



## Modeling non-normal corporate bond yield spreads by copula

Jong-Min Kim<sup>a</sup>, Dong H. Kim<sup>b</sup>, Hojin Jung<sup>c,\*</sup>

<sup>a</sup> Statistics Discipline, Division of Science and Mathematics, University of Minnesota at Morris, Morris, MN 56267, United States

<sup>b</sup> The James F. Dicke College of Business Administration, Ohio Northern University, 525 South Main Street, Ada, OH 45810, United States

<sup>c</sup> Department of Economics, Jeonbuk National University, Jeonju-si, Jeollabuk-do 54896, Republic of Korea

### ARTICLE INFO

#### Keywords:

Copula  
Equity volatility  
Spread

### ABSTRACT

This research focuses on modeling for how corporate bond yield spreads are affected by explanatory variables such as equity volatility, interest rate volatility,  $r$ , slope, rating, liquidity, coupon rate, and maturity. The existing literature assumes normality and linearity in the analysis, which is not the case in our sample. Thus, through a powerful and flexible copula approach, we study the dependence at the mean of the joint distribution by using the Gaussian copula marginal regression method and the dependence structure at the tails by using various copula functions. To our knowledge, this is the first application of the copula marginal regression model to bond market data. In addition, we employ several copula functions to test for the tail dependence between yield spreads and other explanatory variables. We find stronger tail dependence in the joint upper tail for the relation between equity volatility and yield spreads, among others. This result indicates the positive effect of equity volatility on yield spreads in the upper tail is greater than that in the low tail. This finding should be useful to practitioners, such as investors. By relying on better-fitting, more meaningful statistical models, this paper contributes to the extant literature on how corporate bond yield spreads are determined.

### 1. Introduction

Copula has been recently considered to be a flexible way of constructing the dependence of multivariate data in various application fields, such medical research, biostatistics, econometrics, finance, and actuarial science. One of the main reasons why the copula method is popular could be that it does not require independent and identical normal distribution assumption (Kojadinovic & Yan, 2010). Sklar's theorem (Sklar, 1959) states that for any  $n$  continuous random variables, a copula is able to couple  $n$ -univariate marginals into the  $n$ -dimensional distribution. The major advantage of a copula is that no assumption is needed for the variables to be independent or normal. Also, the method has no restrictions on the probability distributions. In particular, the copula regression is more appropriate than a traditional linear model when the response variable is not normally distributed.

During 2007–2009, the global financial system was in extreme turmoil, which drew financial economists' attention to corporate bond yield spreads. Dwyer and Tkac (2009) explore how fixed income markets reacted to the financial crisis. Guidolin and Tam (2013) also explore the effect of the financial crisis on yield spreads and document that bond risk premia rise during the financial crisis. By employing the GARCH-in-mean model, Choudhry (2016) examine how the global financial crisis influences yield spreads in European markets. Contessi, De Pace, and Guidolin (2014) analyze a set of eleven U.S. fixed income yield spreads and find that correlations of this set of yield spreads are much more significant than other times. They attribute their findings to a dramatic increase in exposures of fixed income securities to common risk factors, which was triggered by the financial crisis.

\* Corresponding author.

E-mail addresses: [jongmink@morris.umn.edu](mailto:jongmink@morris.umn.edu) (J.-M. Kim), [d-kim@onu.edu](mailto:d-kim@onu.edu) (D.H. Kim), [hojin.jung@jbnu.ac.kr](mailto:hojin.jung@jbnu.ac.kr) (H. Jung).

<https://doi.org/10.1016/j.najef.2020.101210>

Received 26 June 2019; Received in revised form 3 February 2020; Accepted 13 April 2020

Available online 20 April 2020

1062-9408/ © 2020 Elsevier Inc. All rights reserved.

In order to model corporate bond yield spreads, we apply an innovative copular method, which has been recently introduced into the literature. In particular, this study utilizes the Gaussian copula marginal regression (GCMR) combined with Weibull marginal distributions for variables (see [Masarotto & Varin \(2012\)](#) for some references). We also use various copula functions to test for the asymmetric tail dependences between yield spreads and other explanatory variables. Our sample includes both noncallable and callable corporate bonds. First, we find, in the sample of noncallable corporate bonds, that our GCMR analysis generally shows similar results to the traditional linear regression estimates. However, our GCMR analysis is an improvement over the traditional linear regression analysis in that the Akaike Information Criterion (AIC) has a smaller value for our regression model than the traditional regression model that the existing literature has employed. It should be also noted that the coefficient on the coupon rate is significantly positive as opposed to insignificance of the coefficient on the variable in the traditional regression model. The positive coefficient on the coupon rate is in line with [Elton, Gruber, Agrawal, and Mann \(2001\)](#) and [Longstaff, Mithal, and Neis \(2005\)](#). Investors with corporate bonds face tax disadvantages relative to those with Treasury bonds since interest payments on corporate bonds are subject to state taxes as opposed to Treasury bonds' state tax exemption. Since these tax disadvantages are greater for higher coupon bonds, investors require higher rates of return, thereby leading higher yield spreads.

Using various copula functions, we estimate the tail dependence coefficients for the pairs of yield spreads and other explanatory variables. For example, the relation between yield spreads and equity volatility in the sample of noncallable bonds shows greater upper tail dependence than low tail dependence. This finding represents that the positive impact of equity volatility on yield spreads is higher in the upper tail than that in the lower tail. Practitioners, such as corporate bond investors, can take advantage of this finding since, as pointed out by [Acharya and Carpenter \(2002\)](#), yield spreads are used to quote corporate bond prices in practice. For example, practitioners should note that equity volatility has a stronger effect on yield spreads during a financial crisis, resulting in extreme increases in yield spreads and equity volatility. Then we turn our attention to the callable bond sample. As in our noncallable bond sample, the AIC still has a smaller value for our GCMR analysis combined than the traditional linear regression model. Again, this result indicates that our model is a better-fitting model than our competing model. Unlike our analysis on noncallable bonds, the coefficient on maturity is significantly negative. This appears to be attributed to the fact that the longer the maturity of a bond, the more likely it has a call deferral period.

In addition, the tail dependence analysis on the callable sample provides different results for the equity volatility and yield spreads relation. The relation between yield spreads and equity volatility in the sample of callable bonds no longer shows greater upper tail dependence than lower tail dependence, which is observed in the sample of noncallable bonds. We relate this result to the argument of [Acharya and Carpenter \(2002\)](#), which is that the exercise of the call option can be destroyed by default risk. In the upper tail, extremely high equity volatility likely increases the default risk. This effect weakens the strongly positive effect of equity volatility on yield spreads in the upper tail. This paper contributes to the finance literature by suggesting employing better-fitting and more meaningful statistical models. Unlike other studies assuming normality and linearity, we employ a better-fitting model evidenced by the lower AIC values. In addition, our findings have important implications for practitioners such as hedgers and policymakers in the corporate bond market.

As [Acharya and Carpenter \(2002\)](#) note, yield spreads can be widely and practically used to quote corporate bond prices. Since the 2007–2009 global financial crisis, considerable attention has been paid to corporate bond yield spreads. More specifically, it has attracted a great deal of attention from researchers, practitioners, and policymakers to model and forecast corporate bond yield spreads. Our findings are particularly important to them and will benefit them by demonstrating that the assumptions of normality and linearity are violated, which the previous literature has overlooked. In addition, we find that the degree of dependence for the pairs of (1) equity volatility and yield spreads and (2) interest rate volatility and yield spreads is greater in the upper tail than in the lower tail. The practical implication of this finding is that the aforementioned professionals should place more focus on the effects of equity volatility and interest rate volatility on yield spreads when they face extreme changes in equity volatility and interest rate volatility.

The contribution of this study is twofold. First, we find potential determinants that can explain corporate bond spreads. Second, we investigate the relationship between each individual determinant and the yield spreads. Our empirical analysis is conducted by a powerful and flexible approach, copula. In particular, we study the dependence at the mean level of the joint distribution by using the GCMR method and the dependence structure at the tails by using both various symmetric and asymmetric copula functions. The tail dependence estimation is executed in order to uncover the dependence structure between pairs considered during the extreme financial events.

The remainder of the paper is structured as follows. The next section provides our data selection process. The next section provides a literature review of related work. Section 3 introduces our various copula functions and the GCMR method. We discuss our empirical results on the relationship between yield spreads and other explanatory variables in Section 4. Finally, Section 5 presents concluding remarks.

## 2. Literature review

The invariance of copulas under monotonic transformation is particularly popular in finance. [Kim et al. \(2008\)](#) propose partial correlation with copula modeling. [Kim, Jung, Choi, and Sungur \(2011\)](#) apply a copula function instead of a Bayesian network for constructing directional dependence of genes. In addition, [Kim and Kim \(2014\)](#) suggest using a variation of the standard copula method for modeling directional co-movements of genes. [Kim and Jung \(2016\)](#) employ a copula DCC-GARCH model for predicting volatility. [Pourkhanali, Kim, Tafakori, and Fard \(2016\)](#) present a vine-copula model of systemic risk to analyze the complex interdependencies between different borrowers. In this study, we use the copula method to examine how the target response variable

(corporate bond yield spreads) is related to predictor variables (equity volatility, interest rate volatility,  $r$ , slope, rating, liquidity, coupon rate, and maturity). From our data analysis, we observe that the dependent and some independent variables are not normally distributed, and they do not hold exact linear relations. However, the existing literature on how corporate bond yield spreads are determined assumes normality and linearity.

