Reverse stress testing on non-elliptical jointly distributed multivariate data

Mr. Chevincee Werawanich

An Independent Study Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Financial Engineering
Department of Banking and Finance
FACULTY OF COMMERCE AND ACCOUNTANCY
Chulalongkorn University
Academic Year 2020

การทดสอบภาวะวิกฤตแบบย้อนกลับสำหรับข้อมูลที่มีการแจกแจงร่วมแบบนอนอีลิปติคอล

นายชีวินทรีย์ วีระวานิช

สารนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมการเงิน ภาควิชาการธนาคารและการเงิน
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2563

Independent Study Title	Reverse stress testing on non-elliptical jointly distributed multivariate data
By	Mr. Chevincee Werawanich
Field of Study	Financial Engineering
Thesis Advisor	Associate Professor SIRA SUCHINTABANDID, Ph.D.

Accepted by the FACULTY OF COMMERCE AND ACCOUNTANCY, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science

INDEPENDENT STUDY COMMITTEE

..... Chairman
(Associate Professor THAISIRI WATEWAI, Ph.D.)

..... Advisor
(Associate Professor SIRA SUCHINTABANDID, Ph.D.)

..... Examiner
(Tanawit Sae-Sue, Ph.D.)

ชีวินทรีย์ วีระวานิช : การทดสอบภาวะวิกฤตแบบย้อนกลับสำหรับข้อมูลที่มีการแจกแจงร่วมแบบนอนอีลิปติคอล. (Reverse stress testing on non-elliptical jointly distributed multivariate data) อ.ที่ปรึกษาหลัก : รศ. ดร.สิระ สุจินตะบัณฑิต

-

สาขาวิชา วิศวกรรมการเงิน	ลายมือชื่อ นิสิต .....
ปีการศึกษา 2563	ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

6082985326 : MAJOR FINANCIAL ENGINEERING
KEYWO RD: Reverse stress testing, Non-elliptical distribution, Archimedean copulas

Chevincee Werawanich : Reverse stress testing on non-elliptical jointly distributed multivariate data. Advisor: Assoc. Prof. SIRA SUCHINTABANDID, Ph.D.

-

Field of Study:	Financial Engineering	Student's Signature .....
Academic Year:	2020	Advisor's Signature .....

TABLE OF CONTENTS

1. Introduction

This work deals with quantitative reverse stress testing (RST). Unlike regular stress testing where extreme scenarios are chosen beforehand so that losses given those scenarios can be evaluated, the purpose of RST is to straightforwardly identify the scenarios which push portfolio losses exceeding a predefined threshold. For regular stress testing, its inherent difficulty lies in selecting scenarios which are both sufficiently extreme and sufficiently plausible, yet not subjected to individual's bias nor limited only to historical occurrences. RST here could be considered as a very beneficial tool to assist in a scenario selection process.

One of the main challenges of RST is portfolios in general are sensible to multiple risk factors. For a single risk factor portfolio, it is rather clear in which direction, and even by how much to stress the factor to get an adverse outcome. However, a multiple risk factor portfolio requires more advanced multivariate analysis in order to identify the sensible combinations of stresses applying to each factor, which in concurrence, result in portfolios having large losses.

The role of RST has been emphasized consultatively by many authorities (Financial Services Authority 2009; Committee of European Banking Supervision 2009). However, there is still no industry standard of how the RST should be conducted, even the number of published literatures on this topic is very limited. We observe the fact that the more realistic the framework is to the real-world situations, the more conceptually and mathematically challenging.

Among the recent publications on RST, Glasserman et al. (2015) presented a data-driven framework to identify the most likely scenarios that lead to portfolio losses exceeding a given threshold, referred as the most likely loss scenarios. In this framework, first the conditional mean of portfolio risk factors (given that associated portfolio losses exceeded a predefined threshold) was estimated by nonparametric empirical likelihood algorithm. Then, they derived a scaling procedure based on the tail decay assuming the joint distribution of risk factors and portfolio losses are elliptically contoured, and used it to adjust the confidence regions for the conditional mean to the most likely loss scenarios. Its application is so simple that general practitioners can

conveniently apply the procedure to discover the most likely loss solutions from the raw data they have collected, without having to figure out the exact extreme joint distribution. The concept of conditional expectation is also straightforward enough for practitioners to interpret it from their existing data (comparing to let them figure the most likelihood directly), while Glasserman et al. (2015) provided the missing pieces connecting the conditional expectation to the most likely loss solutions.

Because we view RST for scenario selection as an exploratory process, reliance on a single scenario, even the most likely one, might be misleading and potentially overlook the certainty around the critical scenarios. As highlighted in Glasserman et al. (2015), the main outcomes of their RST method are not only a single most likely scenario, but also the important regions for most likely loss scenarios, where importance reflects both the likelihood of the outcome and the severity of the resulting loss. Those important regions are derived from the confidence regions of the conditional mean. The contours of the regions provide sets of extreme scenarios that are equally plausible.

Our study picked up the most likely loss framework from Glasserman et al. (2015). However, because we believe that financial data in real-world are not necessarily be elliptically distributed, in this work we put our effort to make use of Glasserman et al. (2015)'s framework on general cases where the joint distributions of portfolio risk factors do not necessarily be elliptically contoured.

This study incorporated the concept of copula which brings a benefit from its flexibility in modelling dependency, especially in the tail structure, to govern the joint dependence structure of our portfolio risk factors. Another major advantage from using a copula is that it allows us to alternately derive the representation of the reverse stress testing problem in the form of a copula and a marginal function. This representation plays an important role for estimating the exact solution of reverse stress testing, which in the later stage is used for evaluating the performance of our adapted RST approach.

The objectives of this study are (1) to adapt Glasserman et al. (2015)'s RST framework on non-elliptical data, and (2) to evaluate the performance of our adapted RST framework on portfolios having non-elliptical joint distributions. We selected

three non-elliptical copulas: Frank copula, Clayton copula, and Gumbel copula from Archimedean copula family carrying different dependence characteristics to govern the joint distributions of our portfolio risk factors. We generated portfolio risk factors out from the assigned copula with student-t marginals, combined them into portfolios, then used them as the testing data for our adapted RST.

To adapt Glasserman et al. (2015)'s RST method, we fitted our testing data with an elliptical multivariate t distribution in order to acquire the scaling parameter. We evaluate the performance of our adapted RST method by employing simulation to test for the coverage percentage (counting how many times the resulting important regions we got from our adapted RST method contained the exact solutions to reverse stress testing). We assumed that the adapted scaling approach may exhibit some biases in particular dependence structures, however, we believed that the directions of biases can at least be predicted based on the assumptions we have on the dependence structures. Apart from the coverage results, we also concluded our findings on the characteristics of the resulting important regions for the most likely loss scenarios which vary with copula types and copula's parameters. Since the shapes of those important regions follow their empirical likelihood profiles, we can study the differences in the nature of extreme lower-tail dependence structures as produced by the three copulas within our most likely loss framework.

The organization of this report is arranged as follows. Section 2 walks through the background of RST and reviews some literatures related to our study. Section 3 describes the methodology used in this work. Section 4 shows the results of our adapted RST method. Section 5 provides our conclusion on this study.

2. Background and literature review

2.1 Regular stress testing and reverse stress testing

Apart from maximizing return, what typical investors would expect from their investment is, if losses are to be unavoidable then they shall be under control, or if possible, be bounded at some acceptable levels. This firmly establishes one of the major

goals in investment portfolio management, where several risk management strategies have been studied and developed through decades, and a number of them have been practically employed across the worldwide banking and investment industry. Among those practices, stress testing is one of the very popular and most powerful tools, which could help to give an insight of how a portfolio will fare during bad time or crisis periods. Because of its simplicity in both how it can be performed and how it can be interpreted, stress testing has been using as a standard convention nowadays for risk management practice in investment firms and financial institutions.

In general practice of a regular stress testing, portfolio managers or risk managers will shock their portfolios with some extreme values or off-base rates to see how their portfolio values fluctuate with those predefined assumptions. Those assumptions, or what so-called 'scenarios', might be based on what exactly used to happen in the past, for example, fluctuation of macroeconomic factors during the European sovereign debt crisis between 2009 to 2011. Also, those assumptions could arrive from individual beliefs, on what they might inherently foresee in the upcoming future. For instance, portfolio managers may have their own views on the movement of federal funds rates, therefore they might want to see their portfolios be stressed based on those views.

Even though current practice of stress testing could provide a decent perception for risk assessment, there are a number of things to be concerned. A critically important challenge lies on how to design appropriate input scenarios. We see that selection of input scenarios is extremely severe for the performance of stress testing as a regular stress testing can only be used to foresee how the portfolio values will fare within the sets of the given inputs. If the sets of inputs entirely rely on the historical data, apparently the outputs are likely to underestimate in forecasting overall portfolio risk. Naturally, it is because there were very few occurrences of extreme events by its definition. On the other hand, if the sets of input scenarios largely involve personal judgments, it would then largely rely on individual's experience, and also be likely subjected to biases. Therefore, it is better if we can come up with a standard way to suggest what scenarios should we pick as the inputs of a regular stress testing.

