UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE CIÊNCIAS ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

MARIANA BARTELS

ESTIMAÇÃO DE MEDIDAS DE RISCO PARA PORTFÓLIOS DE ALTAS DIMENSÕES ATRAVÉS DE CÓPULAS FATORIAIS

Porto Alegre
2015

MARIANA BARTELS

ESTIMAÇÃO DE MEDIDAS DE RISCO PARA PORTFÓLIOS DE ALTAS DIMENSÕES ATRAVÉS DE CÓPULAS FATORIAIS

Dissertação submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como quesito parcial para obtenção do título de Mestre em Economia, com ênfase em Economia Aplicada.

Orientador: Prof. Dr. Flávio Augusto Ziegelmann

Porto Alegre
2015

CIP - Catalogação na Publicação

Bartels, Mariana
Estimação de medidas de risco para portfólios de altas dimensões através de cópulas fatoriais / Mariana Bartels. -- 2015.
50 f.

Orientador: Flávio Augusto Ziegelmann.

Dissertação (Mestrado) -- Universidade Federal do Rio Grande do Sul, Faculdade de Ciências Econômicas, Programa de Pós-Graduação em Economia, Porto Alegre, BR-RS, 2015.

1. Econometria. I. Ziegelmann, Flávio Augusto, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

MARIANA BARTELS

ESTIMAÇÃO DE MEDIDAS DE RISCO PARA PORTFÓLIOS DE ALTAS DIMENSÕES ATRAVÉS DE CÓPULAS FATORIAIS

Dissertação submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como quesito parcial para obtenção do título de Mestre em Economia, com ênfase em Economia Aplicada.

Aprovada em Porto Alegre, 11 de maio de 2015.

Banca Examinadora

Prof. Dr. Cristiano Augusto Coelho Fernandes
PUC-Rio

Prof. Dr. Hudson da Silva Torrent
PPGE/UFRGS

Prof. Dr. Osvaldo Candido da Silva Filho
PPGE/UCB

Prof. Dr. Flávio Augusto Ziegelmann - Orientador
PPGE/UFRGS

Agradecimentos

Agradeço aos meus pais e irmãos, pelo apoio que me deram durante todo o curso de mestrado.

Ao meu namorado, Mauro, pela paciência e incentivo.

Aos meus colegas da FEE e do PPGE, pela ajuda nos momentos difíceis.

Aos professores que me deram aula durante o curso, pelo conhecimento que me passaram, indispensável para a minha formação.

À FEE, pela motivação e disponibilidade para realizar o curso de mestrado.

À Secretaria do PPGE pela disponibilidade e prontidão.

Aos professores Oh e Patton, pelo auxílio no desenvolvimento no trabalho.

À banca examinadora, pelas sugestões e contribuições.

E principalmente ao professor Flávio Ziegelmann, pela excelente orientação e conselhos.

Resumo

No mercado financeiro, estimar o risco de portfólios de ações é decisivo, pois indica qual o potencial de perda de um investimento em determinado momento. Quando a quantidade de ações que compõem o portfólio é muito grande, entretanto, a quantidade de modelos disponíveis para esta estimação fica muito reduzida. O principal objetivo deste trabalho é estimar, via cópulas, medidas de risco, como o Valor em Risco (VaR) e o Expected Shortfall (ES), para uma carteira construída a partir de 44 ativos da Bolsa de Valores de São Paulo - BOVESPA, comparando o desempenho de modelos utilizados para tal fim e viáveis para uma dimensão tão alta. Para tanto, no artigo que compõe esta dissertação, utilizamos uma abordagem baseada na estimação de modelos cópula-GARCH para os retornos destas ações. Utilizamos modelos tradicionais de cópulas multidimensionais, bem como modelos mais rebuscados de cópulas fatoriais, explorando três níveis de dependência e a possibilidade de que os parâmetros do modelo variem no tempo. Os resultados indicam que o modelos mais adequados para a estimação do VaR do portfólio, levando em conta a precisão das estimativas e a quantidade de parâmetros a serem estimados, são os baseados em cópulas fatoriais. Em relação ao ES, não encontramos diferenças significativas.

Palavras-chave: Cópulas fatoriais. Cópulas multidimensionais. GARCH. Medidas de risco. Valor em Risco. Expected Shortfall. BOVESPA.

Abstract

In the financial market, estimating the risk of stock portfolios is crucial since it indicates the potential loss of an investment at any given time. When the number of assets that compose the portfolio is very large, however, the number of models available for this estimation is very low. The main objective of this study is to estimate, via copulas, risk measures, such as Value at Risk (VaR) and Expected Shortfall (ES), for portfolios constructed from 44 Brazilian BOVESPA assets, comparing the performance of some available models for such a high dimension. Therefore, in the paper that makes up this dissertation, we use an approach based on the estimation of copula-GARCH models for the returns of these assets. We use traditional multidimensional copula models and more elaborated models of factor copulas, and explore three levels of dependency and the possibility that the model parameters vary in time. The results indicate that the most appropriate models for estimating VaR of this portfolio, taking into account the accuracy of the estimates and the number of parameters to be estimated, are factor copula models. However, regarding ES, we do not find statistically significant superior models.

Keywords: Factor copula. Multidimensional copulas. GARCH. Risk measures. Value at Risk. Expected Shortfall. BOVESPA.

Sumário

1 Introdução

No mercado financeiro, estimar o risco de portfólios de ações é decisivo, pois indica qual o potencial de perda de um investimento em determinado momento. O principal objetivo deste estudo é estimar, via cópulas, medidas de risco, como o Valor em Risco (VaR) e Expected Shortfall (ES), para uma carteira construída a partir de 44 ativos da Bolsa de Valores de São Paulo - BOVESPA, comparando o desempenho de modelos utilizados para tal fim.

Para calcular estas estimativas, é necessário conhecer a distribuição conjunta das séries dos valores das ações ou de seus retornos, o que nem sempre é uma tarefa fácil. Esta tarefa se torna ainda mais árdua à medida que o tamanho do portfólio aumenta.

Muitas variáveis financeiras não seguem distribuições de probabilidade normais univariadas e, portanto, também não seguem conjuntamente distribuições normais multivariadas. Porém, em geral, por simplicidade, estas são as distribuições utilizadas para modelar a distribuição conjunta de séries temporais multivariadas, em especial séries de retornos de ações (BRADLEY; TAQQU, 2003). Este método se torna ainda mais comum quando se trata de uma grande quantidade de séries, pois outros métodos se tornam impraticáveis, devido à alta demanda computacional e à grande quantidade de parâmetros a serem estimados.

A distribuição normal, entretanto, é muito limitada no que diz respeito à representação das características típicas das variáveis financeiras. Elas podem ser assimétricas, a-presentar caudas pesadas e dependência caudal, fatos não contemplados pelas propriedades da distribuição normal, porém de suma importância no mercado financeiro. Para contemplar essas características, recentemente têm-se utilizado o método de cópulas para fazer tal modelagem mais adequadamente. Cópulas são funções que ligam uma função de distribuição multivariada a suas distribuições marginais de qualquer dimensão, possuindo todas as informações relevantes a respeito da estrutura de dependência entre variáveis aleatórias.

A principal vantagem desse tipo de modelagem é permitir separar a estimação das distribuições marginais das séries da estimação de sua estrutura de dependência. Assim, toda a complexidade possivelmente presente na distribuição marginal de cada variável é captada no primeiro passo da estimação, permitindo que a modelagem e estimação da cópula se concentrem apenas na relação entre as variáveis. Isso dá origem aos modelos cópula-GARCH, que se tornaram referência no tratamento deste tipo de problema. A teoria inicial e tradicional de cópulas pode ser encontrada nos livros de Joe (1997) e Nelsen (2006). Para uma revisão da teoria de cópulas mais aplicada a séries financeiras, ver Cherubini, Luciano e Vecchiato (2004).

Para o caso bivariado, ou seja, quando se espera modelar apenas um par de variáveis, há uma grande variedade de famílias de cópulas cujas propriedades já foram bastante estudadas, dentre as quais é possível escolher aquela que se ajusta melhor aos dados. Entretanto, para dimensões maiores, a quantidade de modelos fechados disponíveis na literatura tradicional se reduz. Assim, apesar da recente popularidade dos modelos de cópulas, a maioria dos estudos ainda se limita a estimar a distribuição conjunta de cada par de séries de interesse separadamente, ao invés de procurar um modelo que una todas as séries. Ver, por exemplo, Silva Filho, Ziegelmann e Dueker (2012, 2014).

Tendo em vista superar esta limitação, Joe (1996) propôs a modelagem através de cópulas Vine, a qual foi desenvolvida em maiores detalhes por Bedford e Cooke (2001, 2002) e Kurowicka e Cooke (2006). Essa metodolgia propõe a construção gráfica de uma cascata de cópulas bivariadas, as quais, juntas, formam a distribuição multivariada que se busca modelar. Isto permite uma grande flexibilidade no que tange à estrutura de dependência das séries em casos de dimensão maior que dois. Duas classes de cópulas Vine foram posteriormente desenvolvidas por Aas et al. (2009): as cópulas C-Vine e D-Vine, que nada mais são que diferentes formas de construir a cascata (ou árvore) de cópulas bivariadas, dado que esta pode ser definida de diversas formas. Recentemente, Dissmann et al. (2013) propuseram as cópulas R-Vine (regular Vine), cujos casos especiais são as cópulas C-Vine e D-Vine. As cópulas Vine apresentam bom desempenho quando comparadas a métodos alternativos de modelagem multivariada tradicional, como mostraram Berg e Aas (2009) e Fischer et al. (2009). Czado (2010) e Kurowicka e Joe (2011) apresentam boas revisões sobre cópulas Vine, inclusive com aplicações desta metodologia. As cópulas Vine, portanto, permitem que sejam modelados de forma mais flexível conjuntos de variáveis de tamanhos moderados, como poucas dezenas de séries. Tófoli, Ziegelmann e Silva Filho (2013) introduziram recentemente uma abordagem de cópulas Vine dinâmicas.

Entretanto, para uma quantidade maior de séries a serem modeladas, essa metodologia se torna infactível computacionalmente, pois há muitos parâmetros a serem estimados. Neste sentido, algumas alternativas foram propostas para aumentar o número de dimensões que podem ser incluídas na modelagem. Uma delas é assumir independência condicional nos maiores níveis da cascata da cópula Vine, para reduzir o número de parâmetros a serem estimados, como em Brechmann, Czado e Aas (2012). Outra é utilizar Vines gaussianas em dimensões arbitrárias, como em Kurowicka e Cooke (2006), ou ainda C-vines autorregressivas, como em Heinen e Valdesogo (2009).

Porém, esses métodos ainda se limitam a modelar poucas dezenas de séries. Quando se necessita uma modelagem para um número maior de séries, outra abordagem é necessária. Tendo em vista modelar dados de alta dimensão com esforço computacional plausível, foram recentemente propostos os modelos de cópulas fatoriais. Neste campo, surgiram inicialmente os trabalhos de Hull e White (2004), Andersen e Sidenius (2004), Van der Voort (2005) e McNeil et al. (2005). Mais recentemente, Krupskii e Joe (2013) e Oh e Patton (2012) desenvolveram independentemente dois métodos distintos de cópulas fatoriais, porém ambos com o mesmo propósito: reduzir a dimensão dos dados partindo do pressuposto de que fatores comuns podem representar bem o conjunto original de relações de dependência. Embora desenvolvido recentemente, o modelo proposto por Oh e Patton (2012) já foi utilizado por Creal e Tsay (2014) e Chen, MacMinn e Sun (2013), além de Oh e Patton (2013), e será também explorado neste trabalho.

Quando se trata de cópulas fatoriais, ainda há dois aspectos a se considerar: como se organiza a estrutura de dependência das séries e como esta se comporta através do tempo. A estrutura de dependência entre as séries em um determinado momento no tempo pode ser modelada de diferentes formas, dependendo das suposições que podem ser feitas a respeito das séries. Como utilizado por Oh e Patton (2012, 2013), esta estrutura pode assumir, basicamente, três formas: equidependência, onde todas as séries dependem umas das outras de forma igual; dependência heterogênea, que não supõe nenhuma relação entre as dependências entre as séries; e dependência por blocos, que divide a priori o conjunto de séries temporais em grupos homogêneos em termos de dependência. Estas três diferentes abordagens serão avaliadas neste estudo.

Além disso, esta estrutura de dependência entre as séries pode ser dinâmica. Na literatura que trata de cópulas há diversos estudos sobre a estrutura de dependência das variáveis, porém somente recentemente têm sido desenvolvidos modelos que permitem que esta estrutura mude ao longo do tempo, como é o caso de Mendes (2005), Patton (2006), Almeida e Czado (2012), Silva Filho, Ziegelmann e Dueker (2012) e Oh e Patton (2013). Estes modelos foram utilizados em diversos estudos, como Silva Filho, Ziegelmann (2014), Silva Filho, Ziegelmann e Dueker (2014) e Tófoli, Ziegelmann e Silva Filho (2013). Neste trabalho, utilizaremos em especial a modelagem GAS (Generalized Autoregressive Score) proposta por Creal, Koopman e Lucas (2013) e utilizada em cópulas por Oh e Patton (2013), a qual permite incluir no modelo uma dinâmica temporal nas cargas fatoriais da cópula fatorial.

A estimação dos parâmetros destas cópulas fatoriais, entretanto, não é trivial: pelo fato de sua função de verossimilhança não apresentar forma fechada, seu cálculo é feito numericamente e envolve integrações e inversão de funções com mais de um argumento. Assim, o processo de otimização para obter as estimativas dos parâmetros também é dificultado, pois, neste caso, não é possível utilizar métodos baseados em gradientes, sendo assim mais indicados métodos como o Simplex de Nelder e Mead (1965). Este foi o método de otimização utilizado para a estimação dos parâmetros de todas as cópulas utilizadas neste trabalho.

Dessa forma, os modelos utilizados para estimar as medidas de risco VaR e ES para as carteiras construídas a partir de ativos da BOVESPA serão de cópulas multidimensionais tradicionais e de cópulas fatoriais. Nós especificamente comparamos o desempenho do modelo de cópulas fatoriais, sendo estas dinâmicas ou estáticas e com as três estruturas de dependência expostas acima, com outros modelos de cópulas viáveis para elevadas dimensões, tais como Normal, t-student e Arquimedianas (Gumbel, Clayton e Frank).

