

On some nonlinear dependence structure in portfolio design

Nguyen Phuc Son, Pham Hoang Uyen, Nguyen Dinh Thien

Abstract—Constructing portfolios with high returns and low risks is always in great demand. Markowitz (1952) utilized correlation coefficients between pairs of stocks to build portfolios satisfying different levels of risk tolerance. The correlation coefficient describes the linear dependence structure between two stocks, but cannot capture a lot of nonlinear independence structures. Therefore, sometimes, portfolio performances are not up to investors' expectations. In this paper, based on the theory of copula by Sklar (see [19]), we investigate several new methods to detect nonlinear dependence structures. These new methods allow us to estimate the density of the portfolio which leads to calculations of some popular risk measurements like the value at risk (VaR) of investment portfolios. As for applications, making use of the listed stocks on the Ho Chi Minh city Stock Exchange (HoSE), some Markowitz optimal portfolios are constructed together with their risk measurements. Apparently, with nonlinear dependence structures, the risk evaluations of some pairs of stocks have noticeable twists. This, in turn, may lead to changes of decisions from investors.

Keywords—Portfolio design, data science, dependence structure, copula, risk, stocks, return, measure.

1 INTRODUCTION

SINCE the birth of stock exchanges, investors have been constantly seeking out optimal portfolios. One of the key characteristics of a good

portfolio is a reasonably low risk. All sort of techniques, from conventional wisdoms like "not putting all eggs into one baskets" to highly computational tools like neural network, have been tried to address this issue. Poor descriptions of dependence structures between pairs of stocks or among multiple stocks accounts for unreliable risk estimations of a majority of methods, hence, leads to unsatisfactory portfolios. Markowitz (1952) was probably the first person who incorporated dependences into portfolio designs. However, his work deals with linear dependences only while, in reality, most relations between stocks are nonlinear. In this paper, based on the concept of copula by Sklar (1959), dependence structures of certain pairs of stocks from the Ho Chi Minh Stock Exchange (HoSE) from July 2000 to July 2017 are described, and hence, empirical distributions of portfolios are obtained. That enables us to compute values at risk of various pairs of stocks with nonlinear dependences. For comparison purpose, the traditional values at risk with linear correlations are also included. The VaRs of portfolios from the stocks listed on HoSE show significant differences between the linear-dependence version and the nonlinear-dependence one which indicates the existence of crucial nonlinear structures.

2 LITERATURE REVIEW

In 1994, the concept of Value at Risk, VaR, was introduced with drums and cymbals to answer the question at what value of the investment the risk is equal to a desired percentage point (Szegö, 2002). The aggregate VaR is computed using linear correlation, which means relying on the assumption of multivariate normality of returns (Cherubini and Luciano, 2001). Recall that VaR is the α -quantile of the distribution

$$\text{VaR}_\alpha(S) = F_S^{-1}(\alpha) = \inf \{s : F_S(s) \geq \alpha\}$$

where F_S is the distribution of S . Typical values of the level α are 0.9, 0.95, or 0.99.

Since then, VaR has become a standard measure

Received: 13-10-2017, Accept: 11-12-2017, Published: 15-7-2018.

Author Nguyen Phuc Son, Ho Chi Minh city Institute for Development Studies (e-mail: sonnp@uel.edu.vn).

Author Pham Hoang Uyen, University of Economics and Law, VNUHCM, Viet Nam (e-mail: uyenph@uel.edu.vn).

Author Nguyen Dinh Thien, University of Economics and Law, VNUHCM, Viet Nam (e-mail: thiennd@uel.edu.vn).

of risk in financial markets. Besides, it is being used increasingly by other financial and even nonfinancial firms (Berkowitz and O'Brien, 2002). Although the VaR concept is very simple, its calculation is not easy (Jorion, 2006).

