Macroeconomic and Financial Applications of S-vine Copula Models for Time Series

JIALING HAN

PhD

University of York

Business and Society

January 2023

Abstract

This thesis investigates the modelling and prediction of macroeconomic and financial time series by using S-vine models and vt-S-vine models. The first contribution of this thesis is to demonstrate that non-linear and non-Gaussian models in the class of S-vine processes can fit better than autoregressive moving average (ARMA) models for certain types of macroeconomic time series, such as inflation rates. The S-vine processes generalize the class of Gaussian ARMA models by modelling dependencies using pair copulas and by modelling marginal behaviour using arbitrary continuous distributions. The second part of this thesis concerns the modelling and forecasting the volatile financial returns series by vt-S-vine models. Returns are traditionally estimated by using the generalized autoregressive conditionally heteroscedastic (GARCH) type processes or variants. We show that the vt-S-vine models can compete with GARCH processes in many cases. In order to reveal the statistical properties and structures of GARCH type processes, the vt-S-vine models are applied to replicate GARCH type processes. The best combinations of margins and pair copulas in vt-S-vines for replicating GARCH type processes are determined. The higher-order vt-S-vine processes with a sequence of mixed pair copulas can mimic GARCH type processes more precisely than first-order or second-order S-vine with t copulas as proposed by previous studies. The final part of this thesis is devoted to the application of vt-S-vines in the trading book of banks. The value-at-risk (VaR) and the VaR exceedance probabilities are estimated in vt-S-vine processes, where the best quantile estimation methods are presented in each case. Surprisingly, the quantile estimator that is closest on average to the true value of the quantile of a distribution may not be the one that yields the most accurate value for the exceedance probability. Vt-S-vines are very flexible and promising models for stationary macroeconomic, financial and banking time series, and deserve to be widely used and rapidly developed in future.

Acknowledgement

First of all, I would like to express my appreciation from the bottom of my heart to both my supervisor Prof. Alexander McNeil and my co-supervisor Dr. Alexandra Dias. I am grateful to have them supervising my work from the very beginning during my entire Ph.D. study. Without their guidance, help, and effort, all the work would never be possible to be accomplished. They have been so kind, patient and supportive to me all the time. I also thank both Prof. Alexander McNeil and his lovely wife Janine McNeil for their support and help to my living during the lockdown of the COVID-19 pandemic.

Secondly, I am also thankful that Prof. Jacco Thijssen, my TAP member, has been keen and helpful to my study and given me significantly useful suggestions.

Thirdly, I would like to thank my parents, Mr. Weiyang Han and Ms. Lixia Yu, and other family members for their support and encouragement in my study and life. I am grateful to be loved by them through my life.

Lastly, many thanks to my partner, Dr. Zhuoyi Raymond Ye, for his patience and support to me through my work and life.

I am sincerely grateful to have these people in my research and life. This four-year study has been a significantly valuable and enormously joyful journey because of all mentioned.

Author's Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Contents

Chapter 1

Introduction

Time series data, which are data that have been collected over a period of time on one or more variables, are one of the most widely employed data in quantitative analysis of financial problems. Time series data are associated with a particular frequency of observation or frequency of collection of data points. Generally, all data used in a model should be of the same frequency of observation. For macroeconomic data, the daily, weekly, monthly or quarterly observations are commonly used. The frequency of financial data can be intra-daily and higher. In many studies in finance, the starting point is a time series of prices or rates, although interest centres on the returns. For a number of statistical reasons and practical applications, the academic finance literature generally employs the log-return formulation (which is the difference between log prices of this period's price and the previous period's price).

The serial dependence of financial returns can be captured by copula models. The applications of copulas and their variations in financial time series, such as vine copulas, have recently attracted the attention of academics and practitioners. There are large and growing literature on copula-based models for economic and financial time series; details can found in the review written by Patton [2012]. The copula-based methods can study separately the marginal distributions and the dependence structures that links these distributions to construct a joint distribution. This property gives a greater degree of flexibility in defining and estimating the model, which allows for more selections for the researchers.

1.1 Statistical models for macroeconomic data

Univariate time series models are a class of specifications where one tries to model and forecast financial and economic variables using the information contained their own past values and possibly current and past values of an error term. The time series models are designed to capture the empirically relevant features of the observed data that arise

from a variety of different structural models. A widely used class of time series models is the family of ARIMA models, which were proposed by Box et al. [1970].

ARIMA models are classic models for estimating and predicting macroeconomic data. There are some variants of ARIMA models developed to describe the features in macroeconomic data, such as seasonal autoregressive integrated moving-average (SARIMA) model for macroeconomics time series with seasonality, autoregressive fractionally integrated moving-average (ARFIMA) model for long-memory macroeconomic data. Furthermore, the decomposition of economic time series into trend and cyclical components plays an important role in much of macroeconomics, which can remove the trend or seasonality in the original data and use the ARMA processes to model the remaining stationary time series. More details and references of modeling the macroeconomic data are presented in Chapter 4.

The family of classical ARMA processes is widely used in many traditional applications of time series analysis. These processes describe the linear nature of macroeconomic times series, where the model is linear in the parameters, so there is one parameter multiplied by each variable in the model. However, the ARMA processes do not allow for the volatility in data and many relationships in finance are intrinsically non-linear. They can model the macroeconomic data accurately, since the macroeconomic data usually have stable volatility and linear structures. However, the empirical financial data usually exhibit changing volatility, which is difficult to be modelled by the family of ARMA processes. In this case, the GARCH type models can be used.

1.2 Statistical models for volatile financial returns

For a single time series of financial returns, there are stylized facts existing, such as the serial dependence of the return series and their absolute values, volatility clustering or their heavy tails. The stylized facts of financial time series are based on the empirical observations and their inferences. The stylized facts are typically applied to time series of daily log-returns and often continue to hold in the longer-interval series, such as weekly or monthly returns. It is important to capture these properties of financial returns accurately in empirical application, not only for financial institutions, but also for banking regulators. The widely applied univariate time series models that can mimic the empirical properties of financial returns includes autoregressive integrated moving-average (ARIMA) processes, generalized autoregressive conditionally heteroscedastic (GARCH) processes and their variations. The GARCH-type processes can capture the important phenomenon of volatility, which is very crucial for financial returns. The combinations of the ARMA and GARCH processes, such as ARMA-GARCH models are also widely applied in industry.

Linear structural models such as ARMA model are unable to explain a number of

important features common to much financial data, including:

- Leptokurtosis, which is the tendency for financial asset returns to have distributions that present heavy tails and excess peakness at the mean.

- Volatility clustering, which is the tendency for volatility in financial markets to appear in bunches. Therefore, large returns are expected to follow large returns. Correspondingly, small returns are followed by small returns.

- Leverage effects, which is the tendency for volatility to rise more following a large price fall than following a price rise of the same magnitude.

The details can be found in Cuthbertson [2004].

GARCH models are the most popular non-linear financial models used for modelling and predicting volatility. They are non-linear in variance. The GARCH models, which are extensions of autoregressive conditionally heteroscedastic (ARCH) models proposed by Engle [1982], are developed by Bollerslev [1986]. The GARCH models allows the conditional variance to be dependent on the previous lagged squared values. Generally, the GARCH processes can capture the volatility clustering features in financial returns.

There are a huge number of extensions and variants of the basic GARCH models that have been proposed. For example, the asymmetric GARCH models, including GJR-GARCH proposed by Glosten et al. [1993], or the exponential GARCH (EGARCH) introduced by Nelson [1991]. The two types of extensions of GARCH processes can describe the leverage effects more precisely. It is argued that a negative shock to a financial time series can cause volatility to rise by more than a positive shock of the same magnitude. In the case of equity returns such asymmetries are typically attributed to leverage effects, which can be modelled well by asymmetric GARCH processes.

The GARCH-type models can also be combined with ARMA model, written as ARMA-GARCH models, which are the ARMA models with GARCH errors. These give us a flexible family of ARMA models with GARCH errors that combines the features of both model classes.

The accuracy and efficiency of GARCH type processes in modelling the volatility of financial returns make them very popular in finance, insurance, banks and regulators. However, the marginal distributions of GARCH-type models are unknown. The option of choosing the innovation distributions allow the flexibility of the GARCH models, but the connection between the innovation distributions of GARCH processes and their marginal distributions is not clear so far. Moreover, when setting the regulation for risk measures, such as value-at-risk (VaR), it is difficult to calculate the VaR in GARCH type models, although we can obtain the approximate VaR by Monte Carlo simulation.

1.3 Copula-based methods

The copula-based methods allow the researchers to specify the models for marginal distributions separately from the joint distributions modelling the dependence structures of the transformed time series that follow a uniformly distributed copula process. Unlike the GARCH-type processes, the marginal distributions and the joint distributions are specified, which means we clearly understand the different parts of the models. This property gives more possible combinations of the models. We can choose different margins and copula processes.

There are also extensions and variants of copula-based methods. For the financial time series with volatility, the v-transformation can be combined with copula-based methods to describe the volatility. The definition of v-transformation will be introduced in Chapter 5. The v-transformation is developed in McNeil [2020], who proposed the models are constructed to describe both stochastic volatility in the magnitude of price movements and serial correlation in their directions.

1.3.1 Vine copulas

Among a vast number of variants of copulas, vine copulas proposed by Joe [1996] are widely applied and very popular, because of their flexible structures which allow for the combination of distinctive bivariate copulas in one model. There are many different types of vine structures, such as D-vines, C-vines or R-vines. The type of vine is selected according to the structure of data. The references to vine copulas can be found in Chapter 3.

For financial time series, the D-vine copulas are widely applied, since D-vine structure corresponds well to the time order of each values in time series. D-vine copulas belong to the Markov process. Chen and Fan [2006] investigate the first-order Markov models which are special cases of D-vines in modelling dependence of time series. The research of D-vine copulas extends to higher-order Markov models by Ibragimov [2009]. Nikoloulopoulos et al. [2012] proposed to apply vine copulas with asymmetric tail dependence for financial return data. Other references and properties of D-vine copulas and their variants will be discussed in Chapter 3.