Assim, a Seção 2 deste trabalho consiste em um artigo que descreve a metodologia utilizada no trabalho, e traz uma aplicação desta a dados diários das ações da BOVESPA. A análise realizada indica que as cópulas fatoriais apresentam desempenho superior às demais no que diz respeito à estimação do Valor em Risco para o portfólio de ações da BOVESPA utilizado. Porém não foi possível observar alguma cópula que tenha se destacado positivamente no que se refere à estimação de seu Expected Shortfall. Na Seção 3 fazemos as considerações finais do trabalho, sintetizando os resultados obtidos e explorando possíveis extensões do presente trabalho.

2 Value at Risk and Expected Shortfall Estimation for High Dimensional Portfolios via Factor Copulas: an Application to Brazilian BOVESPA Stock Market

Abstract: In this paper we propose the estimation of risk measures, such as Value at Risk (VaR) and Expected Shortfall (ES) for large dimensional portfolios via copula modeling. For that we compare several high dimensional copula models, from naive ones to complex factor copulas, which are able to tackle the curse of dimensionality whereas simultaneously introducing a high level of complexity into the model. We explore static and dynamic copula fitting and also allow for different levels of flexibility for our model dependence structure, following Oh and Patton (2012). Our results, for assets negotiated at Brazilian BOVESPA stock market, suggest that the factor copula models have superior performance compared to the others with respect to the estimation of Value at Risk. Nevertheless the results are very similar when it comes to estimating the Expected Shortfall.

Keywords: Factor copula. Multidimensional copulas. GARCH. Risk measures. Value at Risk. Expected Shortfall. BOVESPA.

JEL codes: C22, C52, C53, C58.

2.1 Introduction

In the financial market, estimating the risk of stock portfolios is crucial since it indicates the potential loss of an investment at any given time. The main objective of this study is to estimate - via copula modeling - risk measures, such as Value at Risk (VaR) and Expected Shortfall (ES), for portfolios constructed from 44 Brazilian BOVESPA assets, comparing the performance of some available models for such a high dimension.

To calculate these estimates, it is necessary, however, to know the joint distribution of the stock returns, which is not always easy. This task becomes even more difficult as the portfolio size increases. Most financial variables do not have univariate normal probability distributions and therefore do not follow jointly multivariate normal distributions. However, in general, for simplicity, these are the distributions used to model the joint distribution of multivariate time series, in particular those of asset returns (BRADLEY; TAQQU, 2003). This becomes even more popular when it comes to studying a large collection of assets, since other methods can be impractical due to the high computational demand from the large amount of parameters to be estimated.

The normal distribution, nevertheless, is very limited with respect to the representation of the stylized facts about financial variables. These variables can be asymmetric, present heavy tails as well as tail dependence, among other characteristics, facts not reflected by the normal distribution, but of paramount importance in the financial market. To address these features, copulas have recently become popular, modeling these problems more adequately. Copulas are functions that link marginals to their multivariate distribution, thus having all relevant information regarding the dependence structure among random variables.

The main advantage of this type of modeling is to split the estimation of the marginal distributions from the estimation of their dependence structure. Thus, all the complexities possibly present in the marginal distributions of each variable are captured at the first stage of the estimation process, allowing the copula modeling to focus only on the relationship between the variables. This originates the Copula-GARCH modeling approach, which has become a benchmark for dealing with such problems.

For the bivariate case there is a large variety of families of copulas whose properties have been widely studied. In empirical exercises, one can therefore choose that bivariate copula which best fits the data. However, for higher dimensions, the quantity of available copula models in the literature is still small. Thus, despite the recent popularity of copula models, most studies are still limited to estimate the joint distribution of each pair of series of interest separately, as in Patton (2006), Silva Filho and Ziegelmann (2014), Silva Filho, Ziegelmann and Dueker (2012, 2014), among others, rather than searching for a model that links all series.

In order to tackle this drawback, Joe (1996) proposes the vine copulas approach, which is further developed by Bedford and Cooke (2001, 2002) and Kurowicka and Cooke (2006). This method proposes the graphical construction of a cascade of bivariate copulas, which together build on to a multivariate distribution that one seeks for. This allows greater flexibility regarding the variables dependence structure in cases of dimension greater than 2. Two classes of vine copulas are developed in detail by Aas et al. (2009): the C- and D-vine copulas, which are nothing more than different ways to construct the cascade (or tree) of bivariate copulas, as this can be defined in several ways. Recently, Dissmann et al. (2013) propose the R-vine copula, of which the C- and D-vines are special cases. The vine copulas perform well when compared to alternative traditional multivariate modeling methods, as shown by Berg and Aas (2009) and Fischer et al. (2009). Czado (2010) and Kurowicka and Joe (2011) provide good reviews on vine copulas, including applications of this methodology. The vine copulas therefore allow copula modeling in a more flexible way for moderate number of variables, such as a few tens of series, as in Tófoli, Ziegelmann and Silva Filho (2013), who introduce a novel dynamic vine copula approach.

However, for a larger number of variables, this methodology becomes computationally infeasible as there are too many parameters to be estimated. Therefore, some alternatives have been proposed in the literature. One is to assume conditional independence in the highest levels of the cascade, hence reducing the number of parameters to be estimated, as in Brechmann, Czado and Aas (2012). Another is to use Gaussian vines in arbitrary dimensions as Kurowicka and Cooke (2006), or C- autoregressive vines, as in Heinen and Valdesogo (2009).

Nevertheless, these methods are still limited to a few tens of variables. When higher dimensional models are needed, another approach must come into play. Having as a target facing the computational burden imposed by this "curse of dimensionality", factor copula models have been recently proposed. In this field, some papers initially appeared in the literature, such as Hull and White (2004), Andersen and Sidenius (2004), Van der Voort

(2005) and McNeil et al. (2005). Recently, Krupskii and Joe (2013) and Oh and Patton (2012) independently developed two different methods of factor copulas, but both with the same purpose: reducing the dimension of the data assuming that common factors can represent well the original set of dependence relationships. The approach by Oh and Patton (2012) has been used by Creal and Tsay (2014), Chen, MacMinn and Sun (2013), and Oh and Patton (2013), and will be also used in this work.

When it comes to factor copulas, there are two more aspects to consider: how the series dependence structure is organized and how it behaves over time. The dependence structure among variables at a given moment in time can be modeled in different ways, depending on the assumptions that can be made about these variables. In Oh and Patton (2012, 2013), this structure is allowed to assume three different forms: equidependence, where all time series are dependent on each other equally; heterogeneous dependence that assumes no relationship between the dependences among the variables; and block dependence, which divides the initial set of variables in homogenous groups in terms of dependence. These three approaches will be explored in this paper.

Besides that, the dependence structure among the series can be dynamic. In the literature of copulas there are several studies on the structure dependence of the variables, but only recently models which enable this structure to change over time have been developed, such as in Mendes (2005), Patton (2006), Almeida and Czado (2012), Silva Filho, Ziegelmann and Dueker (2012), Tófoli, Ziegelmann and Silva Filho (2013) and Oh and Patton (2013). These models were used in several studies, as Silva Filho, Ziegelmann (2014) and Silva Filho, Ziegelmann and Dueker (2014). In this work, we use particularly GAS (Generalized Autoregressive Score) modeling proposed by Creal, Koopman and Lucas (2013) and applied to copulas by Oh and by Patton (2013). This modeling allows the inclusion of time dynamics in the factor loadings of the factor copula in the model.

Thus, the models used to estimate VaR and ES risk measures for portfolios constructed from assets of BOVESPA will be traditional multidimensional copulas and factor copulas. We specifically compare the performance of the factor copula model, with dynamic or static dependence and using the three dependence structure exposed above, with other feasible high dimensional copula models, such as Normal, Student's t and Archimedean (Gumbel, Clayton and Frank).

The rest of this paper is structured as follows. Section 2.2 presents the basics about copulas, the factor copula methodology itself and risk measures. In Section 2.3 the empirical results are presented; and finally, in Section 2.4, we conclude with our remarks.

2.2 Methodology

Consider the problem of modeling the joint distribution of multiple financial time series using copulas. Here we follow the approach by Joe and Xu (1996), inference from margins, who show that a two-step estimation procedure is consistent. Firstly, one models the conditional marginal distribution of each time series and then, in a second step, one estimates the dependence parameters associated to the copula function. The conditional marginal distributions of each time series are modeled via AR-GARCH processes.

2.2.1 Copula Basics

Copulas can be defined as functions that combine the marginal distributions of a set of variables so as to obtain their multivariate cumulative distribution function (cdf), or otherwise functions that are multivariate cdf's with Uniform [0,1] marginal distributions (Schweizer and Sklar, 2005). Thus, an N-dimensional copula C can be defined by

$$C(u_1, \cdots, u_N) = \mathbb{P}(U_1 \leq u_1, \cdots, U_N \leq u_N), \tag{2.1}$$

where $U_1, \cdots, U_N$ are $U[0,1]$ random variables (r.v.).

Hence, we can jointly model any set of N random variables via N-dimensional copulas, since any random variable can be transformed into a $U[0,1]$ r.v. through its own cdf, generating its probability integral transform.

More formally, we can define a copula as follows.

Definition 1. An N-dimensional copula is a function C with domain $[0,1]^N$, such that:

a) C is grounded and N-increasing.

b) C has margins C_k, $k = 1, 2, \cdots, N$, where $C_k(u) = u$ for every u in $[0,1]$.

Equivalently, an N-copula (or N-dimensional copula) is a function C with the following properties:

c) For all $\mathbf{u}$ in $[0,1]^N$, $C(\mathbf{u}) = 0$ if at least one coordinate of $\mathbf{u}$ is 0, $C(\mathbf{u}) = u_i$ if all the coordinates of $\mathbf{u}$ are 1 except u_i.

d) For all $\mathbf{a}$ and $\mathbf{b}$ in $[0,1]^N$ such that $a_i \leq b_i$ for every i, $V_C([\mathbf{a},\mathbf{b}]) \geq 0$, where V_C is called C-volume.

One of the main results of the theory of copulas is Sklar's Theorem (1956) which we state below.

Theorem 1. *(Sklar's Theorem). Let $Y_1, \cdots, Y_N$ be random variables with distribution functions $F_1, \cdots, F_N$, respectively, and joint distribution function H. Then, there is a C such that*

$$H(y_1, \cdots, y_N) = C\Big(F_1(y_1), \cdots, F_N(y_N)\Big) \tag{2.2}$$

for every $y_1, \cdots, y_N \in \bar{\boldsymbol{R}} = [-\infty, +\infty]$. If $F_1, \cdots, F_N$ are all continuous, then C is unique; otherwise C is determined only on $ImF_1 \times \cdots \times ImF_N$. Reciprocally, if C is an N-copula and $F_1, \cdots, F_N$ are distribution functions, then the function H defined above is an N-dimensional distribution function with margins $F_1, \cdots, F_N$.

Corollary 1. *Let H be an N-dimensional distribution function with continuous margins $F_1, \cdots, F_N$ and copula C (where C is defined as in Theorem 1). Thus, for any $\boldsymbol{u} = (u_1, u_2, \cdots, u_N)$ in $[0,1]^N$,*

$$C(u_1, \cdots, u_N) = H\Big(F_1^{-1}(u_1), \cdots, F_N^{-1}(u_N)\Big), \tag{2.3}$$

where F^{-1} is the generalized inverse.

2.2.1.1 Conditional Copulas

According to Silva Filho, Ziegelmann and Dueker (2014), in many cases, the dependence pattern and the general dynamics of a time series can be captured through conditional distributions on past observations. This has been studied by Mendes (2005) and Patton (2006). Thus, the extension of Sklar's theorem for the conditional case can prove to be very useful. According to Patton (2006), the Sklar's theorem for the conditional case is defined as follows:

Theorem 2. *(Sklar's Theorem - Conditional case). Let $Y_i|\,\boldsymbol{W}$, for $i = 1, \cdots, N$, be random variables with conditional distribution functions F_i, respectively, and conditional* *joint distribution function H of* $\mathbf{Y}|\boldsymbol{W}$*, where* $\boldsymbol{Y} = (Y_1, \cdots, Y_N)$ *and* $\boldsymbol{W}$ *has support* Ω*. Then there is a copula C such that, for any* $\boldsymbol{y} \in \bar{\boldsymbol{R}}^N$ *and* $\boldsymbol{w} \in \Omega$ *,*

$$H(y_1, \cdots, y_N|\boldsymbol{w}) = C\left(F_1(y_1|\boldsymbol{w}), \cdots, F_N(y_N|\boldsymbol{w})\middle|\boldsymbol{w}\right). \tag{2.4}$$

If $F_1, \cdots, F_N$ *are all continuous then C is unique, otherwise it is determined only on* $ImF_1 \times \cdots \times ImF_N$*. Reciprocally, if C is an N-copula and* $F_1, \cdots, F_N$ *are distribution functions, then H as defined above is a conditional N-dimensional distribution function of with margins* $F_1, \cdots, F_N$*.*

Corollary 2. *Let H be an N-dimensional conditional distribution function with continuous conditional margins* $F_1, \cdots, F_N$ *and copula C (where C is defined as in Theorem 2). Thus, for any* $\boldsymbol{u}$ *in* $[0,1]^N$ *and* $\boldsymbol{w} \in \Omega$*,*

$$C(u_1, \cdots, u_N|\boldsymbol{w}) = H\left(F_1^{-1}(u_1|\boldsymbol{w}), \cdots, F_N^{-1}(u_N|\boldsymbol{w})\middle|\boldsymbol{w}\right). \tag{2.5}$$

In the case of time series, in general, $\mathbf{W}$ represents the r.v.´s past, i.e., $\mathbf{W} = Y_{1_{t-1}}, \cdots, Y_{N_{t-1}}, Y_{1_{t-2}}, \cdots, Y_{N_{t-2}}, \cdots$, considering that $Y_1 = Y_{1_t}, \cdots, Y_N = Y_{N_t}$.

2.2.1.2 Estimation

Parameter estimation is performed via maximum likelihood considering the multivariate joint distribution in question, which can be written as

$$H(y_1, \cdots, y_N|\mathbf{w}) = C\left(F_1(y_1|\mathbf{w}), \cdots, F_N(y_N|\mathbf{w})\middle|\mathbf{w}\right). \tag{2.6}$$

As a result, we can show that

$$h(y_1, \cdots, y_N|\mathbf{w}) = \frac{\partial^N H(y_1, \cdots, y_N|\mathbf{w})}{\partial y_1 \cdots \partial y_N} = f_1(y_1|\mathbf{w}) \cdot \cdots \cdot f_N(y_N|\mathbf{w}) \cdot c(u_1, \cdots, u_N|\mathbf{w}), \tag{2.7}$$

where $u_1 = F_1(y_1|\mathbf{w}), \cdots, u_N = F_N(y_N|\mathbf{w})$.