Reverse Stress Testing

Recently, reverse stress testing (RST) has been developed to straightforwardly identify scenarios that lead to severe losses. Hence, it might be considered as a very beneficial tool to assist in scenario selections for the inputs of a regular stress testing. Regardless its benefits and increasing demands from the industry, there are only few academic studies and the current body of literature on RST is still sparse.

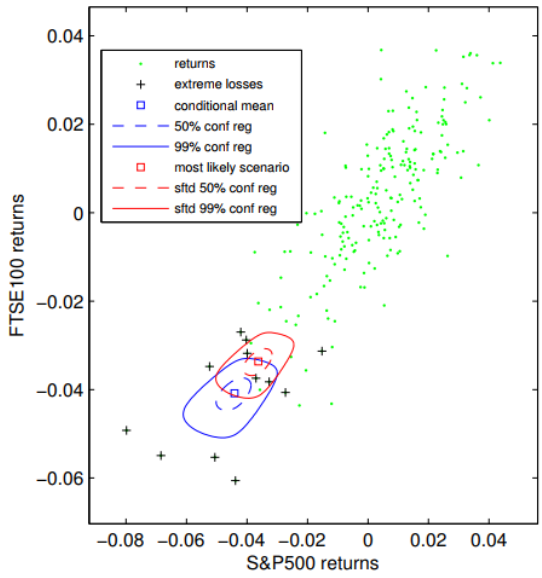
A discussion on qualitative approach for RST presented by Grundke (2012) concluded that qualitative approach alone would not work but at least requires support from quantitative elements. Early studies on quantitative RST conducted on simply structured portfolios with one or two risk factors, for example, Liermann and Klauck (2010). Conceptually, McNeil and Smith (2012) introduced the perception of identifying the most plausible scenario, which is called the most likely ruin event. Many studies rather focused on identifying the worst (in the sense of expected losses for a given portfolio) scenario from a set of scenarios with a given plausibility, for example, Breuer et al. (2012); Breuer and Csiszar (2012).

Recent literatures on methods and algorithms for identifying RST scenarios includes Grundke (2012), Kopeliovich et al. (2015), Flood and Korenko (2015), and Glasserman et al. (2015). Grundke (2012) proposed a framework for a fixed-income portfolio having an exposure to macroeconomic risk factors. Kopeliovich et al. (2015) used principal components of portfolio risk factors to generate candidate scenarios leading to the equal loss. Later, those scenarios are chosen to match with possibility and given constraints. The authors pointed out that selecting only a single scenario or the most plausible one might disregard important information which would affect portfolio risk assessment. Flood and Korenko (2015) suggested an algorithm based on eigenvectors of the variance-covariance matrix for elliptical jointly distributed risk factors to generate equally possible scenarios. This approach allows for a flexible loss function; however, it might overlook some severe but less likely scenarios as the possibility of a scenario is set and not the loss level.

2.2 Reverse stress testing on elliptically distributed portfolio

Among on those recent development in RST studies, Glasserman et al. (2015) came up with a data-driven lightly-parametric framework which can robustly identify 'the most likely scenarios' that lead to portfolio losses exceeding a given threshold. The central idea of this approach is the scaling procedure which derives asymptotically exact scaling multipliers when the joint distribution of market factors and portfolio value falls within a broad family of elliptically contoured distributions.

Figure 1. *Glasserman et al. (2015)'s illustration of an example portfolio having exposure to weekly returns of FTSE100 index and S&P500 index. The blue lines indicated confidence regions of conditional mean, while the red lines indicate important regions for the most likely loss scenario.*



In this scaling framework, first the conditional mean of risk factors (given portfolio having large loss) is estimated by employing nonparametric empirical likelihood (EL) in the sense of Owen (2001). Importantly, this EL estimator can provide confidence regions of conditional mean, and does not rely on significant assumptions about the conditional distribution of the market factors in extremes. The shape of the EL's confidence regions is also able to capture skewness and other features present in extreme outcomes. Then, in the second step, the scaling procedure is applied to adjust the conditional mean and its confidence regions to the most likely loss scenario and the most likely loss regions with respect to given confidence levels. Therefore, not only can we find the most likely combination of stresses which push the portfolio losses

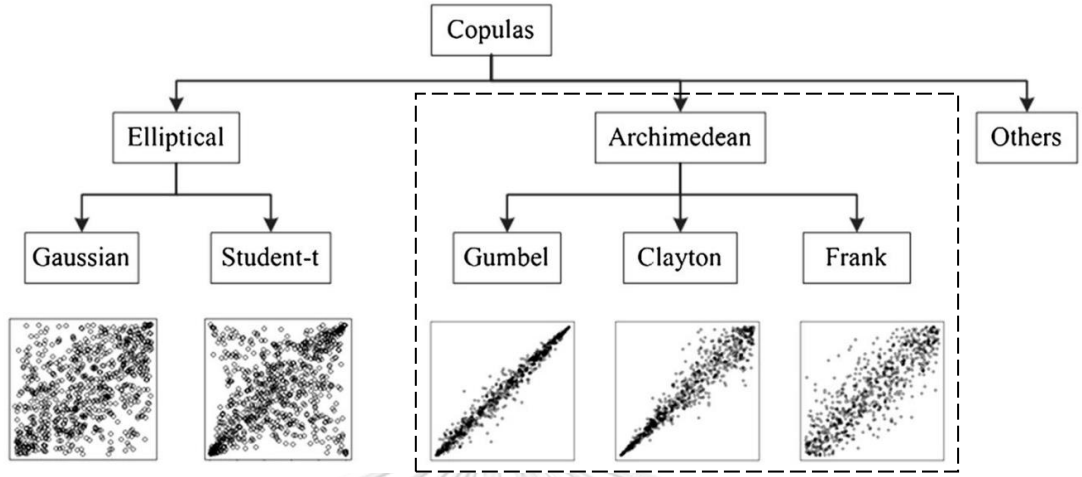
beyond the predefined threshold, the resulting most likely loss regions also provide additional information about uncertainty in extremes which is generally absent from the stress scenario selection process.

2.3 Non-elliptical copula

There were some evidences illustrating that the distribution of market risk factors might not be elliptical, and why reliance on such assumption could bring some failures. For instance, many studies found the equity returns suffer from increased correlations during bear markets, for example, Longin and Solnik (2001, 1995) and Ang and Chen (2002) indicating financial markets exhibit a characteristic known as lower tail dependence. Donnelly and Embrechts (2010) reviewed what occurred during the 2007-2008 financial crisis and gave criticisms on the use of elliptical dependence model such as Gaussian copula. Many studies show that while forecasting models, incorporating asymmetric dependence produce significant gains for the investor (Garcia and Tsafack 2011; Chu 2011).

In our work, we want to introduce the use of copula which having the flexibility in modelling dependence, especially in the tail structure, to reverse stress testing framework. In short, a copula is a multivariate cumulative distribution function for which the marginal probability distribution of each variable is uniformly distributed. With copula modelling, we are able to govern the multivariate dependence structure separately from the marginal distributions. In the previous literature on regular stress testing, there were some studies employing Gaussian copula. However, the Gaussian copula is lacking as it only allows for an elliptical dependence structure. There still be a challenge for integrating copula to reverse stress testing framework, especially in the non-elliptical field.

***Figure 2**. Copulas in the classes of Elliptical copula family and Archimedean copula family (Li et al. 2015).*



In this study, we picked three non-elliptical copulas from Archimedean family which are Frank copula, Clayton copula, and Gumbel copula for modelling dependence structures. Frank is a symmetric copula. Clayton is an asymmetric copula exhibiting greater dependence in the negative tail than in the positive. Gumbel, or also known as Gumbel-Hougard, is an asymmetric copula exhibiting greater dependence in the positive tail than in the negative.

Apart from the nice properties they possess and the ease with which they can be constructed, we also observed many cases that the three copulas from Archimedean family turned out to be most efficient in financial applications and provided some advantages over elliptical-class copulas. For example, Aas et al. (2009) shows that applying Clayton in portfolio optimization and risk management can reduce the effects of extreme downside correlations and improve statistical and economic performance compared to the elliptical dependence copulas such as the Gaussian and Student-t. Hatherley and Alcock (2007) reported that managing asymmetric dependence, using a Clayton standard copula against the benchmark multivariate normal probability model results in reduced downside exposure. Agnihotri (2017) studied the correlation between bond market and stock market in US and China for past between 2006- 2016 and indicated that Frank and Gumbel are the most appropriated copula models.

