APPENDICES

APPENDIX A

Forecasting risk and returns: CAPM model with belief functions

Sutthiporn Piamsuwannakit and Songsak Sriboonchitta

The original paper was published in “Econometrics of Risk”, Studies in Computational Intelligence Volume 583, pp 259– 271.

Forecasting risk and returns : CAPM model with belief functions

Sutthiporn Piamsuwannakit and Songsak Sriboonchitta

Abstract This paper presents a CAPM model with a belief function approach for forecasting the Integrated Oil and Gas Company (CHK) stock and the S&P500 index.The approach composed of two steps. First, we estimate the systematic risk or the beta coefficient in the CAPM model using the maximum likelihood method. Second, to improve the forecasting performance, we incorporate the likelihood-based belief function method. Likelihood-based belief functions are calculated from the historical data.The data set contains of 209 weekly returns during the period of 2010-2013. The finding shows evidence on systematic risk which is associated by the belief function derived from the distribution likelihood function given the market return.Finally, we use the method to predict the return of a particular stock.

1 Introduction

Most investors focus on the stock market return forecasting. The aim is to gain high profit by using the best trading strategies. The more successful in stock return prediction,the more profitable it becomes in stock market investment. The uncertainty and volatility of stock prices have an effect on the investor's decision.The knowledge on the dependence pattern between stock and market returns can help portfolio investors to diversify their assets better as well as reducing their risk at the suitable moments. The Capital Asset Pricing Model(CAPM) is a foundation and widely used model for evaluating the risk of a portfolio of assets with respect to the market risk which was introduced by Sharpe [21]. The CAPM is a linear model that esti-

Sutthiporn Piamsuwannakit
Faculty of Economics, Chiang Mai University
e-mail: sutthiporn_p@cmu.ac.th

Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University
e-mail: songsakecon@gmail.com

mates asset prices using the information on the risk free rate and the market returns. The CAPM takes into account the non- diversifiable risk, which is captured by the parameterβ.The CAPM non-diversifiable risk depends on the correlation between particular stock and overall stock market. Essentially, the standard CAPM model depends on the assumptions of normality of returns and quadratic utility functions of investors.

However, the numerous empirical evidences that have been carried out to analyze the applicability of CAPM in different stock markets have failed to maintain this relationship due to the inadequacy of the market beta alone in explaining the variations in stock returns and the assumptions of CAPM model. For example, Isa, M., Hassan, A. et al [11] applied CAPM in the Malaysian stock market by using the linear regression method,which was carried out on four models. The result indicated that both of the standard CAPM models with constant beta and time varying beta are statistically insignificant. On the other hand, the CAPM models conditional on segregating positive and negative market risk premiums are statistically significant. Nikolaos [18] evaluated of CAPM's validity in the British Stock Exchange. The result showed that under the two steps procedure, the CAPM does not have a statistical significance in portfolio selection. Choudhary, K. and Choudhary, S.[6] applied the CAPM model for the Indian stock.There is a lack of substantiating the theory's basic result illustrating that there is higher risk (beta) is associated with higher levels of return. Masood et al. [15] examined the validity of the CAPM in the capital markets of the Pakistan.The least squares method (OLS) is used to find the beta of the stocks in the first step and then searches for the regression equations in second step.The result showed that there is no support with the CAPM. The intercept term is equal to zero.Also, there is a positive relation between the risk and return.In addition,the market risk premium is a significant explanatory variable for the determining to see if the stock's risk premium are rejected. Zhang and Meng [22] analyzed the CAPM model in the Chinese stock market. The main problem of their studies was found that the effective test method did not exist.

From the above literature reviewed, CAPM is a useful tool to estimate the stock market return in different stock index. It can be concluded that there is no one model that can claim to have the absolute ability to predict the expected stock return by using the standard CAPM model. Then, there is a need of accurate forecast model that consistently predict uncertainty and volatility of the stock market prices. The stock market investor would be able to make decisions on the investment that is more informed and accurate . Therefore, various techniques are used for handling the uncertainty data. One such method applied is the Dempster-Shafer belief function theory , which is a useful tool for forecasting. Many studies have applied the belief function model to predict the uncertainty data. For instance, Nampak et al. [17] used the belief function model in order to forecast groundwater of specific area in Malaysia. Abdallah et al. [3] cooperated the statistical judgements with expert evidence by using belief function for prediction the future centennial sea level which climate change is considered. Kanjanatarakul et al. [13]used the Bass model for innovation diffusion together with past sales data and the formalism of belief functions to quantify the uncertainty on future sales. In their studies, a piece of evidence as a belief function was considered which can be viewed as the distribution of a random set. Furthermore, two main reasons for using the belief function formalism in this paper are the following :

1)The belief function approach does not require the statistician to arbitrarily provide a prior probability distribution when prior knowledge is not available.

2)We wish to measure the weight of statistical evidence that pertains to some specific questions, whereas confidence and prediction intervals are related to sequences of trials.

For more discussion on the comparison bet the belief function approach and classical methods of inference, the reader can find more information with the regards to the work done by Kanjanatarakul et al. [13].

In this contributions,we propose and alternative method for drawing inference via a likelihood based on a belief function approach for estimation of linear regression of CAPM. The objectives of this study are to (1) analyze the dependence pattern between the CHK stock and market returns and to(2) forecast the CHK stock returns using belief functions.

The remainder of the paper is organized as follows. Section 2 provides the Maximum Likelihood Estimation of capital asset pricing model and Section 3 introduces the prediction machinery using belief functions. Section 4 discusses the empirical solutions to the forecasting problem. The last section summarizes the paper.

2 Maximum Likelihood Estimation of capital asset pricing model

The CAPM represents a positive and linear relationship between asset return and systematic risk relative the overall market.The linear regression model is defined as

$$E(R_i) - R_f = \alpha + \beta E(R_m - R_f) \tag{1}$$

where $E(R_i)$ is the expected return of the asset, R_m is the expected market portfolio return, R_f is the risk free rate, α is the intercept and β is the equity beta,representing market risk. The observed the historical returns of stock $R_i = (r_{i1}, \cdots, r_{in})$ and returns from market $R_m = (r_{m1}, \cdots, r_{mn})$.The estimator of β is a measure of risk for financial analysis and also for risk and portfolio managers. The parameter β estimation procedure is defined by Arellano-Valle et al.[2] Let us consider in equation (1)has extended into equation(2) as follow:

$$r_i - r_f = \alpha + \beta(r_{mi} - r_f) + \varepsilon_i \tag{2}$$

or

$$y_i = \alpha + \beta x_i + \varepsilon_i \tag{3}$$

where r_i denotes the return of stock i , r_m is the market return and r_f corresponds to the risk free return , so that

$$y_i = r_i - r_f \tag{4}$$

and

$$x_i = r_m - r_f \tag{5}$$

represent the return of an asset in excess of risk free rate and the excess return of the market portfolio of assets.

The estimation method with the considering in the financial model is based on the least squares theory under the assumption of the random errors $\varepsilon_1, \cdots \varepsilon_n$ are independent and identically distributed according to the normal distribution.

$$N(\varepsilon_i, 0, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{\frac{-1}{2\sigma^2}(y - x\beta)^2\right\} \tag{6}$$

The likelihood function is given by

$$L = \Pi_{n=1}^{n} N(y_i; x_i, \beta, \sigma^2) = (2\pi\sigma^2)^{\frac{-n}{2}} exp\left\{\frac{-1}{2\sigma^2}(y - x\beta)'(y - x\beta)\right\} \tag{7}$$

3 Statistical inference and prediction using Belief functions

3.1 Belief functions

The theory of belief function is a formalism for reasoning with the uncertain, inaccurate and incomplete information.It was developed by Dempster [9] and later formalized by Shafer [20]. The model comprises several functions including Bel (degree of belief), Dis(degree of disbelief), Unc(degree of uncertainty) and Pls(degree of plausibilty), in range of $[0,1]$. Belief function can be defined on finite set and infinite set. Let us begin with finite case.

3.1.1 Belief functions on finite set

In the formalism of belief functions,we assign probabilities to sets (Pearl) [12].The belief model as given below,see Frikha [10], Liu et al [14],Nampak et al [17].

Let Θ be a finite set ,Θ is called frame of discernment of the problem of consideration.The power set of Θ, denoted by 2^{Θ}.

A basic probability assignment (BPA) is a function $m(.)$ from 2^{Θ} to$[0,1]$ that assigns a number $[0,1]$ to each subset A of Θ . The quantity $m(A)$,called the mass of A, which represents the degree of belief attributed exactly to A, and to no one of its subsets. This function satisfies the following condition :

$$0 \leq m(A) \leq 1,\ m(\phi) = 0,\ \sum_{A \subseteq \Theta} m(A) = 1 \tag{8}$$

When $m(A) > 0$, A is called focal element of m. To each BPA, we can associate a belief function and a plausibility function are a mapping $Bel(A) : 2^{\Theta} \rightarrow [0,1]$ and $Pl(A) : 2^{\Theta} \rightarrow [0,1]$ respectively, defined as:

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{9}$$

$$pl(A) = \sum_{A \cap B \neq \phi} m(B) \tag{10}$$

$Bel(A)$ measures the total belief completely attributed to $A \subseteq \Theta$.It is interpreted as the lower bound of probability of A. $Pl(A)$ is interpreted as the upper bound of probability of A.
The two functions satisfied the following properties:

$$Bel(A) \leq Pl(A) \tag{11}$$

$$Pl(A) = 1 - Bel(\bar{A}) \tag{12}$$

Where $\bar{A}$ is the complement of A and $Bel(\bar{A})$ is called a degree of disbelief in A.

$$Pl(A) - Bel(A) = Unc \tag{13}$$

Eq.(13) represents the difference between belief and plausibility.
If $Unc = 0$, then $Bel(A) = Pl(A)$.
Fig.1. shows a schematic description of the relationship between belief,disbelief and uncertain functions.

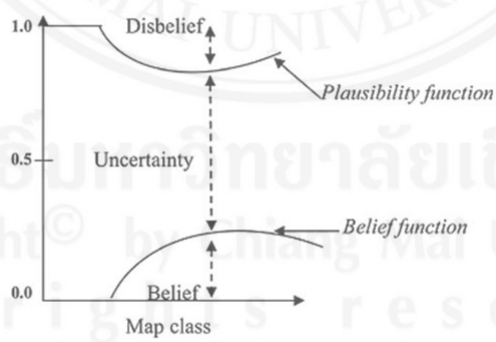


Fig. 1 Schematic description of the relationship between belief,disbelief and uncertainty(Carranza et al.,2005).