To relax normality and linearity assumptions, copula-based methods have been explicitly recognized and employed in the finance literature. For example, [Junker, Szimayer, and Wagner \(2006\)](#) find that a nonlinear relationship in the term structure of interest rates. They investigate tail dependence between factors spanning United States-Treasury yield curve, and a transformed Frank copula turns out to be the best-fitting model in the study. [Al Janabi, Hernandez, Berger, and Nguyen \(2017\)](#) extend a framework for time-varying liquidity-adjusted Value-at-Risk by replacing linear correlations with nonlinear dependence structure. In the framework, they incorporate nonlinear dependence between stock market indices and commodity market indices, such as gold and oil prices by using dynamic conditional correlation  $t$ -copula. They show that the proposed approach provides more efficient frontiers than other competing models. [Mensi, Hammoudeh, Shahzad, and Shahbaz \(2017\)](#) utilize the variational mode decomposition method with copulas to examine dependence structure between oil and major stock market indices of the U.S., Canada, Europe and Pacific basin at short- and long-run investment horizons. They find evidence of tail dependence between the return series, such that risk spillovers are stronger in the long-run than in the short-run investment horizon. [Kahlert et al. \(2017\)](#) concentrate extreme risk factor co-movements of medium-term Eurozone government bonds and find significant upper tail dependence between credit and market liquidity spread changes during the recent financial market crises. [Nguyen and Liu \(2017\)](#) focus on the dynamic dependence among government bonds, gold, and oil by using the time-varying symmetrized Joe-Clayton copula function. Their empirical analysis indicates that government bonds are generally considered as safe haven assets against the equity market index.

A number of studies investigate the dependence structure between various financial assets by using copulas. For example, [Serra and Gil \(2012\)](#) focus on crude oil and biodiesel markets. [Aloui, Aïssa, and Nguyen \(2013\)](#), [Aloui and Aïssa \(2016\)](#), [Reboredo \(2012\)](#), [Wu, Chung, and Chang \(2012\)](#), and more studies investigate the dependence between oil and foreign exchange markets by using a family of copula functions. The relationship between the equity and foreign exchange markets is uncovered by [Ning \(2010\)](#) and [Wang, Wu, and Lai \(2013\)](#). Bond markets have also been extensively studied in the literature; e.g., [Reboredo and Ugolini \(2015\)](#), [Philippas and Siriopoulos \(2013\)](#), [Wu and Lin \(2014\)](#). A time-varying copula approach has also been popular recently. [Wang, Chen, and Huang \(2011\)](#) find the dynamic dependence structure between stock returns across countries. [Patton \(2006\)](#) explores asymmetric dependence and time-varying dependence between exchange rates before and after the introduction of the Euro. In addition, [Jammazi, Tiwari, Ferrer, and Moya \(2015\)](#) also conduct a similar study, but focusing on stock-bond association in European countries.

Spreads have received considerable attention in the literature. [Covitz and Downing \(2007\)](#) focus on commercial paper spreads and show that liquidity is the most important factor in determining very short-term corporate yield spreads. [Gilchrist and Zakrajšek \(2012\)](#) note that credit spreads can be used to predict the economic outlook for the future. They construct their own credit spread index and document that their index is an improvement over the Baa-Aaa spread for predicting future economic activity. [Ang, Boivin, Dong, and Loo-Kung \(2011\)](#) find that an unexpected rise in the Fed's response to inflation raises short-term rates, thereby leading to an increase in the term spread. [Boyarchenko, Fuster, and Lucca \(2019\)](#) propose a new model for pricing mortgage-backed securities (MBS) and they identify key determinants of variations in MBS spreads.

### 3. Measuring dependence using copulas

#### 3.1. Copula function

[Frey and McNeil \(2001\)](#) and [Li \(1999\)](#) introduce copula functions in the financial literature to construct joint multivariate distributions. [Sklar \(1973\)](#) shows that any bivariate distribution function,  $F_{XY}(x, y)$ , can be represented by the marginal distributions of  $F_X(x)$  and  $F_Y(y)$  and a copula function of  $C$  determined in  $[0, 1]^q$ . Thus, a copula function represents the dependence of the marginal distributions. A bivariate copula function holds the following three properties: (i)  $C(\vartheta, 0) = C(0, \xi) = 0$  for all  $\vartheta, \xi \in [0, 1]$ , (ii)  $C(\vartheta, 1) = C(1, \vartheta) = \vartheta$  for all  $\vartheta \in [0, 1]$ , and (iii) given  $\vartheta_1 \leq \vartheta_2$  and  $\xi_1 \leq \xi_2$ ,  $C(\vartheta_1, \xi_1) - C(\vartheta_1, \xi_2) - C(\vartheta_2, \xi_1) + C(\vartheta_2, \xi_2) \geq 0$ , for all  $\vartheta_1, \vartheta_2, \xi_1, \xi_2 \in [0, 1]$ . See [Kim and Jung \(2017\)](#) for more details about the properties of copulas.

A  $q$ -dimensional copula function  $C: [0, 1]^q$  satisfies that  $F(x_1, \dots, x_q) = C(F(x_1), \dots, F(x_q))$ . Then, there exists the density functions of  $F$  and  $C$  such that

$$f(x_1, \dots, x_q) = c(F_1(x_1), \dots, F_q(x_q)) \prod_{i=1}^q f_i(x_i)$$

$$c(\vartheta_1, \dots, \vartheta_q) = \frac{f(F_1^{-1}(\vartheta_1), \dots, F_q^{-1}(\vartheta_q))}{\prod_{i=1}^q f_i(F_i^{-1}(\vartheta_i))}$$

where  $f_i$  and  $F_i^{-1}$  represents the marginal density function and the quantile function of the margins, respectively. Finally, the Gaussian copula density function, which is the most popular among various copular functions is defined by

$$C^{Ga}(\vartheta_1, \dots, \vartheta_q) = \Psi(\Psi^{-1}(\vartheta_1), \dots, \Psi^{-1}(\vartheta_q)), \quad \text{where } z_1 = \Psi(\vartheta_1), \dots, z_q = \Psi(\vartheta_q). \quad (1)$$

A more extensive review of a Gaussian copula function is given in [Kim et al. \(2011\)](#).

### 3.2. Gaussian copula regression method

Our study employs the Gaussian copula regression method to measure the relationship between financial variables. The marginal cumulative distribution for  $x_i$  is  $F(\cdot|x_i)$ . Then, in the Gaussian copula regression, the joint cumulative distribution function is written as follows:

$$\Pr(Y_1 \leq y_1, \dots, Y_q \leq y_q) = \Psi_n(\epsilon_1, \dots, \epsilon_q; \mathbf{P}),$$

where  $\epsilon_i = \Psi^{-1}\{F(y_i|x_i)\}$ . For the Gaussian copula correlation matrix  $\mathbf{P}$ , the standard normal univariate and multivariate cumulative distribution functions are  $\Psi(\cdot)$  and  $\Psi_n(\cdot; \mathbf{P})$ . See [Masarotto and Varin \(2012\)](#) and [Song \(2000\)](#) for more details. In particular, [Masarotto and Varin \(2012\)](#) introduce an alternative formulation of the multivariate Gaussian copula as follows:

$$Y_i = g(\mathbf{x}_i, \epsilon_i),$$

where  $\epsilon_i$  indicates a stochastic error which follows a multivariate normal distribution. Note that  $g(x_i, \epsilon_i) = F^{-1}\{\Psi(\epsilon_i|x_i)\}$  is assumed in GCMR.

With Gaussian copula functions, Weibull marginal distributions for independent variables are considered in our GCMR analysis. The probability density function of the Weibull distribution for  $x \geq 0$  and  $\kappa, \omega > 0$  is

$$f(x) = \frac{\omega}{\kappa} \left( \frac{x - \mu}{\kappa} \right)^{(\omega-1)} \exp \left( - \left( \frac{x - \mu}{\kappa} \right)^\omega \right),$$

where  $\omega$  and  $\mu$  are the shape and location parameters, respectively. The cumulative distribution function of the Weibull distribution for  $x \geq 0$  and  $\omega > 0$  is given as

$$F(x) = 1 - e^{-(x^\omega)}.$$

This Weibull distribution is popularly used in data analysis because of its flexibility of modeling different types of datasets, including skewed data typically observed in the financial datasets.

### 3.3. Copula tail dependence

Tail dependence structure is used to measure the propensity of the extreme co-movement in the lower or upper joint tails of two random variables. Unlike linear correlation, tail dependence by copulas is not required to assume multivariate normality for marginal distributions. The coefficients of left (lower) and right (upper) tail dependence of  $(X, Y)$  are defined in terms of copula by [Nelsen \(2007\)](#) as:

$$\begin{aligned} \tau^L &= \lim_{\varepsilon \rightarrow 0} \text{Prob}[\vartheta \leq \varepsilon | \xi \leq \varepsilon] = \lim_{\varepsilon \rightarrow 0} \text{Prob}[\xi \leq \varepsilon | \vartheta \leq \varepsilon] \\ \tau^R &= \lim_{\delta \rightarrow 1} \text{Prob}[\vartheta > \delta | \xi > \delta] = \lim_{\delta \rightarrow 1} \text{Prob}[\xi > \delta | \vartheta > \delta] \end{aligned}$$

if  $\tau^R \in [0, 1]$  and  $\tau^L \in [0, 1]$  exist. The bivariate Gaussian copula can be written as follows:

$$C^{Ga}(\vartheta, \xi | \rho) = \Psi(\Phi^{-1}(\vartheta), \Psi^{-1}(\xi)), \quad (2)$$

where  $\Psi^{-1}(\vartheta)$  and  $\Psi^{-1}(\xi)$  represents standard normal quantile functions and  $\rho$  is the correlation between the two random variables.  $C^{Ga}$  has zero dependence in tails such that  $\tau^L = \tau^R = 0$ . The Student- $t$  copula is defined by