3. Methodology

To formulate the ideas of reverse stress testing into our analysis, let us first assign $\boldsymbol{Z}$ to be a random d-dimensional vector representing the changes in market risk factors relevant to a portfolio. These factors could be interpreted as rates, prices, returns, or any other economic variables. Suppose $\boldsymbol{Z}$ has a probability density $f_{\boldsymbol{Z}}$ on $\mathbb{R}^{\mathrm{d}}$, which represents the joint probability density of its elements $Z_1, Z_2, \ldots, Z_d$. For any given scenario $\boldsymbol{z}$, the loss of a portfolio being exposed to these factors is denoted by L, which can be written as a function of $\boldsymbol{z}$.

We say that a scenario has a large loss when the portfolio loss L with respect to this scenario is at least, or larger than loss threshold ℓ, which is a predefined value. We write the conditional probability density of $\boldsymbol{Z}$ given $L \geq \ell$ as $f_{\boldsymbol{Z}}(\boldsymbol{z} \mid L \geq \ell)$. The generic problem of reverse stress testing is to find the most likely scenarios $\boldsymbol{z}^*$ given loss L greater than or equal to ℓ; in other words, to solve

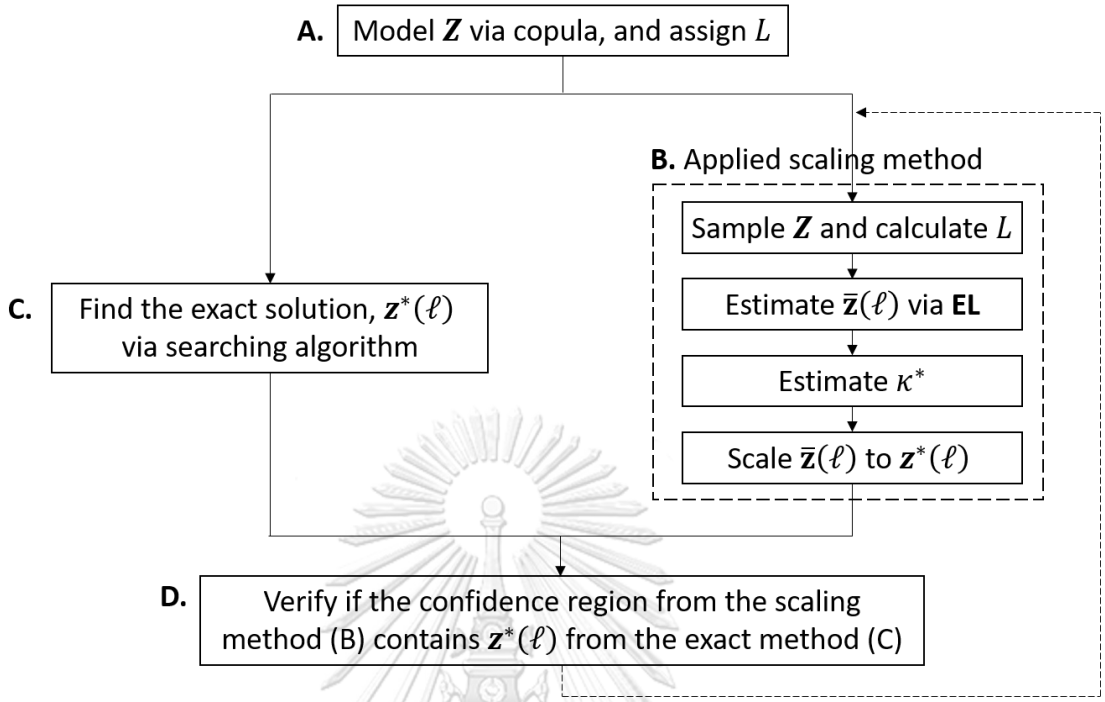
$$\boldsymbol{z}^*(\ell) = \underset{\boldsymbol{z} \in \mathbb{R}^d}{argmax}\ f_{\boldsymbol{Z}}\left(\boldsymbol{z} \mid L \geq \ell\right)$$

The solutions to the previous equation $\left(\boldsymbol{z}^*(\ell)\right)$ are referred as the solutions to the reverse stress test or the most likely loss scenarios.

We see that the solutions to the reverse stress testing problem depend largely on the approximation of the conditional joint probability density, $f_{\boldsymbol{Z}}\left(\boldsymbol{z} \mid L \geq \ell\right)$. However, in reality, most of the time the joint distributions between portfolio risk factors are not certainly known. In this work, we adapted Glasserman et al. (2015)'s RST framework, which is lightly parametric and does not require the exact joint distribution to be known to our synthetic non-elliptical portfolios.

Because Glasserman et al. (2015)'s scaling method is based on the assumption of elliptically contoured distributions, we have to apply some techniques to adjust the method. We tested the performance of our adapted models through simulation, where the exact RST solutions are alternatively acquired from a closed form of copula and marginal. In the later chapter, we discussed the performance or our adapted RST methods. Our research methodology is presented in Figure 3.

Figure 3. *Research methodology*



The methodology in this study is organized into the four main sections:

Section A describes how our synthetic portfolios are constructed. First, we assign a copula to capture the desirable characteristics of the dependence structure of risk factors. Second, we assign marginal distributions for the risk factors. Then, we employ a loss function (L) to combine risk factors into portfolios.

Section B explains how we adapted Glasserman et al. (2015)'s RST approach on our non-elliptical portfolios. This section involves estimation of the conditional mean via Empirical Likelihood algorithm, and the scaling procedure adjusting for the most likelihood solutions. Since the joint distributions of our portfolios are not elliptical, we take some additional steps for determining an appropriate scaling factor based on the characteristic of tail dependence structure as governed by a copula.

Section C shows how we estimate the exact solution to the most likely loss scenario from the closed-form joint pdf deriving from copula and marginals. Section D, we give our design of how we employ simulation to determine the coverage of the most likely loss regions acquiring from our adapted scailing RST method.

Section A. Risk factor modelling and portfolio construction

This section describes in detail how our synthetic portfolios are built, from risk factor modelling to portfolio construction. In Section A1, we present a model which derives the components of the joint probability of portfolio risk factors, $f_{\boldsymbol{Z}}(\boldsymbol{z})$ from their marginals and the governed copula. In Section A2, we determine the function L which is used to transform portfolio risk factors into portfolio values, or portfolio losses. Then, in Section A3, we combine the results in the previous parts to illustrate the conditional joint probability, $f_{\boldsymbol{Z}}(\boldsymbol{z} \mid L \geq \ell)$ which is the probability of scenarios given the associated portfolio losses exceed the given threshold. This provides the basis for the most likelihood estimation which is further explained in section B.

A.1 Risk factor modelling

Our portfolio risk factor is modelled based on the two main steps: (1) construction of marginal, and (2) construction of dependence structure. For ease of interpretation we may think, for example, that our portfolios are equity portfolios consisting of more than one equity indices. Then, returns of equity indices are considered as portfolio risk factors driving the performance of our portfolios.

Because fat tails are generally observed in financial returns, we selected student-t distribution for the marginal distribution of our portfolio risk factors. In other words, $\boldsymbol{Z}_1, \boldsymbol{Z}_2, \dots, \boldsymbol{Z}_d \sim$ Student-t (μ, σ, υ) with the pdf $f_{\boldsymbol{Z}_i}$ and the cdf $F_{\boldsymbol{Z}_i}$ for $i = 1, \dots, d$ as

$$f_{\boldsymbol{Z}_i}(\boldsymbol{z}_i \mid \mu, \sigma, \upsilon) = \frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\Gamma\left(\frac{\upsilon}{2}\right)\sqrt{\pi\upsilon\sigma^2}}\left[1 + \frac{1}{\upsilon}\left(\frac{z_i - \upsilon}{\sigma}\right)^2\right]^{-(\upsilon+1)/2},$$

$$F_{\boldsymbol{Z}_i}(\boldsymbol{z}_i \mid \mu, \sigma, \upsilon) = \int_{-\infty}^{\boldsymbol{z}_i} f_{\boldsymbol{Z}_i}(\boldsymbol{z}_i \mid \mu, \sigma, \upsilon)\, d\boldsymbol{z}_i .$$

Another important and unique characteristic we usually observe from the nature of financial data and is that some of them possess considerable stronger dependency in their lost or gain domains. To allow the flexibility in modelling the joint tail structure, we assigned a copula to govern dependency.