To calculate VaR, Arzac and Bawa (1977), Mittnik and Paoletta (2000) made an assumption that the return of interest is normally distributed, however, this assumption was challenged in some later work. A lot of research has been done on alternative methods. For instance, Jorion (2006) suggested new methodologies to calculate a portfolio VaR are: (i) the variance-covariance approach, also called the parametric method, (ii) the Historical Simulation (Non-parametric method) and (iii) the Monte Carlo simulation, which is a Semiparametric method. In practice, on studying VaR for traders with both long and short positions, Giot and Laurent (2003) found that the returns should not be modeled by either the normal or student t distributions. That leads to the needs to measure portfolio risks where the return distributions are non-normal, see (Artzner et al., 1997, 1999; Favre and Galeano, 2002)

In another line of research, Christoffersen and Pelletier (2004) realized that existing backtesting methods have relatively low power in realistic small sample settings. They explored some new tools to backtest based on the duration of days between the violations of the VaR via Monte Carlo simulation.

On the other hand, Reboredo (2013), using gold as a hedge against inflation, succeeded in reducing portfolio risk via modeling the dependence structure between gold and the USD through copula functions. In the same vein, Markowitz (2014) confirmed that the normality assumption are totally inadequate when applied to distributions of financial return or to distributions of quadratic utility functions in modern portfolio theory. These results strongly support earlier work by Levy and Markowitz (1979); in short, the study had two principal objectives: (1) to see how good mean-variance approximations are for various utility functions and portfolio return distributions; and (2) to test an alternate way of estimating expected utility from a distribution's mean and variance.

Risk analysis, applying copula theory has also been dealt with in the financial literature. Bob (2013) estimated VaR for a portfolio combining copula functions. Siburg et al (2015) proposed to forecast the VaR of bivariate portfolios using

copula which are calibrated to estimates of the coefficient of lower tail dependence.

There were much research centered on the application of a concept in the study of multivariate distributions, the copula, to the investigation of dependent tail events. (Hien et al, 2017).

3 COPULA

Dependence structures and measurements of dependence structures are two mainstreams in correlational study. One of the most notable work is a novel measure of monotonicity dependence by Schweizer and Wolf (1981) with the use of the L^p - metric $d_{L^p}(C, P)$ where C is any copula and P is the product copula. Furthermore, Stoimenov constructed another measure $\omega(C)$ to capture mutual complete dependences (MCD) using a Sobolev metric $d_s(C, P)$. Lately, Hien et al (2017) has devised a new dependence measure which can detect independence, comonotonicity, countermonotonicity relations between random variables.

In another line of research, Hien et al (2015) invented a nonparametric dependence measure for two continuous random variables X and Y with a copula as follows:

$$\lambda(C) = \|C\|_s - 2\|C - M\|_s^2$$

where $\|C\|_s$ is the adjusted Sobolev norm for copula C . In detail, the norm is defined as follows:

$$\|C\|_s = \left(\iint_{[0,1]^2} \left[\frac{\partial C^2(u,v)}{\partial u} + \frac{\partial C^2(u,v)}{\partial v} \right] dudv \right)^{1/2}$$

The authors also provide two practical algorithms to compute $\lambda(C)$ in applications. The first one is used when we have a close-form formula for the copula C , and the second one is for empirical copula C estimated from data.

The recent literature provides strong and rich evidence of return correlations among stocks and between stocks and bonds (Kim et al., 2006). According to Righi et al. (2015), the bivariate copula framework offers more flexibility than the traditional methodologies like the correlation, VaR. In this work, we utilize some of the most popular families of copulas to model dependencies between pairs of stock returns listed on HoSE. Based on these calculations, we evaluate risks of Markowitz optimal portfolios. In the process, we

also explore some new copulas constructed by applying the Wang distortion functions to classical copulas. The values at risk based on linear correlations and the values at risk based on nonlinear correlations show clear differences. That proves the necessity of nonlinear dependence structure on designing portfolios.