3.1.2 Belief functions on infinite set

In an infinite case, there may not be a mass function associated with completely monotone function as in the finite case, Denoeux [7].The definitions are provided which defined by Denoeux [7] as following;

Let (Ω, B) be a measurable space (i.e., B is a sigma–field, that is a non–empty subset of 2^{Θ} closed under complementation and countable union). A belief function on B is a function $Bel : B \to [0,1]$ verifying the following three conditions:

1.$Bel(\phi) = 0$

2.$Bel(\Omega) = 1$

3.For any $k \leq 2$ and any collection $B_1, \cdots, B_k$ of elements of B,

$$Bel(U_{i=1}^{k} B_i) \geq \sum_{\phi \neq I(1,\cdots,k)} (-1)^{|I|+1} Bel(\cap_{i \in I} B_i) \tag{14}$$

Furthermore, a belief function Bel on (Ω, B) is continuous if for any decreasing sequence $B_1 \supset B_2 \supset B_3 \supset \cdots$ of elements of B,

$$\lim_{i \to +\infty} Bel(B_i) = Bel(\cap_{i \in I} B_i) \tag{15}$$

3.2 Likelihood–based belief functions

The likelihood-based belief functions have been derived by Shafer [20].They have been applied by Abdallah et al [3],among others,and justified by Denoeux[8].

Let $x \in X$ be the observable data with a probability density function (pdf) $p_{\theta}X$, where $\theta \in \Theta$ is an unknown parameter. In this paper,we use the method proposed by Shafer [20].The belief function be derived from the Likelihood Principle and Least Commitment Principle(LCP).The information about Θ can be represented by the likelihood function which is defined by $L_x(\theta) = p_{\theta}X$ for all $\theta \in \Theta$. The likelihood ratio is meant to be a ”relative plausibility”, which can be written as:

$$\frac{pl_x(\theta_1)}{pl_x(\theta_2)} = \frac{L_x(\theta_1)}{L_x(\theta_2)} \tag{16}$$

for all $(\theta_1, \theta_2) \in \Theta^2$ or, equivalently, $pl_x(\theta) = cL_x(\theta)$
for all $\theta \in \Theta$ and some positive constant c. From LCP,it can be implied that the highest possible value of c is $\frac{1}{sup_{\theta \in \Theta}} L(\theta|x)$.Thus ,the contour function is defined as follow:

$$pl(\theta;x) = \frac{L(\theta;x)}{sup_{\theta \in \Theta} L(\theta;x)} \tag{17}$$

The information about θ are expressed by the belief function Bel_A^{Θ} with contour function pl_x, i.e., with corresponding plausibility function $pl_x^{\Theta}(A) = sup_{\theta \in A} pl_x(A)$, for all $A \subseteq \Theta$ The focal sets of Bel_A^{Θ} are the levels sets of pl_x defined as follows:

$$\Gamma_x(\omega) = \{\theta \in \Theta | pl_x(\theta) \geq \omega\} \tag{18}$$

for $\theta \in [0,1]$. Equation(18) is called plausibility regions.With the inducing of the Lebesgue measure λ on [0,1] and multi–valued mapping Γ_x from $[0,1] \rightarrow \Theta^2$ the belief function is equivalent to the random set,see Kanjanatarakul et al [13]. We remark that the MLE of θ is the value of θ with highest plausibility.

3.3 Incorporating the belief functions

The objective of this section is to forecast the risk premium of the return of stock i, $y_i = r_i - r_f$. The methodology to incorporate the belief function framework into the prediction procedure follows Kanjanatarakul et al [13]. From the CAPM equation from the previous section, the return equation can be written as:

$$y_i = \alpha + \beta x + \sigma F^{-1}(u) \tag{19}$$

where $F \sim Normal(0,1)$ and $U \sim Uniform(0,1)$

As discussed in Kanjanatarakul et al [13], the forecasting problem is the inverse problem of the regular inference problem. Given the knowledge on the set of parameters $\theta = (\alpha, \beta, \sigma)$ and the distribution $F(.)$, the future value of y_i can be forecasted.

Belief function framework allows us to forecast an interval $[y_i^L, y_i^U]$ for the future value of y_i. The estimation of $[y_i^L, y_i^U]$ can be done using Monte Carlo method. Given a set of two independently $Uniform(0,1)$ random variables (u_s, ω_s), in each simulation s, the lower bound $y_{i,s}^L$ and the upper bound $y_{i,s}^U$ solve the following optimization problems respectively,

$$y_{i,s}^L = min_{\theta} \alpha + \beta x + \sigma F^{-1}(u_s) \tag{20}$$

subject to

$$pl(\theta) \geq \omega_s \tag{21}$$

and

$$y_{i,s}^U = max_{\theta} \alpha + \beta x + \sigma F^{-1}(u_s) \tag{22}$$

subject to

$$pl(\theta) \geq \omega_s \tag{23}$$

In the constraints, the plausibility function $pl(\theta)$ can be derived from the likelihood function. Therefore, using the likelihood function in equation(7), the plausibility function is as follows:

$$pl(\theta) = \frac{L(\theta)}{L(\theta^*)} \tag{24}$$

where θ^* is such that $L(\theta^*) \geq L(\theta)$, $\forall\theta$. The belief and the plausibility functions corresponding to a given set A can be calculated by:

$$Bel(A) = \frac{1}{N}\sharp\{s \in \{1,\cdots,N\} | [y^L_{i,s}, y^U_{i,s}] \subset A\} \tag{25a}$$

$$Pl(A) = \frac{1}{N}\sharp\{s \in \{1,\cdots,N\} | [y^L_{i,s}, y^U_{i,s}] \cap A \neq \emptyset\} \tag{25b}$$

The lower and the upper of the expectation for y_i is, thus,

$$\hat{y^L_i} = E(y^L_{i,s}) = \frac{1}{N}\Sigma y^L_{i,s} \tag{26a}$$

$$\hat{y^U_i} = E(y^U_{i,s}) = \frac{1}{N}\Sigma y^U_{i,s} \tag{26b}$$

4 An application to stock market

4.1 Data

The data contain of 209 weekly returns during the period of 2010–2013 :they were obtained from Yahoo Finance to compute the log returns on integrated oil and gas company (CHK) stock.The log returns prices by using the formula:

$$r_t = ln(\frac{P_t}{P_{t-1}}) \tag{27}$$

where P_t and P_{t-1} are the weekly closing prices at time t and $t-1$ respectively. Mukherji [16] indicated that the treasury bills are better proxies for the risk-free rate than longer–term treasury securities regardless of the investment horizon,which is only related to the U.S. market.In this paper,the treasury bills stand for the risk free rate. The daily returns of the treasury bills are adjusted to the weekly returns and can be used in this manner by using the compound interest that take form:

$$I_{wj} = \{\prod_{i=1}^{N}(1+I_{di})\} - 1 \tag{28}$$

Where $I_{wj}, j = 1,,N$ is the weekly interest rate and $I_{di}, i = 1,,N$ is the daily interest rate. The Maximum Likelihood estimates of the parameters are shown in Table 1

Table 1 Parameter estimation results

Stock name	Parameters	
CHK	β_0	-0.001(0.0031)
	β_1	1.436(0.1417)
	σ^2	.0020(1.91739e^{-4})

[a] * Standard errors in parentheses.

Figure 2 displays two–dimensional marginal contour functions, with one of the three parameters fixed to its MLE

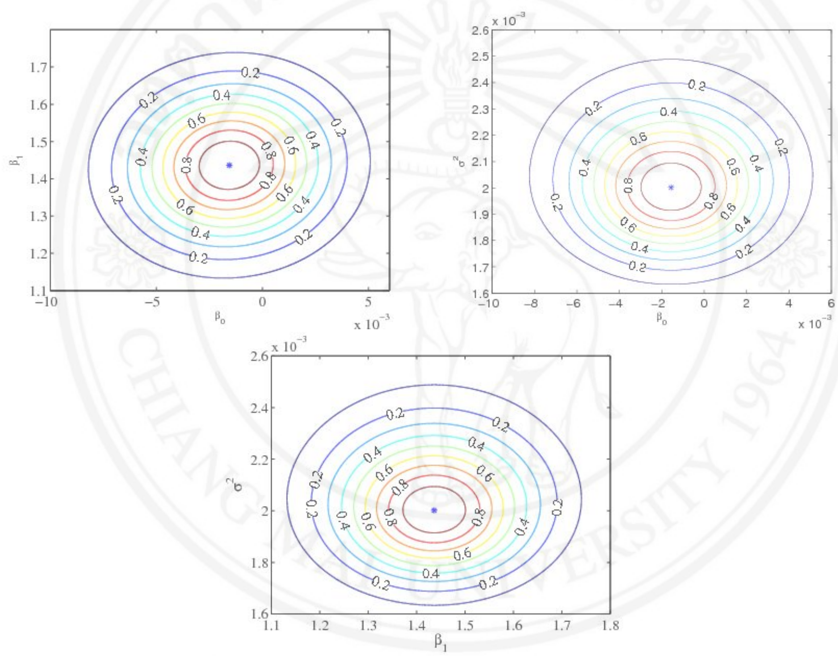


Fig. 2 Displays two-dimensional marginal contour functions

Figure 3 shows the marginal contour functions for parameters β_0, β_1,σ^2. These three plausibilities will be used to perform plausibility intervals for each of the three parameters.