1.4 Outline of the thesis

This thesis consists of three main topics. These topics are mainly around the S-vine copulas, which are the stationary D-vine copulas, including their theory, properties, applications and also their variants for modelling the volatile financial returns. Chapter 2 introduces the fundamental concepts and properties of copulas and vine copulas. Meanwhile, the candidate marginal distributions and bivariate copulas used in S-vine models

are discussed. The methodology of using the S-vine processes to model and predict economic and financial time series is described in Chapter 3, where the simplification of the S-vine models, which reduces the number of parameters remarkably, is explained. In addition, Chapter 3 is also devoted to presenting the techniques and methods employed to improve the partial autocorrelation estimation in vine copulas, such as the rotation of bivariate copulas when the dependence is negative.

Chapter 4 develops the first topic of this thesis, which is applying the S-vine processes in modelling and predicting inflation rates, which embodies the application of S-vine models to macroeconomic time series. This chapter discusses the methods used to adjust raw data to stationary time series in the first part. The results of fitting different S-vine models to the adjusted inflation rates are compared and the prediction results are assessed via the quantile score methods. Chapter 4 is also devoted to the application of both parametric and semi-parametric S-vine models, so as to reveal the influence of marginal distributions.

The second topic is discussed from Chapter 5 to Chapter 7. In Chapter 5, we state the definition and properties of vt-S-vine models, which are the combinations of v-transformation and S-vine models. In order to explore if the vt-S-vine models can model the volatility of GARCH processes, we mimic GARCH processes by vt-S-vines in Chapter 6. The one-step prediction for GARCH processes is exhibited in Chapter 6 as well. Chapter 7 investigate the empirical study of vt-S-vines, where the volatile returns series are modelled and forecasted by selected vt-S-vine models.

The third topic concerns the distribution of quantile exceedances in volatile data generated from vt-S-vine models, which is in the eighth chapter. This has an application to the "traffic light system" used in the regulation of banks' trading books. Chapter 8 is concerned with the VaR exceedances and VaR estimations from vt-S-vine models. In this chapter, we also investigate six possible quantile estimation methods used by banks and regulators. The objective is to find the most suitable methods in different cases for vt-S-vine processes. The distributions of exceedances from vt-S-vine models are also studied in Chapter 8.

1.4.1 New material in the thesis

The background knowledge and fundamental concepts of copulas and S-vine copulas in Chapters 2 and 3 are based on existing literature by a number of authors, as cited in these chapters. Chapter 5 is a review of the concepts and properties of v-transforms underpinning the models used in Chapter 6 and 7; this is also based on published literature. Except for the contents referred to above, the simulations, estimations, predictions and applications of these models and other analyses and explorations of S-vine/vt-S-vine models are new.

Chapter 2

Copulas

This chapter introduces some terminology and concepts in dependence modelling with copulas. We focus on copulas and their properties, especially vine copulas. The key contents and certain properties of copulas in this chapter are based on Chapter 7 in McNeil et al. [2015]. Most of the definitions and expressions of vine copulas in Section 2.6 are referred in Joe [2014].

2.1 Dependence modelling

Suppose we have a sample of size n consisting of vector $\boldsymbol{x}_i = (x_{i1}, ..., x_{id})^\top$ for $i = 1, ..., n$. The $\boldsymbol{x}_i$ are independent and identically distributed (i.i.d) realizations of a random vector $\boldsymbol{X} = (X_1, ..., X_d)^\top$.

There are two components for dependence modelling with copulas. The first is choosing the univariate model for each of the variables $X_1, ..., X_d$, which is referred to as marginal distribution in this chapter. The second is the copula models for the dependence of the d variables.

Generally, the choices for the univariate parametric families in marginal distribution estimation depend on features such as modality, tail weight and asymmetry. Details are discussed in Section 2.2.

After univariate models for each variable have been chosen, copulas can be applied to describe the different types of dependence structure, such as positive or negative dependence, and conditional independence. Also, copulas can embody different joint tail behaviours. The shape and tail behaviour of the density are the reference for selection of parametric copula families and the selection of parametric univariate families.

One of the challenges in copula construction is extending the many available bivariate copulas to multivariate models in order to obtain flexible dependence. Vine copulas are one of the approaches used to express multivariate dependence in terms of combinations of bivariate copula sequences. The details are discussed in Section 2.6.

2.2 Marginal distributions

Before using the copula to describe the dependence between different variables, the marginal distribution is applied to transfer data to uniform distribution in each dimension.

Proposition 1 (Probability Transform). *If X has distribution function F, where F is continuous univariate distribution function, then $F(X) \sim U(0,1)$.*

Proposition 1 is referred to in Proposition 7.2 of McNeil et al. [2015]. The values transformed from probability transform is probability integral transformation (PIT) values. The marginal distributions develop pseudo-observations of the original data as the inputs for copulas. In practice, we do not know F and have to estimate it with model $\hat{F}$. The pseudo observations can be written as: $U_{ij} = \hat{F}_j(x_{ij})$, where $j = 1, ..., d, i = 1, ..., n$. Hence, the selection of marginal distributions exerts influence on the estimation of copulas. The marginal distributions developing PIT values that can pass uniformity tests are chosen in the first step.

There are two main choices of marginal model, parametric and non-parametric margins. The parametric margins are the distributions with parameters, such as normal distribution and student t distribution. The non-parametric margins are usually the empirical distribution function. The two types of margins are both applied in the estimation process.

Many kinds of data found in finance usually have heavier tails than normal distribution; see Ibragimov [2004] for details. There are a number of studies documented in which the distributions of macroeconomics do not follow the normal distribution, including Acemoglu et al. [2017] and Fagiolo et al. [2008]. A recent study, Bladt and McNeil [2022] propose that skewed student distribution from the family of skewed distributions referred to in Fernández and Steel [1998] are more flexible than normal distribution. In order to find the distributions that can describe tails better and transform data to uniform distribution, the skewed distributions are considered in this thesis.

2.3 Copulas

Definition 2.3.1 (copula). *A d-dimensional copula is a distribution function on $[0,1]^d$ with standard uniform marginal distributions.*

Theorem 1 (Sklar 1959). *Let F be a joint distribution function with margins $F_1, ..., F_d$. Then there exists a copula $C : [0,1]^d \to [0,1]$ such that, for all $x_1, ..., x_d$ in $\bar{\mathbb{R}} = [-\infty, \infty]$,*

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)).$$

If the marginal distributions F are continuous, then C is unique; otherwise C is uniquely determined on $RanF_1 \times RanF_2 \times \cdots \times RanF_d$, where $RanF_i = F_i(\bar{\mathbb{R}})$ denotes the range of F_i. Conversely, if C is a copula and $F_1, ..., F_d$ are univariate distribution functions, then the function F defined in Sklar [1959] is a joint distribution function with marginal distributions $F_1, ..., F_d$; see Chapter 7.1 of McNeil et al. [2015] for proofs. The copula of a random vector $(X_1, ..., Xd)$ is the distribution functions of $(U_1, ..., U_d)$ where $U_i = F_i(X_i)$ for all i, i.e. PIT transformation is applied to each variable.

If the marginal distributions $F(x)$ are continuous, we use $f(x_1, ..., x_d)$ to express the density function of $F(x_1, ..., x_d)$.Then, the density function can be written as:

$$f(x_1, x_2, ..., x_d) = c\{F_1(x_1), ..., F_d(x_d)\} \times \prod_{k=1}^{d} f_k(x_k) \qquad for\ all\ \boldsymbol{x} \in \mathbb{R}^d,$$

where c is the density of C, called copula density, and $f_1, ..., f_d$ are the marginal densities. The expression of copula density is given in (7.18) in McNeil et al. [2015],

$$c(u_1, ..., u_d) = \frac{\partial C(u_1, ..., u_d)}{\partial u_1, ..., \partial u_d}. \tag{2.1}$$

Note that not all the copulas have joint densities. The expression 2.1 is the copula density for copulas which are continuous in $[0, 1]^d$. The copula densities are calculated sometimes for fitting copulas to data by the maximum likelihood estimation method.

Proposition 2. *Let $(X_1, ..., X_d)$ be a random vector with continuous margins and copula C and let $T_1, ..., T_d$ be strictly increasing functions. Then, C is the unique copula of $(T_1(X_1), ..., T_d(X_d))$ as well.*

The proof of Proposition 2 is demonstrated in Proposition 7.7 in McNeil et al. [2015].

If C is the distribution function of $(U_1, ..., U_d)$ then the distribution function of $(1 - U_1, ..., 1 - U_d)$ is called the survival copula C. The survival copula is an application of Sklar's identity to multivariate survival functions, proposed in Sklar [1959]. In the case where $F_1, ..., F_d$ are continuous this identity can be written by

$$\begin{aligned} \bar{F}(x_1, ..., x_d) &= \mathbb{P}(X_1 > x_1, ..., X_d > x_d) \\ &= \mathbb{P}(1 - F_1(x_1) \leq \bar{F}_1(x_1), ..., 1 - F_d(x_d) \leq \bar{F}_d(x_d)) \\ &= \hat{C}(\bar{F}_1(x_1), ..., \bar{F}_d(x_d)), \end{aligned} \tag{2.2}$$

where $\bar{F}_i = 1 - F_i$ is the marginal survival function and $\hat{C}$ is the survival copula. Therefore, Equation 2.2 follows by writing $\hat{C}$ for the distribution function $\mathbf{1} - \boldsymbol{U}$, where $\boldsymbol{U} := (F_1(X_1), ..., F_d(X_d))^\top$ and $\mathbf{1}$ is the vector of ones in $\mathbb{R}^d$.

Another important property of copulas is radial symmetry.

Definition 2.3.2. *A random vector $\boldsymbol{X}$ (or its distribution function) is radially sym-*

metric about a point $\boldsymbol{a}$ *if* $(\boldsymbol{X}-\boldsymbol{a})$ *and* $(\boldsymbol{a}-\boldsymbol{X})$ *have the same distribution.*

Chapter 7.1.5 in McNeil et al. [2015] points out that an elliptical random vector $\boldsymbol{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \psi)$ is radially symmetric about $\boldsymbol{\mu}$. If $\boldsymbol{U}$ has a copula C as its distribution function, the only possible symmetric centre should be $(0.5, ..., 0.5)$, so C is radially symmetric if

$$(U_1 - 0.5, ..., U_d - 0.5) \overset{d}{=} (0.5 - U_1, ..., 0.5 - U_d) \Leftrightarrow \boldsymbol{U} \overset{d}{=} \mathbf{1} - \boldsymbol{U}. \tag{2.3}$$

Therefore, if a copula C is radially symmetric, its survival copula $\hat{C}$ is equal to itself. In other words, the 180 degree rotation of a radially symmetric copula C has exactly the same expression as C.

