$, $(z-\bar{z})^3$, and $(z-\bar{z})^4$.[32] In the GARCH Normal case, the correlograms shows that there is no serial correlation in the first four moments. Also, the GARCH-t case of Figure I.8 shows that this density specification may be adequate for the return series. The correlograms remain good with few exceptions. Combining the histogram and the correlogram, we may think GARCH-t specification is slightly better than GARCH Normal specification. Finally, Figure I.9 shows that this may not be the best choice with potential risk, because there exist huge autocorrelations, especially in the 2nd and 4th moments, in almost all

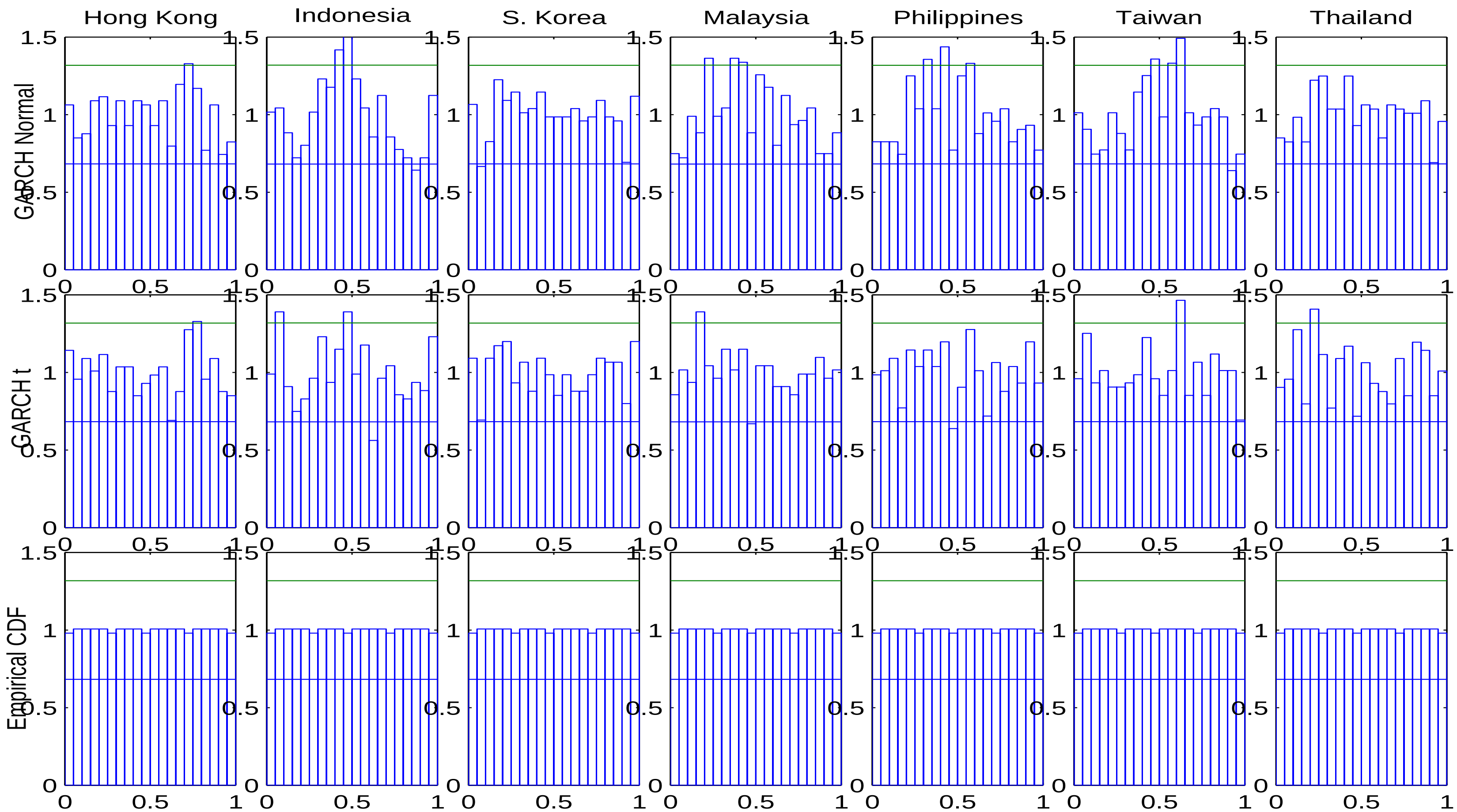
[31] I did the same thing with 40 bins. General results with 40 bins are similar to that with 20 bins.

[32] The two lines superimposed on the correlograms are Bartlett's approximate 95% confidence intervals under the null that z is *i.i.d.*

Table I.6 Model Specification Test using Dynamic Information Matrix Test

	lags	HK	IND	KOR	MAL	PHI	TWN	THA
Panel A: GARCH-Normal distribution								
Autocorrelation	1	0.8503	5.7113	4.0705	2.5521	0.3767	3.2056	10.0336
		(0.3565)	(0.0169)	(0.0436)	(0.1101)	(0.5394)	(0.0734)	(0.0015)
	5	3.7387	13.8467	12.4569	5.8859	2.2733	9.9303	15.3845
		(0.5876)	(0.0166)	(0.0290)	(0.3175)	(0.8102)	(0.0772)	(0.0088)
	10	14.4369	14.9453	13.4853	8.8019	6.1392	26.4379	23.7324
		(0.1540)	(0.1341)	(0.1978)	(0.5510)	(0.8034)	(0.0032)	(0.0083)
	15	23.4912	23.2111	30.1787	16.1799	10.2087	35.4794	28.0403
		(0.0743)	(0.0798)	(0.0113)	(0.3702)	(0.8064)	(0.0021)	(0.0213)
	20	27.2466	27.8538	31.3072	20.4749	20.6087	43.6089	38.4081
		(0.1285)	(0.1129)	(0.0513)	(0.4286)	(0.4205)	(0.0017)	(0.0079)
ARCH	1	1.8195	1.5894	2.4454	0.0572	2.5067	0.0617	4.0945
		(0.1774)	(0.2074)	(0.1179)	(0.8110)	(0.1134)	(0.8038)	(0.0430)
	5	5.1138	15.4305	7.5531	7.4327	9.1431	9.9205	15.7785
		(0.4022)	(0.0087)	(0.1826)	(0.1904)	(0.1035)	(0.0775)	(0.0075)
	10	16.0516	17.1467	10.3724	13.3393	20.3510	12.8134	19.1297
		(0.0982)	(0.0712)	(0.4085)	(0.2053)	(0.0261)	(0.2343)	(0.0386)
	15	26.7145	27.7635	12.9807	22.0452	25.2737	16.9328	33.6799
		(0.0312)	(0.0231)	(0.6038)	(0.1066)	(0.0464)	(0.3229)	(0.0038)
	20	34.0261	31.8023	14.9654	31.1824	41.0788	22.3142	37.3704
		(0.0259)	(0.0455)	(0.7784)	(0.0528)	(0.0036)	(0.3237)	(0.0106)
Panel: GARCH-t distribution								
Autocorrelation	1	1.3594	2.3207	4.9285	2.3232	3.4687	2.0905	7.1426
		(0.2436)	(0.1277)	(0.0264)	(0.1275)	(0.0625)	(0.1482)	(0.0075)
	5	3.6293	9.6214	11.0785	6.5169	3.8932	5.7060	9.9440
		(0.6039)	(0.0867)	(0.0498)	(0.2591)	(0.5649)	(0.3359)	(0.0768)
	10	15.9059	11.9922	12.5937	9.0589	9.7866	15.9688	15.3052
		(0.1024)	(0.2856)	(0.2473)	(0.5265)	(0.4594)	(0.1005)	(0.1213)
	15	24.5899	19.2460	23.8850	10.7577	11.5846	18.8961	21.7799
		(0.0557)	(0.2028)	(0.0671)	(0.7696)	(0.7102)	(0.2185)	(0.1137)
	20	29.5198	22.6549	26.3133	12.2195	16.0939	27.6695	28.1775
		(0.0780)	(0.3061)	(0.1557)	(0.9083)	(0.7108)	(0.1175)	(0.1053)
ARCH	1	0.8209	3.7689	2.9385	0.1762	0.2395	0.6855	0.8722
		(0.3649)	(0.0522)	(0.0865)	(0.6747)	(0.6246)	(0.4077)	(0.3503)
	5	2.0766	13.2397	7.5944	5.6007	6.5345	7.2713	7.6636
		(0.8384)	(0.0212)	(0.1801)	(0.3470)	(0.2576)	(0.2012)	(0.1758)
	10	9.8835	13.0269	10.8593	7.5626	16.6627	9.5036	10.4248
		(0.4508)	(0.2222)	(0.3686)	(0.6715)	(0.0822)	(0.4851)	(0.4040)
	15	19.5135	21.8472	13.0687	15.0509	17.2308	12.7375	26.5267
		(0.1914)	(0.1119)	(0.5970)	(0.4478)	(0.3053)	(0.6226)	(0.0328)
	20	24.6252	23.1322	15.0023	22.7164	27.0544	15.7621	37.4775
		(0.2161)	(0.2823)	(0.7763)	(0.3029)	(0.1337)	(0.7313)	(0.0103)

Note: This statistic follows $\chi^2(p)$, where p is the lags. Parentheses are p-value.



Note: z is the probability integral transform of residuals from each country's model selection.

Figure I.6 Estimate of the Density of Each z

countries[33].

I.4.C Test and Estimation of the Bivariate Model

In this section I present, first, the goodness-of-fit test result suggested by Chen et al. (2004). After that, dynamic dependence structure using the benchmark model, DCC-MVGARCH will be presented. Finally, by presenting another dependence structure using the conditional time-varying copula model, we can see the existence of different patterns in the models.

I.4.C.a Copula Specification test

Chen et al. (2004) develop two simple goodness-of-fit tests for some dependence models. These tests determine whether some dependence models are compatible with given data sets. The first one is a consistent test that requires the kernel estimation of a multivariate density function. This test is easily applied and computed when the dimension of data, d is small. The null hypothesis of this test is

$$H_0 : \Pr(C(U_1, \cdots, U_d) = C_0(U_1, \cdots, U_d; \alpha_0)) = 1 \text{ for some } \alpha_0 \in \mathcal{A}$$

The alternative is

$$H_1 : \Pr(C(U_1, \cdots, U_d) = C_0(U_1, \cdots, U_d; \alpha_0)) < 1 \text{ for all } \alpha_0 \in \mathcal{A}$$

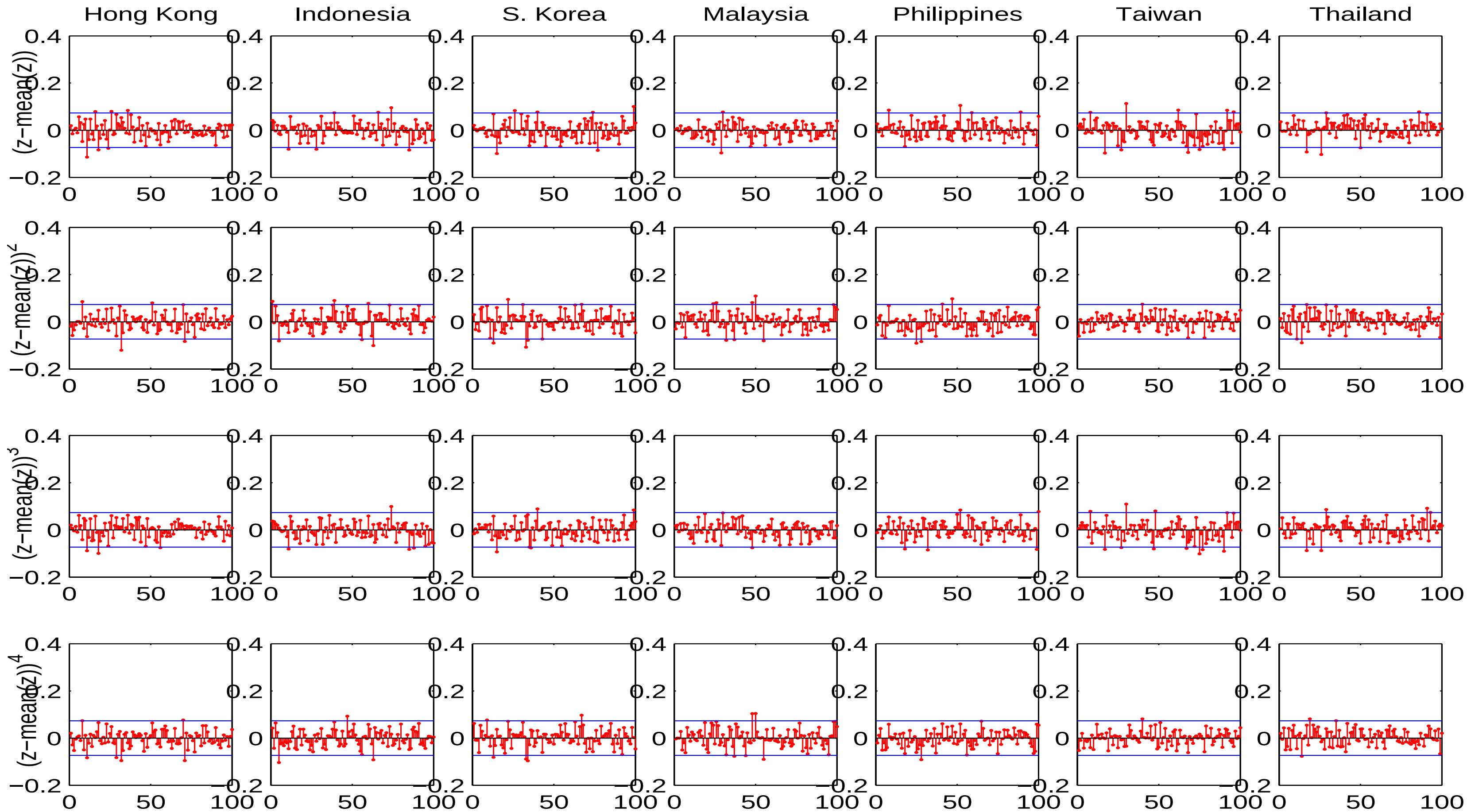
The test statistic can be computed as follows:

$$\begin{aligned} &\hat{Z}_{1,t} = \hat{F}_1(Y_{1,t}) \\ &\hat{Z}_{j,t} = C_{0j}(\hat{F}_j(Y_{j,t}; \hat{\alpha}|\hat{F}_1(Y_{1,t}), \cdots, \hat{F}_{j-1}(Y_{j-1,t}))),\ j = 2, \cdots, d,\ t = 1, \cdots, n \end{aligned} \tag{I.45}$$

where, $\hat{\alpha}$ is a $\sqrt{n}$-consistent estimator of α_0 under the null, and $\hat{F}_j$ is the empirical distribution function. After that, we can calculate test statistic, $\hat{I}_n$, as follows:

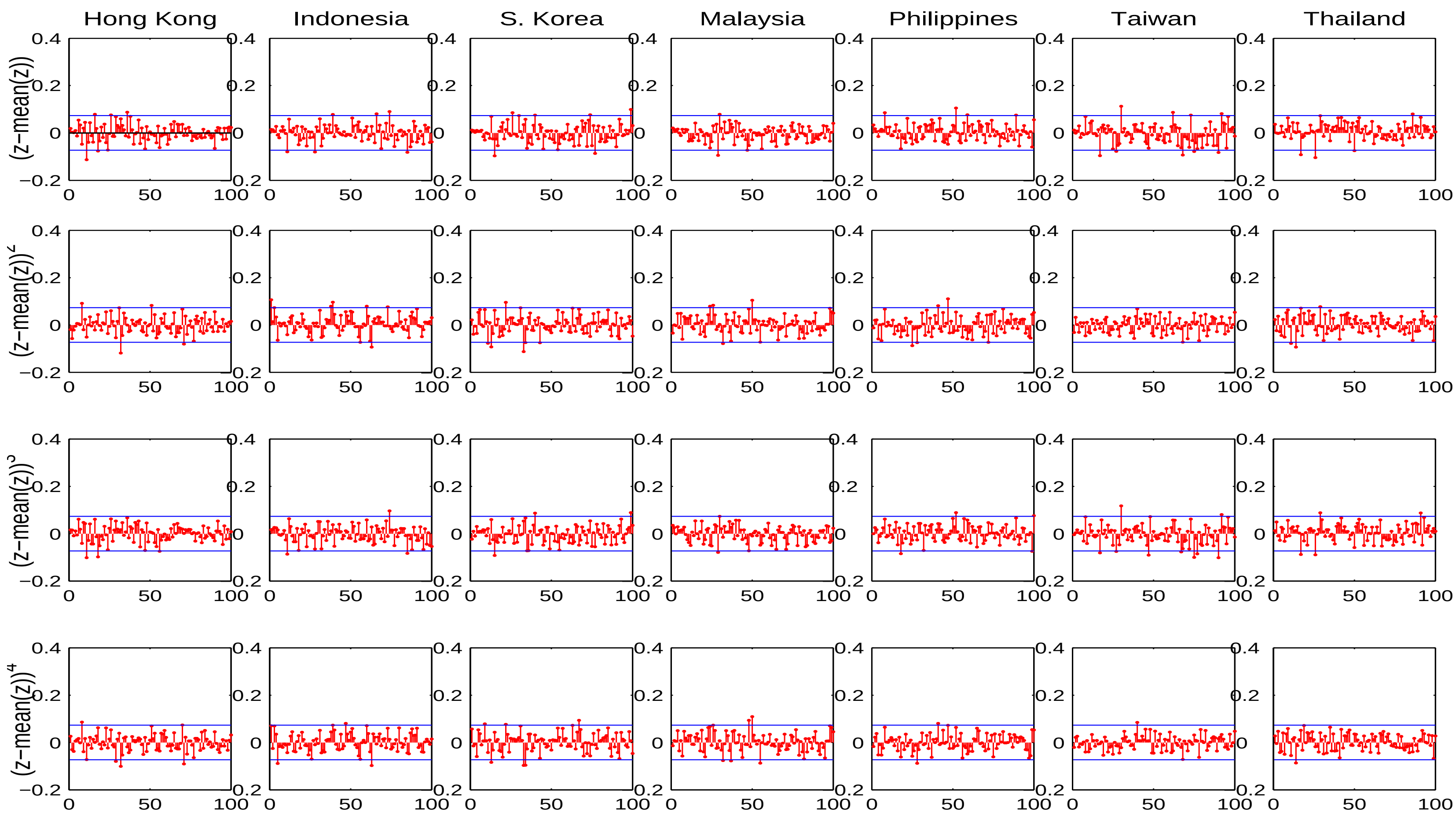
$$\hat{I}_n = \int_0^1 \cdots \int_0^1 [\hat{g}(z_1, \cdots, z_d) - 1]^2 dz_1 \cdots dz_d \tag{I.46}$$

[33] Potential risk of higher moment autocorrelation would be investigated later. I leave this for the future study.



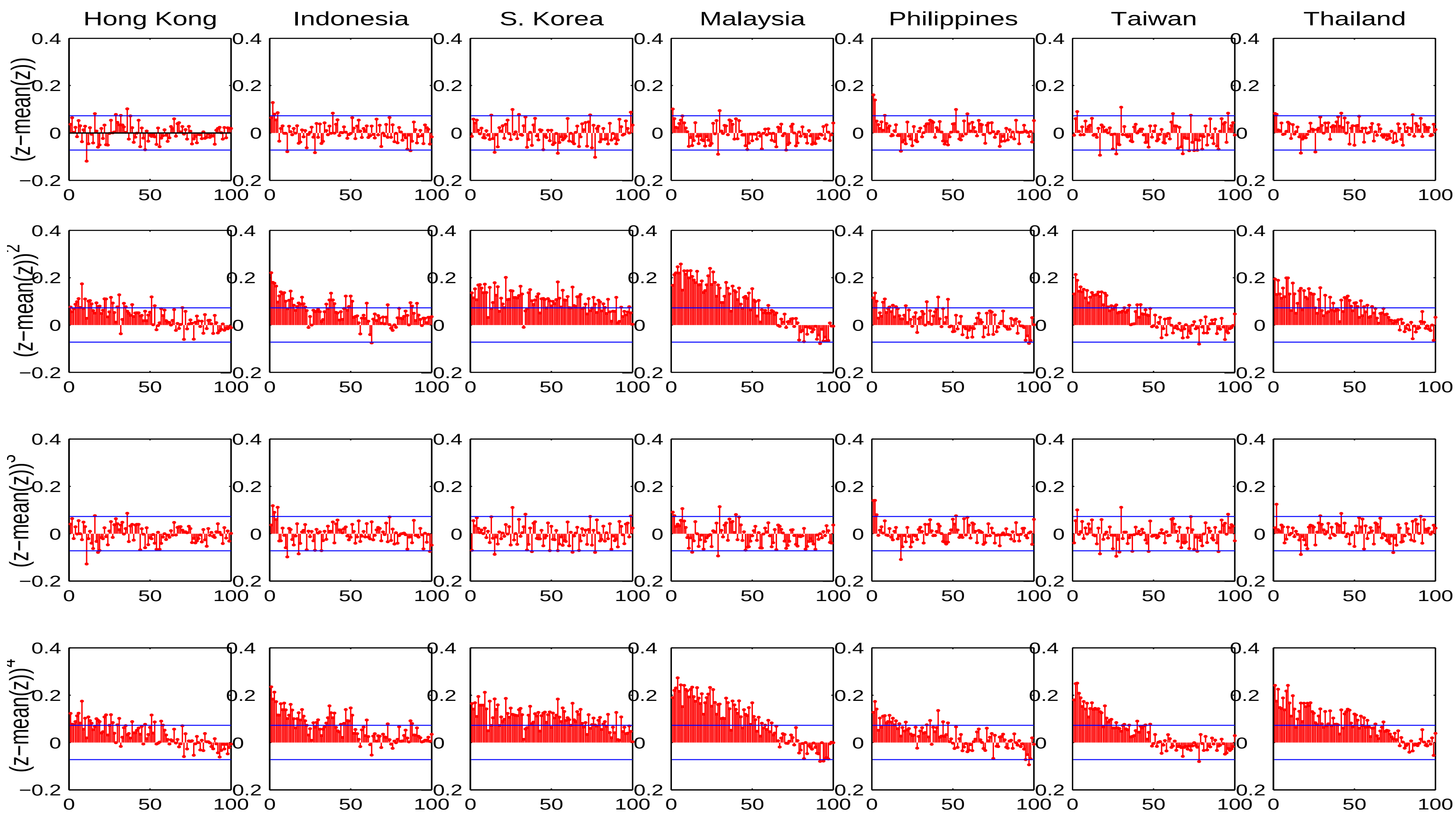
Note: z is the probability integral transform of residuals from each country's model selection. This shows sample autocorrelations of $(z - \bar{z})$, $(z - \bar{z})^2$, $(z - \bar{z})^3$, and $(z - \bar{z})^4$.

Figure I.7 Estimate of the Correlograms of Powers of z (GARCH Normal)



Note: z is the probability integral transform of residuals from each country's model selection. This shows sample autocorrelations of $(z-\bar{z})$, $(z-\bar{z})^2$, $(z-\bar{z})^3$, and $(z-\bar{z})^4$.

Figure I.8 Estimate of the Correlograms of Powers of z (GARCH t)



Note: z is the probability integral transform of residuals from each country's model selection. This shows sample autocorrelations of $(z-\bar{z})$, $(z-\bar{z})^2$, $(z-\bar{z})^3$, and $(z-\bar{z})^4$.

Figure I.9 Estimate of the Correlograms of Powers of z (Empirical CDF) where $\hat{g}(z_1, \cdots, z_d)$ is the kernel estimator from $\{\hat{\mathbf{Z}}_{\mathbf{t}} \equiv (\hat{Z}_{1,t}, \cdots, \hat{Z}_{d,t})\}_{t=1}^{n}$ as follows:

$$\hat{g}(z_1, \cdots, z_d) = \frac{1}{nh^d} \sum_{t=1}^{n} \left[\prod_{j=1}^{d} K_h(z_j, \hat{Z}_{j,t}) \right] \tag{I.47}$$

where K_h is a univariate boundary kernel used in Hong and Li (2002)[34]. Under the some appropriate conditions, the asymptotic null distribution of $\hat{I}_n$ is

$$\frac{nh^{d/2}\hat{I}_n - c_{dn}}{\sigma_d} \longrightarrow N(0,1) \text{ in distribution under } H_0$$

where, $c_{dn} = h^{\frac{d}{2}} \left[(\frac{1}{h} - 2) \int\limits_{-1}^{1} k^2(w)dw + 2 \int\limits_{0}^{1} \int\limits_{-1}^{z} k_z^2(y)dydz \right]^d$, $\sigma_d^2 = 2 \left\{ \int\limits_{-1}^{1} \left[\int\limits_{-1}^{1} k(u+v)k(v)dv \right]^{2d} du \right\}^2$, and, where, $k_z(y) = k(y) / \int_{-1}^{z} k(u)du$.

Table I.7 shows the test results. I examine the goodness-of-fit of some different copulas for the Asian Stock Market indices. Each panel in Table I.7 represents the test results of different marginal distributions. The first and second panels use the data filtered by an AR(p)-GARCH(1,1) normal and an AR(p)-GARCH(1,1)-t process, respectively. The last one is for the unconditional distribution of asset returns. That is, I transform the returns by their empirical distribution functions[35]. Since the models used in this paper are bivariate models, I use the first consistent test[36]. This table shows that it is difficult to reject the null hypothesis in any case. That is, we may say that any copula model tested here may be appropriate or 'right' one with some appropriate dependence parameter, $\hat{\alpha}_0$. The first case results, Normal copula, may be consistent with Chen et al. (2004). In their paper the proportion of rejections of the bivariate Normal copula with GARCH filtering is 0.00 for a randomly selected collection of 100 pairs of equities. One surprising result is that we cannot reject the null with other copula models, either. This result may suggest that using various copula models for asset returns may be appropriate in many cases[37].

[34] To see the functional form, see Hong and Li (2002) and Chen, Fan and Patton (2004)

[35] In this test, we assume that the data are *i.i.d* through time, so GARCH standardized residuals may be appropriate without more assumptions.

[36] When the dimension is large, this test may suffer from the "curse of dimensionality". In such case, we can use the second test proposed in Chen et al. (2004)

[37] This result is only for the constant parameter specification. However, even with the time-varying parameter specification, we cannot reject the null in almost all cases. The results is available upon request.

Table I.7 The Goodness-of-Fit Test with Constant Dependence Assumption

	HK/THA	IND/THA	KOR/THA	MAL/THA	PHI/THA	TWN/THA
Panel A: GARCH-Normal filter						
Normal	-0.2655	-0.6774	-0.7754	-1.1620	-0.7100	-0.5151
	(0.7906)	(0.4982)	(0.4381)	(0.2452)	(0.4777)	(0.6065)
Frank	-0.0922	-0.5132	-0.6264	-1.1566	-0.3333	-0.2588
	(0.9265)	(0.6078)	(0.5311)	(0.2474)	(0.7389)	(0.7958)
Gumbel	-0.1256	-0.4140	-0.6376	-0.6265	-0.2284	-0.1586
	(0.9001)	(0.6789)	(0.5238)	(0.5310)	(0.8193)	(0.8740)
Clayton	0.2825	0.1083	-0.1097	-0.5581	0.1839	-0.4660
	(0.7776)	(0.9138)	(0.9127)	(0.5768)	(0.8541)	(0.6412)
J-C	-0.6146	-0.4593	-0.9041	-1.0311	-0.6170	-0.6875
	(0.5388)	(0.6460)	(0.3660)	(0.3025)	(0.5373)	(0.4917)
Panel B: GARCH-t filter						
Normal	-0.2090	-0.5601	-0.8627	-1.1872	-0.5712	-0.6351
	(0.8344)	(0.5754)	(0.3883)	(0.2351)	(0.5679)	(0.5254)
Frank	-0.0503	-0.4392	-0.7293	-1.1864	-0.2253	-0.3188
	(0.9599)	(0.6605)	(0.4658)	(0.2355)	(0.8218)	(0.7499)
Gumbel	-0.0468	-0.2405	-0.8193	-0.6348	-0.0633	-0.2845
	(0.9626)	(0.8099)	(0.4126)	(0.5256)	(0.9495)	(0.7761)
Clayton	0.2796	0.2039	-0.1929	-0.6293	0.2744	-0.5112
	(0.7798)	(0.4192)	(0.8470)	(0.5292)	(0.7838)	(0.6092)
J-C	-0.6255	-0.3194	-1.0461	-1.1055	-0.4862	-0.8307
	(0.5316)	(0.7494)	(0.2955)	(0.2689)	(0.6268)	(0.4061)
Panel C: Nonparametric						
Normal	0.3946	-0.1995	-0.4500	0.0417	0.3576	-0.4441
	(0.6931)	(0.8419)	(0.6527)	(0.9667)	(0.7206)	(0.6569)
Frank	0.5906	0.1981	-0.0119	0.4763	0.8633	-0.3623
	(0.5548)	(0.8430)	(0.9905)	(0.6339)	(0.3880)	(0.7172)
Gumbel	0.6689	-0.1607	-0.7612	0.5575	0.6884	-0.3552
	(0.5036)	(0.8723)	(0.4466)	(0.5772)	(0.4912)	(0.7225)
Clayton	0.3454	0.7889	0.4268	0.1395	1.1329	-0.0987
	(0.7298)	(0.4302)	(0.6695)	(0.8891)	(0.2573)	(0.9214)
J-C	-0.3866	-0.4923	-1.0626	-0.7329	-0.0383	-0.6047
	(0.6991)	(0.6225)	(0.2880)	(0.4636)	(0.9694)	(0.5454)

Note: Under the null hypothesis, the test statistics follow $N(0,1)$. Parentheses are p-value.

I.4.C.b Dynamic Conditional Correlation Multivariate GARCH

First, I use the DCC-MVGARCH model as a benchmark model. Some recent literature report that conditional correlations between equity returns is not constant. The DCC-MVGARCH model is one of appropriate models to capture varying conditional correlation; for a comprehensive survey of multivariate GARCH modeling, see Bauwens et al. (2003). Here I do not report the parameter estimates using this model, since these parameter estimates are not main concerns. I only present the graph which shows that correlation between the markets estimated by DCC-MVGARCH, more specifically DCC bivariate GARCH model after 1990[38]. Recall that the proposed dynamic correlation structure, Q_t, is equation (I.33).

In order to estimate this model, first we must filter the residuals by the same AR(p) models used in conditional copula models. With these residuals, we can estimate the DCC bivariate model[39]. Figure I.10 shows how the conditional linear dependence changed during the 1990s. The horizontal line in each graph is the constant unconditional linear correlation coefficient shown in Table I.4, and the shaded period is from July 1st, 1997 to June 30th, 1999 for 2 years, the same period shown in I.4 and I.5. This figure has some striking features. In fact Figure I.5 shows the similar pattern of each country's return, having a volatile period around 1997 and some times after that. Interestingly, in the first half of 1997, the dependence decreases almost to the lowest level. But for the next two years after that, the linear dependence seems to increase nearly in all cases, and, eventually, the linear dependence become higher than its constant level. In some cases the linear dependence rises from almost negative to more than 0.5. This finding might be an extension of previous literature, such as King and Wadhwani (1990), Pesaran and Pick (2003), etc. In the 1990s, there was a wide agreement of the existence of correlation breakdown phenomena, increasing dependence after a big shock.

[38] Although DCC-MVGARCH model is very flexible to use many variables, I only report the DCC bivariate case. In fact the difference of result between two is not so much in terms of the shape of dynamics, so I just present the DCC bivariate GARCH model result.

[39] This procedure can be done by the matlab program in UCSD-GARCH toolbox by Sheppard, K.

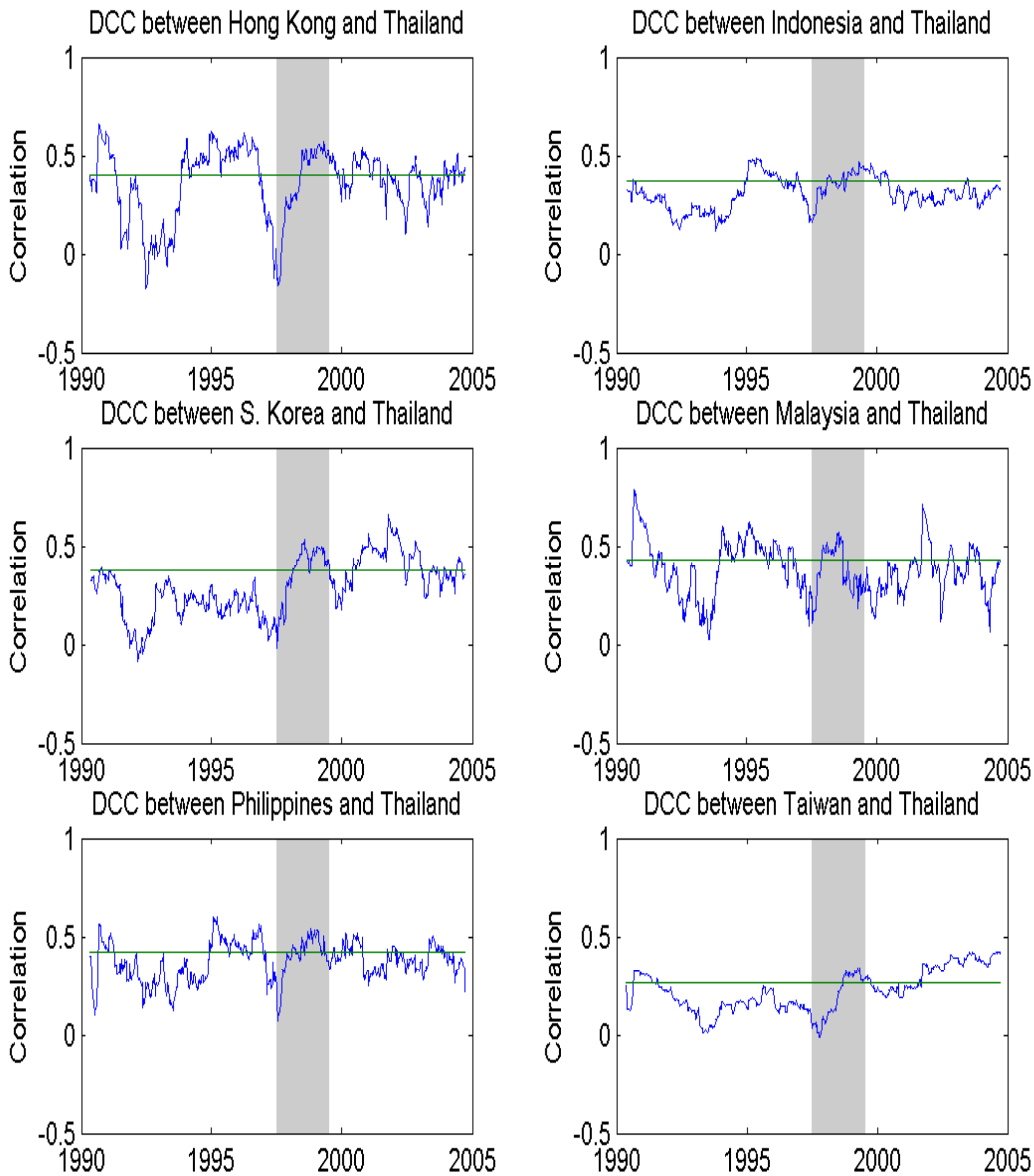
Thus we can say that this paper provides evidence of increasing correlation in the volatile period. However, there is no clear evidence that the highest dependence in the shaded period is the highest in the whole sample period. Therefore, we can say that the volatile return is a sufficient condition for the high linear dependence between the variables, but is not necessary.

Another distinguished mark is that some pairs show increasing dependence even after that period. In the KOR/THA and TWN/THA cases, the dependence is slightly higher than the constant level after 2000. But this fact does not apply to every pair. Generally speaking, we cannot see strong evidence of higher correlation in the recent period(after 2000). The other pairs just show that some higher correlations exist, but those are not huge in the entire sample period.

Another point regarding use of the DCC model should be made before moving to dynamic copula estimation. The persistence of these dynamics can be seen through the parameters in equation (I.33) with similar explanation of just the univariate GARCH model. I do not report the parameter estimates here, but the persistent parameters in DCC model, the sum of the innovation term and the lagged correlation matrices in the DCC estimator, is close to 1. In particular, it ranges from the lowest 0.953 in PHI/THA to the highest 0.995 in TWN/THA, implying a highly persistent processes.

I.4.C.c Time Varying parameter Model with GARCH-t margins

In this section I present the estimation result of the dependence structure using various copula models. Section I.4.C.a justify use of various copula for the Asian Market Stock returns. Recall that we can interpret the different copula parameters differently. The Gaussian copula parameter is very close to the linear correlation. For nondegenerate Gaussian margins with finite variance, this parameter is just the same as the usual linear correlation. The Archimedean copula family has a great variety of different dependence structures. The Frank copula parameter is another dependence measure without tail dependence. Gumbel and Clayton



Note: This graph shows that correlation between the markets estimated by DCC bivariate GARCH model since 1990. The horizontal line in each graph is the constant linear correlation coefficient estimated in the sample period. The shaded period is from July 1st, 1997 to June 30th, 1999 for 2 years.

Figure I.10 DCC-Bivariate GARCH Model

copula parameters represent the upper and lower tail dependence structure, respectively. Also, we cannot compare the degree of dependence just by the parameter magnitude since the parameter spaces are different, although we can convert the parameter value into the Kendall's τ or Spearman's ρ and compare the converted value using Table I.5. Therefore, investigating different time-varying copula parameters would be meaningful work comparing to dynamic linear correlation. First I will present the result with Gaussian and Frank copula model and then I will present the Gumbel, Clayton and Joe copula results. I found that the GARCH-t specification was slightly better than the GARCH normal specification. Therefore I present the GARCH-t specification results as a parametric model. In addition I also present the semi-parametric model results, inspired by section I.2.C as a robust example in the next section. Before showing the results, I briefly summarize the likelihood value of each estimation with a constant dependence parameter in Table I.8. As expected, GARCH-t specification has better likelihood value than GARCH Normal specification with every case. However the log likelihood values with semi-parametric estimation are the best in almost all cases.

Gaussian and Frank result The Gaussian case can be interpreted as close to the conventional dependence measure[40]. The Frank case also can be thought of as abnormal dependence measure without tail dependence. Table I.9 shows the estimation result for the Normal copula and Table I.10 for the Frank copula. Panel A contains results under the constant assumption and panel B with the time-varying assumption shown in equation (I.44). There are a couple of things to notice in general. First, the closest relation is shown in THA/MAL, then THA/PHI, THA/HK, THA/IND, THA/KOR, and THA/TWN in descending order according to panel A of both tables. This is not the same order of simple correlation coefficient shown in Table I.4. THA/KOR has higher correlation than THA/IND in Table I.4. This could be evidence of the failure of linear correlation. However, the

[40] t-copula also has this property. But estimating with t-copula takes more time comparing to Gaussian, since this procedure include a lot of inverse calculation.

Table I.8 Log Likelihood Value of Estimations with Constant Dependence

	Hong Kong	Indonesia	S. Korea	Malaysia	Philippines	Taiwan
Normal						
GARCH-t	-2806.3	-2896.2	-2989.2	-2651.9	-2784.3	-2998.4
GARCH Normal	-2830.6	-2927.7	-3009.7	-2687.8	-2820.5	-3027.6
Semi Parametric	-2648.3	-2770.4	-2654.5	-2746.5	-2641.1	-2805.5
Frank						
GARCH-t	-2811.7	-2897.6	-2993.4	-2657.5	-2789.1	-3001.8
GARCH Normal	-2835.3	-2927.3	-3014.5	-2691.9	-2821.9	-3031.2
Semi Parametric	-2656.5	-2777.5	-2665.2	-2754.7	-2650.7	-2807.8
Gumbel						
GARCH-t	-2811.4	-2904.1	-2991.4	-2665.3	-2793.8	-3003.0
GARCH Normal	-2840.0	-2940.0	-3016.4	-2706.8	-2835.8	-3037.1
Semi Parametric	-2655.9	-2765.3	-2652.2	-2753.2	-2644.2	-2806.7
Clayton						
GARCH-t	-2811.3	-2902.1	-2995.0	-2656	-2787.4	-2997.6
GARCH Normal	-2844.8	-2937.8	-3019.8	-2699.9	-2831.8	-3028.0
Semi Parametric	-2647.4	-2781.6	-2665.1	-2743	-2646.8	-2805.7
Joe-Clayton						
GARCH-t	-2802.9	-2898.4	-2986.4	-2651.7	-2782.7	-2995.9
GARCH Normal	-2835.9	-2935.0	-3011.7	-2696.1	-2828.4	-3027.5
Semi Parametric	-2638.9	-2760.8	-2645.6	-2732.9	-2632.2	-2800.4

Note: Each column represent the log likelihood value with different models where investigate the dependence parameter with Thailand. For example, Hong Kong means the log likelihood value of model using Hong Kong and Thailand.

Spearman's ρ coefficient is the same order as a constant case in panel A. Second, the Likelihood Ratio Test(LRT) statistics show that we can reject the null in all cases at the 10% level . The null hypothesis is $H_0 : \eta_1 = \omega_1 = \lambda_1 = \omega_2 = \lambda_2 = 0$. This infers that the additional parameters are meaningful given our power to detect such differences, and implies that the dynamic specification seems preferable. In Panel B, there are two things to notice. The first thing to notice is the significance of parameters, the λ_1 and λ_2. These parameters represent the asymmetric shock effect on the dependence structure. However, this effect is not significant except a few cases and there is no consistency in their sign. Therefore we cannot say that there is clear evidence of asymmetric dependence. The last thing to notice is

Table I.9 Estimation Results for Normal Copula with GARCH-t Margins

Normal	Hong Kong	Indonesia	S. Korea	Malaysia	Philippines	Taiwan
OBS.	753	748	751	748	752	751
Panel A: Constant Coefficient						
δ	0.3574	0.3034	0.2938	0.3824	0.3717	0.2408
	(0.0299)	(0.0314)	(0.0314)	(0.0291)	(0.0296)	(0.0339)
Likeli	-2806.3	-2896.2	-2989.2	-2651.9	-2784.3	-2998.4
Panel B: Time Varying Coefficient						
η_0	-0.0302	-0.0163	-0.0430	-0.0031	0.0735	0.0392
	(0.0085)	(0.0102)	(0.0101)	(0.0105)	(0.0344)	(0.0489)
η_1	0.9905	1.0000	1.0000	0.9919	0.9430	0.8958
	(0.0018)	(0.0220)	(0.0066)	(0.0053)	(0.0250)	(0.0689)
ω_1	-0.0721	0.0155	0.0406	0.1288	-0.1530	-0.1055
	(0.0546)	(0.0825)	(0.0889)	(0.0396)	(0.0868)	(0.1293)
λ_1	0.0203	-0.0079	-0.1940	-0.2700	0.1862	-0.0496
	(0.0803)	(0.0624)	(0.1394)	(0.0572)	(0.1436)	(0.2226)
ω_2	0.1865	0.0888	0.0836	-0.1103	-0.0879	-0.0417
	(0.0500)	(0.0953)	(0.0672)	(0.0585)	(0.0932)	(0.1312)
λ_2	-0.2865	-0.1249	-0.1458	0.2219	0.2242	0.1676
	(0.0860)	(0.0803)	(0.1260)	(0.0952)	(0.1583)	(0.2564)
Likeli	-2791.7	-2891.4	-2971.8	-2640.4	-2778.9	-2993.1
LRT	29.2**	9.6*	34.8**	23.0**	10.8*	10.6*

Note: The panel A is the model with constant assumption and the panel B is with time-varying assumption. The numbers in parentheses are the standard errors.
LRT statistic follows $\chi^2(5)$ and ** represents the rejection of the null at 5% significant level and *, 10% significant level, respectively.

the persistence parameter. The persistence parameter, η_1 is very high[41]. In some cases it is even 1, implying a unit root process. THA/PHI and THA/TWN have relatively low persistent parameters, though still bigger than 0.88. This implies that the shock effect on the dependence structure lasts for a long time in most cases.

The meaning of such a high persistence is clear in Figure I.11. Figure I.11 shows the time path of the dependence parameters of both the Normal copula and the Frank copula. The solid line represents the normal copula path scaled on the left side axis and the dashed line represents the Frank copula path scaled on

[41] η_1 is the persistence parameter, so the parameter space of η_1 is limited between -1 to 1.