To predict the expected return of the asset $y_{i,n+1}$ for a new market portfolio return $X_{i,n+1}$ we compute the minimum and maximum of $y_{i,n+1}$ given $X_{i,n+1}$ by

$$y_{i,n+1} = \beta_0 + \beta_1 X_{i,n+1} + \sigma F^{-1}(u_s) \quad (29)$$

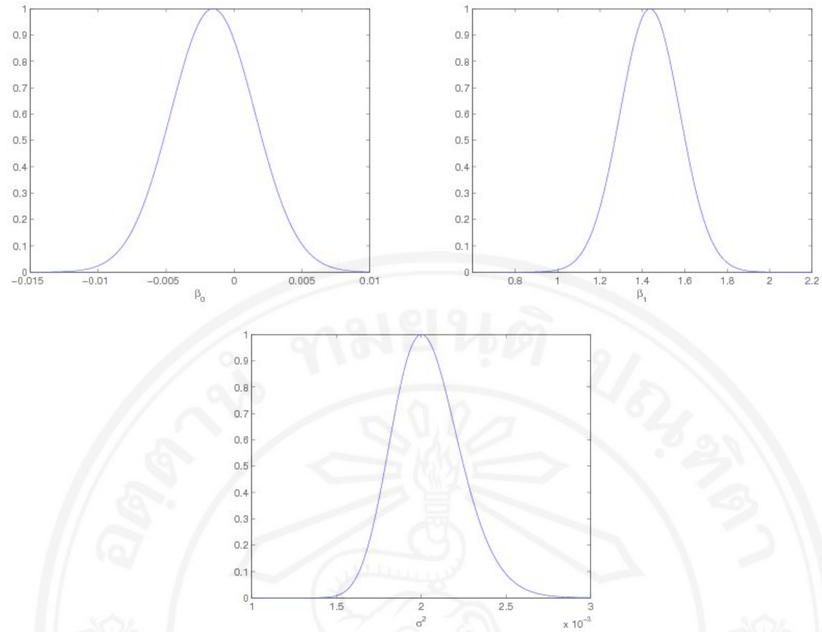


Fig. 3 Marginal plausibility of β_0, β_1 and σ^2.

under the constraint $pl(\theta) \geq \omega_s$, where $F^{-1}(u_s)$is the inverse cumulative distribution function (cdf) of the normal distribution and u, ω are independent random variables with the same uniform distribution U ([0, 1]).Given (29), we randomize independently N pairs of the random number,$(u_s, \omega_s); s = 1, 2, , N$ resulting in N intervals $[y^L_{i,s}(u_s, \omega_s), y^U_{i,s}(u_s, \omega_s]$ For any $A \subset R$, the stock returns $Bel_{y_i}(A)$ and $Pl_{y_i}(A)$ can be estimated by equation (3). The estimated lower and upper expectations of $r_{a,n+1}$ are then:

$$\bar{y}^L_{i,s} = \sum_{s=1}^{N} \frac{y^L_s(u_s, \omega_s)}{N} \tag{30}$$

$$\bar{y}^U_{i,s} = \sum_{s=1}^{N} \frac{y^U_s(u_s, \omega_s)}{N} \tag{31}$$

Figures (4) displays the lower and upper cdfs $Bel_{y_i}([-\infty, y_i])$ and $Pl_{y_i}([-\infty, y_i])$ .This function give us the summary of the predictive belief function Bel_{y_i}.

Figure (5) shows the upper and lower bound of stock return via CAPM using belief function.

The another representation of uncertainty prediction can be defined as the lower–upper expectations of stock returns, the uncertainty and randomness estimation are

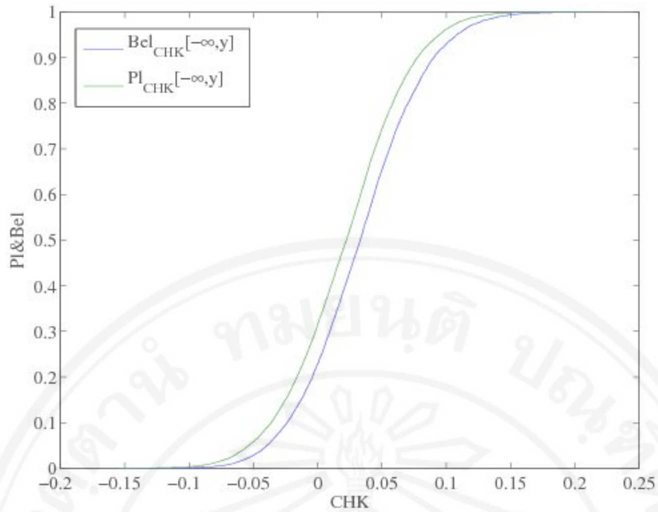


Fig. 4 Lower and Upper cumulative distribution function

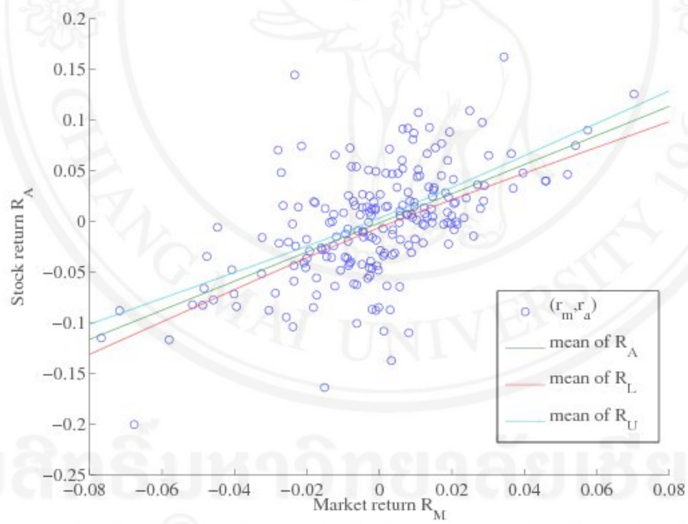


Fig. 5 Lower and Upper interval of stock return via CAPM using belief function

considered. From the empirical result, the gap between the lower and upper cdfs is quite narrw , which shows that estimation uncertainty is small as compared to random uncertainty. Therefore,the investor can use these results to increase the gain of portfolio investment(Autchariyapanikul et al[1].

5 Conclusions

In this paper, we presented the method of standard CAPM with normal distribution for CHK stock in S&P500 in the belief function framework. The Dempster-Shafer belief function theory was used in order to identify the uncertainty .The statistical prediction based on historical data and a financial model. This method consists of two steps.First, a belief function is defined from the normalized likelihood function given the past data which is refered to the uncertainty on the parameter vector θ . Second, the return of stock y_i is illustrated as $\varphi(\theta,u)$,where u is a stochastic variable with known distribution. Then, belief on θ and u are transferred through φ ,resulting in a belief function on y_i. This approach has been adapted to the prediction of the stock returns. A possible extension of this work is to consider uncertainty on the independent variable r_m, which can also be expressed as a belief function and combined with other uncertainties to compute a belief function on y_i.

Acknowledgements
We would like to thank the referees for their comments and suggestions, and Somsak Chanaim,a research assistant at Faculty of Economics , Chaing mai University for his helpful program running output.

References

1. Autchariyapanitkul, K., Chanaim, S. ,Sriboonchitta S.,and Denoeux,T. (2014). Predicting stock returns in the capital asset pricing model using quantile regression and belief functions. In Proceedings of the 3rd Int. Conf. on Belief functions (BELIEF 2014), Oxford, UK. Springer-Verlag.
2. Arellano-Valle, R.B. ,Bolfarine, H. Iglesias, P.L. and Viviani,P.(2010). Portfolio selection : An application to the Chilean stock market. Chilean Journal of Statistics, Vol. 1, No. 2, 315.
3. Ben Abdallah, N., Mouhous-Voyneau, N. and Denoeux, T. (2014). Combining statistical and expert evidence using belief functions: Application to centennial sea level estimation taking into account climate change. International Journal of Approximate Reasoning, 55(1), 341-354.
4. Carranza, E. J. M., Woldai, T. and Chikambwe, E. M. (2005). Application of data-driven evidential belief functions to prospectivity mapping for aquamarine-bearing pegmatites, Lundazi district, Zambia. Natural Resources Research, 14(1), 47-63.
5. Chang, Eric C, Cheng, Joseph W. and Khorna, Ajay.(20005).An Examination of Herd Behavior in Equity Markets: An International Perspective?, Journal of Banking and Finance, Vol. 24, pp. 1651-1679 (2000)

6. Choudary, K., and Choudhary, S. (2010). Testing Capital Asset Pricing Model: Empirical Evidences from Indian Equity Market. Eurasian Journal of Business and Economics, 3(6), 127-138.
7. Denoeux,, T. (2013). Belief functions on infinite spaces[class handout].Faculty of economics,Chiang mai University,Chiang mai, Thailand.
8. Denoeux, T. (2014). Likelihood-based belief function: Justication and some extensions to low-quality data. International Journal of Approximate Reasoning, Volume 55, Issue53 7, 15-1547.
9. Dempster, A. P.,(1967).Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Stat. 38, 325-339.
10. Frikha, A. (2014). On the use of a multi-criteria approach for reliability estimation in belief function theory. Information Fusion, 18, 20-32.
11. Isa, M., Hassan, A., Puah, C. H., and Yong, Y. K. (2008). Risk and return nexus in Malaysian stock market: Empirical evidence from CAPM. Online at http://mpra.ub.uni-muenchen.de/12355/
12. Pearl, J. (1990). Reasoning with belief functions: an analysis of compatibility. International Journal of Approximate Reasoning, 4(5), 363-389.
13. Kanjanatarakul, O., Sriboonchitta, S., and Denoeux, T. (2014). Forecasting using belief functions: An application to marketing econometrics. International Journal of Approximate Reasoning, 55(5), 1113?1128.
14. Liu, Z. G., Pan, Q., Dezert, J., and Mercier, G. (2014). Credal classification rule for uncertain data based on belief functions. Pattern Recognition, 47(7), 2532-2541.
15. Masood, S., Saghir, G., and Muhammad, W. (2012). The Capital Asset Pricing Model: Empirical Evidence from Pakistan.Online at http://mpra.ub.uni-muenchen.de/41961/
16. Mukherji S. (2011).The capital asset pricing model's risk-free rate, International Journal of Business and Finance Research, 5,793-808.
17. Nampak, H., Pradhan, B., and Manap, M. A. (2014). Application of GIS based data driven evidential belief function model to predict groundwater potential zonation. Journal of Hydrology, 513, 283-300.
18. Nikolaos, L. (2009). An empirical evaluation of CAPM?s validity in the British stock exchange. Iinternational Journal of Applied Mathematics and Informatics,3(1), 1-8.
19. Shalit, H., and Yitzhaki, S. (2002). Estimating beta. Review of Quantitative Finance and Accounting, 18(2), 95-118.
20. Shafer, G. (1976). A Mathematical Theory of Evidence, Princeton University Press, Princeton ,NJ.
21. Sharpe, William F. (1964). "Capital Asset Prices ? A Theory of Market Equilibrium Under Conditions of Risk". Journal of Finance XIX (3): 425?42.
22. Zhang, P., and Meng, X. (2013). The market application analysis of CAPM model in China's securities. In 2nd International Conference On Systems Engineering and Modeling (ICSEM-13). Atlantis Press.

