Frontier of error minimization from copula model application: evidence from dependence structure of BRICS’s stock markets

Jittima Singvejsakul1,2, Chukiat Chaiboonsri2 and Songsak Sriboonchitta1,2

1Puey Ungphakorn Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Chiang Mai 50200, Thailand
2Faculty of Economics, Chiang Mai University, Chiang Mai 50200, Thailand
3Email: Jittimasvsk@gmail.com

Abstract. This study is proposed to focus on the comparison between Maximum Likelihood estimation (MLE) and Maximum Entropy bootstrap testing (MEboot) in AR-GARCH model in order to seek the frontier of error minimization or the minimum error terms by employing a cross-entropy selection approach. Empirically here, this paper finds that the error terms estimated by the MEboot are smaller than those from the MLE. Therefore, the MEboot estimator is proposed to estimate time series data like financial data because it is more robust statistically than MLE, thus MEboot can be resistant to errors in the results. Consequently, the MEboot test’s residuals are selected for estimating mathematically an econometric model called D-vine copula which is based on Gaussian distribution to investigate the pattern of dependence structure in BRICS countries. The results showed that Brazil was the major stock market in BRICS that investors will be interested in its financial flows. Accordingly, we had better understand the central point of capital flows inside BRICS’s financial systems. In addition, the dependence structure results show that most financial flows into Brazil come from the Asian continent (India and China). On the other hand, capital out-flows from Brazil have been destined to Russia and South Africa.

1. Introduction

The stock markets play an important role in the economy of many countries for facilitating capital flows from a country to other countries, and thus are especially important for the world’s emerging markets. Interestingly, amidst many emerging markets around the world, one of the important ones is the five largest countries known as the BRICS, comprising Brazil, Russia, India, China and South Africa. This paper aims to study about the dependence structure among these emerging markets which will be useful for investors to plan their investment by tracking how the capital will flow among these countries. In order to find the dependency of these stock markets we need good estimates of the residuals that can capture realistically the structure of dependence. This study also wants to focus on using time series data of these stock markets to establish a decade long relation that a normal non-linear estimation cannot precisely imply. As the data shown in figure 1, non-linear regression resulted in an uninterpretable relations of stock market price indexes. From this issue, a robust estimator is important to explain the distribution and relation of time-series data from stock markets.

Consequently, the main objectives of this research are divided into 2 sections. First, we employ two estimators; Maximum likelihood and Maximum entropy bootstrap to get the frontiers of minimum
error for comparing their robustness statistically. Then we can apply the residuals from the estimator which can give the minimum error to the copula model. Effectively, the structural pattern in the BRICS’s stock markets is investigated using cross entropy to compare the error terms of the Maximum Likelihood estimation (MLE) and Maximum Entropy Bootstrap testing (MEboot). In addition, the data considered in this study is the stock markets indices such as Brazil (BOVESPA Index), Russia (MICEX Index), India (BSE Index), China (SSE composite index) and South Africa (JSE Index) by transforming into standardized residuals of daily log-return data, which has been collected from 2006 to 2016 (2,643 observations).

\[ \text{Figure 1. Daily time series data of BRICS stock markets price indexes from the period of 2006-2016.} \]

2. Literature review

Historically, there have been many studies undertaken using the GARCH or AR model normally based on maximum likelihood testing. For instances, Alizadehet et al. (1999) [1] focuses on volatilities obtained from a price range which is close to the Gaussian properties. Ferenstein and Gasowski (2004) [2] uses the AR-GARCH model to analyze two data sets from stock prices that is able to generalize error distributions. This can be implied as a good model of white noise distributions.

Regarding, the maximum entropy bootstrap estimation, Bartiromo (2013) [3] studied the maximum entropy approach to obtain the distribution of fluctuated stock prices by maximizing information entropies. This can be used to extract expected volatilities from stock options. Ormos and Zibriczky (2014) [4] uses the concept of entropy to explain the equity premium of securities and portfolios, and finds that entropy has a higher explanatory power for the expected return than the capital asset pricing model beta. Fiedor (2015) [5] proposes the maximum entropy production principle which uses the entropy rate to create a general principle underlying the price formation processes. It shows that the predictability of price change is higher at the transaction level than the scale of daily returns. In the studies of applying bootstrapping processes, Ruiz and Pascual (2002) [6] reviews the application of bootstrap procedures for inferencing and predicting financial time-series variables based on bootstrap techniques. Fong and Koh (2011) [7] uses the bootstrap approach to generate returns of U.S. stocks and treasury bills during time horizons from one year to twenty years.