4 METHODOLOGY

In order to compare the portfolio-risk evaluations between the classical method based on linear correlation and our new method based on copulas, we first select weights for stocks to form optimal portfolios using the classical work of Markowitz. Then, we proceed by calculating the classical VaR and estimating the empirical copulas to build the empirical density function for our portfolios, hence obtaining VaR for the nonlinear dependence. Below are the concrete steps:

Step 1: Calculate the log returns of stocks by the formula: $\ln\left(\frac{P_t}{P_{t-1}}\right)$

Step 2: Calculate linear (Pearson) correlation coefficient between each pair of stock returns by the formula:

$$s(X, Y) = \text{Cov}[X, Y] / (s_X^2 \cdot s_Y^2)^{1/2}$$

Where: $\text{Cov}[X, Y]$ is the sample covariance of 2 random variables X and Y.

and s_X and s_Y are the standard deviations of the corresponding stocks X and Y, respectively

Step 3: Select pairs of stocks with the biggest correlation coefficient to construct some portfolios and estimate the empirical density functions of those portfolios.

Step 4: Calculating Values at Risks of the portfolios in both the classical and the copula ways.

5 DATA

Daily adjusted closing prices of listed companies on HoSE from the first day of trading up to 8/8/2017 are used in our research. However, in order to stabilize the time series, we eliminated all "debut" stocks which were listed after 01/01/2016. Thus, there are 313 stocks considered in our study.

6 EMPIRICAL STUDIES

In this part, we illustrate the differences of density functions of portfolios when using linear

correlation (ρ) versus copula (\square). After calculating linear correlations for all possible pairs in our list of 313 stocks, we select two pairs with highest absolute correlations to form portfolios; in particular, one pair has highest positive and one has highest negative correlations. Then, we compare the empirical distribution with the classical linear correlation and the empirical distribution with copula. Finally, remarkable differences in risk measures via VaRs of the two methods are highlighted.

6.1 Types of Graphics

Table 1 below show the ρ and \square value of pairs of stocks which most positive and negative correlation $s(X, Y)$.

TABLE 1
 CORRELATION OF STOCKS

Pair of stocks	Linear correlation coefficient	Copula
SSI (Sai Gon Securities Incorporation) vs HCM (Ho Chi Minh City Securities Corporation)	0.80035	0.690189
ACL (Cuu Long Fish Joint Stock Company) vs STK (Century Synthetic Fiber Corporation)	-0.12196	0.028588

Source: Research result

SSI and HCM are in the same industry and both are big companies in Vietnam. That is the reason why their returns have the same systematic behaviors with respect to the stock market. So, a strong relationship between SSI and HCM comes at no surprise, and this relationship can easily be detected by various means. Meanwhile, ACL and STK belong to two different industries which partially explain why their periodicities and characteristics differ. These, in turn, lead to a much more complex relationship between them. In particular, their returns are negatively associated in the stock exchange.

In the second row, the linear correlation shows a weak negative association between ACL and STK while the copula shows a weak relationship in the opposite direction. This difference definitely affects the portfolio constructions greatly since it changes all the portfolio risk calculations, and hence alters the expected returns and so on. To illustrate the point, we form the portfolios for the two pairs using the Markowitz's method and perform the risk calculations in both the correlation and copula ways.

Table 2 below shows the descriptive statistic of

daily returns of 4 stocks.

TABLE 2
DESCRIPTIVE STATISTIC

	HCM	SSI	ACL	STK
Observations	2053	2053	344	344
Minimum	-0.0695	-0.0683	-0.0697	-0.1176
Median	0.0000	0.0000	0.0000	0.0000
Mean	0.0011	0.0005	0.0001	-0.0008
Quartile 3	0.0138	0.0114	0.0110	0.0158
Maximum	0.0700	0.0692	0.0700	0.1151
Standard deviation	0.0248	0.0219	0.0244	0.0359
Skewness	0.1784	0.2514	-0.0459	0.2307
Kurtosis	0.1590	0.5158	1.0322	1.0759

To better understand the distributions of the returns of four stocks, we plot the four density functions in Figure 1 below.