From the expression above, we can obtain the likelihood function, given by

$$L(\theta_1, \cdots, \theta_N, \theta_C|\mathbf{w}) = \prod_{t=1}^{T} f_1(y_{1_t}|\mathbf{w}, \theta_1) \cdot \cdots \cdot f_N(y_{N_t}|\mathbf{w}, \theta_N) \cdot c(u_{1_t}, \cdots, u_{N_t}|\mathbf{w}, \theta_C), \tag{2.8}$$

where θ_i is vector of parameters from the marginal distribution of Y_i, $i = 1, \cdots, N$, and θ_C is the vector of parameters of the copula that links the marginal distributions of the time series $Y_1, \cdots, Y_N$.

Taking the log, we get

$$l(\theta_1, \cdots, \theta_N, \theta_C|\mathbf{w}) = \sum_{i=1}^{N}\sum_{t=1}^{T} \ln(f_i(y_{i_t}|\mathbf{w}, \theta_i)) + \sum_{t=1}^{T} \ln(c(u_{1_t}, \cdots, u_{N_t}|\mathbf{w}, \theta_C)) =$$

$$= l_{f_1}(\theta_1|\mathbf{w}) + \cdots + l_{f_N}(\theta_N|\mathbf{w}) + l_c(\theta_C|\theta_1, \cdots, \theta_N, \mathbf{w}). \tag{2.9}$$

Estimating jointly all the model parameters by the maximum likelihood method, however, can be a computationally very heavy task. To solve this problem, we use the IFM (Inference From Margins) method, proposed by Joe and Xu (1996), which, instead of maximizing the likelihood of the joint distribution with respect to all parameters, maximizes each of their $(N + 1)$ component parts separately.

Firstly, the marginal parameters from $f_1, \cdots, f_N$, given by $\theta_1, ..., \theta_N$, are estimated as

$$\hat{\theta_i} = \arg\max l_{f_i}(\theta_i|\mathbf{w}), \qquad i = 1, \cdots, N. \tag{2.10}$$

From these estimates, we can obtain estimates of $u_{i_t} = F_i(y_{i_t}|\theta_i, \mathbf{w}), i = 1, \cdots, N$, which in turn can be used to estimate the copula parameters via

$$\hat{\theta_c} = \arg\max l_c(\theta_c|\mathbf{w}, \hat{\theta_1}, \cdots, \hat{\theta_N}) =$$

$$= \arg\max \sum_{t=1}^{T} \ln(c(F_1(y_{1_t}|\hat{\theta_1}, \mathbf{w}), \cdots, F_N(y_{N_t}|\hat{\theta_N}, \mathbf{w})|\mathbf{w}, \theta_C)). \tag{2.11}$$

For copula models that have closed forms, it is easy to estimate the parameters of the copula via maximum likelihood, and this is the procedure adopted in this work. For factor copula models, however, this process becomes much more complicated. Since their likelihood function does not have closed form, Oh and Patton (2013) propose a way to approximate them numerically. This procedure will be further detailed in the Section 2.2.2.3.

Thus, in practice, the complete estimation process consists of the following steps: i) estimating models of the AR-GARCH class for the marginal time series via maximum likelihood, ii) obtaining their standardized residuals and next the estimated cdf's associated to those residuals, iii) with these values, estimating the copula by IFM method, via maximum likelihood.

2.2.2 Factor Copulas

When it comes to high dimensional data, namely a large number of variables, the most common analysis strategies are somehow based on dimension reduction. These strategies include popular techniques such as factor and principal component analyses, for example.

In the context of modeling the joint distribution of a multiple time series, Oh and Patton (2012) propose the factor copula approach, which allows for an acceptable degree of flexibility with a not very high number of parameters. Thus, a factor copula model with one factor for N time series or variables is given by

$$X_{i_t} = \lambda_{i_t} Z_t + \epsilon_{i_t}, \qquad i = 1, 2, \cdots, N, \tag{2.12}$$

where $Z_t \sim F_Z(\gamma_Z), \epsilon_{i_t} \sim F_\epsilon(\gamma_\epsilon), Z \perp \epsilon_i, \forall i$, γ_Z and γ_ϵ are vectors of parameters from the distributions of the common factor Z and idiosyncratic error ϵ, respectively, λ_{i_t} is the i-th factor loading on time t and $\mathbf{X}$ is a latent vector of variables whose copula is the same as the one of $\mathbf{Y}$, i.e.,

$$Y_t \sim H_Y(y_{1_t}, \cdots, y_{N_t}|\mathbf{w}) = C\left(F_{1_t}(y_{1_t}|\mathbf{w}), \cdots, F_{N_t}(y_{N_t}|\mathbf{w})\middle|\mathbf{w}\right) \tag{2.13}$$

and

$$X_t \sim H_X(x_{1_t}, \cdots, x_{N_t}) = C\left(F_{X_{1_t}}(x_{1_t}), \cdots, F_{X_{N_t}}(x_{N_t})\right). \tag{2.14}$$

In this work, we use combinations of Normal, Student's t and asymmetric or skew t-distributions, as parameterized in Hansen (1994), for modeling Z and ϵ, so that important features of financial r.v.'s can be generated, such as asymmetry, heavy tails and tail dependence.

2.2.2.1 Dependence Structure

As the number of time series grows, so does the number of parameters to be estimated by the model as well as the complexity of the optimization process. Oh and Patton (2012) pursue three different levels of dependency structures among the series to achieve a balance between the number of estimated model parameters and flexibility.

The most flexible level is the heterogeneous dependence, which provides that each time series has independent factor loadings from loadings of other time series. This type of structure has many parameters and may become computationally very expensive as the number of time series increases. The least flexible level equidependence, which assumes that all time series depend on each other the same way, presents a unique factor loading for all series, which is a bit too restrictive. On an intermediate level there is the block dependence, which splits a priori the set of time series into homogeneous groups in terms of dependence, that is, it is assumed that any two time series of two different groups are connected to each other in the same way as other two time series of the same two groups. This approach can drastically reduce the number of parameters as each group has a set of factor loadings, and no more each time series.

Thus, in general, the factor copula model can be defined as

$$X_{i_t} = \lambda_{g(i),t}(\gamma_\lambda) Z_t + \epsilon_{i_t}, \qquad i = 1, 2, \cdots, N, \tag{2.15}$$

where $g(i) \in 1, \cdots, G$ is the group to which the i-th time series belongs, G is the number of groups in which the time series are split into, and γ_λ is the set of parameters connected to λ. Note that $G = 1$ is a special case of equidependence, $G = N$ is the particular case of heterogeneous dependence, and $1 < G < N$ is the block dependence case. In this work we explore the three aforementioned approaches.

2.2.2.2 Time-varying parameters

Although dependency parameters of copulas are usually considered fixed in time, this may not properly portray the reality: the dependence coefficients of copulas may vary over time. To capture these dynamics, Patton (2006) proposes the use of time-varying copulas, with parameters following a restricted ARMA (1,10) model. On the other hand, Oh and Patton (2013) apply an adaptation of the GAS (Generalized Autoregressive Score) model, proposed by Creal, Koopman and Lucas (2011, 2013). The last approach will be explored in this work to model the dynamic behavior of the factor loadings of the factor copula.

Thus, the adaptation of the GAS model proposed by Oh and Patton (2013) is as follows:

$$\ln \lambda_{g,t} = \omega_g + \beta \ln \lambda_{g,t-1} + \alpha s_{g,t-1}, \qquad g = 1, \cdots, G, \tag{2.16}$$

where $s_{g,t} = \partial \ln c(\mathbf{u}_t; \lambda_t, \gamma_Z, \gamma_\epsilon)/\partial \lambda_{g,t}$, $\lambda_t = [\lambda_{1,t}, \cdots, \lambda_{G,t}]'$ and so $\gamma_\lambda = [\omega_g, \beta, \alpha]'$, with ω_g having dimension G.

Thus, there are $G + 2$ parameters to be estimated in this part of the model. However, in the heterogeneous dependence case ($G = N$), we have $N + 2$ parameters, which is a huge number for the optimization, with respect to computational effort, when $N = 100$ for example. That is why Oh e Patton (2013) propose that in this case an alternative strategy is used for reducing the number of parameters, based on the assumption that the process λ_t is strictly stationary.

Therefore, when $G = N$ and so $g(i) = i$, we use the following expression:

$$\ln \lambda_{i,t} = \mathbb{E}[\ln \lambda_{i,t}](1 - \beta) + \beta \ln \lambda_{i,t-1} + \alpha s_{i,t-1}, \qquad i = 1, \cdots, N. \tag{2.17}$$

One can then estimate $\mathbb{E}[\ln \lambda_{i,t}]$ in the following way:

$$\mathbb{E}[\ln \hat{\lambda}_{i,t}] = \arg\min_{\mathbf{a}} m_T(\mathbf{a})' m_T(\mathbf{a}), \tag{2.18}$$

where $m_T(\mathbf{a}) = vech\Phi(G(\exp(\mathbf{a}))) - vech\hat{\rho}_u^S$, $\hat{\rho}_u^S$ is the estimated Spearman correlations matrix of $\mathbf{u}$, $\Phi(G(\lambda)) = \rho_X$ and ρ_X is the Spearman correlations matrix of a set of variables $\mathbf{X}$ given by (2.12) with vector of factor loadings λ. Thus we have that the set of parameters governing the process of λ is summarized to $\gamma_\lambda = [\beta, \alpha]'$.

2.2.2.3 Factor Copulas Estimation

Factor copulas implied by (2.12) usually do not have closed form and therefore their estimation is impaired. Oh and Patton (2012) propose a method of moments based parameter estimation for this kind of copulas, which is centered on the comparison of the observed data moments with simulated series moments. This method produces good results. Nevertheless it can not be used for time-varying factor copulas.

Oh and Patton (2013) propose another estimation method, which is based on the maximization of a numerical approximation of the likelihood function. This function can be obtained by the density of the factor copula of X_t:

$$c_t(u_{1_t},\cdots,u_{N_t}) = \frac{f_{X_t}\left(F_{X_{1_t}}^{(-1)}(u_{1_t}),\cdots,F_{X_{N_t}}^{(-1)}(u_{N_t})\right)}{g_{1_t}\left(F_{X_{1_t}}^{(-1)}(u_{1_t})\right)\times\cdots\times g_{N_t}\left(F_{X_{N_t}}^{(-1)}(u_{N_t})\right)}, \tag{2.19}$$

where $f_{X_t}(x_{1_t},\cdots,x_{N_t})$ is the joint pdf of X_t, $g_{i_t}(x_{i_t})$ is the marginal pdf of X_{i_t} and $c_t(u_{1_t},\cdots,u_{N_t})$ is the copula pdf.

In order to obtain this density value, we define:

$$g_{i_t}(x_{i_t}) = \int_0^1 f_\epsilon(x_{i_t} - \lambda_{i_t}F_Z^{(-1)}(m))dm, \tag{2.20}$$

$$F_{X_{i_t}}(x_{i_t}) = \int_0^1 F_\epsilon(x_{i_t} - \lambda_{i_t}F_Z^{(-1)}(m))dm \tag{2.21}$$

and

$$f_{X_t}(x_{1_t},\cdots,x_{N_t}) = \int_0^1 \prod_{i=1}^{N} f_\epsilon(x_{i_t} - \lambda_{i_t}F_Z^{(-1)}(m))dm. \tag{2.22}$$

So, we have the log of the likelihood function of the copula:

$$l_c(\theta_C|\hat{\theta_1},\cdots,\hat{\theta_N},\mathbf{w}) = \sum_{t=1}^{T}\ln\Big(c(F_1(y_{1_t}|\hat{\theta_1},\mathbf{w}),\cdots,F_N(y_{N_t}|\hat{\theta_N},\mathbf{w})|\mathbf{w},\theta_C)\Big), \tag{2.23}$$

where $\theta_C = [\gamma_Z,\gamma_\epsilon,\gamma_\lambda]'$ is the set of copula parameters to be estimated.

2.2.3 Marginal Modeling

For each marginal, we fit a $AR(1) - GARCH(1,1)$ model. In the description below, we suppress the i index of each variable. These models can be described by

$$(Y_t - \mu) = \phi(Y_{t-1} - \mu) + \varepsilon_t, \tag{2.24}$$

$$\varepsilon_t = \sigma_t K_t, \tag{2.25}$$

$$\sigma_t^2 = \omega_1 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2, \tag{2.26}$$

where $K_t \sim Dist(0,1)$ are the standardized residuals, which are independent of ε_{t-l}, $l \geq 1$, for all t. Besides, $\sigma_t^2 = Var(\varepsilon_t|\varepsilon_{t-1},\varepsilon_{t-2},\cdots)$ denotes the conditional variance, $\omega_1 \geq 0$ the intercept and ε_t the conditional mean corrected returns.

Note that the standardized residuals K_t for the ARMA-GARCH models can also be written as

$$K_t = \frac{Y_t - \mathbb{E}(Y_t|\mathbf{W}_{t-1})}{\sqrt{Var(Y_t|\mathbf{W}_{t-1})}} = \frac{Y_t - \mathbb{E}(Y_t|\mathbf{W}_{t-1})}{\sigma_t}, \tag{2.27}$$

where $\mathbf{W}_{t-1} = Y_{t-1}, \varepsilon_{t-1}, Y_{t-2}, \varepsilon_{t-2}, \cdots$.

The standardized residuals K_t can follow any distribution with zero mean and variance 1. Usually, it is assumed that they follow a standard normal distribution. However, they may have heavier tails, then following a Student's t-distribution, for example, which is adopted in this work.

The Student GARCH model was first described by Bollerslev (1987) as an alternative to normal distribution to the standardized residuals. The probability density function of a standardized t-distribution with ν degrees of freedom, is given by

$$f_K(k) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)}\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{k^2}{(\nu-2)}\right)^{-\left(\frac{(\nu+1)}{2}\right)}. \tag{2.28}$$

It is symmetric, but possesses heavier tails than the normal distribution, with kurtosis equal to $6/(\nu-4)$ for $\nu > 4$.

2.2.4 Risk Measures

In this work, we aim to find the best model to fit BOVESPA stock returns joint distribution, in order to best calculate its portfolios risk measures. Several copula models are estimated for our data set, and then, from these estimates, portfolio risk measures are evaluated and compared along with the models. We adopt Value at Risk and Expected Shortfall measures of risk.