Table I.10 Estimation Results for Frank Copula with GARCH-t Margins

Normal	Hong Kong	Indonesia	S. Korea	Malaysia	Philippines	Taiwan
OBS.	753	748	751	748	752	751
Panel A: Constant Coefficient						
δ	2.2008	1.9133	1.7626	2.3672	2.3006	1.3527
	(0.2293)	(0.2259)	(0.2249)	(0.2293)	(0.2293)	(0.2251)
Likeli	-2811.7	-2897.6	-2993.4	-2657.5	-2789.1	-3001.8
Panel B: Time Varying Coefficient						
η_0	-0.2085	-0.1106	-0.2606	-0.0801	0.5229	-0.0569
	(0.0589)	(0.1213)	(0.0830)	(0.0641)	(0.2197)	(0.1661)
η_1	1.0000	0.9910	1.0000	0.9963	0.9526	0.9587
	(0.0039)	(0.0144)	(0.0052)	(0.0056)	(0.0173)	(0.0007)
ω_1	-0.4720	0.3038	0.1109	0.9438	-1.6370	-0.1372
	(0.4035)	(0.3094)	(0.5218)	(0.3112)	(0.6683)	(0.6308)
λ_1	0.3992	-0.3818	-0.8596	-2.1111	2.2039	-0.8968
	(0.5669)	(0.7001)	(0.8533)	(0.5853)	(1.1809)	(1.0214)
ω_2	1.4531	0.3514	0.6494	-0.7079	-0.4537	0.0215
	(0.3459)	(0.4814)	(0.4628)	(0.4087)	(0.6679)	(0.5041)
λ_2	-2.0785	-0.6497	-1.1968	1.4151	1.0525	0.0012
	(0.5999)	(0.7046)	(0.7983)	(0.6415)	(1.1664)	(0.9076)
Likeli	-2798.7	-2894.2	-2980.5	-2648.9	-2782	-2996.6
LRT	26.0**	6.8	25.8**	17.2**	14.2**	10.4*

Note: The panel A is the model with constant assumption and the panel B is with time-varying assumption. The numbers in parentheses are the standard errors.
LRT statistic follows $\chi^2(5)$ and ** represents the rejection of the null at 5% significant level and *, 10% significant level, respectively.

the right side axis. The horizontal line is the constant dependence parameter of both copulas in Panel A of I.9 and I.10. It seems there is more fluctuation in the time-varying parameter of Frank copula, but it is not because of difference in measurement, but because of scaling problems. For instance, at the end of the sample period, in the HK/THA case, parameter of Frank copula almost reaches 5. This implies the Kendall's τ is about 0.45, and this also is about the same when the normal copula parameter is 0.64. For more comparisons, see table I.5. Therefore we can say that the implied dependence transformed to the same measure such as the Kendall's τ or Spearman's ρ is similar in both cases. In Figure I.11, the higher persistence parameter sustains the shock effect longer than the relatively lower persistence parameter case. For example, PHI/THA and TWN/THA have comparatively lower persistence and the time-varying path of dependence parameter is just a small fluctuation along with the constant path. We can clearly see that there are so many differences in the time path pattern within Figure I.11 with different pairs. Remember that from Figure I.4 and Figure I.5 we can see the similarity of the indices and the return paths-similar up and down phases and similar volatile periods. However, we also can see that similarity does not guarantee the similar pattern in the dependence path here.

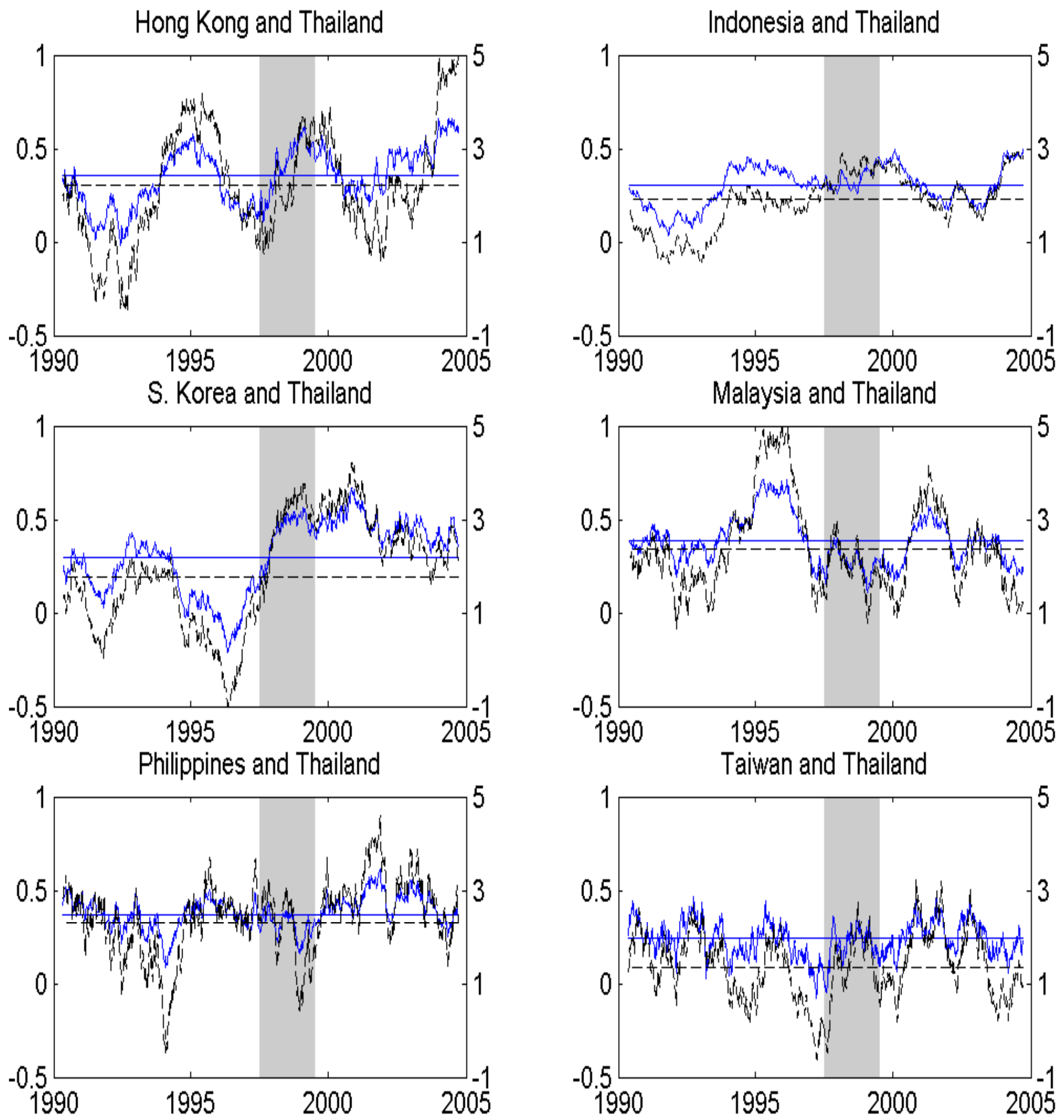
In comparing Figure I.10 to Figure I.11, we immediately see there exist some differences, sometimes substantially, in two different measurements of dependence, while there still exist a lot of similarities. In the first case, HK/THA, the up and down pattern seems to be similar. But recent period dependence is much different. Using a copula parameter shows the stronger relationship of this pair more than using the DCC-MVGARCH model, recently. When the copula model is used, the second case, IND/THA, also shows higher dependence recently and the third case, KOR/THA, shows much lower dependence around 1995. In addition, the fourth case, MAL/THA, seems to be more persistent than when DCC-MVGARCH is used. Needless to say about the last two cases, we can see many different patterns in both models.

Focusing on the shaded period, the difference looks bigger. Unlike Figure I.10, there is no clear and consistent evidence of increasing dependence during the shaded period, although some pairs show the increasing dependence. KOR/THA is in an increasing phase, but the increase began more than one year ago. MAL/THA and PHI/THA even show decreasing dependence for 2 years. In other words, a volatile period does not lead to higher dependence among samples automatically. This is counter evidence of higher volatility results in the higher dependence in King and Wadhwani (1990), Baig and Goldfajn (1999) and so on. As mentioned, these papers used the correlation coefficient as a dependence measure and found a significant increase in cross-market relationship after a shock to one or a group of countries, the so-called ‘Correlation Breakdown.’[42] The previous section of this paper also supports this phenomenon. But this phenomenon is not clear and consistent in every pair when we use a different measure here. Although HK/THA, IND/THA and KOR/THA show increasing relationship after mid-1997, this explanation is not suitable to all cases. Therefore the conjecture that the high volatility results in the higher relationship should be reconsidered. In addition, we cannot see any relation between the bullish or bearish market phase in Figure I.4 and the up and down momentum of the dependence shown in Figure I.11. This could be because of the failure of linear dependence measure, so it should be used with caution. Dependence is one of the important concepts in the academic area as well as in the applied area. However, only the linear dependence measurement is mainly used in both areas. But this empirical finding would justify using this new measure for existing conclusions made by the linear measure.

Gumbel, Clayton and Joe-Clayton result(tail dependence) In this part, I present a different parameter estimation. As mentioned, some Archimedean copulas, such as Gumbel, Clayton and so on, are characterized by upper or lower tail dependence. These tail dependences can be thought of as the probability of events like joint low(high) extreme event happening, given that one has an low(high) extreme event. Therefore these measures could have a special meaning in the risk management area and if these are not similar to usually used dependence measures, then it is also worthwhile to re-examine the conclusions made by using the conventional measure. Recall that Joe-Clayton copula has two parameters and each of them can match to lower tail(θ) and upper tail(δ) dependence. Since empirical results show that estimating Gumbel and Clayton separately and Joe-Clayton alone do not have many differences in the pattern of time path of dependence parameters, I only introduce the Joe-Clayton copula's result here[43].

[42] This in a huge area in economic literature. For more literature review, see Claessens et al. (2001), and for some counter example, see Boyer et al. (1999) and Forbes and Rigobon (2002).

[43] When I estimate the Gumbel and the Clayton separately, the constant parameter estimates are bigger than estimating Joe-Clayton copula solely in every case. For example, the estimate of Gumbel parameter is 1.2826 and that of the upper tail parameter in Joe-Clayton is 1.1996 in THA/HK case. I think it is because of the interaction



Note: This shows the time path of the dependence parameters of both Normal copula and Frank copula. The solid line represents the normal copula path scaled on the left side axis and the dashed line represents the Frank copula path scaled on the right side axis. The horizontal line is the constant dependence parameter in both copulas.

Figure I.11 Time Path of Dependent Parameters of Normal and Frank Copulas with GARCH-t Margins

Table I.11 is the estimation result using the Joe-Clayton copula. Here the first thing to notice is that the higher dependence parameter in normal or Frank copula does not guarantee a higher tail dependence. While the closest relation is shown in THA/MAL in panel A of Table I.9 and Table I.10, the highest estimate of upper tail dependence is THA/HK case in panel A of Table I.11. The descending order of upper tail dependence also is a little bit different, although the lower tail dependence order is not so different. The order is THA/HK, THA/KOR, THA/PHI, THA/MAL, THA/IND and THA/TWN in descending manner. This result may suggest that usual dependence and tail dependence could have special characteristics, and one does not infer any implication of the other.

Next, the Likelihood Ratio Test(LRT) statistics show that we can reject the null in almost all cases at the 5% level. The null hypothesis is $H_0 : \eta_{\delta 1} = \omega_{\delta 1} = \lambda_{\delta 1} = \omega_{\delta 2} = \lambda_{\delta 2} = \eta_{\theta 1} = \omega_{\theta 1} = \lambda_{\theta 1} = \omega_{\theta 2} = \lambda_{\theta 2} = 0$. This also implies that the additional parameters to capture the dynamics are meaningful and implies that the dynamic specification seems preferable in most case. However, the last one, THA/TWN, shows that we cannot reject the null even at the 10% level. This implication will be clear with the next graph, so I will talk about this case later. In Panel B, we can notice two things again. The parameter's significance does not confirm the asymmetric shock effect on the dependence structure. Even though some of $\lambda_{\delta 1}$, $\lambda_{\delta 2}$, $\lambda_{\theta 1}$, and $\lambda_{\theta 2}$ are significant, there is no consistency in their sign. Another evidence of unclear asymmetry effect is in Panel A. When I convert the each constant estimates in Panel A into the implied tail dependence($\lambda_U = 2 - 2^{1/\theta} \; and \; \lambda_L = 2^{-1/\delta}$), again we cannot see evidence of asymmetry. In some cases, the implied upper tail dependence is greater than the implied lower tail dependence, and in some other cases the upper one is smaller than the lower one, so there is no consistency.

between two parameters in Joe-Clayton. Recall that when $\theta = 1$, Clayton copula is obtained, and Joe copula is obtained as $\delta \to 0$. Therefore There could exists the interaction between them when $\theta \neq 1$ or $\delta \nrightarrow 0$. Gumbel and Clayton results are available upon the request. And estimating Joe-Clayton copula is extremely hard and often does not converge well, since the number of parameters is 12. I use Gumbel and Clayton estimation results as the initial values for MLE routine. Even though estimating Gumbel and Clayton is not an easy task, this is much easier than just estimating Joe-Clayton's 12 parameters together in one time.

Another thing to notice is that persistent parameter is also very high, sometimes even 1 in this model again. Note that THA/PHI and THA/TWN again have relatively low persistent parameters, $\eta_{\delta 1}$ and $\eta_{\theta 1}$.

Although the persistent parameters seem to be high and similar, the time path of each parameter does not look similar. Figure I.12 represents the dynamics of each lower and upper tail parameter. The upper line represents the upper tail parameter(δ_t) path and the lower line represents the lower tail parameter(θ_t) path. The horizontal line is the constant dependence parameter of Panel A. First of all, these graphs do not seem to be similar to Figure I.11, although there still exists a similar pattern in some sample period. I think that this is because of the characteristics of tail dependence. As mentioned, this is further evidence that the usual dependence and tail dependence could have their own special characteristics and one cannot infer any implication of the other. Secondly, there is no clear evidence that the volatile period has higher tail dependence at all in general. Again for the shaded period, we cannot see the increasing pattern in general. In many cases, the upper or lower dependence is even lower than the other periods during this period. Another thing to be mentioned is that the parameter space are restricted like $0 < \theta < \infty$, $1 \leq \theta < \infty$. Therefore the time paths of some parameters are bounded by 0 or 1[44]. The bounded parameter period simply represents that there is no lower or upper tail dependence at all, saying that the probability of lower or upper extreme event happening is nothing.

Now I take a look at some individual cases closely. First, in the HK/THA case, two parameters seem to move together, and this pattern is also similar with normal and Frank parameters. But between about 1998 and 2002, the upper one is lower than its constant level, but the lower one is higher than its constant level. The IND/THA case also shows that there is no co-movement of both parameters as a whole. However, KOR/THA and MAL/THA seem to have similar movement patterns of the two parameters in the sample period. The PHI/THA and

[44]Although the parameter space of normal copula is also restricted, there is no case where the parameter is bounded. This make sense, since that is a different dependence measure.

Table I.11 Estimation Results for Joe-Clayton Copula with GARCH-t Margins

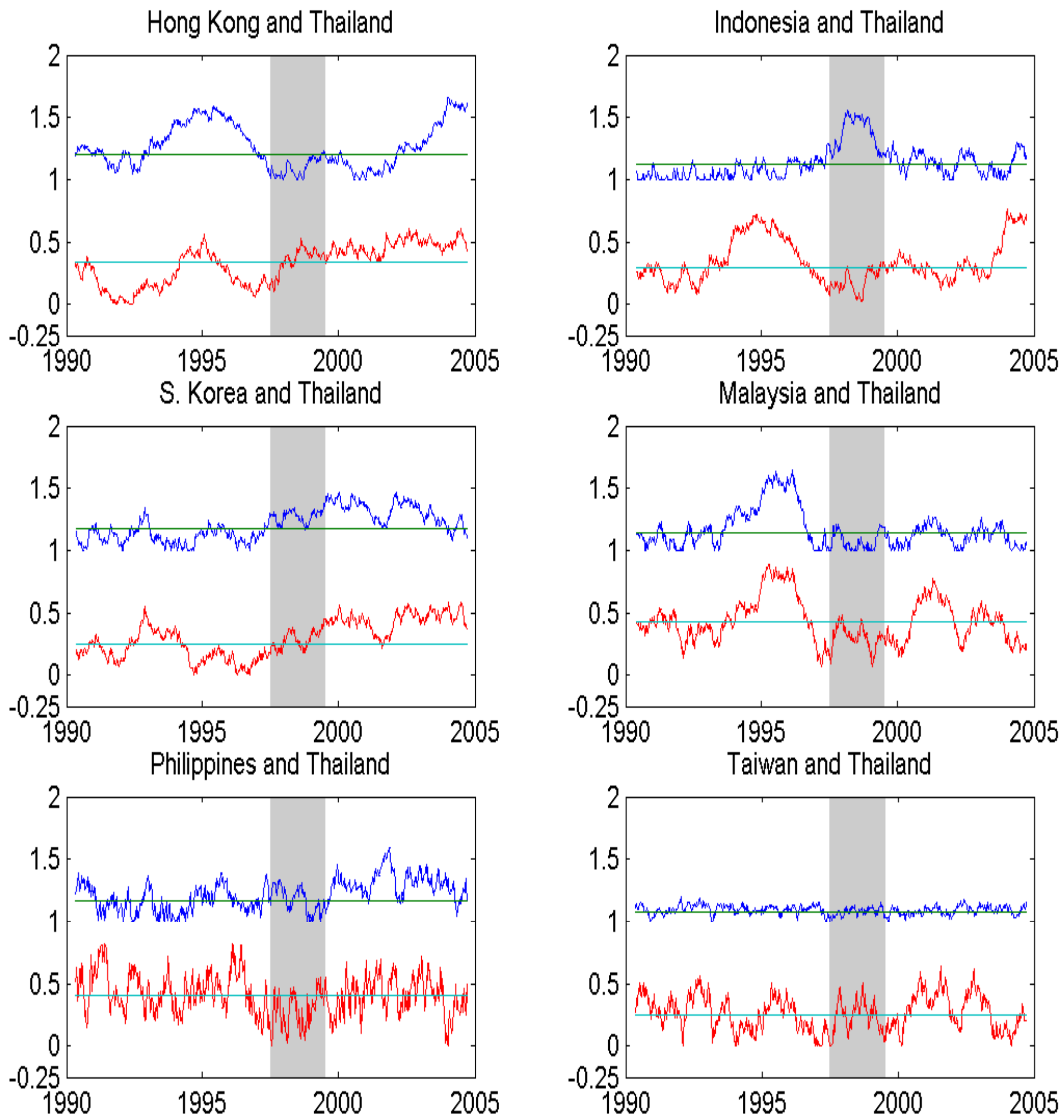
Normal	Hong Kong	Indonesia	S. Korea	Malaysia	Philippines	Taiwan
OBS.	753	748	751	748	752	751
Panel A: Constant Coefficient						
δ	1.1996	1.1205	1.1746	1.1407	1.1678	1.0817
	(0.0582)	(0.0576)	(0.0500)	(0.0545)	(0.0570)	(0.0481)
θ	0.3384	0.2929	0.2425	0.4265	0.4029	0.2502
	(0.0588)	(0.0596)	(0.0551)	(0.0643)	(0.0662)	(0.0561)
Likeli	-2802.9	-2898.4	-2986.4	-2651.7	-2782.7	-2995.9
Panel B: Time Varying Coefficient						
$\eta_{\delta 0}$	-0.0060	-0.0510	-0.0084	-0.0064	0.1015	0.1663
	(0.0177)	(0.0220)	(0.0002)	(0.0088)	(0.0054)	(0.1716)
$\eta_{\delta 1}$	0.9973	1.0000	0.9859	0.9997	0.9582	0.8719
	(0.0059)	(0.0050)	(0.0013)	(0.0088)	(0.0045)	(0.1532)
$\omega_{\delta 1}$	-0.0395	0.2022	0.1904	0.1670	-0.3158	-0.1512
	(0.0465)	(0.0629)	(0.0041)	(0.0309)	(0.0130)	(0.0716)
$\lambda_{\delta 1}$	0.0190	-0.3128	-0.3189	-0.3559	0.3755	0.2275
	(0.0761)	(0.0984)	(0.0098)	(0.0636)	(0.0189)	(0.1097)
$\omega_{\delta 2}$	0.1161	-0.0235	-0.0760	-0.1177	0.0313	0.0297
	(0.0676)	(0.0542)	(0.0026)	(0.0274)	(0.0103)	(0.0353)
$\lambda_{\delta 2}$	-0.0971	-0.0662	0.1231	0.3202	0.0368	-0.0123
	(0.0566)	(0.0981)	(0.0029)	(0.0675)	(0.0019)	(0.0660)
$\eta_{\theta 0}$	-0.0338	-0.0318	-0.0346	-0.0071	0.1639	0.0376
	(0.0416)	(0.0233)	(0.0074)	(0.0198)	(0.0007)	(0.0406)
$\eta_{\theta 1}$	0.9881	1.0000	0.9914	0.9901	0.8922	0.9214
	(0.0209)	(0.0085)	(0.0008)	(0.0059)	(0.0094)	(0.0511)
$\omega_{\theta 1}$	0.0357	-0.0035	0.1559	0.1753	0.1463	-0.1183
	(0.1138)	(0.0352)	(0.0015)	(0.0893)	(0.0051)	(0.1374)
$\lambda_{\theta 1}$	-0.1878	-0.0572	-0.3251	-0.3992	-0.0844	-0.0910
	(0.1762)	(0.0570)	(0.0011)	(0.1663)	(0.0098)	(0.1736)
$\omega_{\theta 2}$	0.0717	0.1754	0.0211	-0.1551	-0.5343	-0.0988
	(0.2078)	(0.0857)	(0.0011)	(0.1208)	(0.0403)	(0.1483)
$\lambda_{\theta 2}$	-0.1085	-0.2013	0.0299	0.3108	1.0299	0.2320
	(0.3920)	(0.1396)	(0.0011)	(0.1901)	(0.0110)	(0.2552)
Likeli	-2786.2	-2888.6	-2973.6	-2639.9	-2773.5	-2989.2
LRT	33.4**	19.6**	25.6**	23.6**	18.4**	13.4

Note: The panel A is the model with constant assumption and the panel B is with time-varying assumption. The numbers in parentheses are the standard errors.
LRT statistic follows $\chi^2(10)$ and ** represents the rejection of the null at 5% significant level and *, 10% significant level, respectively.

TWN/THA cases, where the relatively lower persistent parameter is perceived, show no specific pattern. Especially in the TWN/THA case the upper tail parameter seems to have no variation in the entire sample period. That is why we cannot reject the null of the LRT test. It implies that additional parameters to capture the dynamics are not adequate in TWN/THA case, in particular, the upper tail dependence. As a summary, we cannot find the similarity and pattern of the time path, not only between the upper and lower tail dependence, but also between different pairs.

Recall again that we can convert the estimated parameter into lower and upper tail dependence. The lower tail dependence parameter is $2^{-1/\delta}$, independent of θ, and the upper tail dependence parameter is $2-2^{1/\theta}$, independent of δ. Figure I.13 shows the converted tail dependence. As mentioned earlier, the converted tail dependence can be thought of as the probability of events like joint low(high) extreme event happens, given that one has an low(high) extreme event. As expected, there is nothing in common among sample pairs in general. Sometimes the time path of probability decreases to zero, implying zero probability of extreme realization together given one extreme event. Sometimes the probability increases to more than 0.4. Zero probability of tail dependence is reported in Poon et al. (2002).

Individually, the HK/THA pair shows that the upper tail dependence increases nearly 0.5 around 1995 and in the recent period and the lower tail dependence is moving around 0.25, recently. The IND/THA pair shows that there is clear evidence of increasing upper tail dependence but not in lower tail dependence. The KOR/THA pair has the tendency of increasing in both tail dependence after about 1997, but such an increasing tendency started before 1997. The MAL/THA pair have a quite high lower and upper tail probability around 1995 and even have lower probability than normal for 2 years after 1997, and very low probability recently. In PHI/THA and TWN/THA pairs, both tail dependences are fluctuating in the whole sample period. According to this result there is no clear evidence that Asian



Note: This shows the time path of the dependence parameters of Joe copula. The upper line represents the upper tail parameter(δ) path and the lower line represents the lower tail parameter(θ) path. The horizontal line is the constant dependence parameter of Joe-Clayton copula.

Figure I.12 Time Path of Dependent Parameters of Joe-Clayton copula with GARCH-t Margins financial crisis leads to the higher probability of extreme event together. Therefore we have to be careful to conclude that dependence structures are changed during periods of financial turmoil and that increased tail dependence and asymmetry in times of high volatility characterize Asian countries[45]. In terms of asset allocation, the lower tail dependence can be used to reduce the risk of huge loss in financial investments. Possible strategy is to invest in the pair of countries with high upper tail dependence and low lower tail dependence. I will leave this issue for future study.

I.4.D Robustness

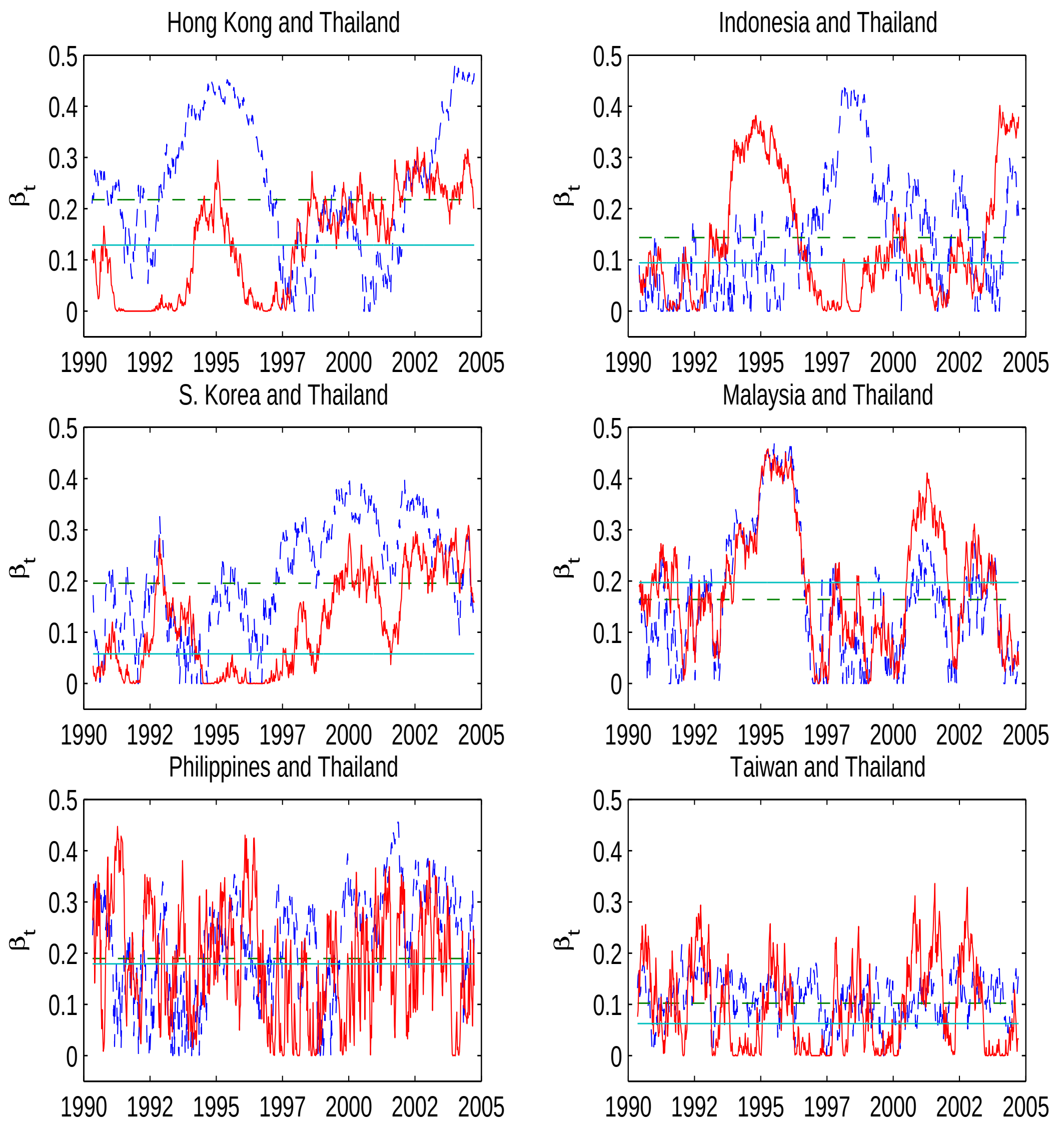
I.4.D.a Nonparametric Two Step Approach

Here I present the results of the semi-parametric method. As shown in section I.2.C, misspecified margins result in (under-estimated) bias. In such situations, semi-parametric estimation is alternative choice. But there is a shortcoming of the semi-parametric method. If the main goal of estimation is ,for example, forecasting, then non-parametric estimation of margins does not give any intuition given information. Therefore I just present the estimation result with non-parametric margins and compare that with the previous results here.

Table I.12 shows the results of constant parameters with different copulas. As we can see, these estimates are greater than the estimates with GARCH-t margins of Table I.9, I.10, and I.11 in most cases. The greatest difference is 0.0667 of THA/KOR case in normal copula, 0.3531 of THA/MAL in Frank copula and the smallest difference is 0.0154 in normal, 0.1696 in Frank of THA/TWN case. Although there exists differences in estimates, they do not seem to be large[46]. And in Joe-Clayton copula the estimates with GARCH-t margins are even greater sometimes. Therefore we can say that the GARCH-t margin assumption is not too bad of a choice. However, since a constant parameter is not the main concern, I

[45] see, Rodriguez (2003)

[46] Table I.2 shows that semi-parametric estimation tends to estimate greater than the true value. Therefore we can say that these differences are not so big.



Note: This shows the time path of implied tail dependence of Joe copula. The dashed line represents the implied upper tail dependence path and the solid line represents the implied lower tail dependence path. The horizontal line is the constant implied upper and lower tail dependence estimated by Joe-Clayton copula.

Figure I.13 Time Path of Implied Tail Dependence of Joe-Clayton Copula with GARCH-t Margins

Table I.12 Estimation Results of Each Copula with Non-Parametric Margins

	Hong Kong	Indonesia	S. Korea	Malaysia	Philippines	Taiwan
OBS.	753	748	751	748	752	751
Normal						
δ	0.3975	0.3577	0.3605	0.4283	0.4091	0.2562
	(0.0293)	(0.0304)	(0.0302)	(0.0278)	(0.0288)	(0.0333)
Likeli	-2648.3	-2770.4	-2654.5	-2746.5	-2641.1	-2805.5
Frank						
δ	2.4605	2.1524	2.0902	2.7203	2.5098	1.5223
	(0.2350)	(0.2314)	(0.2317)	(0.2376)	(0.2339)	(0.2272)
Likeli	-2656.5	-2777.5	-2665.2	-2754.7	-2650.7	-2807.8
Joe-Clayton						
δ	1.1848	1.2751	1.2610	1.2066	1.2515	1.1215
	(0.0502)	(0.0538)	(0.0512)	(0.0527)	(0.0523)	(0.0444)
θ	0.4828	0.2904	0.3037	0.5589	0.4420	0.2630
	(0.0656)	(0.0603)	(0.0599)	(0.0692)	(0.0661)	(0.0634)
Likeli	-2638.9	-2760.8	-2645.6	-2732.9	-2632.2	-2800.4

move on to the time-varying parameter model.

Here I do not report the estimation results in Table format. I only present the time path of dependence(the graphical result) here. The time paths of dependence parameters include all the information about the estimates. Figure I.14 shows the time paths of normal and Frank copulas. Comparing to Figure I.11, we cannot see big differences between them; they are almost identical, though there exist some minor differences, for instance, the recent declining dependence in HK/THA. This is also true with the Gumbel and Clayton copulas, even though I do not present both results here. Therefore we can say that the up and down movement pattern of time-varying parameter is well-captured with the GARCH-t margins and that the GARCH-t margin assumption is not too bad of a choice again. In fact, GARCH-normal margins produce a lot of different results (different parameter estimates and different time paths of dependence parameters), although I do not present that here. As we have seen in section I.4.B, GARCH-normal may not be one of the good choices for each margin. This is evidence that this misspecified margin result in the bad(or wrong) estimation as a whole[47]. These results are

[47] A constant parameter estimates are surprisingly high, for example, the normal copula parameters range from

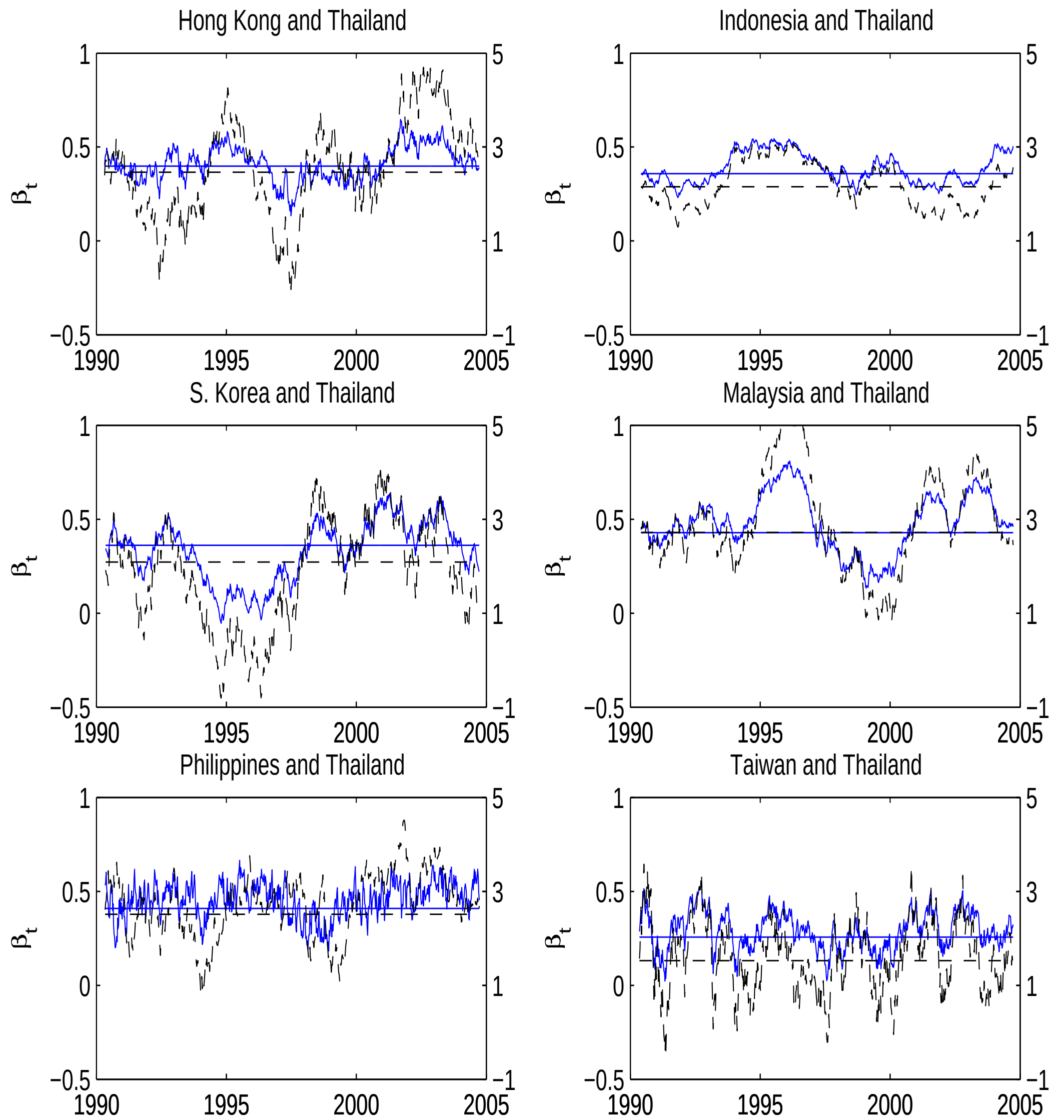
available upon the request.

Finally, Figure I.15 shows the time paths of the dependence parameter in the Joe-Clayton copula. In Figure I.15 we immediately see that it does not seem to be similar to Figure I.12 in some time paths. First, the lower tail of PHI/THA and the upper tail of TWN/THA seem to be more volatile. This might be explained by the low persistence in these time paths. The persistent parameters are only 0.2123 and 0.4995, respectively. Therefore, the previous dependence level does not give any contribution to next period's dependence level, only depending mostly on the forcing variables. The upper tails of IND/THA and MAL/THA seem to be very different. However, this can be explained by the boundary constraint. As we know, the parameter space of δ is restricted between 1 to ∞. In Figure I.12 we can see that these two paths are clearly bounded by 1, since I impose the boundary condition in the estimation procedure. These bounded results could distort the estimation result. Therefore, when we have a lot of bounded results, we should be careful to use that result, or try to estimate another single parameter model to ensure that we have right time paths. Except these examples, we again can see the similar up and down momentum in each time path of dependence parameters, though there is a little bit difference in the magnitude. This implies that GARCH-t assumption with appropriate AR(p) model may not be a bad choice for margins.

I.4.D.b Out of Sample Performance Evaluation

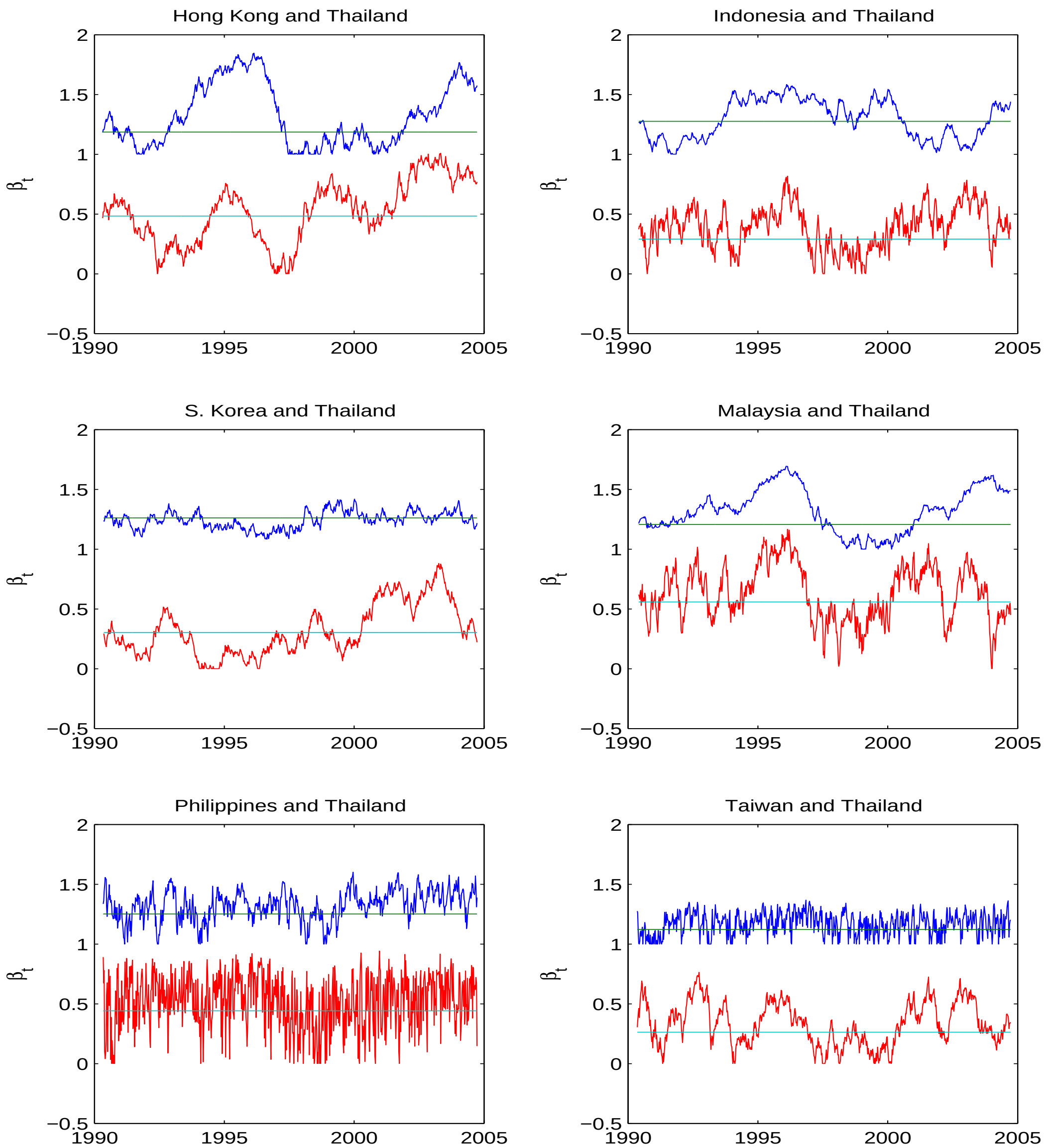
In this section, I present the result of out-of-sample evaluation of the models. A number of recent papers, including Diebold et al. (1998), have gone beyond the traditional evaluation of point forecasts to consider density forecasts (see Christoffersen (1998) for the evaluation of interval forecasts, Clements and Smith (2000) for the multivariate case evaluation, and so on). The basic idea is that when the predicted density of a variable Y_t corresponds to the true predictive density, then $z_t \sim U(0,1)$, $t = 1, \cdots, n$, and the sequence is independently

0.5168 of MAL/THA to 0.6957 of TWN/THA. The time paths of each parameter are not even close to the other two cases.



Note: This shows the time path of the dependence parameters of both Normal copula and Frank copula. The solid line represents the normal copula path scaled in left side axis and the dashed line represents the Frank copula path scaled in right side axis. The horizontal line is the constant dependence parameter in both copulas.

Figure I.14 Time Path of Dependent Parameters of Normal and Frank Copulas with Non-Parametric Margins



Note: This shows the time path of the dependence parameters of Joe copula. The upper line represents the upper tail parameter(δ) path and the lower line represents the lower tail parameter(θ) path. The horizontal line is the constant dependence parameter of Joe-Clayton copula.

Figure I.15 Time Path of Dependent Parameters of Joe-Clayton Copula with Non-Parametric Margins distributed, so $\{Z_t\}_{t+1}^{n}$ is *i.i.d.* $U(0,1)$, where $Z_t = \int_{-\infty}^{y_t} z_t(w)dw$. Therefore we can evaluate the predicted density of forecasting model by assessing uniformity and independence of the probability integral transformation. In the copula model, we also can evaluate the predicted density. Since some copula models have closed form expressions of the joint CDF, it is relatively easy to calculate the probability integral transforms of the realizations of the variables with respect to the predicted densities[48]. With constant dependence parameter assumption[49], first we estimate each model using in-sample data. After that, with the assumption of GARCH-t process for each series, we can calculate the predicted value of $u_{t+1|t}$ and $v_{t+1|t}$ of out-of-sample data at time t easily.

$$u_{t+1} = \int_{-\infty}^{\varepsilon_{t+1|t}} f_t(\omega)d\omega$$

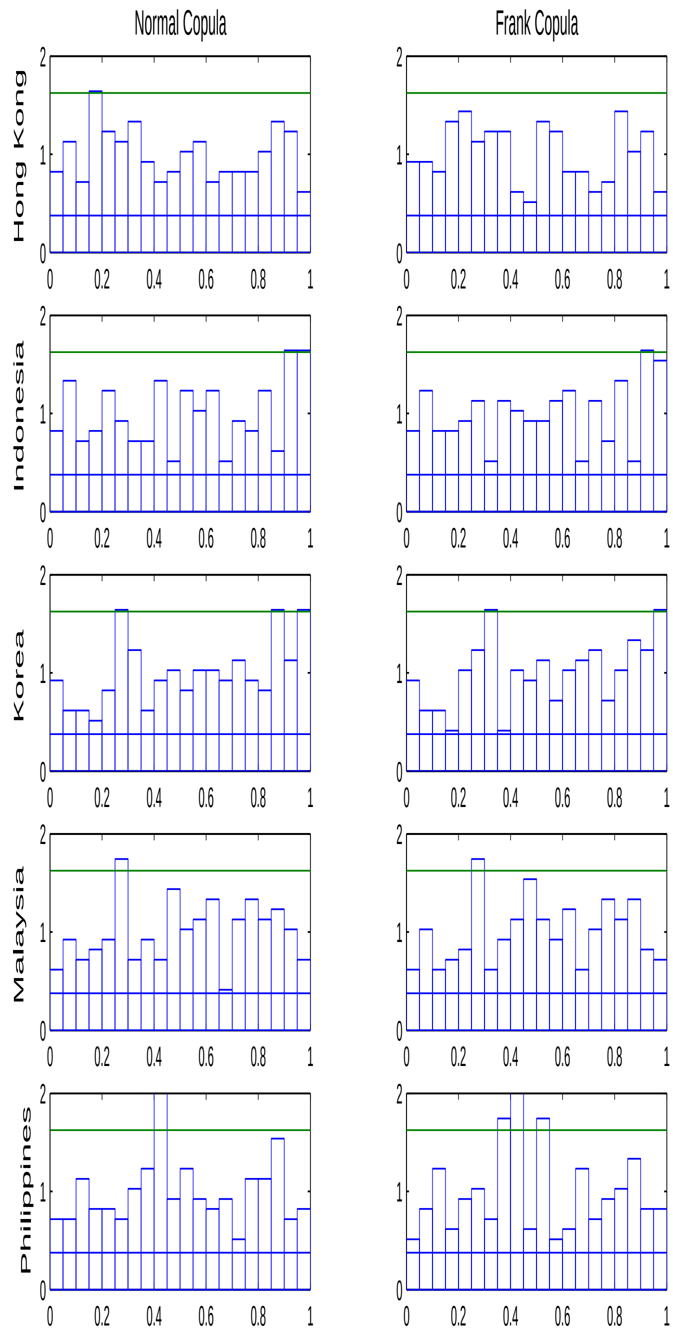
where $f_t(\omega)$ denotes the t distribution density. Given $u_{t+1|t}$ and $v_{t+1|t}$, the sequence of the joint CDF, $\{Z_t^j\}_{t=1}^{n} = \{C(u_{t+1|t}, v_{t+1|t}; \alpha|\Omega_t)\}_{t+1}^{n}$ follows $i.i.d.U(0,1)$ under correct specification of the model. In this application, the 'in-sample' observations are from April 1990 to December 2000, and 'out-of-sample' observations start from January 2001, so there are 195 data points.