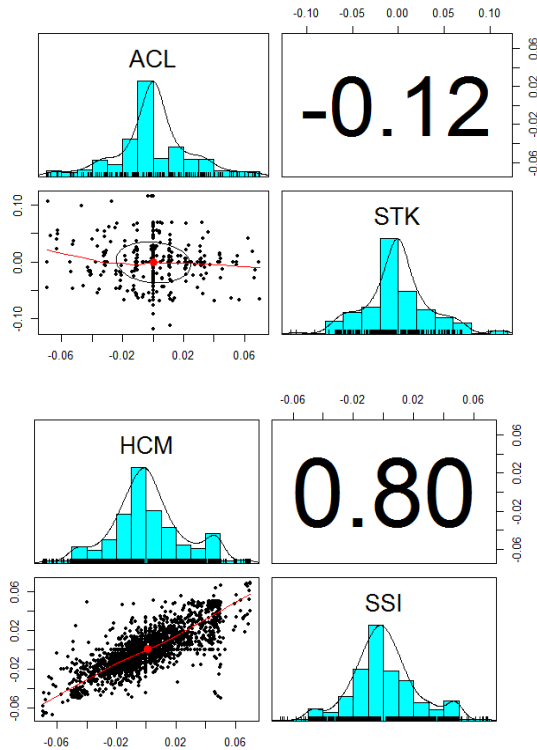


Figure 1. Density functions and correlations of returns
Source: Research result

Note that the empirical distributions of the four returns don't seem to follow normal distributions. Also observe that there is a strong positive correlation between 2 stocks in the financial sector, but there is a negative correlation between ACL (Food Industry) and STK (Textile Industry) in two different industries.

6.2 Values at risk

Below are the stock weights for two portfolios determined by the classic theory of Markowitz.

TABLE 3
STOCK WEIGHTS

	Portfolio 1		Portfolio 2	
Stock	HCM	SSI	ACL	STK
Weight	0.22	0.78	0.29	0.71

Note that in Portfolio 1, HCM and SSI are positively correlated and the weight for HCM is much lower than that of SSI because HCM suffers from much higher volatility. However, in Portfolio 2, although ACL has 47,1% lower volatility than STK, only 29% are allocated to ACL. That means volatility is of paramount importance when the relationship between stocks are positive, but is of little importance when the relationship is negative.

The following charts show the distributions of Portfolio 1 and Portfolio 2 in two scenarios: the left-hand side is the density plots for the portfolios using normal joint densities (with correlation r), and the right-hand side is the plots for the portfolios using empirical densities with copula.