APPENDIX B

Capital Asset Pricing Model with Interval Data

Sutthiporn Piamsuwannakit ,Kittawit Autchariyapanitkul,
Somsak Chanaim and Songsak Sriboonchitta

The original paper was published in " Integrated Uncertainty in Knowledge Modelling and Decision Making" LNAI 9376, pp 163 – 170.

Capital Asset Pricing Model with Interval Data

Sutthiporn Piamsuwannakit[†], Kittawit Autchariyapanitkul [*], Songsak Sriboonchitta[⋆], and Rujira Ouncharoen[‡]

[†,⋆]Faculty of Economics, Chiang Mai University, Chiang Mai, Thailand.
[†]Faculty of Management Science, Chiang Rai Rajabhat University, Chiang Rai, Thailand.
[*]Faculty of Economics, Maejo University, Chiang Mai, Thailand.
[‡]Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand. {Corr.Auther:kittawit_autchariya@cmu.ac.th}

Abstract. We used interval-valued data to predict stock returns rather than just point valued data. Specifically, we used these interval values in the classical capital asset pricing model to estimate the beta coefficient that represents the risk in the portfolios management analysis. We also use the method to obtain a point valued of asset returns from the interval-valued data to measure the sensitivity of the asset return and the market return. Finally, AIC criterion indicated that this approach can provide us better results than use the close price for prediction.

Keywords: CAPM, Interval-Valued Data, Least Squares Method, Linear Regression

1 Introduction

Capital asset pricing model provides a piece of information of asset return related to the market return via its systematic risk. In general, asset returns of any interested asset and market returns are calculated from a single-valued data. Most of the papers in financial econometrics use only closed price taking into account for calculation but in the real world stock price is moving up and down within the range of highest price and lowest price. So, in this paper we intend to use all the points in the range of high and low to improve the results in our calculations. We also put an assumption of a normal distribution on these interval-valued data.

An enormous number of research on CAPM model with single-valued data could be found in much financial research topic, the reader is referred to, e.g., William F. Sharpe [1] and John Lintner [2] only a single-valued of interest was considered. Many various technics were applied to the original CAPM model that we can found in the work from Autchariyapanitkul et al. [3], the authors used quantile regression under asymmetric Laplace distribution (ALD) to quantify the beta of the asset returns in CAPM model. The results showed that this method can capture the stylized facts in financial data to explain the return of stocks under quantile, especially under the middle quantile levels. In Barnes and Hughes [4], the beta risk is significant in both tails of the conditional distribution of returns. In Chen et al. [5], the authors used a couple of methods to obtain the time-varying market betas in CAPM to analyze stock in the Dow Jones Industrial for several quantiles. The results indicated that smooth transition quantile method performed better than others methods.

Interval-valued data has become popular in many research fields especially in the context of financial portfolio analysis. Most of the financial data are usually affected by imprecision, uncertainty, inaccuracy and incompleteness, etc. The uncertainty in the data may be captured with interval-valued data. There are several existing research in the literature for investigating this issue. see Billard [6], Carvalho [7], Cattaneo [8], Diamond [9], Gil [10], Kӧrner [11], Manski [12], Neto [14]. However, In these research papers are lacking in a foundation and theoretical background to support this idea.

The connection between the classical linear regression and the interval-valued data that share the important properties could be found for the work by Sun and Li [15]. In their paper, they provided a theoretical support framework between the classical one and the interval-valued linear regression such as least squares estimation, asymptotic properties, variances estimation, etc. However, in their paper only one of an explanatory variable can use to described the responding variable. In this paper, we intend to apply the concept of the interval-valued data to the CAPM model. We replace a single value of market returns and asset returns with the range of high and low historical data into the model.

The rest of the paper is organized as follows. Section 2 gives a basics knowledge of a linear regression model for interval-valued data. In Section 3 discusses the empirical discovering and the solutions of the forecasting problem. The last section gives the conclusion and extension of the paper.

2 A Review of Real Interval-Valued Data

Now, take a close look at financial data (D_i). Suppose, we have a range of any numbers between a minimum and maximum prices given by $D_i = [\min, \ldots, \max] = [\text{Low},...,\text{High}]$, where the minimum price is the "lowest price", and the maximum price is the "highest price". Certainly, this range contains the point that we called "close price". In many research papers, they are usually using the close price for calculations. A close price is a number that takes any values in the range of D_i between the lowest and the highest prices, $D_i = [\text{Low},...\text{Close},...,\text{High}]$. The close price could be either the lowest price or the highest price.

In this paper, we try to find the better value for calculations rather than a close price that is the best-represented point in the range of D_i to improving our predictions. We considered a normal distribution on this interval-valued data.

3 An interval-valued data in a linear regression model

Suppose we can observe an i.i.d random paired intervals variables $x_i = [\underline{x_i}, \overline{x_i}]$ and $y_i = [\underline{y_i}, \overline{y_i}],\ i = 1,2,\ldots,n$ where $\overline{x_i}, \overline{y_i}$ are the maximum values of x_i and $\underline{x_i}, \underline{y_i}$ are the minimum values of y_i. Additionally, we can rewrite the value of x_i, y_i in the form of intervals as

$$x_i = [x_i^m - x_i^r, x_i^m + x_i^r], \tag{1a}$$

$$y_i = [y_i^m - y_i^r, y_i^m + y_i^r],\ i = 1,2,\ldots,n, \tag{1b}$$

where x_i^m, y_i^m is the mid-points of x_i and y_i and x_i^r, y_i^r is the radii of x_i and y_i, satisfying $x_i^r, y_i^r \geq 0$. Suppose, we consider the following linear regression model given by

$$y_i = ax_i + b + \varepsilon_i,\ i = 1, 2, \cdots, n. \quad (2)$$

Analogously, it is easy to interpret the meaning of x_i, y_i by the distance of centers and radii as the following equations

$$x_i = x_i^m + \delta_{x_i},\ \delta_{x_i} \in N(0, (k_0 \Delta x_i)^2) \quad (3a)$$

$$y_i = y_i^m + \delta_{y_i},\ \delta_{y_i} \in N(0, (k_0 \Delta y_i)^2), \quad (3b)$$

where x_i^m, y_i^m are the centers of x_i and y_i, respectively. Then, $\Delta x_i = \frac{\overline{x_i} - \underline{x_i}}{2}, \Delta y_i = \frac{\overline{y_i} - \underline{y_i}}{2}$ are the radii of x_i and y_i, respectively and $x_i^m = \frac{\overline{x_i} + \underline{x_i}}{2}$, $y_i^m = \frac{\overline{y_i} + \underline{y_i}}{2}$ are the mid-point of x_i and y_i, respectively.Thus, given the linear regression for the interval valued data we have

$$y_i^m + \delta_{y_i} = ax_i^m + a\delta_{x_i} + b \quad (4a)$$

$$y_i^m = ax_i^m + b + (a\delta_{x_i} - \delta_{y_i}), \quad (4b)$$

where $(a\delta_{x_i} - \delta_{y_i}) \sim N(0, \sigma^2) \equiv N(0, k_0^2 a^2 \Delta x_i^2 + \Delta y_i^2)$. Assume that $a\delta_{x_i} - \delta_{y_i}$ is an independence. Thus, we can estimate parameters a, b, k_0 by the maximum likelihood function given by

$$\max_{a,b,k_0} L(a, b, k_0 | ([\underline{x_i}, \overline{x_i}], [\underline{y_i}, \overline{y_i}]), i = 1, \ldots, n)$$
$$= \max_{a,b,k_0} \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi k_0^2 (a^2 \Delta x_i^2 + \Delta y_i^2)}} \exp\left[-\frac{1}{2} \frac{(y_i^m - ax_i^m - b)^2}{k_0^2 (a^2 \Delta x_i^2 + \Delta y_i^2)} \right] \right) \quad (5)$$

This approach was already developed in Sun and Li [15]. And soften the criticisms of lack of theory, Manski has a whole book (see, Manski [12],[13]), this is finance not pure mathematics here. The proof of success is better fit not theorems.

3.1 Goodness of fit in linear regression model for an interval-valued data

In the deterministic linear regression model, we use variance to describe variation of the variable interested and so that as we knew the ratio $\frac{a^2 Var(X)}{Var(Y)} \in [0, 1]$ can be explained as an indication of goodness-of-fit. In this paper, we used the concept of the chi-squared test (χ^2) of the goodness of fit. Recall that $\sigma_{x_i} = k_0 \Delta x_i$ and $\sigma_{y_i} = k_0 \Delta y_i$ given the simple linear regression we have

$$y_i = ax_i + b \quad (6a)$$

$$y_i^m + \delta_{y_i} = ax_i^m + a\delta_{x_i} \quad (6b)$$

$$y_i^m - ax_i^m - b = a\delta_{x_i} - \delta_{y_i}, \quad (6c)$$

where $\delta_{x_i}, \delta_{y_i} \sim N(0, \sigma^2)$. Thus, we have $a^2\sigma_{x_i}^2 + \sigma_{y_i}^2$, by replacing $k_0^2(a^2\Delta x_i^2 + \Delta y_i^2)$ to above equation 6. The empirical χ^2-test is obtained by estimated this following equation

$$\chi_{cal}^2 = \sum_{i=1}^{n} \frac{(y_i^m - a x_i^m - b)^2}{k_0^2(a^2\Delta x_i^2 + \Delta y_i^2)}, \tag{7}$$

where the degree of freedom is $n-2$.

4 An application to the stock market

We consider the following financial model that is so called Capital Asset Pricing Model (CAPM). Only two sets of interval-valued data are used to explain the relationship of the asset. The fitted model is based on the least square estimation.

4.1 Capital Asset Pricing Model

The Capital Asset Pricing Model (CAPM) is a linear relationship that was created by William F. Sharpe [1] and John Lintner [2]. The CAPM use to calculate a sensitivity of the expected return on the asset to expected return on the market. The combination of a linear function of the security market line:

$$E(R_A) - R_F = \beta_0 + \beta_1 E(R_M - R_F), \tag{8}$$

where $E(R_A)$ explains the expected return of the asset, R_M represents the expected market portfolio return, β_0 is the intercept and R_F is the risk-free rate. $E(R_M - R_F)$ is the expected risk premium, and β_1 is the equity beta, denoting market risk. To measure the systematic risk of each stock via the beta takes form:

$$\beta_1 = \frac{cov(R_A, R_M)}{\sigma_M^2}, \tag{9}$$

where σ_M^2 represents the variance of the expected market return. Given that, the CAPM predicts portfolio's expected return should be about its risk and the market returns.