Basically, a copula is a multivariate distribution whose marginals are all uniform over $(0,1)$. For a d −dimensional random vector on the unit cube, a copula C is

$$C(\boldsymbol{u}_1, \dots, \boldsymbol{u}_d) = \Pr(\boldsymbol{U}_1 \leq \boldsymbol{u}_1, \dots, \boldsymbol{U}_d \leq \boldsymbol{u}_d). \quad (1)$$

Combined with the fact that any continuous random variable can be transformed to be uniform over $(0,1)$ by its probability integral transformation, Sklar (1959) showed that there exists a d −dimensional copula C such that for all z in the domain of $F_{\boldsymbol{Z}}$,

$$F_{\boldsymbol{Z}}(\boldsymbol{z}_1, \dots, \boldsymbol{z}_d) = C\{F_{\boldsymbol{Z}_1}(\boldsymbol{z}_1), \dots, F_{\boldsymbol{Z}_d}(\boldsymbol{z}_d)\}. \quad (2)$$

We selected three non-elliptical copulas from Archimedean family including Frank copula, Clayton copula, and Gumbel copula which allow flexibility in modelling the joint structures over elliptical-class copulas. For the bivariate case $(d = 2)$, the simple forms of the three Archimedean copula functions are illustrated in Table 1.

Copula	Bivariate copula cdf, $C(\boldsymbol{u_1}, \boldsymbol{u_2}; \boldsymbol{\theta})$
Frank	$-\frac{1}{\theta}\log\left[1 + \frac{(e^{-\theta \boldsymbol{u}_1}-1)(e^{-\theta \boldsymbol{u}_2}-1)}{(e^{-\theta}-1)}\right]$
Clayton	$\{\max[(\boldsymbol{u}_1^{-\theta} + \boldsymbol{u}_2^{-\theta} - 1), 0]\}^{-\frac{1}{\theta}}$
Gumbel	$\exp\left\{-[(-\ln \boldsymbol{u}_1)^{\theta} + (-\ln \boldsymbol{u}_2)^{\theta}]^{\frac{1}{\theta}}\right\}$

Table 1. *Three copula functions of the three copulas from Archimedean family of bivariate risk factor model.*

In higher dimensions $(d > 2)$, Archimedean copula can be constructed through a generator function φ as

$$C(\boldsymbol{u_1}, \dots, \boldsymbol{u_d}) = \varphi^{-1}\{\varphi(\boldsymbol{u_1}), \dots, \varphi(\boldsymbol{u_d})\}, \quad (3)$$

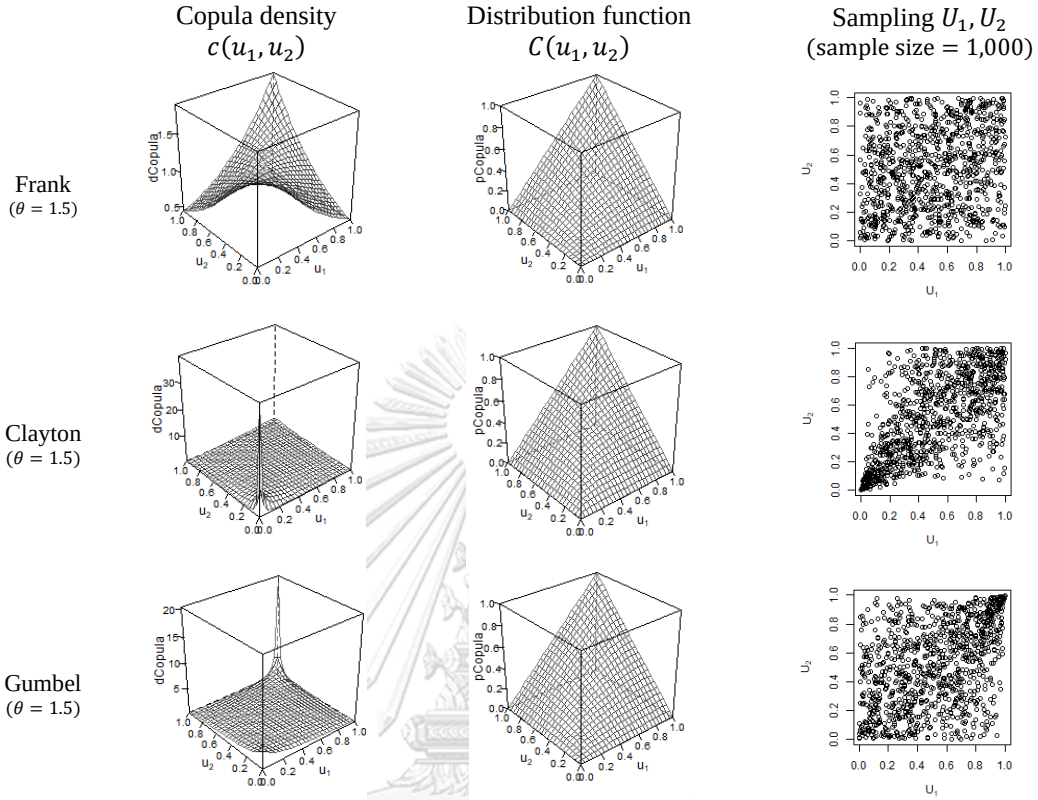
where φ^{-1} is the inverse of the generator φ. The three multivariate Archimedean copulas are summarized in Table 2.

Copula	Parameter space	Generator $\varphi(t)$	Generator Inverse $\varphi^{-1}(t)$
Frank	$\theta \geq 0$	$-\ln\frac{e^{-\theta t}-1}{e^{-\theta}-1}$	$-\alpha^{-1}\ln(1 + e^{-s}(e^{-\theta} - 1))$
Clayton	$\theta \geq 0$	$t^{-\theta} - 1$	$(1 + s)^{-1/\theta}$
Gumbel	$\theta \geq 1$	$(-\ln t)^{\theta}$	$\exp(-s^{1/\theta})$

Table 2. *Archimedean generators φ for multivariate mode, $d > 2$.*

It is worth noting that the parameter value at the boundary of parameter space gives the independent copula after taking the limit.

Figure 4. *Wireframe surfaces of copula density and distribution function, with scatter sampling plots of Frank copula, Clayton copula, and Gumbel copula for bivariate models having copula's* $\theta = 1.5$.



We mentioned that Frank copula is a symmetric copula providing the same degree of dependency for both negative and positive domains. Clayton copula, on the other hand, is an asymmetric copula which exhibits greater dependence in the negative tail than in the positive tail. Gumbel copula is also an asymmetric copula; however, it exhibits greater dependence in the positive tail than in the negative tail. For giving a clearer picture, we illustrated examples of copula density, distribution function, and samplings of the three copulas for bivariate models having $\theta = 1.5$ in Figure 4.

Now that we have captured the essence of how our risk factor modelling are done showing in a straightforward order, in practice of risk factor sampling we rather start from copula then map out to the marginals. An example for a bivariate risk factor sampling case is illustrated in Algorithm 1.

Algorithm 1. Sampling bivariate risk factors

1. Draw i.i.d. samples $u_{1,i}, \dots, u_{1,N}$ from a uniform distribution on $[0,1]$ for the elements of a vector $\boldsymbol{u_1}$ representing the marginal cumulative probability of risk factor $\boldsymbol{Z_1}$.

$$U_{1,i}, \dots, U_{1,N} \sim \text{ i.i.d. Uniform } [0,1]$$

2. Draw i.i.d. samples $w_i, \dots, w_N$ from a uniform distribution on $[0,1]$, which are the values the inverse conditional distribution of a bivariate copula.

$$W_i, \dots, W_N \sim \text{i.i.d. Uniform } [0,1]$$

3. Derive a vector $\boldsymbol{u_2}$ from $\boldsymbol{u_1}$ and $\boldsymbol{w}$ via the inverse conditional distribution function (inverse h-function) of the copula. Note that $\boldsymbol{u_2}$ represents the cumulative probability of risk factor $\boldsymbol{Z_2}$.

$$u_{2,i} = h_1^{-1}\left(w_i \middle| u_{1,i}, \theta\right)$$

4. From the cumulative probability vectors $\boldsymbol{u_1}, \boldsymbol{u_2}$, sample $\boldsymbol{Z_1}, \boldsymbol{Z_2}$ follow their marginal distributions.

$$Z_1 = F_{Z_1}^{-1}(U_1)$$
$$Z_2 = F_{Z_2}^{-1}(U_2)$$

A.2 Portfolio construction

After we have our portfolio risk factors from Section A1, we employed a loss function (L) to combine those portfolio risk factors into a portfolio. The outcomes of L are referred as portfolio values or portfolio losses. As in the same manner as Glasserman et al. (2015), we assumed portfolio risk factors recording the most important factors influencing the portfolio, so that the tail behavior of the portfolio loss consistent with that of the most important factors. Based on Glasserman et al. (2015), the loss function can be any function either deterministic or stochastic. However, to govern the scope of our work, we assigned the portfolio loss to be a deterministic linear function of risk factors, given by

$$L = c^T \boldsymbol{z} \tag{4}$$

where c is a d-dimensional vector of constants sum up to 1. For purposes of illustration, we selected the linear coefficient vector of the loss function c to equally-weight all the relevant factors, therefore $c^T = \left[\frac{1}{d}, \frac{1}{d}, \ldots, \frac{1}{d}\right]$.