Considering a cross-entropy approach, Branger (2004) [8] uses the minimum cross entropy measurement to choose a stochastic discount factor (SDF) given a benchmark SDF for determining the Arrow-Debreu (AD) prices. Additionally, Zhou et al. (2013) [9] reviews the concepts and principles of the cross-entropy method, especially in portfolio selection and asset pricing. Furthermore, they review the effects of the applications of entropy and compare them with the effects using other traditional and new methods.

The final part of the paper is to employ copula models. From the literature survey, Allen et al. (2012) [10] applies regular vine copula to analyze the co-dependencies of 10 major European stock
markets and explore how correlations change in different economic circumstances using three different sample periods. Gruber and Czado (2015) [11] points out that regular vine copulas are novel and very flexible class of dependence models and discusses the general outlines of stepwise tree-by-tree strategies and simultaneous methods for model selection.

Therefore, this paper is the first ever able to employ all of the methods from the literature to find the estimator which is robust statistically in terms of financial data. We use Maximum likelihood estimator (MLE) for comparison with Maximum entropy bootstrap (MEboot) to estimate AR-GARCH and after that we use cross entropy to demonstrate the minimum of the error terms. That means the estimator showing the minimum error term is the robust one.

3. Methodology

3.1. AR-GARCH

AR-GARCH is used to find the residuals term of daily log-return data from BRICS’s stock markets. That consists of AR (Autoregressive) and GARCH (Generalized Autoregressive Conditional Heteroscedasticity). The autoregressive process of order $p$ or AR ($p$) is defined by the equation (2).

$$Y_t = \sum_{j=1}^{p} \phi_j Y_{t-j} + \epsilon_t,$$

where $Y_t$ = the daily log-return data at time $t$

$Y_{t-j}$ = the daily log-return data at time $t-1$

$\epsilon_t \sim N(0, \sigma^2)$

$\phi = (\phi_1, \phi_2, \ldots, \phi_p)$ that is the vector of model coefficient and $p$ is a Non-negative integer.

The AR model establishes that a realization at time $t$ is a linear combination of the $p$ previous realization plus some error term. For GARCH model is an extension of the AR model that the residual term are conditional heteroskedastic by Bollerslev (1986) [12], which is written as

$$h_t = \omega + \sum_{i=1}^{p} b_i h_{t-i} + \sum_{j=1}^{q} \alpha_j \epsilon_{t-j}^2.$$

3.2. The maximum likelihood and maximum entropy bootstrap (MEboot)

We employ 2 estimators for investigating the residuals distribution of AR-GARCH. First, the Maximum likelihood is the method for estimating the distribution that the principle of maximum likelihood yields a choice of the estimator $\hat{\theta}$ as the value for the parameter that makes the observed data most probable by Harris and Stocker (1998) [13].

Second, Maximum entropy is a powerful tool by avoiding all unnecessary distributional assumptions by Vinod (2013) [14]. This computational estimation follows a given finite state space and constraints, which are finite sets of $n$ states with probabilities $q$ and they were defined as a prior estimation form with new information. Accordingly, the entropy maximization is expressed as

$$H(q) = \sum q_i \log(q_i) - \log(n),$$

The bootstrap is a computationally statistical approach regarding the relation between samples and unknown populations by forming a comparable linkage between the samples at hand and appropriately designed observable resamples by Vinod (2013) [14]. Technically, let us compute the bootstrap percentile of the function $y = f(q_i)$, where $q_i$ is the motif statistic: $q \rightarrow \mathbb{R}$. The samples, $q_i, \ldots, q_n$, are assumed to be an identically independent distribution (i.i.d) and the function, $P(q' \mid I)$, stands for the parametric density of choices by O’Neil and Erill (2016) [15]. Hence, we can set $y = f(q_i)$ for each $I$, the bootstrap percentile is given by
Interestingly, the Maximum Entropy and Bootstrapping approaches combined are the powerful technique of modern time series inferences and the algorithms estimated by the Maximum Entropy Bootstrap approach (MEboot) that can generate an ensemble of worldwide changing time by Chaiboonsri and Chaitip (2013) [16]. The overview of the steps in Vinod’s ME bootstrap algorithm that applied seven steps was found and described in the study of Chaitip and Chaiboonsri (2013) [17].

3.3. The conceptual of cross entropy approach

Fundamentally, the cross entropy (CE) is the mutual support degree that can be used to determine the weights of the information sources, where a larger weight represents higher mutual supports by Men et.al (2016) [18]. The cross entropy of two likelihood distributions was expressed as

\[ D(g \parallel f) = \sum_{i}^{n} g_i \ln \left( \frac{g_i}{f_i} \right), \]

and continuous cases was

\[ D(g \parallel f) = \int g(x) \ln \frac{g(x)}{f(x)} dx = \int g(x) \ln g(x) dx - \int g(x) \ln f(x) dx, \]

where \( f \) and \( g \) indicate the probability vector in discrete cases and the probability density function in the continuous case, respectively.