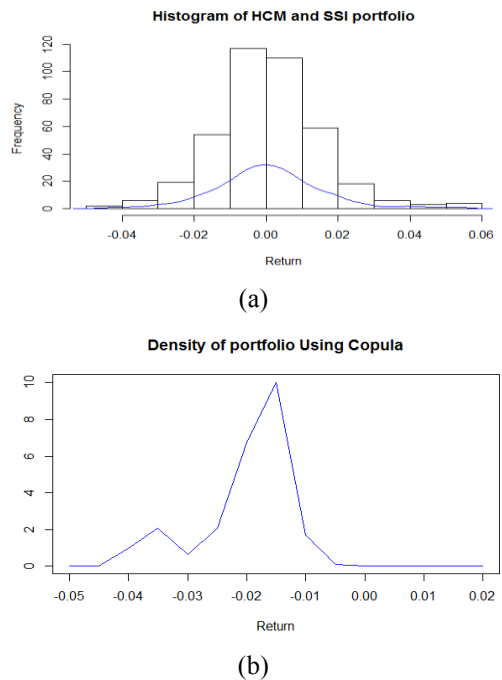


Figure 2. Density distribution of portfolio 1

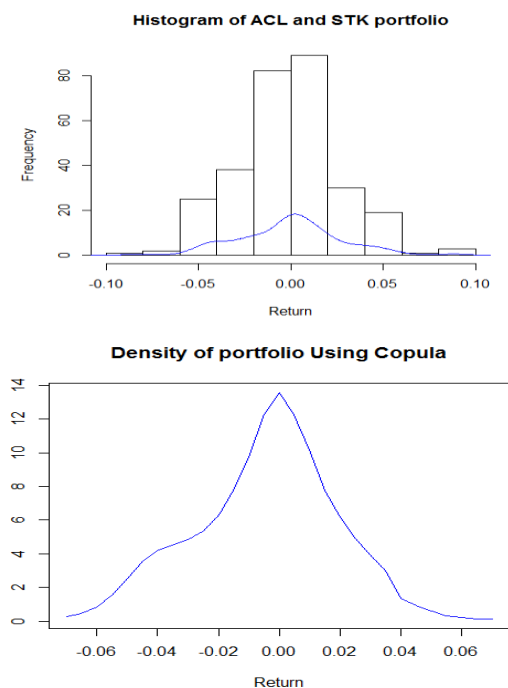


Figure 3: Density distribution of portfolio 2

There are noticeable dissimilarities between the density plots. That explains the differences in the VaR results when the classical normal distributions are used and when the copula-based distributions are used. The detailed calculations for the two portfolios are presented in the Table 4 below.

TABLE 4
 VAR FOR PORTFOLIO 1 AND PORTFOLIO 2 WITH 3 LEVELS OF CONFIDENCE

	Portfolio 1			Portfolio 2		
	90%	95%	99%	90%	95%	99%
Linear correlation (1)	0.0166	0.0219	0.0369	0.0394	0.0456	0.0552
Copula (2)	0.0227	0.0251	0.0304	0.0388	0.0432	0.0534
Different (1) vs (2)	-26.95%	-12.82%	21.39%	1.43%	5.65%	3.48%

Source: Research result

The table 4 presents the values at risk for Portfolio 1 and Portfolio 2 with 3 levels of confidence, namely 90%, 95% and 99%. Each column contains the VaRs computed from linear dependence (correlation) and nonlinear dependence (copula) together with a measurement for the relative differences between the two VaRs. To be precise, the last row of the table is calculated using the following formula.

$$\frac{VaR(\text{correlation}) - VaR(\text{copula})}{VaR(\text{copula})}$$

Recall that, in Portfolio 1, both components belong to the same industry and are positively associated. The volatilities of the two stocks have tremendous impacts on the allocation of weights in the portfolio, and this leads to a fairly large relative difference between the two versions of VaRs. Therefore, investors should take extra precautions when using VaR for risk management

in this case since there are dependences other than the linear one lurking around. For Portfolio 2, the situation is better, the relative differences are smaller since the two stocks are negatively associated so their volatilities tend to cancel each other out. To sum up, in reality, there are always stocks with positive associations in a portfolio. Thus, finding a suitable measures of dependency is fundamentally important for risk management.

7 CONCLUSION

Value at Risk plays a crucial role in financial risk management, especially in portfolio management. In practice, its computations usually relies on the normality assumption of the portfolio distributions with linear dependences between pairs of assets. This, in turn, is used to calculate the optimal weights for the stock components of portfolios. Nevertheless, a number of experts have

pointed out that the linear dependence structures are no longer adequate in modern financial market. Therefore, adoption of the new techniques is inevitable to avoid unwanted consequences later. In this research, we analyse the dependence structure between financial assets and additionally compute the VaR, which is of considerable importance in risk management

Since the pioneering work of Sklar (1959), although having been extensively applied in a large number of applications in business and finance, copulas still have huge potentials for making significant impacts in various problems, especially in risk management. Considerable advancements in computing powers allow copulas to become practical in areas where a deep understanding of dependence structures is crucial, but used to be intractable in the past due to high computational complexity. Our work here provides examples where classical approaches give very different results compared to those obtained via copula. In particular, if two stocks are strongly positively associated, the VaRs of the two methods differ as high as 26.95%

Our plan in the near future is to design portfolios of more than two stock components. It requires pulling in multivariate copulas to describe higher dimensional dependence structures. One of the challenges is how to implement these with real market data. Some important families like the Clayton canonical vine copulas (CVC) which capture lower tail dependence are feasible up to dimension 12. Another direction of research is to bring in other way to model dependence such as the probabilistic graphical model which has been extremely successful in computer sciences and engineering.