2.2.4.1 Value at Risk

A commonly used risk measure in the financial market is the Value at Risk (VaR). Informally, it can be said that, for a level of confidence α and a given portfolio, there is an α probability that there is a loss at time t, in terms of portfolio return, larger than its VaR_t^{α}. That is, the VaR can be seen as the opposite (negative) to the α-th quantile of the portfolio returns distribution. This is the measure used by Basel II and Basel III Accord for setting capital requirements referred to market risk.

Formally, we have the VaR definition as follows:

$$VaR_t^{\alpha} = \arg\min_{x\in\mathbb{R}}\{x : \mathbb{P}(-Y_t \leq x) \geq 1-\alpha\}, \qquad 0<\alpha<1, \tag{2.29}$$

where Y_t is the portfolio return at time t.

The portfolio return distribution H should be known for one to calculate the VaR, so choosing the quantile associated to α. Since this distribution can not be calculated theoretically for factor copulas and it is a difficult task for other copulas, we have to estimate the α quantiles of H via simulation. Hence, we have to adopt a procedure which consists on simulating data sets of the the joint data via the estimated subjacent copula and marginal models for each fixed t. Thus we obtain the simulated distribution of the portfolio returns for each t and then estimate the VaR_t^{α}.

Therefore we can evaluate the fit of each model by calculating the violations of the VaR, i.e., the proportion of values of the observed portfolio return which are smaller than these estimates of $-VaR_t^{\alpha}$. If the model, including the marginal distributions and the copula, is well specified, this proportion should be close to the specified α. To statistically test the VaR estimates we use Christoffersen test (Christoffersen, 1998), whose null hypothesis is that $\mathbb{P}(Y_t < -VaR_t^{\alpha}) = \alpha$ against the alternative $\mathbb{P}(Y_t < -VaR_t^{\alpha}) \neq \alpha$.

2.2.4.2 Expected Shortfall

The measure of risk VaR, however, is not subadditive, i.e., a combination of assets does not necessarily have a lower Value at Risk than the sum of their Values at Risk, and so its said not to be a coherent risk measure. Thereby this risk measure might produce results that go against the principle of finance that investment diversification reduces the risk of a portfolio. Thus the risk measure known as Conditional Value at Risk or Expected Shortfall (ES), proposed by Rockfellar and Uryasev (2000), comes into play. This measure, to a level of confidence α, is given by the negative of the expectation of returns that are below $-VaR_t^{\alpha}$, or, more formally:

$$ES_t^{\alpha} = \mathbb{E}[-Y_t | Y_t \leq -VaR_t^{\alpha}]. \tag{2.30}$$

where Y_t is the portfolio return at time t.

Evaluating the proposed models for the Expected Shortfall, we can compute the ES excess, given by $R_j = Y_j - ES_j^\alpha$, where Y_j corresponds to the j-th violation of the VaR, and ES_j^α is the simulated ES via the assumed model. Thus, it is expected that the closer to zero this excess is, the better the model fits the data. To statistically test these fits we used the test proposed by McNeil and Frey (2010), whose null hypothesis is that the variable R_j has zero mean, against the alternative that its mean is below zero.

2.3 Empirical Results

Data analysis is carried out with MATLAB and R statistical softwares. This section first presents the description of the time series data, then shows the models used for modeling the marginal distributions of the time series, following with the results for the copula estimation and finally the estimated risk measures.

2.3.1 Data Description

Our data is comprised of 44 financial time series of assets that were negotiated by Sao Paulo Stock Exchange, BOVESPA, on June/26/2014 and presented information throughout the study period, which spans from Jan/02/2008 to June/26/2014, corresponding to 1604 observations for each time series.

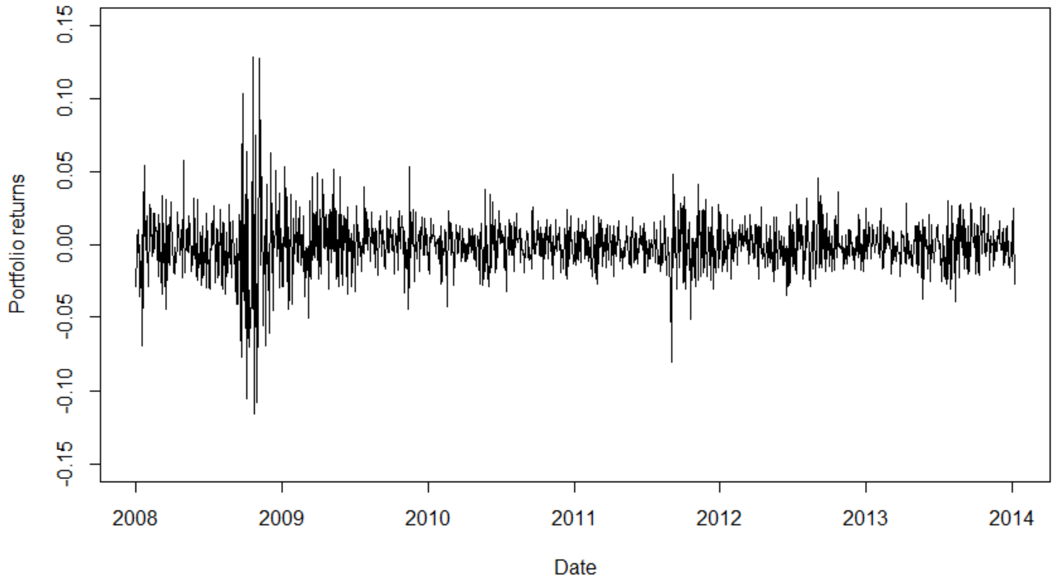
The 44 studied assets were also grouped into 6 blocks according to the market segment in which they operate, for the estimation of factor copula with block dependence. Each block is composed by one or more related economic sectors according to the division made by BOVESPA. The categorization of the assets into blocks is arbitrary and could have been done differently, but this was an ad-hoc choice based on the expectation that companies in the same market segment have similar dependence structures. The names of the assets used in this work as well as the block names in which they were categorized are displayed in Table 2.9, in the Appendix of this paper.

For the analysis and modeling, we use the log-price returns, obtained as

$$Y_{i_t} = \ln(P_{i_t}) - \ln(P_{i_{t-1}}), \tag{2.31}$$

where Y_{i_t} is the asset i return at time t, and P_{i_t} is the price of asset i at time t.

Figure 2.1: Portfolio returns



Source: Elaborated by the author based on data from Bolsa de Valores de São Paulo - BOVESPA (2008-2014).

The portfolio used in this work has equal weights for all assets. Therefore, its approximate return can be given by $Y_t \approx \ln\left(\frac{\sum_{i=1}^{N} \exp(Y_{i_t})}{N}\right)$, where N is the number of assets that comprise the portfolio. This portfolio returns are shown in Figure 2.1 for the period of our analyses. We may note a period of great volatility in the end of 2008, when the US subprime financial crisis took place.

The rows of Table 2.1 show the statistics describing the stock returns we are working with, and Figures 2.6 to 2.8 in the Appendix show their behavior over time. In the columns of the table we can see the mean, minimum, maximum and some other descriptive statistics calculated for all 44 assets and also these same statistics calculated for the portfolio returns. For example, the mean of the sample standard deviations of the 44 assets is 0.0325, which is displayed in the second row of the second column of the table. Here we highlight the fact that, on average, these assets present negative asymmetry. More specifically, 80% of them have this feature, mainly caused by certain points in time when the value plummeted. Another featured property is the kurtosis, with an average of approximately 127, and only one of the assets presented returns with kurtosis not larger than 3, which corresponds to a Normal distribution. This indicates the strong presence of heavy tails in the distributions of these returns, which will be modeled using Student's t-distribution for the error of the AR-GARCH model used to adjust their marginals.

Table 2.1: Summary statistics for stock returns

STATISTIC	MEAN	MIN.	1^{st} Q	MEDIAN	3^{rd} Q	MAX.	PORTF.
Mean	-0,0005	-0,0038	-0,0008	-0,0004	-0,0001	0,0005	-0,0001
Std. Dev.	0,0325	0,0183	0,0256	0,0286	0,0353	0,1025	0,0182
Skewness	-4,3668	-28,7708	-6,1240	-0,7511	-0,0629	0,5153	0,0976
Kurtosis	126,9888	1,6336	5,4137	14,3257	133,9783	1032,1558	7,7852
Minimum	-0,5257	-2,9855	-0,6976	-0,2827	-0,1817	-0,0662	-0,1159
5%	-0,0418	-0,0740	-0,0455	-0,0399	-0,0353	-0,0286	-0,0273
25%	-0,0150	-0,0258	-0,0168	-0,0140	-0,0119	-0,0092	-0,0092
Median	-0,0003	-0,0024	-0,0004	0,0000	0,0000	0,0008	-0,0001
75%	0,0143	0,0102	0,0120	0,0135	0,0156	0,0214	0,0092
95%	0,0419	0,0273	0,0353	0,0399	0,0467	0,0649	0,0256
Maximum	0,2144	0,1094	0,1410	0,1728	0,2131	1,7391	0,1289

Source: Elaborated by the author.

2.3.2 Marginal Distributions

The distributions of all sets of returns are modeled using an $AR(1)-GARCH(1,1)$ model, for which the estimated parameters are shown in Table 2.2. As we can see, the estimated mean values are close to zero for all series, indicating that all series revolve around zero, as expected. Is is also important to note that most degrees of freedom values are small, indicating that it is actually useful to use a heavy-tailed distribution for the residuals, rather than a Normal distribution.

To validate the copula analysis, it is necessary to test whether the standardized residuals follow the assumed distribution (Student's t) to ensure that their integral transform probability have Uniform $[0,1]$ distributions. To do so, we perform Kolmogorov-Smirnov (KS) tests in the standardized residuals for which the p-values are also shown in Table 2.2. In all cases, the p-value was above 0.1, indicating that these residuals distributions are close to Student's t.

2.3.3 Copulas

After estimating the chosen models for each time series, we use the integral probability transforms of the standardized residuals, which follow Uniform $[0,1]$ distributions, for the copula analysis.

The copulas modeling all 44 assets in question are estimated using several specifications, namely: Normal, Student's t, Gumbel, Clayton and Frank, described in Cherubini,

Table 2.2: Models fitted to the BOVESPA asset returns time series

TIME SERIES	PARAMETERS						KS (p-value)
	μ	ϕ	ω_1	α_1	β_1	ν	
CCRO3	7.1E-4	-0.063	1.0E-4	0.229	0.584	4.482	0.940
CYRE3	-3.0E-4	0.004	5.4E-6	0.086	0.910	19.647	0.640
EVEN3	-2.9E-4	0.034	3.0E-5	0.112	0.853	9.971	0.580
GFSA3	-1.4E-3	-0.004	6.7E-5	0.142	0.821	7.740	0.565
GOLL4	-1.1E-3	-0.005	2.2E-5	0.062	0.920	6.898	0.699
MRVE3	1.7E-4	0.049	1.7E-4	0.166	0.711	6.505	0.664
PDGR3	-7.5E-4	0.014	1.3E-4	0.161	0.749	6.267	0.472
RSID3	-1.6E-3	0.031	5.2E-5	0.079	0.885	7.272	0.756
CSAN3	7.7E-4	-0.001	7.0E-6	0.083	0.906	8.572	0.562
JBSS3	1.5E-5	-0.062	1.8E-5	0.069	0.915	8.934	0.865
MRFG3	-6.7E-4	0.074	1.3E-5	0.058	0.933	4.528	0.979
CRUZ3	3.5E-4	-0.033	1.0E-4	0.263	0.522	5.460	0.129
NATU3	5.4E-4	-0.045	6.1E-6	0.059	0.927	15.870	0.832
LAME4	3.6E-4	-0.026	1.5E-5	0.072	0.905	6.103	0.671
LREN3	6.6E-4	-0.001	7.3E-6	0.073	0.917	11.644	0.109
BBAS3	1.8E-4	0.072	1.1E-5	0.084	0.897	10.636	0.415
BBDC3	1.6E-5	0.003	2.7E-5	0.079	0.860	4.993	0.745
BBDC4	1.0E-4	0.020	9.7E-6	0.047	0.929	5.609	0.949
BRAP4	-2.4E-4	0.042	5.3E-6	0.052	0.940	7.481	0.644
BRML3	1.7E-4	-0.003	9.1E-5	0.182	0.695	6.132	0.307
ITSA4	-2.4E-5	0.009	9.5E-6	0.067	0.912	6.850	0.553
BRKM5	7.7E-5	0.003	1.7E-5	0.091	0.888	6.897	0.581
CSNA3	-8.3E-4	0.066	6.2E-5	0.173	0.773	4.544	0.553
GGBR4	-7.6E-4	0.030	3.4E-6	0.009	0.983	7.350	0.601
GOAU4	-8.0E-4	0.017	3.7E-6	0.009	0.984	7.467	0.372
MMXM3	-1.9E-3	0.021	6.1E-4	0.566	0.342	3.656	0.143
SUZB5	-6.6E-4	0.035	9.3E-6	0.052	0.934	7.356	0.856
USIM5	-1.2E-3	0.084	6.4E-5	0.093	0.843	5.166	0.839
VALE3	-1.6E-4	0.019	4.9E-6	0.048	0.943	7.652	0.609
VALE5	1.1E-4	0.012	5.6E-6	0.059	0.931	6.248	0.694
EMBR3	1.6E-5	-0.042	8.5E-6	0.053	0.932	5.259	0.605
RENT3	5.9E-4	0.007	1.5E-5	0.096	0.880	9.070	0.802
PETR3	-6.9E-4	0.009	2.8E-5	0.091	0.866	5.364	0.703
PETR4	-3.9E-4	0.010	2.6E-5	0.119	0.841	5.679	0.721
CESP6	5.9E-4	-0.018	1.7E-5	0.096	0.877	5.942	0.442
CMIG4	7.6E-4	-0.021	2.0E-5	0.040	0.918	3.691	0.869
CPFE3	1.5E-4	-0.094	4.6E-5	0.187	0.686	5.743	0.881
CPLE6	3.4E-4	-0.035	1.3E-5	0.081	0.886	8.834	0.585
ELET3	-8.7E-4	0.044	2.0E-5	0.151	0.838	4.673	0.961
ELET6	-5.3E-4	0.032	1.2E-5	0.082	0.901	4.610	0.914
ENBR3	3.6E-4	-0.098	4.8E-5	0.155	0.745	4.974	0.817
LIGT3	2.5E-4	-0.040	7.9E-6	0.056	0.928	5.350	0.597
SBSP3	1.0E-3	-0.104	7.0E-5	0.184	0.710	5.088	0.775
TBLE3	4.1E-4	-0.121	5.2E-6	0.060	0.923	7.156	0.933

Source: Elaborated by the author.