3.4. The vine pairs-copulas construction

The copula is a multivariate distribution function defined on the unit cube \([0, 1]^n\) with uniformly distributed marginal. This definition is very natural if one considers how a copula is derived from a continuous multivariate distribution function by Nelsen (2006) [19].

A principle for constructing multivariate copula from the product of bivariate pair copula with description of the statistical inference techniques for the specific vines which are called (D-) vines was provided by Nikoloulopoulos and Joe (2012) [20]. This constitutes very flexible models, since bivariate copulas can easily accommodate complex dependence structures such as asymmetric dependence or strong joint tail behavior. This can also be written as

\[ f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2), \]

where \( C \) is the copula associated with \( F \) via Sklar’s Theorem. From equation (7) the conditional density of \( X_2 \) given \( X_1 \) can be determined as

\[ f_{2|1}(x_2 | x_1) = \frac{f(x_1, x_2)}{f_1(x_1)} = c_{12}(F_1(x_1), F_2(x_2))f_2(x_2). \]

This formula in its general form

\[ f_{j|i}(x_j | x_i) = \frac{f(x_i, x_j)}{f_i(x_i)} = c_{ij}(F_i(x_i), F_j(x_j))f_j(x_j). \]

Therefore, this is the dimensional joint which can be shown in terms of multivariate copula.

4. Conceptual framework

Figure 2 provides the conceptual estimation steps of research processes. There are 5 steps for estimating the results.
Figure 2. The conceptual estimation steps of research processes.

5. Data description and empirical results

5.1. The data description

Figure 3 provides the index return of daily data in BRICS during the period of 2006 to 2016. Table 1 presents the descriptive statistics of the stock return from data collected on a log daily data in BRICS’s stock markets over the course of the specified period.

Figure 3. The log return of daily data in BRICS during period of 2006 to 2016.
**Table 1.** Descriptive statistics of the stock return of daily data in BRICS during period of 2006 to 2016.

| Variable | Obs | Mean  | Min  | Max  | Std.dev | Dicky-Fuller Unit root test |
|----------|-----|-------|------|------|---------|-----------------------------|
| Brazil   | 2643| 0.0004| -0.1209 | 0.1367 | 0.0180 | -50.9675 (0.0000) |
| Russia   | 2643| 0.0002| -0.5397 | 0.6540 | 0.0343 | -25.4991 (0.0000) |
| India    | 2643| 0.0006| -0.1180 | 0.1599 | 0.0161 | -44.7465 (0.0000) |
| China    | 2643| 9.5E-05| -0.0925 | 0.0903 | 0.0164 | -51.0775 (0.0000) |
| South Africa | 2643| 0.0007| -0.0105 | 0.1397 | 0.0180 | -41.0929 (0.0000) |

Source: Computation

5.2. The estimated results of the cross entropy selecting model.

The results in table 2 were estimated based on the Cross entropy which showed the error terms from the AR-GARCH model with MLE and the AR-GARCH model with the MEboot. This presents that the overall entropy value of residuals terms calculated from MEboot testing was smaller than the counterpart from the MLE estimation, which were 397.8729 and 396.8297, respectively. The details provided in table 2 showed that cross-entropy values estimated by the MLE were 397.8206, 397.8181, 397.7879, 398.1679 and 397.7929. These stand for South Africa, Brazil, India, Russia and China, respectively. In addition, the cross-entropy values estimated by the MEboot of South Africa, Brazil, India, Russia and China were 396.8042, 396.8452, 396.9863, 397.2071 and 396.8565, consecutively. From the results, all of error terms in MEboot are lower than those from MLE. This can therefore be used to imply that the residuals estimated by the MEboot are more precise than those from the MLE test.

**Table 2.** The estimated results of the cross entropy selecting model in the BRICS’s countries.

| | AR-GARCH (MLE) | AR-GARCH (MEBoot) |
| | (Cross-Entropy) | (Cross-Entropy) |
| All entropy | 397.8729 | 396.8297 |
| South Africa | 397.8206 | 396.8042 |
| Brazil | 397.8181 | 396.8452 |
| India | 397.7879 | 396.9863 |
| Russia | 398.1679 | 397.2071 |
| China | 397.7929 | 396.8565 |

Source: Computation

5.3. The results of estimation of D-vine copula tree based on Maximum Entropy bootstrap

For investigating the dependence structure correlation in financial economics, the residuals from the AR-GARCH model testing by the MEboot estimator was used to graphically estimate D-vine copula tree. Technically, D-vine trees based on Gaussian copula were employed to clarify the major markets in stock markets of BRICS countries. The relations were represented in figure 4. Considering deeply in the estimated D-vine tree, there are four dependence structures which explain the relation among BRICS’ stock markets. These results show that most capital inflows were directed from India (BSE Index) and China (SSE composite index) to Brazil (BOVESPA Index). This can be used to explain that when the stock market systems in Brazil is stimulated, this will induce a cross-country fund from India and China. Conversely, there are capital outflows, which is the transfer from Brazil (BOVESPA Index) to Russia (MICEX Index) and South Africa (JSE Index). The details are graphically shown in figure 4.
Figure 4. The estimation results of structural relation among stock markets of BRICS from D-Vine trees in 2006-2016.