This research is funded by University of Economics and Law Ho Chi Minh City research with contract number CS/2017-08.

REFERENCES

- [1]. Artzner, P., Delbaen F., Eber, J. M. and Heath, D. (1997). Thinking Coherently. *Risk*, 10, November, pp 68–71.
- [2]. Arzac, E., and V. Bawa, (1977). Portfolio Choice and Equilibrium in Capital Markets with Safety-First Investors. *Journal of Financial Economics*, 4, 277–288.
- [3]. Artzner, P., Delbaen F., Eber, J. M. and Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9, pp 203–228.
- [4]. Berkowitz, J. and O'Brien, J. (2002). How Accurate Are Value-at-Risk Models at Commercial Banks? *The Journal Of Finance*, Vol. LVII, No. 3, June 2002, pp 1093 – 1111.
- [5]. Bob, N. K. (2013). Value at Risk Estimation. A GARCH-EVT-Copula Approach. *Mathematical Statistics Stockholm University*.
- [6]. Cherubini, U. and Luciano, E. (2001). Value-at-risk Trade-off and Capital Allocation with Copulas. *Economic Notes*, 30: 235–256
- [7]. Christoffersen, P. and Pelletier, D. (2004). Backtesting Value-at-Risk: A Duration-Based Approach. *Journal of Financial Econometrics*, Volume 2, Issue 1, 1 January 2004, pp 84–108
- [8]. Favre, L. and Galeano, J. (2002). Mean-modified value-at-risk optimization with hedge funds. *Journal of Alternative Investments*, 5, 21–25.
- [9]. Giot, P. and Laurent, S. (2003). Value-at-risk for long and short trading positions. *Journal of Applied Econometrics*, Volume 18, Issue 6, November/December 2003, pp 641–663
- [10]. Hien Duy Tran, Uyen Hoang Pham, Sel Ly and T. Vo-Duy. (2017). Extraction dependence structure of distorted copulas via a measure of dependence. *Annals of Operations Research*. DOI: 10.1007/s10479-017-2487-2
- [11]. Jorion, P. (2006). Value at Risk: The New Benchmark for Managing Financial Risk. *McGraw-Hill Education; 3rd Edition (November 9, 2006)*
- [12]. Kim, S. J., Moshirian, F., & Wu, E. (2006). Evolution of international stock and bond market integration: influence of the European Monetary Union. *Journal of Banking & Finance*, 30(5), 1507–1534.
- [13]. Laurent, F. and Galeano, J.A. (2002). Mean-Modified Value-at-Risk Optimization with Hedge Funds. *The Journal of Alternative Investments*, Fall 2002, Vol. 5, No. 2: pp. 21–25
- [14]. Levy, H. and Markowitz, H.M. (1979). Approximating expected utility by a function of mean and variance. *American Economic Review* 69 (3), 308–317.
- [15]. Markowitz, H.M. (1952). Portfolio Selection. *The Journal of Finance*, Vol. 7, No. 1. (Mar., 1952), pp. 77–91 .
- [16]. Markowitz, H.M. (2012). Mean–variance approximations to expected utility. *European Journal of Operational Research*, Volume 234, Issue 2, 16 April 2014, Pages 346–355 .
- [17]. Markowitz, H. (2014). Mean–variance approximations to expected utility. *European Journal of Operational Research*, 234:346–355
- [18]. Mittnik, S. and Paoletta, M. S. (2000). Conditional density and value-at-risk prediction of Asian currency exchange rates. *Journal of Forecast*, Vol 19, Issue 4, July 2000, pp 313–333
- [19]. Reboredo, J. C. (2013). Is gold a safe haven or a hedge for the US dollar? Implications for risk management. *Journal of Banking & Finance* 37 (2013), pp 2665–2676 .
- [20]. Righi, M. B., Schlender, S. G., & Ceretta, P. S. (2015). Pair copula constructions to determine the dependence structure of Treasury bond yields. *IIMB Management Review*, 27(4), 216–227 .
- [21]. Schweizer, B., and Wolff, E. F. (1981). On nonparametric measures of dependence for random variables. *The Annals of Statistics*, 9(4), pp 879–885.
- [22]. Siburg, K.F., Stoimenov, P., Weiß, G.N.F. (2015). Forecasting Portfolio-Value-at-Risk with Nonparametric Lower Tail Dependence Estimates. *Journal of Banking & Finance*, Volume 54, May 2015, Pages 129–140.
- [23]. Sklar, A., "Fonctions de répartition à n dimensions et leurs marges", Publ. Inst. Statist. Univ. Paris, 8, 1959, page 229–231