Luciano and Vecchiato (2004). Besides, the factor copulas "Normal-Normal", "t-t" and "skew t-t" (where K is the common factor with skewed Student's t-distribution and ε is the error, which follows a Student's t-distribution), proposed by Oh and Patton (2013), are also used. We also explore the 3 different dependence structures for the factor copulas as mentioned in Section 2.2.2.1 and the possibility of whether or not they vary in time, as in Section 2.2.2.2. The copulas that do not vary in time are called static, and the ones that vary are named dynamic. Static copulas are estimated in the same manner as dynamic ones, but constraining their parameters α and β to $\alpha = 0$ and $\beta = 0$.

The model of static factor copula with heterogeneous dependence has too many parameters and we can not use the same approach used in the estimation of the dynamic copula with heterogeneous dependence (see Section 2.2.2.2), so we decide not to estimate it because of the huge computational burden.

The estimated parameters, the log-likelihood (LL) and the Akaike Information Criterion (AIC) for each copula are presented in Tables 2.3 to 2.6. Parameterization of the non-factor copula is performed according to Cherubini, Luciano and Vecchiato (2004). Note that for the Normal and Student's t copulas, the parameter matrix R is a matrix where all the off-diagonal values are given by ρ. Also note that for Student's t copula we present the estimated inverses of the degrees of freedom instead of the degrees of freedom themselves, because for the maximization of the likelihood process the inverses are more convenient, since they must lie within the interval $[0, 0.5)$ so that the distribution has finite variance.

The AIC presented in the tables below is calculated as $AIC = \frac{-2LL+2p}{T}$, where LL is the log-likelihood, p is the number of parameters of the copula, and T is the number of unities of time. This criterion shows the balance between fit and number of parameters. The lower the AIC, the better the model balances fit and number of parameters.

One can note that the Normal copula and the Normal factor copula are equal according to the following equation involving its parameters: $\rho = \frac{(\exp\omega)^2}{1+(\exp\omega)^2}$. Despite this equality, when it comes to estimates, it becomes only an approximation, since the calculation of the likelihood of the factor copula is only approximate, due to the various simplifications made in order to make it feasible.

The traditional copula that best fits the data, according to the AIC, is the Student's t copula, then the Normal copula. This indicates that adding the possibility of heavy tails to the Normal model in fact improves its fit to the data. On the other hand, Gumbel, Clayton and Frank copulas found to be poor to adjust the dependence for this set of BOVESPA assets.

Table 2.3: Parameters Estimates of the Standard Copulas

MODEL	PARAMETER	ESTIMATE	LL	AIC
Normal	ρ	0.3543	12,906.5	-16.10
Student's t	$1/\nu$	0.0319	13,612.4	-16.98
	ρ	0.3682		
Gumbel	α	1.2221	9,770.1	-12.19
Clayton	α	0.1875	8,922.1	-11.13
Frank	α	1.8645	9,337.2	-11.65

Source: Elaborated by the author.

Among the factor copula, the combination of distributions that best fits the data, according to the AIC, is the "skew t-t" and therefore the subsequent analyzes only take into account these versions of the factor copulas. Its estimated inverse of degrees of freedom, both of the common factor and the idiosyncratic error, are greater than zero, which is equivalent to small degrees of freedom, indicating that the Normal copula is in fact inappropriate. The asymmetry of the common factor is also different from zero, specifically lower than zero, which is consistent with lower tail dependence, i.e., the assets are more dependent to each other when the prices are going down, which is a stylized fact of stock returns.

Obviously, the dependence structure that best fits the data is the heterogeneous dependency, and this remained the best model even penalizing for its huge number of parameters, as shown by the values of the AIC. Likewise, it is observed that the dynamic models present best fits to the data than the static ones. When it comes to dynamic models we see that the β parameter is often close to 1, and greater in the block dependence case, indicating strong temporal persistence, specially in the latter case.

Figure 2.2 shows the evolution of the equidependent dynamic skew t-t factor copula loadings. The equidependent static skew t-t factor copula loading is also shown in this figure as the red line. We can see that the dynamic loadings vary mostly between 0.5 and 1, around 0.7415, the value of the static factor loading. The biggest exception is the value correspondent to Aug/09/2011. At this date, the BOVESPA index had one of its biggest peaks in recent times, and all the assets performed better than its marginal models would predict, generating very large residuals. Since all the 44 assets behaved this way, they turned out to be very correlated, which resulted in a great value of λ_t.

Table 2.4: Parameters Estimates of the Equidependent Factor Copulas

PARAMETER	MODEL Factor Normal - Normal		Factor t - t		Factor Skew t - t	
	Static	Dynamic	Static	Dynamic	Static	Dynamic
ω	-0.2581	-0.2844	-0.2777	-0.0646	-0.2991	-0.0419
α	-	0.0388	-	0.0168	-	0.0166
β	-	0.9520	-	0.7719	-	0.8326
$1/\nu_Z$	-	-	0.1777	0.1932	0.1436	0.2432
$1/\nu_\epsilon$	-	-	0.1473	0.1502	0.1503	0.1503
ϕ	-	-	-	-	-0.1555	-0.1109
LL	13,032.7	13,045.6	13,589.1	13,670.2	13,622.7	13,706.0
AIC	-16.26	-16.28	-16.95	-17.06	-17.00	-17.10

Source: Elaborated by the author.

Table 2.5: Parameters Estimates of the Block Dependent Factor Copulas

PARAMETER	MODEL Factor Normal - Normal		Factor t - t		Factor Skew t - t	
	Static	Dynamic	Static	Dynamic	Static	Dynamic
ω_1	-0.2659	-0.0197	-0.2660	-0.0182	-0.2786	-0.0132
ω_2	-0.5123	-0.0531	-0.4805	-0.0325	-0.5067	-0.0283
ω_3	-0.0065	-0.0036	0.0134	0.0020	-0.0055	0.0000
ω_4	-0.0550	-0.0097	-0.0496	-0.0021	-0.0455	-0.0024
ω_5	-0.2804	-0.0255	-0.2821	-0.0158	-0.2884	-0.0144
ω_6	-0.5473	-0.0485	-0.5612	-0.0410	-0.5636	-0.0305
α	-	0.0210	-	0.0149	-	0.0114
β	-	0.9073	-	0.9317	-	0.9480
$1/\nu_Z$	-	-	0.1622	0.1546	0.1349	0.1430
$1/\nu_\epsilon$	-	-	0.1504	0.1472	0.1503	0.1510
ϕ	-	-	-	-	-0.1565	-0.1742
LL	13,585.4	13,755.2	14,165.5	14,314.0	14,207.3	14,370.7
AIC	-16.94	-17.16	-17.66	-17.86	-17.71	-17.93

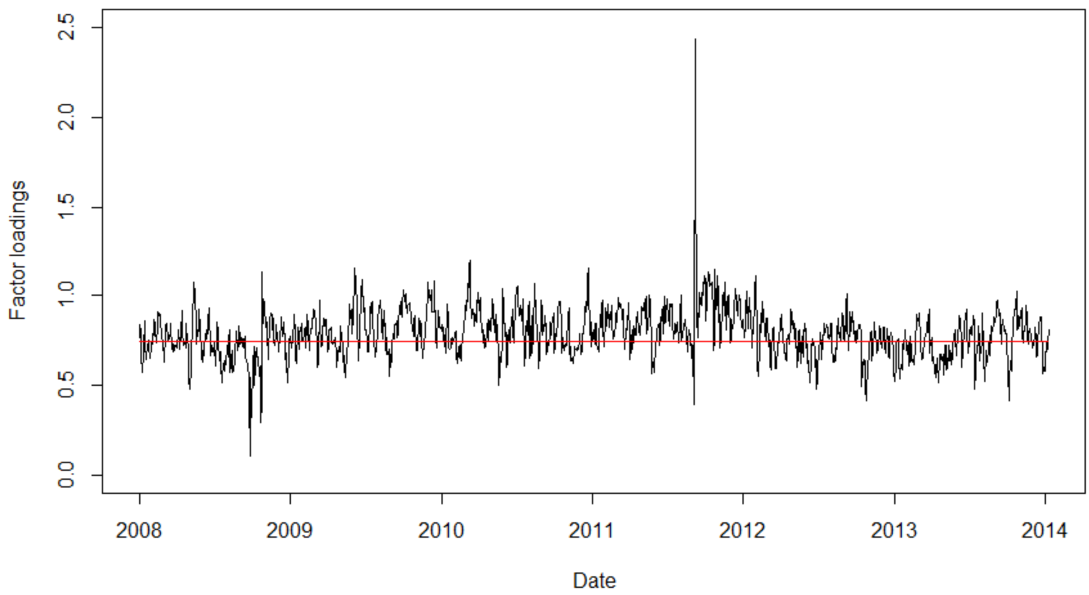
Source: Elaborated by the author.

Table 2.6: Parameters Estimates of the Heterogeneous Dependent Factor Copulas

PARAMETER	MODEL		
	Factor Normal - Normal	Factor t - t	Factor Skew t - t
ω_1	-0.0092	-0.0082	-0.0082
ω_2	-0.0002	-0.0001	-0.0001
ω_3	-0.0052	-0.0047	-0.0047
ω_4	-0.0007	-0.0006	-0.0006
ω_5	-0.0062	-0.0056	-0.0056
ω_6	-0.0030	-0.0027	-0.0027
ω_7	-0.0028	-0.0025	-0.0025
ω_8	-0.0010	-0.0009	-0.0009
ω_9	-0.0069	-0.0062	-0.0062
ω_{10}	-0.0057	-0.0051	-0.0051
ω_{11}	-0.0099	-0.0089	-0.0089
ω_{12}	-0.0103	-0.0092	-0.0092
ω_{13}	-0.0097	-0.0087	-0.0087
ω_{14}	-0.0014	-0.0012	-0.0012
ω_{15}	-0.0035	-0.0031	-0.0031
ω_{16}	-0.0012	-0.0010	-0.0010
ω_{17}	0.0009	0.0008	0.0008
ω_{18}	0.0027	0.0024	0.0024
ω_{19}	0.0019	0.0017	0.0017
ω_{20}	-0.0080	-0.0071	-0.0071
ω_{21}	0.0023	0.0020	0.0020
ω_{22}	-0.0039	-0.0035	-0.0035
ω_{23}	0.0009	0.0008	0.0008
ω_{24}	0.0019	0.0017	0.0017
ω_{25}	0.0018	0.0016	0.0016
ω_{26}	-0.0057	-0.0051	-0.0051
ω_{27}	-0.0058	-0.0052	-0.0052
ω_{28}	-0.0011	-0.0010	-0.0010
ω_{29}	0.0019	0.0017	0.0017
ω_{30}	0.0021	0.0018	0.0018
ω_{31}	-0.0092	-0.0082	-0.0082
ω_{32}	-0.0059	-0.0053	-0.0053
ω_{33}	-0.0001	-0.0001	-0.0001
ω_{34}	0.0001	0.0001	0.0001
ω_{35}	-0.0076	-0.0068	-0.0068
ω_{36}	-0.0067	-0.0060	-0.0060
ω_{37}	-0.0056	-0.0050	-0.0050
ω_{38}	-0.0051	-0.0046	-0.0046
ω_{39}	-0.0044	-0.0039	-0.0039
ω_{40}	-0.0046	-0.0041	-0.0041
ω_{41}	-0.0088	-0.0079	-0.0079
ω_{42}	-0.0082	-0.0074	-0.0074
ω_{43}	-0.0066	-0.0059	-0.0059
ω_{44}	-0.0152	-0.0136	-0.0136
α	0.0292	0.0337	0.0340
β	0.9858	0.9873	0.9873
$1/\nu_Z$	-	0.1148	0.0931
$1/\nu_\epsilon$	-	0.0890	0.0887
ϕ	-	-	-0.1370
LL	14,762.6	15,282.8	15,311.3
AIC	-18.42	-19.01	-19.10

Source: Elaborated by the author.

Figure 2.2: Factor loadings according to the equidependent dynamic and static factor "skew t-t" copula



Source: Elaborated by the author.

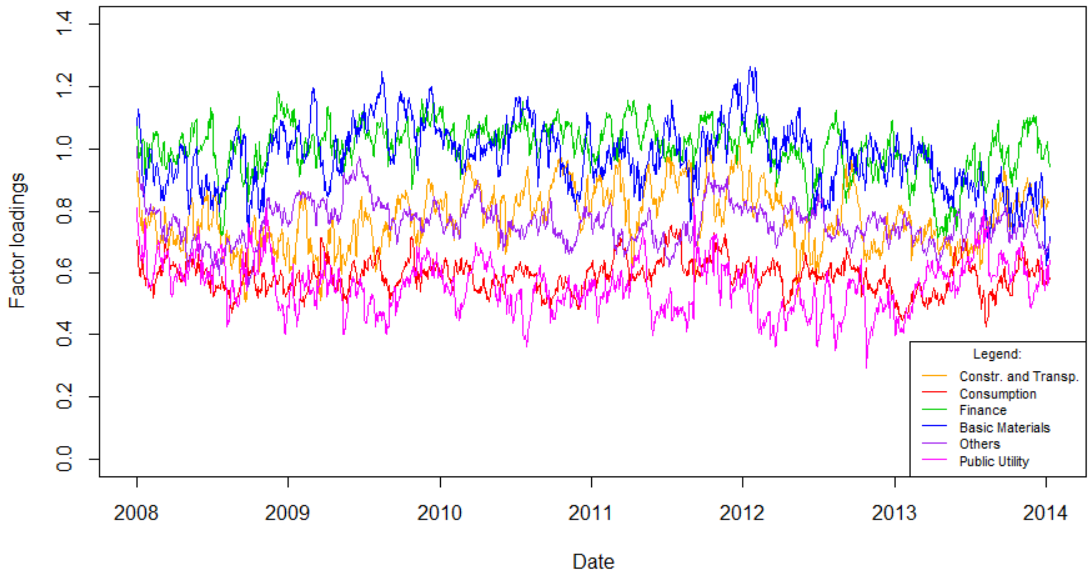
Figure 2.3 presents the evolution of the block dependent dynamic skew t-t factor copula loadings. Each block presents a different level around which their factor loadings vary over time. Finance and Basic Materials are the groups with the biggest levels, while Public Utility and Consumption are the ones with the smallest levels. Finance is a group composed only by banks and Basic Materials have mostly steel and mining companies. They are, in fact, the most homogeneous groups and therefore it is expected that they have bigger correlations and, thus, larger factor loadings.