$$C^{St}(\vartheta, \xi | \rho, \nu) = H(h^{-1}(\vartheta), h^{-1}(\xi)), \quad (3)$$

where  $h^{-1}(\vartheta)$  and  $h^{-1}(\xi)$  are the Student- $t$  quantile distribution functions and  $\nu$  indicates the degree-of-freedom parameter. Unlike the Gaussian copula, it has nonzero tail dependence,  $\tau^L = \tau^R > 0$ . For  $\alpha > 0$  and  $\alpha \neq 1$ , the Plackett copula is defined as:

$$C^{Pl}(\vartheta, \xi | \alpha) = \frac{[1 + (\alpha - 1)(\vartheta + \xi)] - \sqrt{[1 + (\alpha - 1)(\vartheta + \xi)]^2 - 4\vartheta\xi(\alpha - 1)}}{2(\alpha - 1)}, \quad (4)$$

where  $\alpha$  is the cross product ratio.  $C^{Pl}$  has lower and upper tail dependencies of  $\tau^L = 0$  and  $\tau^R = 0$ . The last symmetric copula function that our study considers is the Frank copula. For  $\theta \in (-\infty, \infty) \setminus \{0\}$ , the copula function is given by

$$C^{Fr}(\vartheta, \xi | \psi) = -\frac{1}{\psi} \log \left( 1 + \frac{(e^{-\psi\vartheta} - 1)(e^{-\psi\xi} - 1)}{e^{-\psi} - 1} \right). \quad (5)$$

$C^{Fr}$  has zero dependence in tails such that  $\tau^L = \tau^R = 0$ .

We also consider Clayton, Gumbel and SJC copulas to investigate possible asymmetric tail dependence. For a review of various copula functions, we refer to [Joe \(1997\)](#) and [Patton \(2006\)](#). We can define the Clayton copula as

$$C^{Cl}(\vartheta, \xi | \theta) = \max \left\{ (\vartheta^{-\theta} + \xi^{-\theta} - 1)^{\frac{1}{\theta}}, 0 \right\}. \quad (6)$$

$C^{Cl}$  is asymmetric such that the coefficient of the left tail is higher than zero in the right tail. Similarly, the Gumbel copula is given

by

$$C^{Gu}(\vartheta, \xi|\varphi) = \exp\left(-((- \log \vartheta)^\varphi + (- \log \xi)^\varphi)^{\frac{1}{\varphi}}\right). \quad (7)$$

The right tail is higher than zero in the left tail in  $C^{Gu}$ . Patton (2006) proposes the symmetrized Joe-Clayton (SJC) copula, which enables to capture asymmetric tail dependence and symmetric dependence as a special case,  $\tau^L = \tau^R$ . This copula is given by

$$C^{SJC}(\vartheta, \xi|\tau^L, \tau^R) = 0.5 \times [C^{JC}(\vartheta, \xi|\tau^L, \tau^R) + C^{JC}(1 - \vartheta, 1 - \xi|\tau^L, \tau^R) + \vartheta + \xi - 1], \quad (8)$$

where  $C^{JC}$  is the Joe-Clayton copula (Joe, 1997), defined as  $C^{JC}(\vartheta, \xi|\tau^L, \tau^R) = 1 - (1 - \{[1 - (1 - \vartheta)^\gamma]^{-\gamma} + [1 - (1 - \xi)^\gamma]^{-\gamma} - 1\}^{-1/\gamma})^{1/\eta}$ , where  $\eta = 1/\log_2(2 - \tau^R)$ ,  $\gamma = -1/\log_2(\tau^L)$ ,  $\tau^L \in (0, 1)$ , and  $\tau^R \in (0, 1)$ .

## 4. Empirical result

### 4.1. Data

Our transaction data from 2003 to 2009 are collected from the Trade Reporting and Compliance Engine (TRACE). This is the same database used by Kim and Stock (2014). Following Kim and Stock (2014), we employ several selection criteria. First, we drop canceled, corrected, and repeated interdealer trades from the sample. Then we exclude bonds that did not trade during the five days between the five business days before the last transaction date and the last transaction date of the month. We take the volume-weighted average values by using transaction prices on the last transaction date each month and use them as the month-end bond prices.

Following Kim and Stock (2014), we also consider bond characteristics extracted from the Fixed Income Securities Database (FISD). Our sample data do not include floating-rate bonds, putable bonds, convertible bonds, make-whole bonds, or bonds with sinking fund provisions. Furthermore, we eliminate bonds with an unusual frequency of coupon payments as in Elton et al. (2001). Following the variable construction used by Duffee (1999) and Eom, Helwege, and Huang (2004), we drop bonds with maturity less than a year due to the fact that they are unlikely to trade. We employ the H.15 release of the Federal Reserve System to extract the Treasury constant maturity yields. Yield spreads are measured as the daily corporate bond yield minus the Treasury constant maturity rate with the same maturity.

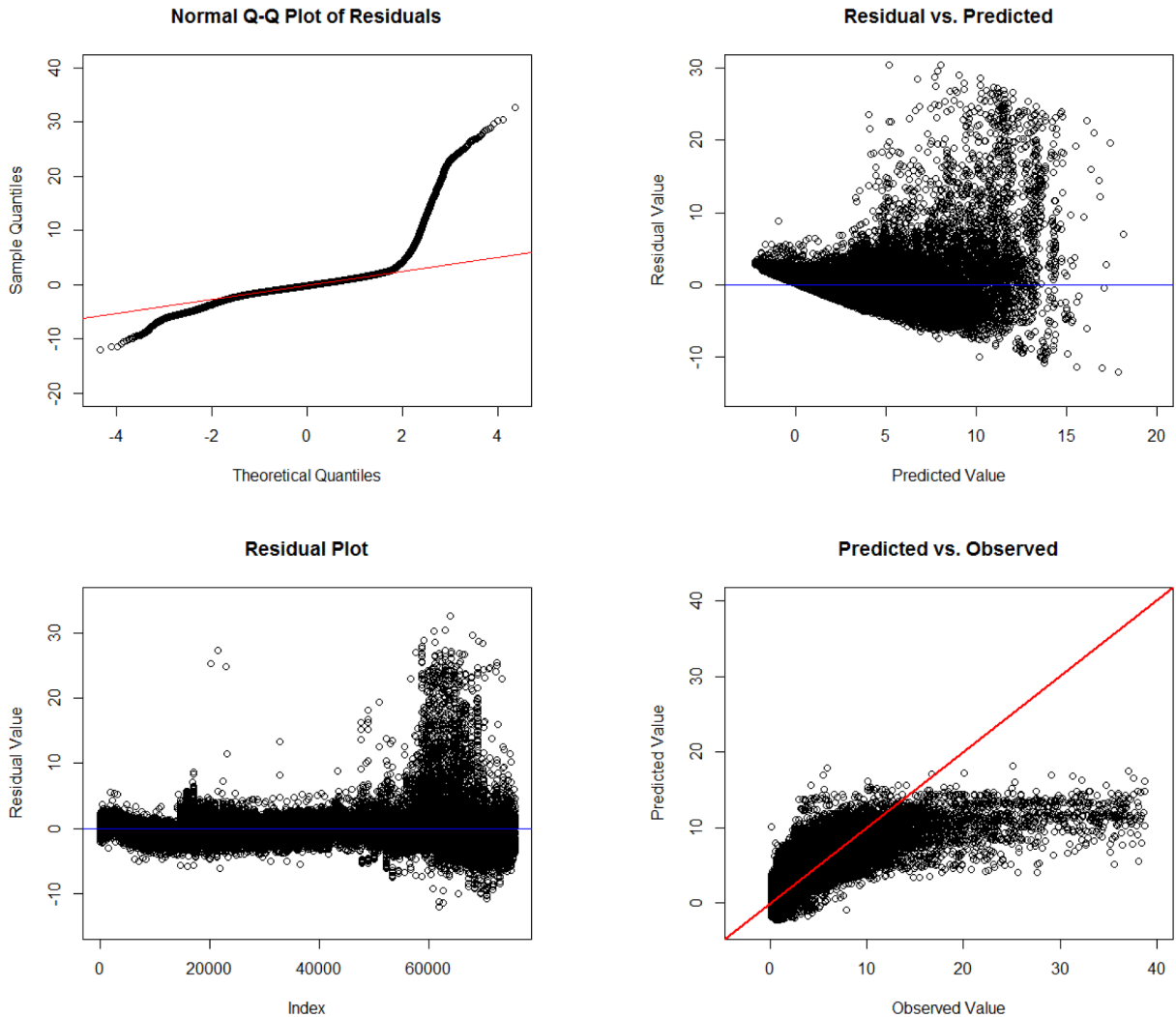
We also follow Kim and Stock (2014) to construct explanatory variables. Campbell and Taksler (2003) show that most variations in yield spreads can be explained by equity volatility. Equity volatility is constructed by using the standard deviation of excess returns on a daily basis over the one month proceeding the bond transaction date. Kim and Stock (2014) find that interest rate volatility increases yield spreads. Interest rate volatility is measured by calculating the standard deviation of the one-month Treasury constant maturity rate over the one month proceeding the bond transaction date. Collin-Dufresne and Goldstein (2001) and Longstaff and Schwartz (1995) theoretically suggest that short-term interest rates are negatively correlated with yield spreads, and Bao, Pan, and Wang (2011), Chen, Lesmond, and Wei (2007), Duffee (1999), Longstaff and Schwartz (1995) empirically document this negative correlation between them. We consider the one-month Treasury constant maturity rate as short term interest rate ( $r$ ) in this study. As in Breeden (2011), Ederington and Stock (2002), Estrella and Hardouvelis (1991), Estrella and Mishkin (1996), the slope of the yield curve reflects the market's expectations about the economic outlook for the future. Then it likely decreases yield spreads. We measure Slope by computing the difference between the 10-year constant maturity rate and its 1-year counterpart. The credit rating has been documented to be the most common measure of default risk. There should be a negative correlation between credit quality and yield spreads. We assign Rating a cardinalized S&P rating, where AAA is coded as 1, ..., D is coded as 22.<sup>1</sup> Bao et al. (2011), Chen et al. (2007), Güntay and Hackbarth (2010), stress the importance of liquidity in the determination of yield spreads. We measure Liquidity as the ratio of the number of days on which the bond traded over the one month preceding the bond trading day to the total number of business days for the same month. We also control for coupon rate and maturity. Each bond-month transaction is included in the sample if it has data on all explanatory variables. Our restriction results in a sample of 126,041 bond-month transactions, of which 75,742 are noncallable bonds and 50,299 are callable bonds.