Testing for Uniformity Now we assess the uniformity aspect, conditioning on no serial correlation, by plotting the graphical density estimate of the probability integral transform. Figure I.16 and fig:outsample2 shows the density of $\{Z_t^j\}_{t=1}^{n}$ for different models. Although there are a few out-of-bound cases[50], the histograms in figure I.16 and fig:outsample2 are close to uniform. A misspecified likelihood can lead to poor forecasts, therefore density forecast evaluation can help us to flag misspecified likelihoods. In general, we can conclude that any copula model used in this paper is well-specified.

[48] Although there is no closed form of the joint CDF, we can calculate the probability integral transforms. Clements and Smith (2000) shows $\{Z_t^j\}_{t=1}^{n} = \{Z_{1|2,t}^{c} \times Z_{2,t}^{m}\}_{t=1}^{n}$, so we can use the predicted densities for multivariate cases.

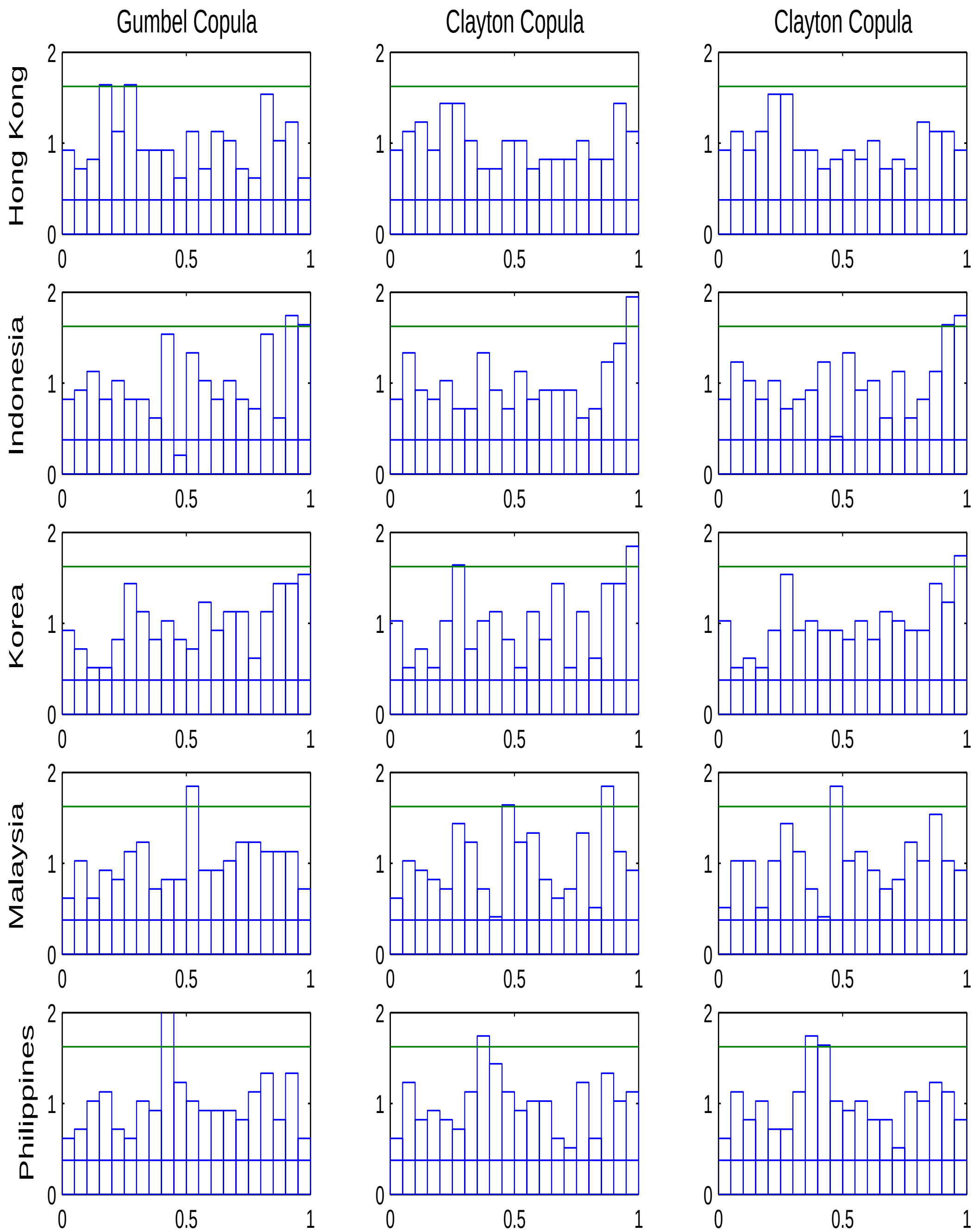
[49] This assumption could be expended to the time-varying dependence assumption. However, the dependence variable is an unobservable variable, so I assume constant dependence parameter for this paper.

[50] If we consider the time-variation of dependence parameter, then we may improve the shape of histograms.



Note: $\{Z_t^j\}_{t=1}^n$ is the probability integral transform of different joint distribution assumption. The first column is for the Normal copula and the second one is for the Frank copula.

Figure I.16 The Density of $\{Z_t^j\}_{t=1}^n$ for Normal and Frank Model



Note: $\{Z_t^j\}_{t=1}^n$ is the probability integral transform of different joint distribution assumption. The first column is for the Gumbel copula, the second one is for the Clayton copula, and the last one is for the Joe-Clayton copula.

Figure I.17 The Density of $\{Z_t^j\}_{t=1}^n$ for Gumbel, Clayton and Joe-Clayton Model

Testing for Independence The *i.i.d.* assumption is tested using LM tests for serial correlation. In order to check dependence in higher moments, I also consider $(Z - \bar{Z})^j$ for j up to 4, suggested in Diebold et al. (1998). Table I.13 reports LM tests for autocorrelation. Here we are unable to reject the null of no serial correlation in any case at the conventional 5% level.

Table I.13: LM tests for *i.i.d.* of out-of-sample

	Lag	Moment	HongKong	Indonesia	Korea	Malaysia	Philippines	Taiwan
Normal	1	1	0.951	0.704	0.516	0.342	0.680	0.623
		2	0.719	0.094	0.480	0.390	0.810	0.953
		3	0.867	0.483	0.471	0.676	0.766	0.804
		4	0.914	0.226	0.637	0.523	0.857	0.582
	5	1	0.948	0.863	0.482	0.610	0.981	0.821
		2	0.972	0.473	0.120	0.932	0.854	0.657
		3	0.967	0.616	0.532	0.972	0.909	0.896
		4	0.985	0.513	0.474	0.963	0.939	0.846
	10	1	0.927	0.653	0.814	0.381	0.858	0.948
		2	0.954	0.839	0.214	0.387	0.815	0.722
		3	0.997	0.753	0.828	0.950	0.891	0.989
		4	0.997	0.734	0.855	0.951	0.955	0.954
Frank	1	1	0.962	0.744	0.483	0.330	0.793	0.610
		2	0.676	0.102	0.517	0.422	0.931	0.946
		3	0.799	0.468	0.468	0.679	0.748	0.792
		4	0.847	0.245	0.633	0.567	0.803	0.582
	5	1	0.962	0.872	0.488	0.658	0.964	0.831
		2	0.970	0.475	0.142	0.955	0.968	0.681
		3	0.963	0.595	0.552	0.982	0.873	0.907
		4	0.963	0.485	0.488	0.976	0.885	0.863
	10	1	0.935	0.652	0.806	0.411	0.831	0.953
		2	0.970	0.846	0.237	0.419	0.816	0.718
		3	0.998	0.761	0.834	0.951	0.728	0.990
		4	0.995	0.714	0.857	0.964	0.738	0.957
Gumbel	1	1	0.867	0.675	0.560	0.329	0.598	0.582
		2	0.951	0.140	0.455	0.357	0.596	0.859
		3	0.828	0.542	0.526	0.684	0.799	0.938
		4	0.792	0.262	0.704	0.479	0.952	0.651
	5	1	0.913	0.861	0.467	0.567	0.988	0.776
		2	0.974	0.586	0.100	0.876	0.496	0.526
		3	0.953	0.729	0.546	0.930	0.949	0.857
		4	0.973	0.674	0.572	0.922	0.975	0.784

Note: The table reports the p-value for χ^2 LM tests (Q statistics) of serial correlation up to 10th order.

Continued...

Table I.13: (Continued) LM tests for *i.i.d.* of out-of-sample

	Lag	Moment	HongKong	Indonesia	Korea	Malaysia	Philippines	Taiwan
Gumbel	10	1	0.894	0.648	0.827	0.332	0.903	0.930
		2	0.954	0.875	0.192	0.394	0.557	0.691
		3	0.992	0.813	0.824	0.945	0.964	0.981
		4	0.992	0.848	0.901	0.939	0.992	0.952
Clayton	1	1	0.906	0.734	0.608	0.297	0.725	0.563
		2	0.410	0.062	0.523	0.454	0.772	0.814
		3	0.619	0.404	0.455	0.687	0.606	0.854
		4	0.623	0.189	0.581	0.529	0.695	0.613
	5	1	0.984	0.843	0.489	0.597	0.941	0.847
		2	0.892	0.422	0.174	0.853	0.975	0.698
		3	0.905	0.575	0.522	0.987	0.737	0.921
		4	0.827	0.416	0.411	0.979	0.779	0.866
	10	1	0.938	0.723	0.791	0.398	0.770	0.959
		2	0.883	0.849	0.244	0.349	0.815	0.753
		3	0.995	0.790	0.783	0.944	0.584	0.991
		4	0.985	0.743	0.759	0.894	0.446	0.957
Joe-Clayton	1	1	0.865	0.692	0.575	0.298	0.639	0.570
		2	0.773	0.102	0.468	0.416	0.843	0.836
		3	0.987	0.484	0.502	0.692	0.720	0.861
		4	0.983	0.227	0.665	0.491	0.853	0.605
	5	1	0.948	0.855	0.458	0.573	0.977	0.828
		2	0.975	0.524	0.107	0.870	0.828	0.651
		3	0.960	0.672	0.500	0.962	0.910	0.903
		4	0.989	0.588	0.483	0.944	0.962	0.838
	10	1	0.917	0.680	0.807	0.342	0.862	0.953
		2	0.955	0.859	0.194	0.343	0.783	0.730
		3	0.996	0.785	0.795	0.939	0.933	0.988
		4	0.998	0.786	0.854	0.900	0.984	0.948

Note: The table reports the p-value for χ^2 LM tests (Q statistics) of serial correlation up to 10th order.

In summary, we can say that these results all show that we have substantial evidence of qualitative and quantitative time-varying dependence. Time variation in the conditional copula seems significant. When the parameters of the conditional copula are allowed to vary through time, they deviate quite substantially from the constant parameter. In addition, we have to notice that this variation has also some deviation from the time-varying linear dependence measure. Sometimes these deviations are substantial, too. This may be because of the failure of the linear dependence measure to capture the true dependence among variables. Although there exists evidence of time variation, there is little evidence of a significant increasing change in dependence structure among Asian stock markets after 1997, when we use copula models. Some results show that there exists a transition from low to high dependence with market crisis, but it is not true in all pairs. In other words, dependency after 1997 and 1998 does not seem high, and is sometimes even lower than the constant level.

We also see that there are different movements in general dependence parameters and tail dependence parameters. Tail dependence can be thought to be the probability of joint extreme realization given one extreme event. We can easily convert to that using the tail dependence parameter. If every single parameter has its own interpretation, then best model selection among many alternative copula models could be meaningless. Also this paper showed that various copula models are proper for analyzing the Asian stock markets. Thus, we have to be careful to use the right dependence measure depending on the purpose and to interpret the implicit meaning. For example, the parameter of Plackett copula introduced in section I.2.B.c can be interpreted as the cross product ratio or odds ratio. Therefore I can guess that the time path of this parameter to be different from the others.

The last thing to mention is in regards to asymmetric dependency. Some of the previous literature found that there exists an asymmetric feature of dependence, such as Longin and Solnik (2001) and Patton (2001). However, we cannot see clear evidence of asymmetric dependency here. The parameters to capture this asymmetry are not significant in all cases, and the sign of the parameters are not even consistent through the sample pairs. This could be because I include two forcing variables, u_t^* and v_t^*, into the evolution equation (I.44) and there is some interaction between those two variables. Therefore another model of the evolution equation would capture asymmetry well. In addition to that, we cannot see any consistent pattern of dependence paths between the bullish phase and the bearish phase, although some pairs have higher dependence during the bullish phase. This suggests unclear evidence of asymmetric dependence in general at least among

Asian stock markets.

I.5 Conclusion

Dependence is one of the most important measures in the economic or financial literature. One of the most frequently used dependence measures is, of course, the linear correlation coefficient. This linear correlation provides a very simple and convenient summary of two associated variables. However, it is not a complete measure of dependence in many situations. The multivariate distribution, which is the main measure of dependence has complicated features, making it a not so easy task to capture summary measure of dependence. Thus, the linear measure is not suitable generally. Recently, some economists have paid attention to the theory of copulas as an alternative to linear correlation, and the use of copulas also is growing in the applied literature. Although copula was introduced a long time ago, it has not been widely used in the economic literature. But some nice features of copulas promise its wider use in the future.

In order to use a copula-based model, first of all, we have to consider the right marginal distribution for each variable. Without the correct marginal distribution, dependence parameter estimates will be biased downward. Simulation results also show that when we are not sure about the parametric form of the margins, non-parametric estimation of the margins provide a lot of gain, with little bias. In order to test the misspecification of margins, I employed the information matrix test of White (1982) and the nonformal test suggested by Diebold et al. (1998). In this paper GARCH normal and GARCH-t assumption are tested using those tests, and it turns out that GARCH-t is better than GARCH normal.

In this paper, I showed how the time-varying conditional dependence of Asian markets has changed since 1990, using various copula models. During that period, Asian financial (stock) markets experienced a deep financial crisis around 1997 and the crisis lasted for some time after that. There has been a lot of literature devoted to finding evidence of a significant increase in cross-market relationships after the financial crisis, the so-called "correlation breakdown." In the 1990s there was wide agreement of the existence of the correlation breakdown phenomenon. But some recent work by Boyer et al. (1999) and Forbes and Rigobon (2002) found almost no evidence of correlation breakdown using bias-adjusted correlations. First, this paper shows that such a correlation breakdown is not clear using time-varying copula models. As expected, using the DCC-MVGARCH as a benchmark model provides an increasing dependence for 2 years after mid-1997 in all pairs, although that higher dependence period is not the only higher period in the entire sample. However, copula models, which are shown to be appropriate to the Asian stock market, do not provide an increasing dependence result for the crisis period in general. This could be because the linear measure is inadequate in capturing the real dependence.

Second, the time paths of the general dependence and the tail dependence do not show the same or similar movements. In general, the higher dependence in normal or Frank copula does not guarantee higher tail dependence. This could be because both have different characteristics. Another empirical finding of this paper is unclear evidence of asymmetric dependence. Although previous papers, such as Longin and Solnik (2001), seem to find this asymmetry, that is not clear here. This could be due to the difference in model construction or the different interpretations of the result. All of these results are similar with non-parametric margins.

Finally, we detect the relatively high persistence in the time path of dependence. In most cases the persistence parameter exceeds 0.95, sometimes as high as 1, implying high persistence in dependence. This might provide further possibilities and applications of this modeling. For example, the asset allocation based on the conditional time-varying copula parameter could provide better performance. And the value-at-risk analysis related to the conditional tail dependence could provide some insight to risk management. On a practical level, the finding of this paper suggests that utilization of portfolio allocation weight among many choices from copula estimates could reduce the exposure of downside risk and could provide information about what proportion to hedge actively. This work will be extension of some theoretical and applied literature (Ang and Chen (2002), Ang et al. (2006), and so on). Further, this conditional model method may be useful to construct models for forecasting purpose in multivariate case, the focus of many recent works.

References

Ang, A., and Chen, J., 2002: Asymmetric correlation of equity portfolios. *Journal of Financial Economics*, **63**(2), 442–294.

Ang, A., Chen, J., and Xing, Y., 2006: Downside risk. *Review of Financial Studies*, **19**(1), 1–40.

Baig, T., and Goldfajn, I., 1999: Financial market contagion in the asian crisis. *IMF Staff Papers*, **46**(2), 167–195.

Bauwens, L., Laurent, S., and Rombouts, J. V. K., 2003: Multivariate garch models: A survey. *CORE Discussion paper*, **2003/13**.

Bollerslev, T., 1987: A conditional heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, **69**, 542–547.

Boyer, B. H., Gibson, M. S., and Loretan, M., 1999: Pitfalls in tests for changes in correlations. *International Finance Discussion Paper*, **597**.

Breymann, W., Dias, A., and Embrechts, P., 2003: Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance*, **3**(1), 1–16.

Chen, X., and Fan, Y., 2004: Estimation of copula-based semiparametric time series models. *Manuscript, NYU and Vanderbilt University*.

Chen, X., Fan, Y., and Patton, A., 2004: Simple tests for models of dependence between financial time series: with applications to us equity returns and exchange rates. *FMG Discussion Papers*, **dp483**.

Christoffersen, P., 1998: Evaluating interval forecasts. *International Economic Review*, **39**(4), 841–862.

Claessens, S., Dornbusch, R., and Park, Y. C., 2001: Contagion: How it spreads and how it can be stopped?". In *International Financial Contagion*, editors S. Claessens, and K. Forbes, 19–41. Kluwer Academic Publishers, Boston, MA.

Clements, M. P., and Smith, J., 2000: Evaluating the forecast densities of linear and non-linear models: Applications to output growth and unemployment. *Journal of Forecasting*, **19**, 255–276.

Davidson, R., and MacKinnon, J. G., 1993: *Estimation and Inference in Econometrics*. Oxford University Press, New York.

Dias, A., and Embrechts, P., 2003: Dynamic copula models for multivariate high-frequency data in finance. *ETH-Zentrum working paper*.

Diebold, F. X., Gunther, T. A., and Tay, A. S., 1998: Evaluating density forecasts with applications to financial risk management. *International Economic Review*, **39**(4), 863–883.

Durrleman, V., A., N., and Roncalli, T., 2000: Which copula is the right on? *Working paper, Credit Lyonnais*.

Embrechts, P., Lindskog, F., and McNeil, A., 2003: Modelling dependence with copulas and applications to risk management. In *Handbook of Heavy Tailed Distributions in Finance*, editor S. Rachev, 329–384. Elsevier.

Embrechts, P., McNeil, A., and Straumann, D., 2002: Correlation and dependence in risk management: properties and pitfalls. In *Risk Management: Value at Risk and Beyond*, editor M. Dempster, 176–223. Cambridge University Press, Cambridge.

Engle, R., 2002: Dynamic conditional correlation-a simple class of multivariate garch models. *Journal of Business and Economic Statistics*, **20**(3), 339–350.

Engle, R., and Sheppard, K., 2001: Theoretical and empirical properties of dynamic conditional correlation multivariate garch. *UCSD Discussion paper 2001-15, UCSD*.

Fermanian, J.-D., 2003: Goodness-of-fit tests for copulas. *Working paper Crest*.

Fermanian, J.-D., Radulovic, D., and Wegkamp, M., 2002: Weak convergence of empirical copula processes. *Working Paper CREST*, **2002-06**.

Fermanian, J.-D., and Scaillet, O., 2003: Nonparmetric estimation of copulas for time series. *Journal of Risk*, **5**(4), 25–54.

Forbes, K., and Rigobon, R., 2002: No contagion, only interdependence: Measuring stock market co-movements. *Journal of Finance*, **57**(2), 2223–62.

Genest, C., Ghoudi, K., and Rivest, L.-P., 1995: A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, **82**(3), 543–552.

Genest, C., and Rivest, L.-P., 1993: Statistical inference procedures for bivariate archimedean copulas. *Journal of the American Statistical Association*, **88**(423), 1034–1043.

Glosten, L., Jagannathan, R., and Runkle, D., 1993: Relationship between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, **48**, 1779–1801.

Hamilton, J. D., 1994: *Time Series Analysis*. Princeton University press, Princeton, New Jersey.

Hamilton, J. D., 1996: Specification testing in markov-switching time-series models. *Journal of Econometrics*, **70**, 127–157.

Joe, H., 1997: *Multivariate models and dependence concepts*. Chapman & Hall, London.

Joe, H., and Xu, J. J., 1996: The estimation method of inference functions for margins for multivariate models. *Technical Report 166, Department of Statistics, University of British Columbia*.

Jondeau, E., and Rockinger, M., 2001: Conditional dependency of financial series: An application of copulas. *HEC working paper*, **723**.

King, M., and Wadhwani, S., 1990: Transmission of volatility between stock markets. *Review of Financial Studies*, **3**(1), 5–33.

Longin, F., and Solnik, B., 2001: Extreme correlation of international equity markets. *Journal of Finance*, **56**(2), 649–676.

Loretan, M., and English, W. B., 2000: Evaluating 'correlation breakdowns' during periods of market volatility. In *International financial markets and the implications for monetary and financial stability*, 214–31. Bank for International Settlements, Basel, Switzerland.

Nelsen, R., 1999: *An Introduction to Copulas*. Springer-Verlag, New York.

Patton, A., 2001: Modelling time-varying exchange rate dependence using the conditional copula. *UC San Diego working paper*, **09**.

Patton, A., 2004: On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics*, **2**(1), 130–168.

Pesaran, H., and Pick, A., 2003: Econometric issues in the analysis of contagion. *mimeo, University of Cambridge*.

Poon, S.-H., Rockinger, M., and Tawn, J. A., 2002: Modelling extreme-value dependence in the international financial market. *Working paper, University of Lancaster*.

Rodriguez, J. C., 2003: Measuring financial contagion: A copula approach. *EURANDOM working paper*.

Tse, Y. K., 2000: A test for constant correlations in a multivariate garch model. *Journal of Econometrics*, **98**, 107–127.

White, H., 1982: Specification testing in dynamic models. In *Advances in Econometrics - Fifth world congress, Vol. 1*, editor T. F. Bewley, 1–58. Cambridge University Press, Cambridge.

II

Return Distributions and Linear Portfolio Choice Rules

II.1 Introduction

The classic mean-variance asset allocation problem studied by Markowitz (1952) has given rise to a whole new area of modern portfolio theory. Under the assumption of "compactness" or "small risk" of the distribution of stock returns, Samuelson (1970) shows that the importance of all moments beyond the variance is much smaller than that of the expected value and variance, and the variance is as important as the mean to investors' welfare. Even though mean-variance analysis is consistent with expected utility maximization only when distributions are normal or utility functions are quadratic, more general problems involving expected utility optimization can be approximated through tradeoff between mean and variance. (Levy and Markowitz, 1979; Kroll et al., 1984).

Recent papers, however, extend the Mean-Variance analysis to higher-order moments and emphasize the need to consider moments beyond the mean and variance in portfolio choice and asset pricing.[1] The fundamental reason for

[1] Kraus and Litzenberger (1976) and Harvey and Siddique (2000) find that the market takes account of skewness in the valuation of assets using, namely 3-moments CAPM, so that investors require compensation for the holding of negatively skewed assets. Dittmar (2002) provides nonlinear pricing kernels, which is restricted by moment preference. Also, see Harvey et al. (2004); Jondeau and Rockinger (2004); Guidolin and Timmermann (2005b), etc.

extending interest to higher moments departs from non-normality of distributions.[2] Lai (1991) and Chunhachinda et al. (1997) use polynomial goal programming to show that investors trade expected return of the portfolio for skewness. Athayde and Flôres Jr (2004) give methods for determining higher dimensional 'efficient frontiers.' Jondeau and Rockinger (2005) investigate the conditional allocation of wealth, departing from non-normality.

It is difficult, however, to know the effective trade-offs among moments and what the preference of investors is over the higher moments. We generally believe that investors prefer higher expected returns and lower risk, consistent with Markowitz's insight. However, higher moments such as skewness have received relatively little attention, apparently because of the difficulty of dealing with higher moments. Markowitz's insight leads to a trade-off relationship between mean and variance in the simple mean-variance cases, but this offers no insight into people's reaction to higher moments in general.

This paper proposes one possible explanation for a trade-off relationship among moments under general utility functions[3] using flexible joint distributions based on copula representations. In general, consideration of higher moment preferences is intricately linked to the form of the utility function. Optimal asset allocation problems can be very sensitive to the choice of the utility function in the expected utility framework. Therefore, the choice of the utility function is an important factor. A family of commonly used utility functions is the HARA (Hyperbolic Absolute Risk Aversion) or LRT (Linear Risk Tolerance) class. The problem of using that class of utility functions is that it is very difficult to get closed form solutions for investors who maximize their expected utility.

One way to consider higher moments is to include these higher moments

[2]Recently, a great deal of empirical evidence suggests that characteristics of returns on financial assets can be captured by predictability, time variation of moments, volatility clustering, and asymmetric correlation depending on the underlying regimes. See Longin and Solnik (1995, 2001); Kandel and Stambaugh (1996); Ang and Bekaert (2002a); Ang and Chen (2002); Guidolin and Timmermann (2005a), etc.

[3]Literature exists about the form of utility functions. As almost everyone agrees, the assumption of standard agents' preferences, and the standard arguments of positive marginal utility and risk aversion lead to $U' > 0$ and $U'' < 0$. As shown in Arditti (1967), decreasing absolute risk aversion implies $U''' > 0$. In order to rule out certain counterintuitive risk-taking behavior with decreasing absolute risk aversion, Kimball (1993) and Pratt and Zeckhauser (1987) show that $U'''' < 0$ is sufficient condition. Also, see Scott and Horvath (1980).

directly into the objective (utility) function and assign weights to each moment. See Harvey et al. (2004). Another way is to expand a generic utility function to an m-th order Taylor series, and then take expectations. For this method, one has to define the relevant moments of the return distribution to analytically study the effect of each moment on investment decisions. Also, one should very carefully choose the approximation point. In practice, up to third or forth moments are considered depending on the researcher's interest and expansion has been done around the expected final wealth from the previous period or the wealth of the previous period.[4] From this expansion, of course, we can see that higher expected returns and right-skewed distributions lead to higher expected utility, while the second and fourth moments decrease in expected utility: see Guidolin and Timmermann (2005b). However, this fact does not give any direct information about the relationship between the optimal asset allocation decision and each moment, because the optimal asset allocation decision is an unknown non-linear function of every moment and is not related to the level of the expected utility. The other way is to calculate the portfolio choice decisions using numerical optimization procedures. See, Barberis (2000); Balduzzi and Lynch (1999), etc. However, this method is a black box, thus it does not provide any idea about the relationship between higher-order moments and portfolio choice decisions.

The first contribution of this paper is to investigate the average relationship between a CRRA investor's decision and moments (or moment parameters) in a given data set. This relationship provides a simple tool to advise the investors how to allocate wealth among assets. We use the Emerging Market (EM) index and the S&P 500 index as our data set. That the Emerging Markets have non-normal distributions is well known (Harvey, 1995), so it is difficult to apply the standard mean-variance criterion. Here, we propose a useful approach, linear projection, to investigate the marginal effect of each moment parameter. Additionally we use copulas to represent flexible joint distributions. The use of copulas allows

[4]In terms of expected utility, Loistl (1976) provides the idea of where to expand in Taylor series expansion for convergence of expected utility with exponential, power and logarithmic utility function cases. For recent applications, see Jondeau and Rockinger (2005), and Guidolin and Timmermann (2005b).

us to disentangle the effects from margins and dependence, therefore we can easily capture more precise characteristics of the return distribution found in empirical data.

We find that these simple linear projections explain from 75% in very poor cases and up to 99% of CRRA investors' portfolio investment behavior. Also we find that the first moment is a great deal more important than what we expected from the mean-variance criterion. For the second moment, we find that the explanatory power is extremely low with lower risk aversion, although it increases with higher risk aversion. One striking feature about the second moment is that CRRA investors' optimal choice leads to holding more of assets which become more volatile in the presence of a riskless asset. Also, we find that skewness is marginally important, but kurtosis information is not important to a CRRA investor's decision.

The second contribution is that we show that a linear decision rule can be a preferable alternative decision rule to the Mean-Variance portfolio (MV) decision rule. A linear decision rule can easily co-operate with the current information about moments. Also we can easily include or exclude some uncertainty information due to additivity of linear projection. We conduct performance comparisons between CRRA investors' decision rules with various levels of risk aversion under various joint distributions and other decision rules (a MV decision rule and a naive decision rule). We measure the portfolio's performance using the Sharpe ratio, realized wealth and realized utility. The CRRA decision rule under a multivariate normal assumption outperforms the MV decision rule in terms of realized wealth. In addition, we compare the decision rule based on different joint distributions. The results suggest that more sophisticated models do not perform better, at least in our sample, possibly because of forecasting errors.

This paper is organized as follows. In the next section, we set up the 'pure' asset allocation problem and describe various utility functions and distributions. In Section II.3, we examine the asset allocation problem under the regime switching model as a special case of a dynamic portfolio selection problem. Then, in Section II.4, we present a simple linear projection based on the estimates of the regime switching model processes. We provide the results of various specifications. The results from Section II.4 are applied to investigate the case of time-varying investment opportunities, time-varying mean and variance, and out-of-sample performance test in Section II.5. Section II.5 provides out-of-sample performance results and compares performances. Section II.6 offers our conclusions. Appendix Apendix A provides details of the skewed Generalized t distribution used to consider the flexible joint distribution, and Appendix Apendix B provides details of using a Griddy-Gibbs Sampler.

II.2 Asset Allocation Problem

The 'pure'[5] asset allocation problem is to find the allocation where the investor's expected utility is maximized with some constraints. Suppose that the investor's utility function is $U(W_{t+T};\theta)$, which depends on wealth at time $t+T$, W_{t+T}, and θ describes the characteristics of the utility function. The investors choose among N assets, whose returns, $r_{i,t},\ i = 1, 2, \cdots, N$, are continuously compounded. The Nth asset would be a riskless asset. Portfolio weights are represented in the vector, $\boldsymbol{\alpha}_{N,t} \equiv (\alpha_{1,t}\ \ \alpha_{2,t}\ \ \cdots\ \ \alpha_{N,t})'$. One constraint about weights is $\sum_{i=1}^{N} \alpha_{i,t} = 1$.[6] In this setting, the asset allocation problem of buy-and-hold investors who want to maximize the terminal expected utility at time $t+T$ conditional on the information up to t period, becomes

$$\begin{aligned} \max_{\boldsymbol{\alpha}_{N,t}} E[U(W_{t+T};\theta)|Z_t] \\ s.t.\ 1 - \boldsymbol{\alpha}_{N,t}\boldsymbol{i}_N = 0 \\ W_{t+T} = \boldsymbol{\alpha}'_{N,t}\exp(\mathbf{R}_{N,t+T})W_t \end{aligned} \tag{II.1}$$

[5]This means there is no interim consumption

[6]If there is a riskless asset, the Nth asset is considered as a riskless asset. Therefore, $(1-\boldsymbol{\alpha}'_{N-1,t}\boldsymbol{i}_{N-1})$, where $\boldsymbol{i}_{N-1}$ is a vector of $(N-1)$ ones, is invested in a riskless asset. And return for a riskless asset is $r^f_t = r_{N,t}$.

where, $\mathbf{R}_{N,t+T} = \mathbf{r}_{N,t+1} + \mathbf{r}_{N,t+2} + \cdots + \mathbf{r}_{N,t+T}$, and $\mathbf{r}_{N,t+j}$ is the $(N \times 1)$ vector of returns at $t+j$ period, where $j = 1, 2, \cdots, T$. Therefore, $\mathbf{r}_{N,t+j}$ is the vector of N assets' continuously compounded returns from $t+j-1$ to $t+j$ period, and $\mathbf{R}_{N,t+T}$ is the vector of continuously compounded returns over the T-period horizon. Consequently, $\exp(\mathbf{R}_{N,t+T})$ is the vector of continuously compounded gross returns of N assets. Depending on the problem, we can impose a short-selling or borrowing constraint with $\alpha_{i,t} \in [0,1],\ i = 1, 2, \cdots, N$.

At this point, we face two decision problems in order to solve the optimization problem. The first one is the choice of the utility function and the other is the joint distribution in a parametric setting. These two elements are very important and the result of asset allocation is very sensitive to these choices. To see how the choice of utility function, or different preferences, affects the result, we consider two different utility models: Mean-Variance preferences and power utility functions. Also, to see the effect of different distributions on the expected utility, we consider flexible distributions, applying copula functions.

II.2.A Utility functions

II.2.A.a Mean-Variance Preference

Consider an investor with standard mean-variance preferences. These preferences assign an uncertain prospect of W_{t+T} over the mean and variance. In general, we can express the investor's objective function in (II.1) as the following

$$E[U(W_{t+T};\theta)|Z_t] = E[W_{t+T}|Z_t] - \frac{\gamma^{MV}}{2} V[W_{t+T}|Z_t] \tag{II.2}$$

where $\gamma^{MV} \geq 0$ and measures the coefficient of absolute risk aversion.

This Mean-Variance (MV) preferences have, however, a major theoretical drawback: they may fail to be monotone. It may happen that an agent with mean-variance preferences strictly prefers less to more, thus violating one of the most compelling principles of economic rationality.[7] In addition, another defect of the

[7] Due to this violation of monotonicity, it also violates the independence axiom of expected utility theory.

MV preferences could be that it does not count on higher moments of the distributions, which may sometimes matter to investors. Although this objective function seems to be problematic, practitioners seem to use this MV very often. The main virtue of this specification of preferences is due to its analytically simple tractability, clear intuitive meaning, and ease of interpretation: investors are assumed to maximize a preference functional defined simply over the mean and variance of their portfolios. Due to this appealing feature of MV preferences, we use MV preferences here. Thus, we can directly compare this result to other alternatives' results.

Given this objective function, the optimal solution without a riskless asset will be

$$\boldsymbol{\alpha}^*_{N,t} = \frac{1}{\gamma^{MV} W_t} \Sigma^{-1}_{t+T} \mu_{t+T} - \frac{\boldsymbol{i}'_N \Sigma^{-1}_{t+T} \mu_{t+T}}{\gamma^{MV} W_t \boldsymbol{i}'_N \Sigma^{-1}_{t+T} \boldsymbol{i}_N} \Sigma^{-1}_{t+T} \boldsymbol{i}_N + \frac{1}{\boldsymbol{i}'_N \Sigma^{-1}_{t+T} \boldsymbol{i}_N} \Sigma^{-1}_{t+T} \boldsymbol{i}_N \quad \text{(II.3)}$$

where $\mu_{t+T} = E[\exp(\mathbf{R}_{N,t+T})|Z_t]$ and $\Sigma_{t+T} = Var[\exp(\mathbf{R}_{N,t+T})|Z_t]$[8]. Furthermore if the portfolio choice includes a riskless asset,

$$\boldsymbol{\alpha}^*_{N-1,t} = \frac{1}{\gamma^{MV} W_t} {\Sigma^{ex}_{N-1,t+T}}^{-1} \mu^{ex}_{N-1,t+T} \quad \text{(II.4)}$$

where $\mu^{ex}_{N-1,t+T} = E[\mathbf{R}^{ex}_{N-1,t+T}|Z_t]$ and $\Sigma_{t+T} = Var[\mathbf{R}^{ex}_{N-1,t+T}|Z_t]$. $\mathbf{R}^{ex}_{N-1,t+T} = \exp(\mathbf{R}_{N-1,t+T}) - \boldsymbol{i}_{N-1} \exp(Tr^f_t)$ and $Tr^f_t = \sum_{j=1}^{T} r^f_{t+j}$, thus, $\mathbf{R}^{ex}_{N-1,t+T}$ is excess gross return of risky assets. Therefore, the optimal choice is to allocate $\alpha^*_{N-1,t}$ of wealth to the risky assets and to invest the remainder, $(1 - {\alpha^*_{N-1,t}}' \boldsymbol{i}_{n-1})$ in the riskless asset.

[8] For the distributions rather than the multivariate normal distribution, μ_{T+T} and Σ_{t+T} can be calculated by the numerical approximation, which is used in the simulation part of this paper.

II.2.A.b Power Utility

We also consider an investor with power or constant relative risk aversion (CRRA) utility function[9]. The power utility function is

$$U(W_{t+T}) = \begin{cases} \dfrac{1}{(1-\gamma^P)} {W_{t+T}}^{(1-\gamma^P)}, & \text{if } \gamma^P \neq 1 \text{ and } \gamma^P > 0 \\ \log W_{t+T}, & \text{if } \gamma^P = 1 \end{cases} \tag{II.5}$$

where, $\gamma^P = -W \dfrac{U''(W)}{U'(W)}$, thus γ^P is a relative risk aversion measure. In other words, it represents decreasing absolute risk aversion. The power utility function is by far one of the most popular utility functions in the portfolio selection problem. This is largely because the investor's portfolio policy is proportional to wealth and the value function is homothetic in wealth. However, there is generally no closed form solution for the optimization problem (II.1) with the this form of utility function[10]. Due to this limitation, recent papers generally use approximate solutions[11] by Campbell and Viceira (1999, 2001) or numerical techniques, for example, Monte Carlo simulations by Barberis (2000) and Detemple et al. (2003) or Gaussian Quadrature methods by Balduzzi and Lynch (1999) and Ang and Bekaert (2002a). In this paper, we simply apply the Monte Carlo method to find the optimal portfolio weight.

Following Barberis (2000), we approximate the integral using the Monte-Carlo methods and find the optimal portfolio weight in which the approximation value is maximized.

$$\begin{gathered} \max_{\boldsymbol{\alpha}_t} \frac{1}{I} \sum_{i=1}^{I} U(W^i_{t+T}) \\ s.t.\ W^i_{t+T} = \boldsymbol{\alpha}'_{N,t} \exp(\mathbf{R}^i_{N,t+T} | Z_t) W_t \end{gathered} \tag{II.6}$$

where, $\mathbf{R}^i_{N,t+T}$ is the $(N \times 1)$ vector of returns in the i-th Monte Carlo simulation.

[9] As a special case of the HARA (hyperbolic absolute risk aversion) or LRT (linear risk tolerance) class of utility functions, we consider power utility in this paper. Log utility function is considered one special case of the power utility function.

[10] See, Geweke (2001) for more discussion on general cases. In a dynamic asset allocation framework, however, Kim and Omberg (1996) provide closed form solutions with HARA preferences under mean-reverting asset return assumption without interim consumption. Also, see Wachter (2002) and Liu (2006).

[11] Campbell and Viceira (1999, 2001) derive closed form expressions using log-linear approximations. Also, see Campbell et al. (2003).

Although quadrature method suffers from the curse of dimension, Monte Carlo simulation can reduce this problem.

II.2.B Multivariate Joint Distribution

In order to investigate the impact of flexible distributions on the optimal asset allocation rather than the multivariate normal distribution, we use the copula function. Copula[12] is a special multivariate joint distribution which separates the joint distribution into two parts: the dependence structure and the marginal distribution. This feature makes the construction of the joint distribution among the random variables a bit easier and more flexible. The first step in constructing the copula is to transform the random variables $X_1, ..., X_n$ to standard uniform marginal distribution. Specifically, suppose each of random variables $X_1, ..., X_n$ has a continuous cumulative marginal distribution (CDF), $F_1, ..., F_n$. Then, by the following transformation $T : R^n \to R^n, (x_1, ..., x_n) \mapsto (F_1(x_1), ...F_n(x_n))$, we can get uniformly distributed variable. Then the joint distribution function C of $(F_1(x_1), ..., F_n(x_n))$ is the so-called copula of the random variables $X_1, ..., X_n$. It follows that

$$\begin{aligned} F(x_1, ..., x_n) &= \Pr[F_1(X_1) \leq F_1(x_1), ..., F_n(X_n) \leq F_n(X_n)] \\ &= C(F_1(x_1), ..., F_n(x_n)) \end{aligned} \tag{II.7}$$

The following is the formal definition of copula

Definition II.2.1. *An n-dimensional copula is a function $C : [0,1]^n \to [0,1]$ such that*

1. C is grounded and n-increasing.

2. C has margins C_k, $k = 1, 2, ..., n$, which satisfy $C_k(u) = u$ for all u in $[0,1]$.

Equivalently, an n-copula is a function $C : [0,1]^n \to [0,1]$ with the following properties.

1. For every $\mathbf{u}$ in $[0,1]^n$, $C(\mathbf{u}) = 0$ if at least one coordinate of $\mathbf{u}$ is 0, and $C(\mathbf{u}) = u_k$

[12] For a more formal definition, concepts and examples, see Nelsen (1999) and Embrechts et al. (2003).

if all coordinates of $\mathbf{u}$ are equal to 1 except $u_{k.}$

2. For all $(a_1, ..., a_n), \geq 0$, where $V_C([\mathbf{a}, \mathbf{b}])$ is the C-volume of $[\mathbf{a}, \mathbf{b}]$

In this paper we use various copula functions (Normal, Gumbel, and Clayton) and two marginal distributions (Normal and Skewed Generalized T (Skewed GT) distributions) in order to consider various specifications of joint distribution. Skewed GT distribution by Theodossiou (1998) is an extended version of the Generalized t distribution, and several well-known distributions are nested within the skewed GT. These include the symmetric t, Hansen's (1994) skewed t, the power exponential, the Laplace, the normal, and the uniform distributions. Appendix Apendix A provides details on this distribution and some characteristics. Also we derive a Cumulative Density Function (CDF) of this density function in Equation (II.25). And since the bivariate joint normal distribution is a special case of normal copula under normal margins, later we will consider the normal copula with normal margins as the benchmark.

II.3 Asset Allocation under Regime Switching

The presence of persistent regimes in various financial series is well reported in empirical financial literature. For example, Perez-Quiros and Timmermann (2000); Ang and Bekaert (2002a) and Guidolin and Timmermann (2005a) provide evidence of persistency in stock market returns. Asymmetry of correlation under different regimes also is well reported (Longin and Solnik, 1995, 2001; Ramchand and Susmel, 1998; Ang and Chen, 2002; Ang and Bekaert, 2002a). Since a large number of papers, Perez-Quiros and Timmermann (2000); Ang and Chen (2002); Guidolin and Timmermann (2005a), suggest that Markov Switching (MS) process can capture these properties of return process very well, we adopt an MS process for the return process as well as using flexible joint distributions. Ang and Bekaert (2002a) use a MS model in an asset allocation problem for the first time, and recently Guidolin and Timmermann (2005b,c) also adopted a Markov Switching model in various optimal decision problems.

Suppose that the vector of continuously compounded returns, $\mathbf{r}_{N,t} = (r_{1,t}\ r_{2,t}\ \cdots\ r_{N,t})'$ is generated by a Markov Switching process, that is, the mean and covariance in returns are driven by a common state variable, $s_t \in \{1, 2, \cdots, S\}$.

$$\mathbf{r}_{N,t} = \boldsymbol{\mu}_{s_t} + \mathbf{A}\mathbf{r}_{N,t-1} + \boldsymbol{\Sigma}_{s_t}^{\frac{1}{2}}\boldsymbol{\varepsilon}_t,\ \boldsymbol{\varepsilon}_t \sim IIN(\mathbf{0}, \mathbf{I_N}) \tag{II.8}$$

where $\mathbf{A}$ is a diagonal matrix to capture the first-order autocorrelation in each series and the elements are $a_j,\ j = 1, 2, \cdots, N$[13]. $\boldsymbol{\mu}_{s_t}$ is a $n \times 1$ vector of means, depending on the regimes. And, in general, $\boldsymbol{\Sigma}_{s_t}$ depends on the regimes and can be expressed as

$$\boldsymbol{\Sigma}_{s_t} = \begin{pmatrix} \sigma^2_{1,s_t} & \sigma_{12,s_t} & \cdots & \sigma_{1N,s_t} \\ \sigma_{21,s_t} & \sigma^2_{2,s_t} & \cdots & \sigma_{2N,s_t} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1,s_t} & \sigma_{N2,s_t} & \cdots & \sigma^2_{N,s_t} \end{pmatrix} \tag{II.9}$$

When we use a multivariate normal distribution, there are $\frac{N(N+1)}{2}$ parameters in the covariance structure. In this paper, the covariance structure is governed by the copula functions, therefore $N + 1$ parameters are needed to construct the covariance structure. Also we consider 2 risky assets, so the number of parameters are the same. The regimes, s_t, follow a S-state Markov chain with the transition probability

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1S} \\ p_{21} & p_{22} & \cdots & p_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ p_{S1} & p_{S2} & \cdots & p_{SS} \end{pmatrix} \tag{II.10}$$

where $p_{ij} = Pr[S_{t+1} = j | S_t = i; Z_t]$. In this setting, the future return distribution is a mixture of copula functions with weights, which depend on the current state probability and the transition probabilities. Marron and Wand (1992) point out that mixtures of normal distributions provide a flexible family, therefore we can think that mixtures of copula functions provide even more flexible joint distributions. This model allows variation in means, variances and dependence structures

[13]When we consider $\mathbf{A}$ also as a regime switching variable, $\mathbf{A}_{s_t}$, there is no considerable gain in terms of likelihood value. Therefore we just consider $\mathbf{A}$ is not switching with regimes.

of asset returns depending on states. One more attractive feature of this model is that a change in probability can be easily interpreted as a change in distributions. We will discuss this issue in more detail later. Since state variable, s_t, is unobservable, estimation procedures follow the EM algorithm: see Hamilton (1989); Kim and Nelson (1999).

II.3.A Data Description

This paper considers the effect of higher moments on the asset allocation problem. We use two risky assets in this paper. Therefore this analysis uses monthly return on S&P 500 Composite Price Index and S&P/IFC Emerging Markets Composite.[14] A large number of papers has shown that market returns in emerging markets strongly depart from non-normality (see, Jondeau and Rockinger (2003), etc). In fact, the emerging market has experienced huge financial crises, such as the Mexican crisis in 1994, the Russian default crisis in 1998, and the Brazilian devaluation in 1999. These provide useful comparisons between relatively volatile and stable returns' effect of higher moments on the asset allocation. The T-bill rate is used for a riskless asset, obtained from the CRSP dataset (Fama/Bliss riskfree rate). The data covers the period from January 1967 to December 2003, for a total of 444 monthly observations. Table II.1 summarizes the statistics of the series. As noticed, the Emerging Market (EM) was characterized by a higher mean and higher volatility, and the max and min value of EM is higher and lower, respectively. Also the return of EM shows that there is a little autocorrelation. Returns are strongly non-normal and strongly skewed with fat tails.