Definition 2.3.3 (Exchangeability)**.** *A random vector* $\boldsymbol{X}$ *is exchangeable if* $(X_1, ..., X_d)$ *has the same distribution as* $(X_{\Pi(1)}, ..., X_{\Pi(d)})$ *for any permutation* $(\Pi(1), ..., \Pi(d))$ *of* $(1, ..., d)$*.*

If a copula C is an exchangeable copula of an exchangeable random vector of uniform variates $\boldsymbol{U}$, it must satisfy that

$$C(u_1, ..., u_d) = C(u_{\Pi(1)}, ..., u_{\Pi(1)}) \tag{2.4}$$

for possible permutations; details may be found in Definition 7.16 in McNeil et al. [2015].

The conditional distributions of copulas are useful concepts of copulas. We focus on the bivariate copulas.

Definition 2.3.4 (Conditional distributions of copulas)**.** *If* (U_1, U_2) *has copula* C*, then its conditional distribution is*

$$C_{U_2|U_1}(u_2|u_1) = \mathbb{P}(U_2 \leq u_2 | U_1 = u_1) = \frac{\partial}{\partial u_1} C(u_1, u_2), \tag{2.5}$$

where the partial derivative exists almost everywhere (see Nelsen [2007] for details).

The conditional distribution is a distribution on the interval $[0, 1]$, which is only a uniform distribution when C is the independence copula. According to Definition 2.3.4 and the concept of exchangeability, if the distribution function of (U_1, U_2) is an exchangeable bivariate copula, then

$$\mathbb{P}(U_2 \leq u_2 | U_1 = u_1) = \mathbb{P}(U_1 \leq u_2 | U_2 = u_1), \tag{2.6}$$

which implies symmetry.

2.3.1 Tail dependence

Tail dependence is a measure of strength of dependence in the joint lower or joint upper tail of a multivariate distribution (Chapter 2.13 in Joe [2014]). This section focuses on the tail dependence of bivariate copulas.

The tail dependence coefficient is a measure of the strength of dependence in the tails of bivariate distributions. It is a concept of extremal dependence. The tail dependence coefficient is derived via a conditional probability. Hence, the range of the coefficient is between 0 and 1.

Assuming there are two random variables X_1 and X_2 with marginal distributions F_1 and F_2, the coefficient of upper tail dependence of X_1 and X_2 can be written as

$$\lambda_U := \lambda_U(X_1, X_2) = \lim_{q \to 1^-} \mathbb{P}(X_2 > F_2^{-1}(q) | X_1 > F_1^{-1}(q)),$$

where q is the quantile, provided $\lambda_U \in [0, 1]$ exists. Hence, the upper tail dependence is the probability that X_2 exceeds its q-quantile under the condition that X_1 exceeds its q-quantile. If $\lambda_u \in (0, 1]$, then we can say there is an upper tail dependence or extremal dependence of X_1 and X_2 in the upper tail; if $\lambda_U = 0$, then X_1 and X_2 are asymptotically independent in the upper tail. Similarly, the lower tail dependence coefficient can be written as

$$\lambda_L := \lambda_L(X_1, X_2) = \lim_{q \to 0^+} \mathbb{P}(X_2 \leq F_2^{-1}(q) | X_1 \leq F_1^{-1}(q)),$$

provided $\lambda_L \in [0, 1]$ exists.

If F_1 and F_2 are continuous, then we can obtain simple expressions for λ_U and λ_L in terms of the unique copula C of the bivariate distribution. The lower tail dependence can be written as

$$\lambda_L = \lim_{q \to 0^+} \frac{C(q, q)}{q}. \tag{2.7}$$

Copula C has lower tail dependence if $\lambda_L \in (0, 1]$ and no lower tail dependence when $\lambda_L = 0$. For upper tail dependence, the coefficient is

$$\lambda_U = \lim_{q \to 0^+} \frac{\hat{C}(q, q)}{q}, \tag{2.8}$$

where $\hat{C}$ is the survival copula of C. Copula C has upper tail dependence if $\lambda_U \in (0, 1]$ and no upper tail dependence when $\lambda_U = 0$. For radially symmetric copulas, the survival copula is equal to the copula. Hence, the upper and lower tail dependence are equal for radially symmetric copulas.

The concepts of tail dependence of bivariate copulas can be extended to multivariate

copulas, but this is beyond the scope of this thesis. Some details can be found in Chapter 2.13 in Joe [2014].

2.3.2 Rank correlations

Rank correlations are measures of dependence that depend only on the copula of a bivariate distribution and not on the marginal distributions. The rank correlation coefficients calculated from data only concern the ordering of the sample for each variable and not the actual numerical values.

The main reason for considering the rank correlations is that they can be used to calibrate copulas to empirical data. Generally, there are two main varieties of rank correlation, Kendall's and Spearman's. Both of them can be taken as a measure of concordance for bivariate random vectors. Assuming there are two points in $\mathbb{R}^2$, (x_1, x_2) and $(\tilde{x}_1, \tilde{x}_2)$, they are concordant if $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) > 0$ and they are discordant if $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) < 0$. This thesis applies the Kendall's tau to describe the rank correlations between variables.

Assume there are two random vectors (X_1, X_2) and $(\tilde{X}_1, \tilde{X}_2)$, which are independent of each other and have exactly the same distribution. If X_1 and X_2 have the same changing trend, then the probability of concordance is expected to be high compared to the probability of discordance. Conversely, if X_1 and X_2 have the opposite changing trend, the conclusion is expected to be opposite. Moreover, the probability of concordance will be 0.5, which is the same as the probability of discordance when X_1 and X_2 are independent. Hence, the expression of Kendall's tau can be simply written as

$$\tau(X_1, X_2) = \mathbb{P}((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0) - \mathbb{P}((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0). \tag{2.9}$$

Definition 2.3.5 (Kendall's tau)**.** *For continuous random variables X_1 and X_2, Kendall's tau is given by*

$$\tau(X_1, X_2) = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1, \tag{2.10}$$

where C is the unique copula of X_1 and X_2.

Details of Definition 2.3.5 are in Definition 7.31 of McNeil et al. [2015]. According to Fredricks and Nelsen [2007], via integration by parts,

$$\tau(X_1, X_2) = 1 - 4 \int_0^1 \int_0^1 C_{U_2|U_1}(u_2|u_1) C_{U_1|U_2}(u_1|u_2) du_1 du_2. \tag{2.11}$$

A proof of Formula 2.10 can be found in Proposition 7.32 of McNeil et al. [2015]. If we know the copula of the random variables, we can apply Formula 2.10 to calculate $\tau(X_1, X_2)$.

The sample version of Kendall's tau for data $(x_{i1}, x_{i2}), i = 1, ..., n$, (see Section 2.12.1 in Joe [2014]) is

$$\hat{\tau}(X_1, X_2) = \frac{4}{n(n-1)} \sum_{1 \le i < j \le n} I((x_{i1} - x_{j1})(x_{i2} - x_{j2}) > 0) - 1. \tag{2.12}$$

2.4 Examples of copulas

After introducing the properties of copulas, we present two commonly used parametric families of copulas, Gaussian copula and Archimedean copula. The applications of the properties are proposed in this section.

2.4.1 Gaussian copula

If $\boldsymbol{Y} \sim N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is a multivariate normal random vector, then its copula is called Gaussian copula. According to Proposition 2, the copula of $\boldsymbol{Y}$ is exactly the same as the copula of $\boldsymbol{X} \sim N_d(\mathbf{0}, \mathrm{P})$, where P is the correlation matrix of $\boldsymbol{Y}$. It is because the operation of standardizing the margins amounts to applying a series of strictly increasing transformations $X_i = \frac{Y_i - \mu_i}{\sqrt{\Sigma_{ii}}}$, where $i = 1, ..., d$.

The Gaussian copula of $\boldsymbol{X}$ is given by

$$\begin{aligned} C_P^{Ga}(\boldsymbol{u}) &= \mathbb{P}(\Phi(X_1) \le u_1, ..., \Phi(X_d) \le u_d) \\ &= \boldsymbol{\Phi}_P(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)), \end{aligned} \tag{2.13}$$

where Φ denotes the standard univariate normal distribution function and $\boldsymbol{\Phi}_P$ denotes the joint normal distribution function of $\boldsymbol{X}$. The notation C_P^{Ga} emphasizes that the copula is parameterized by the $\frac{1}{2}d(d-1)$ parameters of the correlation matrix (McNeil et al. [2015], p. 227). In two dimensions, the Gaussian copula can be written as C_ρ^{Ga}, where $\rho = \rho(X_1, X_2)$ and ρ denotes the Pearson correlation coefficients between X_1 and X_2.

The bivariate Gaussian copula can be expressed as an integral over density of $\boldsymbol{X}$,

$$C_\rho^{Ga}(u_1, u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi(1-\rho^2)^{1/2}} exp\left(\frac{-(s_1^2 - 2\rho s_1 s_2 + s_2^2)}{2(1-\rho^2)}\right) ds_1 ds_2. \tag{2.14}$$

The conditional distribution of Gaussian copula is:

$$C_{U_2|U_1}(u_2|u_1; \rho) = \Phi\left(\frac{\Phi^{-1}(u_2) - \rho\Phi^{-1}(u_1)}{\sqrt{1-\rho^2}}\right). \tag{2.15}$$

The Gaussian copula is radially symmetric. Hence, the 180 degree rotation of Gaussian copula is equal to itself, which means the survival copula of the Gaussian copula

is identical to the Gaussian copula itself.