A linear model imposes the following four fundamental assumptions regarding the data:

- Each data point is independent from every other point.
- The residuals follow a normal distribution.
- Homogeneity of variance is assumed across the fitted values.
- There is a linear relationship between the dependent and the independent variables.

The previous studies, including Campbell and Taksler (2003) and Kim and Stock (2014), truncate the top and bottom 1% of yield spreads to satisfy the above assumptions in their data analysis. However, the trimmed sample by the ad hoc approach still violates the strong assumptions. In this study, we still use the same dataset employed by Kim and Stock (2014), but we drop only the missing values from the complete bond-month transactions for noncallable and callable bonds of 222,440 during the sample period. Thus, our study considers the entire sample available, including outliers. The sample consists of 126,041 different bond-month transactions

<sup>1</sup> Following Chen et al. (2007), we exclude bonds for which S&P ratings are unavailable.



**Fig. 1.** Checking linear model assumptions. Note: This figure displays the plots for the normal Q-Q, for the residual and predicted values, for the residual values by a linear model, and for the predicted and observed values (left, right, up, and down, respectively).

including 75,742 noncallable bonds and 50,299 callable bonds.

Fig. 1 presents whether our sample dataset satisfies the linear model assumptions. The normal quantile-quantile (Q-Q) plot is typically used to check the normal distribution of the residuals. Note that residuals are measured as the difference between the observed  $y$  and model-predicted  $\hat{y}$ . The Q-Q plot in Fig. 1 shows that the residuals at the lower and upper quantiles deviate from the straight line, which strongly implies that our sample data violate the normality assumption. The top-right plot depicts the residuals and the predicted values. To satisfy the assumption of linearity, we need to ensure that residuals are not too far away from 0. But this plot provides strong evidence that our sample data violate the assumption of linearity. The plot clearly displays that residuals are positively and negatively dispersed due to the unexpected financial crisis event. The plot is also used to assess if the homoscedasticity assumption is met. We find that the assumption of homoscedasticity is also violated in that there is a pattern in the residuals, which are not equally spread around the  $y = 0$  line. The bottom-left plot presents that the residual values are increasing, which indicates that our sample data do not follow the independence assumption. It also shows the extremely large positive residuals, which means the unexpected observed values appeared over the crisis periods from mid-2007 to mid-2008. The last plot in Fig. 1 depicts the predicted and observed values to show the model fit. The plot indicates that a linear model does not fit well with our sample data and a nonlinear regression model is a better alternative in our empirical analysis.

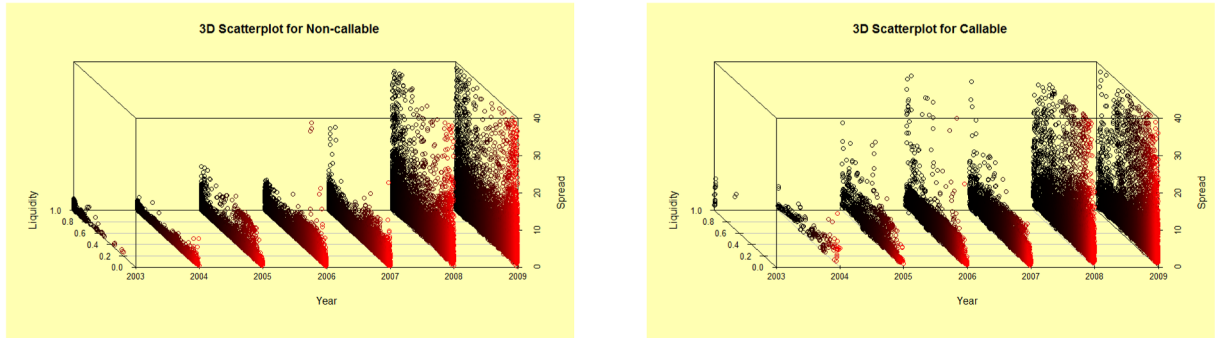
The violations of the linear model assumptions in Kim and Stock (2014) motivate us to employ the copula regression model as an alternative. Table 1 shows the advantages of using the copular method compared with the multivariate linear regression model employed in Kim and Stock (2014) (see Parsa & Klugman, 2011; Masarotto & Varin, 2012 for further discussion).

Fig. 2 displays the 3D scatter plots for Liquidity and Spread of noncallable and callable corporate bonds over the sample period. Each bond type shows a different relationship between the two variables before and after the economic crisis. For example, we



**Table 1**  
Advantages by using Copula regression versus linear regression model.

Assumption	Multivariate Linear Model in Kim and Stock (2014)	Copula Regression
Normality	Required	Free
Linearity	Required	Free
Independence	Required	Free
Homoscedasticity	Required	Free



**Fig. 2.** 3D scatter plots for Liquidity and Spread over the sample period.

observe a negative relationship between the two variables for noncallable corporate bonds over the sample period, which is consistent with our expectation. However, there is a negative relationship before 2008 while a positive relationship appears for callable corporate bonds from 2008 to 2009. The scatter plot on the right panel shows the majority of sample observations during the crisis. A possible explanation is that as interest rates rise, call features become costly but benefit from hedging the interest rate risk (Booth, Gounopoulos, & Skinner, 2014). To verify our findings, we conduct our empirical analysis by using the GCMR model in the next section.<sup>2</sup>

#### 4.2. Noncallable data case

Table 2 presents summary statistics and distributional characteristics for the variables used in our empirical analysis. It shows that over the sample period, on average, Spread is 2.48 while Equity volatility is close to the zero mean. The average values of Interest rate volatility ( $\sigma_r$ ), Rating and Maturity are 0.53%, 4.62 and 5.96, respectively. The table also shows that the variables are skewed and exhibit excess kurtosis, especially for Equity volatility, Maturity, Spread, and Rating. In particular, the positive skewness indicates that the variables have fatter upper tail density functions. Note that if the kurtosis statistic is higher than three, then the random variable is not normally distributed. This article considers the regression analysis to investigate the degree of interdependence among the financial variables.

$$\text{Spread}_t = \beta_0 + \beta_1 \times \text{Equityvolatility}_t + \beta_2 \times \text{Interstratevolatility}_t + \beta_3 \times r_t + \beta_4 \times \text{Slope}_t + \beta_5 \times \text{Rating}_t + \beta_6 \times \text{Liquidity}_t + \beta_7 \times \text{Couponrate} + \beta_8 \times \text{Maturity} + e_t,$$

where  $e_t = \phi e_{t-1} + \omega_t$ , and  $\omega_t \sim iidN(0, \sigma_\omega^2)$  for GCMR.

Table 3 presents the estimation results. The first column is our baseline specification with a linear regression model, and the second is the same specification with the GCMR model. Copulas can split marginals from the dependence structure, and a Weibull marginal distribution is considered in this study. In addition, we specify the ARMA (1,0) error correlation structure. Thus,  $e_t$  follows an AR(1) process as shown in the regression equation. Note that the AR(1) model can be used as the Gaussian copula correlation matrix. In particular, we utilize the *gcmr* R package and perform inference by using a likelihood approach. More details for the computational process can be found in Masarotto and Varin (2012). The estimated coefficients can be interpreted as the strength of dependence between each covariate and Spread. There are two differences between the first and second columns. First and most importantly, the AIC value is smaller in the second column than in the first column. Since a model with the smallest AIC value among competing models is the best fitting model, the GCMR model is a better- fitting model than the traditional linear regression model.

<sup>2</sup> As an alternative method, wavelet analysis can be used to capture the co-movements between our sample variables in the time-space domain. However, as Parikh and Baraniecki (1996) point out, the implementation of the approach is computationally intensive. Also, there is no consensus so far in the literature that the discrete wavelet transform is more efficient than the others. In our empirical analysis, as a result, we favor the copula method, which is easy to apply and does not require the following assumptions: multivariate normality, linear relationship, and homoscedasticity of residuals.

**Table 2**

Summary statistics: noncallable case.

	Obs.	Mean	Median	Minimum	Maximum	St.D	Skewness	Kurtosis
Spread	75,742	2.48	1.26	0.11	38.80	3.56	4.48	31.27
Equity volatility	75,742	0.02	0.01	0.00	0.15	0.02	2.81	12.06
Interest rate volatility	75,742	0.53	0.53	0.05	1.28	0.31	0.50	2.76
$r$	75,742	2.71	2.79	0.00	5.27	1.90	-0.12	1.47
Slope	75,742	1.11	0.77	-0.47	3.39	1.23	0.36	1.67
Rating	75,742	6.42	6.00	1.00	22.00	3.42	1.04	4.62
Liquidity	75,742	0.51	0.47	0.00	1.00	0.34	0.13	1.53
Coupon rate	75,742	6.03	6.00	0.94	12.40	1.55	0.29	2.75
Maturity	75,742	5.96	3.80	1.00	91.28	6.32	4.20	40.84

**Table 3**

Estimation results: noncallable case.