4.2 Beta estimation with interval data

From the deterministic model in equation (8), we calculate the β coefficient through the likelihood by equation (5) instead. Suppose we have observed the realization interval stock return $[\overline{R_A i}, \underline{R_A i}] = [(\overline{r_{a1}}, \underline{r_{a1}}), \ldots, [\overline{r_{an}}, \underline{r_{an}})], i = 1, 2, \ldots, n$ and return from market $[\overline{R_M i}, \underline{R_M i}] = [(\overline{r_{m1}}, \underline{r_{m1}}), \ldots, (\overline{r_{mn}}, \underline{r_{mn}})], i = 1, 2, \ldots, n$ over the past N years. These observations will be assumed an independent random. From likelihood for an interval values we have

$$\max_{a,b,k_0} L(a, b, k_0 | ([\underline{R_{Mi}}, \overline{R_{Mi}}], [\underline{R_{Ai}}, \overline{R_{Ai}}]), i = 1, \ldots, n)$$

$$= \max_{a,b,k_0} \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi k_0^2(a^2\Delta Rm_i^2 + \Delta Ra_i^2)}} \exp\left[-\frac{1}{2} \frac{(Ra_i^m - aRm_i^m - b)^2}{k_0^2(a^2\Delta Rm_i^2 + \Delta Ra_i^2)} \right] \right) \tag{10}$$

4.3 Empirical results

Our data contains 259 weekly interval-valued returns in total during 2010-2015 are obtained from Yahoo. We compute the log returns on the following stock, namely, Chesapeake Energy Corporation (CHK)and Microsoft Corporation (MSFT). Due to significant capitalization and high turnover volume.

In this paper, we use Treasury bills as a proxy. From Autchariyapanitkul et al. [3] and Mukherji [16] suggested that Treasury bills are better proxies for the risk-free rate, only related to the U.S. market.

Table 1. Estimated parameter results for CHK

	Interval-Valued data		Point-Valued data	
parameters	values	std. Dev.	values	std. Dev.
β_0	–0.0021	0.0233	–0.0191	0.0055
β_1	0.9873	0.0914	0.7226	0.0713
k	0.4472	0.0845	-	-
MSE	-	-	0.036	
LL	525.7021	-	361.1400	-
χ^2	259.00	-	-	-
AIC	–1045.04	-	–716.28	-

Table 2. Estimated parameter results for MSFT

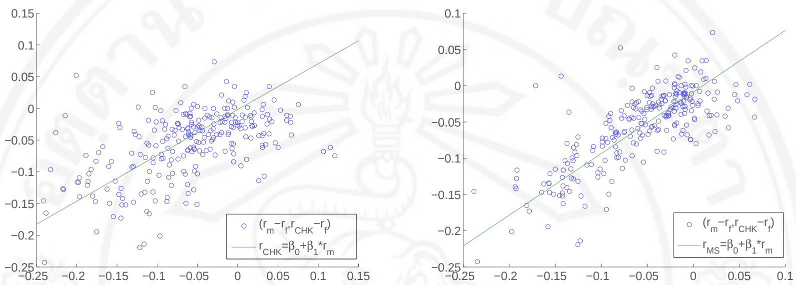
	Interval-Valued data		Point-Valued data	
parameters	values	std. Dev.	values	std. Dev.
β_0	–0.0004	0.0015	–0.0088	0.0035
β_1	1.0086	0.0220	0.8489	0.0005
k	0.4017	0.0170	-	-
MSE	-	-	0.0025	-
LL	692.3808	-	478.9365	-
χ^2	259.00	-	-	-
AIC	–1378.76	-	–951.87	-

Table 1 and Table 2 report the estimated results from equation (5). For example, the simple linear regression model for the asset returns (Y) and the market returns (X) for interval valued data for CHK is written to be

$$R_A = -0.0021 + 0.9873 R_M. \tag{11}$$

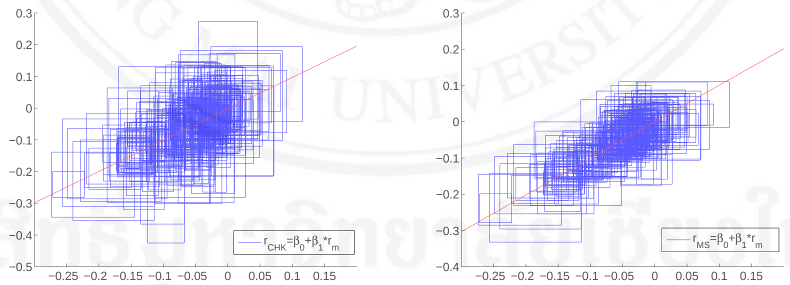
From the above linear equation, the return of a stock is likely to increase less than the return from the market. A non-parametric chi-square test is used to validate the method of interval-valued data. The theoretical χ^2_{n-2} gives the value of CHK, $\chi^2_{n-2} = 303.2984$ compare with the empirical value $\chi^2_{emp} = 259.00$ confirm that the market returns can be used to explain the asset returns. The model selection criteria Akaike information criterion (AIC) was employed to compare these two techniques. The AIC of interval-valued data gives a value of –1051.4402 is smaller than the AIC of pointed-valued data, which indicate that the results from the interval-valued method is more prefer than the deterministic one.

The relationship between market return and asset return are plotted in Figure 1 and Figure 2 for pointed-valued data and interval-valued data, respectively.



(a) CHK (b) MSFT

Fig. 1. : Securities characteristic line for point valued data



(a) CHK (b) MSFT

Fig. 2. : Securities characteristic line for interval valued data

The rectangular are the high and low interval-valued data, and the straight line is the securities characteristic line, the slope of this straight line represent the systematic risk beta. All investments and portfolio of investments must lie along a straight line in the return beta space.

5 Conclusions and Extension

The systematic risk has played as the critical role of financial measurement in capital asset pricing model. Academic and practitioners attempt to estimate its underlying value accurately. Fortunately, there have been the novel approaches to evaluating the beta with interval-valued data. We used every price range of real world data to obtained the single value of the systematic risk same as the results from the conventional CAMP model.

In this paper, we use our approach to an interval-valued data in CAPM for only one stock in *S&P*500 for a demonstration. With this, a method can be used to investigate the linear relationship between the expected asset returns and its asymmetric market risk by including all of the levels of prices in the range of an interval-valued data. The results clearly show that the beta can measure the responsiveness to the asset returns and market returns. However, only a systematic risk is calculated through the model, and we neglect the unsystematic risk under CAPM assumption. CAPM concludes that the expected return of a security or a portfolio equals the rate on a risk-free security plus a risk premium.

By AIC criterion, it should be noticed that the estimation by using interval-valued data more reasonable than just used the single valued in the calculations. Not only one explanatory variable can be used to explain the outcome variable but with this method also allowed us to use more than one covariate in the model.

For future research, we are interested to use this method to the time series models such as ARMA, GARCH model. Additionally, we can use this method to the model with more than one explanatory variables such as Fama and French (1993). A three-factor model can be extended the CAPM by putting size and value factors in the classical one.

Acknowledgement(s) : The authors thank Prof. Dr. Vladik Kreinovich for his suggestions for an interval-valued data computation method. We would like to thank the referee for giving comments on manuscript.

References

1. William, F., Sharpe. Capital Asset Prices A Theory of Market Equilibrium Under Conditions of Risk, *Journal of Finance*, Vol. 19(3), pp. 425–442, 1964.
2. John, Lintner. The Valuation of Risk Assets and the Selection of Risky Investments in Stock. *The Review of Economics and Statistics*. Vol. 47(1), pp. 13–37, 1965.
3. Autchariyapanitkul, K., Chainam, S. and Sriboonchitta, S. Quantile regression under asymmetric Laplace distribution in capital asset pricing model, *Econometrics of Risk* , Springer International Publishing, 583, pp. 219–231, 2015.
4. Barnes, L.M. and Hughes, W.A. A Quantile Regression Analysis of the Cross Section of Stock Market Returns. Federal Reserve Bank of Boston, *working paper*, 2002.
5. Chen, W.S.C., Lin, S. and Yu, L.H.P. Smooth Transition Quantile Capital Asset Pricing Models with Heteroscedasticity. *Computational Economics*, 40, pp. 19–48, 2012.

6. Billard, L. Dependencies and variation components of symbolic interval-valued data. Selected Contributions in Data Analysis and Classification, pp.3–12, 2007.
7. Carvalho, F.A.T, Neto, E.A.L. and Tenorio, C.P. A new method to fit a linear regression model for interval-valued data. *Lecture Notes in Computer Science*, 3238, pp.259–306, 2012.
8. Cattaneo, M.E.G.V. and Wiencierz, A. Likelihood-based imprecise regression. *International Journal of Approximate Reasoning*, 53, pp.1137–1154, 2012.
9. Diamond, P. Least square fitting of compact set-valued data. *J. Math. Anal. Appl.*, 147, pp.531–544, 1990.
10. Gil, M.A., Lubiano, M.A., Montenegro, M., and Lopez, M.T. Least squares fitting of an affine function and strength of association for interval-valued data. *Metrika*, 56, pp.97–101, 2002.
11. K*ö*rner, R. and N*ä*ther, W. Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates. *Information Sciences*, 109, pp.95–118, 1998.
12. Manski, C.F. and Tamer, T. Inference on regressions with interval data on a regressor or outcome. *Econometrica*, 70, pp.519–546, 2002.
13. Manski, C.F. Partial Identification of Probability Distributions, New York:Springer-Verlag, 2003.
14. Neto, E.A.L and Carvalho, F.A.T. Centre and range method for fitting a linear regression model to symbolic interval data. *Computational Statistics & Data Analysis*, 52, pp.1500–1515, 2008.
15. Sun, Y. and Li, C. Linear regression for interval-valued data: a new and comprehensive model, under review. arXiv: 1401.1831, 2015.
16. Mukherji, S. The capital asset pricing model's risk-free rate, *International Journal of Business and Finance Research*, 5, pp.793–808, 2011.