2.3.4 Risk Measures

After estimating the copulas, we compute the Value at Risk and Expected Shortfall. These risk measures are calculated for a portfolio with equal weights for each component asset. For the calculation of risk measures, though, we dismissed the first year of the time series, 2008, in order to avoid problems caused by the influence of the chosen points to start the simulated series.

To assess the quality of the estimated Value at Risk under a specific copula model, we compute the proportion of observed values along the time series that violate the simulated dynamic VaR_t^α from the assumed model. We run 500 simulations of each model for each

Figure 2.3: Factor loadings according to the block dependent dynamic factor “skew t-t” copula



Source: Elaborated by the author.

time period t. The better the fit of the model, the closer this proportion should be to the α value. Levels of violation for each tested model are presented in Table 2.7 with their respective p-values from the Christoffersen test in parentheses.

Gumbel, Frank and Clayton copulas have very poor performance with regard to the calculation of VaR, as it can be seen in the Table 2.7: their proportion of VaR violations are far superior to α values. So we can see that they underestimate the Value at Risk of the portfolio and this is statistically significant according to the Christoffersen test, that rejects its null hypothesis in all these cases.

The Normal and Student's t copulas have a slightly better performance, but overestimate increasingly the VaR, as the α value increases. We can even see that the Christoffersen test does not reject the null hypothesis that $\mathbb{P}(Y_t < -VaR_t^\alpha) = \alpha$ for $\alpha = 0.01$, but rejects to $\alpha = 0.10$.

Skew t-t factor copulas show similar performance to the Normal and Student's t copulas for the levels $\alpha = 0.01$ and $\alpha = 0.025$, but have better results for $\alpha = 0.05$ and $\alpha = 0.1$. In addition, the null hypothesis of Christoffersen test is not rejected for any α level of any of the skew t-t factor copulas, reinforcing its quality to estimate the Value at Risk of this portfolio.

Among the factor copulas, however, there seems to be no appreciable differences in the results obtained from the different tested dependence structures or between their static and dynamic versions. What can be seen is only a slight decline in the VaR estimation with $\alpha = 0.01$ and a slight improvement in VaR estimation with $\alpha = 0.10$ when they are estimated by the equidependent dynamic factor "skew t-t" copula in relation to the other factor copula models. However, in general, all factor copulas tested here perform very similarly.

As we can see in Figure 2.3, there seems to be a difference among the levels of the factor loadings of the different blocks, and we can see the same pattern in the factor loadings of the heterogeneous model. This fact could bring on improvements on the VaR estimation. However, if there was an improvement, this must have been so small that it was not captured in the analysis presented here.

As regards to the dynamic models, it is expected that they do not bring very large improvements to the model, although they fit better the data, according to the AIC. The level of factor loadings in all types of dependence does not change much over time, and so the estimated Value at Risk does not undergo large changes to adjust the characteristics of each period in time. Thus the dynamic models have very similar behavior to static ones.

Table 2.7: VaR violations for each model at level α

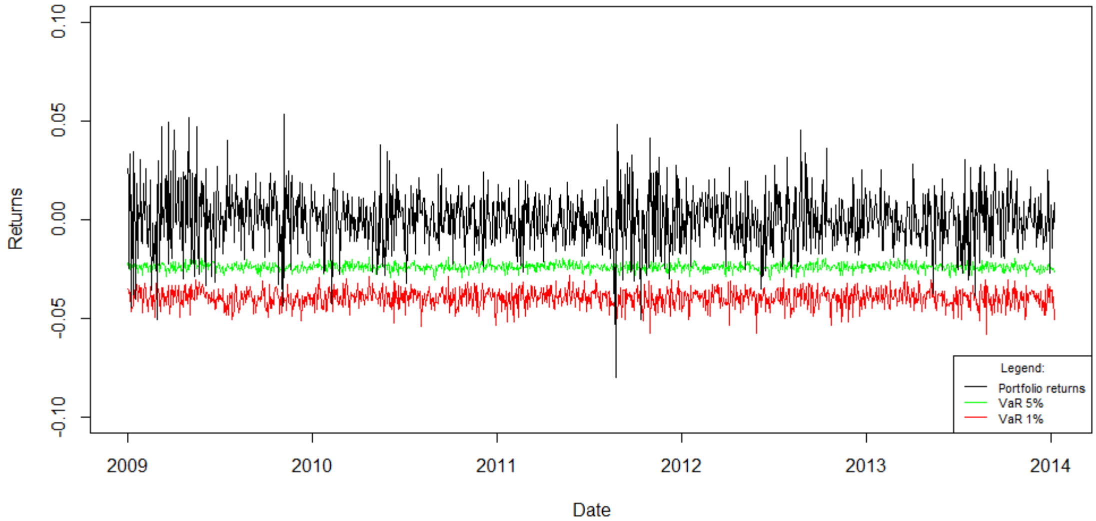
MODEL	LEVEL α			
	0.010	0.025	0.050	0.100
Normal	0.0088 (0.812)	0.0133 (0.008)	0.0376 (0.069)	0.0768 (0.009)
Student's t	0.0081 (0.702)	0.0162 (0.061)	0.0362 (0.007)	0.0753 (0.003)
Gumbel	0.0487 (0.000)	0.0716 (0.000)	0.0975 (0.000)	0.1404 (0.000)
Clayton	0.0177 (0.022)	0.0450 (0.000)	0.0827 (0.000)	0.1574 (0.000)
Frank	0.0613 (0.000)	0.0739 (0.000)	0.0938 (0.000)	0.1212 (0.031)
Factor (equidep.)	0.0059 (0.248)	0.0192 (0.215)	0.0421 (0.350)	0.0916 (0.277)
Factor (block dep.)	0.0066 (0.392)	0.0177 (0.125)	0.0406 (0.251)	0.0909 (0.321)
Dyn. factor (equidep.)	0.0036 (0.027)	0.0184 (0.167)	0.0428 (0.424)	0.1012 (0.438)
Dyn. factor (block dep.)	0.0066 (0.392)	0.0162 (0.061)	0.0413 (0.305)	0.0909 (0.321)
Dyn. factor (heter. dep.)	0.0066 (0.392)	0.0162 (0.061)	0.0436 (0.358)	0.0894 (0.297)

Source: Elaborated by the author.

To illustrate the Value at Risk behavior over time, we use Figures 2.4 and 2.5, which show the portfolio returns and their risk measures as estimated by the block dependent dynamic and static factor "skew t-t" copula models.

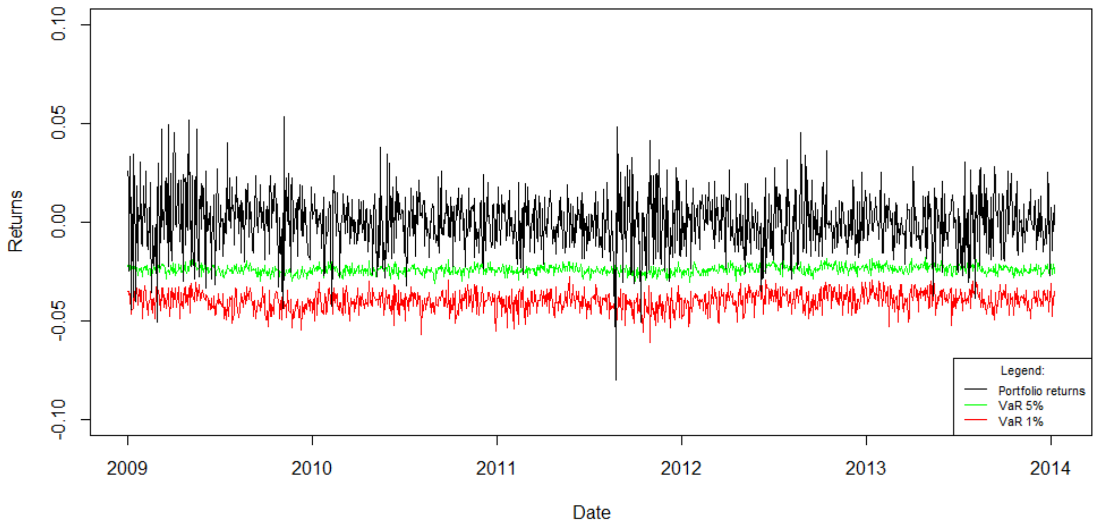
Evaluating these same models for the Expected Shortfall, we compute the ES excess, given by $R_j = Y_j - ES_j^{\alpha}$. The better the fit of the model, the closer this value should be

Figure 2.4: Value at Risk estimated by the block dependent static factor “skew t-t” copula



Source: Elaborated by the author.

Figure 2.5: Value at Risk estimated by the block dependent dynamic factor “skew t-t” copula



Source: Elaborated by the author.

to zero. The excess levels of tested ES for each model are listed in Table 2.8 whit their respective p-values from the McNeil and Frey test in parentheses.

With regard to the ES we can not detect any clear difference among the performances of different copula models. It is noted only that, according to McNeil and Frey test, Gumbel and Frank copulas perform worse, with ES excess values significantly different from zero. All the other copulas perform well and we can not differentiate them according to their ES estimates: in order to do that, a larger sample would be needed for more accurate estimates.

Table 2.8: ES Excess for each model at level α

MODEL	LEVEL α			
	0.010	0.025	0.050	0.100
Normal	-0.0006 (0.516)	-0.0043 (0.087)	0.0007 (0.687)	0.0006 (0.699)
Student's t	-0.0023 (0.391)	-0.0024 (0.256)	0.0002 (0.567)	0.0005 (0.690)
Gumbel	-0.0045 (0.000)	-0.0045 (0.000)	-0.0043 (0.000)	-0.0040 (0.000)
Clayton	-0.0015 (0.364)	-0.0007 (0.360)	-0.0005 (0.349)	-0.0007 (0.223)
Frank	-0.0059 (0.000)	-0.0063 (0.000)	-0.0058 (0.000)	-0.0055 (0.000)
Factor (equidep.)	-0.0058 (0.140)	0.0011 (0.647)	0.0014 (0.788)	0.0010 (0.824)
Factor (block dep.)	-0.0047 (0.182)	0.0002 (0.553)	0.0012 (0.757)	0.0012 (0.860)
Dyn. factor (equidep.)	-0.0090 (0.302)	0.0013 (0.687)	0.0026 (0.874)	0.0023 (0.964)
Dyn. factor (block dep.)	-0.0040 (0.254)	-0.0007 (0.451)	0.0014 (0.779)	0.0010 (0.828)
Dyn. factor (heter. dep.)	-0.0048 (0.173)	-0.0011 (0.417)	0.0009 (0.721)	0.0009 (0.800)

Source: Elaborated by the author.

2.4 Concluding Remarks

In this paper, we try and estimate risk measures to a portfolio comprised by BOVESPA assets, evaluating the performance of different copula models. We are particularly interested in high dimensional problems, which bring along the very well known "curse of dimensionality". A very recent area of copula research, the factor copula modeling, is explored and its performance is compared to traditional multivariate approaches in terms of accuracy for evaluating risk measures, in particular VaR and ES.

The analysis performed here indicates that the Archimedean copulas are worse than the Normal and Student's t copulas with respect to the estimation of the Value at Risk for a portfolio of 44 assets of BOVESPA. These, in turn, have underperformed the factor copula models. Among the factor copula, however, there seems to be no appreciable difference in the results obtained from the different tested dependence structures or between their static and dynamic versions. Regarding the estimation of this portfolio Expected Shortfall, most of the tested models show good results, but none of them appears to be better than the other ones.

Thus, the factor copula approach proved to be useful in estimating more accurately the portfolio VaR, and the model that shows better results with fewer parameters to be estimated is the equidependent static factor copula with skewed Student's t and Student's t-distributions to the common factor and the idiosyncratic error, respectively. The biggest improvement of this model compared to normal multivariate model, which is traditionally used, is the inclusion of features such as heavy tails and asymmetry, which proved to be decisive on the estimation of VaR. This model also preformed well in the estimation of the portfolio Expected Shortfall. Although computationally expensive, this model is still feasible to be estimated and used to calculate risk measures.

Possible extensions tackling these issues are diversification of these distributions, the increase of the number of factors and the estimation of these measures of risk for different sets of data.

References

AAS, K. et al. Pair-Copula Constructions of Multiple Dependence. **Insurance**: Mathematics and Economics, Amsterdam, v. 44, n. 2, p. 182-198, 2009.

ALMEIDA, C.; CZADO, C. Efficient Bayesian Inference for Stochastic Time-varying Copula Models. **Computational Statistics & Data Analysis**, Amsterdam, v. 56, n. 6, p. 1511-1527, 2012.

ANDERSEN, L.; SIDENIUS, J. Extensions to the Gaussian Copula: Random Recovery and Random Factor Loadings. **Journal of Credit Risk**, London, v. 1, p. 29-70, 2004.

BEDFORD, T.J.; COOKE, R.M. Probability Density Decomposition for Conditionally Dependent Random Variables Modeled by Vines. **Annals of Mathematics and Artificial Intelligence**, Dordrecht, v. 32, n. 1, p. 245-268, 2001.

BEDFORD, T.J.; COOKE, R.M. Vines - A New Graphical Model for Dependent Random Variables. **The Annals of Statistics**, Beachwood, v. 30, n. 4, p. 1031-1068, 2002.

BERG, D.; AAS, K. Models for Construction of Higher-Dimensional Dependence: A Comparison Study. **European Journal of Finance**, Abinfdon, v. 15, p. 639-659, 2009.

BOLLERSLEV, T. A. Conditionally Heteroskedasticity Time Series Model for Speculative Prices and Rates of Returns. **The Review of Economics and Statistics**, Cambridge, v. 69, n. 3, p. 542-547, 1987.

BOLSA DE VALORES DE SÃO PAULO. **Série Histórica de Cotações**. São Paulo, 2008-2014. Available at: http://www.bmfbovespa.com.br/shared/iframe.aspx?idioma=pt-brurl=http://www.bmfbovespa.com.br/pt-br/cotacoes-historicas/FormSeriesHistoricas.asp. Accessed 28 may 2015.

BRADLEY, B.O; TAQQU, M.S. Financial Risk and Heavy Tails. In: RACHEV, S.T. (Ed.). **Handbook of Heavy Tailed Distributions in Finance**. Amsterdam: Elsevier/ North-Holland, 2003. p. 35-103.

BRECHMANN, E.C.; CZADO, C.; AAS, K. Truncated regular vines in high dimensions with applications to financial data. **Canadian Journal of Statistics**, Ottawa, v. 40, n. 1, p. 68-85, 2012.