[14] The International Finance Corporation (IFC) introduced their indices back in 1988 with data calculated back to 1984. Global Financial Data Inc. (www.globalfindata.com) added this data back to 1920. From 1967 on, the following weights are used: Africa/Asia (20%), Latin America (40%) and East Asia (40%). In March 1998, the indices included countries for Latin America (Argentina (32 stocks), Brazil (75), Chile (50), Colombia (16), Mexico (61), Peru (24), and Venezuela (12)), East Asia (China (43), Korea (184), Philippines (49), and Taiwan (98), South Asia (India (72), Indonesia (61), Malaysia (157), Pakistan (24), Sri Lanka (5), and Thailand (65)), Europe (Czech Republic (6), Greece (54), Hungary (13), Poland (29), Portugal (23), Russia (30), Slovakia (5), and Turkey (58)), Middle East/Africa (Egypt (28), Israel (46), Jordan (6), Morocco (11) South Africa (76) and Zimbabwe (10)). In 1999, Standard and Poors took over the calculation of the IFC indices.

Table II.1 Descriptive Statistics

	S&P 500 Index	Emerging Market	US 1M T-bills
Mean	0.0059	0.0068	0.0048
Median	0.0085	0.0096	0.0044
Max	0.1510	0.2303	0.0126
Min	-0.2454	-0.2951	0.0007
St. Dev.	0.0449	0.0571	0.0021
Skewness	-0.5977	-0.6460	0.9110
Kurtosis	5.3647	6.6376	4.5058
J-B stat.	127.61**	271.46**	101.83**
1st autocorr.	0.0096	0.2017	0.9614

Note: ** denotes 1% significance.
The sample period is 1967:01-2003:12.

II.3.B Estimation Results under the Regime Switching Process

Tables II.2 and II.3 summarize parameter estimates of the regime switching return process under various joint distribution assumptions. Subscript 1 represents the Emerging Market return. Superscript represents the regimes. The margins are two cases: Normal and Skewed GT distributions. We consider various copula functions: the mixture of the Normal, the Gumbel and the Clayton. The normal copula represents the symmetric distribution, the Gumbel represents the asymmetry and more probability mass to the upper right tail distribution, and the Clayton represents the asymmetry and more mass to the lower left tail[15]. For example, the distribution of the Normal copula with normal margins represents the same as does the regime switching bivariate normal distribution. The mixture of copulas implies that regime switching occurs in the joint distribution form between the Normal and the Gumbel or between the Clayton and the Normal.

As demonstrated in Tables II.2 and II.3, the returns of regime 1 have a negative mean, higher volatility, but lower dependence[16]. For example, ρ^1 are

[15] Of course it is possible to use another copula function. However, the regime switching model itself is flexible enough to capture complicated forms of heteroskedasticity, fat tails, and skews in the underlying distribution of returns. See Timmermann (2000). To consider the positive or negative asymmetry in the joint distribution, we include the Gumbel and Clayton copulas. In addition we use the Skewed Generalized T distribution, which may give more flexibility. In fact when we use t copula, it is not a better model in terms of likelihood value.

[16] This is a bit different from what most empirical papers found. Generally asymmetry in dependence implies higher volatility and higher dependence. See Longin and Solnik (2001); Ang and Bekaert (2002a).

0.2936 and 0.1625, while ρ^2 are 0.3862 and 0.3354 when normal and skewed GT are used, respectively[17]. One thing to notice is that the Gumbel generally puts more mass on the upper right tail. According to the estimation results, when we use the mixture of the Normal and Gumbel copulas for the joint distribution, the Normal detects regime 1, lower mean and higher volatility regime, and the Gumbel detects regime 2, which might be thought of as the normal period in other specifications. Therefore it may seem that this distributional assumption in the mixture of Normal and Gumbel copulas is rather inappropriate to capture the characteristics of the return processes. Another interesting result is that when we use skewed GT distribution, it classified regimes better. The RCM values of Table II.3 are less than those in Table II.2.

Once we use the regime switching model, one possible interpretation of its results is that at any time during the period, the actual parameter values can be between the parameter value of regime 1 and 2, given the data set. With this interpretation, we can find the moment effect in the next section. We discuss this issue in detail below. Furthermore, in order to compare the performance of various specifications, we will consider some specifications in Tables II.2 and II.3. Table II.2 demonstrates that the Normal and Normal specification is the best in terms of the log likelihood value, therefore, we use that as the benchmark, Model 1, from here forward. In Table II.3, the mixture of Normal and Clayton copulas and Skewed-GT margins is the best, therefore we use that as one of comparison models, Model 3. To consider the margin effect, we consider the Normal copula and Skewed-GT specification as another comparison model, Model 2.

II.3.C Asset Allocation under Regime Switching Assumption

In this section we provide the asset allocation results of the Monte Carlo simulation under the Regime Switching assumption. We follow Barberis (2000)

[17] One difficulty in using the copula function is that it is difficult to compare with conventional dependence measures, for example, correlation. Kendall's τ is used for this conversion. Since there exists one to one mapping between the copula dependence measure and Kendall's τ, first we convert the dependence measure to Kendall's τ, then convert again to corresponding ρ.

Table II.2 Regime Switching Model Estimation I

	Copula	Normal		Normal/Gumbel		Clayton/Normal	
	Margins	Normal		Normal		Normal	
Regime 1	μ_1^1	-0.0158	(0.0020)	-0.0134	(0.0144)	-0.0178	(0.0125)
	μ_2^1	-0.0120	(0.0080)	-0.0125	(0.0086)	-0.0125	(0.0094)
	σ_1^1	0.0906	(0.0091)	0.0947	(0.0110)	0.0921	(0.0094)
	σ_2^1	0.0634	(0.0060)	0.0644	(0.0058)	0.0638	(0.0060)
	ρ^1 (or δ^1)	0.2936	(0.1138)	0.3755	(0.0633)	0.2541	(0.1338)
Regime 2	μ_1^2	0.0115	(0.0027)	0.0108	(0.0030)	0.0115	(0.0029)
	μ_2^2	0.0109	(0.0024)	0.0108	(0.0029)	0.0105	(0.0026)
	σ_1^2	0.0396	(0.0024)	0.0399	(0.0031)	0.0405	(0.0025)
	σ_2^2	0.0367	(0.0022)	0.0371	(0.0023)	0.0372	(0.0021)
	ρ^2 (or δ^2)	0.3862	(0.0566)	1.2369	(0.1529)	0.4078	(0.0579)
Prob.	p	0.6596	(0.1096)	0.9164	(0.0350)	0.6387	(0.0937)
	q	0.9070	(0.0451)	0.6687	(0.1009)	0.9117	(0.0297)
Auto.	a_1	0.1865	(0.0478)	0.1996	(0.0493)	0.1801	(0.0501)
	a_2	-0.0328	(0.0442)	-0.0357	(0.0490)	-0.0251	(0.0474)
Log Likelihood		1479.93		1478.58		1479.48	
RCM		35.2706		33.122		33.1175	
Selection		(Model 1)					

Note: Dependence measures are different depending on the copulas. $0 \leq \rho^i \leq 1,\ i = 1, 2$ is for the Normal copula. $0 < \delta_C^1 \leq \infty$ is for the Clayton copula. $1 \leq \delta_G^2 \leq \infty$ is for the Gumbel copula.
RCM is the regime classification measure from Ang and Bekaert (2002b). $0 \leq RMC \leq 100$ and lower values denote better regime classification.

and approximate the integral using equation (II.6) for the CRRA utility function. In order to comply with the regime-switching model, the approximation form of the objective function is changed to the following form

$$\begin{aligned} &\max_{\boldsymbol{\alpha}_t} \frac{1}{I} \sum_{i=1}^{I} U(W_{t+T}^i; \gamma^P) \\ s.t.\ W_{t+T}^i &= \sum_{s_t=1}^{S} \left\{ \boldsymbol{\alpha}_{N,t}^{'} \exp(\mathbf{R}_{N,t+T}^i | \hat{\theta}_{s_t}) W_t \right\} \times p_{s_t} \end{aligned} \tag{II.11}$$

where, $\mathbf{R}_{N,t+T}^i$ is the $(N \times 1)$ vector of returns in the i-th Monte Carlo simulation conditioning on $\hat{\theta}_{s_t}$. Given p_{s_t}, state probabilities, we can calculate W_{t+T}^i. Finally,

Table II.3 Regime Switching Model Estimation II

	Copula Margins	Normal Skew-GT		Normal/Gumbel Skew-GT		Clayton/Normal Skew-GT	
Regime 1	μ_1^1	-0.0542	(0.0001)	-0.0542	(0.0000)	-0.0563	(0.0000)
	μ_2^1	-0.0392	(0.0213)	-0.0558	(0.0037)	-0.0353	(0.0179)
	σ_1^1	0.0826	(0.0166)	0.0889	(0.0271)	0.0801	(0.0158)
	σ_2^1	0.0651	(0.0150)	0.3808	(0.0062)	0.0633	(0.0080)
	ν_1^1	50.141	(4.2376)	17.089	(0.3690)	118.21	(9.5313)
	ν_2^1	3.7028	(2.2473)	2.0159	(0.0025)	3.7812	(0.6492)
	λ_1^1	0.0864	(0.1191)	0.0423	(0.1183)	0.1053	(0.1106)
	λ_2^1	0.1297	(0.1963)	0.2480	(0.0114)	0.0968	(0.1882)
	κ_1^1	0.7036	(0.2058)	0.6259	(0.2338)	0.7381	(0.1915)
	κ_2^1	3.1599	(2.5725)	88.188	(2.4705)	3.1294	(1.4963)
	ρ^1 (or δ^1)	0.1625	(0.1089)	0.3253	(0.0558)	0.2126	(0.1451)
Regime 2	μ_1^2	0.0116	(0.0026)	0.0118	(0.0027)	0.0110	(0.0032)
	μ_2^2	0.0144	(0.0052)	0.0131	(0.0058)	0.0148	(0.0063)
	σ_1^2	0.0426	(0.0031)	0.0444	(0.0030)	0.0417	(0.0032)
	σ_2^2	0.0364	(0.0018)	0.0368	(0.0016)	0.0364	(0.0018)
	ν_1^2	620.03	(55.124)	1476.1	(35.749)	475.36	(39.867)
	ν_2^2	5.7825	(2.7981)	5.7149	(1.4373)	5.8569	(3.5523)
	λ_1^2	0.0843	(0.0692)	0.0581	(0.0664)	0.1139	(0.0782)
	λ_2^2	-0.0234	(0.0949)	-0.0092	(0.1008)	-0.0235	(0.1061)
	κ_1^2	1.3375	(0.1616)	1.2807	(0.1606)	1.3621	(0.1973)
	κ_2^2	3.3892	(1.0758)	3.4682	(0.1372)	3.3352	(1.1958)
	ρ^2 (or δ^2)	0.3354	(0.0603)	1.0676	(0.1134)	0.3352	(0.0562)
Prob.	p	0.5886	(0.0850)	0.9270	(0.0271)	0.5990	(0.0868)
	q	0.9120	(0.0320)	0.5796	(0.0874)	0.9079	(0.0292)
Auto.	a_1	0.1319	(0.0019)	0.1291	(0.00000	0.0992	(0.0001)
	a_2	-0.0568	(0.0422)	-0.0445	(0.0334)	-0.0567	(0.0425)
Log Likelihood		1491.12		1490.39		1491.62	
RCM		24.6461		20.2765		24.6747	
Selection		(Model 2)				(Model 3)	

Note: Dependence measures are different depending on the copulas. $0 \leq \rho^i \leq 1,\ i = 1,2$ is for the Normal copula. $0 < \delta_C^1 \leq \infty$ is for the Clayton copula. $1 \leq \delta_G^2 \leq \infty$ is for the Gumbel copula.
RCM is the regime classification measure from Ang and Bekaert (2002b). $0 \leq RMC \leq 100$ and lower values denote better regime classification.

given γ^P and $I = 100,000$[18], we can optimize the objective function numerically to get $\boldsymbol{\alpha}_t^*$. The optimal asset allocation is the function of the utility function, the distribution, and the state probability. At this point, we only care about the investors who use buy-and-hold strategy, and the investment horizon is only one period, $T = 1$. For Models 2 and 3, we use the truncated version of the skewed generalized T distribution only for this exercise, because the original skewed GT distribution tends to generate an substantial amount of extreme samples[19]. We choose the truncation points where the cumulative probability to the right or left side of the density is 0.1%. That way, the maximum or minimum samples are very close to the those of the historical data set[20]. For the both utility cases, we set the initial wealth as 1[21]. Finally we repeat the experiment 10 times and take the average of the optimal asset allocation results of each experiment. In this experiment, we do not consider the first order autocorrelation effect, assuming present returns, $\mathbf{r}_{N,t}$, of equation (II.8) is $\mathbf{0}$. We, also, do not place any constraints on the optimal asset allocation. As expected, the optimal asset allocation results vary widely, depending on the objective function and distribution. This is confirmed by the following figures. Figure II.1 show the optimal asset allocation results with the CRRA utility function. For details of the graphs, see the note below each graph.

As mentioned, the results are very sensitive to distributions. Especially when the skewed GTs are used as the marginal distributions (Models 2 and 3), we can see more variation than when the normal distribution is the margin (Model 1). Comparing Models 2 and 3, in which different joint distributions are used, the results do not change very much.[22] This might indicate that the asymmetry in the

[18] Guidolin and Timmermann (2005c) show that sampling errors and random variation are substantial for sample size, I, is less than 20,000. Also $I = 30,000$ guarantees a substantial reduction in the incidence of sampling error, therefore $I = 100,000$ is ample to ensure negligible sample errors.

[19] These extreme samples seem to generate very different optimal asset allocation results. Guidolin and Timmermann (2005c)'s experiment only consider the multivariate normal case, therefore it is not excusable to apply their argument for the skewed GT distribution case. However, we try to use an ample sample later so that we assume that we can reduce simulation errors even with the skewed GT marginal distribution.

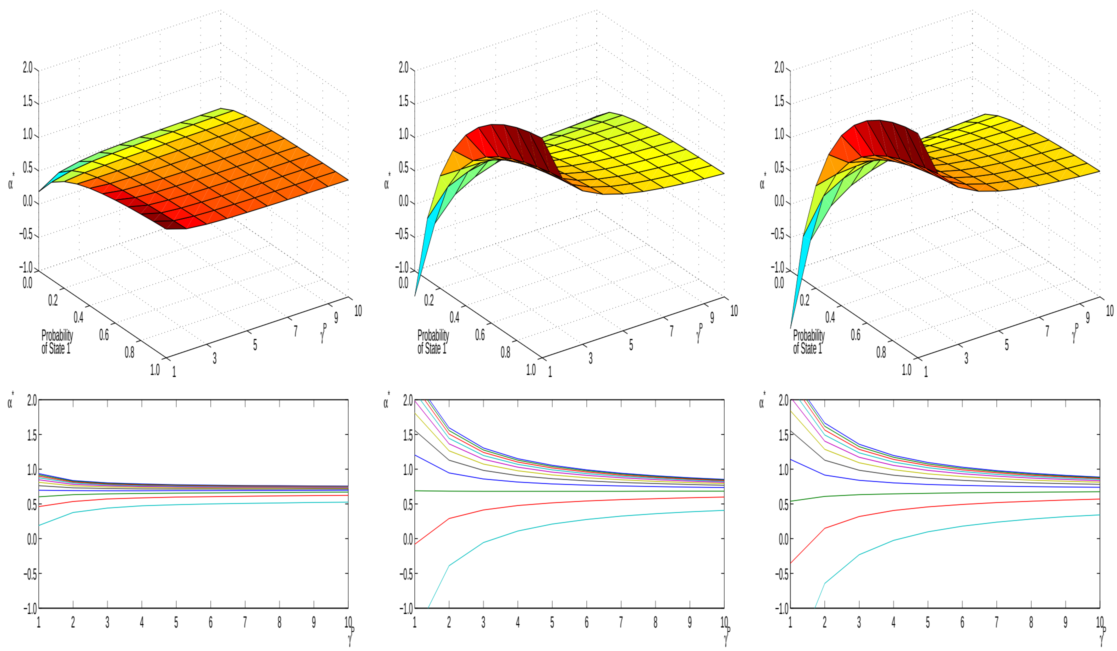
[20] This truncation can ensure the existence of the moments with any distributions. Then we can approximate the moments by the numerical integration. However, there exist the mth non-centered moment for the skewed GT distribution.

[21] The optimal asset allocation result is the function of the initial wealth for the MV utility case, therefore different initial wealth produces different results. However, we set 1 for convenience.

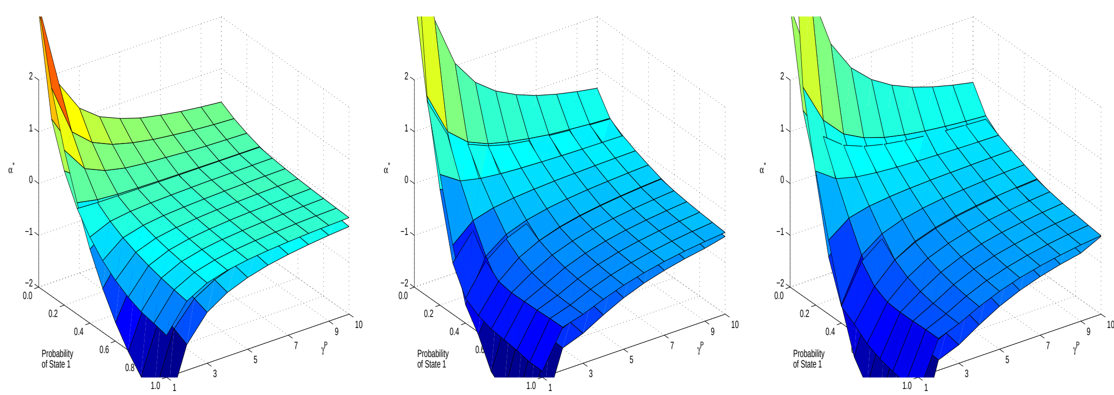
[22] Generally we can see the convergence of α^* with higher risk averseness from the bottom rows of Panel A in Figure II.1. Specifically, when the probability of state 1 is 1.0, α^* tends to decrease and converge to a certain

(a) Model 1 (b) Model 2 (c) Model 3

Panel A. Without a riskless asset



Panel B. With a riskless asset



Note: This is the optimal asset allocation results with the CRRA utility function among 2 risky assets without and with a riskless asset. The first column is the result of Model 1, the second is for Model 2, and the third is for Model 3. Panel A is for the case without a riskless asset, and α^* indicates the weight on asset 2, the S&P 500 index. In Panel A, the lower row is the view from γ^P axis of the corresponding upper row, therefore, the top line represents the optimal asset allocation when the probability of state 1 is 1 and the bottom represents it when the probability is 0. The lines from the bottom represent when the probability increases by 0.1 increments. Panel B is for the case with riskless asset. The top layer, when the probability is high, is the optimal allocation of the first risky asset, the emerging market index.

Figure II.1 Optimal Asset Allocation with CRRA Utility joint distribution does not play a very important role in the investors' decision. Also, more information about distribution may lead to more aggressive investment decisions. The skewed GT margins can capture the fat-tail, skewness and kurtosis of the marginal distributions, therefore this information might help people know the risk factors better and invest aggressively.

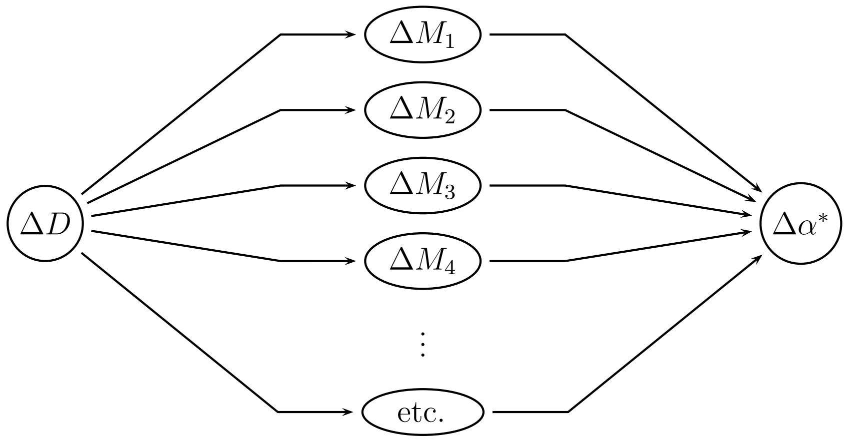
In addition, less risk averse investors' optimal asset allocation tends to be more sensitive to the change in state probability than more risk averse investors. This is very consistent with general asset allocation results, since more risk averse investors tend to diversify their investment than less risk averse investors, intuitively. Therefore more risk averse investors are less sensitive to the state probability change. The difference between the CRRA and the MV utility function is not notable. As shown in the CRRA utility case, the MV utility case also shows similar results. However it could be that the unknown policy functions with the CRRA utility are different from those with the MV utility.

II.4 Do Higher Moments Matter?

The previous results are rather predictable, because the optimal decision is the function of the distribution and the objective function and in the previous case, a change in the state probability means a change in the distribution. Now, we have to ask what kind of factors make a difference in the asset allocation decision. One of the questions we can ask here is 'Do moments matter?' The answer seems to be 'Yes', because we already saw that different distributions result in different optimal asset allocations and the different distributions can be characterized by the difference in moments of distributions. Now, we can ask other questions, 'Do moments higher than the second moment matter?', and 'If the answer is yes, how much and how many higher moments matter?' Figure II.2 helps us conceptualize

level. However, if the risk aversion is extremely high, such as more than 25, and the Clayton copula, or the skewed GTs, or both are used for state 1, α^* tends to increase with non-zero probability of state 1 in the CRRA utility case as a result of the optimization. Therefore, the asymmetry of joint distribution can be considered with really extreme risk averse investors. We do not show this result in this paper, since an extreme risk averse investor is not realistic.

how to answer these questions.



Note: This graph shows that the change in distribution changes at each moment, and the change in each moment changes the optimal asset allocation. Here, ΔD represents the change in distribution, ΔM_i, $i = 1, 2, \cdots$, represents the change in each moment, and $\Delta\alpha^*$ represents the change in the optimal asset allocation.

Figure II.2 Effect of Changing Distribution

Now we define each moment as the conventional mean, covariance, co-skewness, co-kurtosis, and so on. Following Jondeau and Rockinger (2005), we can define covariance, co-skewness and co-kurtosis as the following:

$$\begin{aligned} M_1 &= E_t[\mathbf{r}_N] \\ M_2 &= E_t[(\mathbf{r}_N - \boldsymbol{\mu})(\mathbf{r}_N - \boldsymbol{\mu})'] \\ M_3 &= E_t[(\mathbf{r}_N - \boldsymbol{\mu})(\mathbf{r}_N - \boldsymbol{\mu})' \otimes (\mathbf{r}_N - \boldsymbol{\mu})'] \\ M_4 &= E_t[(\mathbf{r}_N - \boldsymbol{\mu})(\mathbf{r}_N - \boldsymbol{\mu})' \otimes (\mathbf{r}_N - \boldsymbol{\mu})' \otimes (\mathbf{r}_N - \boldsymbol{\mu})'] \end{aligned} \tag{II.12}$$

where $\mu_{t+1} = E_t[\mathbf{r}_{N,t+1}]$. In a parametric model for the joint distribution, however, each moment is the non-linear function of the parameters of the joint distribution. For example, assume the underlying distribution is the joint normal distribution and it has $2N + \frac{N(N-1)}{2}$ parameters. In the bivariate case, there are 5 parameters. In this case, the change in μ_1 alters every moment, implying that 5 parameters fully capture the characteristics of the bivariate normal distribution. Therefore, it is necessary to see how the parameters affect the asset allocation decision. To see the higher moments' effects, we can use a simple projection method.

For the MV utility case, the choice of portfolio is determined by the first two moments. In this case, closed form solutions exist (see, equation (II.3) and (II.4)). Therefore, α^* is the non-linear function of μ_1, μ_2, σ_1, σ_2 and ρ, for the bivariate case. For the CRRA utility case, we do not know the form of the policy function, however, we assume that the policy function, α^*, is the linear function of the parameters, in order to see how higher moments matter.

$$\alpha_j^* = \omega + \sum_{j=1}^{N}\sum_{m=1}^{M_j} \beta_{\theta_{j,m}}\theta_{j,m} + \sum_{d=1}^{D}\beta_{\theta_d}\theta_d + \varepsilon_j \tag{II.13}$$

where, subscript j is the asset index, N is the number of assets, m is the parameter index and M_j is the number of parameters of jth asset distribution. Subscript d is the dependence parameter index and D is the number of dependence parameters. The sources of ε_j are linear approximation errors and simulation errors. The coefficient parameters, $\beta_{\theta_{j,m}},\ m = 1, \cdots, M_j,\ j = 1, \cdots, N$ and $\beta_{\theta_d},\ d = 1, \cdots, D$, represent the marginal effect of each distributional parameter on the investor's decision. Furthermore, in order to see the level effect of distributional parameters, we can expand equation (II.13) to include the linear parameterizations of the non-linear function of the distribution parameters, $\boldsymbol{\theta} \equiv (\theta_{1,1}, \cdots, \theta_{1,M_1}, \cdots, \theta_{N,1}, \cdots, \theta_{N,M_N}, \theta_1, \cdots, \theta_D)'$, following the suggestion by White (2006). Then, equation (II.13) becomes:

$$\alpha_j^* = \omega + \sum_{j=1}^{N}\sum_{m=1}^{M_j} \beta_{\theta_{j,m}}\theta_{j,m} + \sum_{d=1}^{D}\beta_{\theta_d}\theta_d + \sum_{q=1}^{Q}\beta_{\psi_q}\psi_q(\boldsymbol{\theta}) + \varepsilon_j \tag{II.14}$$

where Q is some finite integer number and ψ_q are nonlinear functions of $\boldsymbol{\theta}$.[23]

As mentioned before, we can reasonably think that the range of parameter values is between two regimes, from the regime-switching estimates in Tables II.2 and II.3 for each model. Thus, based on the regime-switching model estimates, we generate 500 of each parameter value through random sampling from a uniform distribution between two estimate values of each regime. For example, the value

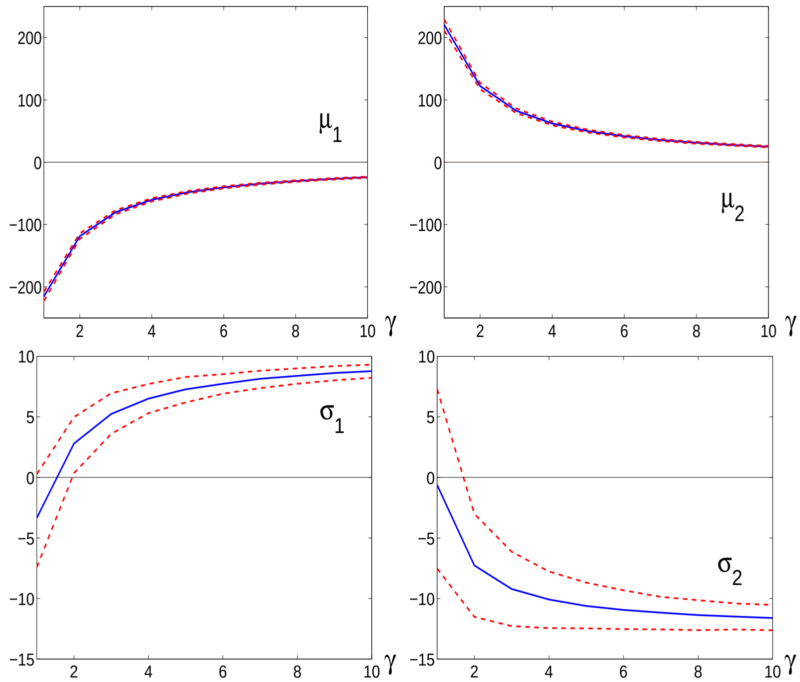
[23] White (2006) says that this specification delivers flexibility while simultaneously eliminating the computational challenges arising from nonlinearity in the parameters.

of μ_1 is -0.0158 in regime 1 and 0.0115 in regime 2, so we generate uniform random numbers between -0.0158 and 0.0115. Of course, this selection criterion might seem to be very ad hoc. Another possible way to generate parameter samples is using a bayesian framework. A bayesian framework makes it possible to get posterior densities of each parameter. Then we can randomly sample from its posterior distribution. However, in this paper, we simply use uniform random sampling. Once we get the random parameter values, we generate the return data for the Monte Carlo simulation given those parameter values. 500,000 return data are generated for model 1, and 100,00 for Models 2 and 3. Finally, given the return data, we find the optimal asset allocation at which the objective function is numerically maximized. Then we can have the entire data set for the linear projection. Next, we run the regression to see the parameter effect on the policy function. Here, we use the bootstrapping method, whose resampling size is 1000, to get a robust confidence interval of coefficients.

II.4.A Without a Riskless Asset

II.4.A.a Linear Marginal Effect

Figure II.3 shows the regression results based on the equation (II.13) for Model 1, which is equivalent to the bivariate normal distribution. The coefficient estimates can be interpreted as the average partial effect of each distribution parameter on the optimal decision. Here, we present the case in which the dependent variable is the optimal asset weight of asset 2, α_2^*. In the case without a riskless asset and with only 2 risky assets, the regression analysis for asset 1 is redundant, since $\alpha_1^* = 1 - \alpha_2^*$. This result is quite consistent with our conventional thought. The weight on asset 2, α_2^*, increases, when asset 2's expected return, μ_2, increases and asset 1's volatility, σ_1, increases. Also the magnitude of the first order moment's partial effect decreases toward 0 and the second order moment's partial effect increases with higher risk aversion. At the lower level of risk aversion, the partial effect of the first order moment parameters is extremely high. For instance, the expected return of asset 2 increases by 1%, the investor with $\gamma = 2$ increases the weight on asset 2 almost by 1.

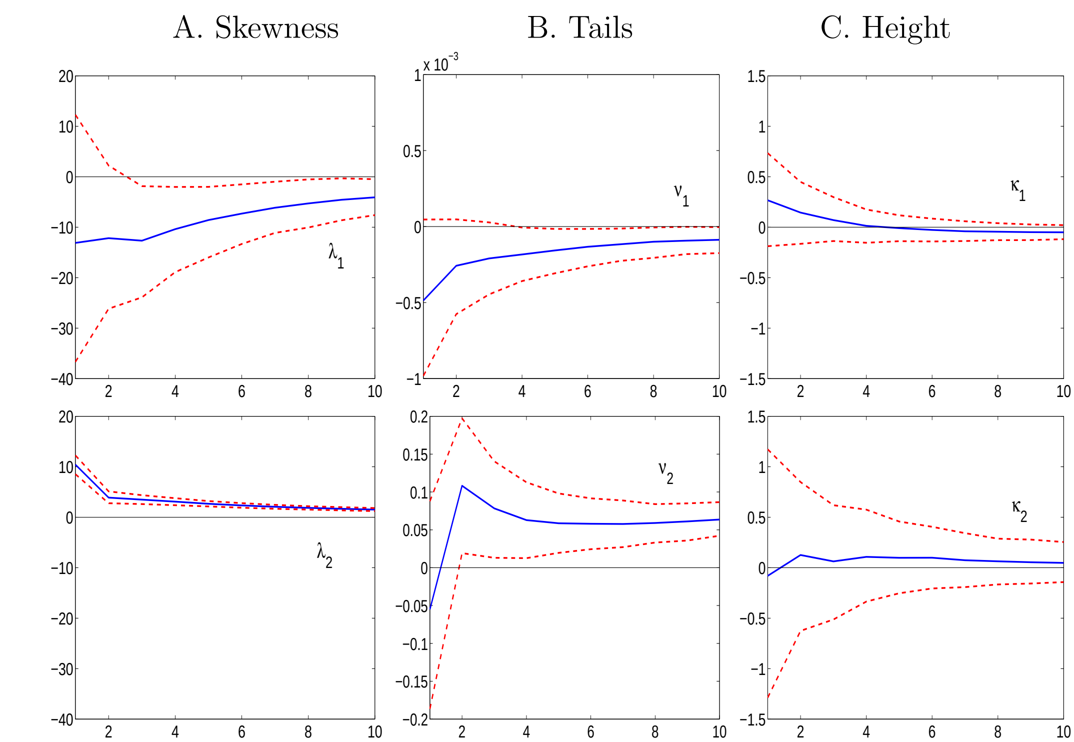


Note: This graph shows the coefficient estimates based on equation (II.13) with Model 1, depending on different risk aversion level, γ. Dependent variable is α_2^*, the Emerging Market Index. Dotted lines are 95% confidence intervals.

Figure II.3 Coefficients on the Optimal Decision w/o Riskless asset of Model 1

Now, we can investigate the higher order moment's marginal effect, using Model 2 and Model 3. Figure II.4 summarizes the regression results for Model 3, only focusing on the higher moments, as the results for the first and second order moments' effects are similar to Model 1, and are all significant. However, difference exists in the marginal effect magnitude. Although we do not report it here, the magnitude of the first moment's partial effect, β_{μ_1} and β_{μ_2}, becomes significantly smaller, compared to Model 1. This may be explained by the existence of higher order moment risk. Model 2 also provides similar results to Model 3, therefore we do not report the results from Model 2.

As mentioned above, the margins used for Model 2 and Model 3 are the skewed GT distribution. The skewed GT distribution is controlled by 4 parameters after considering μ. As shown in Appendix A, each moment is the complicated function of every parameter. However, in general, we can say that σ governs the variance, ν controls the tails, λ controls the skewness, and κ, the heights of the density. Unlike Model 1, the decision makers have to consider various risks, variance as well as skewness, kurtosis, and so on. For the present task, we will define risk as higher than the second order moment. From Figure II.4, we can see what kinds of risk are very important to decision makers. The coefficients for $\kappa_j,\ j = 1, 2$, controlling the heights, do not seem to be significant in almost any case. The coefficients for $\lambda_j,\ j = 1, 2$, controlling the skewness, and $\nu_j,\ j = 1, 2$, controlling the tails, are significant. Therefore, we can say that the investors with the CRRA type utility function react to the change in skewness and tailness, which is not the factor which the investors with the MV type utility function consider.

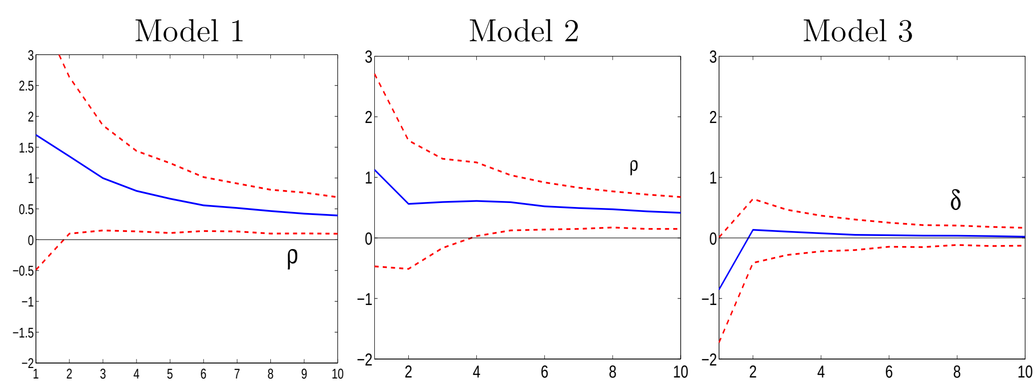


Note: This graph shows the coefficient estimates of higher order moment parameter based on the equation (II.13) with Model 2, depending on different risk aversion level, γ. Dependent variable is α_2^*, the Emerging Market Index. Dotted lines are 95% confidence intervals.

Figure II.4 Coefficients of Higher Order Moments Parameter on the Optimal Decision w/o Riskless Asset of Model 2

II.4.A.b Dependence Structure

We want to see whether dependence structure matters. Figure II.5 reports the effects of dependence parameters. Recall that the dependent variable is α_2^*. We can see the significant positive marginal effect for Model 1, implying that the preference is for asset 2 (U.S. market) over asset 1 (Emerging Market) in the presence of a higher dependent relationship. However, the marginal effects using Model 3 are not significant, and Model 2 is in-between. The main difference between Model 1 and Models 2 and 3 is the assumption of the marginal distributions, and the big difference between Models 1 and 2 and Model 3 is the assumption of the joint density. Recall that we use the normal distribution for the margins for Model 1. For Models 2 and 3, we alternatively assume the skewed GT distribution for the margins. Once we consider the higher order moments, such as skewness, in the margins and the asymmetric joint distribution, the dependence parameter loses importance. This result continues to be the case with a riskless asset later. This may provide a very important implication in the portfolio choice problem. Depending on the assumption on marginal distributions and joint distribution, the dependence structure may or may not be very important in the asset allocation problem.



Note: This graph shows the coefficient estimates of dependence parameter based on equation (II.13), depending on different risk aversion level, γ. Dependent variable is α_2^*, the Emerging Market Index. Dotted lines are 95% confidence intervals.

Figure II.5 Coefficients of Dependence Parameter on the Optimal Decision w/o Riskless Asset

II.4.A.c Nonlinearity

In this section, we investigate the relationship between the CRRA investor's decision and the uncertainty characterized by the distributional parameters more deeply using equation (II.14). With the liner relationship framed in this way, an important next question is, "What choices of basis functions, ψ_q, are available?" Of course, there is a vast range of possible choices of basis functions. However, we select some basis functions based on trial-and-error. In this paper, the set of basis functions is $\boldsymbol{\psi} = (\frac{\mu_1}{\sigma_1}, \frac{\mu_2}{\sigma_2}, \rho\sigma_1\sigma_2)'$, only considering the nonlinearity of first two moments parameters for all models. The reason for choosing $\boldsymbol{\psi}$ is for comparison. This will be discussed in detail later in this paper.

Table II.4 shows the linear projection results based on equation (II.14) when $\gamma = 3$. When we include $\boldsymbol{\psi}$, we can see the level effect of parameters. For example, we can interpret the estimation results based on equation (II.13) as a constant partial (marginal) effect, $\frac{\partial \alpha_j^*}{\partial \boldsymbol{\theta}}$. However, inclusion of $\boldsymbol{\psi}$ changes the partial effect as $\boldsymbol{\psi}$ is the function of other variables, so it is not constant any more. From the estimation results of Model 1, we observe that $\frac{\partial \alpha_2^*}{\partial \mu_1} = 19.38 - 6.06\frac{1}{\sigma_1}$, implying that the magnitude of the marginal effect of μ_1 on α_2 is getting smaller with higher σ_1. Also $\frac{\partial \alpha_2^*}{\partial \mu_2} = 15.42 + 3.14\frac{1}{\sigma_2}$ implies that the magnitude of the marginal effect of μ_2 on α_2^* becomes smaller with higher σ_2. For second order moment parameters, we can get the following: $\frac{\partial \alpha_2^*}{\partial \sigma_1} = -0.61 + 6.06\frac{\mu_1}{\sigma_1^2} + 484.14\rho\sigma_2$, $\frac{\partial \alpha_2^*}{\partial \sigma_2} = -19.89 - 3.13\frac{\mu_2}{\sigma_2^2} + 484.14\rho\sigma_1$, and $\frac{\partial \alpha_2^*}{\partial \rho} = -1.05 + 484.14\sigma_1\sigma_2$. And the coefficients of nonlinear terms are all significant, therefore we might say that a level effect exists.

We also extend this interpretation to Models 2 and 3. However, the coefficients of $\rho\sigma_1\sigma_2$ for Model 2 and $\delta\sigma_1\sigma_2$ for Model 3 are not significant. This may be explained by insignificance of ρ (or δ) even in linear regression based on equation (II.13). Also none of the nonlinear basis functions using higher order moment parameters, $\psi(\nu_1, \nu_2, \lambda_1, \lambda_2, \kappa_1, \kappa_2)$, seems to be significant or add more explanatory power. White (2006) states that basis functions should deliver a good

Table II.4 Linear Parameterization of Non-Linear Functions of Distribution Parameters

	Model 1			Model 2			Model 3		
Const.	1.459	[0.672	2.235]	-1.297	[-7.127	3.860]	0.343	[-2.541	3.349]
μ_1	19.38	[10.29	28.52]	-19.82	[-30.15	-10.17]	-18.60	[-29.24	-8.354]
μ_2	15.42	[-0.928	32.59]	19.64	[5.801	33.19]	26.24	[11.07	41.38]
σ_1	-0.614	[-6.437	5.465]	6.157	[-4.825	17.22]	5.061	[-5.054	16.62]
σ_2	-19.90	[-27.68	-11.94]	-7.748	[-20.95	7.029]	-2.842	[-15.37	10.51]
ρ (or δ)	-1.047	[-2.307	0.199]	-0.267	[-2.910	2.468]	0.645	[-0.794	2.144]
ν_1				0.000	[-0.000	0.000]	0.000	[-0.000	0.000]
ν_2				0.047	[-0.016	0.106]	0.014	[-0.046	0.077]
λ_1				17.39	[-40.43	79.80]	-8.686	[-24.11	6.641]
λ_2				4.229	[3.368	5.088]	3.660	[2.591	4.751]
κ_1				-0.040	[-0.244	0.156]	0.145	[-0.070	0.350]
κ_2				-0.088	[-0.668	0.498]	0.073	[-0.532	0.712]
μ_1/σ_1	-6.062	[-6.614	-5.530]	-1.522	[-2.108	-0.923]	-1.723	[-2.327	-1.123]
μ_2/σ_2	3.137	[2.320	3.902]	1.127	[0.461	1.805]	0.878	[0.170	1.613]
$\rho\sigma_1\sigma_2$	484.1	[119.4	833.4]	118.1	[-715.8	882.2]	-168.8	[-671.4	296.5]

Note: The dependent variable is α_2^*, which represents the weight for the S&P 500 index. Square brackets show 95% confidence interval.

approximation using as small a value for Q as possible. This is one reason not to add more basis functions of higher order moment parameters.

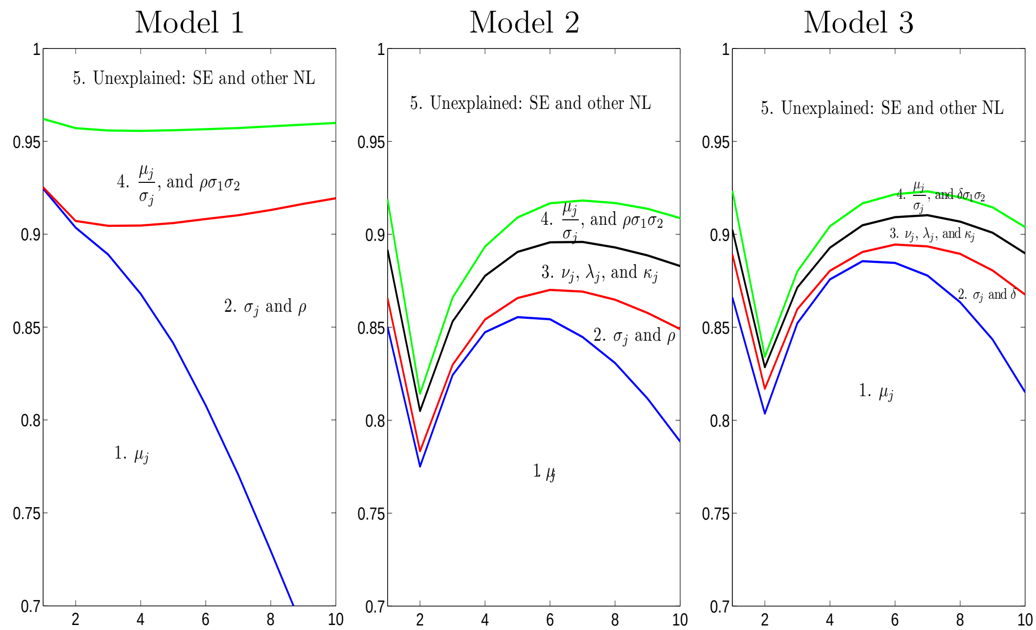
II.4.A.d Explanatory Power

In this section, we investigate which elements of joint density function has explanatory power, using the information from $\bar{R}^2$. For model 1, the first order moment parameters, μ_j, $j = 1, 2$, seem to have strong explanatory power to α_2^*. When γ is low, the explanatory power of μ_j is more than 90%, but it decreases along with higher γ. When we include the second order moment parameters, σ_j, $j = 1, 2$, and ρ, their explanatory power increases with higher γ. For example, the marginal $\bar{R}^2$ of σ_j, $j = 1, 2$, and ρ is almost negligible with lower γ, while it increases sharply with higher γ. As a whole, the explanatory power of the first two moment paramters is always over 90%. Also some non-liner terms add 4∼5% more explanatory power. About 4% of that not explained may come from two sources:

first, the other non-linearities of the policy function, and second, the Monte Carlo simulation error.[24] However, over 95% explanatory power means that linear forms fit well enough to use these results as the approximation of the policy function. Also assuming that the risk aversion of people in general is less than 5, we can say that the first order moment parameters' explanatory power dominates over that of the second order moment paramter's.