	Linear Regression	Gaussian Copula Regression
Equity volatility	100.300*** (0.601)	21.699*** (0.166)
Interest rate volatility	1.480*** (0.032)	0.788*** (0.008)
$r$	-1.282*** (0.020)	-0.320*** (0.005)
Slope	-1.702*** (0.030)	-0.321*** (0.007)
Rating	0.323*** (0.003)	0.104*** (0.001)
Liquidity	-0.371*** (0.025)	-0.349*** (0.006)
Coupon rate	-0.005 (0.006)	0.021*** (0.001)
Maturity	-0.000 (0.001)	0.007*** (0.000)
Constant	3.295*** (0.105)	0.357*** (0.025)
Log likelihood	170,690	93,316
AIC	341,405	186,652

\*\*\*, \*\*, and \* indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.

Second, every coefficient estimate from the GCMR model is statistically significant at the 1% significance level. The two variables that are not significant in the first column, but significant in the second column are Coupon rate and Maturity. [Elton et al. \(2001\)](#) and [Longstaff et al. \(2005\)](#) suggest that coupon rates should be positively correlated with yield spreads due to taxation. Investors pay state taxes on interest earned on corporate bonds, but not on Treasury bonds. This tax disadvantage is greater for higher coupon rate corporate bonds, thus investors demanding a higher rate of returns on these bonds. Therefore, coupon rates should be positively related to yield spreads. Our result is consistent with this line of reasoning. A significantly positive sign on Coupon rate in the first column, in contrast to the insignificance in the second column, is also evidence that the GCMR model is an improvement over the traditional linear regression model. In addition, Maturity is significantly positive in the second column while it is not in the first column. A positive sign on Maturity can be explained if the corporate bond has a steeper yield curve than that of the government bond.

Since GCMR is a better fitting model, our interpretations of other variables are based on the estimation results reported in the second column of [Table 3](#).<sup>3</sup> Results on other variables are consistent with the corporate finance literature. As in [Campbell and Taksler \(2003\)](#) and [Kim and Stock \(2014\)](#), equity volatility and interest rate volatility are positively correlated with yield spreads, respectively. As in [Duffee \(1998\)](#) and [Longstaff and Schwartz \(1995\)](#),  $r$  has a negative sign. A negative sign on the slope can be explained by the reflection of the slope of the yield curve on how positively the market views the future economy, suggested by [Breedon \(2011\)](#), [Ederington and Stock \(2002\)](#), [Estrella and Hardouvelis \(1991\)](#), [Estrella and Mishkin \(1996\)](#). A positive sign on Rating means that there is a negative association between credit quality and yield spreads. A negative sign on Liquidity is consistent with [Bao et al. \(2011\)](#), [Chen et al. \(2007\)](#), [Güntay and Hackbarth \(2010\)](#), [Rossi \(2014\)](#).

We also investigate the co-movements between corporate bond yield spreads and other explanatory variables by using various

<sup>3</sup> [Parsa and Klugman \(2011\)](#) recommend copula regression as an alternative to a linear model, such as ordinary least squares (OLS) and generalized linear model (GLM) for actuarial applications.



**Table 4**

Statistics for copulas between standardized uniform pairs: noncallable case.

Copula type	Equity volatility-Spread		Interest rate volatility-Spread		r-Spread		Slope-Spread	
	$\tau^L$	$\tau^R$	$\tau^L$	$\tau^R$	$\tau^L$	$\tau^R$	$\tau^L$	$\tau^R$
Gaussian	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Clayton	<b>0.437</b>	<b>0.000</b>	<b>0.087</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.179</b>	<b>0.000</b>
Rotated Clayton	<b>0.000</b>	<b>0.628</b>	<b>0.000</b>	<b>0.262</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.240</b>
Plackett	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Frank	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Gumbel	<b>0.000</b>	<b>0.558</b>	<b>0.000</b>	<b>0.290</b>	<b>0.000</b>	<b>0.122</b>	<b>0.000</b>	<b>0.291</b>
Rotated Gumbel	<b>0.495</b>	<b>0.000</b>	<b>0.234</b>	<b>0.000</b>	<b>0.122</b>	<b>0.000</b>	<b>0.277</b>	<b>0.000</b>
Student's <i>t</i>	0.022	0.022	0.000	0.000	0.000	0.000	0.000	0.000
SJC	<b>0.009</b>	<b>0.637</b>	<b>0.000</b>	<b>0.305</b>	<b>0.000</b>	<b>0.000</b>	<b>0.080</b>	<b>0.232</b>

Copula type	Rating-Spread		Liquidity-Spread		Coupon rate-Spread		Maturity-Spread	
	$\tau^L$	$\tau^R$	$\tau^L$	$\tau^R$	$\tau^L$	$\tau^R$	$\tau^L$	$\tau^R$
Gaussian	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Clayton	<b>0.169</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.134</b>	<b>0.000</b>	<b>0.126</b>	<b>0.000</b>
Rotated Clayton	<b>0.000</b>	<b>0.381</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.041</b>	<b>0.000</b>	<b>0.033</b>
Plackett	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Frank	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Gumbel	<b>0.000</b>	<b>0.354</b>	<b>0.000</b>	<b>0.122</b>	<b>0.000</b>	<b>0.169</b>	<b>0.000</b>	<b>0.159</b>
Rotated Gumbel	<b>0.296</b>	<b>0.000</b>	<b>0.122</b>	<b>0.000</b>	<b>0.203</b>	<b>0.000</b>	<b>0.199</b>	<b>0.000</b>
Student's <i>t</i>	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SJC	<b>0.001</b>	<b>0.411</b>	<b>0.000</b>	<b>0.000</b>	<b>0.180</b>	<b>0.010</b>	<b>0.177</b>	<b>0.002</b>

Gaussian, Plackett, Frank, and Student-*t* copulas show symmetry while Clayton, rotated Clayton, Gumbel, rotated Gumbel, and SJC copulas show asymmetry in the tail dependence. Note that a highlighted value indicates the statistic for asymmetric copula functions.

copula functions. Table 4 presents the asymmetry of tail dependence between pairs considered in this study by using a broad class of copula functions. Embrechts, McNeil, and Straumann (2002) show the Pearson correlation coefficient can be misleading if each paired sample does not follow the bivariate normal distribution. As our key variables follow the skewed heavy-tailed distributions, the Pearson correlation cannot describe the relationship between variables appropriately. A copula function can be a useful measure of dependence for non-normal pairs. The relative magnitude between the lower ( $\tau^L$ ) and upper ( $\tau^R$ ) tail dependence coefficients indicates the possibility of asymmetry in the tail dependence. For example, our statistics show that  $\tau^R$  is higher than  $\tau^L$  for the pair of Equity volatility and Spread and the pair of Interest rate volatility and Spread in the copula functions that allow asymmetric tail dependence. In other words, the degree of dependence for these pairs is higher in the upper tail than in the lower tail. This finding is practically applicable to various financial institutions, such as investment groups, banks, and mutual funds, in the sense that they prepare better for extreme changes in equity volatility and interest rate volatility. For example, equity volatility and yield spreads skyrocket in the event of an extremely adverse change such as a financial crisis. Investors know this. However, they may not know that they should put more focus on the effects of equity volatility and interest rate volatility on yield spreads in the presence of this extreme adverse change. Therefore, our paper is useful to practitioners in the sense that it suggests that they pay more than the usual attention to the positive effect of equity volatility on yield spreads when equity volatility and yield spreads are both extremely high.

In the same manner, we can easily find stronger tail dependence in the right tail for the Rating and Spread pair over the sample period. Weaker credit quality is associated with higher yield spreads and stronger credit quality is associated with lower yield spreads. However, our finding suggests that the link between extremely weak credit quality and extremely high yield spreads is much stronger than that between extremely strong credit quality and extremely low yield spreads. This is consistent with Kim and Stock (2014), who notes that the difference in yield spreads between AAA bonds and AA+ bonds should be lower than that between C bonds and D bonds. With this in mind, they use an opposite measure of credit ratings where D is coded as 1, ..., AAA is coded as 22, and then take logs of all credit ratings as opposed to their original measure of credit ratings where AAA = 1, ..., D = 1 without using logs of ratings.

The Pearson product moment correlation is valid under the following assumptions: normal distribution of variables, measuring level of the linearly related variables without outliers, and homoscedasticity. But the non-callable corporate bond yield spreads and other bond market data violate absence of outliers, normality of variables assumptions because there exist skewness and high kurtosis to all variables as shown in Table 2. Therefore, we transform data to standard uniform variables through its empirical distribution function and then we compute Pearson correlations.

The Pearson correlation coefficients are presented in Table 5. We first note that the dependence between Slope and *r* is the strongest while the Coupon rate-Interest rate volatility pair has the least dependence in a pairwise comparison. It appears that the Coupon rate and Interest rate volatility are statistically independent over the sample period. Another interesting observation is that the correlation of any pair with *r* is always negative.

**Table 5**

Correlation between standardized uniform pairs: noncallable case.