The Gaussian copula is asymptotically independent in both the lower and upper tails. We calculate the coefficients of tail dependence for two-dimension copula as an example. By applying L'Hôpital's rule and using equation 2.5, the coefficients of the lower tail dependence are

$$\lambda_L = \lim_{q\to 0^+} \frac{dC(q,q)}{dq} = \lim_{q\to 0^+} \mathbb{P}(U_2 \le q|U_1 = q) + \lim_{q\to 0^+} \mathbb{P}(U_1 \le q|U_2 = q).$$

Since Gaussian copula is exchangeable, using formula 2.6, the expression can be written as

$$\lambda_L = 2 \lim_{q\to 0^+} \mathbb{P}(U_2 \le q|U_1 = q). \tag{2.16}$$

Due to the copula C is an elliptical distribution, its coefficient of upper tail dependence is equal to the coefficient of lower tail dependence. In order to evaluate the tail-dependence coefficient for Gauss copula C_ρ^{Ga}, let $(X_1, X_2) := (\Phi^{-1}(U_1), \Phi^{-1}(U_2))$, so (X_1, X_2) has a bivariate normal distribution with standard margins and correlation ρ. It follows from equation 2.16 that

$$\begin{aligned}\lambda_L = \lambda_U &= 2 \lim_{q\to 0^+} \mathbb{P}(\Phi^{-1}(U_2) \le \Phi^{-1}(q)|\Phi^{-1}(U_1) = \Phi^{-1}(q)) \\ &= 2 \lim_{x\to -\infty} \mathbb{P}(X_2 \le x|X_1 = x).\end{aligned}$$

According to Equation 2.15, using the fact that $X_2|X_1 = x \sim N(\rho x, 1-\rho^2)$, it can be computed that

$$\lambda_L = \lambda_U = 2 \lim_{x\to -\infty} \Phi(x\sqrt{1-\rho}/\sqrt{1+\rho}) = 0,$$

provided $|\rho| \le 1$. Therefore, the Gaussian copula is asymptotically independent in both tails; see Example 7.38 in McNeil et al. [2015] for details.

Furthermore, if random variables (X_1, X_2) have Gaussian copula function C, the Kendall's tau rank correlation for Gaussian copula is

$$\tau(X_1, X_2) = \frac{2}{\pi} \arcsin \rho. \tag{2.17}$$

Gaussian copula can describe both positive and negative dependence.

2.4.2 T copula

The t copula can be written as

$$C_{\nu,P}^t(\mathbf{u}) = \mathbf{t}_{\nu,P}(t_\nu^{-1}(u_1), ..., t_\nu^{-d}), \tag{2.18}$$

where t_ν is the df of a standard univariate t distribution with ν degrees of freedom. Similar to Gaussian copula, the P is a correlation matrix. The $\mathbf{t}_{\nu,P}$ is the joint df of the vector $\mathbf{X} \sim t_d(\nu, \mathbf{0}, P)$. When the df ν goes to infinity, the t copula will converge to the Gaussian copula. However, the t copula will not be independent copula when $P = I_d$ if $\nu < \infty$.

We focus on the bivariate t copula in this thesis. There are two parameters in the t copula, df (ν) and correlation coefficient (ρ). The Kendall's tau of the t copula is exactly the same as the one of Gaussian, which can be found in Equation 2.17. The tail dependence of the t copula is different with the Gaussian copula. An asymptotic tail dependence formula for the t copula can be found in Example 7.39 of McNeil et al. [2015], which is

$$\lambda = 2t_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right), \tag{2.19}$$

where $\rho > -1$. The lower and upper tail dependence are the same in the t copula, because the t copula is radially symmetric.

2.4.3 Archimedean copulas

Let $\psi[0,\infty) \to [0,1]$ be a decreasing, continuous function that satisfies the conditions $\psi(0) = 1$ and $\lim_{t\to\infty} \psi(t) = 0$. A d-dimensional Archimedean copula can be written as

$$C(u_1, ..., u_d) = \psi(\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)) \tag{2.20}$$

which is a copula if and only if ψ is convex. The function ψ is called the Archimedean generator and function.

In this section, we concentrate on the bivariate Archimedean copulas, which are written as

$$C(u_1, u_2) = \psi(\psi^{-1}(u_1) + \psi^{-1}(u_2)).$$

The well-known one-parameter Archimedean copulas include Joe, Gumbel, Frank and Clayton copulas, whose generators are shown in Table 2.1. We assume the parameter in the one parameter bivariate Archimedean copulas is θ.

Kendall's rank correlations can be computed for Archimedean copulas directly from the generator inverse using Proposition 3 below. The formula obtained can be applied to calibrate Archimedean copulas to empirical data using the sample version of Kendall's tau, as Equation 2.12. The estimation approach is applying the principle of the method of moments.

Proposition 3. *Let X_1 and X_2 be continuous random variables with a unique*

Copula	Parameter Range	Generator $\psi(t)$
Joe	$[1, \infty)$	$1 - (1 - e^{-t})^{1/\theta}$
Gumbel	$[1, \infty)$	$\exp -t^{1/\theta}$
Frank	R	$-\frac{1}{\theta}\ln(1 - (1 - e^{-\theta})e^{-t})$
Clayton	$[-1, \infty)$	$\max((1 + \theta t)^{-1/\theta}, 0)$

Table 2.1: Parameter ranges and copula generators of one parameter bivariate Archimedean. θ denotes the parameter in each copula. $\psi(t)$ represents the Archimedean copula generators.

Copula	τ	λ_L	λ_U
Joe	$1 + 2(2 - \theta)^{-1}[DG(2) - DG(2/\theta + 1)]$	0	$2 - 2^{1/\theta}$
Gumbel	$(\theta - 1)/\theta$	0	$2 - 2^{1/\theta}$
Frank	$1 + 4\theta^{-1}(D_1(\theta) - 1)$	0	0
Clayton	$\theta/(\theta + 2)$	$\begin{cases} 2^{-1/\theta}, & \theta > 0 \\ 0, & \theta \leq 0 \end{cases}$	0

Table 2.2: Kendall's tau and coefficients of tail dependence for one parameter bivariate Archimedean. θ denotes the parameter in each copula. τ represents the Kendall's tau and λ_L and λ_U denote the coefficients of lower and upper tail dependence, separately. DG is the *digamma* function which is introduced in the Section 4.7.1 in Joe [2014]. $D_1(\theta)$ is the Debye function of order one, defined by $D_1(\theta) = \int_0^\theta t/(exp(t) - 1)\ dt/\theta$.

Archimedean copula C generated by ψ. Then

$$\tau(X_1, X_2) = 1 + 4\int_0^1 \frac{\psi^{-1}(t)}{d\psi^{-1}(t)/dt}dt. \tag{2.21}$$

Proof. See Corollary 5.1.4 in Nelsen [2007].

The coefficients of tail dependence are easily calculated, since the Archimedean copulas are closed-form copulas. Values for Kendall's tau and the coefficients of tail dependence for the copulas in Table 2.1 are presented in Table 2.2.

The Kendall's tau of Joe copula in Table 2.2 is a simplification of a result of Schepsmeier [2010]. The Joe and Gumbel copulas can only describe positive dependence. Theoretically, Clayton copula can manifest negative dependence when $\theta < 0$, but it is hardly implemented in software. Because the density is 0 on the set $(u_1, u_2) : u_1^{-\theta} + u_2^{-\theta} < 1$, the extension is not useful for statistical modelling; see Section 4.6 of Joe [2014] for details. Hence, the Clayton copula can only describe positive dependence in practice. Frank copula contains negative dependence as well. The survival copula of Joe, Gumbel and Clayton copulas have exactly the same Kendall's tau as the copulas themselves, according to Equation 2.10. Hence, the 180 degree rotation of Joe, Gumbel and Clayton copulas share the Kendall's tau with themselves.

According to the coefficients of lower and upper dependence, the Joe and Gumbel

copula are asymptotically independent in lower tails, but have upper tail dependence, whose strength depend on the parameter θ. The Frank copula like the Gaussian copula, is independent in both tails. The Clayton copula has the opposite tail dependence of Gumbel copula. However, when $\theta \leq 0$, the Clayton copula can be asymptotically independent in both tails.

The Joe, Gumbel and Clayton copulas are not radially symmetric, which can be seen in their tail dependence. Frank is the only radially symmetric Archimedean copula as was proved in Frank [1979].

Furthermore, since the Joe, Gumbel and Clayton copulas only have positive dependence, we can use them to model negative dependence by rotating them at 90 or 270 degree. In addition, the Clayton copula can present the similar tail dependence coefficient as Joe and Gumbel via its 180-degree rotation. Therefore, we can model the tail dependence of data more accurately by applying their rotations.

2.4.4 Asymmetric bivariate copulas

The copulas referred to in Sections 2.4.1, 2.4.2 and 2.4.3 are symmetric bivariate copulas. Vine copulas with asymmetric tail dependence are introduced by Nikoloulopoulos et al. [2012]. In this section, the bivariate BB1 copula proposed by Joe and Hu [1996] is introduced as an example of an asymmetric bivariate copula. There are some other families of two-parameter copulas, such as the BB7 copula with a complex expression of Kendall's tau; these are beyond the scope of this section. We focus on the BB1 copula in this thesis.

The BB1 copula is a mixture of the Gumbel family and gamma Laplace transformation (deltails see Chapter 4.17 in Joe [2014]). The expression of this two-parameter copula family is

$$C(u, v; \theta, \delta) = \{1 + [(u^{-\theta} - 1)^{\delta} + (v^{-\theta} - 1)^{\delta}]^{1/\delta}\}^{-1/\theta}, \tag{2.22}$$

where $0 \leq u, v \leq 1$ and $\theta > 0, \delta \geq 1$.

The BB1 copula with two parameters, θ and δ, has asymmetric lower and upper tail dependence, λ_L and λ_U, which can be expressed as

$$\lambda_L = 2^{-1/(\delta\theta)} \text{ and } \lambda_U = 2 - 2^{1/\delta}, \tag{2.23}$$

where $0 < \lambda_L, \lambda_U < 1$.

The Kendall's tau of the BB1 copula is

$$\tau = 1 - \frac{2}{\delta(\theta + 2)}, \tag{2.24}$$

where $\theta > 0, \delta \geq 1$. When $\delta = 1$, the BB1 copula is the Clayton copula. This property

will help us to estimate the parameters in applications.

The BB1 copula is positively dependent, so we rotate it by 90 or 270 degrees when it is required to model negative dependence. Meanwhile, the asymmetry of the upper and lower tail dependence can improve the modelling of the time series with asymmetric tail dependence features. The 180-degree rotation of the BB1 copula can be used to make the copula more suitable for real data in some cases.