Recall that the difference between Model 1 and Models 2 and 3 is the assumption of the marginal distribution. For Models 2 and 3, $\bar{R}^2$s of μ_j, $j = 1, 2$, are still high, about 80% within γ from 1 to 10. However, we can see a bit different pattern. It may seem to be that $\gamma = 1$ is a very special case, therefore we do not discuss it. With this exceptiont, the marginal $\bar{R}^2$ of the first order moment parameters does not monotonically decrease, unlike Model 1. It increases with higher risk aversion from $\gamma = 2$ to 6, then decreases later. As a whole, the marginal $\bar{R}^2$ of σ_j, $j = 1, 2$, and ρ (or δ) increases with higher γ, however not as much as Model 1. The linear information of higher order moments, ν_j, λ_j, and κ_j, $j = 1, 2$, explains only 1~3% of the variation in α_2^*. Of course, when we define all higher moment paramters including σ and ρ as risk factors, the explanatory power of risk factors increases with higher risk aversion, consistant with our thought. However, it is not as much as Model 1. The commnon feature is that the portion of the unexplained sector becomes bigger, compared to Model 1. We believe there are many other non-linear combinations of distributional parameters, which may be important to explain the optimal policy function. Also, simulation errors based on the skewed GT marginal distribution seem to be bigger than the normal marginal distribution, although we do not present any evidence of it in this paper.

[24] As pointed out in Judd (1998) (See p.303), we must be much more careful in evaluating integrals when they are not of direct interest but computed as part of solving an optimization problem. Especially when γ is low, the optimal solutions vary in each simulation trial.



Note: This graph shows the mean of $\bar{R}^2$ of linear projection including each independent variable from bootstrapping. Dependent variable is α_2^*. The bottom sector represents $\bar{R}^2$ of a regression including only a constant and μ_j, $j = 1, 2$ as independent variables. The second bottom sector represents $\bar{R}^2$ of a regression including σ_j, $j = 1, 2$ and ρ (or δ) as well as dependent variables.

Figure II.6 $\bar{R}^2$ based on Independent Variables

II.4.B With a Riskless Asset

II.4.B.a General Results

Now we consider the asset allocation problem in the presence of a riskless asset. Figure II.7 shows the regression results based on equation (II.13) for Model 1 in the presence of a riskless asset. Here, the dependent variable is α_1^*. The existence of a riskless asset changes the magnitude of the partial effect. The magnitude of $\frac{\partial \alpha_1^*}{\partial \mu_1}$ (or $\frac{\partial \alpha_1^*}{\partial \sigma_1}$) is almost more than twice (3 times for $\frac{\partial \alpha_1^*}{\partial \sigma_1}$) bigger than the magnitude of $\frac{\partial \alpha_1^*}{\partial \mu_2}$ (or $\frac{\partial \alpha_1^*}{\partial \sigma_2}$). The magnitude of $\frac{\partial \alpha_2^*}{\partial \mu_2}$ (or $\frac{\partial \alpha_2^*}{\partial \sigma_2}$) is almost more than 4 times (3.5 times for $\frac{\partial \alpha_2^*}{\partial \sigma_2}$) bigger than the magnitude of $\frac{\partial \alpha_2^*}{\partial \mu_1}$ (or $\frac{\partial \alpha_2^*}{\partial \sigma_1}$), although we do not report the result of the case with α_2^* dependent variable. Compared to the MV case, this result may seem to be a bit similar. Recall that the optimal asset allocation with the MV case is shown in equation (II.4). In the case of 2 risky assets, the optimal asset allocation becomes:

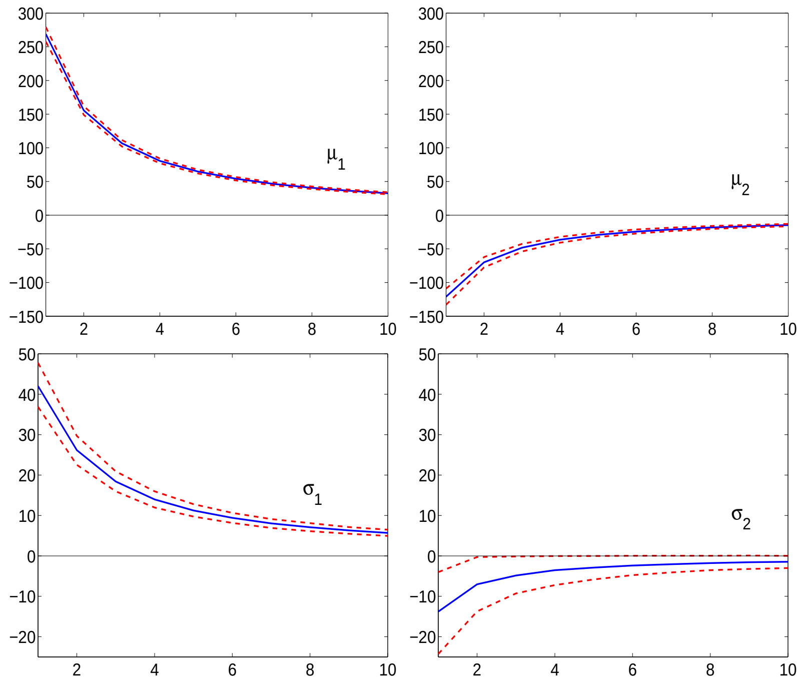
$$\boldsymbol{\alpha}_t^* = \frac{1}{\gamma^{MV} W_t} {\Sigma_{t+T}^{ex}}^{-1} \mu_{t+T}^{ex} = \frac{1}{\gamma^{MV} W_t} \frac{1}{(1-\rho^2)\sigma_1^2\sigma_2^2} \begin{bmatrix} \sigma_2^2\mu_1 - \rho\sigma_1\sigma_2\mu_2 \\ -\rho\sigma_1\sigma_2\mu_1 + \sigma_1^2\mu_2 \end{bmatrix} \tag{II.15}$$

where, $\Sigma_{t+T}^{ex} \equiv \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$ and $\mu_{t+T}^{ex} \equiv \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$. In this case, the magnitude ratio of $\frac{\partial \alpha_1^*}{\partial \mu_1}$ and $\frac{\partial \alpha_1^*}{\partial \mu_2}$ depends on the ratio of σ_2^2 and $\rho\sigma_1\sigma_2$, which is $\frac{\sigma_2}{\rho\sigma_1}$. Given the fact that $\frac{\sigma_2}{\rho\sigma_1} \approx 2$ with this data set, it may imply that the CRRA utility case seems to be similar to be the MV case. We also can extend this interpretation to α_2^* case. However, Figure II.7 shows results which are a bit strange regarding to σ_1 and σ_2. According to Figure II.7, the optimal choice of investors with the CRRA utility function suggests that when the volatility of asset 1 increases, investors should hold more of asset 1. This seems to be counter-intuitive. In the standard portfolio choice problem, we assume that investors dislike risk, so that when volatility (or risk) of a certain asset increases, investors should not hold more of that asset. In the next section, we will discuss about that.

II.4.B.b Higher Risk and Hold More?

In the 2 risky assets' MV case, $\frac{\partial \alpha_1^*}{\partial \sigma_1} = \frac{1}{1-\rho^2} \frac{\rho\sigma_1\mu_2 - 2\sigma_2\mu_1}{\sigma_1^3\sigma_2}$ and $\frac{\partial \alpha_1^*}{\partial \sigma_2} = \frac{1}{1-\rho^2} \frac{\rho\mu_2}{\sigma_1\sigma_2^2}$, assuming $\frac{1}{\gamma^{MV} W_t} = 1$. In general, $\frac{\partial \alpha_1^*}{\partial \sigma_1}$ is negative, unless $\rho\sigma_1\mu_2 - 2\sigma_2\mu_1$ is positive. Also, $\frac{\partial \alpha_1^*}{\partial \sigma_2}$ is positive, assuming that $\rho\mu_2$ is positive.[25] However, this result suggests the opposite interpretation. $\frac{\partial \alpha_1^*}{\partial \sigma_1}$ is positive, and $\frac{\partial \alpha_1^*}{\partial \sigma_2}$ is negative. As the policy function is unknown, it is very difficult to imagine and analyze this phenomenon in the Mean-Variance space. Figure II.8 shows the possible analogue of only the case of two risky assets without a riskless asset. In this case, the weight on asset 1 increases, even when σ_1 increases. However, in the presence of a riskless asset, the case where $\frac{\partial \alpha_1^*}{\partial \sigma_1}$ is positive, $\frac{\partial \alpha_1^*}{\partial \sigma_2}$ is negative, so that the weight

[25] In general, the case of $\rho\sigma_1\mu_2 - 2\sigma_2\mu_1 > 0$ is a very rare occurance in the real world, as $-1 < \rho < 1$. The case of $\rho\mu_2 < 0$ may happen. However, given the data set of this paper, $\rho > 0$. Therefore we may say that the possibility of $\rho\mu_2 < 0$ can not be too high.



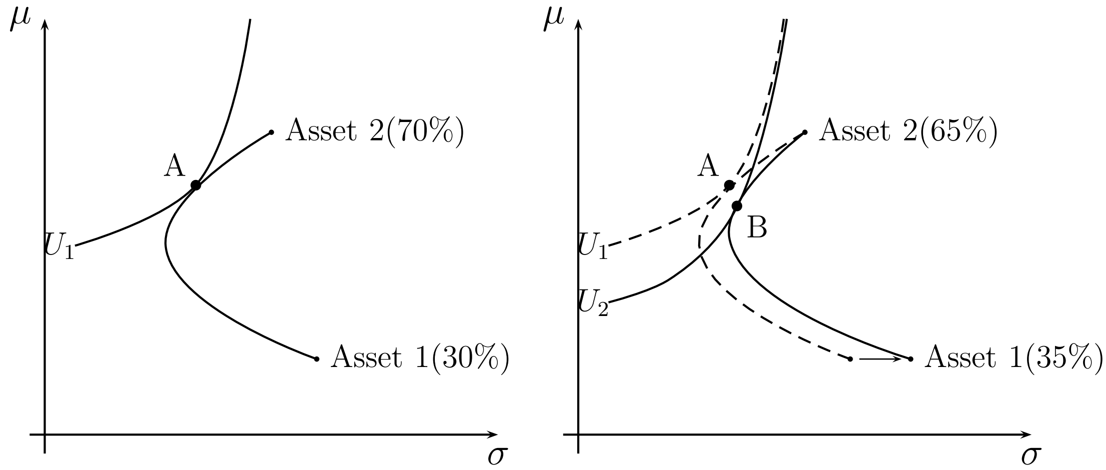
Note: This graph shows the coefficient estimates based on the equation (II.13) with Model 1, depending on different risk aversion level, γ. Dependent variable is α_1^*, the Emerging Market Index. Dotted lines are 95% confidence intervals.

Figure II.7 Coefficients on the Optimal Decision w/ Riskless Asset of Model 1

on a riskless asset increases with any increase in σ_1 is very hard to resemble in the Mean-Variance space. Therefore, positive $\dfrac{\partial\alpha_1^*}{\partial\sigma_1}$ and negative $\dfrac{\partial\alpha_1^*}{\partial\sigma_2}$ (or negative $\dfrac{\partial\alpha_2^*}{\partial\sigma_1}$ and positive $\dfrac{\partial\alpha_2^*}{\partial\sigma_2}$) of the CRRA utility function seems to be very unique and different from the case with the MV utility function. This discussion will extend to higher order moment parameters in the next section.

II.4.B.c Higher Order Moment Parameters

Table II.5 summarizes the coefficient estimates of higher order moment parameters based on equation (II.13). In the presence of a riskless asset, the magnitude of $\dfrac{\partial\alpha_1^*}{\partial\lambda_1}$ (or $\dfrac{\partial\alpha_2^*}{\partial\lambda_2}$) is bigger than $\dfrac{\partial\alpha_1^*}{\partial\lambda_2}$ (or $\dfrac{\partial\alpha_2^*}{\partial\lambda_1}$) and also significant in almost all γ, although it is not shown in Table II.5. This difference in the magnitude can be explained by the existence of a riskless asset. Suppose $\dfrac{\partial\alpha_1^*}{\partial\lambda_1}$ is positive, implying that when λ_1 increases, then the weight on asset 1, α_1^*, increases. However, this



Note: This graph shows the possibility of positive $\frac{\partial \alpha_1^*}{\partial \sigma_1}$ and negative $\frac{\partial \alpha_1^*}{\partial \sigma_2}$ in Mean-Variance Space in the absence of a riskless asset. When σ_1 increases, the weight on asset 1 increases. However, we have to notice that $U_1 > U_2$

Figure II.8 Positive $\frac{\partial \alpha_1^*}{\partial \sigma_1}$ and Negative $\frac{\partial \alpha_1^*}{\partial \sigma_2}$ in Mean-Variance Space without a Riskless Asset

increase in α_1^* comes not only from the weight of asset 2 but also from the weight of a riskless asset. Therefore, the magnitude of $\frac{\partial \alpha_1^*}{\partial \lambda_1}$ (or $\frac{\partial \alpha_2^*}{\partial \lambda_2}$) and $\frac{\partial \alpha_1^*}{\partial \lambda_2}$ (or $\frac{\partial \alpha_2^*}{\partial \lambda_1}$) is different. From Table II.5, we can see that positive and significant $\frac{\partial \alpha_1^*}{\partial \lambda_1}$ and $\frac{\partial \alpha_2^*}{\partial \lambda_2}$ may be used for indirect evidence that people with the CRRA utility function prefer positive skewness. Of course, this may not be true as σ_j, $j = 1, 2$ case. Therefore we may need further study on this issue.

The estimation results on other higher order moment parameters, μs and κs, show the mixed results. First, for Model 2, ν_2 is significant, while ν_1 and κ_2 are not significant at all. Model 3 also shows that only κ_1 on α_1^* and ν_2 on α_2^* are significant. Of course, it is hard to extend this result globally, because our space for sampling parameters is bounded based on the Regime-Switching estimation. However, this may suggest a simple and robust procedure for portfolio choice. The information about tailness and peak of distribution may not be very important. Another thing to notice is that the sign of $\frac{\partial \alpha_2^*}{\partial \nu_2}$ seems to be counter-intuitive as

$\dfrac{\partial \alpha_1^*}{\partial \sigma_1}$ and $\dfrac{\partial \alpha_2^*}{\partial \sigma_2}$. In our study, significant $\dfrac{\partial \alpha_2^*}{\partial \nu_2} < 0$ and $\dfrac{\partial \alpha_1^*}{\partial \nu_2} > 0$ without a riskless asset, and $\dfrac{\partial \alpha_2^*}{\partial \nu_2} > 0$ with it as shown in Figure II.4. $\dfrac{\partial \alpha_2^*}{\partial \nu_2} < 0$ means that when fat-tailness of asset 2 decreases (ν_2 increases), the CRRA investors decrease the weight on asset 2, α_2^*. Based on the general consensus that people may dislike fat-tailness, this seems to be counter-intuitive as the variance parameter case.

Table II.5 Summary of Partial Effect of Higher Order Moment Parameters

Model	Dep.	Indep.	Model 1	Model 2		Model 3	
Equation (II.13)	α_1^*	ν_1	N.A.	+ (9)	Insig (10)	- (9)	Insig (10)
		ν_2	N.A.	+ (10)	Sig (9)	+ (10)	Insig (10)
		λ_1	N.A.	+ (10)	Sig (9)	+ (10)	Sig (6)
		λ_2	N.A.	- (10)	Sig (9)	- (8)	Insig (10)
		κ_1	N.A.	- (10)	Sig (9)	- (10)	Sig (9)
		κ_2	N.A.	+ (5)	Insig (10)	- (10)	Insig (10)
	α_2^*	ν_1	N.A.	- (10)	Insig (10)	+ (10)	Insig (10)
		ν_2	N.A.	- (10)	Sig (9)	- (10)	Sig (9)
		λ_1	N.A.	- (9)	Insig (10)	+ (10)	Insig (10)
		λ_2	N.A.	+ (10)	Sig (10)	+ (10)	Sig (10)
		κ_1	N.A.	+ (10)	Insig (6)	- (10)	Insig (10)
		κ_2	N.A.	- (10)	Insig (10)	- (6)	Insig (10)

Note: +/- represent the dominant signs of coefficients on each distributional parameter. Sig/Insig represent the dominant significance on each distributional parameter. The number in parentheses represents the number of certain sign or significance within $\gamma = 1$ to 10. For example, '+ (9) and Insig (10)' means that there are 9 positive coefficients out of 10 and 10 out of 10 cases are statistically insignificant.

II.4.B.d Dependence Structure and Nonlinearity

Table II.6 summarizes the linear regression results on the dependence parameter based on equation (II.13) and on nonlinear terms based on equation (II.14). First, the coefficient for the dependence parameter is not significant except for one. This result may suggest an important message about the portfolio choice problem regarding the role of the dependence structure, which may not important. Although many of them are not significant, the signs of the dependence parameter on Models 2 and 3 are positive, implying that more risky assets are held when dependence is high. The MV utility case, $\dfrac{\partial \boldsymbol{\alpha}^*}{\partial \rho}$ is negative, assuming a general condition. This means that people with MV utility function tend to hold fewer risk assets when ρ increases. However, it is not true with CRRA utility function, reflecting a different decision rule of CRRA utility investors.

Regarding non-linear terms, we find mixed results. However, notice that $\dfrac{\mu_1}{\sigma_1}$ on α_1^* and $\dfrac{\mu_2}{\sigma_2}$ on α_2^* are significant. Remember that the existence of a riskless asset makes the partial effect of an other asset's distributional parameter smaller. Therefore, significant $\frac{\mu_1}{\sigma_1}$ on α_1^* and $\frac{\mu_2}{\sigma_2}$ on α_2^* still indicate that there is the level effect of first and second order moment parameters.

Table II.6 Summary of Partial Effect of Dependence Parameter and Nonlinear Terms

Model	Dep.	Indep.	Model 1		Model 2		Model 3	
Eq (II.13)	α_1^*	ρ (or δ)	- (10)	Insig (10)	+ (10)	Insig (10)	+ (10)	Insig (10)
	α_2^*	ρ (or δ)	+ (10)	Insig (10)	+ (5)	Insig (10)	+ (10)	Sig (10)
Eq (II.14)	α_1^*	$\dfrac{\mu_1}{\sigma_1}$	+ (10)	Sig (10)	+ (10)	Sig (10)	+ (10)	Sig (10)
		$\dfrac{\mu_2}{\sigma_2}$	- (10)	Sig (10)	- (9)	Insig (10)	- (7)	Insig (10)
		$\rho\sigma_1\sigma_2$	+ (10)	Insig (10)	- (7)	Insig (10)	- (8)	Insig (10)
	α_2^*	$\dfrac{\mu_1}{\sigma_1}$	- (10)	Sig (10)	- (9)	Insig (10)	- (9)	Insig (10)
		$\dfrac{\mu_2}{\sigma_2}$	+ (10)	Sig (10)	+ (10)	Sig (10)	+ (10)	Sig (10)
		$\rho\sigma_1\sigma_2$	+ (10)	Sig (10)	- (9)	Insig (10)	- (10)	Insig (10)

Note: +/- represent the dominant signs of coefficients on each distributional parameter. Sig/Insig represent the dominant significance on each distributional parameter. The number in parentheses represents the number of certain sign or significance within $\gamma = 1$ to 10. For example, '+ (9) and Insig (10)' means that there are 9 positive coefficients out of 10 and 10 out of 10 cases are statistically insignificant.

II.4.B.e Explanatory Power

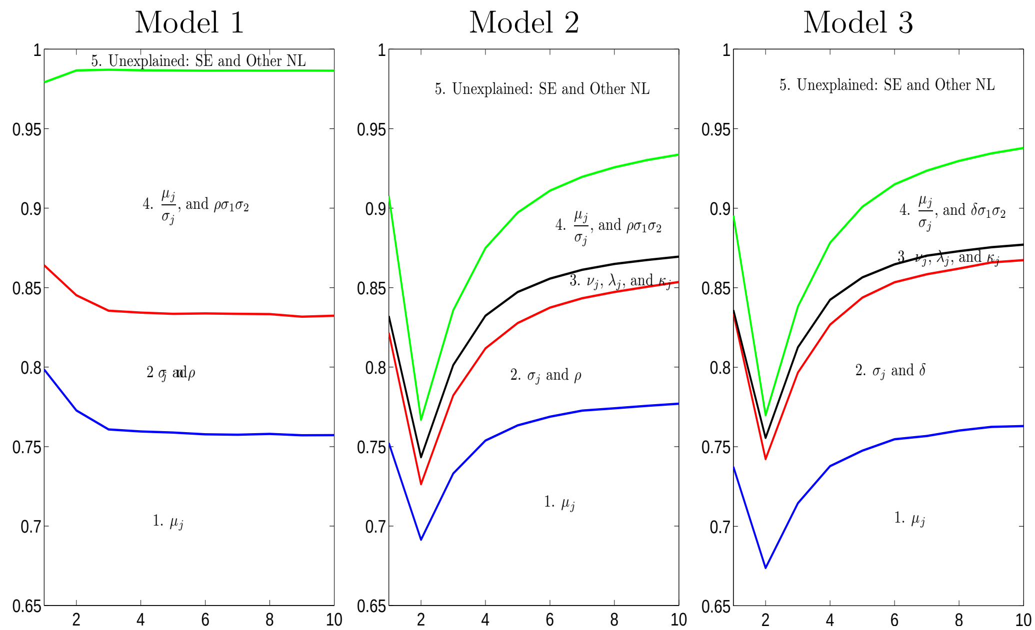
Figure II.9 shows the $\bar{R}^2$ of linear projection in various settings. In the existence of a riskless asset, the first order moment parameters, μ_j, $j = 1, 2$, again have the greatest explanatory power. On average, 70∼80% of variation in the

CRRA investor's decision is explained by first order moment parameters. For α_2^*, the explanatory power of first order moment is 65$\sim$80%. (Not reported here.) One interesting result is that the $\bar{R}^2$ of the linear model including non-linear terms is almost 98.5$\sim$99.5%. If we consider that simulation errors still exist, this result seems to almost perfectly resemble the CRRA investor's portfolio choice decision. The portfolio choice problem of the MV investor without a riskless asset has one more constraint than with a riskless asset. Therefore, the MV investor's decision rule without a riskless asset shown in equation (II.3) seems to be complicated compared to equation (II.4). As an analogue of the MV investor's decision rule shown in the equations (II.3) and (II.4), we may think that the number of the factors in a CRRA investor's portfolio choice problem is smaller in the presence of a riskless asset than in the absence of it. This is another reason to include $\frac{\mu_1}{\sigma_1}$, $\frac{\mu_2}{\sigma_2}$, and $\rho\sigma_1\sigma_2$ as independent variables. Based on this reasoning, we may say that the top portion in Models 2 and 3, which is unexplained, mainly comes from the non-linearity of higher order moment parameters rather that that of first and second order moment parameters.

The linear information of higher order moments from ν_j, λ_j, and κ_j, $j = 1, 2$, explains only less than 2%. Assuming that the portions 3 and 5 are due to the higher order moment parameters, their explanatory power is from 8% to at most 25%, depending on γ, in the CRRA investor's decision. In the next section, we will compare various portfolio choice rules' out-of-sample performance to check whether or not higher moment information helps to improve performance.

II.5 Out-of-Sample Performance

Once we have the linear approximated policy function, we can easily apply these results to the portfolio choice problem when the parameters characterizing investment opportunities vary over time. Also, we can eliminate a lot of the computational burden using the liner approximated policy function. Another good



Note: This graph shows the mean of $\bar{R}^2$ of linear projection including each independent variable from bootstrapping. Dependent variable is α_1^*. The bottom sector represents $\bar{R}^2$ of a regression including only a constant and μ_j, $j = 1, 2$ as independent variables. The second bottom sector represents $\bar{R}^2$ of a regression including σ_j, $j = 1, 2$ and ρ (or δ) as well as dependent variables.

Figure II.9 $\bar{R}^2$ based on Independent Variables

feature of the linear approximation is that we can check which elements of joint distribution play an important role and which elements help to improve the portfolio performance. In this section, we will discuss the possible dynamics of parameters to apply the conditional portfolio choice problem first. Then we will provide the results of the portfolio performance comparison. Finally, we will provide the possible connection between the MV utility and the CRRA utility function.

II.5.A Dynamics of Parameters

For model 1, first, we focus on the model with time-varying expected returns and a covariance matrix. One possible model considering time-varying first and second moment parameters is the Dynamic Conditional Correlation Multivariate GARCH model (DCC-MVGARCH), proposed by Engle and Sheppard (2001) and Engle (2002). In the DCC-MVGARCH model, univariate GARCH models are estimated for each series, then a time varying conditional correlation matrix is estimated using a simple specification. We assume AR(1) process for each univariate series. In many criteria such as AIC or BIC information criterion or the likelihood ratio test, AR(1) seems to be the best-fit model. Similar to equation (II.8), AR(1) process describes the conditional mean equation to capture the possible first-order serial correlation as follows:

$$\mathbf{r}_{N,t} = \boldsymbol{\mu}_N + \mathbf{A}\mathbf{r}_{N,t-1} + \mathbf{u}_t \tag{II.16}$$

where $\mathbf{A}$ is a diagonal matrix and the elements are $a_j,\ j = 1, 2, \cdots, N$[26]. $\boldsymbol{\mu}_N$ is a $n \times 1$ vector of means. Following the DCC-MVGARCH model, the error term is modeled as:

$$\begin{gathered} \mathbf{u}_t | \mathcal{F}_{t-1} \sim N(0, H_t) \\ H_t \equiv D_t R_t D_t \end{gathered} \tag{II.17}$$

where D_t is the $k \times k$ diagonal matrix of time varying standard deviations from univariate GARCH models with $\sqrt{h_{it}}$ on the ith diagonal, and R_t is the time varying correlation matrix. The proposed dynamic correlation structure is:

$$\begin{gathered} Q_t = \left(1 - \sum_{m=1}^{M} \alpha_m - \sum_{n=1}^{N} \beta_n\right) \overline{Q} + \sum_{m=1}^{M} \alpha_m (\epsilon_{t-m} \epsilon'_{t-m}) + \sum_{n=1}^{N} \beta_n Q_{t-n} \\ R_t = Q_t^{*-1} Q_t Q_t^{*-1} \end{gathered} \tag{II.18}$$

where $\overline{Q}$ is the unconditional covariance of the standardized residuals resulting from the first stage estimation, and Q_t^* is a diagonal matrix composed of the square root of the diagonal elements of Q_t.

For Models 2 and 3, there are 3 more parameters for each margin. Time variation in higher moments has been analyzed in Hansen (1994), and Harvey and Siddique (1999). In this paper, we use recent contributions by Jondeau and Rockinger (2003), and Jondeau and Rockinger (2005). Here, we only consider the first lag of error term as the time variation force, therefore each parameter is

[26] When we take $\mathbf{A}$ also as a regime switching variable, $\mathbf{A}_{s_t}$, there is no considerable gain in terms of likelihood value. Therefore we consider $\mathbf{A}$ as not switching with regimes.

modeled as:

$$\begin{aligned}
\frac{1}{\nu_{i,t}} &= \frac{1}{\nu_{i,t-1}} + \alpha_\nu |u_{i,t-1}| \\
\lambda_{i,t} &= \lambda_{i,t-1} + \alpha_\lambda u_{i,t-1} \\
\frac{1}{\kappa_{i,t}} &= \frac{1}{\kappa_{i,t-1}} + \alpha_\kappa |u_{i,t-1}|
\end{aligned} \tag{II.19}$$

for $i = 1, 2$. In the estimation, we do not use any function to restrict the values of each parameter, but we estimate parameters by adding time-varying parameters one by one. Table II.7 reports the likelihood ratio test result for various models. We find that considering time-varying variance, GARCH(1,1), is very helpful to capture the dynamics of distribution, but the time-variation of higher moments, $\nu_{i,t}$, $\lambda_{i,t}$, and $\kappa_{i,t}$, $i = 1, 2$ are not useful for both margins. Therefore, we use the time-varying variance and constant higher moment parameters for models 2 and 3, that is, AR(1)-GARCH(1,1)-Constant ν_i, λ_i, and κ_i, $i = 1, 2$. As shown in a previous section, first and second order moment parameters are the most important factors and the importance of other higher moment information may not very great. Therefore, this model selection conforms with the results from a previous section.

Once we select the marginal distribution model, the next step is to ensure that it cooperates with the DCC model. In order to do that, we need two transformations based on the copula model. In equation (II.16), $\hat{\mathbf{u}}_t$ is converted to get uniformly distributed variables using the skewed GT CDF, $F_{skewedGT}(u_i)$. Then we convert again to get standard normally distributed variables, $\Phi^{(-1)}(F_{skewedGT}(u_i))$. From the full sample estimation not shown in this paper, we can see that there might be a structural break in the correlation structure between the EM index and the S&P 500 index. At roughly 2000, the correlation between the two indices jumps dramatically. In this paper, the structural break in the parameters is not a central issue. Therefore, we consider the out-of-sample performance comparison only before 2000. Once we eliminate the latest data and only include the data before 2000, the persistence of the DCC model, captured by β_1 in equation (II.18),

Table II.7 Likelihood Ratio Test Results for Constant and Time-varying Models

		LL	d.f.	Stat.
EM Index				
Constant: $\sqrt{h_1}$, ν_1, λ_1, κ_1	Time-Varying: NO	676.78		
Constant: ν_1, λ_1, κ_1	Time-Varying: $h_{1,t}$	699.69	1	45.82**
Constant: λ_1, κ_1	Time-Varying: $h_{1,t}$, $\nu_{1,t}$	700.39	1	1.40
Constant: κ_1	Time-Varying: $h_{1,t}$, $\nu_{1,t}$, $\lambda_{1,t}$	700.40	1	0.02
Constant: NO	Time-Varying: $h_{1,t}$, $\nu_{1,t}$, $\lambda_{1,t}$, $\kappa_{1,t}$	700.47	1	0.15
S&P 500 Index				
Constant: $\sqrt{h_2}$, ν_2, λ_2, κ_2	Time-Varying: NO	761.38		
Constant: ν_2, λ_2, κ_2	Time-Varying: $h_{2,t}$	769.95	1	17.15**
Constant: λ_2, κ_2	Time-Varying: $h_{2,t}$, $\nu_{2,t}$	769.97	1	0.05
Constant: κ_2	Time-Varying: $h_{2,t}$, $\nu_{2,t}$, $\lambda_{2,t}$	770.64	1	1.34
Constant: NO	Time-Varying: $h_{2,t}$, $\nu_{2,t}$, $\lambda_{2,t}$, $\kappa_{2,t}$	770.65	1	0.01

Note: 1% level critical value for $\chi^2(1) = 6.63$. 5% level critical value for $\chi^2(1) = 3.84$.
** denotes 1% significant level and * denotes 5% significant level.

becomes small, almost zero for both models. The likelihood ratio statistic of Model 1 for the constant correlation matrix, $\bar{R}$, is 4.0253, and the statistic of Model 2 is 5.0601, however the 5% level critical value for $\chi^2(2)$ is 5.99. Therefore we can reject the time-varying correlation model.

Based on this result, we use AR(1)-GARCH(1,1)-Constant other variables for the out-of-sample performance comparison for both models.[27] Also, in order to account for estimation errors, we apply the one of Bayesian approaches,[28] Griddy Gibbs Sampling (GGS) method, proposed by Ritter and Tanner (1992), Bauwens et al. (1999) is used. GGS is a very useful method when we do not have any idea what the prior distributions are. In this case, we assume that we do not have any priors of higher order moment parameters and GARCH parameters, therefore we

[27] Of course, we can use AR(1)-GARCH(1,1)-Time-varying other variables model. However, from section II.4, we saw that some of the higher moment parameters are not significant, and the variation in the higher moment parameters is very small even with a time-varying Model. Therefore it is quite reasonable to use AR(1)-GARCH(1,1)-Constant other variables model.

[28] Considerable effort has been devoted to the issue of handling estimation error with the goal of improving the performance of asset allocation models. See DeMiguel et al. (2006).

use flat priors for many parameters except two: ν_j and κ_j. For those, we use the Half-Cauchy prior, such as $(1-\nu^2)^{-1}$, following the suggestion by Bauwens et al. (1999). Details of GGS is shown in Appendix Apendix B. Table II.8 shows the estimation results of Griddy Gibbs-Sampling.

Table II.8 In-sample Posterior Results using GGS

	Model 1			Model 2 and Model 3		
	Mean	95% C.I.	Prior Interval	Mean	95% C.I.	Prior Interval
μ_1	0.007	[0.002 0.013]	(-0.030 0.030)	0.011	[-0.004 0.027]	(-0.030 0.060)
a_1	0.194	[0.066 0.323]	(-0.200 0.500)	0.246	[0.055 0.405]	(-0.100 0.700)
ω_{1,h_t}	0.001	[0.000 0.001]	(0.000 0.003)	0.010	[0.000 0.050]	(0.000 0.050)
α_{1,h_t}	0.283	[0.044 0.521]	(0.000 0.850)	0.557	[0.057 0.975]	(0.000 0.990)
β_{1,h_t}	0.641	[0.346 0.913]	(0.005 0.990)	0.140	[0.002 0.543]	(0.000 0.990)
ν_1				7.763	[2.114 5.964]	(2.010 200.0)
λ_1				-0.053	[-0.298 0.160]	(-0.800 0.800)
κ_1				1.738	[0.173 3.353]	(0.100 50.00)
μ_2	0.005	[-0.001 0.011]	(-0.020 0.030)	0.012	[-0.016 0.042]	(-0.030 0.060)
a_2	0.017	[-0.107 0.142]	(-0.250 0.350)	-0.005	[-0.213 0.208]	(-0.500 0.600)
ω_{2,h_t}	0.001	[0.000 0.002]	(0.000 0.004)	0.010	[0.000 0.049]	(0.000 0.050)
α_{2,h_t}	0.078	[0.006 0.202]	(0.000 0.700)	0.367	[0.011 0.969]	(0.000 0.990)
β_{2,h_t}	0.389	[0.016 0.858]	(0.000 0.990)	0.187	[0.003 0.592]	(0.000 0.990)
ν_2				4.008	[2.054 3.933]	(2.010 100.0)
λ_2				-0.103	[-0.539 0.346]	(-0.700 0.700)
κ_2				2.591	[0.233 4.950]	(0.100 75.00)
ρ	0.191	[0.069 0.309]	(-0.100 0.500)	0.243	[0.024 0.815]	(-0.300 0.900)
δ				0.418	[0.020 2.518]	(0.010 5.000)

Note: These are the posterior results of Griddy Gibbs-Sampling for each model in the in-sample (Jan. 67 - Dec. 90). For the posteriors, we report posterior means, 95% confidence intervals in square brackets, and prior intervals in parentheses. We use the flat priors for most cases. For ν_j and κ_j, $j = 1, 2$, we use the half-Cauchy prior. The number of draws is 1000 after discarding 100 initial draws. For Models 2 and 3, we use the same marginal distribution, therefore we only report δ for Model 3.

II.5.B Performance Comparison

For buy-and-hold investors, the optimal asset allocation is the function of expected moments. Since moments are the function of parameters, we can say that the decision is the function of expected value of the parameters, $\boldsymbol{\alpha}^*_{t+T}(E(\boldsymbol{\theta}_{t+T}|Z_t))$. Estimation results represent low persistency in expected means and relatively high persistency in volatility. Therefore, the expected return is very important to short term buy-and-hold investors. Of course, the expected return is important to long term investors as well. However, since the expected volatility has longer persistence in our sample, after the expected return converges to the steady-state value in the very short term, the level of expected volatility becomes the most important factor. That creates different optimal asset allocation decisions between long and short term investors. Given the estimates of the selected model, we can forecast the expected values of parameters. Since we assume the AR(1) process for the mean equation, the expected value of return at time $t+T$ is

$$E_t[\mathbf{r}_{t+T}|Z_t] = \sum_{i=0}^{T-1} \mathbf{A}^i \boldsymbol{\mu}_N + A^T \mathbf{r}_t, \quad T = 1, 2, \cdots . \tag{II.20}$$

The T-step ahead forecast of a GARCH(1,1) for ith asset is given by:

$$h_{i,t+1} = \omega_{i,h_t} + \alpha_{i,h_t} u_{i,t}^2 + \beta_{i,h_t} h_{i,t}$$

$$h_{i,t+T} = \omega_{i,h_t} \textstyle\sum_{i=0}^{T-2} (\alpha_{i,h_t} + \beta_{i,h_t})^i + (\alpha_{i,h_t} + \beta_{i,h_t})^{T-1} h_{t+1}, \quad T = 2, 3, \cdots . \tag{II.21}$$

Based on these forecasting equations and results from Table II.8, we can compare out-of-sample performances of various portfolio choice models. We use mainly three sets of comparisons models. The first set of comparison models uses full information, that is done by the Monte Carlo simulation. The second and third set of comparison models uses the linear approximation results of the previous section. The first usefulness of the linearly approximated decision rule is that it does not need to rely on computational power, although recently computational cost has decreased dramatically. The second useful aspect is that we can include and exclude each moment's linear and non-linear information on the portfolio choice problem, depending on how useful it is in terms of performance or investor's preference. For example, if we believe that only the sum of portions 1, 2 and 4 in Figures II.5 and II.9 is useful, then we can use only that, excluding other information in a general portfolio choice problem. The linear approximation results of the previous section make it possible. In this paper, the second set uses the portions 1 through 5 in Figures II.5 and II.9, excluding simulation error and all non-linear information of moment parameters except $\frac{\mu_j}{\sigma_j}$, $j = 1, 2$ and $\rho\sigma_1\sigma 2$. The last set excludes the higher order moment information, shown in portion 3, therefore, it only uses the linear information of the first and second order moment information, the sum of portions 1, 2, and 4 in Figures II.5 and II.9. This is only applied to Models 2 and 3. In addition, we include a naive decision (1/N) rule and the Mean-Variance utility case for performance comparison.[29]

II.5.B.a Sharpe Ratio

First, we use the Sharpe Ratio (SR), which is the most common performance measure of evaluation, although it may not be an appropriate measure with the CRRA utility function. We measure the out-of-sample Sharpe ratio of each decision rule d as the sample mean of out-of-sample excess returns (over the riskless asset), $\hat{\mu}_{ex,d}$, divided by their sample standard deviation, $\hat{\sigma}_{ex,d}$:

$$\hat{SR}_d = \frac{\hat{\mu}_{ex,d}}{\hat{\sigma}_{ex,d}} \tag{II.22}$$

First, we report the Sharpe ratios of each index and a naive decision (1/N) rule for comparison purposes. Table II.9 shows those results. It is interesting that during the out-of-sample period, the Emerging Market index performs very poorly; the Sharpe ratio is negative. This might be explained by a series of financial crises in the Emerging Markets during the 1990s. The poor performance of the EM index causes the poor performance of the naive decision rule.

[29] Of course, the risk aversion of the MV utility function and power utility function is different, therefore they can not map on to each other directly. However in order to compare their performances depending on the risk aversion level, we just set γ^{MV} equal to 1 through 10.

Table II.9 Statistics of Out-of-Sample Performance

	Mean	Std	Sharpe Ratio
EM Index	0.0032	0.0616	**-0.0250**
S&P 500 Index	0.0124	0.0385	**0.1964**
Naive (1/2) w/o RF	0.0078	0.0438	**0.0688**
Naive (1/3) w/ RF	0.0068	0.0292	**0.0690**

Note: This table presents the means, standard deviations and Sharpe ratios of out-of-sample returns of each index and a naive decision (1/N) rule for the last 10 years (from Jan. 1991 to Dec. 2000).

Table II.10 shows sample means, standard deviations, and the Sharpe ratios of various settings in the absence of a riskless asset. According to Table II.10, MV allocation seems to be the best in terms of the Sharpe ratio. However, its expected mean and standard deviation are extremely small. In other words, MV strategies produce relatively conservative portfolio choice rules compared to CRRA strategies, therefore, in the absence of a riskless asset, MV stategies seem to have a higher Sharpe ratio. However, they do not perform well in terms of realized wealth. This is discussed in more detail later in the paper.

When we compare the full models (M1 full, M2 full, and M3 full), Model 1 performs better than others. This is because when we consider higher order moment parameters to capture characteristics of distribution well, the optimal decision rules lead to holding more of the Emerging Market Index, at least in my sample. As mentioned, the performance of the Emerging Market Index during 1990s is very poor. Therefore, the Sharpe ratios of M2 full and M3 full are lower. Of course, the skewness and kurtosis information does not help to improve the Sharpe ratio measure, because it only considers the mean and volatility of return distribution. Therefore it is not reasonable to say that MV strategies perform better than CRRA strategies.

The comparison of full, linear, and linear without higher order moment information cases shows the best performance of a full case. The decision rules of the linear model tend to put even more weight on the EM index than full cases,

Table II.10 Statistics of Out-of-Sample Performance w/o a Riskless Asset

	γ	1	2	3	4	5	6	7	8	9	10
	Mean	0.0095	0.0088	0.0085	0.0083	0.0081	0.0080	0.0080	0.0079	0.0079	0.0078
MV	Std	0.0068	0.0042	0.0033	0.0028	0.0025	0.0022	0.0021	0.0020	0.0019	0.0018
	SR	**0.6956**	**0.9707**	**1.1380**	**1.2657**	**1.3693**	**1.4552**	**1.5268**	**1.5870**	**1.6377**	**1.6805**
	Mean	0.0350	0.0231	0.0185	0.0163	0.0150	0.0141	0.0135	0.0130	0.0127	0.0124
M1 Full	Std	0.1362	0.0840	0.0664	0.0582	0.0537	0.0510	0.0491	0.0479	0.0469	0.0461
	SR	**0.2216**	**0.2182**	**0.2071**	**0.1972**	**0.1901**	**0.1832**	**0.1771**	**0.1719**	**0.1680**	**0.1644**
	Mean	0.0388	0.0256	0.0205	0.0179	0.0164	0.0153	0.0145	0.0140	0.0135	0.0132
M1 Linear	Std	0.1610	0.1008	0.0770	0.0657	0.0591	0.0551	0.0524	0.0505	0.0490	0.0479
	SR	**0.2111**	**0.2069**	**0.2044**	**0.1998**	**0.1956**	**0.1909**	**0.1858**	**0.1816**	**0.1781**	**0.1750**
	Mean	0.0128	0.0107	0.0099	0.0094	0.0090	0.0087	0.0084	0.0081	0.0078	0.0076
M2 Full	Std	0.0681	0.0544	0.0500	0.0479	0.0468	0.0462	0.0459	0.0457	0.0457	0.0458
	SR	**0.1169**	**0.1079**	**0.1032**	**0.0960**	**0.0897**	**0.0837**	**0.0780**	**0.0724**	**0.0669**	**0.0618**
	Mean	-0.1261	-0.0268	-0.0102	-0.0125	-0.0115	-0.0094	-0.0073	-0.0066	-0.0045	-0.0040
M2 Linear	Std	0.8819	0.2873	0.1808	0.1840	0.1717	0.1552	0.1399	0.1330	0.1194	0.1146
	SR	**-0.1484**	**-0.1099**	**-0.0829**	**-0.0939**	**-0.0950**	**-0.0916**	**-0.0866**	**-0.0854**	**-0.0779**	**-0.0764**
	Mean	-0.0672	-0.0002	0.0039	-0.0010	-0.0020	-0.0010	0.0001	0.0002	0.0020	0.0019
M2 L noH	Std	0.5398	0.1385	0.1059	0.1197	0.1186	0.1087	0.0993	0.0963	0.0850	0.0836
	SR	**-0.1333**	**-0.0359**	**-0.0083**	**-0.0480**	**-0.0572**	**-0.0534**	**-0.0468**	**-0.0476**	**-0.0328**	**-0.0344**
	Mean	0.0129	0.0102	0.0098	0.0089	0.0081	0.0073	0.0065	0.0060	0.0055	0.0052
M3 Full	Std	0.0706	0.0551	0.0511	0.0490	0.0484	0.0489	0.0501	0.0516	0.0533	0.0547
	SR	**0.1149**	**0.0978**	**0.0983**	**0.0848**	**0.0682**	**0.0507**	**0.0351**	**0.0226**	**0.0135**	**0.0070**
	Mean	-0.1195	-0.0200	-0.0090	-0.0118	-0.0114	-0.0091	-0.0068	-0.0057	-0.0032	-0.0033
M3 Linear	Std	0.8469	0.2496	0.1759	0.1799	0.1709	0.1533	0.1373	0.1286	0.1128	0.1110
	SR	**-0.1468**	**-0.0993**	**-0.0783**	**-0.0920**	**-0.0946**	**-0.0904**	**-0.0845**	**-0.0818**	**-0.0711**	**-0.0724**
	Mean	-0.0672	-0.0002	0.0039	-0.0010	-0.0020	-0.0010	0.0001	0.0002	0.0020	0.0019
M3 L noH	Std	0.5398	0.1385	0.1059	0.1197	0.1186	0.1087	0.0993	0.0963	0.0850	0.0836
	SR	**-0.1333**	**-0.0359**	**-0.0083**	**-0.0480**	**-0.0572**	**-0.0534**	**-0.0468**	**-0.0476**	**-0.0328**	**-0.0344**

Note: This table presents the means, standard deviations and Sharpe ratios of out-of-sample returns of various models without a riskless asset for the last 10 years (from Jan. 1991 to Dec. 2000).