CHEN, H.; MACMINN, R.; SUN, T. **Multi-Population Mortality Models**: A Factor Copula Approach. Available at: http://www.fma.org/Nashville/Papers/MultiPopulationMortalityModelAFactorCopulaApproachChenMacMinnSun.pdf. Accessed 17 may 2014.

CHERUBINI, U.; LUCIANO, E.; VECCHIATO, W. **Copula Methods in Finance**. Chichester: John Willey & Sons, 2004.

CREAL, D.D.; KOOPMAN, S.J.; LUCAS, A. A Dynamic Multivariate Heavy-tailed Model for Time-varying Volatilities and Correlations. **Journal of Business and Economic Statistics**, Amsterdam, v. 29, n. 4, p. 552-563, 2011.

CREAL, D.D.; KOOPMAN, S.J.; LUCAS, A. Generalized Autoregressive Score Models with Applications. **Journal of Applied Econometrics**, Hoboken, v. 28, n. 5, p. 777-795, 2013.

CREAL, D.D.; TSAY, R.S. **High Dimensional Dynamic Stochastic Copula Models**. 2014. Available at: http://faculty.chicagobooth.edu/drew.creal/research/papers/crealTsay2014.pdf. Accessed 17 may 2014.

CHRISTOFFERSEN, P.F. Evaluating Interval Forecasts. **International Economic Review**, Philadelphia, v. 39, n. 4, p. 841-862, 1998.

CZADO, C. Pair-Copula Constructions of Multivariate Copulas. In: JAWORSKI, P. et al. (Ed.). **Copula Theory and Its Applications**. Berlin: Springer-Verlag, 2010. p. 93-109.

DISSMANN, J. et al. Selecting and Estimating Regular Vine Copulae and Application to Financial Returns. **Computational Statistics & Data Analysis**, Amsterdam, v. 59, n. 1, p. 52-69, 2013.

FISCHER, M. et al. An Empirical Analysis of Multivariate Copula Models. **Quantitative Finance**, Abingdon, v. 9, n. 7, p. 839-854, 2009.

HANSEN, B.E. Autoregressive Conditional Density Estimation. **International Economic Review**, Philadelphia, v. 35, n. 3, p. 705-730, 1994.

HEINEN, A.; VALDESOGO, A. **Asymmetric CAPM Dependence for Large Dimensions**: The Canonical Vine Autoregressive Model. 2009. Available at: http://www.ub.edu/economiaempresa/jobmarket/papers/Alfonso/ApplicantFile View.pdf. Accessed 17 may 2014.

HULL, J.; WHITE, A. Valuation of a CDO and an n^{th} to Default CDS without Monte Carlo Simulation. **Journal of Derivatives**, New York, v. 12, p. 8-23, 2004.

JOE, H. Families of m-Variate Distributions with Given Margins and $m(m-1)/2$ Bivariate Dependence Parameters. In: RÜSCHENDORF, L.; SCHWEIZER, B.; TAYLOR, M.D. (Ed.). **Distributions with Fixed Marginals and Related Topics**. Hayward: Institute of Mathematical Statistics, 1996. p. 120-141.

JOE, H. **Multivariate Models and Dependence Concepts**. London: Chapman & Hall, 1997.

JOE, H.; XU, J.J. **The Estimation Method of Inference Functions for Margins for Multivariate Models**. Vancouver: University of British Columbia, Department of Statistics, 1996. (Technical Report, 166)

KRUPSKII, P.; JOE, H. Factor Copula Models for Multivariate Data. **Journal of Multivariate Analysis**, Amsterdam, v. 120, p. 85-101, 2013.

KUROWICKA, D.; COOKE, R.M. **Uncertainty Analysis with High Dimensional Dependence Modelling**. Chichester: John Wiley & Sons, 2006.

KUROWICKA, D.; JOE, H. **Dependence Modeling**: Vine Copula Handbook. Singapore: World Scientific Publishing, 2011.

MENDES, B.V.M. Computing Conditional VaR Using Time-varying Copulas. **Revista Brasileira de Finanças**, Rio de Janeiro, v. 3, n. 2, p. 251-265, 2005.

MCNEIL, A.J.; FREY, R.; EMBRECHTS, P. **Quantitative Risk Management**. New Jersey: Princeton University Press, 2005.

MCNEIL, A.J.; FREY, R. Estimation of Tail-related Risk Measures for Heteroscedastic Financial Time Series: an extreme value approach. **Journal of Empirical Finance**, Amsterdam, v. 7, n. 3-4, p. 271-300, 2000.

NELDER, J.A.; MEAD, A. A Simplex Method for Function Minimization. **Computer Journal**, Oxford, v. 7, p. 308-313, 1965.

NELSEN, R. **An Introduction to Copulas**. New York: Springer-Verlag, 2006.

OH, D.H.; PATTON, A.J. **Modelling Dependence in High Dimensions with Factor Copulas**. Durham: Duke University, 2012. (Working Paper).

OH, D.H.; PATTON, A.J. **Time-Varying Systemic Risk**: Evidence from a Dynamic Copula Model of CDS Spreads. Durham: Economic Research Initiatives at Duke (ERID), 2013. (Working Paper, n. 167).

PATTON, A.J. Modeling Asymmetric Exchange Rate Dependence. **International Economic Review**, Philadelphia, v. 2, n. 47, p. 527-556, 2006.

ROCKFELLAR, R.T.; URYASEV, S. Optimization of Conditional Value-at-Risk. **Journal of Risk**, London, v. 2, n. 3, p. 21-41, 2000.

SCHWEIZER, B.; SKLAR, A. **Probabilistic Metric Spaces**. Mineola, New York: Dover, 2005.

SILVA FILHO, O.C.; ZIEGELMANN, F.A. Assessing Some Stylized Facts About Financial Market Indexes: a Markov copula approach. **Journal of Economic Studies**, Bingley, v. 41, n. 2, p. 253-271, 2014.

SILVA FILHO, O.C.; ZIEGELMANN, F.A.; DUEKER, M. Modeling Dependence Dynamics through Copulas with Regime Switching. **Insurance**: Mathematics & Economics, Amsterdam, v. 50, p. 346-356, 2012.

SILVA FILHO, O.C.; ZIEGELMANN, F.A.; DUEKER, M. Assessing Dependence Between Financial Market Indexes using Conditional Time-varying Copulas: Applications to Value at Risk (VaR). **Quantitative Finance**, Abingdon, v. 14, n. 12, p. 2155-2170, 2014.

TOFOLI, P.; ZIEGELMANN, F.A.; SILVA FILHO, O.C. Dynamic D-vine Copula Model with Applications to Value-at-Risk (VAR). In: BRAZILIAN CONFERENCE ON STATISTICAL MODELING IN INSURANCE AND FINANCE, 6th., 2013, Maresias. **Proceedings...** Maresias: ABE, 2013. Não paginado.

VAN DER VOORT, M. Factor Copulas: **Totally External Defaults**. Rotterdam: Erasmus University Rotterdam, 2005. (Working paper)

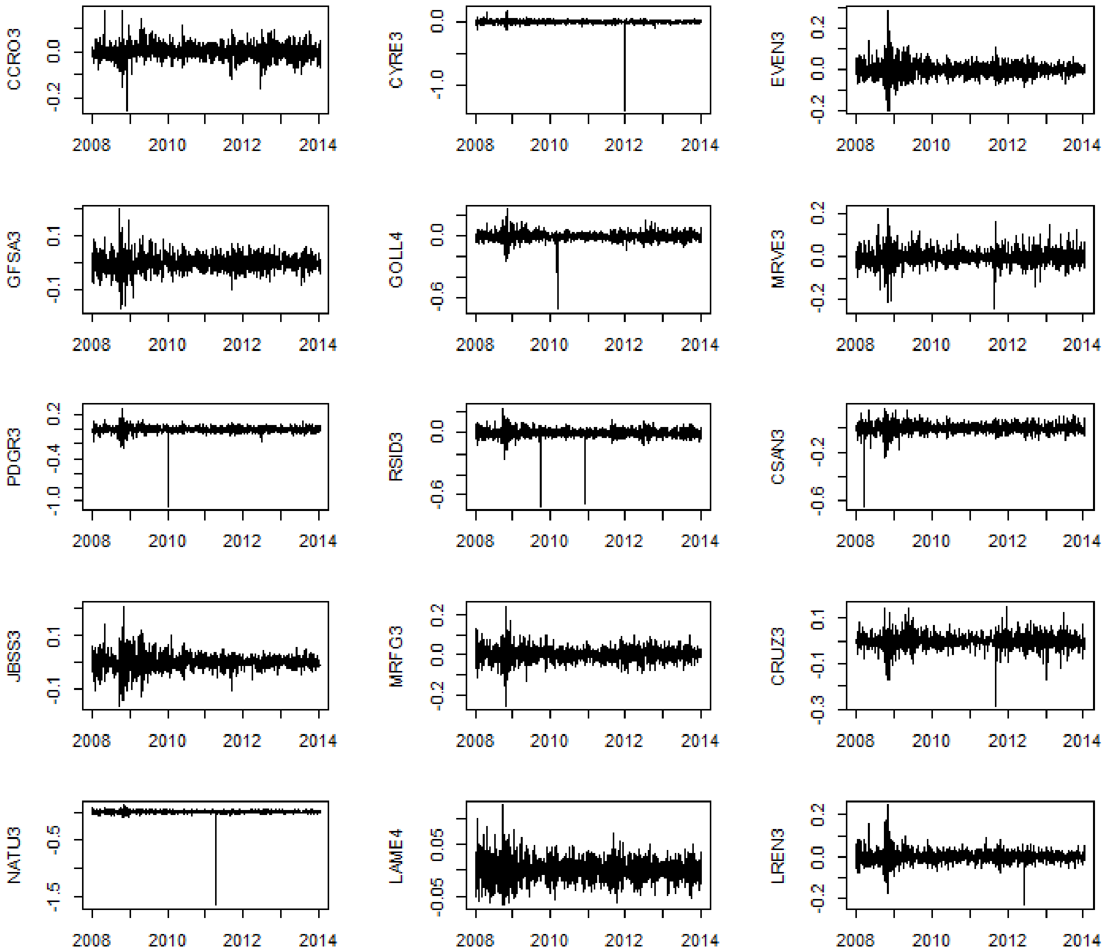
Appendix

Table 2.9: BOVESPA assets and their blocks

CODE	ASSET	BLOCK	BLOCK CODE
BISA3	BROOKFIELD	Construction and Transportation	1
CCRO3	CCR SA	Construction and Transportation	1
CYRE3	CYRELA REALT	Construction and Transportation	1
EVEN3	EVEN	Construction and Transportation	1
GFSA3	GAFISA	Construction and Transportation	1
GOLL4	GOL	Construction and Transportation	1
MRVE3	MRV	Construction and Transportation	1
PDGR3	PDG REALT	Construction and Transportation	1
RSID3	ROSSI RESID	Construction and Transportation	1
CSAN3	COSAN	Consumption	2
JBSS3	JBS	Consumption	2
MRFG3	MARFRIG	Consumption	2
CRUZ3	SOUZA CRUZ	Consumption	2
NATU3	NATURA	Consumption	2
LAME4	LOJAS AMERIC	Consumption	2
LREN3	LOJAS RENNER	Consumption	2
BBAS3	BRASIL	Finance	3
BBDC3	BRADESCO	Finance	3
BBDC4	BRADESCO	Finance	3
BRAP4	BRADESPAR	Finance	3
BRML3	BR MALLS PAR	Finance	3
ITSA4	ITAUSA	Finance	3
BRKM5	BRASKEM	Basic Materials	4
CSNA3	SID NACIONAL	Basic Materials	4
GGBR4	GERDAU	Basic Materials	4
GOAU4	GERDAU MET	Basic Materials	4
MMXM3	MMX MINER	Basic Materials	4
SUZB5	SUZANO PAPEL	Basic Materials	4
USIM5	USIMINAS	Basic Materials	4
VALE3	VALE	Basic Materials	4
VALE5	VALE	Basic Materials	4
EMBR3	EMBRAER	Others	5
RENT3	LOCALIZA	Others	5
PETR3	PETROBRAS	Others	5
PETR4	PETROBRAS	Others	5
CESP6	CESP	Public Utility	6
CMIG4	CEMIG	Public Utility	6
CPFE3	CPFL ENERGIA	Public Utility	6
CPLE6	COPEL	Public Utility	6
ELET3	ELETROBRAS	Public Utility	6
ELET6	ELETROBRAS	Public Utility	6
ENBR3	ENERGIAS BR	Public Utility	6
LIGT3	LIGHT S/A	Public Utility	6
SBSP3	SABESP	Public Utility	6
TBLE3	TRACTEBEL	Public Utility	6

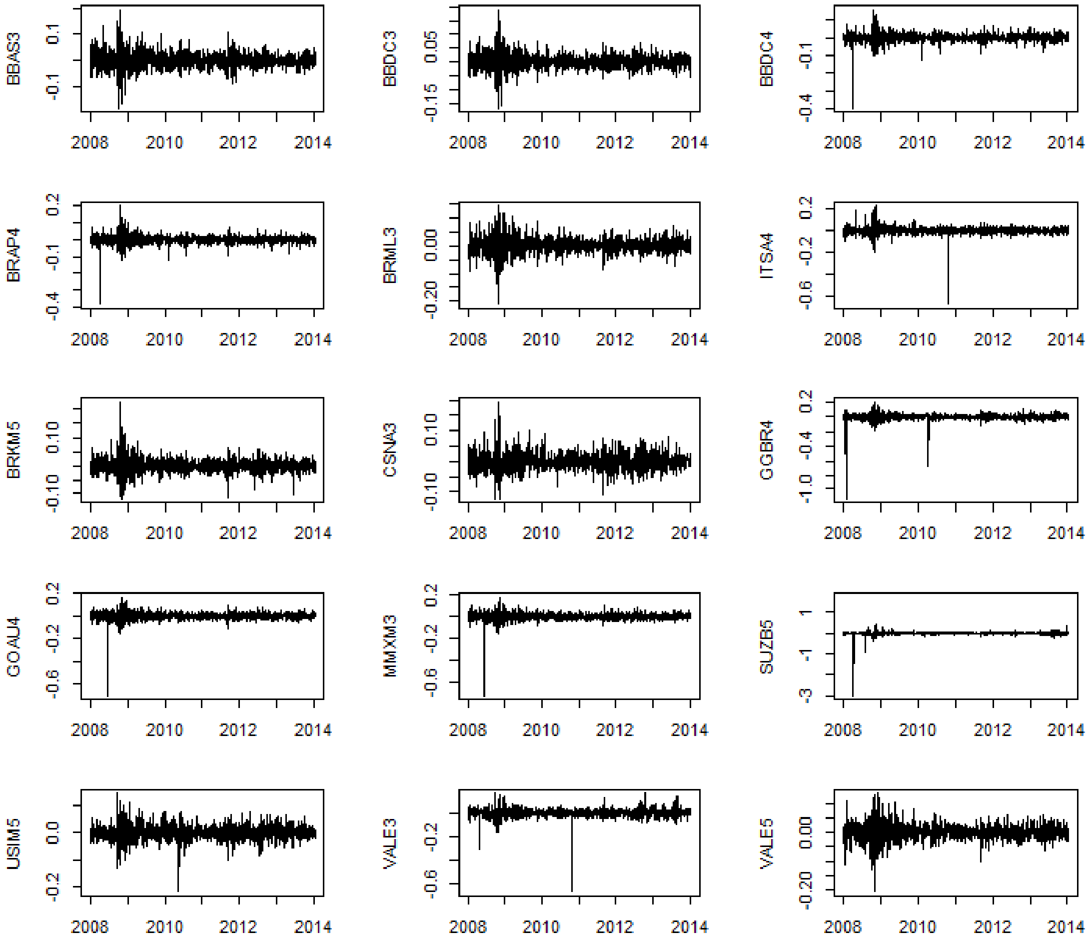
Source: Elaborated by the author.