2.5 Partial copulas and partial correlations

The partial copula is a complex concept. It is an approach to remove the dependence on the conditional value, which is discussed in Bergsma [2004]. Before introducing the partial copula, we consider the random variables which are expressed by conditional distribution of copula

$$U_{1|2} := C_{1|2}(U_1|U_2) \qquad and \qquad U_{3|2} := C_{3|2}(U_3|U_2), \tag{2.25}$$

where U_1, U_2, U_3 are three variables whose joint distribution is a copula. These random variables are also called conditional probability integral transforms (CPIT), see the formula (3.32) in Czado [2019]. The distribution of $U_{1|2}$ is uniform since

$$\begin{aligned}
\mathbb{P}(U_{1|2} \leq u_{1|2}) &= \int_0^1 \mathbb{P}(C_{1|2}(U_1|U_2) \leq u_{1|2}|U_2 = u_2)du_2 \\
&= \int_0^1 \mathbb{P}(U_1 \leq C_{1|2}^{-1}(u_{1|2}|U_2)|U_2 = u_2)du_2 \\
&= \int_0^1 C_{1|2}(C_{1|2}^{-1}(u_{1|2}|u_2)|u_2)du_2 \\
&= \int_0^1 u_{1|2}du_2 = u_{1|2}
\end{aligned}$$

holds. Similarly, the random variable $U_{3|2}$ has a uniform distribution as well. The joint distribution of $(U_{1|2}, U_{3|2})$ is a copula; this copula is called the partial copula and is denoted by C^P. It is worth noting that the partial copula does not depend on a specific value of the conditional variable U_2 ([Czado, 2019, p. 66]). Besides, Gijbels et al. [2015] proves that the partial copula is a copula by definition.

To explain the definition of partial copula in a different way, we use the Gaussian copula as an example. Assume we have three random variables X_1, X_2, X_3 with multivariate Gaussian distribution. The copula of the distributions of X_1 and X_3 given X_2 is a Gaussian copula, which connects $X_1|X_2$ to $X_3|X_2$ in "partial copula". The correlation between the two conditional variables is partial correlation.

In the case of d variables, we can start by considering the dependence of any pair of copulas. The partial correlations can be used to study the dependence of two variables

after the effects of the remaining variables are removed.

Partial correlations can be interpreted as the correlations between the variables X_i and X_j conditional on the third variable X_l, where $0 \leq i, j, l \leq d$ and they are different to each other. For the calculation of the partial correlations, Yule and Kendall [1950] defined the following recursive formula.

Definition 2.5.1 (Recursion for partial correlations)**.** *Let $X_1, ..., X_d$ be random variables with mean zero and variance σ_i^2. Further denote by $I^d_{-(i,j)}$ the set $1, ..., d$ with indices i and j for $i \neq j$ removed. The corresponding $\binom{d}{2}$ partial correlations $\rho_{i,j;I^d_{-(i,j)}}$ satisfy the following recursions:*

$$\rho_{i,j;I^d_{-(i,j)}} = \frac{\rho_{i,j;I^{d-1}_{-(i,j)}} - \rho_{i,d;I^{d-1}_{-(i,j)}} \rho_{j,d;I^{d-1}_{-(i,j)}}}{\sqrt{1 - \rho^2_{i,d;I^{d-1}_{-(i,j)}}} \sqrt{1 - \rho^2_{j,d;I^{d-1}_{-(i,j)}}}}. \tag{2.26}$$

See Anderson [1958] for a derivation of this recursion.

After calculating the partial correlations, the partial Kendall's tau can be computed in Gaussian copula using the equation 2.17. Non-Gaussian copulas, such as Archimedean copulas, share the Kendall's tau with Gaussian copula. Then, we can use the expression of Kendall's tau in Table 2.2 and apply it to calculate the parameters for Archimedean copulas.

2.6 Vine copulas

The vine structure is introduced in this section. Vine models of dependence have been developed in a series of publications including Chapter 3 of Joe [2014], Cooke [1997], Kurowicka and Cooke [2006], Aas et al. [2009] and Smith et al. [2010]. Vine copulas are a flexible class of copula models. Joe [1996] and Joe [1997] propose that copula can be decomposed into a sequences of bivariate copulas. The decomposition is not unique. However, all possible decomposition can be expressed by a graphical model, called regular vine (R-vine), referred to in Bedford and Cooke [2002] and Bedford and Cooke [2001].

2.6.1 Vine structure

Vine structure is a complex concept. In order to explain it clearly, we start from a simple example. Assume there are four variables U_1, U_2, U_3, U_4, where U_i follows uniform distribution for $i = 1, ..., 4$. To study their dependence, we can use vine copulas in the class of pair copulas used to model each two variables. There are six options for pair copulas, $U_1U_2, U_1U_3, U_1U_4, U_2U_3, U_2U_4, U_3U_4$. In vine copulas, the pairs of variables will appear in different orders according to the shape of the vine. There are many

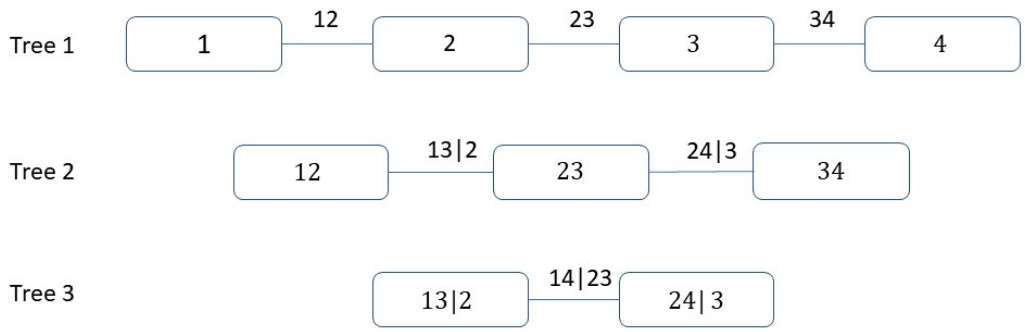


Figure 2.1: Vine Example 1: D-vine of four variables

distinctive types of vine structures, which are selected according to the features of data. We enumerate two typical and classical vines as examples.

Figure 2.1 and 2.2 are the examples of D-vine and C-vine, which are truncated at 4. In the two graphs, the vines consist of four hierarchies, which are called trees in vine structures. The elements in each tree are nodes, for example, variables $1, 2, 3, 4$ are the four nodes in Tree One. The connection of any pairs of nodes is called an edge. In Figure 2.1, the first tree has three edges and the edges of the first tree become the nodes of the second tree. The situation is exactly the same as the Figure 2.2. Furthermore, all the possible combinations of variable pairs appear in a vine without repetition. Moreover, these vine structures are not the only possible forms. There are other reasonable forms to construct the vine models, such as Figure 2.3, another form of C-vine. The construction of each tree and the order of elements in each tree are varied. Therefore, the vine models are very flexible.

To summarize the properties of vine, we follow the conclusion of Joe [2014]. Suppose we have a d-dimensional vine ($\mathcal{V}$), the vine should have the properties

- The vine consists of $d-1$ trees ($\mathcal{T}$), $\mathcal{V} = \{\mathcal{T}_1, ..., \mathcal{T}_{d-1}\}$.
- The number of edges is $d(d-1)/2$, which equals to the number of pairs of variables.
- Each pair of variables appears once as a conditioned set. The size of the conditioning set for any edge in $\mathcal{T}_l$ is $l-1$.
- Each edge $\boldsymbol{e}$ in $\mathcal{T}_l$ is a node in $\mathcal{T}_{l+1}$; the d nodes of $\mathcal{T}_1$ can be written as $1, ...d$.
- If edges are used to connect the nodes $\mathcal{T}_l$, then the result is a tree, which is a connected graph with no cycles.

The degree of a node is defined as the number of edges attached to that nodes. A regular vine is called a D-vine if all nodes in $\mathcal{T}_1$ have the degree not higher than two. We concentrate on the D-vine models in the following section.