Table II.11 Statistics of Out-of-Sample Performance w/ a Riskless Asset

	γ	1	2	3	4	5	6	7	8	9	10
	Mean	0.0285	0.0159	0.0122	0.0104	0.0093	0.0086	0.0081	0.0077	0.0074	0.0071
MV	Std	0.1407	0.0610	0.0399	0.0299	0.0239	0.0200	0.0172	0.0151	0.0134	0.0121
	SR	**0.1684**	**0.1813**	**0.1852**	**0.1874**	**0.1887**	**0.1897**	**0.1904**	**0.1909**	**0.1914**	**0.1918**
	Mean	0.0419	0.0253	0.0186	0.0151	0.0131	0.0120	0.0111	0.0102	0.0098	0.0093
M1 Full	Std	0.1985	0.1103	0.0743	0.0554	0.0443	0.0363	0.0310	0.0269	0.0242	0.0218
	SR	**0.1872**	**0.1861**	**0.1855**	**0.1862**	**0.1886**	**0.1990**	**0.2028**	**0.2026**	**0.2068**	**0.2079**
	Mean	0.0455	0.0286	0.0211	0.0171	0.0147	0.0131	0.0119	0.0110	0.0103	0.0098
M1 Linear	Std	0.2010	0.1163	0.0811	0.0614	0.0493	0.0411	0.0353	0.0310	0.0275	0.0248
	SR	**0.2024**	**0.2044**	**0.2013**	**0.2012**	**0.2012**	**0.2020**	**0.2013**	**0.2012**	**0.2015**	**0.2016**
	Mean	0.0119	0.0082	0.0072	0.0066	0.0063	0.0061	0.0059	0.0058	0.0057	0.0056
M2 Full	Std	0.0529	0.0275	0.0185	0.0133	0.0110	0.0095	0.0081	0.0071	0.0063	0.0056
	SR	**0.1338**	**0.1240**	**0.1305**	**0.1352**	**0.1326**	**0.1386**	**0.1392**	**0.1421**	**0.1422**	**0.1389**
	Mean	-0.0854	-0.0458	-0.0390	-0.0314	-0.0260	-0.0197	-0.0163	-0.0132	-0.0115	-0.0104
M2 Linear	Std	0.9373	0.4447	0.3279	0.2376	0.2037	0.1626	0.1409	0.1229	0.1086	0.1045
	SR	**-0.0962**	**-0.1139**	**-0.1334**	**-0.1522**	**-0.1509**	**-0.1508**	**-0.1494**	**-0.1461**	**-0.1497**	**-0.1453**
	Mean	-0.0357	-0.0212	-0.0164	-0.0115	-0.0081	-0.0050	-0.0035	-0.0019	-0.0014	-0.0011
M2 L noH	Std	0.5292	0.2469	0.1815	0.1246	0.1050	0.0804	0.0692	0.0599	0.0529	0.0514
	SR	**-0.0764**	**-0.1054**	**-0.1170**	**-0.1310**	**-0.1231**	**-0.1222**	**-0.1205**	**-0.1109**	**-0.1172**	**-0.1151**
	Mean	0.0121	0.0085	0.0072	0.0067	0.0063	0.0060	0.0059	0.0057	0.0056	0.0055
M3 Full	Std	0.0563	0.0287	0.0194	0.0147	0.0118	0.0098	0.0084	0.0073	0.0066	0.0058
	SR	**0.1295**	**0.1277**	**0.1245**	**0.1269**	**0.1292**	**0.1193**	**0.1292**	**0.1293**	**0.1228**	**0.1238**
	Mean	-0.0339	0.0054	0.0047	0.0045	0.0035	0.0023	0.0024	0.0025	0.0019	0.0023
M3 Linear	Std	0.1835	0.0516	0.0632	0.0563	0.0583	0.0537	0.0485	0.0442	0.0409	0.0374
	SR	**-0.2109**	**0.0125**	**-0.0016**	**-0.0056**	**-0.0222**	**-0.0460**	**-0.0487**	**-0.0521**	**-0.0702**	**-0.0666**
	Mean	0.0205	0.0365	0.0262	0.0212	0.0168	0.0136	0.0120	0.0107	0.0093	0.0088
M3 L noH	Std	0.1636	0.2003	0.1290	0.0921	0.0627	0.0444	0.0340	0.0272	0.0207	0.0178
	SR	**0.0958**	**0.1584**	**0.1660**	**0.1786**	**0.1921**	**0.1987**	**0.2134**	**0.2164**	**0.2200**	**0.2253**

Note: This table presents the means, standard deviations and Sharpe ratios of out-of-sample returns of various models with a riskless asset for the last 10 years (from Jan. 1991 to Dec. 2000).

making the performance of the linear model even worse as a whole. When we exclude the higher order moment information, the weight on the EM index is in between. This is exactly reflected in the performance of each decision rule. Later, the same pattern is shown the case with a riskless asset again.

Table II.11 shows the results of the case with a riskless asset. In this case, the performance of Model 1 is dominant over others, unlike in the absence of a riskless asset. However, Models 2 and 3 tend to invest in the EM index excessively, like in the absence of a riskless asset. In turn, the performances of those models are not particularly good. In fact, our samples in this paper are a little extreme cases. During the 1990s, the US market was bullish; the Emerging Market was in a bearish mood. Therefore, we need to extend this study to other cases to confirm whether Models 2 and 3, which include higher order moment information, still perform poorly or not. Also we need to check which moment information is helpful in terms of performance of a portfolio choice problem. This will provide robust conclusions about the usefulness of higher order moment information. We will leave this for future research. In the next section, we will use another performance measure, the realized wealth and the realized utility.

II.5.B.b Realized Wealth and Realized Utility

In this section, we compare the performance of Models by realized wealth and realized utility. First, we assume that the investors have one dollar at the beginning of the period (Jan. 1991), and that they invest one dollar according to the optimal decision rule based on each comparison model and roll-over each period. Another assumption here is that once the investors are bankrupt, $W_t < 0$, they can finance by borrowing money from other people or other ways, and invest one dollar and roll-over again.[30] Figure II.10 shows the realized wealth of an

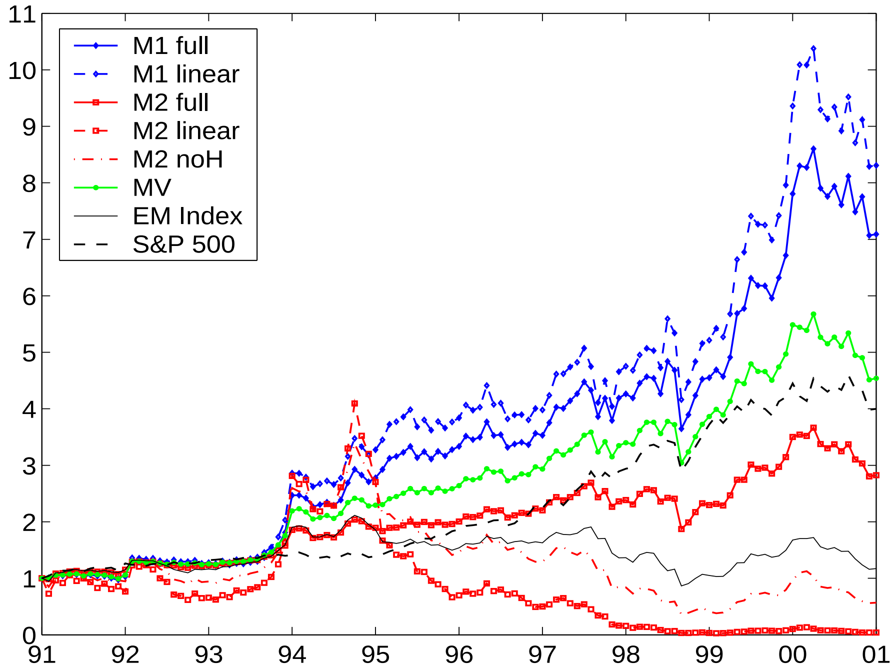
[30] Bankruptcy cases happen with low risk aversion, such as $\gamma = 1$. The following is the summary of the number of bankruptcies: (1) In the absence of a riskless asset: M2 Linear (16 bankruptcies with $\gamma = 1$), M2 L noH (5 with $\gamma = 1$), M3 Linear (15 with $\gamma = 1$), and M3 L noH (5 with $\gamma = 1$). (2) In the presence of a riskless asset: M2 Linear (15 with $\gamma = 1$, 2 with $\gamma = 2$), M2 L noH (2 with $\gamma = 1$). Two reasons are suggested. Low risk aversion investors tend to invest extremely, and the poor performance of the EM index in our sample. Therefore they may experience bankruptcies many times during 1990s. However, the assumption that general investors' risk aversion is higher than 1 or even 2 may be reasonable. Therefore, we can disregard extremely low risk aversion cases.

investor with $\gamma = 3$ who rolls-over his or her wealth during the out-of-sample period for both with and without a riskless asset cases. Of course, the level of realized wealth really depends on the degree of the investor's risk aversion. Higher γ means more conservative portfolio choice, so that the level of realized wealth becomes lower. One example is that if we assume γ is bigger than 6, the level of realized wealth at the last data point is lower than the case of investing only on the S&P 500 Index, although volatility also is lower. Assumption of $\gamma = 3$ is not strange and is supported by a great deal of financial literature, therefore, it may be a reasonable assumption. See, Bollerslev et al. (2004). In figure II.10, we do not include the results based on Model 3, since they are very similar to the results of Model 2. This might be futher evidence that the assumption of dependence structure does not affect the optimal portfolio choice rule, and, in turn, the performance of the decision rules.

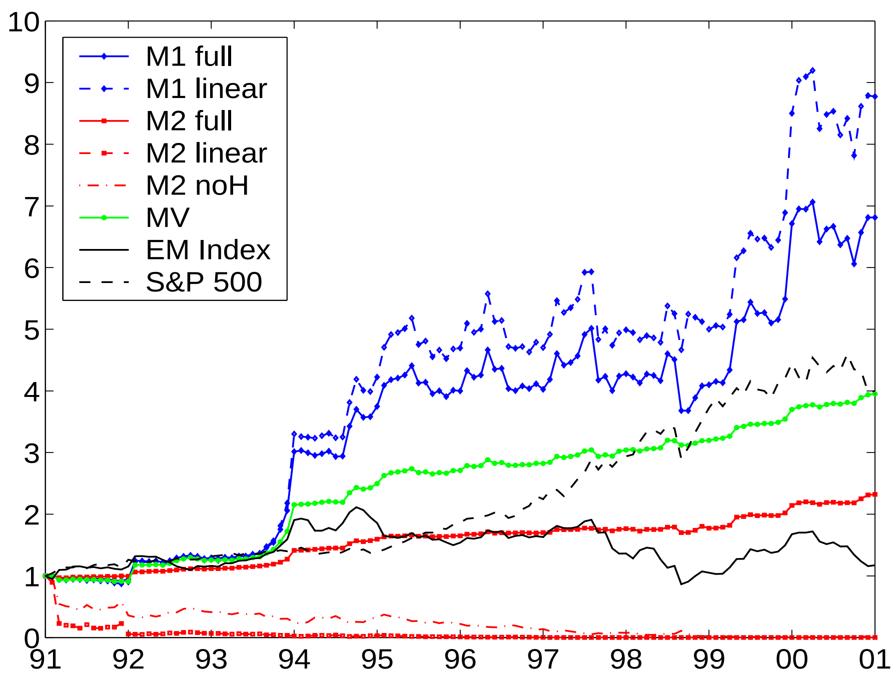
However, we can see that Model 1, especially the linearly approximated decision rule based on Model 1, is dominant. In the comparison to the Sharpe ratio, the results from MV show that it outperforms others. However, MV strategy seems to be very conservative so that it generates higher a Sharpe ratio, but not higher realized wealth. The poor performance of Model 2 is confirmed here again. Consideration of higher moment effects provides substantial losses in the rolling-over realized wealth in our sample. As mentioned, this is because the decision rule of Model 2 (or 3) place too much weight on the Emerging Market Index. However, the better performance of by Model 1 might be due to a few isolated extremely underperforming or outperforming months in our sample.

In order to eliminate that possibility, we perform the following test. First, we assume that we invest one dollar in each period following the optimal strategies, and then we can calculate the realized wealth, $RW_{i,t}^{\gamma}$, and the realized utilities, $RU_{i,t}^{\gamma}$, of each model. Here, i is a model index. For MV preference (quadratic utility) investors, the realized utilities are sometimes expressed in terms of Certainty Equivalent (CEQ) returns: see, DeMiguel et al. (2006). However, in the CRRA

(a) In the Absence of a Riskless Asset



(b) In the Presence of a Riskless Asset



Note: This graph shows the realized wealth of rolling-over strategies based on several portfolio choice models. The results based on Model 3 are not here, since those are very similar to Model 2's results.

Figure II.10 Realized Wealth of Rolling-Over Strategy ($\gamma = 3$)

Table II.12 Performance Comparison Test: Realized Wealth

	γ	1	2	3	4	5	6	7	8	9	10
Without a Riskless Asset	$E(\eta^{\gamma}_{1F-2F,t})$	0.0222***	0.0125***	0.0086***	0.0069***	0.0060***	0.0055***	0.0051***	0.0049***	0.0048***	0.0048***
	SE	(0.0076)	(0.0041)	(0.0028)	(0.0022)	(0.0019)	(0.0017)	(0.0015)	(0.0014)	(0.0014)	(0.0014)
	$E(\eta^{\gamma}_{1F-3F,t})$	0.0321***	0.0202***	0.0156***	0.0133***	0.0121***	0.0112***	0.0106***	0.0101***	0.0097***	0.0095***
	SE	(0.0103)	(0.0056)	(0.0042)	(0.0036)	(0.0034)	(0.0032)	(0.0032)	(0.0031)	(0.0031)	(0.0031)
	$E(\eta^{\gamma}_{2F-3F,t})$	0.0098***	0.0077***	0.0070***	0.0065***	0.0061**	0.0057**	0.0054**	0.0052**	0.0049**	0.0047**
	SE	(0.0031)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0023)
	$E(\eta^{\gamma}_{1F-1L,t})$	-0.0038	-0.0025	-0.0020	-0.0016	-0.0014	-0.0012*	-0.0010*	-0.0009*	-0.0009*	-0.0008**
	SE	(0.0033)	(0.0021)	(0.0014)	(0.0010)	(0.0008)	(0.0007)	(0.0006)	(0.0005)	(0.0004)	(0.0004)
	$E(\eta^{\gamma}_{2F-2L,t})$	0.0774	0.0374	0.0201	0.0219	0.0205	0.0181	0.0157	0.0147	0.0123	0.0116
	SE	(0.0676)	(0.0228)	(0.0134)	(0.0139)	(0.0129)	(0.0114)	(0.0100)	(0.0093)	(0.0080)	(0.0075)
	$E(\eta^{\gamma}_{3F-3L,t})$	0.0693	0.0229	0.0119	0.0147	0.0143	0.0120	0.0097	0.0086	0.0061	0.0062
	SE	(0.0652)	(0.0180)	(0.0112)	(0.0115)	(0.0106)	(0.0090)	(0.0075)	(0.0067)	(0.0052)	(0.0050)
	$E(\eta^{\gamma}_{3F-3nH,t})$	0.0621	0.0031	-0.0010	0.0039	0.0049	0.0039	0.0028	0.0027	0.0009	0.0010
	SE	(0.0432)	(0.0078)	(0.0049)	(0.0059)	(0.0058)	(0.0048)	(0.0039)	(0.0036)	(0.0026)	(0.0024)
With a Riskless Asset	$E(\eta^{\gamma}_{1F-2F,t})$	0.0301**	0.0171**	0.0114**	0.0085**	0.0069**	0.0059**	0.0052**	0.0045**	0.0041**	0.0037**
	SE	(0.0138)	(0.0078)	(0.0053)	(0.0040)	(0.0031)	(0.0025)	(0.0022)	(0.0019)	(0.0017)	(0.0015)
	$E(\eta^{\gamma}_{1F-3F,t})$	0.0299**	0.0169**	0.0114**	0.0084**	0.0068**	0.0061**	0.0052**	0.0045**	0.0042**	0.0038**
	SE	(0.0135)	(0.0078)	(0.0052)	(0.0039)	(0.0031)	(0.0025)	(0.0021)	(0.0019)	(0.0017)	(0.0015)
	$E(\eta^{\gamma}_{2F-3F,t})$	-0.0002	-0.0003	0.0000	-0.0001	-0.0001	0.0001	0.0000	0.0001	0.0001	0.0001
	SE	(0.0005)	(0.0004)	(0.0003)	(0.0003)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
	$E(\eta^{\gamma}_{1F-1L,t})$	-0.0035	-0.0032	-0.0026	-0.0020	-0.0016	-0.0011	-0.0008	-0.0008	-0.0005	-0.0005
	SE	(0.0039)	(0.0025)	(0.0018)	(0.0013)	(0.0011)	(0.0009)	(0.0008)	(0.0007)	(0.0006)	(0.0006)
	$E(\eta^{\gamma}_{2F-2L,t})$	0.0259	0.0535	0.0462	0.0380*	0.0322*	0.0258*	0.0222*	0.0190*	0.0172*	0.0160
	SE	(0.0773)	(0.0413)	(0.0304)	(0.0219)	(0.0188)	(0.0150)	(0.0130)	(0.0114)	(0.0100)	(0.0096)
	$E(\eta^{\gamma}_{3F-3L,t})$	0.0460**	0.0030	0.0025	0.0022	0.0028	0.0036	0.0034	0.0032	0.0037	0.0032
	SE	(0.0165)	(0.0053)	(0.0063)	(0.0056)	(0.0057)	(0.0052)	(0.0047)	(0.0043)	(0.0039)	(0.0036)
	$E(\eta^{\gamma}_{3F-3nH,t})$	-0.0084	-0.0281	-0.0190*	-0.0146*	-0.0105*	-0.0077*	-0.0062**	-0.0049**	-0.0038**	-0.0033**
	SE	(0.0135)	(0.0174)	(0.0113)	(0.0081)	(0.0055)	(0.0039)	(0.0030)	(0.0024)	(0.0018)	(0.0016)

Note: *** denotes 1% significant level, ** denotes 5% significant level, and * denotes 10% significant level.

Table II.13 Performance Comparison Test: Realized Utility

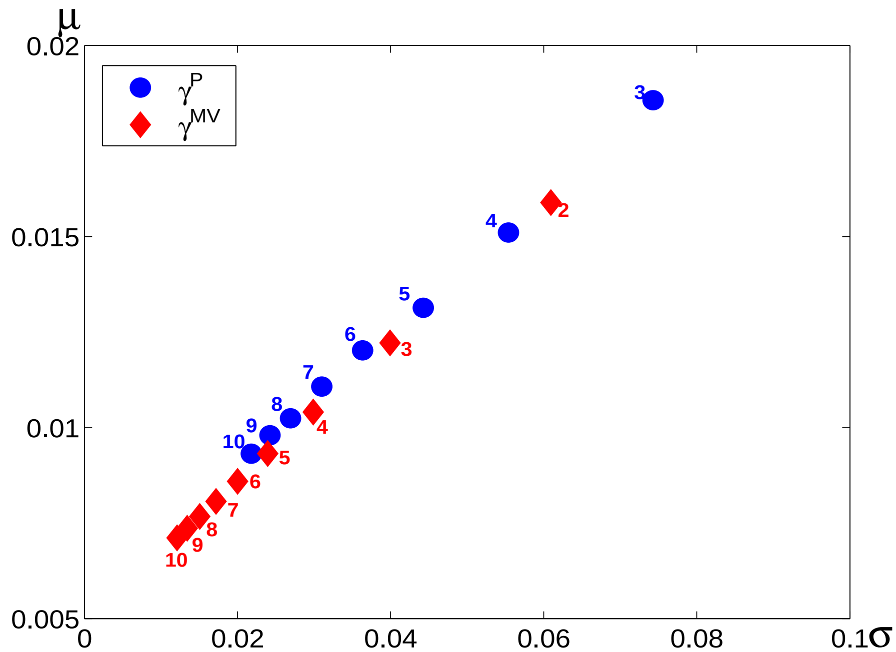
	γ	1	2	3	4	5	6	7	8	9	10
Without a Riskless Asset	$E(\zeta^{\gamma}_{1F-2F,t})$	0.0163*	0.0090*	0.0062	0.0051	0.0047	0.0045	0.0046	0.0048	0.0053	0.0059
	SE	(0.0090)	(0.0051)	(0.0041)	(0.0039)	(0.0041)	(0.0045)	(0.0051)	(0.0060)	(0.0074)	(0.0091)
	$E(\zeta^{\gamma}_{1F-3F,t})$	0.0259	0.0177	0.0155	0.0152	0.0158	0.0168	0.0183	0.0201	0.0225	0.0254
	SE	(1.1165)	(0.0395)	(0.0198)	(0.0303)	(0.0378)	(0.0413)	(0.0433)	(0.0535)	(0.0520)	(0.0642)
	$E(\zeta^{\gamma}_{2F-3F,t})$	0.0095	0.0088	0.0093	0.0101	0.0111	0.0122	0.0136	0.0153	0.0172	0.0196
	SE	(1.0852)	(0.0262)	(0.0168)	(0.0262)	(0.0347)	(0.0366)	(0.0369)	(0.0428)	(0.0365)	(0.0496)
	$E(\zeta^{\gamma}_{1F-1L,t})$	-0.0010	-0.0001	-0.0002	-0.0002	-0.0002	-0.0002	-0.0001	-0.0002	-0.0001	-0.0002
	SE	(0.0026)	(0.0016)	(0.0010)	(0.0007)	(0.0006)	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0002)
	$E(\zeta^{\gamma}_{2F-2L,t})$	5.0195***	0.1477***	0.0733***	0.1090***	0.1260***	0.1278***	0.1239***	0.1391**	0.1261**	0.1421**
	SE	(1.1165)	(0.0395)	(0.0198)	(0.0303)	(0.0378)	(0.0413)	(0.0433)	(0.0535)	(0.0520)	(0.0642)
	$E(\zeta^{\gamma}_{3F-3L,t})$	4.7219***	0.0946***	0.0590***	0.0926***	0.1134***	0.1108***	0.1031***	0.1091**	0.0866**	0.1069**
	SE	(1.0852)	(0.0262)	(0.0168)	(0.0262)	(0.0347)	(0.0366)	(0.0369)	(0.0428)	(0.0365)	(0.0496)
	$E(\zeta^{\gamma}_{3F-3nH,t})$	1.6573**	0.0181**	0.0097*	0.0266***	0.0353***	0.0341***	0.0309***	0.0349**	0.0239**	0.0277**
	SE	(0.6559)	(0.0081)	(0.0053)	(0.0086)	(0.0110)	(0.0113)	(0.0111)	(0.0135)	(0.0104)	(0.0130)
With a Riskless Asset	$E(\zeta^{\gamma}_{1F-2F,t})$	0.0144	0.0077	0.0049	0.0037	0.0030	0.0028	0.0025	0.0021	0.0020	0.0018
	SE	(0.0125)	(0.0065)	(0.0043)	(0.0031)	(0.0025)	(0.0019)	(0.0016)	(0.0014)	(0.0012)	(0.0011)
	$E(\zeta^{\gamma}_{1F-3F,t})$	0.0143	0.0075	0.0049	0.0037	0.0030	0.0030	0.0026	0.0022	0.0021*	0.0019*
	SE	(0.0123)	(0.0065)	(0.0043)	(0.0031)	(0.0024)	(0.0019)	(0.0016)	(0.0014)	(0.0012)	(0.0011)
	$E(\zeta^{\gamma}_{2F-3F,t})$	-0.0001	-0.0002	0.0000	0.0000	0.0000	0.0002	0.0001	0.0001	0.0001	0.0001
	SE	(0.0005)	(0.0004)	(0.0003)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
	$E(\zeta^{\gamma}_{1L-1F,t})$	-0.0040	-0.0024	-0.0015	-0.0010	-0.0007	-0.0002	0.0000	-0.0001	0.0001	0.0001
	SE	(0.0041)	(0.0025)	(0.0018)	(0.0011)	(0.0010)	(0.0008)	(0.0007)	(0.0007)	(0.0006)	(0.0005)
	$E(\zeta^{\gamma}_{2F-2L,t})$	4.4080***	1.1035*	0.3638***	0.2190***	0.1936***	0.1361***	0.1155***	0.0965***	0.0841***	0.0859***
	SE	(1.0639)	(0.6297)	(0.1050)	(0.0567)	(0.0490)	(0.0333)	(0.0279)	(0.0231)	(0.0199)	(0.0204)
	$E(\zeta^{\gamma}_{3F-3L,t})$	0.0636***	0.0048	0.0077	0.0078	0.0106*	0.0116**	0.0110**	0.0103**	0.0105***	0.0095***
	SE	(0.0176)	(0.0052)	(0.0062)	(0.0055)	(0.0057)	(0.0052)	(0.0047)	(0.0043)	(0.0039)	(0.0036)
	$E(\zeta^{\gamma}_{3F-3nH,t})$	0.0066	0.1620	0.0184	0.0070	0.0007	-0.0013	-0.0019	-0.0019	-0.0018	-0.0017
	SE	(0.0166)	(0.1667)	(0.0250)	(0.0137)	(0.0074)	(0.0047)	(0.0033)	(0.0026)	(0.0019)	(0.0016)

Note: *** denotes 1% significant level, ** denotes 5% significant level, and * denotes 10% significant level.

case, the realized utilities depend on the realized wealth and the risk aversion, γ^P, and it is a completely different measure than the CEQ of MV preference investors. Therefore, we only test the decision rules with the CRRA utility function, which is indexed by Models 1, 2 and 3. In order to test whether Model i performs differently than Model j, we define $\eta^{\gamma}_{i-j,t} = RW^{\gamma}_{i,t} - RW^{\gamma}_{j,t}$, and $\zeta^{\gamma}_{i-j,t} = RU^{\gamma}_{i,t} - RU^{\gamma}_{j,t}$. Then we test the nulls, $E(\eta^{\gamma}_{i-j,t}) \neq 0$ and $E(\zeta^{\gamma}_{i-j,t}) \neq 0$. As mentioned, this test eliminates the possibility where one or a few extremely good or bad samples distort the final wealth results.

Table II.12 summarizes the results based on the realized wealth. Even with this criterion, we can see that Model 1 performs better than Models 2 and 3 in both cases, supporting previous results. However, we can not see clear and consistent evidence that full information models outperform liner approximation based decision rules. Instead, there are some cases where linear approximated rules perform better than full information decision rules. However, when we change the performance measure to the realized utility shown in Table II.13, the results look different. Full information models perform better than linearly approximated decision rules. Also we do not observe strong evidence that Model 1 outperforms Models 2 and 3. This is because of the shape of utility function.

We can summarize the results on the whole as following. If we consider more sophisticated (joint or marginal) distributions, we might well get more information about return distributions. In our sample, more information leads investors to give more weight to the Emerging Market index to maximize their expected utility. However, the poor performance of the EM index seems to negatively affect portfolio performance. In other words, a large forecasting error in the EM index process may result in the poor performance of decision rules of sophisticated models. Therefore, at least in our sample, the performance of decision rules from the skewed GT marginal distribution does not seem to produce a dominant strategy over the others. That is, aggressive decision rules may sometimes ruin the performance of an investment along with a gross forecasting error. Of course, the EM



Note: This shows the means and standard deviations of MV decision rule and CRRA decision rule based on Model 1. Circled dots represent the combination of expected mean and volatility of decision rules based on the CRRA utility function under the multivariate normal distribution. Diamond dots represent the combination of expected mean and volatility of decision rules based on the MV utility function under the multivariate normal distribution.

Figure II.11 The Mean-Variance in Out-of-Sample

index in our out-of-sample seems to be very risky. As observed in the previous section, CRRA investors tend to invest more on higher volatility risky assets. It could be a main reason for the underperformance of Models 2 and 3.

II.5.C γ^{MV} vs γ^{P}

Here, γ^{MV} indicates the MV investor's degree of risk aversion and γ^{P} represents the CRRA investor's risk aversion. Superscript P indicates the power utility function. In this section, we present the possible connection between γ^{MV} and γ^{P} induced by out-of-sample performance test results. As known, it is not easy to compare the risk aversion level of MV utility and CRRA utility investors directly as the utility functions are different. However, Figure II.11 provides an indirect relationship between them. Circled dots represent the combination of expected mean and volatility of decision rules based on the CRRA utility function under the multivariate normal distribution. Diamond dots represent the combination of expected mean and volatility of decision rules based on the MV utility function under the multivariate normal distribution. The decision rules of The CRRA utility investors under non-multivariate normal distributions is also governed by other factors. Therefore it is meaningless to plot them on the mean-variance space, so we exclude those cases. As seen in Figure II.11, $\gamma^P \neq \gamma^{MV}$. We may say that $\gamma^P = 6 \approx \gamma^{MV} = 3$, $\gamma^P = 8 \approx \gamma^{MV} = 4$, and so on. This implies that investors' risk taking with the CRRA utility function is greater than the MV utility function at the same level of numerical value.

II.6 Conclusion

This paper examines how investors with the CRRA utility function react to changes in the return distribution as capturred by its moments. In other words, we disentangle how investors' decision rule changes depending on changes in each moment. To do that, we generate a data set by simulation, then we use a simple projection method based on the regime switching model. Further, based on the projection results, we propose a linearly approximated portfolio choice rule by CRRA utility investors, which makes it possible to include or exclude uncertainty information. Finally, we apply these decision rules to compare their performances under various distributional assumptions and the Mean-Variance decision rule, taking into consideration the dynamics of parameters during out-of-sample.

As everyone agrees, the expected mean is the most important factor for investors to make a decision on how to allocate and invest their wealth. Given the data set used in this paper, the expected return can explain from about 70% to upwards of 90% of their investment behavior, depending on the degree of risk aversion and distributional assumptions. Of course, this interpretation comes from an approximated linear decision rule, therefore we can not consider all possible non-linearities of the optimal policy function. When we define risk as all other factors which describe the uncertainty of investment opportunities, its explanatory power is modest at best, from nearly negligible with lower risk aversion to at most about 10%, even with higher risk aversion. In one extreme case, Model 1 and $\gamma = 10$ without a riskless asset, it can explain more than 25%. However, in general, we can say that the risk factors explain a relatively small part of CRRA investors' decisions.

One seemingly striking result is the relationship between volatility and investment weight in the presence of a riskless asset. It may seem that a risk averse investor, who wants to maximize expected utility, tries to avoid volatility risk. However, the truth is that he tries to maximize his expected utility by increasing the weight of the riskier asset. Another interesting result is the effect of the dependence structure on the asset allocation problem. With a naive multivariate normal distribution, the dependence structure has a statistically significant effect on the decision rule. However, once we use a more sophisticated distributional model, which has more parameters to capture higher order moment characteristics such as the skewed GT distribution, the dependence structure becomes insignificant in the absence of a riskless asset. However, with a riskless asset, the dependence structure does not seem to have significant explanatory power on the optimal portfolio choice. Multivariate joint distributions are relatively harder to handle in many problems than univariate distribution. Therefore, once we consider each univariate distribution more carefully, information from the dependence structure may not be very important, at least, in an asset allocation problem. Also, among higher moment parameters, the skewness parameter is important in many cases, while other higher moment parameters are not statistically significant.

Although we linearly approximate the allocation decision, it fits the unknown true allocation function very well. Based on the return process, we can facilitate the linearly approximated allocation function. To do that, we use the time-varying parameter model. As one possible model, we use an AR(1)-GARCH(1,1)-skewed GT model. Estimation results represent low persistency in expected means and high persistency in volatility. Therefore, the expected return is very important to short term buy-and-hold investors. Of course, the expected return is important to long-term investors as well. However, since the expected volatility has longer persistence in our sample, after the expected return converges to the steady-state value in the very short term, the level of expected volatility becomes the most important factor. That creates different optimal asset allocation decisions between long and short term investors.

Finally, we apply a linear decision rule in an out-of-sample performance test for a 1-month buy-and-hold investor. In order to compare performances, we use various decision rules and risk aversions along with different distributional assumptions. Compared to the MV investment rule, the performance of the decision rule from a multivariate normal distribution (Normal copula and Normal margins), Model 1, performs well. Especially, the performance of the linear decision rule based on Model 1 is dominant in our sample in terms of realized wealth. However, when we compare the performances from the multivariate normal distribution and the Normal copula with the skewed GT margins specification, there is no clear evidence that the sophisticated distribution model performs better. It even seems that the simpler model performs better. Additionally, more information about risk factors may lead people to invest more aggressively. However, this might ruin the performance of the investment, perhaps due to a forecasting error.

There are some limitations. First, we do not consider all possible nonlinearity of decision rules. Second, the linear portfolio choice rules could yield local solutions, depending on the possible range of parameter values. Therefore, we could extend it to find globally approximated rules under different sampling schemes. Also, most obviously, we could extend it to include many other asset cases. Emerging Markets are very volatile. If we apply this method to more stable markets, we may well get a better-fitted linear approximated decision function. Further, this framework easily cooperates with many other features of return process so that we can create various portfolio choice rules, which may use selective information about risk factors. Possible extensions are the use of predictability or others, such as a jump component of the return process. One assumption of this framework is that the linearly approximated decision rule fits well enough to use in this way. We may need to have some supporting research to make such conclusions firm. Finally, this work does not impose any constraints, such as no short selling or borrowing. We could extend this framework to include such constraints.

Apendix A Skewed Generalized T Distribution

AA.1 Probability Density Function

$$f(x|\sigma, \nu, \lambda, k) = C \left(1 + \frac{k}{(\nu - 2)} \left\{ \frac{|x|}{\sigma\theta(1 + sign(x) \cdot \lambda)} \right\}^k \right)^{-\frac{\nu+1}{k}} \tag{II.23}$$

where, $C = \frac{kS(\lambda)}{2\sigma} B\left(\frac{1}{k}, \frac{\nu}{k}\right)^{-\frac{3}{2}} B\left(\frac{3}{k}, \frac{\nu-2}{k}\right)^{\frac{1}{2}}$, $\theta = \frac{1}{S(\lambda)} \left(\frac{k}{\nu-2}\right)^{\frac{1}{k}} B\left(\frac{1}{k}, \frac{\nu}{k}\right)^{\frac{1}{2}} B\left(\frac{3}{k}, \frac{\nu-2}{k}\right)^{-\frac{1}{2}}$, and $S(\lambda) = \left(1 + 3\lambda^2 - 4\lambda^2 B\left(\frac{2}{k}, \frac{\nu-1}{k}\right)^2 B\left(\frac{1}{k}, \frac{\nu}{k}\right)^{-1} B\left(\frac{3}{k}, \frac{\nu-2}{k}\right)^{-1}\right)^{\frac{1}{2}}$. $k > 0$ and $\nu > 2$ control the height and tails, the skewness parameter $-1 < \lambda < 1$ controls the rate of descent of density around $x = 1$, and σ^2 is the variance. ν has the degree of freedom interpretation in case $\lambda = 0$ and $k = 2$.

The skewed GT distribution generates the Generalized t distribution for $\lambda = 0$, Hansen's skewed t for $k = 2$, the students' t for $\lambda = 0$ and $k = 2$, the power exponential for $\lambda = 0$ and $\nu = \infty$, the Laplace for $\lambda = 0$, $k = 1$ and $\nu = \infty$, the Cauchy for $\lambda = 0$, $k = 2$ and $\nu = 1$, the normal for $\lambda = 0$, $k = 2$ and $\nu = \infty$, finally the uniform for for $\lambda = 0$, $k = \infty$ and $\nu = \infty$. This is from Theodossiou (1998).

AA.2 *m*th Non-Centered Moment

Theorem Apendix A.1. *The mth non-centered moment is*

$$\begin{aligned} M_m \equiv E(x^m) = \frac{1}{2}\left((-1)^m(1-\lambda)^{(m+1)} + (1+\lambda)^{(m+1)}\right)\left(\frac{\sigma}{S(\lambda)}\right)^m \\ \times B\left(\frac{m+1}{k}, \frac{\nu-m}{k}\right) B\left(\frac{1}{k}, \frac{\nu}{k}\right)^{\frac{m-2}{2}} B\left(\frac{3}{k}, \frac{\nu-2}{k}\right)^{-\frac{m}{2}} \end{aligned} \tag{II.24}$$

Proof. see Theodossiou (1998) □

AA.3 Cumulative Density Function

Theorem Apendix A.2. *The Cumulative Density Function (CDF) is*

$$F(x) = \begin{cases} C\left(\frac{1}{q}\right)^{\frac{1}{k}} \frac{1}{k} B_y\left(\frac{\nu}{k}, \frac{1}{k}\right), & for x < 0 \\ 1 - C\left(\frac{1}{q}\right)^{\frac{1}{k}} \frac{1}{k} B_y\left(\frac{\nu}{k}, \frac{1}{k}\right), & for x \geq 0 \end{cases} \tag{II.25}$$

where, $y = \left(1 + q \cdot |x|^k\right)^{-1}$ *and* $q \equiv \frac{k}{(\nu - 2)} \left(\frac{1}{\sigma\theta(1 + sign(x) \cdot \lambda)}\right)^k$

Proof. let's $q \equiv \frac{k}{(\nu - 2)} \left(\frac{1}{\sigma\theta(1 + sign(x) \cdot \lambda)}\right)^k$ and $m \equiv \frac{\nu + 1}{k}$. Then $q > 0$ and $m > 0$, further $f(x|k, \nu, \lambda, \sigma) = C\left(1 + q \cdot |x|^k\right)^{-m}$.

For $x < 0$, CDF of this probability function is

$$\int_{-\infty}^{x} C\left(1 + q \cdot (-u)^k\right)^{-m} du \tag{II.26}$$

Let's $v = \left(1 + q \cdot (-u)^k\right)^{-1}$, then $u = -\left(\frac{1}{q}\right)^{\frac{1}{k}} \left(\frac{1 - v}{v}\right)^{\frac{1}{k}}$. Therefore,

$$\frac{du}{dv} = \left(\frac{1}{q}\right)^{\frac{1}{k}} \frac{1}{k} (1 - v)^{\left(\frac{1-k}{k}\right)} v^{\left(-\frac{1+k}{k}\right)} \tag{II.27}$$

By substitution, the original integration becomes

$$\begin{aligned} \int_{-\infty}^{x} C\left(1 + q \cdot (-u)^k\right)^{-m} du &= C \int_{0}^{y} v^m \left(\frac{1}{q}\right)^{\frac{1}{k}} \frac{1}{k} (1 - v)^{\left(\frac{1-k}{k}\right)} v^{\left(-\frac{1+k}{k}\right)} dv \\ &= C\left(\frac{1}{q}\right)^{\frac{1}{k}} \frac{1}{k} \int_{0}^{y} (1 - v)^{\left(\frac{1}{k}-1\right)} v^{\left(\frac{\nu}{k}-1\right)} dv \\ &= C\left(\frac{1}{q}\right)^{\frac{1}{k}} \frac{1}{k} B_y\left(\frac{\nu}{k}, \frac{1}{k}\right) \end{aligned} \tag{II.28}$$

where $y = \left(1 + q \cdot |x|^k\right)^{-1}$, and $0 \leq y < 1$. And $B_y\left(\frac{\nu}{k}, \frac{1}{k}\right)$ is the incomplete beta function.

For $x \geq 0$, CDF of this probability function is

$$1 - \int_{x}^{\infty} C\left(1 + q \cdot u^k\right)^{-m} du \tag{II.29}$$

Similarly, let's $\tilde{v} = \left(1 + q \cdot u^k\right)^{-1}$, then $u = \left(\frac{1}{q}\right)^{\frac{1}{k}} \left(\frac{1-\tilde{v}}{\tilde{v}}\right)^{\frac{1}{k}}$. Therefore,

$$\frac{du}{d\tilde{v}} = -\left(\frac{1}{q}\right)^{\frac{1}{k}} \frac{1}{k}(1-\tilde{v})^{(\frac{1-k}{k})}\tilde{v}^{(-\frac{1+k}{k})} \tag{II.30}$$

By substitution, the original integration becomes

$$\begin{aligned}
\int_x^\infty C\left(1 + q \cdot u^k\right)^{-m} du &= -\int_\infty^x C\left(1 + q \cdot u^k\right)^{-m} du \\
&= -C\int_0^{\tilde{y}} -v^m \left(\frac{1}{q}\right)^{\frac{1}{k}} \frac{1}{k}(1-\tilde{v})^{(\frac{1-k}{k})}\tilde{v}^{(-\frac{1+k}{k})}dv \\
&= C\left(\frac{1}{q}\right)^{\frac{1}{k}} \frac{1}{k}\int_0^{\tilde{y}} (1-\tilde{v})^{(\frac{1}{k}-1)}\tilde{v}^{(\frac{\nu}{k}-1)}dv \\
&= C\left(\frac{1}{q}\right)^{\frac{1}{k}} \frac{1}{k}B_{\tilde{y}}\left(\frac{\nu}{k}, \frac{1}{k}\right)
\end{aligned} \tag{II.31}$$

where $\tilde{y} = \left(1 + q \cdot |x|^k\right)^{-1}$, and $0 \le \tilde{y} \le 1$. Therefore, CDF for $x \ge 0$ is

$$1 - C\left(\frac{1}{q}\right)^{\frac{1}{k}} \frac{1}{k}B_{\tilde{y}}\left(\frac{\nu}{k}, \frac{1}{k}\right) \tag{II.32}$$

$\square$

Apendix B Griddy Gibbs Sampling

Suppose we have the simple ARCH model

$$\begin{aligned}
y_t &= \mu + \epsilon_t\sqrt{h_t} \\
\epsilon_t|\mathcal{F}_{t-1} &\sim N(0,1) \\
h_t &= \omega + \alpha(y_{t-1} - \mu)^2
\end{aligned} \tag{II.33}$$

Let us assume a flat prior, then the conditional posterior density $\varphi(\mu|\omega,\alpha)$ has a kernel as the following

$$\kappa(\mu|\omega,\alpha) = \Pi_t[h_t(\omega,\alpha,\mu)]^{-\frac{1}{2}} \exp -\frac{(y_t-\mu)^2}{2h_t(\omega,\alpha,\mu)} \tag{II.34}$$

When ω and α are a given fixed value. If h_t were fixed, equation (II.34) would be a normal density. As h_t is a function of ω, α, and μ as well, the conditional posterior density of μ contains h_t, which is also a function of μ. Consequently, it cannot be a normal or any other well known density from which random numbers could be easily generated. However, the kernel of $\varphi(\mu|\omega, \alpha)$ conditioning on a previous draw of the conditioning parameter, can be evaluated over a grid of points. One can compute the corresponding distribution function using numerical integration techniques. Afterward, one can generate a draw of μ by inversion of the distribution at a random value sampled from a uniform distribution in [0,1]. This technique is called the Griddy Gibbs Sampler. The GGS procedure follows for nth draw of the posterior.

1. Given values of ω^{n-1} and α^{n-1}, compute $\kappa(\mu|\omega, \alpha)$ over the G grid points $(\mu_1, \mu_2, \cdots, \mu_K)$, so we can get the vector $\phi_\kappa = (\kappa_1, \kappa_2, \cdots, \kappa_G)$.

2. Using one of numerical integration techniques, compute the vector of the values $\Phi_\Phi = (\Phi_1, \Phi_2, \cdots, \Phi_G)$ by $\Phi_i = \displaystyle\int_{\mu_1}^{\mu_i} \kappa(\mu|\omega^{n-1}, \alpha^{n-1}, y)d\mu$, $i = 1, 2, \cdots, G$.

3. Compute the normalized cdf values $\Phi_N = \Phi_\Phi/\Phi_G$

4. Generate $u \sim U[0, 1]$ and invert by numerical interpolation to get a draw of $\mu|\omega^{n-1}, \alpha^{n-1}, y$. Index μ^n and store this draw.

5. Repeat 1-4 for $\omega^n|\alpha^{n-1}, \mu^n, y$. And repeat again for $\alpha^n|\mu^n, \omega^n, y$.

References

Ang, A., and Bekaert, G., 2002a: International asset allocation with regime shifts. *The review of financial studies*, **15**(4), 1137–1187.

Ang, A., and Bekaert, G., 2002b: Regime switches in interest rates. *Journal of Business and Economic Statistics*, **20**(1), 163–182.

Ang, A., and Chen, J., 2002: Asymmetric correlation of equity portfolios. *Journal of Financial Economics*, **63**(2), 442–294.

Arditti, F. D., 1967: Risk and the required return on equity. *Journal of Finance*, **22**(1), 19–36.

Athayde, G. M., and Flôres Jr, R. G., 2004: Finding a maximum skewness portfolio - a general solution to three-moments portfolio choice. *Journal of Economic Dynamics and Control*, **28**(7), 1335–1352.

Balduzzi, P., and Lynch, A. W., 1999: Transaction costs and predictability: Some utility cost calculations. *Journal of Financial Economeics*, **52**(1), 47–78.

Barberis, N., 2000: Investing for the long run when returns are predictable. *Journal of Finance*, **55**(1), 225–264.

Bauwens, L., Lubrano, M., and Richard, J.-F., 1999: *Bayesian Inference in Dynamic Econometric Models*. Advanced Texts in Econometrics. Oxford University Press, Oxford.

Bollerslev, T., Gibson, M., and Zhou, H., 2004: Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. *Working Paper*.

Campbell, J. Y., Chan, Y. L., and Viceira, L. M., 2003: A multivariate model of strategic asset allocation. *Journal of Financial Economeics*, **67**(1), 41–80.

Campbell, J. Y., and Viceira, L. M., 1999: Consumption and portfolio decisions when expected returns are time varying. *Quarterly Journal of Economics*, **114**, 433–495.

Campbell, J. Y., and Viceira, L. M., 2001: Who should buy long-term bonds? *American Economic Review*, **91**, 99–127.

Chunhachinda, P., Dandapani, K., Hamid, S., and Prakash, A. J., 1997: Portfolio selection and skewness: Evidence from international stock markets. *Journal of Banking and Finance*, **21**(2), 143–167.

DeMiguel, V., Garlappi, L., and Uppal, R., 2006: 1/n. *Working Paper*.

Detemple, J. B., Garcia, R., and Rindisbacher, M., 2003: A monte carlo method for optimal portfolios. *Journal of Finance*, **58**(1), 401–446.

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

Embrechts, P., Lindskog, F., and McNeil, A., 2003: Modelling dependence with copulas and applications to risk management. In *Handbook of Heavy Tailed Distributions in Finance*, editor S. Rachev, 329–384. Elsevier.