A.3 Reverse stress testing

Combining what we have done in Section A1 and Section A2 together, now we are able to write the conditional pdf of our portfolio risk factors as $f_{\mathbf{Z}}(\boldsymbol{z}|L \geq \ell)$. Therefore, to solve for the most likely loss scenarios is to find the solutions of

$$\boldsymbol{z}^*(\ell) = \underset{\boldsymbol{z} \in \mathbb{R}^d}{\operatorname{argmax}} \; f_{\mathbf{Z}}\left(\boldsymbol{z} \mid L \geq \ell\right) \tag{5}$$

where L follows equation (4). Because of using copula, we can also derive $f_{\mathbf{Z}}$ in terms of copula and marginal functions. This plays a significant role in verifying the results of our adapted scaling reverse stress testing approach which will be explained later in more detail in Section C. Before that, let us show how we adapted scaling method to estimate the most likely loss regions without knowing the actual closed-form of the risk factor distribution in Section B.

Section B. Reverse stress testing on non-elliptical distribution (adapted scaling method)

After having all the models readily set up in section A, we now proceed to perform a reverse stress test to find the most likely loss scenarios. In this section, we present how we adapt Glasserman et al. (2015)'s RST on our synthetic portfolios.

In brief, first we estimated the conditional mean of the underlying market factors given large losses ($\mathbb{E}[\boldsymbol{z}|L \geq \ell]$) by a nonparametric Empirical Likelihood (EL) algorithm. Then, we fitted a multivariate student t distribution to estimate an index variable υ in order to apply scaling method to adjust the conditional mean to the most likely loss scenarios. Expecting this might established some biases in particular dependence structures, in the next chapter we will test for the performance of our RST method in each copula setting with the exact solutions acquired from Section C. Apart from the coverage results, we also study the characteristics of the resulting important regions for the estimated most likely loss scenarios, and the behavior of coverages (or

non-coverages). In the end of this study, we conclude whether our adapted RST method can be used on non-elliptical portfolios constructed through the three different copulas. In the case that having biases on a particular dependence characteristic, we discuss whether we could at least predict the direction of biases based on the assumption we have on the dependence structure.

B.1 Empirical likelihood estimation of the conditional mean

For our goal to find the solution to RST $\left(\boldsymbol{z}^*(\ell)\right)$, our first intermediate step is to focus on estimating $\bar{\boldsymbol{z}}(\ell)\ \ (= \mathbb{E}[\boldsymbol{z}|L \geq \ell])$, which is the conditional mean of the portfolio risk factors given portfolio having large losses.

Before moving toward the estimation process, we mentioned that our objective of RST is not to identify only a single most likely scenario, but its whole important regions where the contours of the regions provide the sets of equally plausible scenarios. For that, in this section, we need to find not only the conditional mean, but also its whole confidence region to be used as the inputs for the next section.

Owen (2001)'s Empirical likelihood (EL) is a nonparametric estimation procedure through which we get confidence regions for the conditional mean. Importantly, the EL estimator does not rely on any significant assumption about the conditional distribution of the market factors. The shape of the resulting confidence regions is able to capture skewness and other features presented in extreme outcomes.

This method considers convex combinations of the observations as candidate estimates of the mean:

$$w_1\boldsymbol{z}_1 + w_2\boldsymbol{z}_2 + \cdots + w_n\boldsymbol{z}_n,$$

$$\sum_{i=1}^{n} w_i = 1\,, w_i \geq 0, i = 1, \ldots, n\,. \tag{6}$$

The profile empirical likelihood associated with a candidate value **x** is

$$\mathcal{R}(\mathbf{x}) = \max\left\{\prod_{i=1}^{n} nw_i : \sum_{i=1}^{n} w_i\boldsymbol{z}_i = \mathbf{x}\,,\right.$$

$$\sum_{i=1}^{n} w_i = 1 \ , w_i \geq 0 \ i = 1, \dots, n \Bigg\} \tag{7}$$

The product inside the braces is the likelihood ratio of the probability vector $(w_1, w_2, \dots, w_n)$ to the uniform distribution $(1/n\,, 1/n\,, \dots, 1/n)$. $\mathcal{R}(\mathbf{x})$ is larger when $\mathbf{x}$ is a more uniform convex combination of the weights on the observations, and it is maximized at the sample mean of the observations.

The maximization problem defining the profile empirical likelihood is easy to solve by first reformulating it as

$$\max_{w_1,\dots,w_2} \sum_{i=1}^{n} \log w_1 \quad \text{subject to} \quad \sum_{i=1}^{n} w_i = 1 \ , \sum_{i=1}^{n} w_i \boldsymbol{z}_i = \mathbf{x} \tag{8}$$

Suppose that the observations are i.i.d. with mean $\boldsymbol{\mu_0} (= \mathbb{E}[\boldsymbol{Z} | L \geq \ell])$, and suppose that the convex hull of the observations contains $\boldsymbol{\mu_0}$ with probability approaching 1 as the number of observations increases. Then, Owen (2001)'s Theorem 3.2 states that $-2 \log \mathcal{R}_\ell(\boldsymbol{\mu_0})$ has an asymptotic χ_d^2 distribution for large n. This provides the basis for EL confidence regions. By fixing a confidence level $1 - \alpha$ and finding the quantile χ_α for which $\mathbb{P}(\chi_d^2 \geq \chi_\alpha) = \alpha$; the corresponding $1 - \alpha$ confidence region for $\boldsymbol{\mu_0}$ is the set

$$\mathcal{C}_{1-\alpha,n} = \Bigg\{ \sum_{i=1}^{n} w_i \boldsymbol{z}_i : \prod_{i=1}^{n} n w_i \geq \exp\left(-\frac{\chi_\alpha}{2}\right), \sum_{i=1}^{n} w_i = 1 \ , w_i \geq 0 \ i = 1, \dots, n \Bigg\} \tag{9}$$

Because the algorithm is empirical based, it takes the inputs from the samples of factors and portfolio values. Therefore, we first have to sample the risk factors based on our risk factor model presented in Section A1. From these samples, we formed our portfolios based on equation (4) and calculated the realization of portfolio losses (L) for every sample scenario. As a result, we had observations $(\boldsymbol{z}_i, L_i)$, $i = 1, 2, \dots$ of scenarios $\boldsymbol{z}_i$ and corresponding loss L_i. From these, we assigned our threshold for large loss, ℓ to be equal to the value of the p^{th} percentile of portfolio values. Then, we discard all observations except those for which the loss is at least ℓ. Consequently, we

were left with n large loss observations, $(\boldsymbol{z}_1, L_1), (\boldsymbol{z}_2, L_2), \ldots, (\boldsymbol{z}_n, L_n)$ where $L_i \geq l$ for $i = 1,2, \ldots, n$. Now we see that by estimating the remaining scenarios, the original problem of estimating a conditional mean reduces to one of estimating an unconditional mean.

After applying an empirical likelihood algorithm described in this section to the large loss scenarios, we can get the conditional expectation $\mathbb{E}(\boldsymbol{z}|L \geq \ell)$ together with its confidence regions. As mentioned earlier, the resulting confidence regions from EL algorithm are favoring because they give the information about the uncertainty around the critical points, and are able to capture skewness and other notable shape characteristics in the extremes with minimal assumptions on the distribution of the underlying data.

B.2 Adapting the scaling method on non-elliptical distribution

Now we proceed to the procedure which converts the conditional mean $\bar{\boldsymbol{z}}(\ell)$ to the most likely loss scenario $\boldsymbol{z}^*(\ell)$. Glasserman et al. (2015) shows that there exists a positive scalar sequence κ such that

$$\boldsymbol{z}^*(\ell) = \kappa_\ell \bar{\boldsymbol{z}}(\ell), \text{ and } \kappa_\ell \to \kappa \text{ as } \ell \to \infty \tag{10}$$

where $\kappa = (\upsilon - 1)/\upsilon$ for an elliptical distribution family regularly varying with an index $\upsilon > 1$. This υ represents the tail distribution decays at the power $x^{-\upsilon}$, thus a smaller index ν thus indicates a heavier tail.