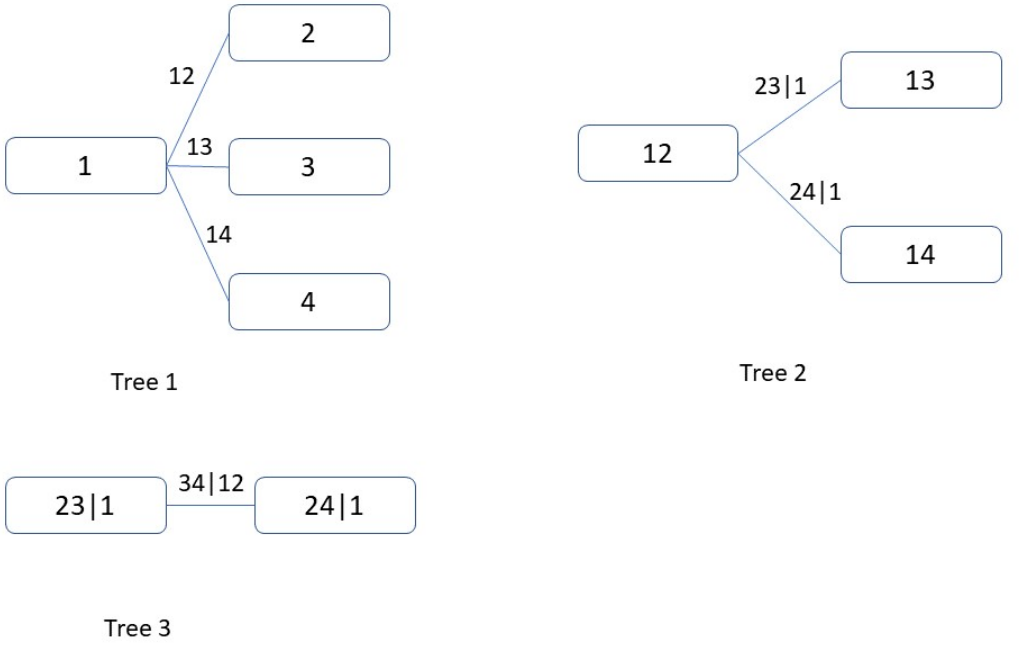


Figure 2.2: Vine Example 2: C-vine of four variables

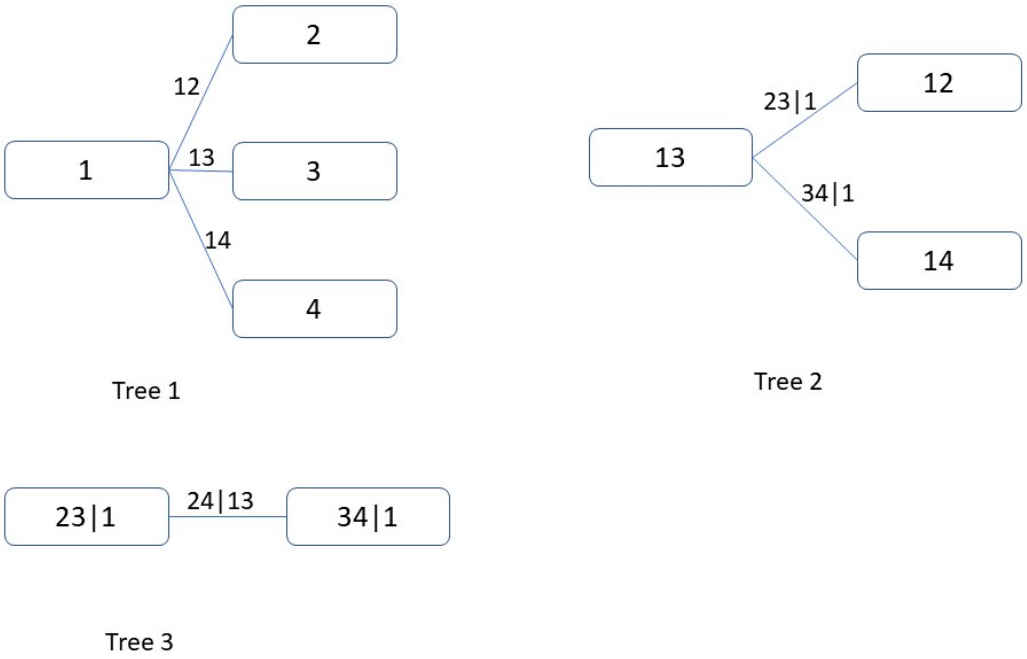


Figure 2.3: Vine Example 3: C-vine of four variables

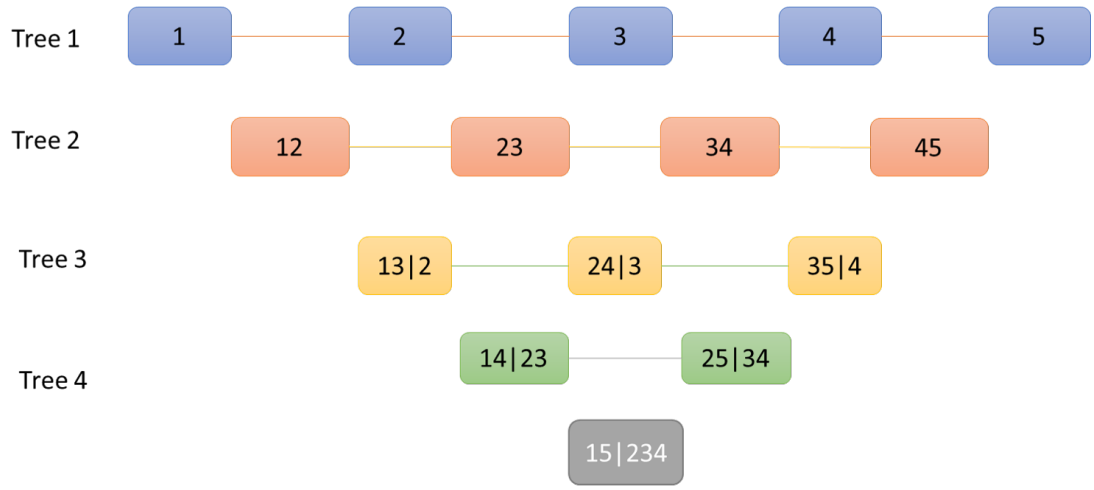


Figure 2.4: D-vine tree sequence in 5 dimensions.

Let S be a non-empty subset of $\{1, ..., d\}$, which is a conditioning set of variables. Given $1 \leq i, j \leq d$ and $i \neq j$, which is the conditioned set of variables. Besides, i, j are not included in set S. The copulas for the conditional distributions do not depend on the values $\boldsymbol{u}_S$ of the conditioning variables, the expression can be simplified by the assumption

$$C_{ij|S}(u_{i|S}, u_{j|S}; \boldsymbol{u}_S) = C_{ij|S}(u_{i|S}, u_{j|S}), \tag{2.27}$$

which means these pair copulas are independent of the conditioning variables; see Haff et al. [2010] for details.

A vine copula or pair copula construction is the resulting copula when a set of $\binom{d}{2}$ bivariate copulas are applied sequentially for a d-dimensional copula, constructed to satisfy the simplifying assumption 2.27 (See Section 3.9.3 in Joe [2014]).

2.6.2 D-vine copulas

Figure 2.4 demonstrates an example of a D-vine with five variables. The key feature of D-vine is that the edges of each tree only connect adjacent nodes, which make it simple to understand and apply.

In Figure 2.4, there are five variables in tree one. In the second tree, four copulas are required to describe dependence between adjacent variables $(1, 2), (2, 3), (3, 4), (4, 5)$. In tree three, three copulas depict the dependence of variables $(1, 3), (2, 4), (3, 5)$ conditional on $2, 3, 4$, respectively. The copulas in tree four accommodate the dependence of $(1, 4), (2, 5)$ conditional on $(2, 3)$ and $(3, 4)$. Finally, a copula describing dependence of $(1, 5)$ under condition of $(2, 3, 4)$ is required. Totally, ten pair copulas are required in a D-vine with five variables. The candidate copulas can be any types of bivariate copulas, such as Gaussian copula and the Gumbel copula. If the ten copulas are all Gaussian copulas, then the vine copulas will be Gaussian copulas as well, and vice versa.

Generally, each pair copulas in D-vine analyze a pair of variables conditional on the variables in-between. Each tree consists of a single path in D-vine copulas. The structure of D-vine is the most suitable for time series, because D-vine can describe strict stationarity of a random vector under some additional translation-invariance restrictions on the vine structures (Bladt and McNeil [2022]).

If a random vector has a joint density function $f(x_1, ..., x_n)$, the density function may be decomposed as D-vine. The decomposition (Bladt and McNeil [2022]) can be written as

$$f(x_1, ..., x_n) = \Big(\prod_{i=1}^{n} f_{X_i}(x_i)\Big) \prod_{k=1}^{n-1} \prod_{j=k+1}^{n} c_{j-k,j|S_{j-k,j}}(F_{j-k|S_{j-k,j}}(x_{j-k}), F_{j|S_{j-k,j}}(x_j)) \tag{2.28}$$

where $S_{j-k,j} = j-k+1, ..., j-1$ is sets including the variables in between X_{j-k} and X_j, f_{X_i} represents the marginal density of X_i, $c_{j-k,j|S_{j-k,j}}$ is the density of the bivariate copula $C_{j-k,j|S_{j-k,j}}$ of the joint distribution function of X_{j-k} and X_j conditional on the intermediate variables $X_{j-k+1}, ..., X_{j-1}$, and the conditional distribution function could be expressed by

$$F_{i|S_{j-k,j}}(x) = P(X_i \leq x | X_{j-k+1} = x_{j-k+1}, ..., X_{j-1} = x_{j-1}), i \in \{j-k, j\}. \tag{2.29}$$

The decomposition above implies a decomposition of the density $c(u_1, ..., u_n)$ of the unique copula of $(X_1, ..., X_n)$ which could be written by

$$c(F_1(x_1), ..., F_n(x_n)) = \prod_{k=1}^{n-1} \prod_{j=k+1}^{n} c_{j-k,j|S_{j-k,j}}(F_{j-k|S_{j-k,j}}(x_{j-k}), F_{j|S_{j-k,j}}(x_j)). \tag{2.30}$$

In practical applications, the simplified D-vine decomposition attracts more attention from researchers, where assuming the copula densities $c_{j-k,j|S_{j-k,j}}$ do not depend on the values of variables in the conditioning set $S_{j-k,j}$. Hence, we follow the notation used in Bladt and McNeil [2022] and write $c_{j-k,j}$. The expression of D-vine decomposition can be further simplified by additional translation invariant restriction.

2.6.3 S-vine copulas

S-vine copula is a stationary D-vine copula. In order to allow for simplifying the expression of S-vine, we impose strict stationary conditions (details in Nagler et al. [2022]) to D-vine to obtain stationary D-vine (S-vine). Under these conditions, the dependence among variables does not change with time shifting.

Let $(\mathcal{V}, C(\mathcal{V}))$ be a vine copula model. The vine copula model is called translation invariant if the partial copulas of variables in edge e are the same as the partial copulas of variables in the edge $e + \tau$, where τ indicates a shift in time by τ steps. Translation invariance is formally defined by Beare and Seo [2015] and used by Nagler et al. [2022], which proposes that translation invariance can guarantee stationarity.

Figure 2.4 exhibits an example of D-vine copula. According to the translation invariance condition for S-vine, the D-vine copula will be an S-vine copula if:

- Copulas describing dependence of variable pairs $(1, 2), (2, 3), (3, 4)$ and $(4, 5)$ are exactly the same;
- Partial copulas describing variable pairs $(1, 3), (2, 4), (3, 5)$ conditional on variables $2, 3, 4$, respectively, are the same;
- Partial copulas describing variable pairs $(1, 4), (2, 5)$ conditional on variable pairs $(2, 3), (4, 5)$, respectively, are the same.

Totally, the S-vine copula consists of four pair copulas. Therefore, compared with the D-vine copula requiring ten pair copulas, the translation invariance condition simplifies the vine models. If you make copulas in each tree of the D-vine copula equal, you will obtain an S-vine copula. This is useful for time series as we explained in the next chapter.

Chapter 3

S-vine Models for Time Series

This chapter presents a methodology for modelling and forecasting time series using S-vine models. The parametric S-vine model consists of an S-vine copula with Gaussian or non-Gaussian pair copulas and a parametric margin. In the semi-parametric model, the parametric margin is replaced by a non-parametric margin. The autoregressive moving-average (ARMA) model is widely applied in modelling many kinds of time series. Hence, we compare results from parametric S-vine models with non-Gaussian copulas and margins to ARMA models. As explained in Bladt and McNeil [2022], an ARMA model can be represented as an S-vine model (or process) with Gaussian pair copulas and a normal margin. Therefore, the comparison between non-Gaussian S-vines and Gaussian S-vines is a key concern is in this chapter.