[24]. Stoimenov, P. A. (2008). A measure of mutual complete dependence. *Ph.D.Thesis, TU.Dortmund*.

[25]. Szego, G. (2002). Measures of risk. *Journal of Banking & Finance* 26, (2002), pp 1253–1272

[26]. Tran, H. D., Pham, U. H., Ly, S., and Vo-Duy, T. (2015). A New Measure of Monotone Dependence by Using Sobolev Norms for Copula. In *Integrated Uncertainty in Knowledge*

Modelling and Decision Making. Springer International Publishing, pp. 126-137.

[27]. Wang, R., Peng, L. & Yang, J. (2013). Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. *Finance and Stochastics, April 2013, Vol 17, Issue 2, pp 395–417.*

Nghiên cứu về một số cấu trúc phụ thuộc phi tuyến tính trong thiết kế danh mục đầu tư

Nguyễn Phúc Sơn^{1,*}, Phạm Hoàng Uyên², Nguyễn Đình Thiên²

¹Viện Nghiên cứu phát triển TP.HCM

²Trường Đại học Kinh tế - Luật, ĐHQG-HCM

*Tác giả liên hệ: sonnp@uel.edu.vn

Ngày nhận bản thảo: 21-8-2017, Ngày chấp nhận đăng: 13-10-2017; Ngày đăng: 15-7-2018

Tóm tắt—Thiết kế các danh mục đầu tư có lợi nhuận cao và rủi ro thấp luôn là đối tượng của các nhà nghiên cứu. Markowitz (1952) sử dụng các hệ số tương quan giữa các cặp cổ phiếu để xây dựng các danh mục thỏa mãn các mức rủi ro có thể chấp nhận được. Hệ số tương quan mô tả cấu trúc phụ thuộc tuyến tính giữa hai cổ phiếu nhưng không thể tích hợp được các cấu trúc độc lập phi tuyến tính. Vì vậy, hiệu quả của danh mục đầu tư đôi khi không đáp ứng được kỳ vọng của nhà đầu tư. Trong bài viết này, dựa trên lý thuyết copula của Sklar (xem [19]), chúng tôi kiểm tra một số phương pháp mới để xác định các cấu trúc phụ thuộc phi tuyến tính. Những

phương pháp mới này giúp chúng tôi ước lượng được phân bố của các danh mục, từ đó cho phép áp dụng các phương pháp ước lượng rủi ro phổ biến của các danh mục đầu tư như VaR. Chúng tôi áp dụng phương pháp này đối với các cổ phiếu niêm yết trên Sàn Giao dịch Cổ phiếu TP.HCM (HoSE), xây dựng một số danh mục tối ưu theo phương pháp của Markowitz cùng với các phương pháp ước tính rủi ro. Kết quả cho thấy, với các cấu trúc phụ thuộc phi tuyến tính, ước tính rủi ro của một số cặp cổ phiếu có những tác động đáng chú ý đến danh mục đầu tư. Kết quả này dẫn đến thay đổi các quyết định của nhà đầu tư.

Từ khóa—Thiết kế danh mục đầu tư, khoa học dữ liệu, cấu trúc phụ thuộc, copula, rủi ro, cổ phiếu, lợi nhuận, phương pháp.