APPENDIX C

Optimizing Stock Returns Portfolio Using the Dependence Structure Between Capital Asset Pricing Models : A Vine Copula-based approach

Kittawit Autchariyapanitkul, Sutthiporn Piamsuwannakit
Somsak Chanaim and Songsak Sriboonchitta

This is the original paper that was published in "Causal Inference in Econometrics", Studies in Computational Intelligence Volume 622, pp 319 – 331.

Optimizing Stock Returns Portfolio Using the Dependence Structure Between Capital Asset Pricing Models: A Vine Copula-based approach

K. Autchariyapanitkul*, S. Piamsuwannakit†, S. Chanaim, and ‡ S. Sriboonchitta⋆

Abstract We applied the vine copulas, which can measure the dependence structure of uncertainty in portfolio investments. C-vine and D-vine copulas based on capital asset pricing models were used to exhibit portfolio risk structure in the content of asset allocation. With this approach, we employed the Monte Carlo simulation and the empirical results of C-vine and D-vine copulas to determine the expected shortfall of an optimally weighted portfolio. Furthermore, we used the condition Value-at-Risk (CVaR) model with the assumption of C-vine and D-vine joint distribution to gain the maximum returns in portfolios.

Keywords: CAPM, Vine-Copulas, CVaR, Conditional Value at Risk.

1 Introduction

An important task of financial institutions is evaluating the exposure to market and credit risks. Market risks arise from variations in prices of equities, commodities, exchange rates, and interest rates. Credit risks refer to potential losses that might occur because of a change in the counterparty's credit quality such as a rating migration or a default. The dependence on market or credit risks can be measured by changes in the portfolio value, or gains and losses.

The classical portfolio theory was originally conceived by Markowitz in 1952, the idea that explained the return of the portfolio by mean and variance. Since econo-

*Kittawit Autchariyapanitkul
Faculty of Economics, Maejo University, Chiang Mai, Thailand.
e-mail: Kittawit_a@mju.ac.th,

†Sutthiporn Piamsuwannakit, ⋆Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 52000, Thailand.

‡Somsak Chanaim
Faculty of Science, Chiang Mai University, Chiang Mai 52000, Thailand.

metrics concerns quantitative relations in modern economic life, its analysis consists mainly of determining the impact of a set of variables on some other variable of interest. For example, we wish to determine how return on market X affects return on asset Y in a stock exchange. Now this problem is a regression problem, namely, capital asset pricing model (CAPM). We regress the values of the variable of interest Y, usually called the dependent variable in the explanatory variable X, often called the independent variable. This regression problem is formulated by Sharpe [1] and Lintner [2].

Many pieces of research on the CAPM model is used to explain the diversification of the risk parameter and the performance of portfolios. The investigated issue from Zabarankin et al. [3] purposed drawdown parameter in CAPM model to provide tools for hedging against market drawdowns. Fabozzi and Francis, Levy used CAPM measure risk parameter for a various period. The contributions to the CAMP are the papers of Vassilios [4], Chochola et al. [5], Zhi et al. [6].

A typical risk assessment situation is this. Consider a portfolio consisting of n assets whose possible losses are random variables $X_1, X_2, ..., X_n$. We are interested in the overall risk of the portfolio at some given time, i.e., the total loss $Y = X_1 + X_2 + ... + X_n$. The value-at-risk (VaR) is a commonly used methodology for estimating of risks. The essence of the VaR computations is an estimation of high quantiles (see, Autchariyapanitkul et al. [7]) in the portfolio return distributions. Usually, these computations are based on the assumption of normality of the financial return distribution. However, financial data often reveal that the underlying distribution is not normal. The standard value-at-risk is $F_Y^{-1}(\alpha)$, the maximum possible total loss at level $\alpha \in [0,1]$, i.e.,

$$P(Y > F_Y^{-1}(\alpha)) \leq 1 - \alpha$$

In order to obtain the distribution F_Y of Y, we need the joint distribution of $(X_1, X_2, ..., X_n)$, since, clearly, we cannot assume that the $X_i's$ are mutually independent. A multivariate normal distribution will not work, since empirical work of Mandelbrot and Fama showing that financial variables are rather heavy-tailed. Not only we need copulas to come up with a realistic multivariate model (i.e., a joint distribution for $(X_1, X_2, ..., X_n)$), but we also need copulas to describe quantitatively the dependence among assets.

Vine copulas started with Harry Joe in 1996. He gave a construction of multivariate copulas in terms of bivariate copulas, expressed in terms of distribution functions. Thus, it suffices, besides estimating the marginals, to come up with a high dimensional copula to arrive at a joint distribution for the marginal. In one hand, while lots of parametric bivariate copulas models exist in the literature, there seems not to be the case for higher dimensional copulas. On the contrary, we want a high dimensional copula to capture, say, pairwise dependencies between capital asset pricing models. First, We modeled pairwise dependencies by bivariate copulas and then glue them together to obtain the global high dimensional copula. Zhang et al. [8] used vine copula methods estimate CVAR of the portfolio based on VaR measurement, and showed that D-vine copula model is superior to C-vine and R-vine copulas. Also, to study construct dependence structure, So and Yeung [9] used the time varying vine copulas based GARCH model to show that Kendalls tau and linear correlation of the stock return change over time. Moreover, an enormous number of papers about vine copulas that we can found in a study of Aas et al. [10], Gugan and Maugis [11], Roboredo and Ugolini [12].

In this paper, we intend to use C-vine and D-vine copulas to examine the dependence structure between CAPM models. Then, use the joint distribution that minimize expected shortfall with respect to the expected returns to show the optimal weight of stocks in portfolios. Similarly to the work of Autchariyapanitkul [13] introduced multivariate t-copula to optimize stock returns in portfolio analysis.

This study concentrated on the top 50 largest companies by market capitalization on the Stock Exchange of Thailand ($SET50$). With this method, we used it to measure the risk of a multi-dimensional stock returns in portfolios. Thus, the primary benefaction of this paper can be reviewed as follows: First, we emphasize that the dependence structure is determined by vine copulas and evaluates the complicated nonlinear relations among financial portfolio management. Second, we use the high-dimensional of bull ship stocks show the notable proportion of stocks to the returns of the portfolios. In this studied the selection of the optimal portfolio depends on the underlying assumption on the behavior of the assets under various situations. An unreliable model for dependence structure can cause the damage on portfolios.

The remains of this paper is designed as follows: Section 2 provides a short theoretical framework of copulas, covering C-vine and D-vine copulas. Section 3 conducts the empirical results, and final Section gives the conclusion and extension.

2 Copulas and Vine Copulas

Consider the situation where we know the marginal distributions F and G of the random variables X and Y, respectively (or to be more realistic in term of their estimates). We wish to model and quantify, among other things, the correlation between X and Y. So far, It is all about Sklar's theorem that says: If H is the joint distribution of (X,Y), then there is a copula C such that

$$H(x,y) = C(F(x), G(y))$$

for $(x,y) \in \mathbb{R}^2$.

However, everybody only looks as the "nice" case where both F and G are *continuous*. It is a nice case since the Sklar's theorem becomes:

(i) The copula C is *unique*.

(ii) It can be extracted as

$$C(u,v) = H(F^{-1}(u), G^{-1}(v))$$

(iii) C characterizes dependence structures and dependence measures (with desirable properties). For example,

$$C(u,v) = uv \Longleftrightarrow X \perp Y$$

$$C(u,v) = u \wedge v \Longleftrightarrow F(X) = G(Y)$$

$$C(u,v) = (u+v-1) \vee 0 \Longleftrightarrow F(X) = 1 - G(Y)$$

and dependence measures for (X,Y) can be defined nicely in terms of C (with invariant property).

2.1 Vine Copulas

Suppose, we have a data set on random vector of interest, let say, $X = (X_1, X_2, \cdots, X_d)$. We are focused in making inference about some function of X, e.g., $Y = \varphi(X) = \sum_{i=d}^{d} \alpha_i X_i$ (say, in financial (portfolio) investments), where, e.g., the interest is on deriving the value-at-risk $VaR_\alpha(F_Y)$.

We need the joint distribution H_X of X to determine the distribution F_Y of Y in order to derive

$$VaR_\alpha F(Y) = F_Y^{-1}(\alpha) = inf\{y \in R : F_Y(y) \geq \alpha\} \tag{1}$$

The accurate specification of F_Y is crucial! It comes from the specification of H_X. Now, we have data on X and wish to specify a joint H_X which seems to generate the observed data (a problem of curve fitting). Moreover, since the dependencies among the components $X_i, i = 1, 2, \cdots, d$ are of enormous importance, they should be captured as accurate as possible. Thus, the problem of specifying H_X should take into account, at least, two things in mind: generating the observed data, and modeling pairwise dependencies faithfully.

To accomplish the above program, first recall that, according to Sklar's theorem, we have

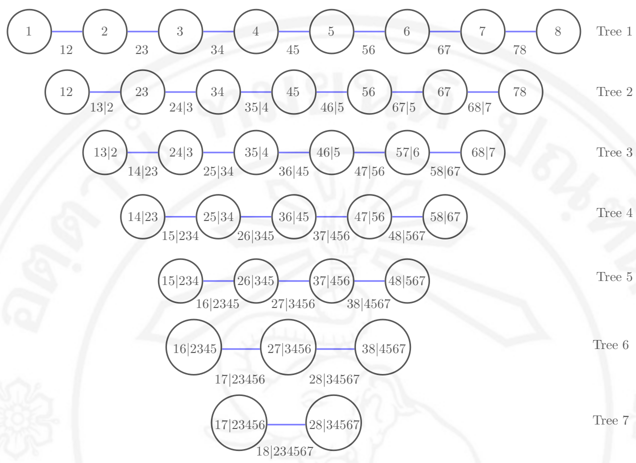
$$H_X(x_1, x_2, \cdots, x_d) = C(F_1(x_1), F_2(x_2), \cdots, F_d(x_d)) \tag{2}$$

2.2 Drawable Vine (D-vine)

The decomposition of the joint density in terms of bivariate (pairwise)copulas and marginals is Drawable and hence is called a D-vine. With this drawable vine copula, the joint density is obtained simply by multiplying all (bivariate) copula densities appeared in the tree together with all marginal densities.