Engle, R., 2002: Dynamic conditional correlation-a simple class of multivariate garch models. *Journal of Business and Economic Statistics*, **20**(3), 339–350.

Engle, R., and Sheppard, K., 2001: Theoretical and empirical properties of dynamic conditional correlation multivariate garch. *UCSD Discussion paper 2001-15, UCSD*.

Geweke, J., 2001: A note on some limitations of crra utility. *Economics Letters*, **71**, 341–345.

Guidolin, M., and Timmermann, A., 2005a: An econometric model of nonlinear dynamics in the joint distribution of stock and bond returns. *Journal of Applied Econometrics*, **Forthcoming**.

Guidolin, M., and Timmermann, A., 2005b: International asset allocation under regime switching, skew and kurtosis preferences. *Working Papers, Federal Reserve Bank of St. Louis*.

Guidolin, M., and Timmermann, A., 2005c: Strategic asset allocation and consumption decisions under multivariate regime switching. *Working Papers, Federal Reserve Bank of St. Louis*.

Hamilton, J., 1989: A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, **57**(357-384).

Hansen, B. E., 1994: Autoregressive conditional density estimation. *International Economic Review*, **35**(3), 705–730.

Harvey, C., 1995: Predictable risk and returns in emerging markets. *Review of Financial Studies*, **8**, 773–816.

Harvey, C., Liechty, J., Liechty, M., and Müller, P., 2004: Portfolio selection with higher moments. *mimeo, Duke University*.

Harvey, C., and Siddique, A., 1999: Autoregressive conditional skewness. *Journal of Financial and Quantitative Analysis*, **34**(4), 465–487.

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

Jondeau, E., and Rockinger, M., 2003: Conditional volatility, skewness, and kurtosis: Existence, persistence and comovements. *Journal of Economic Dynamics and Control*, **27**, 1699–1737.

Jondeau, E., and Rockinger, M., 2004: Optimal portfolio allocation under higher moments. *Jornal of the European Financial Management Association*, **Forthcoming**.

Jondeau, E., and Rockinger, M., 2005: Conditional asset allocation under non-normaility: How costly is t the mean-variance criterion? *Working Paper*, **2005**.

Judd, K. L., 1998: *Numerical Methods in Economics*. The MIT Press, Cambridge, MA.

Kandel, S., and Stambaugh, R., 1996: On the predictability of stock returns: An asset allocation perspective. *Journal of Finance*, **51**, 385–424.

Kim, C. J., and Nelson, C. R., 1999: *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. The MIT Press, Cambridge, MA.

Kim, T. S., and Omberg, E., 1996: Dynamic nonmyopic portfolio behavior. *Review of Financial Studies*, **9**(1), 141–161.

Kimball, M. S., 1993: Standard risk aversion. *Econometrica*, **61**(3), 589–611.

Kraus, A., and Litzenberger, R. H., 1976: Skewness preference and the valuation of risk assets. *Journal of Finance*, **31**(4), 1085–1100.

Kroll, Y., Levy, H., and Mariowitz, H. M., 1984: Mean-variance versus direct utility maximization. *Journal of Finance*, **39**(1), 47–61.

Lai, T. Y., 1991: Portfolio selection with skewness: A multiple-objective approach. *Review of Quantitative Finance and Accounting*, **1**, 293–305.

Levy, H., and Markowitz, H. M., 1979: Approximating epected utility by a function of mean and variance. *American Economic Review*, **69**(3), 308–317.

Liu, J., 2006: Portfolio selection in stochastic environments. *Review of Financial Studies*, **forthcoming**.

Loistl, O., 1976: The erroneous approximation of expected utility by means of a taylor's series expansion: Analytic and computational results. *The American Economic Review*, **66**(5), 904–910.

Longin, F., and Solnik, B., 1995: Is the correlation is international equity returns constant: 1960-1990? *Journal of International Money and Finance*, **14**(1), 3–26.

Longin, F., and Solnik, B., 2001: Extreme correlation of international equity markets. *Journal of Finance*, **56**(2), 649–676.

Markowitz, H., 1952: Portfolio selection. *The Journal of Finance*, **7**(1), 77–91.

Marron, J., and Wand, M., 1992: Exact mean integrated squared error. *Annals of Statistics*, **20**(2), 712–736.

Nelsen, R., 1999: *An Introduction to Copulas*. Springer-Verlag, New York.

Perez-Quiros, G., and Timmermann, A., 2000: Firm size and cyclical variations in stock returns. *Journal of Finance*, **55**, 1229–1262.

Pratt, J. W., and Zeckhauser, R. J., 1987: Proper risk aversion. *Econometrica*, **55**(1), 143–154.

Ramchand, L., and Susmel, R., 1998: Volatility and cross correlation across major stock markets. *Journal of Empirical Finance*, **5**, 397–416.

Ritter, C., and Tanner, M. A., 1992: Facilitating the gibbs sampler: The gibbs stopper and the griddy-gibbs sampler. *Journal of American Statistical Association*, **87**(419), 861–868.

Samuelson, P., 1970: The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *Review of Economic Studies*, **37**, 537–542.

Scott, R. C., and Horvath, P. A., 1980: On the direction of preference for moments of higher order than the variance. *Journal of Finance*, **35**(4), 915–919.

Theodossiou, P., 1998: Financial data and the skewed gneralized t distribution. *Management Science*, **44**(12, Part 1 of 2), 1650–1661.

Timmermann, A., 2000: Moments of markov switching models. *Journal of Econometrics*, **96**(1), 75–111.

Wachter, J., 2002: Optimal consumption and portfolio allocation under mean-reverting returns: An exact solution for complete markets. *Journal of Financial and Quantitative Analysis*, **37**, 63–91.

White, H., 2006: Approximate nonlinear forecasting methods. In *Handbook of Economic Forecasting*, editors G. Elliott, C. Granger, and A. Timmermann, volume 1, 459–512. Elsevier North Holland, Amsterdam.

III

Quantile Forecasts and Dependence Structure

III.1 Introduction

Recent work on economic forecasting has focused on evaluating *interval* forecasts rather than *point* forecasts as measures of uncertainty. However, there are relatively fewer studies of interval forecasts, although Christoffersen (1998) indicates the usefulness of interval forecasts which indicate the likely range of outcomes, while point forecasts are of limited value since they only describe one possible outcome[1]. Although the usefulness of interval forecasts is frequently advocated, it has only been in limited use: applied only to univariate cases, although some insights into multivariate extensions are provided in Christoffersen (1998). In this paper, we extend our interest of interval forecasts to multivariate cases and to portfolios of multivariate cases[2].

Most of the interest in interval forecasts arises from the viewpoint of risk management. Risk management for financial market participants usually concerns portfolio risk. Therefore, it is natural to extend interest in interval forecasts to

[1] Clemen et al. (1995) provide an application involving precipitation probability forecasts. Crnkovic and Drachman (1996) propose to evaluate forecast models based on their forecasted distributions.

[2] There is similar extension of density forecasts from univariate to multivariate cases. Diebold et al. (1998) illustrate the framework with a detailed application to density forecasting of asset returns in environments with time-varying volatility. They devote their concern only to univariate cases. Later, Diebold et al. (1999), and Clements and Smith (2002) extend density forecasting to multivariate cases.

multivariate cases. When we extend our interest to multivariate cases, one of the main concerns is how important the specification of the dependence structure is. Unless each variable is independent, dependence structures matter for multivariate models. Recent developments of copula methods help researchers model multivariate cases in an easy and flexible way, and a growing number of research uses copulas to model multivariate density functions. As a consequence of this development, we try to address how dependence structure matters to interval forecasts of multivariate cases.

Chatfield (1993) mentions that model misspecification is a much more important source of poor interval forecasting than is estimation error. Although dependence structure is one element of the multivariate density, little attention has been paid to importance of dependence structure in the context of conditional quantile forecasting. Since we can decompose multivariate density functions into marginal distributions and joint dependence structures, it is possible to observe whether a specification of dependence structures could be important in terms of multivariate interval forecasts. Also we can use interval forecasts as a tool to evaluate the model specifications. In this paper, we provide a simple, but attractive way to assign interval of multivariate cases. After defining an interval, we provide the simulation results and evaluate the interval forecasts of three international stock markets. The evaluation is based on the likelihood ratio (LR) test by Christoffersen (1998). Christoffersen (1998) first advocates the use of conditional expectations in the evaluation of quantile forecasts, and his 'conditional coverage' test has now become standard practice in the interval forecasting literature.

The upshot of this paper is that dependence structures in multivariate cases is of little importance to interval forecasts, so that more thought must be given to marginal distributions. First, we perform a Monte Carlo simulation to study the power of the LR test. Simulation results suggest that the LR test is hard to reject under an alternative dependence structure but under the same margin assumption, as indicated in the literature, such as Berkowitz (2001), Lopez

(1999), etc. This result may imply that there is little difference in the outcome of adopting different dependence structure models. Through an application to three international stock markets, we confirm this result. There is little difference in the test result under different dependence structures, if margins are correctly specified. And no dependence structure seems to dominate others. However, when we turn our interest to the lower left tail case, none of the models used in this paper are suitable to perform interval forecasting well, at least not in our weekly return samples.

Finally, we apply the LR test to the portfolio interval forecasts as well as the conditional quantile forecast evaluation (CQFE) test by Giacomini and Komunjer (2005). After constructing portfolios, multivariate joint interval forecasting problems become univariate problems. Among the GARCH-Normal(1,1) and the GARCH-t(1,1) models, the GARCH-t(1,1) specification passes the conditional coverage tests in all cases. On the other hand, the GARCH-Normal(1,1) fails the tests with extremely low coverage rates, generally used in value-at-risk (VaR) applications, while it passes the tests with moderate coverage rates. The test results of the CQFE test show that combinations of individual forecasts may improve the performance of the interval forecasts of portfolios in many cases, although sometimes the individual forecasts from the GARCH-t(1,1) outperform the forecast combinations.

The remainder of the paper is structured as follows. Section III.2 gives an overview of the LR test of Christoffersen (1998), and then shows how we can define intervals in multivariate cases. In Section III.3, we present simulation results based on various combinations of marginal distributions and dependence structures. Section III.4 presents an empirical application to international stock markets, and Section III.5 is devoted to portfolio interval forecasts using the LR test as well as the CQFE test. Finally, we conclude the paper in Section III.6.

III.2 Framework

In this section, we describe the conditional coverage test of Christoffersen (1998) used in univariate cases. Next, we define the interval which provides a way to evaluate the interval forecast of multivariate cases.

III.2.A Univariate Case

The indicator variable, I_t, given interval forecast, $(L_{t|t-1}(p), U_{t|t-1}(p))$ for time t, made at time $t-1$, is defined as the following,

$$I_t = \begin{cases} 1, & \text{if } y_t \in \left[L_{t|t-1}(p), U_{t|t-1}(p)\right] \\ 0, & \text{if } y_t \notin \left[L_{t|t-1}(p), U_{t|t-1}(p)\right] \end{cases}$$

A testing criterion is to test $E[I_t|\Psi_{t-1}] = E[I_t|I_{t-1}, I_{t-2}, I_{t-3}, \cdots, I_1] = p$, for all t. And this test is equivalent to testing that the sequence I_t is identically and independently distributed Bernoulli with parameter p, i.e., $I_t \overset{iid}{\sim} \text{Bern}(p)$.

III.2.A.a The LR Test of Unconditional Coverage

The hypothesis that $E[I_t] = p$ should be tested against the alternative $E[I_t] \neq p$, given independence. The likelihood under the null is

$$L(p; I_1, I_2, \cdots, I_T) = (1-p)^{n_o} p^{n_1}$$

and under the alternative

$$L(\pi; I_1, I_2, \cdots, I_T) = (1-\pi)^{n_o} \pi^{n_1}$$

Testing for unconditional coverage can be formulated as a standard likelihood ratio test,

$$LR_{uc} = -2 \log \left[\frac{L(p; I_1, I_2, \cdots, I_T)}{L(\hat{\pi}; I_1, I_2, \cdots, I_T)} \right] \overset{asy}{\sim} \chi^2(s-1) = \chi^2(1)$$

where, $\hat{\pi} = \dfrac{n_1}{n_0 + n_1}$ is the maximum likelihood estimate of π, and $s = 2$ is the number of possible outcomes of the sequence. However, this test does not have any power against the alternative that the zeros and ones come clustered together in time-dependent fashion.

III.2.A.b The LR Test of Independence

Independence is tested against an explicit first-order Markov chain alternative[3]. A binary first-order Markov chain, I_t, with transition probability matrix

$$\Pi_1 = \begin{bmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{bmatrix}$$

where $\pi_{ij} = \Pr[I_t = j | I_{t-1} = i]$. The approximate likelihood function for this process is

$$L(\Pi_1; I_1, I_2, \cdots, I_T) = (1-\pi_{01})^{n_{00}} (\pi_{01})^{n_{01}} (1-\pi_{11})^{n_{10}} (\pi_{11})^{n_{11}}$$

where n_{ij} is the number of observations with value i followed by j. The parameter estimates by MLE are simply ratios of the counts of the appropriate cells, $\hat{\pi}_{01} = \dfrac{n_{01}}{n_{00}+n_{01}}$ and $\hat{\pi}_{11} = \dfrac{n_{11}}{n_{10}+n_{11}}$. Under the null that the sequence is independent, the transition probability matrix becomes

$$\Pi_0 = \begin{bmatrix} 1-\pi_0 & \pi_0 \\ 1-\pi_0 & \pi_0 \end{bmatrix}$$

and corresponds to independence. The likelihood under the null becomes

$$L(\Pi_0; I_1, I_2, \cdots, I_T) = (1-\pi_0)^{(n_{00}+n_{01})} (\pi_0)^{(n_{01}+n_{11})}$$

and the ML estimate is $\hat{\pi}_0 = \dfrac{n_{01}+n_{11}}{n_{00}+n_{10}+n_{01}+n_{11}}$. From Hoel (1954), the LR test of independence is asymptotically distributed as a following

$$LR_{ind} = -2\log\left[\frac{L(\hat{\Pi}_0; I_1, I_2, \cdots, I_T)}{L(\hat{\Pi}_1; I_1, I_2, \cdots, I_T)}\right] \overset{asy}{\sim} \chi^2((s-1)^2) = \chi^2(1)$$

[3]Christoffersen and Diebold (2000) suggest the runs test, which does not depend on the nominal coverage of the intervals, q, and is exact and uniformly most powerful against a first-order Markov alternative. Also, Pesaran and Timmermann (2006) propose a new test of independence based on the maximum canonical correlation between pairs of discrete variables. However, we simply use the likelihood ratio test to be consistent with Christoffersen (1998).

III.2.A.c The Joint Test of Coverage and Independence

The tests for unconditional coverage and independence are now combined to form a complete test of conditional coverage. From Christoffersen (1998), the LR test of conditional coverage is distributed as

$$LR_{cc} = -2\log\left[\frac{L(p; I_1, I_2, \cdots, I_T)}{L(\hat{\Pi}_1; I_1, I_2, \cdots, I_T)}\right] \overset{asy}{\sim} \chi^2(s(s-1)) = \chi^2(2).$$

This, in turn, implies that three LR tests are numerically related by the following identity,

$$LR_{cc} = LR_{uc} + LR_{ind}$$

This LR framework enables joint testing of randomness and correct coverage while retaining the individual hypotheses as subcomponents. See Christoffersen (1998) for more details[4].

III.2.B The Multivariate Case

As an extension of the univariate case, evaluation of multivariate interval[5] forecasts provides no conceptual difficulties. Define the indicator variable, I_t, for a given a sample of an m-variate time series, $\{Y_t\}_{t=1}^{T}$, and a sequence of out-of-sample region forecasts, $\{R_{t|t-1}(q)\}_{t=1}^{T}$, where $R_{t|t-1}(q) \in \mathfrak{R}^m$ for time t, made at time $t-1$, with the prespecified desired coverage, p, of the region as a following,

$$I_t = \begin{cases} 1, & \text{if } Y_t \in R_{t|t-1}(q) \\ 0, & \text{if } Y_t \notin R_{t|t-1}(q) \end{cases}$$

However, if the region is the Cartesian product of m closed intervals as in Christoffersen (1998),

$$R_{t|t-1}(q) = \big\{\mathbf{x}_t \mid \Pr\big[\mathbf{x}_t \in (L_{1,t|t-1}, U_{1,t|t-1}) \times \cdots \times (L_{m,t|t-1}, U_{m,t|t-1})\big] = q\big\}$$

[4]Christoffersen (1998) and Christoffersen and Diebold (2000) also broaden the test methods to allow for multivariate and higher-order dependence using a regression method. There is some usefulness of the regression representation. It can test the null hypothesis of correct conditional coverage by a simple F-test. Second, we can include predictor variables, such as lagged squared returns. Finally, higher-order dependence can be tested via simple inclusion of additional lags of the predictor vairables.

[5]The term, 'interval,' may only be appropriate when we use it for the univariate case. The term, 'area,' may be more adequate to use for the multivariate case. However, we do not distinguish between these two terms in this paper.

where, $\mathbf{x_t}$ is an $(m \times 1)$ vector, the region forecasts could be difficult because computing a joint forecast region, given p, from m individual interval forecasts $\{(L_{n,t|t-1}, U_{n,t|t-1})\}_{n=1}^{m}$ is not a easy task, according to the Bonferroni general inequality, $\Pr\left(\bigcap_{n=1}^{g} A_n\right) \geq 1 - \sum_{n=1}^{g} \Pr\left(\bar{A}_n\right)$, where A_n and its complement $\bar{A}_n$ are any events[6]. However, in a multivariate setting, as seen in Case A in Figure III.1, we can calculate the probability of the region exactly as the following:

$$\begin{aligned} R_{t|t-1}(q) = \Big\{\mathbf{x}_t \mid\ & F(U_{1,t|t-1}, U_{2,t|t-1}) - F(L_{1,t|t-1}, U_{2,t|t-1}) \\ & - F(U_{1,t|t-1}, L_{2,t|t-1}) + F(L_{1,t|t-1}, L_{2,t|t-1}) = q\Big\} \end{aligned}$$

where, $F(\cdot)$ is the cumulative density function (CDF) of $\mathbf{x}_t$.

The choice of region, $\{(L_{n,t|t-1}, U_{n,t|t-1})\}_{n=1}^{m}$, given q, may depend on the objective of the interval forecasts. There are some special cases, as seen in Case B and Case C in Figure III.1. As one of special cases, if $L_{n,t|t-1} = -\infty$ for all n, the region becomes the low-left tail (in a 2-dimension plane), compatible with univariate Value-at-Risk (VaR) applications. If $L_{n,t|t-1} = -\infty$ and $U_{n,t|t-1} = \infty$ for $n = 1, \cdots, i-1, i+1, \cdots, m$, then

$$R_{t|t-1}(q) = \Big\{\mathbf{x}_t \mid \Pr\big[x_{i,t} \in [L_{i,t|t-1}, U_{i,t|t-1}]\big] = q\Big\}$$

which is equivalent to the conditional coverage test of univariate case for ith variable.

To investors who have higher weight on ith asset and are concerned with big losses in their wealth, a choice of high value of $U_{i,t|t-1}$ and relatively small value of $U_{n,t|t-1}$ for $n = 1, \cdots, i-1, i+1, \cdots, m$ would be suitable with $L_{n,t|t-1} = \infty$ for all n. Case C can be used if we are concerned with one market crash given the other market's return in the international market, since there are discrepancies in markets opening in international markets. In this paper, we assume, for simplicity, $L_{n,t|t-1} = L_{t|t-1}$ and $U_{n,t|t-1} = U_{t|t-1}$ for all n.

Computing $\mathbf{x}$, such that $F(\mathbf{x}) = \Pr[\mathbf{X} \leq \mathbf{x}] = \int_{\mathbf{X}\leq\mathbf{x}} f(\mathbf{x})d\mathbf{x} = q$, given $0 \leq q \leq 1$, is computationally burdensome in many cases because of integrals.

[6] If A_n are disjoint sets for all n, then the inequality becomes an equality.

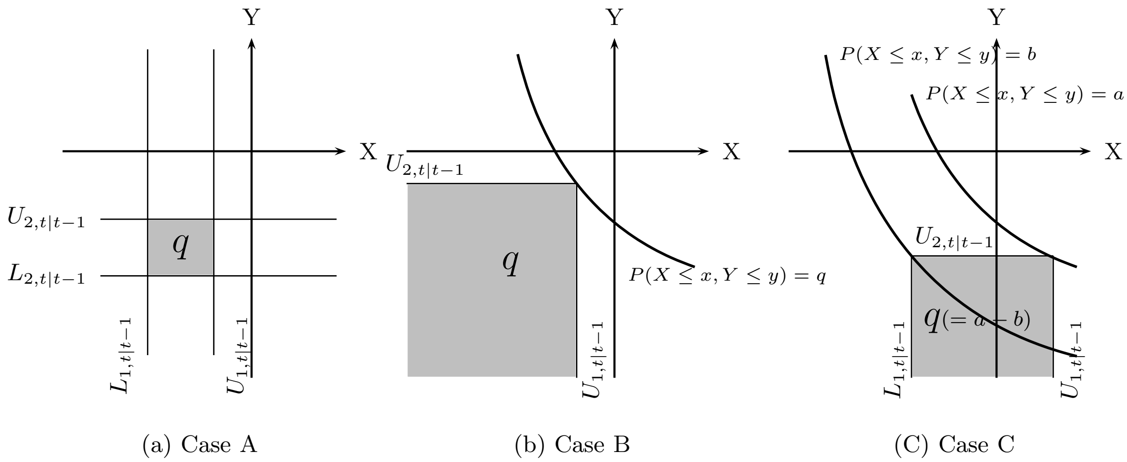


Figure III.1 Comparison of forecasting regions

One example is a multivariate normal distribution, which does not have a closed form cumulative density function (CDF). In this case, $f(\mathbf{x})$ is the probability density function (PDF) of a multivariate normal distribution. However, we can apply the closed form CDF using copulas. Copulas are distribution functions whose one-dimensional margins are uniform. Some copulas, such as the Gumbel, Clayton copula[7], not only have a closed form, but have many additional beneficial properties, including easy construction, asymmetric dependence structure, etc. In this paper, we will apply some copulas to test whether or not asymmetry of the dependence structure is important in terms of multivariate interval forecasting. After defining the indicator function I_t, the testing procedure is identical to the univariate case. Therefore, there is nothing further to be said at the theoretical level.

III.3 Simulation of Multivariate GARCH Process

Now we turn to an empirical application of the conditional coverage test of multivariate cases using simulated data. One of the main points of this paper is to see the importance of dependence structure in multivariate interval forecasting.

[7] These copulas belong to the, so called, Archimedean copula. For more information, see Joe (1997) and Nelsen (1999). Embrechts et al. (2003) provide some applications of copulas in risk management.

To do this, assume a bivariate time series generated by bivariate normal-Gaussian GARCH(1,1) model[8] as the following:

$$\mathbf{y}_t|\Omega_{t-1} \sim N(\mathbf{0}, \Sigma_t)$$
$$\Sigma_t = \begin{pmatrix} h_{1,t} & \rho\sqrt{h_{1,t}}\sqrt{h_{2,t}} \\ \rho\sqrt{h_{1,t}}\sqrt{h_{2,t}} & h_{2,t} \end{pmatrix}$$
$$h_{i,t} = \omega_i + \alpha_i y_{i,t-1}^2 + \beta_i h_{i,t-1}, \text{ for } i = 1, 2$$

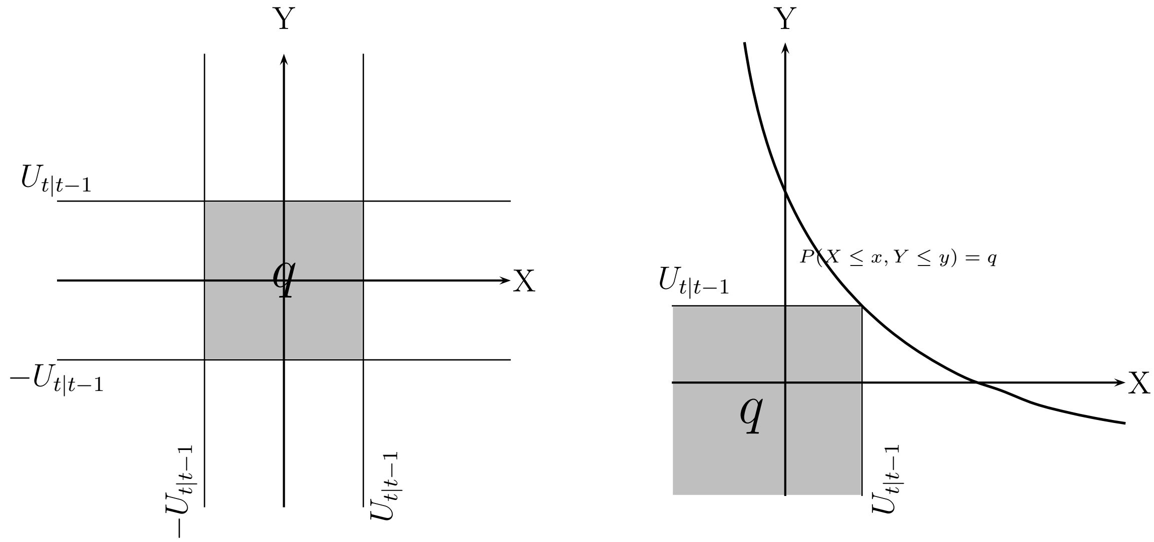
where, $(\rho, \omega_1, \alpha_1, \beta_1, \omega_2, \alpha_2, \beta_2) = (0.30, 0.05, 0.10, 0.85, 0.10, 0.20, 0.70)$. The sample size is 2,000. Using first 1,000 of the sample, we estimate parameters of various models, based on different dependence structure assumptions, so that the margins' assumption is GARCH-Normal(1,1) through all models. In this paper, we use three different models: Normal, Gumbel and Clayton copula. With parameter estimates, we can find $\{(L_{t|t-1}(q), U_{t|t-1}(q))\}_{t=1}^{1000}$ for the remaining 1,000 in the sample. There are two cases to find $(L_{t|t-1}(q), U_{t|t-1}(q))$. The first set of $(L_{t|t-1}(q), U_{t|t-1}(q)) \equiv (-U_{t|t-1}(q), U_{t|t-1}(q))$, given q, is the one that satisfies

$$\begin{aligned} & C_c\left(\Phi\left(\frac{U_{t|t-1}}{\sqrt{h_{1,t}}};\hat{\theta}_1\right), \Phi\left(\frac{U_{t|t-1}}{\sqrt{h_{2,t}}};\hat{\theta}_2\right);\hat{\theta}_c\right) - C_c\left(\Phi\left(\frac{U_{t|t-1}}{\sqrt{h_{1,t}}};\hat{\theta}_1\right), \Phi\left(\frac{-U_{t|t-1}}{\sqrt{h_{2,t}}};\hat{\theta}_2\right);\hat{\theta}_c\right) \\ & - C_c\left(\Phi\left(\frac{-U_{t|t-1}}{\sqrt{h_{1,t}}};\hat{\theta}_1\right), \Phi\left(\frac{U_{t|t-1}}{\sqrt{h_{2,t}}};\hat{\theta}_2\right);\hat{\theta}_c\right) + C_c\left(\Phi\left(\frac{-U_{t|t-1}}{\sqrt{h_{1,t}}};\hat{\theta}_1\right), \Phi\left(\frac{-U_{t|t-1}}{\sqrt{h_{2,t}}};\hat{\theta}_2\right);\hat{\theta}_c\right) = q \end{aligned}$$

where, $\hat{\theta}_i$ is the parameter estimates of ith variable for $i = 1, 2$, and $\hat{\theta}_c$ is the parameter estimates of the dependence structure of C_c, where a subscript c is a copula model index. The second case is the low-left tail case, so that $L_{t|t-1} = -\infty$ and $U_{t|t-1}$ is the one satisfying $C_c\left(\Phi\left(\frac{U_{t|t-1}}{\sqrt{h_{1,t}}};\hat{\theta}_1\right), \Phi\left(\frac{U_{t|t-1}}{\sqrt{h_{2,t}}};\hat{\theta}_2\right);\hat{\theta}_c\right) = q$, given q[9]. Figure III.2 shows the difference between cases. As seen in Figure III.2, Case

[8] In terms of copula, this means that the dependence structure is normal copula and the margins are Gaussian GARCH(1,1).

[9] In some cases, such as multivariate normal, finding $L_{t|t-1}$ or $U_{t|t-1}$ incurs a computation burden, since there is no closed form CDF. In this paper, we apply an algorithm by Drezner and Wesolowsky (1990) to find $L_{t|t-1}$ or $U_{t|t-1}$ with a reduced computation burden.



(a) Case 1 (b) Case 2

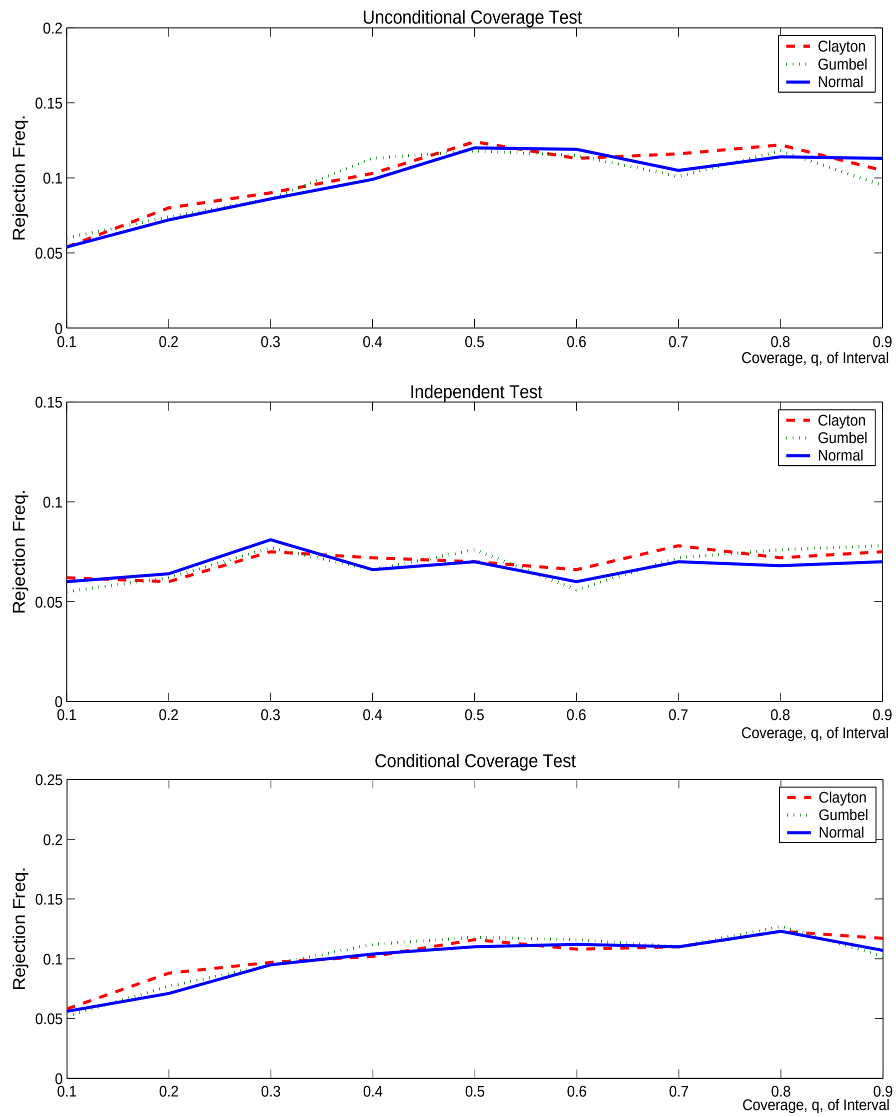
Figure III.2 Difference between Cases

1 could be called a body-interval and Case 2 is low-tail interval case used in this simulation section.

Finally, given $\{(L_{t|t-1}, U_{t|t-1})\}_{t=1}^{1000}$, we can test the conditional coverage test to see whether the dependence structure is important by the Monte Carlo simulation. The number of Monte Carlo replications is 1,000. Figures III.3 and III.4 show the power of the tests to reject the different dependence interval forecasts when the data are generated by a Normal dependence structure.

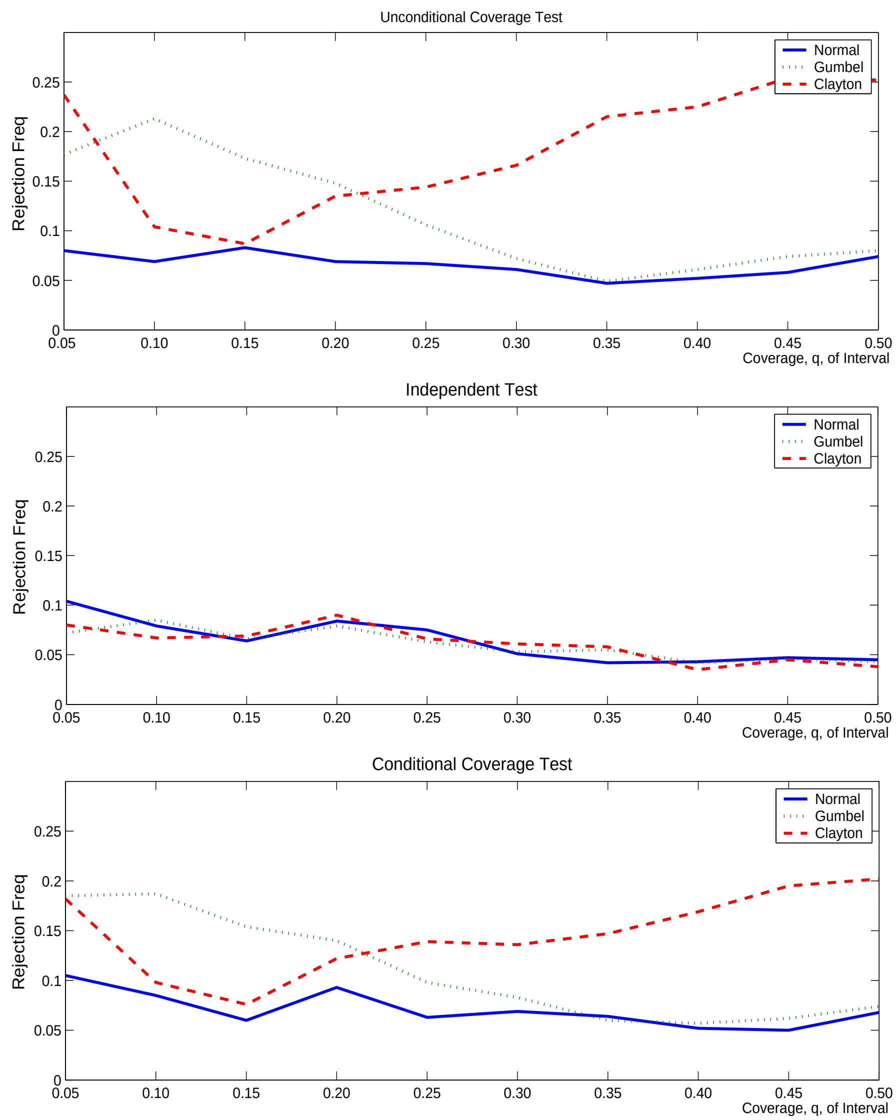
As seen in Figures III.3 and III.4, the power of the tests is not high enough to reject the alternative dependence structure outright[10]. Especially, the power of the test in the case 1 is almost the same as the size of the test. In Case 2, the power of the Gumbel dependence interval forecasts is the highest when q is small. This might be because of the definition of the interval. In a body-interval case, as q gets bigger, the area covered by this test gets larger toward the low-left tail and upper-right tail. However, the coverage area still includes the body parts. Inclusion

[10]Berkowitz (2001) shows the low power of the Christoffersen (1998) Bernoulli test in the VaR case. The LR test suggested by Berkowitz (2001) based on the truncated normal distribution would do a better job to evaluate the interval forecasting with the higher power. Also, we may apply a loss function evaluation method. Lopez (1999) also indicates that an alternative evaluation method using loss function based on probability forecasts is only as capable of differentiating between forecasts from accurate and inaccurate models as the other methods.



Note: This graph shows the power of the LR tests to reject the different dependence interval forecasts of Case 1, when the data are generated by a Normal dependence structure. The solid lines are the Type I error. Significance level of the test is 5%. The top graph is the unconditional coverage LR test. The middle graph is the independence LR test. And the bottom graph is the conditional coverage LR test.

Figure III.3 Power of Tests against Normal Dependence Structure of Case 1



Note: This graph shows the power of the LR tests to reject the different dependence interval forecasts of Case 1, when the data are generated by a Normal dependence structure. The solid lines are the Type I error. Significance level of the test is 5%. The top graph is the unconditional coverage LR test. The middle graph is the independence LR test. And the bottom graph is the conditional coverage LR test.

Figure III.4 Power of Tests against Normal Dependence Structure of Case 2 of body parts can make it a little bit difficult to differentiate the asymmetry in the dependence structure. And Case 2 still covers the low-left tail from lower q, so the power of the test is higher with a lower q, generally. Since we assume the GARCH-Normal(1,1) process for both the generating and estimating series, there is nothing to say about the margin effect nor the independence. This may imply that it is very hard to distinguish the difference of the interval forecasting in the different dependence structures if margins are correctly specified. The next section is devoted to interval forecasting of international stock markets.

III.4 Application: International Stock Markets

In this section, we apply the LR tests proposed by Christoffersen (1998) to the international stock markets for multivariate cases. We employ weekly stock market index returns[11] of U.S. (S&P 500), France (FTSE100), and Japan (Nikkei225). The weekly return is calculated using Wednesday closing prices. If Wednesday is not a trading day, then we use Tuesday, Thursday, Monday, and so on. From January 1984 to March 2007, we have a total of 1209 observations. Table III.1 summarizes the descriptive statistics of the series. As evident in the Table III.1, the stock market returns are strongly not-normal and strongly skewed with fat tails.

The experiment entails, first, estimating the model parameters necessary to form the forecasts on the first 626 observations (from January 1984 to December 1995). Then we calculated $\{(L_{t|t-1}(q), U_{t|t-1}(q))\}_{t=627}^{1209}$ based on the fixed estimates, as seen in Figure III.2. Finally we do out-of-sample interval forecasting and forecast evaluation on the last 583 observations. In this application, we use various underlying models. Recall that the main purpose of this paper is to check whether the dependence structure is important in multivariate interval forecasting. Therefore, we use various combinations of margins (the GARCH-Normal(1,1) and

[11] Some mismatches of trading dates exist among international stock markets. This is the reason why, instead of daily returns, we employ weekly returns.

Table III.1 Descriptive Statistics

	S&P 500	FTSE 100	Nikkei 225
Mean	0.0018	0.0015	0.0004
Median	0.0032	0.0027	0.0023
Max	0.1018	0.1359	0.1214
Min	-0.1666	-0.1782	-0.1089
St. Dev.	0.0214	0.0225	0.0281
Skewness	-0.6443	-0.6292	-0.2386
Kurtosis	7.7250	10.5710	4.6221
J-B stat.	1202.03**	2954.01**	142.74**

Note: ** denotes 1% significance. The sample period is 1984:01-2007:03.

GARCH-t(1,1)) and copulas (the Normal, Gumbel and Clayton copulas)

$$\begin{aligned}\mathbf{r}_t &= \boldsymbol{\alpha} + \boldsymbol{\varepsilon}_t \\ &= \boldsymbol{\alpha} + \sqrt{\boldsymbol{h}_t} \cdot \boldsymbol{v}_t\end{aligned}$$

$$\sqrt{\boldsymbol{h}_t} \cdot \boldsymbol{v}_t | \Omega_{t-1} \sim C_c \left(F_m \left(\frac{v_{1,t}}{\sqrt{h_{1,t}}}, \theta_1 \right), \cdots, F_m \left(\frac{v_{N,t}}{\sqrt{h_{N,t}}}, \theta_N \right); \theta_c \right)$$

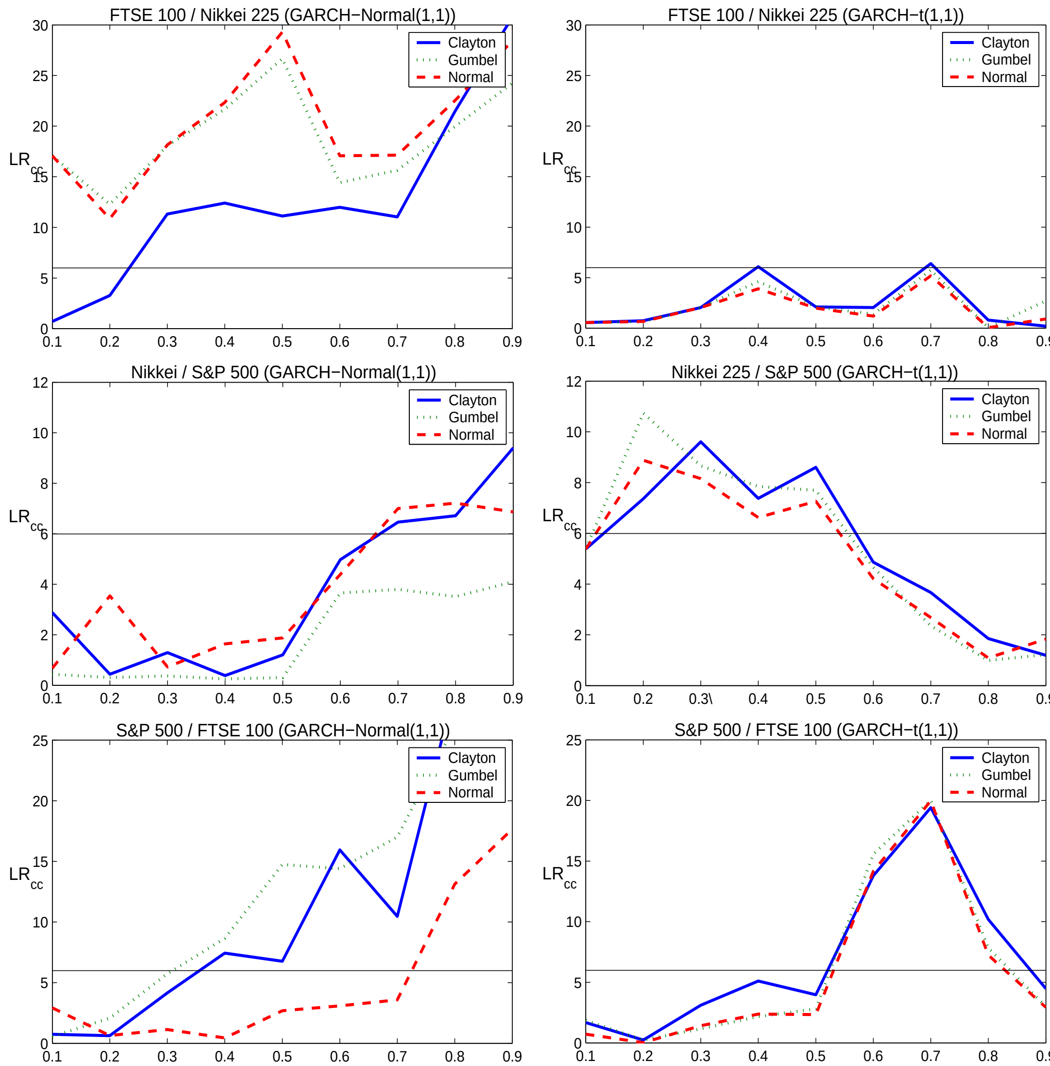
$$h_{i,t} = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{i,t-1}, \text{ for } i = 1, \cdots, N$$

where, $F_m(\cdot)$ is the CDF of normalized (standardized) margins, where a subscript m is a margin index. And the bold characters are $(N \times 1)$ vectors[12]. In this application, N is 2. The results of the testing are presented in Sections III.4.A and III.4.B.

III.4.A Case 1: Body-Interval

Figure III.5 shows the LR$_{cc}$ test statistics of many cases. This body-interval is based on Case 1 in Figure III.2, so first, we find $U_{t|t-1}$ satistying $C_c(U_{t|t-1}, U_{t|t-1}) - C_c(U_{t|t-1}, -U_{t|t-1}) - C_c(-U_{t|t-1}, U_{t|t-1}) + C_c(-U_{t|t-1}, -U_{t|t-1}) = q$, given q and $C_c(\cdot)$. Then we can calculate the indicator variable. Each panel in Figure III.5 represents different combinations of three international stock markets. Clearly, the margin specification is important. In this application, we use

[12] We do not include the lagged returns terms, such as $\mathbf{r}_{t-1}$, in this specification. This might be fruitful for future research.