	Pearson correlations								
	Spread	Equity volatility	Interest rate volatility	$r$	Slope	Rating	Liquidity	Coupon	Maturity
Spread	1.000								
Equity volatility	0.713	1.000							
Interest rate volatility	0.403	0.210	1.000						
$r$	-0.469	-0.563	-0.147	1.000					
Slope	0.419	0.529	0.070	-0.958	1.000				
Rating	0.424	0.426	0.005	-0.012	0.007	1.000			
Liquidity	-0.164	-0.057	-0.080	-0.094	0.107	-0.100	1.000		
Coupon rate		0.225	0.134	-0.016	-0.113	0.118	0.359	0.007	1.000
Maturity	0.243	0.062	0.034	-0.075	0.072	0.081	0.000	0.317	1.000

#### 4.3. Callable data case

Next we turn our attention to callable bonds (see [Tables 6–9](#)). We run the two different regressions again: traditional linear regression and GCMR. As in the sample of noncallable bonds, we still find that GCMR is a better-fitting model to our sample data based on the model selection criterion. The estimation results are generally consistent with those in [Table 3](#), but apparent differences are also observed. Compared with the callable sample, the coefficient on Maturity becomes negative and statistically significant at 1% significance level. We interpret this finding as evidence that Maturity behaves differently for noncallable and callable bonds. The negative sign for Maturity seems to be attributed to the fact that the maturity of a bond is positively associated with the likelihood of having a call deferral period, thereby reducing call spreads. In addition, Coupon rate is no longer statistically significant and surprisingly, Liquidity has a positive sign.

The tail dependence analysis on the callable sample reveals that the extreme co-movements for most pairs are consistent with the previous findings in [Table 4](#). However, we find that the tail dependence exhibits strong changes in some pairs, such as Equity volatility-Spread and Slope-Spread. For the Equity volatility and Spread pair, their left-tail dependence is much stronger than right-tail dependence with the callable bond sample while they have a strong right tail dependence with the noncallable sample, even though their Pearson correlations are almost constant across sample datasets. This result is likely attributed to the fact that default risk resulting from equity volatility decreases the probability of call options being exercised, thus reducing call spreads. This finding is useful to practitioners, such as investors, banks, mutual funds, and other financial institutions in the sense that for a period such as a financial crisis, where equity volatility and yield spreads are unusually high, they should pay more attention to the magnitude of the effect of equity volatility on yield spreads for noncallable bonds than for callable bonds. This finding is consistent with the argument of [Acharya and Carpenter \(2002\)](#) and the findings of [Kim and Stock \(2014\)](#).

#### 4.4. Nonparametric D-vine copula

We also investigate the co-movements between our dependent variable and several covariates by using the nonparametric D-vine copula based quantile regression approach introduced in [Kraus and Czado \(2017\)](#). The flexible method is based on fitting a regression D-vine copula to the sample dataset. In particular, it sequentially adds independent variables until the likelihood on the model prediction power conditional on the independent variables is maximized. The algorithm ends when an additional variable does not increase the conditional likelihood of the regression model. In this manner, the algorithm automatically select the parsimonious covariates (see [Kraus & Czado, 2017](#) for more details). A vine copula structure introduced in [Bedford and Cooke \(2002\)](#) is capable for modeling the dependence structure given multivariate data by building blocks of pairwise copula construction with various copulas. Specifically, we use the programming language R package *vinereg* for the regression. In the estimation process, the marginal distributions are estimated by a nonparametric approach. It is worth noting that we select the best fitting pair copular functions among parametric and nonparametric copulas by utilizing the R-function *BiCopSelect* of the package *VineCopula*.

**Table 6**

Summary statistics: callable case.

	Obs.	Mean	Median	Minimum	Maximum	St.D	Skewness	Kurtosis
Spread	50,299	4.04	2.55	0.14	38.79	4.54	3.68	20.33
Equity volatility	50,299	0.03	0.02	0.00	0.23	0.02	1.84	8.02
Interest rate volatility	50,299	0.58	0.57	0.05	1.28	0.34	0.31	2.26
$r$	50,299	2.45	2.07	0.00	5.27	2.05	0.05	1.31
Slope	50,299	1.26	1.57	-0.45	3.39	1.30	0.18	1.53
Rating	50,299	7.35	6.00	1.00	22.00	5.07	0.64	2.25
Liquidity	50,299	0.31	0.23	0.00	1.00	0.24	0.99	3.09
Coupon rate	50,299	6.10	5.80	1.25	13.50	1.25	1.60	6.90
Maturity	50,299	11.75	10.71	1.00	86.28	7.02	0.95	6.17

**Table 7**  
Estimation results: callable case.

	Linear Regression	Gaussian Copula Regression
Equity volatility	34.171*** (0.977)	6.014*** (0.133)
Interest rate volatility	1.177*** (0.051)	0.335*** (0.008)
$r$	-2.145*** (0.033)	-0.455*** (0.004)
Slope	-2.523*** (0.051)	-0.466*** (0.007)
Rating	0.473*** (0.005)	0.099*** (0.001)
Liquidity	0.134 (0.071)	0.176*** (0.010)
Coupon rate	-0.270*** (0.016)	-0.003 (0.002)
Maturity	-0.057*** (0.002)	-0.012*** (0.000)
Constant	9.717*** (0.181)	1.979*** (0.024)
Log likelihood	130,740	83,770
AIC	261,490	167,560

\*\*\*, \*\*, and \* indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.

**Table 8**  
Statistics for copulas between standardized uniform pairs: callable case.

Copula type	Equity volatility-Spread		Interest rate volatility-Spread		$r$ -Spread		Slope-Spread	
	$\tau^L$	$\tau^R$	$\tau^L$	$\tau^R$	$\tau^L$	$\tau^R$	$\tau^L$	$\tau^R$
Gaussian	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Clayton	<b>0.584</b>	<b>0.000</b>	<b>0.047</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.297</b>	<b>0.000</b>
Rotated Clayton	<b>0.000</b>	<b>0.461</b>	<b>0.000</b>	<b>0.073</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.145</b>
Plackett	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Frank	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Gumbel	<b>0.000</b>	<b>0.498</b>	<b>0.000</b>	<b>0.175</b>	<b>0.000</b>	<b>0.122</b>	<b>0.000</b>	<b>0.263</b>
Rotated Gumbel	<b>0.538</b>	<b>0.000</b>	<b>0.153</b>	<b>0.000</b>	<b>0.122</b>	<b>0.000</b>	<b>0.311</b>	<b>0.000</b>
Student's $t$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SJC	<b>0.578</b>	<b>0.151</b>	<b>0.017</b>	<b>0.086</b>	<b>0.000</b>	<b>0.000</b>	<b>0.339</b>	<b>0.001</b>
Copula type	Rating-Spread		Liquidity-Spread		Coupon rate-Spread		Maturity-Spread	
	$\tau^L$	$\tau^R$	$\tau^L$	$\tau^R$	$\tau^L$	$\tau^R$	$\tau^L$	$\tau^R$
Gaussian	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Clayton	<b>0.494</b>	<b>0.000</b>	<b>0.020</b>	<b>0.000</b>	<b>0.381</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
Rotated Clayton	<b>0.000</b>	<b>0.533</b>	<b>0.000</b>	<b>0.055</b>	<b>0.000</b>	<b>0.267</b>	<b>0.000</b>	<b>0.000</b>
Plackett	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Frank	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Gumbel	<b>0.000</b>	<b>0.498</b>	<b>0.000</b>	<b>0.155</b>	<b>0.000</b>	<b>0.338</b>	<b>0.000</b>	<b>0.122</b>
Rotated Gumbel	<b>0.489</b>	<b>0.000</b>	<b>0.133</b>	<b>0.000</b>	<b>0.371</b>	<b>0.000</b>	<b>0.122</b>	<b>0.000</b>
Student's $t$	0.000	0.000	0.000	0.000	0.002	0.002	0.001	0.001
SJC	<b>0.329</b>	<b>0.489</b>	<b>0.006</b>	<b>0.087</b>	<b>0.368</b>	<b>0.119</b>	<b>0.000</b>	<b>0.000</b>

Gaussian, Plackett, Frank, and Student- $t$  copulas show symmetry while Clayton, rotated Clayton, Gumbel, rotated Gumbel, and SJC copulas show asymmetry in the tail dependence. Note that a highlighted value indicates the statistic for asymmetric copula functions.

Table 10 presents the estimation results reached by using the D-vine regression algorithm. Note that we report only the conditional log-likelihood values, which are used as a measure of the model's fit in Kraus and Czado (2017). The missing values for some variables indicate that Coupon rate (in the noncallable case) and Slope and Liquidity (in the callable case) do not increase the model's conditional log-likelihood. Thus, the algorithm automatically excludes the covariates to provide the parsimonious model. The covariates chosen by the algorithm are significantly related with the response variable, Spread.

Given the selected variables, Fig. 3 and Fig. 4 display the dependence between each independent variable and Spread across the quantiles ( $q = 0.1$  (red line),  $0.5$  (green line), and  $0.9$  (blue line)). Note that any variables on both horizontal and vertical axes are

**Table 9**

Correlation between standardized uniform pairs: callable case.