The definition of stationary time series and application of S-vine models for time series are discussed in Section 3.1. The theory of S-vine processes is introduced in Section 3.2. The theory in the Sections 3.1 and 3.2 is mainly based on the theory proposed by Bladt and McNeil [2022, 2021]. The estimation methods of S-vine model are illustrated in Section 3.3. The maximum likelihood estimation method is used to estimate the parameters in Section 3.3. Section 3.4 presents the methodology and theory of one-step prediction. Moreover, simulations are required before the prediction of empirical study to compare the power of the tests and find the effective methods to distinguish between "good" and "bad" models.

3.1 Time series modelled by S-vine copulas

S-vine copulas provide a flexible framework for constructing stationary time series models to permit both non-Gaussian marginal behaviour and non-linear and non-Gaussian serial dependence behaviour. D-vine is the most suitable vine structure for time series data and longitudinal data. Nagler et al. [2022] study the D-vine structures that can be used to construct stationary multivariate time series. D-vines are the only vines for

which a translation invariance assumption is sufficient to construct a stationary time series, called S-vines. In our research, we follow them to refer to the restricted D-vines as stationary vines or S-vines.

S-vines have a distinctive structure. There is literature which investigates the application of S-vine in serial dependence of time series, including Darsow et al. [1992], Chen and Fan [2006], Domma et al. [2009], Beare [2010], Joe et al. [2010] and Nikoloulopoulos et al. [2012], who apply the first order Markov copula models as the simple examples of S-vine processes. Among other contributions, Joe et al. [2010] introduce the tail dependence functions of vine copula and Nikoloulopoulos et al. [2012] propose the vine copula with asymmetric tail dependence for time series. The higher order Markov models for univariate series or multivariate series are introduced in Brechmann and Czado [2015], Beare and Seo [2015], Nagler et al. [2022] and Loaiza-Maya et al. [2018]. Bladt and McNeil [2021] suggest that S-vine models could be generalized to infinite order and present how the finite or infinite copula sequences may be used to develop a non-linear structure to compete with linear processes, such as ARMA model.

There are two types of time series, univariate time series and multivariate time series. Univariate time series data are data for a single entity collected at multiple time periods. Multivariate time series contains multiple entities at multiple time periods. The multivariate time series considers not only temporal dependence but also the cross-sectional dependence. The temporal dependence exists within each component univariate time series. The application of vine copula used to model univariate time series includes Joe [1996], Aas et al. [2009] and Bedford and Cooke [2002]. There are also certain studies that employ vine copula to model temporal dependence in longitudinal data, such as Smith et al. [2010], Shi and Yang [2018] and Zhao et al. [2022]. In our study, we consider univariate time series data. The cross-sectional dependence is beyond the scope of our study. Hence, we focus on the higher order copula-based univariate time series models.

A time series (X_t) is stationary if its probability distribution does not change over time ([Stock et al., 2003, p. 587]). The formal definition of stationarity is as follows.

Definition 3.1.1 (strict stationary)**.** *The time series $(X_t)_{t\in\mathbb{Z}}$ is strictly stationary if*

$$(X_{t_1}, ..., X_{t_n}) \stackrel{d}{=} (X_{t_1+k}, ..., X_{t_n+k})$$

for all $t_1, ..., t_n, k \in \mathbb{Z}$ and for all $n \in \mathbb{N}$.

According to the stationary definition, the distribution of a stationary time series is invariant in time. The invariance reduces the complexity of models, but not all vines structures guarantee the stationary condition. If the vine copula $\mathcal{V}$ for time series $U_{t_1}, ..., U_{t_m}$ is the same as the vine copula $\mathcal{V}^k$ for time series $U_{t_1+k}, ..., U_{t_m+k}$, then the vine copula model $(\mathcal{V}, C(\mathcal{V}))$ is stationary for all translation invariant choices of $C(\mathcal{V})$, where $C(\mathcal{V})$ is the copula of vine $\mathcal{V}$ defined in Section 2.6.3.

Bladt and McNeil [2022] propose a strictly stationary stochastic processes, whose higher-dimensional marginal distributions are simplified D-vines. They force the marginal distributions of all variables to be identical and the D-vine copulas they apply are S-vine copulas, which requires that $c_{j-k,j}$, defined in Section 2.6.2, is the same for all $j \in \{k+1, ..., n\}$ and so each pair copula density in the model can be identified with its lag k and we can write $c_k := c_{j-k,j}$, where c_k is the density of some bivariate copula C_k. We apply the formula and method referred to by Bladt and McNeil [2022] to express the S-vine copula.

An S-vine copula density in dimension $d \geq 2$ can be written as

$$c_{(d)}(u_1, ..., u_d) = \prod_{k=1}^{d-1} \prod_{j=k+1}^{d} c_k \left(R^*_{k-1}(\boldsymbol{u}_{[j-k+1:j-1]}, u_{j-k}), R_{k-1}(\boldsymbol{u}_{[j-1:j-k+1]}, u_j) \right) \quad (3.1)$$

where $(c_k)_{k\in\mathbb{N}}$ are the densities of the bivariate copulas in the sequence of $(C_k)_{k\in\mathbb{N}}$. $R_k : (0,1)^k \times (0,1) \to (0,1)$ and $R^*_k : (0,1)^k \times (0,1) \to (0,1)$ are families of functions defined from $R_k : (0,1)^k \times (0,1) \to (0,1)$ in a recursion, where $R^*_1(u,x) = h^{(1)}_1(u,x)$, $R_1(u,x) = h^{(2)}_1(x,u)$ and

$$\begin{aligned} R_k(\boldsymbol{u}, x) &= h^{(1)}_k \left(R^*_{k-1}(\boldsymbol{u}_{[k-1:1]}, u_k), R_{k-1}(\boldsymbol{u}_{[1:k-1]}, x) \right) \\ R^*_k(\boldsymbol{u}, x) &= h^{(2)}_k \left(R^*_{k-1}(\boldsymbol{u}_{[1:k-1]}, x), R_{k-1}(\boldsymbol{u}_{[k-1:1]}, u_k) \right) \end{aligned} \quad (3.2)$$

for $k \geq 2$, where $h^{(i)}_k(u_1, u_2) = \frac{\partial}{\partial u_i} C_k(u_1, u_2)$ and $R_0(\cdot, u) = R^*_0(\cdot, u) = u$ for all u. The sequences of function R_k and R^*_k for $k \geq 1$ are named as forward and backward Rosenblatt functions by Bladt and McNeil [2022]. The function can be simplified if the copulas C_k are exchangeable for $k = 1, ..., d$ and $d \geq 1$. The forward and backward Rosenblatt functions are identical under this condition, where $R_k(\boldsymbol{u}, x) = R^*_k(\boldsymbol{u}, x)$. In this case, the recursions in Equation 3.2 can be written as

$$R_k(\boldsymbol{u}, x) = h^{(1)}_k \left(R_{k-1}(\boldsymbol{u}_{[k-1:1]}, u_1), R_{k-1}(\boldsymbol{u}_{[1:k-1]}, x) \right). \quad (3.3)$$

The Rosenblatt functions are applied in prediction. Hence, to explain them in more details, we demonstrate their expression in the conditional probability

$$\begin{aligned} R_k(\boldsymbol{u}, x) &= \mathbb{P}(U_t \leq x | U_{t-1} = u_1, ..., U_{t-k} = u_k) \\ R^*_k(\boldsymbol{u}, x) &= \mathbb{P}(U_t \leq x | U_{t+1} = u_1, ..., U_{t+k} = u_k). \end{aligned} \quad (3.4)$$

The forward functions are the conditional distribution functions of U_t given the previous values $\boldsymbol{U}_{[(t-1):(t-k)]}$. The derivatives of forward function $r_k(\boldsymbol{u}, x) = \frac{\partial}{\partial x} R_k(\boldsymbol{u}, x)$ are written as

$$
\begin{aligned}
r_k(\boldsymbol{u}, x) &= \frac{c_{(k+1)}(u_1, ..., u_k, x)}{c_{(k)}(u_1, ..., u_k)} \\
&= c_1(u_k, x) \prod_{j=2}^{k} c_j \left(R^*_{j-1}(\boldsymbol{u}_{[k-j+2:k]}, u_{k-j+1}), R_{j-1}(\boldsymbol{u}_{[k:k-j+2]}, x) \right).
\end{aligned} \tag{3.5}
$$

According to Bladt and McNeil [2022], the Rosenblatt forward functions have unique inverses $Q_k(\boldsymbol{u}, z)$ satisfying $R_k(\boldsymbol{u}, Q_k(\boldsymbol{u}, z)) = z$ for all $(\boldsymbol{u}, z) \in (0,1)^k \times (0,1)$. These inverse functions are called Rosenblatt quantile functions, which can be applied to sequentially generate the realizations from S-vine copulas.

3.2 S-vine processes

Definition 3.2.1 (S-vine process.)**.** *A strictly stationary time series $(X_t)_{t\in\mathbb{Z}}$ is an S-vine process if for every $t \in \mathbb{Z}$ and $n \geq 2$ the marginal distribution of the vector $(X_t, ..., X_{t+n-1})$ is absolutely continuous and admits a unique copula $C_{(n)}$ with a joint density $c_{(n)}$ of the form in Equation 3.1. An S-vine process $(U_t)_{t\in\mathbb{Z}}$ is an S-vine copula process if its univariate marginal distribution is standard uniform.*

The S-vine process is defined in Bladt and McNeil [2022]. In order to simplify the processes, we follow the paper's idea to set an identical bivariate copula sequence in S-vine models. However, as we referred in Section 2.4.3, if the bivariate copula are not comprehensive and cannot describe the negative dependence, then we use a rotation of the copula or replace it with a copula that can model negative dependence.

We can choose any continuous univariate marginal distribution for S-vine processes to satisfy the condition of S-vine copula. In this chapter, we focus on the estimation and prediction of S-vine processes. Examples of the S-vine processes are introduced in the following sections.

3.2.1 Gaussian processes

Gaussian processes are processes whose finite-dimensional marginal distributions are multivariate Gaussian. Bladt and McNeil [2022] prove that every stationary Gaussian process is an S-vine process. Every S-vine process in which the pair copulas of the sequence $(C_k)_{k\in\mathbb{N}}$ are Gaussian and the marginal distribution F_X is Gaussian, is a Gaussian process. Therefore, every Gaussian process has a unique S-vine copula representation.