As we mentioned earlier that in addition to scale a single scenario $\bar{\boldsymbol{z}}(\ell)$, more importantly we want to scale whole confidence regions of $\bar{\boldsymbol{z}}(\ell)$ to get confidence regions for $\boldsymbol{z}^*(\ell)$. Such a procedure associates two limits. Firstly, equation (10) applies as $\ell \to \infty$. Secondly, the chi-square limit that underpins the EL method holds as the number of observations grows $(n \to \infty)$. Based on Owen's (2001) Theorem 4.1, Glasserman et al. (2015) shows that,

$$-2 \log \mathcal{R}_\ell\big(\bar{\mathbf{z}}(\ell)\big) = -2 \log \mathcal{R}_\ell\left(\kappa_\ell^{-1} \mathbf{z}^*(\ell)\right) \to \chi_\alpha^2 \tag{11}$$

in distribution, and $\kappa_\ell \mathcal{C}_{1-\alpha, n_\ell}$ is an asymptotic $100(1-\alpha)\%$ confidence region for the most likely loss scenario $\mathbf{z}^*(\ell)$; i.e.

$$\mathbb{P}\left(\mathbf{z}^*(\ell) \in \kappa_\ell \mathcal{C}_{1-\alpha,n_\ell}\right) \to 1-\alpha \tag{12}$$

$as\ \ell \to \infty, where\ \kappa_\ell \to \kappa, with\ \kappa$ as stated in equation (10).

Estimating the scaling parameter for non-elliptical distribution

The scaling parameter κ in equation (10) assumes portfolio joint distribution falls within a broad family of elliptically contoured distributions. However, to adapt this scaling method to non-elliptical portfolios, we employ a multivariate t_v distribution to fit into our portfolio risk factor samples. Such procedure allows us to roughly acquire a tail heaviness parameter, υ which is the degree of freedom of multivariate t_v to be used as the input for $\kappa = (\upsilon - 1)/\upsilon$.

The density of multivariate t_v with parameters $\mu, \mathbf{\Sigma}, \upsilon$ is given by

$$f(\boldsymbol{z}|\mu,\ \mathbf{\Sigma},\ \upsilon) = \frac{\Gamma\left(\frac{1}{2}(\upsilon+d)\right)}{\Gamma\left(\frac{1}{2}\upsilon\right)(\pi\upsilon)^{d/2}|\mathbf{\Sigma}|^{1/2}} \times \left(1 + \frac{(\mathbf{z}-\mu)^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{z}-\mu)}{\upsilon}\right)^{-(\upsilon+d)/2}$$

$$\text{for } \boldsymbol{z} \in \mathbb{R}^{\boldsymbol{d}}$$

where mean and variance of the distribution are given by

$$\mathbb{E}[\boldsymbol{Z}] = \mu, \quad \mathbb{V}[\boldsymbol{Z}] = \frac{\upsilon}{\upsilon-2}\boldsymbol{\Sigma}$$

assuming $\upsilon > 2$. In order to estimate υ, we first estimate the sample mean and covariance and then maximize the likelihood over υ.

Before moving to the next section, let us point out that the purpose of this study is to explore whether Glasserman et al. (2015)'s RST approach could be adapted to use in general cases where the joint distribution of portfolio risk factors does not necessary be elliptically distributed. One of the major features of Glasserman et al. (2015)'s work is that it is very lightly parametric in a way that only a single parameter required is κ, which is determined by the tail decay of the factors. In the sense that we want to preserve the feature without adding too much complexity, we first thus employing the above technique to roughly get the scaling parameter κ.

Incorporating a copula within this study, we presume that the technique above should reliably perform for the cases that portfolio risk factors are constructed through Elliptical copula family such as Gaussian copula and Student-t copula. Therefore, we

instead picked three non-elliptical copulas: Frank, Clayton, and Gumbel from Archimedean family having different dependence structures to test the performance of our adapted scaling method. We assume our scaling method may give some biases in particular dependence structures; however, we assume the direction of biases could at least be predictable based on the assumption we have on the dependence structure.

Section C. Estimating the exact solution of reverse stress testing

This section shows how we proceed to estimate the exact solution to the problem of reverse stress testing. From equation (2), we can write the joint cdf for our portfolio risk factors as

$$F_{\boldsymbol{Z}}(\boldsymbol{z}_1, \dots, \boldsymbol{z}_d) = C(\boldsymbol{u}_1, \dots, \boldsymbol{u}_d) = C\left(F_{\boldsymbol{Z}_1}(\boldsymbol{z}_1), \dots, F_{\boldsymbol{Z}_d}(\boldsymbol{z}_d)\right)$$

$$\text{where } (\boldsymbol{u}_1, \dots, \boldsymbol{u}_d) \sim Uniform\,[0,1]_d.$$

From the relationship between pdf and cdf, we are able to derive the components of the join pdf as

$$f_{\boldsymbol{Z}}(\boldsymbol{z}_1, \dots, \boldsymbol{z}_d) = C\left(F_{\boldsymbol{Z}_1}(\boldsymbol{z}_1), \dots, F_{\boldsymbol{Z}_d}(\boldsymbol{z}_d)\right) \cdot f_{\boldsymbol{Z}_1}(\boldsymbol{z}_1) \cdots f_{\boldsymbol{Z}_d}(\boldsymbol{z}_d).$$

Hence, the generic problem of reverse stress testing in equation (8) turns into

$$\boldsymbol{z}^*(\ell) = (\boldsymbol{z}_1^*, \dots, \boldsymbol{z}_d^*) = \underset{\boldsymbol{z} \in \mathbb{R}^d, L(\boldsymbol{z}) \geq \ell}{\operatorname{argmax}} \left\{ C\left(F_{\boldsymbol{Z}_1}(\boldsymbol{z}_1), \dots, F_{\boldsymbol{Z}_d}(\boldsymbol{z}_d)\right) \cdot f_{\boldsymbol{Z}_1}(\boldsymbol{z}_1) \cdots f_{\boldsymbol{Z}_d}(\boldsymbol{z}_d) \right\}.$$

From what given, we are able to search for the solutions to the reverse stress testing problem through appropriated optimization techniques. More details on the derivation steps and an example for a bivariate Clayton copula case are provided in Appendix A.

Section D. Performance evaluation

Now that we have our adapted RST procedure from section B and the formula to estimate the exact solution to RST from section C, we can proceed to evaluate the performance of our adapted RST procedure through simulation.

As our study focuses on evaluating the performance of our adapted RST method on non-elliptical dependence structures, we select copula types and copula parameters to be our main independence variables. Then, for each copula setting, the marginals are sampled through i.i.d. student t distributions with degree of freedom $\upsilon = 7$. Results will be evaluated as per equal loss threshold ℓ at 0.1^{th} percentile of the sampled scenarios.

From Section A, we selected three non-elliptical copula types from Archimedean family: 1. Frank copula, which is a symmetric copula; 2. Clayton copula, an asymmetric copula exhibiting greater dependence in the negative tail; and 3. Gumbel copula, an asymmetric copula exhibiting greater dependence in the positive tail.

As the three copula yields the independent copula after taking the limit of the copula's parameter to the boundary of the parameter space ($\theta \geq 0$ for Frank copula and Clayton copula, and $\theta \geq 1$ for Gumbel copula), we set the initial value of copula's θ for each copula closing to its boundary as a starting point. From that, we expect to see the consequences from the characteristic of each copula type getting stronger with the increasing value of copula's θ.

In section B, we mentioned that the scaling method gives asymptotic support for confidence regions as the sample size increase. However, in practice, extreme data in the real world are limitedly available by their nature. Therefore, for our evaluation here, we want to see the performance for both the cases having a small number of large-loss observations ($n = 10$), knowing that the actual coverage of the estimated confidence regions may differ from the nominal coverage; together with the cases having a larger number of large-loss observations ($n = 50$), where the resulting estimated coverages are expected to go closer to the nominal coverages; in the circumstances that our adapted RST approach performs well.

We conduct our study on bivariate risk factor portfolios ($d = 2$) together with three-dimensional risk factor portfolios ($d = 3$), expecting the estimated coverages for portfolios with higher dimension may reach the nominal coverages slower than those for portfolios with lower dimension, at the equal number of input large-loss observation n. We carry out the important regions for the most likely loss scenario at 50% and 95% confidence for comprehensive assessment. In summary, our parameter settings for simulation are designed as shown in Table 3 below.

Parameter	**Setting**
Copula	1. Frank copula: $\theta = 0.2,\ 1.5,\ 3.0$ 2. Clayton copula: $\theta = 0.2,\ 1.5,\ 3.0$ 3. Gumbel copula: $\theta = 1.2,\ 1.5,\ 3.0$
Dimension of portfolio risk factors	$d = 2, 3$
Sample size of large-loss scenarios	$n = 10, 50$
Confidence level	at 50% and 95% confidence

Table 3. *Parameter setting for simulation.*

To test performance our adapted RST method, we generate enough sample points of $\boldsymbol{Z}$ for having n large-loss observations which the loss $L = c^T\boldsymbol{Z}$ are at least ℓ. We keep these n large-loss observations, then construct the confidence regions for the conditional mean $\bar{\boldsymbol{z}}(\ell)$ $(= \mathbb{E}[\boldsymbol{z}|L \geq \ell])$ via an EL algorithm. We convert the confidence regions to the important regions for the most likely loss scenario $\boldsymbol{z}^*(\ell)$ following the instructions given in section B. Then, we can examine whether this most likely loss boundary contains the exact most likely loss point $\mathbf{z}^*(\ell)$ that we acquired from section C.