6. Conclusions
The study of comparable estimated error terms of time-series data in the financial variables of BRICS countries are successfully clarified in this paper. Empirically, the residuals term estimated by the MEboot testing contain the minimum error terms in relation to the MLE by using the cross entropy-value calculations. That means the Maximum Entropy bootstrap (MEboot) is the optimal estimator for the financial data or time series because of its lower minimum error terms. Therefore, this estimator is considered as a robust statistic which can be resistant to errors occurring in the results, and it is not affected by the outlier and thus provides good statistical estimation for financial data. Consequently, this implies that the dependence structure of D-Vine tree is better to estimate by using MEboot technique more than MLE, and potentially provides a precise parameter, especially of the financial data.

For investigating the directions of capital flows in BRICS countries, this dependence structure pattern is successfully expressed by the D-vine copula model based on MEboot. The results are graphically presented in figure 4 and indicate that most of the financial flows transferring into Brazil have come from the Asian continent (India and China). On the other hand, the capital outflows from Brazil have been destined to Russia and South Africa. Consequently, from the empirical computational results, we are now able to better verify the financial flows in BRICS countries. This can strongly confirm that applied mathematics in econometrics has not only been extremely important for seeking precise forecasting in financial areas, but this analysis can be efficiently adapted to predictive studies in many academic subjects.

References
[1] Alizadeh S, Brandt M and Diebold F 1999 Range-based estimation of stochastic volatility Models Journal of Finance 57 1047-1092
[2] Ferenstein E and Gasowski M 2004 Modelling stock returns with AR-GARCH processes Statistics and Operations Research Transactions 28 55-68
[3] Bartiromo R 2013 Maximum entropy distribution of stock price fluctuations Statistical finance Cornell university 4 1106-1134
[4] Ormos M and Zibriczky D 2014 Entropy-Based Financial Asset Pricing Journal pone 742
[5] Fiedor P 2015 Maximum Entropy Production Principle for Stock Returns Quantitative finance 11-14
[6] Ruiz E and Pascual L 2002 Bootstrapping Financial Time Series Journal of economic surveys 16 271-300
[7] Fong W and Koh T 2011 Strategic asset allocation with low-risk stocks: a bootstrap analysis Journal of investment management 13 39-58
[8] Branger N 2004 Pricing derivative securities using cross-entropy: an economic analysis
   *International journal of theoretical and applied finance* 7(1) 63-81

[9] Zhou R, Cai R and Tong G 2013 *Applications of Entropy in Finance* (School of Economics and
   Management, Beijing University of Chemical) 15(11) 4909-493

[10] Allen D, McAleer M and Singh A 2012 Risk measurement and risk modelling using application
    of vine copulas *Journal of economics and finance* 47 25-49

[11] Gruber F and Czado C 2015 Bayesian model selection of regular vine copulas (Springer Verlag)

[12] Bollerslver T 1986 Generalized Autoregressive Conditional Heteroskedasticity *Journal of
    Economics* 31 307-327

[13] Harris J W and Stocker H Maximum Likelihood Method *Handbook of Mathematics and
    Computational Science* (New York: Springer-Verlag) 824

[14] Vinod H D 2013 Maximum entropy bootstrap simulations for variance estimation Available at
    SSRN: https://ssrn.com/abstract=2295723

[15] O’Neil P K and Erill I 2016 Parametric bootstrapping for biological sequence motifs *BMC
    Bioinformatics* 17(406)

[16] Chaiboonsri C and Chaitip P 2013 A boundary analysis of ICT firms on Thailand stock market:
    a maximum entropy bootstrap approach and Highest Density Regions (HDR) approach *Int. J.
    Computational Economics and Econometrics* 314-26

[17] Chaitip P and Chaiboonsri C 2013 AEC’s demand for ICT: maximum entropy bootstrap
    approach in panel data *Procedia Economics and Finance* 5 125-132

[18] Men B, Long R and Zhang J 2016 Combined forecasting of streamflow based on cross entropy
    *Entropy* 18(336)

[19] Nelsen R 2006 An Introduction to Copulas( New York: Springer)

[20] Nikoloulopoulos A and Joe H 2012 Vine Copulas with Asymmetric Tail Dependence and
    Applications to Financial Return Data *Computational Statistics & Data Analysis* 56(11)
    3659-3673