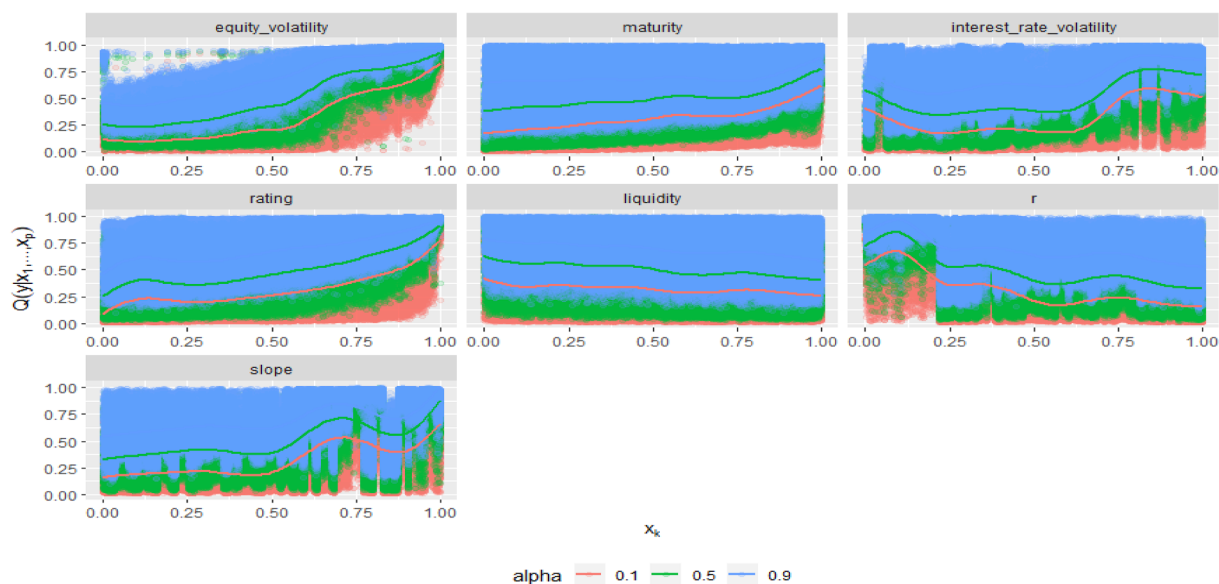
Pearson correlations									
		Equity volatility	Interest rate volatility	$r$	Slope	Rating	Liquidity	Coupon rate	Maturity
Spread	1.000								
Equity volatility	0.702	1.000							
Interest rate volatility	0.286	0.047	1.000						
$r$	-0.452	-0.662	-0.189	1.000					
Slope	0.415	0.665	0.147	-0.947	1.000				
Rating	0.631	0.435	-0.083	0.018	-0.011	1.000			
Liquidity	0.225	0.000	-0.021	0.154	-0.162	0.310	1.000		
Coupon rate	0.458	0.272	0.003	0.000	0.001	0.568	0.519	1.000	
Maturity	-0.353	-0.087	0.020	-0.121	0.125	-0.464	-0.168	-0.060	1.000

**Table 10**

Estimation results by using vine regression models.

	Noncallable case	Callable case
Equity volatility	30,773.781***	22,085.521***
Interest rate volatility	5,985.705***	5,353.505***
$r$	760.506***	898.520***
Slope	1,055.187***	
Rating	7,183.113***	7,119.322***
Liquidity	1,475.476***	
Coupon rate		1,909.574***
Maturity	5,970.329***	1,527.681***

\*\*\* indicates statistical significance at the 1% level.

**Fig. 3.** Dependence between each covariate and Spread: noncallable case.

transformed to a uniform distribution. We can interpret from the shape of the lines that there are non-linear and asymmetric relationships between the pairs of interest. This finding provides additional justification for our use of copula functions.

We run the vine copula-based mean regression and vine copula-based median regression models with the selected covariates reported in Table 10. We take the predicted values from each regression. Then, we obtain residuals of each regression by subtracting the predicted values from the actual values of Spread in the sample. Table 11 presents the central tendency and dispersion of the residual values. The columns of 'mean vine' show the statistics obtained from the vine copula-based mean regression, and the

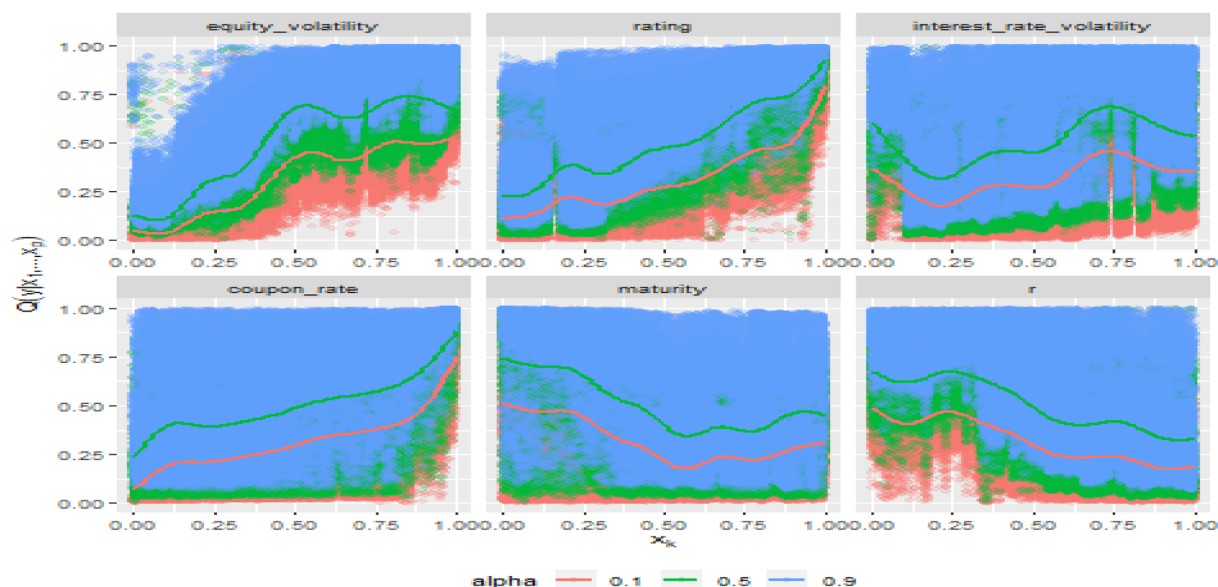


Fig. 4. Dependence between each covariate and Spread: callable case.

Table 11

Residual values by using the vine copular regression model.

	noncallable case		callable case	
	mean vine	median vine	mean vine	median vine
1 <sup>st</sup> Q	-0.0970	-0.0909	-0.0897	-0.0879
Median	-0.0004	-0.0014	0.0014	-0.0001
Mean	-0.0032	-0.0017	-0.0038	-0.0068
3 <sup>rd</sup> Q	0.0887	0.0858	0.0880	0.0768
IQR	0.1858	0.1767	0.1776	0.1647

1<sup>st</sup>Q, 3<sup>rd</sup>Q, and IQR represent the first quartile, third quartile, and interquartile range values, respectively.

columns of 'median vine' are based on the vine copula-based median regression. When using the subsample for the noncallable case, we find that, in terms of the IQR measure, the residuals from the mean vine copula regression are more spread out than the residuals from the median vine copula regression model. Although the median (mean) from the mean vine copula is greater (less) than that from the median vine copula regression, the difference is negligible. Regarding the subsample for the callable case, a similar pattern of the dispersion for the middle 50% of the residuals is observed. However, unlike the noncallable case, the sign of inequality becomes reversed for the median and mean statistics. The findings from Table 11 also justify our use of copula functions in this study.

## 5. Conclusion

In this paper, our corporate bond data do not follow normal distribution and have outliers. Therefore, we employ the copula methods to examine the relationship between yield spreads and other explanatory variables. In particular, we utilize the GCMR model with the Weibull marginal distribution to analyze how corporate bond yield spreads are determined. Furthermore, our empirical analysis uses various copula functions to test for the tail dependence structure. Based on our statistical setup, we show that our regression model is a better-fitting model than the one based on the lower AIC value. In the GCMR estimation results, coupon rates increase noncallable bond yield spreads, while coupon rates do not affect callable bond yield spreads in the traditional regression. The positive sign of coupon rates suggest that investors demand higher rates of return since they face greater tax disadvantages for higher corporate coupon bonds relative to their Treasury counterparts.

Our tail dependence results are useful to practitioners. For example, greater upper tail dependence than low tail dependence for the spread and equity volatility pair in the sample of noncallable bonds suggests practitioners pay more than usual attention to the positive effect of equity volatility on yield spreads. Interestingly, our callable sample reveals the opposite results. We find greater low tail dependence than upper tail dependence for the spread and equity volatility pair in the sample of callable bonds. We attribute this to the likelihood that yield spreads decline due to the reduction of call option values driven by a rise in default risk following an increase in equity volatility.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## CRediT authorship contribution statement

**Jong-Min Kim:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation. **Dong H. Kim:** Data curation, Writing - original draft, Writing - review & editing. **Hojin Jung:** Investigation, Writing - original draft, Writing - review & editing, Supervision, Project administration.