We repeat the above process by 500 times and record the percentage of how many times that $\mathbf{z}^*(\ell)$ are within the coverage of our estimated most likely loss important regions. We also observe the converging behaviors, especially for the cases having high non-coverage percentage, to see whether there is any pattern of how the estimated boundaries shifted away from the exact $\mathbf{z}^*(\ell)$ points in that particular dependence structures.

4. Results

Our adapted RST method were tested on non-elliptically distributed portfolios constructed through on Frank copula, Clayton copula, and Gumbel copula with varying parameter following the instructions given in the previous section. Our main results are presented in Table 4. Also, we provided illustration examples for all bivariate case in Figure B1-1:4, Appendix B.

4.1 Main results

Frank	$\theta = 0.2$		1.5		3	
95% confidence	$n = 10$	50	$n = 10$	50	$n = 10$	50
$d = 2$	86.4	94.4	82.2	94.2	79.8	95.0
3	73.4	92.6	64.2	86.8	63.4	85.8
50% confidence						
$d = 2$	44.6	59.2	48.6	56.6	46.8	56.2
3	40.0	47.2	32.2	42.4	29.8	33.4
Clayton	$\theta = 0.2$		1.5		3	
95% confidence	$n = 10$	50	$n = 10$	50	$n = 10$	50
$d = 2$	84.8	88.0	83.0	84.0	84.6	81.4
3	72.8	93.4	75.8	87.0	77.0	26.0
50% confidence						
$d = 2$	39.8	39.4	37.0	27.4	37.8	8.6
3	34.4	43.2	33.6	4.4	16.2	2.8
Gumbel	$\theta = 1.2$		1.5		3	
95% confidence	$n = 10$	50	$n = 10$	50	$n = 10$	50
$d = 2$	84.4	93.8	85.2	94.8	84.0	93.4
3	67.8	90.4	74.0	92.0	75.4	94.2
50% confidence						
$d = 2$	51.4	58.8	41.6	61.0	44.4	51.2
3	36.2	51.4	35.4	51.4	32.4	52.2

Table 4. *Estimated coverage percentage of the most likely loss scenario for portfolio dimension d, sample size of large-loss n on Frank, Clayton, and Gumbel copulas with student-t marginals having degrees of freedom $\nu = 7$, at loss level $\ell = 0.1$ percentile and at confidence levels of 95% and 50% respectively for 500 iterations.*

Coverage

On the bottom line, the coverage results indicated that our adapted RST method reliably performed on portfolios having the join dependence structures as governed by Frank copula and Gumbel Copula. This is from the observation that the coverage percentages of our estimated important regions reach to the nominal coverages when the number of input large-loss observation (n) increased, regardless of the fact that the sizes of the regions become much smaller (see Figure B1-1:4, Appendix B, for references) when $n = 10$ increased from to $n = 50$. However, for Clayton copula, we observed the performance slightly inferior to those of Frank copula and Gumbel Copula, particular when the values of θ increased.

Estimated important regions for the most likely loss scenarios

We mentioned that apart from the coverage results, we also monitored the coverage behaviors of our estimated important regions to see whether there is any pattern of how the regions shifted away from the exact solutions of $\mathbf{z}^*(\ell)$ for each dependence structures. Since the shapes of the important regions follow the empirical likelihood profiles of the extreme lower tails, we also observed different characteristics given by the three copulas. The results are as followings.

- For portfolios with underlying Frank copula, the exact solutions $\mathbf{z}^*(\ell)$ located at around the center of the estimated important regions. The shapes of the important regions exhibited greater variance in the direction parallel to the loss function contours, and lower variance in the perpendicular to the loss function contours. The shapes and sizes did not significantly vary with the values of θ. However, at $\theta = 3$, the locations of the important regions slightly shifted upright toward the origin, comparing to the locations of the exact solutions.

- For portfolios with underlying Clayton copula, we observed three major differences from those of Frank copula and Gumbel Copula. First, most of the time, the exact solutions $\mathbf{z}^*(\ell)$ located at top right corner of the estimated important regions. Second, the shapes of the important regions significantly exhibited lower variance in direction parallel to the loss function contours, while exhibited higher variance in the direction perpendicular to the loss function contours. Third, the sizes of the

important regions were smaller comparing to those of Frank copula and Gumbel copula at the same confidence level and the same number of large-loss observations. Importantly, we observed all the three characteristics went stronger with the increasing value of θ.

- For portfolios with underlying Gumbel copula, similar to Frank copula, the exact solutions $\mathbf{z}^*(\ell)$ located at around the center of the estimated important regions. The shapes of the important regions also exhibited greater variance in the direction parallel to the loss function contours, and lower variance in the perpendicular direction. However, at $\theta = 3$, the locations of the important regions slightly moved to the bottom-left direction, comparing to the locations of the exact solutions. We observed that with the increasing value of θ, the variance in the direction parallel to the loss function contours reduced, while the variance in the perpendicular direction rose, in a sense that the shapes of the important regions became rounder.

In efforts to support our observation on the locations of the resulting important regions, we plotted the histograms of the estimated most likely loss solutions acquiring from our adapted RST method, separated by risk factor. After that, we ranked the exact solutions among the estimated results, and located them inside the histograms. The results for the three copula cases are in the same lines our observations above, as illustrated in Figure B2-1:3, Appendix B for bivariate risk factor portfolio examples.

4.2 Additional results

Form our main results, we got a rough picture on the effects of copula type and copula's parameter θ on the performance of our adapted RST method. In this section, we continued further to explore what factors which can influence the performance of our adapted RST method on non-elliptical bivariate risk factor portfolios.

Kendall's tau rank correlation

First, we selected *Kendall's tau*, which is one of the best-known rank correlation measures and suitable for multivariate data with underlying copula. We described the meaning of Kendall's tau and the relationship to our copula's θ in Appendix C.1. Again, we employed simulation to test the performance of our adapted RST model on the

varying values of Kendall's tau on Frank copula, Clayton copula, and Gumbel copula. The results are shown on Table 5 below.

Kendall's tau	Frank	Clayton	Gumbel
$\tau = 0.001$	96.2	94.8	96.2
0.1	94.2	86.2	97.0
0.2	95.8	77.4	96.4
0.3	94.2	87.6	95.8
0.4	93.2	83.8	97.6
0.5	90.8	83.8	92.8
0.6	88.2	81.2	94.4
0.7	89.2	57.2	93.0
0.8	88.4	13.4	91.2
0.9	85.4	9.0	83.4

Table 5. *Estimated coverage percentage of the most likely loss scenario for portfolio dimension $d = 2$, sample size of large-loss $n = 50$ on Frank copula, Clayton copula, and Gumbel copula, with student-t marginals having degrees of freedom $\nu = 7$, at loss level $\ell = 0.1$ percentile, and at confidence levels of 95% for 500 iterations, varying on Kendall's tau rank coefficient τ.*

Still, at first glance we observed the performance of our adapted RST method differed with different type of copulas. In the similar manner with the main results, our method can moderately perform on portfolios constructed through Frank copula and Gumbel copula. While in Clayton copula's cases, the performance swiftly dropped with the increasing values of Kendall's tau.

Since we already detected the patterns of how the estimated important regions of the most likely loss scenarios shifted away from the exact solutions when the values of the copula's θ increases for different types of copula, we can measure the Euclidean distances between the estimated RST solutions and the exact solutions. We found that by using the Kendall's tau parameter instead of the copula's θ, we could much better explain the resulting Euclidean distances in a sense of a linear relationship between those two factors (see Figure B3, Appendix B, for reference).

Tail dependence coefficient

Apart from Kendall's tau rank correlation which can give the information on the degree of concordance for the entire distribution, we also want to see whether if our performance varies over the tail-specific dependence measure. For multivariate data with underlying copula, tail dependence coefficient provides a handy measure to assess the degree of dependency separately for the lower tail and the upper tail (see definition and the relationship with and copula's θ in Appendix C.2).

Because Frank copula always gives 0 tail dependence coefficient for both tails (both upper-tail independent and lower-tail independent), we conducted our study on Clayton copula which provides solely the lower tail dependence, and Gumbel copula which provides solely the upper tail dependence on varying degrees of tail dependence coefficient. The results are shown on Table 6.