The usefulness of graphical displays is this. When trying to model dependencies in a multivariate model (i.e., we do not know the joint distribution!), we choose a D-vine, according to important pairwise dependencies of interest. We have a "formula" to arrive at the joint distribution, i.e., to come to a model capturing the dependencies of interest. How to use D-vine copulas to build multivariate models? In general, we should figure out that, any d-dimensional copula density can be decomposed in $\frac{d(d-1)}{2}$ different ways. $d = 8, X = (X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)$ a possible D-vine is

Fig. 1 Dvine



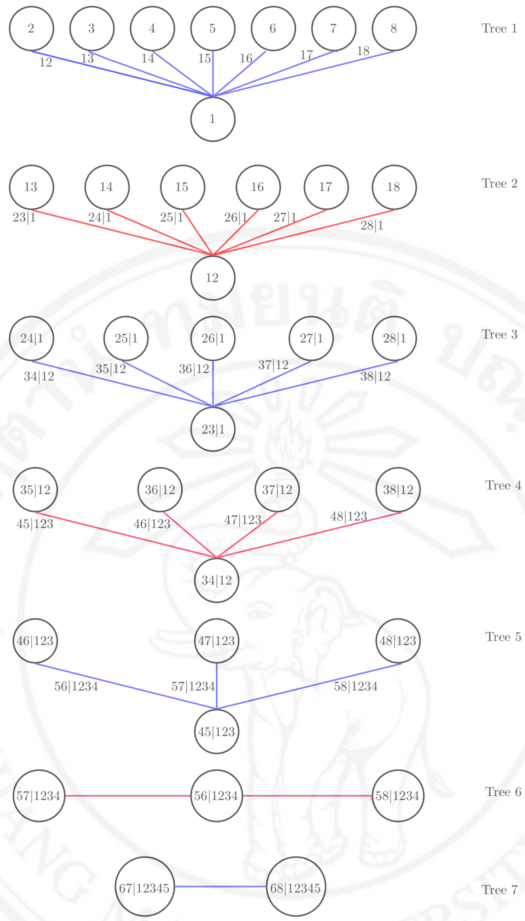
resulting in the multivariate (density) model

$$
\begin{aligned}
f(x_1, x_2, \cdots, x_8) = \prod_{i=1}^{8} f_i(x_i) \cdot & c_{12} c_{23} c_{34} c_{45} c_{56} c_{67} c_{78} \\
& \cdot c_{13|2} c_{24|3} c_{35|4} c_{46|5} c_{57|6} c_{68|7} \\
& \cdot c_{14|23} c_{25|34} c_{36|45} c_{47|56} c_{58|67} \\
& \cdot c_{15|234} c_{26|345} c_{37|456} c_{48|567} \\
& \cdot c_{16|2345} c_{27|3456} c_{38|4567} \\
& \cdot c_{17|23456} c_{28|34567} \\
& \cdot c_{18|234567}
\end{aligned}
\tag{3}
$$

2.3 *Canonical Vine (C-vine)*

A C-vine is a regular vine such that each tree T_j has a unique node of degree d-j. The node with maximal degree in T_1 is the root, for eight-dimension (d=8) C-vine copulas can written as

Fig. 2 Cvine



The decomposition of joint densities in terms of C-vines copulas is illustrated as follows

$$
\begin{aligned}
f(x_1,x_2,\cdots,x_8) = \prod_{i=1}^{8} f_i(x_i) & \cdot c_{12}(F_1,F_2)\cdot c_{13}(F_1,F_3)\cdot c_{14}(F_1,F_4)\cdot c_{15}(F_1,F_5) \\
& \cdot c_{16}(F_1,F_6)\cdot c_{17}(F_1,F_7)\cdot c_{18}(F_1,F_8)\cdot c_{23|1}(F_{2|1},F_{3|1}) \\
& \cdot c_{24|1}(F_{2|1},F_{4|1})\cdot c_{25|1}(F_{2|1}\cdot F_{5|1})\cdot c_{26|1}(F_{2|1}\cdot F_{6|1}) \\
& \cdot c_{27|1}(F_{2|1}\cdot F_{7|7})\cdot c_{28|1}(F_{2|1}\cdot F_{8|1})\cdot c_{34|12}(F_{3|12},F_{4|12}) \\
& \cdot c_{35|12}(F_{3|12},F_{5|12})\cdot c_{36|12}(F_{3|12},F_{6|12})\cdot c_{37|12}(F_{3|12},F_{7|12}) \\
& \cdot c_{38|12}(F_{3|12},F_{8|12})\cdot c_{45|123}(F_{4|123},F_{5|123})\cdot c_{46|123}(F_{4|123},F_{6|123}) \\
& \cdot c_{47|123}(F_{4|123},F_{7|123})\cdot c_{48|123}(F_{4|123},F_{8|123})\cdot c_{56|1234}(F_{5|1234},F_{6|1234}) \\
& \cdot c_{57|1234}(F_{5|1234},F_{7|1234})\cdot c_{58|1234}(F_{5|1234},F_{8|1234}) \\
& \cdot c_{67|12345}(F_{6|12345},F_{7|12345})\cdot c_{68|12345}(F_{6|12345},F_{8|12345}) \\
& \cdot c_{78|123456}(F_7,F_8)
\end{aligned} \tag{4}
$$

3 An Application and Empirical Results

3.1 Capital Asset Pricing Model:CAPM

The Capital Asset Pricing Model (CAPM) was formerly conceived by William F. Sharpe [1] and John Lintner [2]. The CAPM is the linear combination of the expected excess return on asset and expected market returns. A linear function of CAPM model can be addressed as follows:

$$E(R_A) - R_F = \beta_0 + \beta_1 E(R_M - R_F), \tag{5}$$

where $E(R_A)$ and R_M describe the expected return on stock and the expected market returns, sequentially, β_0 show the intercept and R_F is the risk-free rate. $E(R_M - R_F)$ is the expected risk premium, and β_1 is the risk parameter. We can calculate the systematic risk of each stock by this mathematical statement

$$\beta_1 = \frac{cov(R_A, R_M)}{\sigma_M^2}, \tag{6}$$

where σ_M^2 is the variance of the expected market returns. Given CAPM equation of each stock returns, we can calculate the joint dependency structure via C-vine and D-vine to carry out the optimization process.

3.2 Optimal Portfolio with Conditional Value at Risk via Vine-Copulas

We start our calculation of VaR and CVaR of an equally weighted portfolio and then, the optimal portfolio can be constructed by minimizing CVAR subject to maximum returns. The procedure of optimization, we refer to the paper from Autchariyapanitkul [13]. The following formula can show as below:

$$\text{Min } CVAR = E[r_p | r \leq r_\alpha], \tag{7a}$$

$$\text{subject to } E(r_p) = w_1 E(r_1) + w_2 E(r_2) + \cdots + w_n E(r_n), \tag{7b}$$

$$w_1 + w_2 + \cdots + w_n = 1, \tag{7c}$$

$$0 \leq w_i \leq 1, \; where \; i = 1, 2, \cdots, n,$$

where r_α is the lower $\alpha - quantile$, and r_p is the return on individual asset at time t.

We use vine copulas to extract dependence structure between CAPM equations and then use the solutions of C-vine and D-vine copulas parameters to create an efficient portfolio and find the optimal solutions for the expected returns with minimum lost.

Now, we simulate the error terms of each stock form the CAPM equations by using the estimated vine-copulas to generate a set of 1,000,000 samples. Then, we obtained a possible price of each stock under CAPM models and vine-copulas to optimization problem.

3.3 Data

The data contains 260 weekly returns during 2010-2014 are retrieved from DataStream, we calculate the log returns on the tracking stocks. The data consist of the returns from the 8 big capitalization companies such as Banpu Public Company Limited (BANPU), Bank of Ayudhya Public Company Limited (BAY), Bangkok Bank Public Company Limited(BBL), Central Pattana Public Company Limited (CPN), Land and Houses Public Company Limited (LH), Pruksa Real Estate Public Company Limited (PS), Thanachart Capital Public Company Limited (TCAP) and Thai Oil Public Company Limited (TOP). Table 1 supplies a summary of the variables.

Table 1 Summary statistics

	SET50	BANPU	BAY	BBL	CPN	LH	PS	TCAP	TOP
Mean	0.0025	-0.0036	0.0028	0.0020	0.0055	0.0015	0.0020	0.0016	-0.0002
Median	0.0041	-0.0051	0	0.0025	0.0048	0.0000	0.0000	0.0000	0.0000
Max.	0.0706	0.1802	0.1341	0.1002	0.1268	0.1638	0.1650	0.1475	0.1377
Min.	-0.0766	-0.1324	-0.1658	-0.1039	-0.1406	-0.1947	-0.1926	-0.1581	-0.2173
SD.	0.0253	0.0422	0.0425	0.0344	0.0426	0.0503	0.0567	0.0382	0.0413
Skew.	-0.3412	0.1823	-0.1674	0.1789	-0.1494	0.1803	-0.3417	-0.2854	-0.3675
Kurt.	3.8347	4.6789	4.2171	3.2542	3.6822	4.2344	3.8459	4.6276	6.2557
J.B.	12.5927	31.9757	17.2625	2.0867	6.0089	17.9156	12.8106	32.2261	120.6802
PROB.	0.0080	0.0010	0.0034	0.3097	0.0454	0.0030	0.0077	0.0010	0.0010

All values are the log return.

3.4 Experimental results

Given equations from (5) and (6), we can estimate parameters of CAPM models as the following

Table 2 Parameters estimation from CAPM models

	BANPU	BAY	BBL	CPN	LH	PS	TCAP	TOP
β_0	–0.0060	0.0003	-0.0005	0.0030	–0.0014	–0.0009	–0.0008	–0.0028
	(0.0021)	(0.0021)	(0.0014)	(0.0022)	(0.0024)	(0.0029)	(0.0018)	(0.0020)
β_1	0.9653	1.0098	1.0248	0.9695	1.2592	1.2617	0.9556	1.0543
	(0.0846)	(0.0835)	(0.0554)	(0.0855)	(0.0956)	(0.1153)	(0.728)	(0.0773)
σ^2	0.0012	0.0012	0.00005	0.0012	0.0015	0.0022	0.0009	0.0010
R^2	0.3350	0.3620	0.5700	0.3320	0.4020	0.3170	0.4000	0.4190
KS test	0.0811	0.7856	0.4211	0.8055	0.4854	0.6835	0.4326	0.0678

Table 3 and Table 4 show the estimation results for C-vine and D-vine copulas, respectively.