## References

- Acharya, V. V., & Carpenter, J. N. (2002). Corporate bond valuation and hedging with stochastic interest rates and endogenous bankruptcy. *Review of Financial Studies*, 15(5), 1355–1383.
- Al Janabi, M. A., Hernandez, J. A., Berger, T., & Nguyen, D. K. (2017). Multivariate dependence and portfolio optimization algorithms under illiquid market scenarios. *European Journal of Operational Research*, 259(3), 1121–1131.
- Aloui, R., & Aïssa, M. S. B. (2016). Relationship between oil, stock prices and exchange rates: A vine copula based GARCH method. *The North American Journal of Economics and Finance*, 37, 458–471.
- Aloui, R., Aïssa, M. S. B., & Nguyen, D. K. (2013). Conditional dependence structure between oil prices and exchange rates: A copula-GARCH approach. *Journal of International Money and Finance*, 32, 719–738.
- Ang, A., Boivin, J., Dong, S., & Loo-Kung, R. (2011). Monetary policy shifts and the term structure. *The Review of Economic Studies*, 78(2), 429–457.
- Bao, J., Pan, J., & Wang, J. (2011). The illiquidity of corporate bonds. *The Journal of Finance*, 66(3), 911–946.
- Bedford, T., & Cooke, R. M. (2002). Vines – A new graphical model for dependent random variables. *The Annals of Statistics*, 30(4), 1031–1068.
- Booth, L., Gounopoulos, D., & Skinner, F. (2014). The choice between callable and noncallable bonds. *Journal of Financial Research*, 37(4), 435–460.
- Boyarchenko, N., Fuster, A., & Lucca, D. O. (2019). Understanding mortgage spreads. *The Review of Financial Studies*, 32(10), 3799–3850.
- Breeden, D. (2011). A stocks, bonds and consumers leading index (SBCLI): Consumer behavior as a leading indicator. *World Finance Conference, Rhodes, Greece*.
- Campbell, J. Y., & Taksler, G. B. (2003). Equity volatility and corporate bond yields. *The Journal of Finance*, 58(6), 2321–2350.
- Chen, L., Lesmond, D. A., & Wei, J. (2007). Corporate yield spreads and bond liquidity. *The Journal of Finance*, 62(1), 119–149.
- Choudhry, T. (2016). Time-varying risk premium yield spread effect in term structure and global financial crisis: Evidence from Europe. *International Review of Financial Analysis*, 48, 303–311.
- Collin-Dufresne, P., & Goldstein, R. S. (2001). Do credit spreads reflect stationary leverage ratios? *The Journal of Finance*, 56(5), 1929–1957.
- Contessi, S., De Pace, P., & Guidolin, M. (2014). How did the financial crisis alter the correlations of US yield spreads? *Journal of Empirical Finance*, 28, 362–385.
- Covitz, D., & Downing, C. (2007). Liquidity or credit risk? The determinants of very short-term corporate yield spreads. *The Journal of Finance*, 62(5), 2303–2328.
- Duffee, G. R. (1998). The relation between treasury yields and corporate bond yield spreads. *The Journal of Finance*, 53(6), 2225–2241.
- Duffee, G. R. (1999). Estimating the price of default risk. *Review of Financial Studies*, 12(1), 197–226.
- Dwyer, G. P., & Tkac, P. (2009). The financial crisis of 2008 in fixed-income markets. *Journal of International Money and Finance*, 28(8), 1293–1316.
- Ederington, L. H., & Stock, D. R. (2002). Impact of call features on corporate bond yields. *The Journal of Fixed Income*, 12(2), 58–68.
- Elton, E. J., Gruber, M. J., Agrawal, D., & Mann, C. (2001). Explaining the rate spread on corporate bonds. *The Journal of Finance*, 56(1), 247–277.
- Embrechts, P., McNeil, A., & Straumann, D. (2002). Correlation and dependence in risk management: Properties and pitfalls. *Risk Management: Value at Risk and Beyond*, 176–223.
- Eom, Y. H., Helwege, J., & Huang, J. Z. (2004). Structural models of corporate bond pricing: An empirical analysis. *Review of Financial Studies*, 17(2), 499–544.
- Estrella, A., & Hardouvelis, G. A. (1991). The term structure as a predictor of real economic activity. *The Journal of Finance*, 46(2), 555–576.
- Estrella, A., & Mishkin, F. S. (1996). The Yield Curve as a Predictor of US Recessions. *Current Issues in Economics and Finance*. Vol. 2, No.7. Federal Reserve Bank, New York.
- Frey, R., & McNeil, A. J. (2001). *Modelling dependent defaults*. ETH, Eidgenössische Technische Hochschule Zürich, Department of Mathematics.
- Gilchrist, S., & Zakrajšek, E. (2012). Credit spreads and business cycle fluctuations. *American Economic Review*, 102(4), 1692–1720.
- Guidolin, M., & Tam, Y. M. (2013). A yield spread perspective on the great financial crisis: break-point test evidence. *International Review of Financial Analysis*, 26, 18–39.
- Güntay, L., & Hackbarth, D. (2010). Corporate bond credit spreads and forecast dispersion. *Journal of Banking & Finance*, 34(10), 2328–2345.
- Jammazi, R., Tiwari, A. K., Ferrer, R., & Moya, P. (2015). Time-varying dependence between stock and government bond returns: International evidence with dynamic copulas. *The North American Journal of Economics and Finance*, 33, 74–93.
- Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. CRC Press.
- Junker, M., Szimayer, A., & Wagner, N. (2006). Nonlinear term structure dependence: Copula functions, empirics, and risk implications. *Journal of Banking & Finance*, 30(4), 1171–1199.
- Kahlert, D., & Wagner, N. (2017). Dependence between extreme risk factor movements of euro area sovereign bonds. Working Paper.
- Kim, D., & Kim, J. M. (2014). Analysis of directional dependence using asymmetric copula-based regression models. *Journal of Statistical Computation and Simulation*, 84(9), 1990–2010.
- Kim, D. H., & Stock, D. (2014). The effect of interest rate volatility on corporate yield spreads on both noncallable and callable bonds. *Journal of Corporate Finance*, 26, 20–35.
- Kim, J. M., & Jung, H. (2016). Linear time-varying regression with copula-DCC-GARCH models for volatility. *Economics Letters*, 145, 262–265.
- Kim, J. M., & Jung, H. (2017). Can asymmetric conditional volatility imply asymmetric tail dependence? *Economic Modelling*, 64, 409–418.
- Kim, J. M., Jung, Y. S., Choi, T., & Sungur, E. A. (2011). Partial correlation with copula modeling. *Computational Statistics & Data Analysis*, 55(3), 1357–1366.
- Kim, J. M., Jung, Y. S., Sungur, E. A., Han, K. H., Park, C., & Sohn, I. (2008). A copula method for modeling directional dependence of genes. *BMC Bioinformatics*, 9(1), 225.
- Kojadinovic, I., & Yan, J. (2010). Modeling multivariate distributions with continuous margins using the copula R package. *Journal of Statistical Software*, 34(9), 1–20.
- Kraus, D., & Czado, C. (2017). D-vine copula based quantile regression. *Computational Statistics & Data Analysis*, 110, 1–18.
- Li, D. X. (1999). On default correlation: A copula function approach. *Journal of Fixed Income*, 9(4), 43–54.
- Longstaff, F. A., Mithal, S., & Neis, E. (2005). Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. *The Journal of Finance*, 60(5), 2213–2253.
- Longstaff, F. A., & Schwartz, E. S. (1995). A simple approach to valuing risky fixed and floating rate debt. *The Journal of Finance*, 50(3), 789–819.
- Masarotto, G., & Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics*, 6, 1517–1549.
- Mensi, W., Hammoudeh, S., Shahzad, S. J. H., & Shahbaz, M. (2017). Modeling systemic risk and dependence structure between oil and stock markets using a variational mode decomposition-based copula method. *Journal of Banking & Finance*, 75, 258–279.
- Nelsen, R. B. (2007). *An introduction to copulas*. Springer Science & Business Media.



- Nguyen, P., & Liu, W. h. (2017). Time-varying linkage of possible safe haven assets: A cross-market and cross-asset analysis. *International Review of Finance*, Forthcoming.
- Ning, C. (2010). Dependence structure between the equity market and the foreign exchange market – A copula approach. *Journal of International Money and Finance*, 29(5), 743–759.
- Parikh, V., & Baraniecki, A. (1996). Computationally efficient discrete wavelet transform model. *International Journal of Modelling and Simulation*, 16(3), 142–150.
- Parsa, R. A., & Klugman, S. A. (2011). Copula regression. *Variance Advancing and Science of Risk*, 5, 45–54.
- Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2), 527–556.
- Philippas, D., & Siriopoulos, C. (2013). Putting the “C” into crisis: Contagion, correlations and copulas on EMU bond markets. *Journal of International Financial Markets, Institutions and Money*, 27, 161–176.
- Pourkhanali, A., Kim, J. M., Tafakori, L., & Fard, F. A. (2016). Measuring systemic risk using vine-copula. *Economic Modelling*, 53, 63–74.
- Reboredo, J. C. (2012). Modelling oil price and exchange rate co-movements. *Journal of Policy Modeling*, 34(3), 419–440.
- Reboredo, J. C., & Ugolini, A. (2015). Systemic risk in European sovereign debt markets: A CoVaR-copula approach. *Journal of International Money and Finance*, 51, 214–244.
- Rossi, M. (2014). Realized volatility, liquidity, and corporate yield spreads. *The Quarterly Journal of Finance*, 4(01), 1450004.
- Serra, T., & Gil, J. M. (2012). Biodiesel as a motor fuel price stabilization mechanism. *Energy Policy*, 50, 689–698.
- Sklar, A. (1973). Random Variables, Joint Distribution Functions, and Copulas, *Kybernetika* (Prague), 9, pp. 449–460.
- Sklar, M. (1959). Fonctions de Répartition À N Dimensions et Leurs Marges. Université Paris 8.
- Song, P. X. K. (2000). Multivariate dispersion models generated from Gaussian copula. *Scandinavian Journal of Statistics*, 27(2), 305–320.
- Wang, K., Chen, Y. H., & Huang, S. W. (2011). The dynamic dependence between the chinese market and other international stock markets: A time-varying copula approach. *International Review of Economics & Finance*, 20(4), 654–664.
- Wang, Y. C., Wu, J. L., & Lai, Y. H. (2013). A revisit to the dependence structure between the stock and foreign exchange markets: A dependence-switching copula approach. *Journal of Banking & Finance*, 37(5), 1706–1719.
- Wu, C. C., Chung, H., & Chang, Y. H. (2012). The economic value of co-movement between oil price and exchange rate using copula-based GARCH models. *Energy Economics*, 34(1), 270–282.
- Wu, C. C., & Lin, Z. Y. (2014). An economic evaluation of stock–bond return comovements with copula-based GARCH models. *Quantitative Finance*, 14(7), 1283–1296.