		Clayton			Gumbel
$\lambda_L = 0.001$	$\lambda_U = 0$	95.8	$\lambda_L = 0$	$\lambda_U = 0.001$	96.2
0.1	0	80.4	0	0.1	95.6
0.2	0	79.6	0	0.2	95.8
0.3	0	80.2	0	0.3	95.6
0.4	0	77.0	0	0.4	96.6
0.5	0	81.2	0	0.5	95.6
0.6	0	81.2	0	0.6	95.4
0.7	0	83.2	0	0.7	95.0
0.8	0	79.4	0	0.8	90.2
0.9	0	18.2	0	0.9	89.0

Table 6. *Estimated coverage percentage of the most likely loss scenario for portfolio dimension $d = 2$, sample size of large-loss $n = 50$ on Clayton copula and Gumbel copula, with student-t marginals having degrees of freedom $\nu = 7$, at loss level $\ell = 0.1$ percentile, and at confidence levels of 95% for 500 iterations, varying on tail dependence coefficients $\boldsymbol{\lambda}_L$ and $\boldsymbol{\lambda}_U$*

Now we observed that for our non-elliptically distributed portfolios, our adapted RST method gave stronger biases at the higher degree of lower tail dependence coefficients. On the other hand, while the lower tails remain independent, the varying degrees of the upper tail dependence coefficient did not significantly affect the performance of our adapted RST method.

5. Conclusion

Aggregating what we have learned from our results, as for the coverage performance of how our estimating important regions from our adapted RST method can capture the exact RST solutions, we observed the followings.

- Our adapted RST method can perform well on non-elliptical portfolios having the dependence structures constructed through Frank copula, a symmetric and tail independent copula, and portfolios with underlying Gumbel copula, a non-symmetric copula carrying dependent upper tail and independent lower tail.
- The performance of our adapted RST method slightly dropped when performing on portfolios having the joint dependence governed by Clayton copula, a non-symmetric copula carrying dependent lower tail and independent upper tail. However, we observed the performance significantly dropped at very high degree of lower tail dependence ($\lambda_L = 0.9$) .

Apart from the coverage performance, we also observed the characteristics of the resulting important regions which follow empirical profiles of extreme dependence structures. Importantly, we found that the shapes of the regions were the key factors affected the performance of our adapted RST model. For the three non-elliptical copulas applied in our work, we concluded the followings.

- For Gumbel copula, the shape of the important regions displayed higher variance in the direction parallel to the loss function contour and lower variance in the direction perpendicular to the loss function contour. The regions captured the most likely loss scenarios at around their centers. With increasing values of θ, locations of the regions slightly shifted to the bottom-left direction, at the same time that the shapes became rounder (as the variances in the parallel direction diminished, and the variances in the perpendicular direction rose). As a result, the regions persistently captured the exact RST solutions, making our adapted RST method generally performed well in portfolios with underlying Gumbel copula.
- For Frank copula, the shape of the important regions also exhibited greater variance in the direction parallel to the loss function contour, and lower variance in the perpendicular direction to the loss function contour. However, this characteristic

was consistent over the varying θ. While θ increases, the location of the regions slightly shifted upright toward the origin comparing to the location of the exact RST solutions.

- For Clayton copula, there are three main differences from those of Frank copula and Gumbel Copula. First, the shape of the important regions significantly exhibited lower variance in direction parallel to the loss function contour, while exhibited higher variance in the direction perpendicular to the loss function contour. Second, the regions captured the exact RST solutions at around their top right corner of the estimated important regions. Third, the size of the important regions was smaller comparing to those of Frank copula and Gumbel copula at the same confidence level and the same number of large-loss observations. As a result, the performance of our RST method on Clayton copula was inferior to those of Frank copula and Gumbel copula, especially when the degree of lower tail dependency increases.

Apart from the effects of copula as listed above, the sizes of the important regions also varied with the number of large-loss observations (n) and the confidence levels we applied during the estimation of the conditional mean via an EL algorithm. With the larger number of n being the input, the size of the important regions for the most likely loss solution becomes smaller in the sense that the model becomes more confident. The confidence level, on the other hand, provides the approaching boundary for the estimating coverage to reach the nominal coverage at big n, assuming the model performs well on that data structure. Rationally, the sizes of the important regions are bigger at the higher confidence levels.

In summary, our work illustrated the potential to adapt Glasserman et al. (2015)'s RST framework on non-elliptical portfolios by fitting an elliptical distribution (multivariate t) on portfolio risk factors. The performance testing on three non-elliptical copulas from Archimedean family (Frank, Clayton, and Gumbel) showed that the estimated most likely loss regions from our adapted RST method can robustly capture the exact solutions to RST in most of the cases.

We note that the performance of our adapted RST approach depends on (1) the dependence structure of extreme lower tail and (2) the quality of fitting elliptical

distribution into non-elliptical distribution. For example, the method can perform well with data underlying Frank (symmetric) copula and Gumbel (non-symmetric, lower tail independent) copula, while being slightly inferior with data underlying Clayton (non-symmetric, lower tail dependent) copula. We also note that our adapted RST method can give a drawback when performing on non-symmetric data having particularly strong dependence in the left tail comparing to the right tail dependence (for example, Clayton copula with lower tail dependent coefficient above 0.8).

On the bottom line, our RST method can still give a reliable guideline when estimating the most likely scenarios leading to losses exceeding a given threshold (solutions to RST) for portfolios with underlying non-elliptically distributed risk factor, which is useful for financial risk assessment considering there is not yet any proper RST framework has been developed on non-elliptical distribution framework. Also, as we summarized what we observed on the different characteristics of the resulting estimated most likely loss regions arriving from the three underlying copulas, practitioners who want to follow our practice can interpret and adjust the results, accordingly, based on the assumption on the joint dependence structure.

REFERENCES

Aas, K., C. Czado, A. Frigessi, and H. Bakken. 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics* 44 (2):182-198.

Agnihotri, S. 2017. Copula Approach: Correlation Between Bond Market and Stock Market, Between Developed and Emerging Economies. *International Journal for Innovative Research in Multidisciplinary Field* 3 (2):203-209.

Ang, A., and J. Chen. 2002. Asymmetric correlations of equity portfolios. *Journal of Financial Economics* 63 (3):443-494.

Breuer, T., and I. Csiszar. 2012. Systematic Stress Tests with Entropic Plausibility Constraints. *Journal of Banking & Finance* 37.

Breuer, T., M. Jandačka, J. Mencia, and M. Summer. 2012. A systematic approach to multi-period stress testing of portfolio credit risk. *Journal of Banking & Finance* 36 (2):332-340.

Chu, B. 2011. Recovering copulas from limited information and an application to asset allocation. *Journal of Banking & Finance* 35 (7):1824-1842.

Committee of European Banking Supervision. 2009. Guidelines on stress testing (CP32): European Banking Authority.

Donnelly, C., and P. Embrechts. 2010. The devil is in the tails. *Actuarial mathematics and the subprime mortgage crisis* 40 (1):1-33.

Financial Services Authority. 2009. Stress and scenario testing: feedback on CP08/24 and final rules. London, UK.

Flood, M. D., and G. G. Korenko. 2015. Systematic scenario selection: stress testing and the nature of uncertainty. *Quantitative Finance* 15 (1):43-59.

Garcia, R., and G. Tsafack. 2011. Dependence structure and extreme comovements in international equity and bond markets. *Journal of Banking & Finance* 35 (8):1954-1970.

Glasserman, P., C. Kang, and W. Kang. 2015. Stress scenario selection by empirical likelihood. *Quantitative Finance* 15 (1):25-41.

Grundke, P. 2012. Further recipes for quantitative reverse stress testing. *The Journal of Risk Model Validation* 6:81-102.

Hatherley, A., and J. Alcock. 2007. Portfolio construction incorporating asymmetric dependence structures: A user's guide. *Accounting and Finance* 47:447-472.

Hofert, M., I. Kojadinovic, M. Mächler, and J. Yan. 2018. *Elements of Copula Modeling with R*.

Kopeliovich, Y., A. Novosyolov, D. Satchkov, and B. Schachter. 2015. Robust Risk Estimation and Hedging: A Reverse Stress Testing Approach. *The Journal of Derivatives* 22:10 - 25.

Li, J., X. Zhu, C.-F. Lee, D. Wu, J. Feng, and Y. Shi. 2015. On the aggregation of credit, market and operational risks. *Review of Quantitative Finance & Accounting* 44 (1):161-189.

Liermann, V., and K. O. Klauck. 2010. Bank in stress test: Die Bank 52-55.

Longin, F., and B. Solnik. 1995. Is the Correlation in International Equity Returns Constant: 1960– 1990? *Journal of International Money and Finance* 14:3-26.

———. 2001. Extreme Value Correlation of International Equity Markets. *The Journal of Finance* 56:649-676.

McNeil, A., and A. Smith. 2012. Multivariate Stress Scenarios and Solvency. *Insurance Mathematics and Economics* 50.

Nelsen, R. B. 2006. *An introduction to copulas*. New York: Springer.

Owen, A. B. 2001. *Empirical Likelihood*. Boca Raton, Florida: Chapman & Hall/CRC.

Sklar, A. 1959. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris* 8:229-231.