Table 3 Estimated Results of C-vine copula

Pairs	Families	Parameter1	Parameter2	AIC
*1,2	Frank	-1.1185 (0.4130)	-	-5.2231
1,3	Gumbel	1.0630 (0.0370)	-	-9.6632
1,5	Frank	-1.2610 (0.4220)	-	-6.7918
2,4|1	Clayton	0.2010 (0.0697)	-	-10.3094
2,5|1	Clayton	0.2151 (0.0784)	-	-8.1258
3,4|1,2	Rotated BB8	-1.2777 (0.2005)	-0.9540 (0.0753)	-4.8876
3,5|1,2	Gaussian	-0.1833 (0.0595)	-	-6.7201
3,7|1,2	Gaussian	-0.1952 (0.0593)	-	-7.8331
4,5|1,2,3	Rotated Gumbel	1.1671 (0.0527)	-	-19.7762
4,6|1,2,3	Rotated BB8	1.2728 (0.1545)	0.9608 (0.0572)	-4.9537
4,7|1,2,3	Frank	0.8335 (0.4005)	-	-2.3229
7,8|1,2,3,4,5,6	Frank	0.7930 (0.4031)	-	-1.8701

() standard error is in parenthesis, 5% level of significant. *1=BANPU, 2=CPN, 3=TOP, 4=PS, 5=LH, 6=TCAP, 7=BBL, 8=BAY.

Table 4 Estimated Results of D-vine copula

Pairs	Families	Parameter1	Parameter2	AIC
*1,2	Gaussian	-0.1928 (0.0586)	-	-7.8570
2,3	Frank	0.8544 (0.4129)	-	-2.2602
4,5	Frank	-1.1185 (0.4133)	-	-5.2231
6,7	Clayton	0.1536 (0.0737)	-	-3.9440
1,3|2	Rotated BB8	-1.3060 (0.2259)	-0.9489 (0.0833)	-5.9783
3,5|4	Clayton	0.2010 (0.0696)	-	-10.3093
5,7|6	Survival BB8	1.4622 (0.2565)	0.9138 (0.0972)	-9.5663
1,4|2,3	Gumbel	1.05627 (0.0349)	-	-9.0128
3,6|4,5	Survival BB8	1.2374 (0.1098)	0.9864 (0.0202)	-7.9990
4,7|5,6	Frank	-0.9626 (0.4092)	-	-3.4873
3,7|4,5,6	Survival Gumbel	1.1833 (0.0530)	-	-20.7914
1,7|2,3,4,5,6	Rotated Clayton	-0.1281 (0.0648)	-	-3.4141
2,8|3,4,5,6,7	Frank	0.8412633 (0.4084)	-	-2.2348

() standard error is in parenthesis, 5% level of significant. *1=TOP, 2=BBL, 3=PS, 4=BANPU, 5=CPN, 6=TCAP, 7=LH, 8=BAY.

Given a market return $R_M = 0.01$ and a risk free rate $R_F = 0$, we considered all possible ordered of vine-copulas with the lowest AIC. Note that, this method does not guaranteed the best ordered of vine-copulas but in this paper we only have one set of vine-copulas with the lowest AIC. In general, it is possible to have many set of vine-copulas with the same minimum AIC values. Then, we compare the AIC values of the C-vine and D-vine copulas, we found that the D-vine copula structure gives a better results. We can use values of the D-vine copula to estimate the CVAR and efficient portfolio with the maximum expected return for a minimum loss.

Table 5 shows the expected returns of VaR and CVaR at levels of 1%, 5% and 10% with an equally weighted stock. We notice that the estimated CVaR converges to -1.4289, -1.8687 and -2.7599 at 10%, 5% and 1% levels in period $t+1$, respectively.

We applied the Monte Carlo simulation to produce a set of 1,000,000 samples. Then, provided a significant level of 5%, we optimized the portfolio by employing the mean-CVaR model and received the efficient frontier of the portfolio under different expected returns, as displayed in Fig. 3.

Table 5 Expected shortfall of equally weighted portfolios

	Expected Returns	VaR	CVaR
10%	0.9405	-0.7537	-1.4289
5 %	0.9405	-1.2657	-1.8687
1 %	0.9405	-2.2458	-2.7599

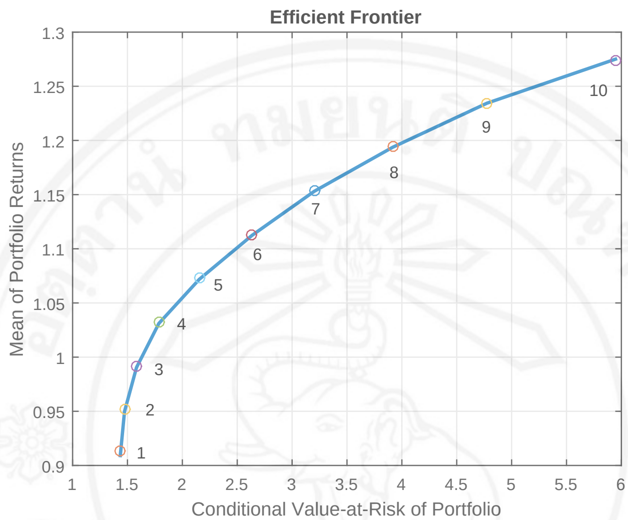


Fig. 3 : The efficient frontiers of CVaR under mean

Eventually, we also obtained the optimal weight of the portfolios varies to the CVAR. Table 6 exposes some of the results of optimal weight with the expected returns in the frontier.

Table 6 Optimal weighted portfolios for CVAR 5 %

Portfolios	$BANPU_{w_1}$	BAY_{w_2}	BBL_{w_3}	CPN_{w_4}	LH_{w_5}	PS_{w_6}	$TCAP_{w_7}$	TOP_{w_8}	Returns
1	0.1086	0.1024	0.2725	0.1179	0.0759	0.0154	0.1229	0.1845	0.9129
2	0.0604	0.1152	0.2866	0.1415	0.0715	0.0246	0.1187	0.1814	0.9522
3	0.0130	0.1259	0.2979	0.1662	0.0695	0.0339	0.1153	0.1782	0.9916
4	0.0000	0.1449	0.2857	0.2270	0.0695	0.0467	0.0873	0.1391	1.0320
5	0.0000	0.1676	0.2652	0.3020	0.0667	0.0618	0.0500	0.0867	1.0727
6	0.0000	0.1872	0.2512	0.3798	0.0626	0.0729	0.0143	0.0320	1.1133
7	0.0000	0.1985	0.1801	0.4810	0.0530	0.0874	0.0000	0.0000	1.1539
8	0.0000	0.1961	0.0530	0.6114	0.0329	0.1065	0.0000	0.0000	1.1942
9	0.0000	0.1164	0.0000	0.7714	0.0000	0.1123	0.0000	0.0000	1.2340
10	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.2736

4 Concluding Remarks

In this paper, we have determined the risk in portfolio management by employing CVaR and used the mean-CVaR model to optimize portfolios. We used the C-vine and D-vine copula to measured dependence structure between capital asset pricing model (CAPM) affects the returns of portfolios. We carried our analysis in two steps. First, we examined the dependence structure of stock returns obtained from CAPM equations. Second, we investigated how the dependence structure of the asset pricing model influences portfolio optimization. We used an optimization procedure to allocate risk in the portfolios. It is feasible to reason that vine copulas can be explained dependency structure of the asset in the portfolio management.

Acknowledgements The authors thank Prof. Dr. Hung T. Nguyen for his helpful comments and suggestions.

References

1. Sharp, William F. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, Vol.19(3), pp. 425-442, 1964.
2. Lintner, J. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47(1), pp.12-37, 1965.
3. Zabarankin, Michael, Konstantin Pavlikov, and Stan Uryasev. Capital asset pricing model (CAPM) with drawdown measure. *European Journal of Operational Research*, 234(2), 508-517, 2014.
4. Papavassiliou, Vassilios G. A new method for estimating liquidity risk: Insights from a liquidity-adjusted CAPM framework. *Journal of International Financial Markets, Institutions and Money*, Vol.24, pp. 184-197, 2013.
5. Chochola, O., Huskova, M., Praskova, Z., & Steinebach, J. G. Robust monitoring of CAPM portfolio betas II. *Journal of Multivariate Analysis* Vol.132, pp.58-81, 2014.
6. Da, Zhi, Re-Jin Guo, and Ravi Jagannathan. CAPM for estimating the cost of equity capital: Interpreting the empirical evidence. *Journal of Financial Economics*, 103(1), pp. 204-220, 2012.
7. Autchariyapanitkul, K., Chainam, S. and Sriboonchitta, S. Quantile regression under asymmetric Laplace distribution in capital asset pricing model, *Econometrics of Risk* , Springer International Publishing, Vol. 583, pp. 219-231, 2015.
8. Zhang, Bangzheng, et al. Forecasting VaR and CVAR of stock index portfolio: A Vine copula method. *Physica A: Statistical Mechanics and its Applications*, 416, 112-124, 2014.
9. So, Mike KP, and Cherry YT Yeung. Vine-copula GARCH model with dynamic conditional dependence. *Computational Statistics & Data Analysis*, 76, 655-671, 2014.
10. Aas, K., Czado, C., Frigessi, A., & Bakken, H. Pair-copula constructions of multiple dependence. Insurance: Mathematics and economics, 44(2), 182-198, 2009.
11. Guegan, Dominique, and Maugis, P.A. An econometric study of vine copulas. *International Journal of Economics and Finance*. 2(5), 2-14, 2011.
12. Reboredo, Juan C., and Ugolini, A. A vine-copula conditional value-at-risk approach to systemic sovereign debt risk for the financial sector. *The North American Journal of Economics and Finance*, 32, 98-123, 2015.
13. Autchariyapanitkul, K., Chainam, S. and Sriboonchitta, S. Portfolio optimization of stock returns in high-dimensions: A copula-based approach. *Thai Journal of Mathematics*, Special Issue on: Copula Mathematics and Econometrics, 11-23, 2014.

CURRICULUM VITAE

Author's Name Miss Sutthiporn Piamsuwannakit

Date of Birth 25th July 1969

Education 2003 Master of Economics, Chiang Mai University, Chiang Mai, Thailand

1991 Bachelor of Science (Agricultures), Chiang Mai University, Chiang Mai, Thailand

Scholarships Funded by Chiang Rai Rajabhat University-PhD Scholarship

Experience Lecturer, Department of Business Economics, Faculty of Management of Sciences, Chiang Rai Ratjabhat University, Chiang Rai, Thailand