Figure 2.6: Stock returns of BOVESPA assets (part 1)



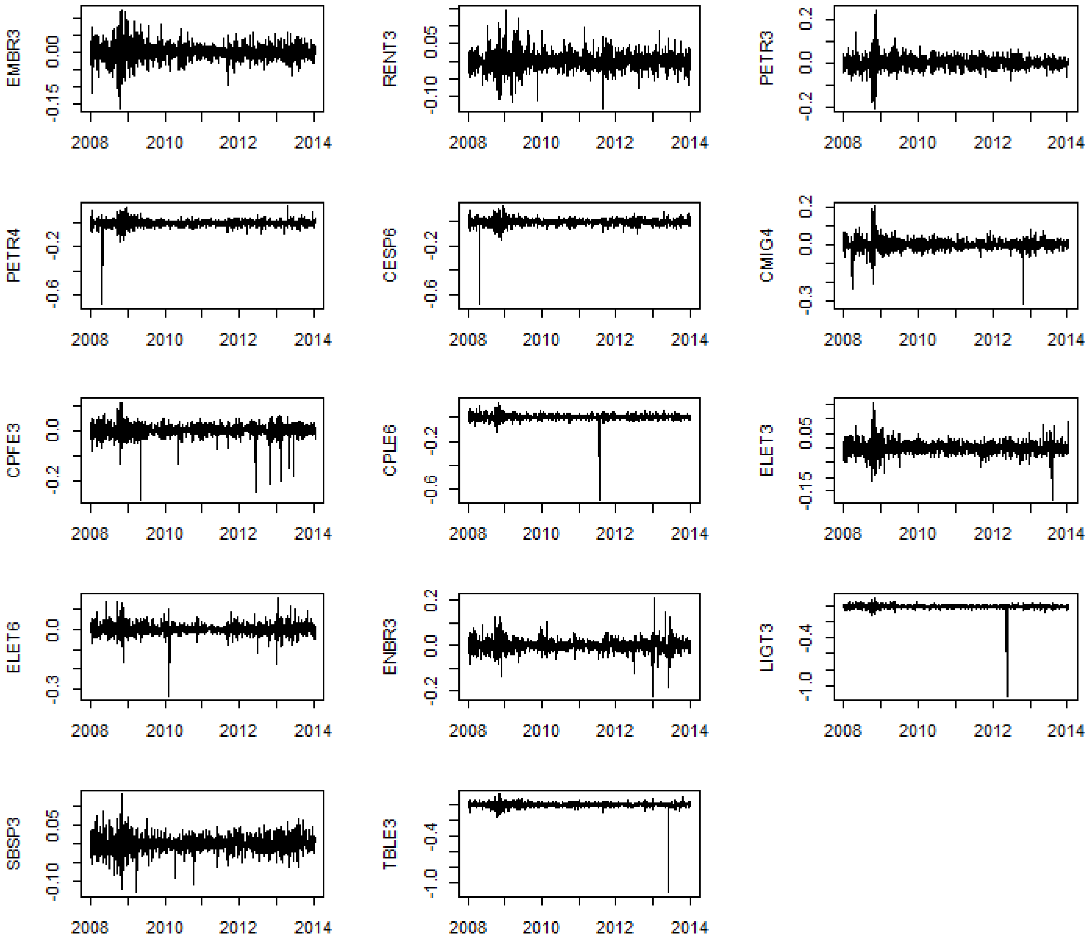
Source: Elaborated by the author based on data from BOVESPA (2008-2014).

Figure 2.7: Stock returns of BOVESPA assets (part 2)



Source: Elaborated by the author based on data from BOVESPA (2008-2014).

Figure 2.8: Stock returns of BOVESPA assets (part 3)



Source: Elaborated by the author based on data from BOVESPA (2008-2014).

3 Considerações Finais

Este trabalho teve como objetivo estimar medidas de risco de uma carteira composta por ativos da BOVESPA, avaliando o desempenho de diferentes modelos de cópulas. Estamos particularmente interessados em problemas de alta dimensão, que trazem junto a muito conhecida "maldição da dimensionalidade". Uma área muito recente da pesquisa em cópulas, a modelagem de cópulas fatoriais, é explorada, e seu desempenho é comparado com abordagens multivariadas tradicionais, em termos de precisão para avaliar medidas de risco, em especial VaR e ES.

Na Seção 2, apresentamos um artigo em que utilizamos um portfólio composto por 44 ações da BOVESPA com esta finalidade. A análise indica que as cópulas Arquimedianas são inferiores às cópulas Normal e t-Student no que diz respeito à estimação do Valor em Risco para este portfólio. Estas, por sua vez, apresentam desempenho inferior às cópulas fatoriais. Dentre as cópulas fatoriais, entretanto, não parece haver diferença sensível nos resultados obtidos a partir das diferentes estruturas de dependência testadas, nem entre suas versões estática e dinâmica. No que diz respeito à estimação do Expected Shortfall deste portfólio, a maioria dos modelos testados apresentou bons resultados, sem que nenhum tenha se destacado dos outros positivamente.

Assim, a metodologia de cópulas fatoriais se mostrou útil no cálculo do Valor em Risco do portfólio, sendo que o modelo que apresentou melhores resultados com menos parâmetros a serem estimados foi o de cópulas fatoriais estáticas equidependentes com distribuições t-Student assimétrica e t-Student simétrica para o fator comum e o erro idiossincrático, respectivamente. A melhor contribuição deste modelo em relação ao modelo normal multivariado tradicionalmente utilizado é a inclusão de características como caudas pesadas e assimetria, que se mostraram decisivas na estimação do VaR. Este modelo também apresentou bons resultados no que diz respeito à estimação do Expected Shortfall. Apesar de computacionalmente custoso, este modelo ainda é factível de ser estimado e utilizado para o cálculo de medidas de risco.

Possíveis extensões que abordam estas questões são a diversificação dessas distribuições, o aumento do número de fatores e a estimativa destas medidas de risco para conjuntos diferentes de dados.

Referências

AAS, K. et al. Pair-Copula Constructions of Multiple Dependence. **Insurance**: Mathematics and Economics, Amsterdam, v. 44, n. 2, p. 182-198, 2009.

ALMEIDA, C.; CZADO, C. Efficient Bayesian Inference for Stochastic Time-varying Copula Models. **Computational Statistics & Data Analysis**, Amsterdam, v. 56, n. 6, p. 1511-1527, 2012.

ANDERSEN, L.; SIDENIUS, J. Extensions to the Gaussian Copula: Random Recovery and Random Factor Loadings. **Journal of Credit Risk**, London, v. 1, p. 29-70, 2004.

BEDFORD, T.J.; COOKE, R.M. Probability Density Decomposition for Conditionally Dependent Random Variables Modeled by Vines. **Annals of Mathematics and Artificial Intelligence**, Dordrecht, v. 32, n. 1, p. 245-268, 2001.

BEDFORD, T.J.; COOKE, R.M. Vines - A New Graphical Model for Dependent Random Variables. **The Annals of Statistics**, Beachwood, v. 30, n. 4, p. 1031-1068, 2002.

BERG, D.; AAS, K. Models for Construction of Higher-Dimensional Dependence: A Comparison Study. **European Journal of Finance**, Abingdon, v. 15, p. 639-659, 2009.

BRADLEY, B.O; TAQQU, M.S. Financial Risk and Heavy Tails. In: RACHEV, S.T. (Ed.). **Handbook of Heavy Tailed Distributions in Finance**. Amsterdam: Elsevier/ North-Holland, 2003. p. 35-103.

BRECHMANN, E.C.; CZADO, C.; AAS, K. Truncated Regular Vines in High Dimensions with Applications to Financial Data. **Canadian Journal of Statistics**, Ottawa, v. 40, n. 1, p. 68-85, 2012.

CHEN, H.; MACMINN, R.; SUN, T. **Multi-Population Mortality Models**: A Factor Copula Approach. Disponível em: http://www.fma.org/Nashville/Papers /MultiPopulationMortalityModelAFactorCopulaApproachChenMacMinnSun.pdf. Acesso 17 maio 2014.

CHERUBINI, U.; LUCIANO, E.; VECCHIATO, W. **Copula Methods in Finance**. Chichester: John Willey & Sons, 2004.

CREAL, D.D.; KOOPMAN, S.J.; LUCAS, A. Generalized Autoregressive Score Models with Applications. **Journal of Applied Econometrics**, Hoboken, v. 28, n. 5, p. 777-795, 2013.

CREAL, D.D.; TSAY, R.S. **High Dimensional Dynamic Stochastic Copula Models**. 2014. Disponível em: http://faculty.chicagobooth.edu/drew.creal/ research/papers/crealTsay2014.pdf. Acesso em 17 maio 2014.

CZADO, C. Pair-Copula Constructions of Multivariate Copulas. In: JAWORSKI, P. et al. (Ed.). **Copula Theory and Its Applications**. Berlin: Springer-Verlag, 2010. p. 93-109.

DISSMANN, J. et al. Selecting and Estimating Regular Vine Copulae and Application to Financial Returns. **Computational Statistics & Data Analysis**, Amsterdam, v. 59, n. 1, p. 52-69, 2013.

FISCHER, M. et al. An Empirical Analysis of Multivariate Copula Models. **Quantitative Finance**, Abingdon, v. 9, n. 7, p. 839-854, 2009.

HEINEN, A.; VALDESOGO, A. **Asymmetric CAPM Dependence for Large Dimensions**: The Canonical Vine Autoregressive Model. 2009. Disponível em: http://www.ub.edu/economiaempresa/jobmarket/papers/Alfonso/ApplicantFileView.pdf. Acesso 17 maio 2014.

HULL, J.; WHITE, A. Valuation of a CDO and an n^{th} to Default CDS without Monte Carlo Simulation. **Journal of Derivatives**, New York, v. 12, p. 8-23, 2004.

JOE, H. Families of m-Variate Distributions with Given Margins and $m(m-1)/2$ Bivariate Dependence Parameters. In: RÜSCHENDORF, L.; SCHWEIZER, B.; TAYLOR, M.D. (Ed.). **Distributions with Fixed Marginals and Related Topics**. Hayward: Institute of Mathematical Statistics, 1996. p. 120-141.

JOE, H. **Multivariate Models and Dependence Concepts**. London: Chapman & Hall, 1997.

KRUPSKII, P.; JOE, H. Factor Copula Models for Multivariate Data. **Journal of Multivariate Analysis**, Amsterdam, v. 120, p. 85-101, 2013.

KUROWICKA, D.; COOKE, R.M. **Uncertainty Analysis with High Dimensional Dependence Modelling**. Chichester: John Wiley & Sons, 2006.

KUROWICKA, D.; JOE, H. **Dependence Modeling**: Vine Copula Handbook. Singapore: World Scientific Publishing, 2011.

MENDES, B.V.M. Computing Conditional VaR Using Time-varying Copulas. **Revista Brasileira de Finanças**, Rio de Janeiro, v. 3, n. 2, p. 251-265, 2005.

MCNEIL, A.J.; FREY, R.; EMBRECHTS, P. **Quantitative Risk Management**. New Jersey: Princeton University Press, 2005.

NELDER, J.A.; MEAD, A. A Simplex Method for Function Minimization. **Computer Journal**, Oxford, v. 7, p. 308-313, 1965.

NELSEN, R. **An Introduction to Copulas**. New York: Springer-Verlag, 2006.

OH, D.H.; PATTON, A.J. **Modelling Dependence in High Dimensions with Factor Copulas**. Durham: Duke University, 2012. (Working Paper).

OH, D.H.; PATTON, A.J. **Time-Varying Systemic Risk**: Evidence from a Dynamic Copula Model of CDS Spreads. Durham: Economic Research Initiatives at Duke (ERID), 2013. (Working Paper, n. 167).

PATTON, A.J. Modeling Asymmetric Exchange Rate Dependence. **International Economic Review**, Philadelphia, v. 2, n. 47, p. 527-556, 2006.

SILVA FILHO, O.C.; ZIEGELMANN, F.A. Assessing Some Stylized Facts About Financial Market Indexes: a Markov copula approach. **Journal of Economic Studies**, Bingley, v. 41, n. 2, p. 253-271, 2014.

SILVA FILHO, O.C.; ZIEGELMANN, F.A.; DUEKER, M. Modeling Dependence Dynamics through Copulas with Regime Switching. **Insurance**: Mathematics &

Economics, Amsterdam, v. 50, p. 346-356, 2012.

SILVA FILHO, O.C.; ZIEGELMANN, F.A.; DUEKER, M. Assessing Dependence Between Financial Market Indexes using Conditional Time-varying Copulas: Applications to Value at Risk (VaR). **Quantitative Finance**, Abingdon, v. 14, n. 12, p. 2155-2170, 2014.

TOFOLI, P.; ZIEGELMANN, F.A.; SILVA FILHO, O.C. Dynamic D-vine Copula Model with Applications to Value-at-Risk (VAR). In: BRAZILIAN CONFERENCE ON STATISTICAL MODELING IN INSURANCE AND FINANCE, 6th., 2013, Maresias. **Proceedings...** Maresias: ABE, 2013. Não paginado.

VAN DER VOORT, M. Factor Copulas: **Totally External Defaults**. Rotterdam: Erasmus University Rotterdam, 2005. (Working paper)