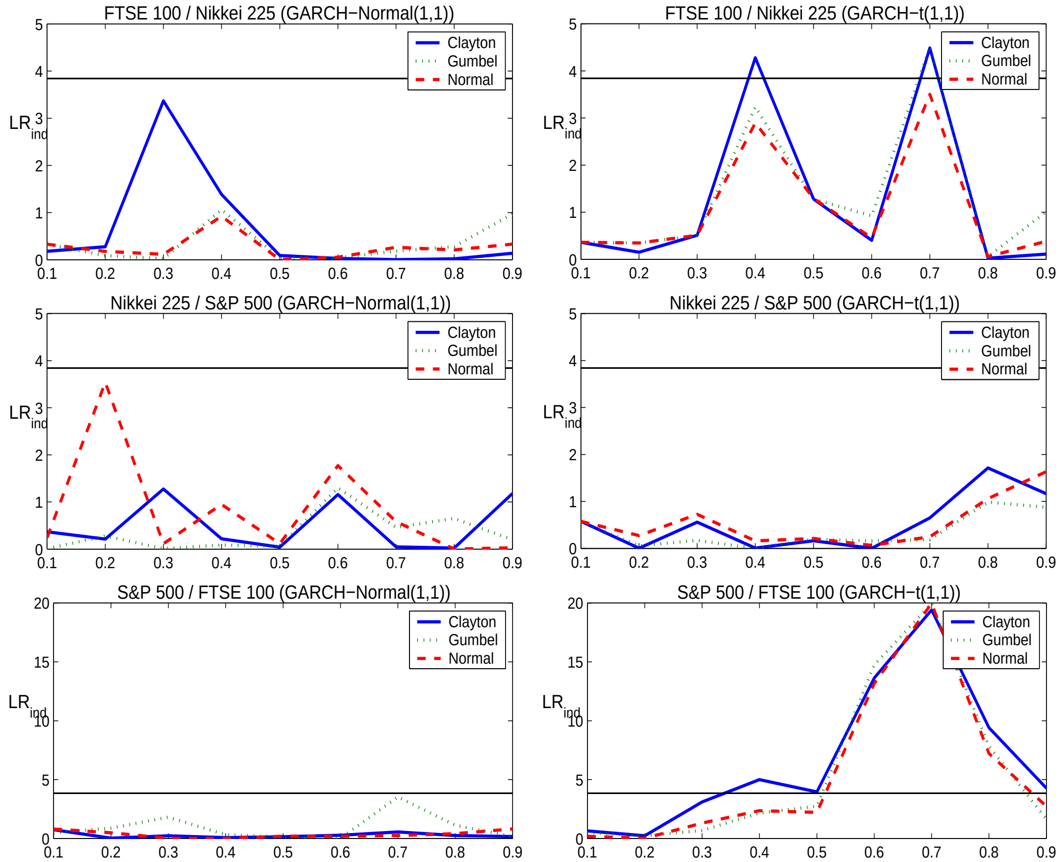
Note: The left column is GARCH-Normal(1,1) margin forecasts of different copula assumptions, and the right column is GARCH-t(1,1) margin forecasts of different copula assumptions. The solid line is the Clayton, the dashed line is the Gumbel, and the dotted line is the Normal copula. The solid horizontal line represents the 5 per cent significance level of the appropriate $\chi^2(2)$ distribution. The test values are plotted for coverages ranging between 10 and 90 per cent.

Figure III.5 LR$_{cc}$ Statistics of Case 1 two different margins. The first one is the GARCH-Normal(1,1) assumption (left column) and the second one is the GARCH-t(1,1) (right column).

In the FTSE100/Nikkei225 case, forecasts based on the GARCH-Normal (1,1) margin assumption do not pass the conditional coverage test in most cases. The GARCH-t(1,1) margins, regardless of the copula assumptions, pass the tests in all coverage rates. However, in the Nikkei225/S&P500, we can see the mixed test results. The GARCH-Normal assumption passes the tests with lower coverage rates regardless of the copula assumptions, while The GARCH-t assumption passes the tests with higher coverage rates. In the S&P500/FTSE100 case, the GARCH-t assumption passes the tests except when q is 60%, 70%, and 80%. This result implies that if we are interested in the partial model misspecification, we can conclude that a universally well-specified model may not exist. Also, recall that when q becomes higher, the coverage area becomes bigger including tail parts. The test results show that the GARCH-t assumption passes the conditional coverage test, while the GARCH-Normal does not. Therefore, the outskirts of joint density seem to fit with the fat tailed model.

Another important finding is that the test results seem to depend on the marginal distribution assumption rather than the dependence structure. We observe similar test statistics under the same marginal distribution. Therefore, no dependence structrue seems to be dominant over others, in general. These results are consistent with the simulation results in the previous section. From the previous section, we observed that it is hard to differentiate the performance of interval forecasting from different dependence structures. One implication of this application is that the most important part to model for body-interval forecasts is the marginal distributions, not the joint dependence structure. Now we turn our attention to see what causes the conditional coverage tests to be rejected in some cases.

Figure III.6 shows the LR_{ind} statistics. As evident in the Figure III.6, the LR_{ind} statistics show that the independence tests are passed across the joint



Note: The left column is GARCH-Normal(1,1) margin forecasts of different copula assumptions, and the right column is GARCH-t(1,1) margin forecasts of different copula assumptions. The solid line is the Clayton, the dashed line is the Gumbel, and the dotted line is the Normal copula. The solid horizontal line represents the 5 per cent significance level of the appropriate $\chi^2(1)$ distribution. The test values are plotted for coverages ranging between 10 and 90 per cent.

Figure III.6 LR$_{ind}$ Statistics of Case 1

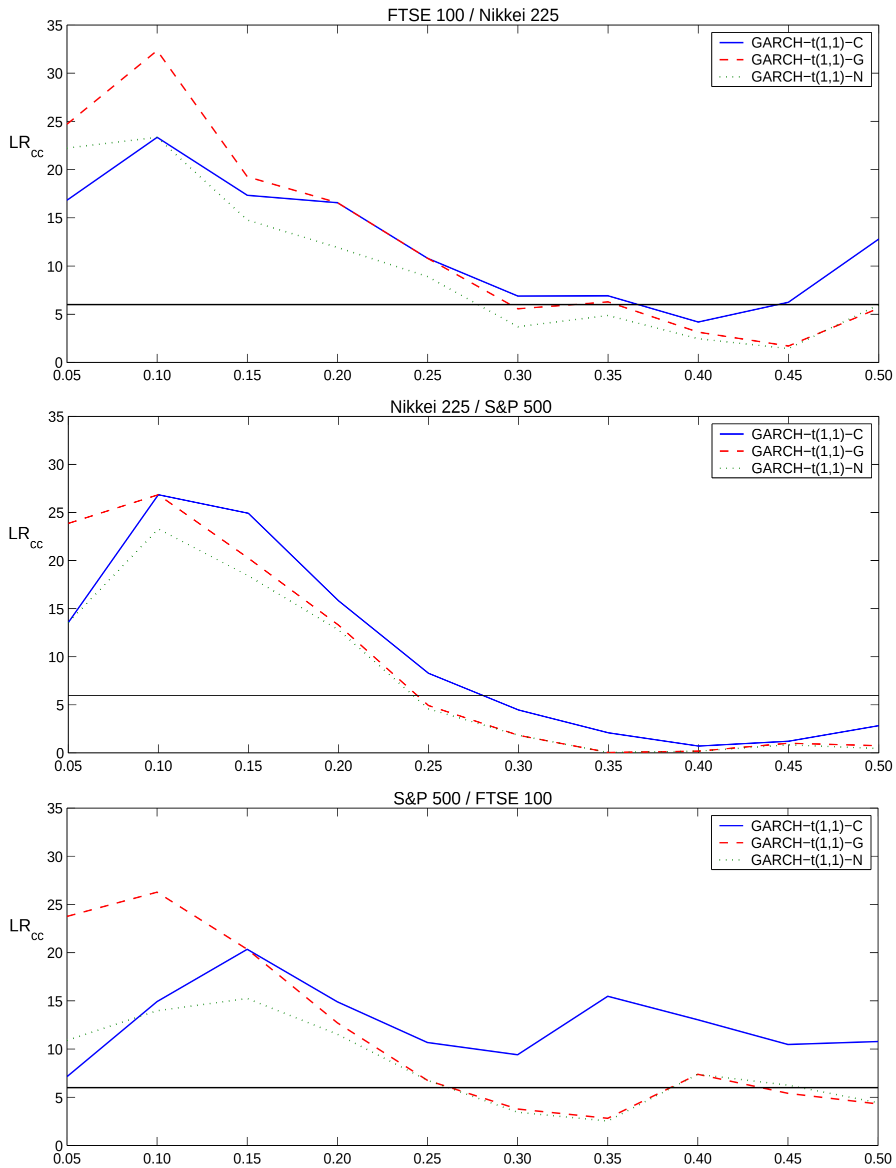
dependence structures and across country pairs, except a few cases in which q is 60%, 70%, and 80% under the GARCH-t assumption. Recall that LR$_{cc}$ statistic is the sum of LR$_{ind}$ and LR$_{uc}$. And LR$_{ind}$ and LR$_{uc}$ statistics follow $\chi^2(1)$ in the bivariate case. Therefore, this result implies that the rejection of unconditional coverage, that is, misspecified models, lead to rejection of the complete test of conditional coverage according to the decomposition property of these statistics. Therefore, we may need to model with caution in order to get the univerally well-specified model.

III.4.B Case 2: Low-tail Interval

Figure III.7 shows the low-tail interval case in Case 2 of Figure III.2. This case is the one which most people are interested in, similar to VaR in the univariate case. We do not include the LR_{cc} statistics under GARCH-Normal(1,1) margins, since they do not pass the test in all cases with large statistic values. However, unlike the body-interval forecast cases, the complete conditional coverage test results show that the low-tail interval forecasts fail the test when q is small, even under GARCH-t(1,1) margin assumption. Also, this failure prevails across the joint dependence structures and across the country pairs. There may be a serious loss of information from having only a low-left tail (one-sided) interval forecast when volatility dynamics are present. This problem will be discussed briefly later. However, when q gets bigger, we can see that it tends to pass the tests. Yet, there is no clear evidence that one among the many dependence structures dominates others in conditional coverage tests.

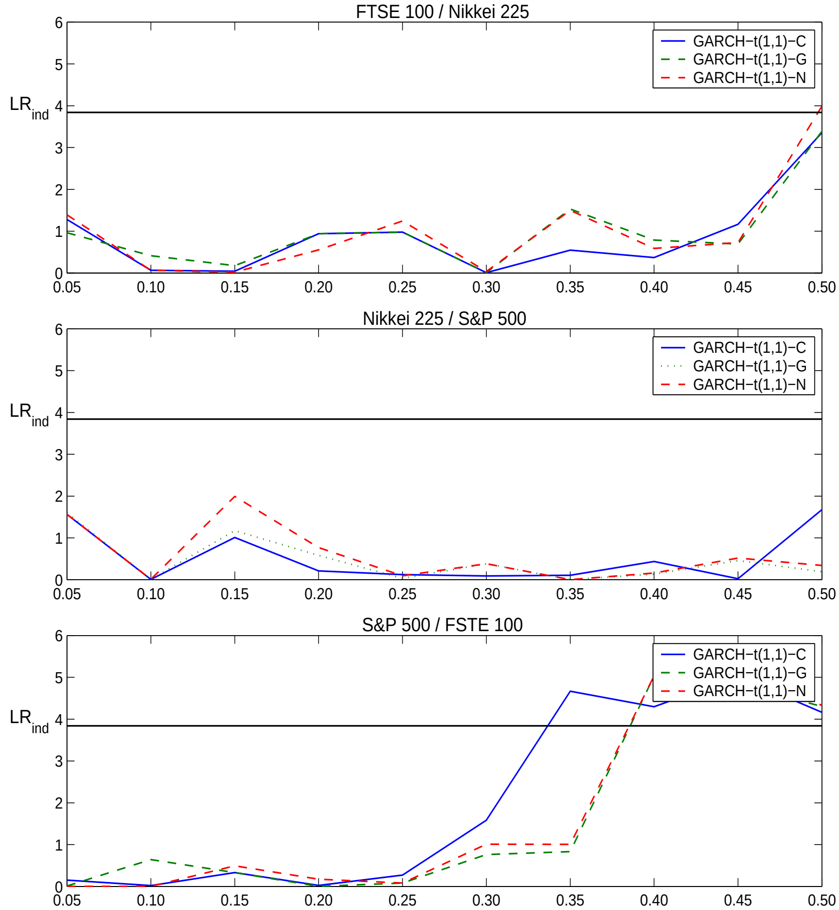
This conflicts somewhat with previous literature. Longin and Solnik (2001) find that international stock markets are more correlated in bear markets. Ang and Chen (2002) and Ang et al. (2006) find that stock return correlation increases during market downturns, and that downside correlations are related to standard size and value factors, as well as momentum. Generally, the Clayton or Gumbel copula can capture the asymmetry dependence, so it is adequate to model more probability mass to events in the lower-left or upper-right tail. However, the results of this paper suggest that none of these models are able to do a better job in the low-left tail interval forecasting, at least our weekly return samples. If, for example, the Clayton copula is dominant over others in the low-left tail interval forecasts, we may say that asymmetry in dependence is a very important factor and that there is a stronger correlation in downside markets. In this paper, we can not find strong evidence that dependence structure is important in the interval forecasts, which is consistent with the low power of simulation results.

Figure III.8 shows the decomposition of the complete conditional coverage



Note: The lines are GARCH-t(1,1) margin forecasts of different copula assumptions. The solid line is the Clayton, the dashed line is the Gumbel, and the dotted line is the Normal copula. The solid horizontal line represents the 5 per cent significance level of the appropriate $\chi^2(1)$ distribution. The test values are plotted for coverages ranging between 10 and 90 per cent.

Figure III.7 LR$_{cc}$ Statistics of Case 2



Note: The solid line is the Clayton, the dashed line is the Gumbel, and the dotted line is the Normal copula. The solid horizontal line represents the 5 per cent significance level of the appropriate $\chi^2(1)$ distribution. The test values are plotted for coverages ranging between 10 and 90 per cent.

Figure III.8 LR$_{ind}$ Statistics under GARCH-t(1,1) margins

tests by LR$_{ind}$ statistics. As in Case 1, the dependence tests are passed almost everywhere, especially when q is small. Therefore, it is clear that the failure of the unconditional coverage tests leads to the failure of the complete conditional coverage tests.

We may suggest some reasons for this failure. First, the sample size of this test is relatively small. The test statistics used in this paper are derived from their asymptotic distribution. As seen in Figures III.3 and III.4, the probability of the type I error is a little bit higher. Therefore, it is important to understand the finite sample properties of these test[13]. Second is model misspecification. If there is

[13]Finding an answer to this question is outside the scope of this paper, so we do not perform any practice to

room for improving in specifying the dependence structure, for example, using the empirical quantiles, as is done in Engle and Gonzalez-Rivera (1991) for univariate cases and Fermanian and Scaillet (2003) for multivariate cases, performance of interval forecasts might be better. Other possible improvement can be achieved by including the dynamics dependence structure specification, as suggested by Engle and Sheppard (2001), Engle (2002) and Patton (2006). The last one is about extreme quantiles. Extreme quantiles are very sensitive to the few observations in the tails of the empirical distribution of the sample. In multivariate cases, we define extreme quantiles as the low-left tail of distributions. However, if we decompose multivariate variables into each single variable, extreme quantiles are the left end or the low end of density. This discrepancy of defining extreme quantiles could cause the failure of the tests.

III.5 Portfolio Interval Evaluation

In the previous section, none of dependence structures pass the test in Case 2 with smaller coverage rate, q. This might be due to model misspecification. Now, we turn our attention to the portfolio interval evaluation. Construction of portfolio may make the tail of portfolio density fatter, depending on the dependence structures. In general, the linear combinations of normally distributed random variables are normally distributed. However, this statement is only true when the dependence structure is normal. In order to check it, we perform simple simulation. First, we generate three bivariate series with different dependence assumptions as the following:

$$\mathbf{r}_t = \sqrt{\boldsymbol{h}_t} \cdot \boldsymbol{v}_t$$

$$\sqrt{\boldsymbol{h}_t} \cdot \boldsymbol{v}_t | \Omega_{t-1} \sim C_c \left(F_m \left(\frac{v_{1,t}}{\sqrt{h_{1,t}}}, \theta_1 \right), F_m \left(\frac{v_{2,t}}{\sqrt{h_{2,t}}}, \theta_2 \right); \theta_c \right)$$

$$h_{i,t} = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{i,t-1}, \text{ for } i = 1, 2$$

answer this question.

where, $(\omega_1, \alpha_1, \beta_1, \omega_2, \alpha_2, \beta_2) = (0.05, 0.70, 0.15, 0.05, 0.30, 0.40)$ and F_m is the CDF of the GARCH-Normal(1,1). θ_c is set to 0.3000 for the Normal, to 0.4813 for the Clayton, and to 1.2407 for the Gumbel copula[14]. Then, construct the equal weighted portfolios. Finally we estimate parameters using the GARCH-t(1,1) model and repeat 1,000 times. The median values of ν parameter estimates, representing the thickness of tails, are 27.21, 25.72, and 115.72 for the Clayton, Gumble, and Normal copulas, respectively. The tail of portfolio constructed from assymetric dependence structures seems to be thicker than normal dependence structure. This result implies that the dependence structures could alter the tail behaviors of portfolios even with the same margins[15].

In this setion, first, we perform the LR tests of the portfolio interval evaluation at moderate coverage rates, $0.05 \leq q \leq 0.5$. Next, we focus on the general VaR setting, q is 1%, 2.5%, and 5%. In this section, we include the conditional quantile forecasts test by Giacomini and Komunjer (2005) to see whether one forecasts outperforms the other forecasts as well as the general LR test.

III.5.A LR test at $0.05 \leq q \leq 0.5$

Portfolio construction reduces the multivariate interval evaluation problem to the univariate problem. There is some literature on interval evaluation of portfolios. Christiansen (1999) presents a methodology for calculating the VaR of portfolios of Danish zero-coupon bonds using the factor-ARCH model. Lopez and Walter (2001) evaluate the relative accuracy of different covariance matrix forecasts using a foreign exchange portfolio under a VaR framework. Generally, the portfolio return is $r_{p,t} = \sum_{i=1}^{N} \omega_i r_{i,t}$, where $\sum_{i=1}^{N} \omega_i = 1$. In this paper, we construct equal weighted portfolios, so $\omega_i = \frac{1}{N}$. Table III.2 shows the descriptive statistics of the equal weighted portfolios. By constructing portfolios, in general, we observe

[14] These value have the same Kendall's rank correlation.

[15] We can easily show that $Y \sim N(\mu, \Sigma)$ *if and only if* for each $b \equiv (b_1, \cdots, b_m)' \in R^m$, $b'Y$ has a univariate normal distribution with mean $b'\mu$ and variance $b'\Sigma b'$. Therefore, by contrapositive, we know that if Y is not distributed as $N(\mu, \Sigma)$, then $b'Y$ is not distributed normally, for some b.

higher negative skewness with smaller median values of each index, compared to Table III.1 of each index.

Table III.2 Descriptive Statistics of Equal Weighted Portfolio

	SP500/FTSE100	FTSE100/Nikkei225	Nikkei225/SP500	All 3 indexes
Mean	0.0016	0.0010	0.0011	0.0012
Median	0.0029	0.0021	0.0020	0.0025
Max	0.1096	0.0747	0.0730	0.0776
Min	-0.1724	-0.1425	-0.1367	-0.1505
St. Dev.	0.0199	0.0209	0.0207	0.0191
Skewness	-0.8698	-0.7092	-0.5993	-0.8657
Kurtosis	11.0630	6.3380	5.7687	8.0518
J-B stat.	3412.52**	658.86**	455.69**	1429.49**

Note: ** denotes 1% significance. The sample period is 1984:01-2007:03.

Now we can do the LR test of each portfolio as a univariate case. Model specification is the following:

$$\begin{aligned} r_{p,t} &= \mu_p + \varepsilon_{p,t} \\ &= \mu_p + \sqrt{h_{p,t}} \cdot v_{p,t} \end{aligned}$$
$$h_{p,t} = \omega_p + \alpha_p \varepsilon_{p,t-1}^2 + \beta_p h_{p,t-1}$$

We assume that the distributions of $v_{p,t}$ are standard normal and standardized t. Like the multivariate case, the in-sample period is from Jan. 1984 to Dec. 1995. We report the parameter estimates of the models in Table III.3. One notable result is that the log-likelihood values of GARCH-t(1,1) are improved, having $\hat{\nu}_p$ values of 5.6 to 7.4 . Therefore, we can select the GARCH-t(1,1) in terms of information criteria such as AIC or BIC. Next, we apply these models and parameter estimates to interval forecasting.

Figure III.9 shows the LR test results using one-sided quantiles, following the procedure by Christoffersen (1998). According to Figure III.9, construction of portfolios helps to pass the conditional coverage tests under different conditional variance assumptions in the highest number of cases, although we fail to pass the tests in the multivariate cases under various dependence structure assumptions. If we only are concerned with the interval forecasting of constructed portfolio,

Table III.3 Parameter Estimates of Equal Weighted Portfolios

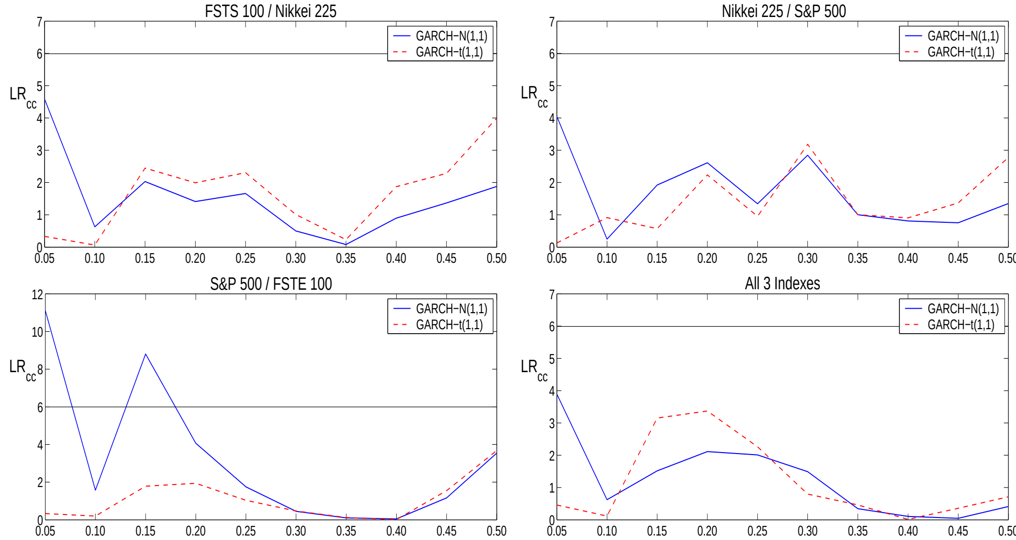
	SP500/FTSE100	FTSE100/Nikkei225	Nikkei225/SP500	All 3 indices
Panel A. GARCH-Normal (1,1)				
μ_p	0.0029 (0.0006)	0.0024 (0.0007)	0.0023 (0.0007)	0.0024 (0.0007)
ω_p	0.0001 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)
α_p	0.3048 (0.0000)	0.8560 (0.0052)	0.8844 (0.0021)	0.8700 (0.0012)
β_p	0.4112 (0.0768)	0.1337 (0.0000)	0.1020 (0.0021)	0.1215 (0.0010)
LL	1649.57	1607.12	1614.12	1667.22
Panel A. GARCH-t (1,1)				
μ_p	0.0029 (0.0006)	0.0029 (0.0001)	0.0028 (0.0007)	0.0027 (0.0006)
ω_p	0.0001 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)
α_p	0.7036 (0.1228)	0.8024 (0.0343)	0.8725 (0.0001)	0.8246 (0.0509)
β_p	0.1123 (0.0579)	0.1516 (0.0439)	0.1019 (0.0025)	0.1197 (0.0403)
ν_p	7.4012 (1.4838)	7.4468 (2.3098)	5.6521 (0.0002)	7.2623 (1.6312)
LL	1674.13	1630.34	1639.16	1698.48

the misspecification problem of joint interval forecasts becomes negligible. Also, interval forecasting results does not show the dominance of the GARCH-t(1,1) model. The results show that the LR tests are passed in both models, implying that there is little difference between the two models in the interval forecasting, unlike the model selection problem. In the section III.4, we observe that there are noticeable differences between the GARCH-Normal(1,1) and the GARCH-t(1,1) assumptions as margins. However, these differences are mitigated in the portfolio interval forecasting problem, when the coverage rates are between 5% and 50%. The observation made in Christoffersen (1998) that the Gaussian innovation assumption fails at small coverage rates, as well as at large coverage rates, does not hold in this portfolio application.

III.5.B VaR Evaluation of Equal Weighted Portfolios

Now, we turn our attention to the general VaR setting. In the general VaR setting, the interesting quantiles are 1%, 2.5%, and 5%[16]. In this section,

[16] In the VaR literature, popular confidence levels usually are 99% and 95%, but we include 2.5% confidence level.



Note: The solid line is the GARCH-Normal(1,1), and the dashed line is the GARCH-t(1,1). The solid horizontal line represents the 5 per cent significance level of the appropriate $\chi^2(2)$ distribution. The test values are plotted for coverages ranging between 5 and 50 per cent.

Figure III.9 LR$_{cc}$ Statistics of Equal Weighted Portfolios

first, we perform the LR test by Christoffersen (1998). Then, we perform the conditional quantile forecasts encompassing (CQFE) test by Giacomini and Komunjer (2005) for comparing conditional quantile forecasts in an out-of-sample framework. The general idea of the encompassing test is to test whether forecast $\hat{q}_{1,t}$ (or $\hat{q}_{2,t}$) encompasses $\hat{q}_{2,t}$ (or $\hat{q}_{1,t}$). The test statistics are

$$ENC1_n = n\left((\hat{\theta}_{1n}, \hat{\theta}_{2n}) - (1,0)\right)\hat{\Omega}_n^{-1}\left((\hat{\theta}_{1n}, \hat{\theta}_{2n}) - (1,0)\right)'$$
$$ENC2_n = n\left((\hat{\theta}_{1n}, \hat{\theta}_{2n}) - (0,1)\right)\hat{\Omega}_n^{-1}\left((\hat{\theta}_{1n}, \hat{\theta}_{2n}) - (0,1)\right)'$$

where $(\hat{\theta}_{1n}, \hat{\theta}_{2n})$ is the GMM estimates of forecasting combination and $\hat{\Omega}_n$ is some consistent estimate of $\Omega \equiv (\boldsymbol{\gamma}' S^{-1} \boldsymbol{\gamma})^{-1}$. Then under H_{10}: forecast $\hat{q}_{1,t}$ encompasses $\hat{q}_{2,t}$, $ENC1_n \xrightarrow{d} \chi^2(2)$, as $n \longrightarrow \infty$, and under H_{1a}: forecast $\hat{q}_{1,t}$ does not encompasses $\hat{q}_{2,t}$, $ENC1_n \longrightarrow +\infty$, as $n \longrightarrow \infty$. This holds for H_{20} and H_{2a}. For more information, see Giacomini and Komunjer (2005).

In this application, we include one more forecast: the conditional quantile forecasts based on the quantile regressions as the following

$$Q_{r_{p,t}}(q|\sigma_{p,t}) = \mu_q + \beta_{q,1}\sigma_{p,t}$$

Table III.4 Estimation of Quantile Regression

	$q = 1\%$		$q = 2.5\%$		$q = 5\%$	
	μ_q	$\beta_{q,1}$	μ_q	$\beta_{q,1}$	μ_q	$\beta_{q,1}$
S&P500/	-2.0676	-1.1051	-1.0646	-1.2060	-0.2882	-1.2842
FTSE100	(7.32E-05)	(4.57E-06)	(4.83E-05)	(1.17E-05)	(6.85E-05)	(1.30E-05)
FTSE100/	-0.6994	-2.3289	-0.2165	-1.9071	0.1547	-1.6476
Nikkei225	(2.91E-06)	(2.45E-06)	(5.88E-05)	(1.78E-05)	(0.0002)	(0.0001)
Nikkei225/	-2.4563	-1.2203	-0.9347	-1.4573	-0.7465	-1.3033
S&P500	(0.0002)	(0.0001)	(4.58E-05)	(1.11E-05)	(0.0008)	(0.0004)
All 3	-0.8731	-1.9024	0.0293	-1.9095	0.1034	-1.5604
indices	(0.0007)	(0.0003)	(3.78E-05)	(1.19E-05)	(0.0004)	(0.0004)

Note: These are the parameter estimates of quantile regression. μ_q is a constant and $\beta_{q,1}$ is a coefficient of the conditional volatility based on BARCH-Normal(1,1).

where $\sigma_{p,t}$ is the conditional volatility based on GARCH-Normal(1,1). Estimating of $\theta_q = [\mu_q, \beta_q]'$ is based on Komunjer (2005). Komunjer (2005) proposed estimators which belong to the family of quasi-maximum likelihood estimators and are based on a new family of densities which we call 'tick-exponential.' The tick-exponential family is

$$\varphi_t^q(y, \eta) = \exp\left(-(1-q)[a_t(\eta) - b_t(y)]1_{y\leq\eta} q[a_t(\eta) - c_y(\eta)]1_{y>\eta}\right),$$

and we set $a_t(\eta) = \dfrac{1}{q(1-q)}\eta$ and $b_t(y) = c_t(y) = \dfrac{1}{q(1-q)}y$ as a special case. In this case, $\ln \varphi_t^q$ is proportional to the 'tick' function, known in the literature as the asymmetrical slope or check function. Table III.4 summarizes the conditional quantile regression estimation.

The LR test results of Christoffersen (1998) are summarized in Table III.5. As seen in Figure III.9, the GARCH-Normal(1,1) assumption seems to be good to pass the tests, when q is higher. However, when an interesting interval is an extreme low-tail, such as 1% or 2.5%, the GARCH-Normal(1,1) fails the conditional coverage test across portfolios, while the GARCH-t(1,1) assumption passes the test with out of sample empirical coverage rates, $\hat{q}$, which are very

Table III.5 Conditional Coverage Test - Out of Sample

	$q = 1\%$		$q = 2.5\%$		$q = 5\%$	
	$\hat{q}$	LR_{cc}	$\hat{q}$	LR_{cc}	$\hat{q}$	LR_{cc}
Panel A. GARCH-Normal (1,1)						
S&P500/FTSE100	0.0377	28.3360*	0.0463	11.3890*	0.0823	11.1220*
FTSE100/Nikkei225	0.0206	5.5784	0.0412	7.3456*	0.0703	4.5837
Nikkei225/S&P500	0.0257	10.9830*	0.0480	10.1610*	0.0669	4.0476
All 3 indices	0.0274	13.0870*	0.0429	8.6028*	0.0686	3.8933
Panel B. GARCH-t (1,1)						
S&P500/FTSE100	0.0172	2.8458	0.0292	1.4249	0.0549	0.3305
FTSE100/Nikkei225	0.0172	2.8458	0.0274	1.0483	0.0549	0.3305
Nikkei225/S&P500	0.0034	3.4065	0.0206	0.9923	0.0480	0.1261
All 3 indices	0.0120	0.3974	0.0326	2.5579	0.0532	0.4569
Panel C. Quantile Regression (QS)						
S&P500/FTSE100	0.0223	7.2188*	0.0497	14.5170*	0.0840	12.3890*
FTSE100/Nikkei225	0.0086	0.2092	0.0257	0.8079	0.0532	0.4569
Nikkei225/S&P500	0.0154	1.7870	0.0326	2.5579	0.0429	2.8821
All 3 indices	0.0189	4.1156	0.0395	6.1832*	0.0686	3.8933

Note: Out of sample empirical coverage $\hat{q} = \sum \frac{I_{p,t+1}}{T}$ and likelihood ratio LR_{cc} for different coverages. Marked values with * are rejected at 5% confidence level.

close to theoretical coverage rates, q. Also, the forecasts based on the quantile regression model pass the conditional coverage test in almost all cases except the S&P500/FTSE100 pair.

Now, we turn into the conditional quantile forecast test to see whether one forecast encompasses the other forecasts. Given the forecasts, we estimate the optimal combination weights $(\theta_0^*, \theta_i^*, \theta_j^*)'$ in the forecast combination $\theta_0 + \theta_i VaR_{i,t} + \theta_j VaR_{j,t}$ using the GMM approach. For the purposes of this empirical application, we use $\mathbf{W}^* \equiv (1, r_{p,t}, VaR_{i,t}, VaR_{j,t})'$ as instrument variables. Table III.6 presents some results of the CQFE test.

As can be seen from Table III.6, test results are mixed. In some cases, for example, 2.5% and 5% of S&P500/FTSE100 and 5% of FTSE100/Nikkei225, GARCH-t forecasts encompass GARCH-Normal forecasts. These results suggest that the individual forecast by the GARCH-t model outperforms the forecast combination. However, in some cases, neither forecast encompasses its competitor for given levels of q in some cases: for example, GARCH-t and Quantile Regression forecasts of 1% of S&P500/FTSE100, 5% of FTSE100/Nikkei225, 2.5% of Nikkei225/S&P500, and 5% of all indices. This implies that the forecast combination outperforms the individual forecasts.

Table III.6: Conditional Quantile Forecast Encompassing Test for VaR Measures

q		$\hat{\theta}_{0n}$	$\hat{\theta}_{1n}$	$\hat{\theta}_{2n}$	$ENC1_n$	$ENC2_n$	J-stat.
S&P 500 / FTSE 100							
1%	G-N vs G-t	0.8802	0.1560	0.8488	5.0207	0.1792	10.0615
		(2.4928)	(0.4695)	(0.8420)			
	G-t vs QR	-1.8628	1.4884	0.1714	17.0430*	39.8610*	8.5779
		(1.0264)	(0.5899)	(0.8284)			
2.5%	G-N vs G-t	-0.1793	-0.0663	1.2803	50.5173*	0.0541	9.9943
		(4.9816)	(0.5588)	(1.8500)			
	G-t vs QR	-0.0816	0.3848	0.8100	1.8553	0.0886	6.3645
		(3.1652)	(2.0028)	(1.3762)			
5%	G-N vs G-t	0.0090	-0.0292	1.0813	57.5479*	0.0126	4.7939
		(1.4253)	(0.2863)	(0.7310)			
	G-t vs QR	-0.0996	1.1562	-0.0828	0.0476	24.4700*	4.7935
		(1.2852)	(0.7506)	(0.3797)			
FTSE 100 / Nikkei 225							
1%	G-N vs G-t	-1.3215	0.4575	0.7787	3.2458	1.6022	4.3218
		(0.9449)	(1.2296)	(1.1342)			
	G-t vs QR	-1.1221	-0.0474	1.1753	0.7448	0.4134	2.3827
		(1.3114)	(3.7012)	(3.5782)			
2.5%	G-N vs G-t	-0.8051	0.7499	0.6680	3.3232	3.0526	6.6212
		(1.0008)	(1.5809)	(1.5489)			
	G-t vs QR	-0.4630	0.6057	0.7479	3.2587	2.2284	5.3642
		(2.1274)	(4.4600)	(4.1680)			
5%	G-N vs G-t	-0.2017	0.5380	0.6449	7.9069*	3.0781	7.0080
		(0.2731)	(0.5259)	(0.4768)			
	G-t vs QR	-0.4340	0.3473	0.9086	6.0478*	8.8147*	7.5050
		(0.2947)	(0.7621)	(0.8127)			

Note: Out-of-sample CQFE test for VaR measure for equal weighted portfolios of a long position with an investment horizon of 1 week. The consistent standard errors of the GMM estimator $(\theta_{0n}, \theta_{1n}, \theta_{2n})'$ were computed with the smoothing parameter $\tau = .01$ and are reported in parentheses. J-stat. is the value of the J-test statistics. The values marked with * are significant at the 5% level.

Continued...

Table III.6: (Continued) CQFE Test for VaR Measures

q		$\hat{\theta}_{0n}$	$\hat{\theta}_{1n}$	$\hat{\theta}_{2n}$	$ENC1_n$	$ENC2_n$	J-stat.
Nikkei 225 / S&P 500							
1%	G-N vs G-t	-0.5130 (1.4948)	0.0983 (4.9560)	0.8318 (3.8468)	1.8836	0.8960	4.0388
	G-t vs QR	-1.5501 (16.154)	1.1812 (2.7315)	-0.01828 (6.6748)	3.3543	65.024*	3.2491
2.5%	G-N vs G-t	-0.0419 (1.5580)	0.7404 (6.7990)	0.4268 (5.7313)	1.5152	0.1017	7.7490
	G-t vs QR	-0.97643 (9.5137)	1.6029 (6.2149)	-0.46832 (9.9390)	6.1821*	30.165*	8.3602
5%	G-N vs G-t	-0.7873 (0.4795)	0.5710 (1.6717)	0.7659 (1.5057)	11.3189*	6.1537*	5.1227
	G-t vs QR	-0.6572 (1.9436)	1.1096 (1.7185)	0.1534 (2.4060)	4.7782	25.6080*	6.4837
All 3 indices							
1%	G-N vs G-t	-1.7299 (5.9222)	0.5816 (5.1065)	0.8440 (5.4395)	4.5888	6.2000*	4.5466
	G-t vs QR	-0.5629 (0.9365)	1.0342 (3.3354)	0.1129 (3.7906)	0.4159	1.5736	3.6859
2.5%	G-N vs G-t	-0.6795 (0.6716)	0.6548 (0.9324)	0.6510 (0.9600)	6.6460*	7.9908*	3.3987
	G-t vs QR	-0.3841 (0.6047)	1.5219 (0.7609)	-0.4000 (0.7367)	0.9615	4.1193	1.2461
5%	G-N vs G-t	0.5810 (0.4334)	0.5257 (0.9512)	0.3583 (1.0070)	5.3740	2.1720	2.9759
	G-t vs QR	-1.3148 (0.5002)	1.3417 (0.8988)	0.2724 (0.9394)	12.2800*	15.4030*	9.7833

Note: Out-of-sample CQFE test for VaR measure for equal weighted portfolios of a long position with an investment horizon of 1 week. The consistent standard errors of the GMM estimator $(\theta_{0n}, \theta_{1n}, \theta_{2n})'$ were computed with the smoothing parameter $\tau = .01$ and are reported in parentheses. J-stat. is the value of the J-test statistics. The values marked with * are significant at the 5% level.

In conclusion, if we are concerned about events, such as markets crashing together, the dependence structures should be modelled with caution. However, we observe that the misspecification problems in joint interval forecasting of low-left tail might be reduced by portfolio construction, so that we do not need to pay too much attention to the dependence structure modelling if we construct portfolios assuming portfolio weights are given. Also, any model specification used in the paper seems to be fit to forecast intervals well if interesting intervals are in the moderate ranges. However, if concerns are focused on the extreme left tails, such as q is 1%, 2.5%, and 5%, the fat-tailed model specification of the GARCH-t(1,1) passes the conditional coverage tests, while GARCH-Normal(1,1) does not pass the test in almost all cases. Therefore, GARCH-t can be said to be the 'acceptably accurate' VaR forecasts model. This result indicates that we need to assume fat-tailed models for extremely low coverage rates, generally used in the VaR applications. Therefore, it is important to set up the interval forecasting models carefully depending on the interesting intervals to forecast. However, when we include one other forecast from the quantile regression, the results of conditional quantile forecasts test suggest that combination of individual forecasts may improve the performance of the interval forecasts of portfolios in many cases, although sometimes the individual forecasts outperform the forecast combinations.

III.6 Conclusion

This paper has extended the conditional coverage test of Christoffersen (1998) to the multivariate cases to answer whether the dependence structure is important for evaluating interval (area) forecasts. We provide two different interval forecasts: body interval and low-left tail interval forecasts. The complete conditional coverage test is a likelihood ratio test and decomposes easily into subtests of independence and unconditional coverage tests. The extension to multivariate cases is done by using the copulas, which are simply the flexible cumulative density functions (CDFs) with closed forms in some cases.

In Monte Carlo simulations, it is shown that the power of the test to reject the different dependence interval forecasts is very low. This implies that the dependence structure is not very important, which makes it difficult to discriminate among the interval forecasts from different dependence structures, if marginal distributions are correctly specified.

An application to weekly returns of three international stock markets confirms this conclusion. The body-interval forecast (Case 1) from the GARCH-Normal(1,1) margins passes the conditional coverage test for certain coverage rates of Nikkei225/S&P500 pair, but fails the test for most others. However, no dependence structure seems to be dominant. The GARCH-t(1,1) margins pass the tests in most cases across the various dependence structures, passing both the independent dynamics test and the unconditional nominal coverage test. Therefore, we can conclude that it is more important to specify margins correctly rather than correctly specifying dependence structures even for modelling the multivariate joint interval (area) forecasts.

Financial market participants have recently shown increasing interest in one-sided interval forecasting, known as Value-at-Risk (VaR) measure. Therefore it is natural to pay attention to low-left tail interval forecast in multivariate cases. However, it is interesting that all dynamic and parametric forecast models used in this paper are rejected when the interval is low-left tail (Case 2) with low coverage rates. These failures are mainly due to the failures of the unconditional nominal coverage tests. Considering a nonparametric multivariate specification or dynamics of dependence structure may likely present favorable alternatives.

Finally, we perform the portfolio interval evaluation using the LR test as well as the CQFE test by Giacomini and Komunjer (2005). Portfolio construction reduces the multivariate interval evaluation problem to the univariate problem. Selected forecasting models are the GARCH-Normal(1,1), the GARCH-t(1,1), and the forecasts from the quantile regression. The test results indicate that the GARCH-t(1,1) model passes the tests across the samples under the various moderate coverage rates, so it can be said to be 'acceptably accurate', while the GARCH-Normal(1,1) fails the tests under extremely low coverage rates. This result indicates that we need to assume fat-tailed models for extremely low coverage rates, generally used in the VaR applications. The test results of the CQFE test shows that a combination of individual forecasts may improve the performance of the interval forecasts of portfolios in many cases, although sometimes the individual forecasts from the GARCH-t(1,1) outperform the forecast combinations.

Recently, a number of academic studies and applications have paid attention to modelling the multivariate cases beyond the univariate cases. The results of this paper suggest a couple of new directions for future research. Although this paper does not provide a strong support for importance of modelling dependence structure, especially after constructing portfolios, it might be important depending on the magnitude of dependency. Or it might be because of the trading (long) horizon effect as pointed out in Christoffersen and Diebold (2000) for univariate volatility forecasts.

References

Ang, A., and Chen, J., 2002: Asymmetric correlation of equity portfolios. *Journal of Financial Economics*, **63**(2), 442–294.

Ang, A., Chen, J., and Xing, Y., 2006: Downside risk. *Review of Financial Studies*, **19**(1), 1–40.

Berkowitz, J., 2001: Testing density forecasts, with applications to risk management. *Journal of Business and Economic Statistics*, **19**(4), 465–474.

Chatfield, C., 1993: Calculating interval forecasts. *Journal of Business and Economic Statistics*, **11**, 121–135.

Christiansen, C., 1999: Valu-at-risk using the factor-arch model. *Journal of Risk*, **1**(2), 65–86.

Christoffersen, P., 1998: Evaluating interval forecasts. *International Economic Review*, **39**(4), 841–862.

Christoffersen, P., and Diebold, F. X., 2000: How relevant is volatility forecasting for financial risk management? *Review of Economics and Statistics*, **82**, 12–23.

Clemen, R., Murphy, A., and Winkler, R., 1995: Screening probability forecasts: Contrasts between choosing and combining. *International Journal of Forecasting*, **11**, 133–146.

Clements, M. P., and Smith, J., 2002: Evaluating multivariate forecast densities: a comparison of two approaches. *International Journal of Forecasting*, **18**, 397–407.

Crnkovic, C., and Drachman, J., 1996: Quality control. *Risk*, **9**, 138–143.

Diebold, F. X., Gunther, T. A., and Tay, A. S., 1998: Evaluating density forecasts with applications to financial risk management. *International Economic Review*, **39**(4), 863–883.

Diebold, F. X., Hahn, J., and Tay, A. S., 1999: Multivariate density forecast evaluation and calibration in financial risk management: Hig-frequency returns on foreign exchange. *Review of Economics and Statistics*, **81**(4), 661–673.

Drezner, Z., and Wesolowsky, G. O., 1990: On the computation of the bivariate normal integral. *Journal of Statistical Computation and Simulation*, **35**, 101–107.

Embrechts, P., Lindskog, F., and McNeil, A., 2003: Modelling dependence with copulas and applications to risk management. In *Handbook of Heavy Tailed Distributions in Finance*, editor S. Rachev, 329–384. Elsevier.

Engle, R., 2002: Dynamic conditional correlation-a simple class of multivariate garch models. *Journal of Business and Economic Statistics*, **20**(3), 339–350.

Engle, R., and Gonzalez-Rivera, G., 1991: Semiparametric arch models. *Journal of Business and Economic Statistics*, **9**, 345–359.

Engle, R., and Sheppard, K., 2001: Theoretical and empirical properties of dynamic conditional correlation multivariate garch. *UCSD Discussion paper 2001-15, UCSD.*

Fermanian, J.-D., and Scaillet, O., 2003: Nonparmetric estimation of copulas for time series. *Journal of Risk*, **5**(4), 25–54.

Giacomini, R., and Komunjer, I., 2005: Evaluation and combination of conditional quantile forecasts. *Journal of Business and Economic Statistics*, **23**(4), 416–431.

Hoel, P., 1954: A test for markov chains. *Biometrika*, **41**, 430–433.

Joe, H., 1997: *Multivariate models and dependence concepts.* Chapman & Hall, London.

Komunjer, I., 2005: Quasi-maximum likelihood estiamtion for conditional quantiles. *Journal of Econometrics*, **128**(1), 137–164.

Longin, F., and Solnik, B., 2001: Extreme correlation of international equity markets. *Journal of Finance*, **56**(2), 649–676.

Lopez, J., 1999: Regulatory evaluation of value-at-risk models. *Journal of risk*, **1**, 37–64.

Lopez, J., and Walter, C., 2001: Evaluating covariance matrix forecasts in a value-at-risk framework. *Journal of risk*, **3**(3), 69–98.

Nelsen, R., 1999: *An Introduction to Copulas.* Springer-Verlag, New York.

Patton, A., 2006: Modelling asymmetric exchange rate dependence. *International Economic Review*, **47**(2), 527–556.

Pesaran, M. H., and Timmermann, A., 2006: Testing dependence among serially correlated multi-category variables. *Working paper.*