Let $(X_t)_{t\in\mathbb{N}}$ be a stationary Gaussian process with autocorrelation function (acf) $(\rho_k)_{k\in\mathbb{N}}$. The mean, variance and acf uniquely determine a Gaussian process. Meanwhile, the acf uniquely determines the partial autocorrelation function (pacf) $(\alpha_k)_{k\in\mathbb{N}}$

via a one-by-one transformation (Barndorff-Nielsen and Schou [1973], Ramsey [1974]). Hence, the Gaussian process has a unique pacf, which is the correlation of the conditional distribution of (X_{t-k}, X_t) given the intermediate variables. The pair copulas in the S-vine copula representation are given by $C_k = C^{Ga}_{\alpha_k}$, where α_k is the pacf of Gaussian copula, following the expression in Bladt and McNeil [2022]. The one-to-one series of recursive transformations relating $(\alpha_k)_{k\in\mathbb{N}}$ to $(\rho_k)_{k\in\mathbb{N}}$ is $\alpha_1 = \rho_1$ and, for $k > 1$,

$$\alpha_k = \frac{\rho_k - \boldsymbol{\rho}_{k-1}^\top P_{k-1}^{-1} \overline{\boldsymbol{\rho}}_{k-1}}{1 - \boldsymbol{\rho}_{k-1}^\top P_{k-1}^{-1} \boldsymbol{\rho}_{k-1}}, \qquad \rho_k = \alpha_k(1 - \boldsymbol{\rho}_{k-1}^\top P_{k-1}^{-1} \boldsymbol{\rho}_{k-1}) + \boldsymbol{\rho}_{k-1}^\top P_{k-1}^{-1} \overline{\boldsymbol{\rho}}_{k-1}; \tag{3.6}$$

where the $\boldsymbol{\rho}_k = (\rho_1, ..., \rho_k)^\top$ is acf of Gaussian process, $\bar{\boldsymbol{\rho}}_k = (\rho_k, ..., \rho_1)^\top$ and P_k denotes the correlation matrix of $(X_1, ..., X_k)$. $P_1 = 1$ and for $k > 1$, P_k is asymmetric Toeplitz matrix whose diagonals are filled by the first $k-1$ elements of $\boldsymbol{\rho}_k$; in addition, P_k is non-singular for all k. For details, see Bladt and McNeil [2022], Joe [2006] or the Durbin-Levinson Algorithm (Brockwell and Davis [1991], Proposition 5.2.1). The restriction to non-singular Gaussian processes ensures that $|\rho_k| < 1$ and $|\alpha_k| < 1$ for all $k \in \mathbb{N}$.

ARMA processes are stationary and a Gaussian ARMA is the ARMA process with Gaussian innovations. Hence, the Gaussian ARMA process is a stationary Gaussian process, which can be represented as an S-vine process, according to Bladt and McNeil [2022]. Therefore, any causal Gaussian ARMA(p,q) model can be estimated by full maximum likelihood estimation using the joint density decomposition of an S-vine. Let $\boldsymbol{x} = \{x_1, ..., x_n\}$ be a realization from a strictly stationary process with parametric marginal distribution $F_X(x; \boldsymbol{\theta}_m)$ and joint copula density $c_{\boldsymbol{U}}(u_1, ..., u_n; \boldsymbol{\theta})$. The full log-likelihood can be calculated via

$$\begin{aligned} ln\ L(\boldsymbol{\theta}_c, \boldsymbol{\theta}_m; \boldsymbol{x}) &= ln(c_{\boldsymbol{U}}(F_X(x_1; \boldsymbol{\theta}_m), ..., F_X(x_n; \boldsymbol{\theta}_m); \boldsymbol{\theta}_c)) + \sum_{i=1}^{n} ln(f_X(x_i; \boldsymbol{\theta}_m)) \\ &= \sum_{i=1}^{k} \sum_{t=1}^{n-i} ln\ c_i \left(u_{t|S_{t,t+i}}, u_{t+i|S_{t,t+i}}; \boldsymbol{\theta}_i \right) + \sum_{i=1}^{n} ln(f_X(x_i; \boldsymbol{\theta}_m)), \end{aligned} \tag{3.7}$$

where $\boldsymbol{\theta}_i = (\theta_1, ..., \theta_i)$ and θ_i is the parameter of the ith pair copula. The terms $c_i \left(u_{t|S_{t,t+i}}, u_{t+i|S_{t,t+i}}; \boldsymbol{\theta}_i \right)$ depend on θ_i via the copula c_i and on $\theta_1, ..., \theta_{i-1}$ through the conditional distributions $u_{t|S_{t,t+i}}$ and $u_{t+i|S_{t,t+i}}$, which are defined in Section 2.5 and 2.6. In Gaussian ARMA model, $F_X(x_i) = \Phi(x_i)$ and $C_{\boldsymbol{U}}(; \boldsymbol{\theta}_c) = \boldsymbol{\Phi}(; P_k)$. Hence, $c_{\boldsymbol{U}}(; \boldsymbol{\theta}_c)$ can be computed via Equation 2.1.

For ARMA(p,q) with $q > 0$, the sequence of Gaussian pair copulas is infinite and hence, the order of the S-vine is infinite. For AR(p), the order is finite. There is an assumption for acf of Gaussian ARMA to ensure the Gaussian process to be a mixing

process and therefore ergodic (Cornfeld et al. [2012]; Maruyama [1970]), which is

Assumption 1. *The acf* $(\rho_k)_{k\in\mathbb{N}}$ *satisfies* $\rho_k \to 0$ *as* $k \to \infty$.

If $\boldsymbol{\phi} = (\phi_1, ..., \phi_p)^\top$, $\boldsymbol{\psi} = (\psi_1, ..., \psi_q)^\top$ and $\rho_k(\phi, \psi)$ denote the AR and MA parameters and the acf, then we apply the transformation between acf and pacf to parameterize in terms of ϕ and ψ using Gaussian pair copula $C_k = C^{Ga}_{\alpha_k(\phi,\psi)}$. The approach uses Gaussian innovations to estimate the parameters of Gaussian ARMA processes, where we select sequences of Gaussian pair copulas (C_k) parameterized by the sequences of partial correlations (α_k); see Bladt and McNeil [2022] for details. The estimation process in this approach can reduce the amount of the parameters. For example, an S-vine process with ARMA(p,q) has $p + q$ estimated parameters. If $q = 0$, the ARMA model will become autoregressive model with order p and the order of S-vine will be p. If $q > 0$, the process will include the moving-average part with order q and the order of S-vine will be infinity. In practice, the order of ARMA will be estimated from the time series first. Then, we fix the order p, q of ARMA in Gaussian S-vine process. The parameters $(\phi_1, ..., \phi_p, \psi_1, ..., \psi_q)$ are calculated via the relationship between $\boldsymbol{\phi}$, $\boldsymbol{\psi}$ and acf.

3.2.2 Non-Gaussian S-vine processes

The Gaussian ARMA process in Section 3.2.1 is a special case of S-vine models. In this section, we consider infinite-order S-vine copula processes constructed from general sequences $(C_k)_{k\in\mathbb{N}}$ of pair copulas that are not Gaussian. The marginal distributions of the non-Gaussian S-vine can be any continuous distributions introduced in Section 2.2. The pair copulas can be any continuous copulas. We take the sequence of pair copulas from parametric family and parameterize them to satisfy two requirements. The Kendall partial autocorrelation (kpacf) is identified to an ergodic Gaussian process and the copulas should converge uniformly to the independence copula as $k \to \infty$.

A practical approach is introduced in Bladt and McNeil [2022], which applies a well-behaved Gaussian process to approximate the pattern of decay of dependence in data and estimates the parameters in ARMA process via acf or pacf formula. Then, in order to construct a stable causal process, we select the non-Gaussian copulas in the S-vine process and calculate their parameters by connecting the parameters to the Kendall rank correlation function.

For all Archimedean copulas, the expression between the acf or pacf and parameters in ARMA is complex. However, the Kendall rank autocorrelation function of the copula sequence is explicit.The Kendall rank autocorrelation function of the copula sequence is defined in the following way.

Definition 3.2.2. *The Kendall partial autocorrelation function (kpacf)* $(\tau_k)_{k\in\mathbb{N}}$ *associated with a copula sequence* $(C_k)_{k\in\mathbb{N}}$ *is given by*

$$\tau_k = \tau(C_k), k \in \mathbb{N},$$

where $\tau(C_k)$ denotes the Kendall's tau coefficient for copula C_k.

In addition, the bivariate copulas selected in the Table 2.2 are one parameter copulas, which means the copula can be uniquely defined by one parameter.

For a Gaussian copula sequence with $C_k = C^{Ga}_{\alpha_k}$, the expression between kpacf and pacf is in Equation 2.17. As in the previous section, suppose that $(\alpha_k(\boldsymbol{\phi}, \boldsymbol{\psi}))_{k\in\mathbb{N}}$ is the pacf of a stationary and ergodic model Gaussian process parameterized by the parameters $\boldsymbol{\phi}, \boldsymbol{\psi}$ in ARMA model. Hence, this implies a parametric form for kpacf $(\tau_k(\boldsymbol{\phi}, \boldsymbol{\psi}))_{k\in\mathbb{N}}$ as well. And then, the non-Gaussian pair copulas sequences share the kpacf and their parameters can be calculated and the whole model is specified.

However, there is a practical problem caused by the range of kpacf τ_k, which should take any value in $(-1, 1)$. Only certain copulas, such as Gaussian or Frank satisfy the requirement. The copula sequences with Gumbel, Joe or Clayton copulas can only give the positive kpacf. In order to solve this problem, Bladt and McNeil [2022] suggest two approaches, one is to allow 90 or 270 degree rotations of the copulas at negative values of τ_k, the other is to substitute a comprehensive copula, such as Gaussian or Frank, at any position k in the sequence where τ_k is negative.

The ARMA processes used to estimate the pacf of the time series can be ARMA processes or seasonal ARMA processes if there are seasonality features in data. Assuming the order of ARMA is (p, q) and the order of S-vine is k, if $q = 0$, the finite set of values $\{\alpha_1, ..., \alpha_p\}$ yields an AR(p) model which is a special case of the finite S-vine models. The order k of S-vine will be equal to p in this case, because the level of dependence as measured by the Kendall correlation when $k > p$ will be equal to zero. Otherwise, if q is not equal to zero, the level of dependence as measured by the Kendall correlation will converge to zero with $k \to \infty$. The speed of convergence depends on the selected copula families. There is a comparison between speed of convergence of Gaussian, Frank, Clayton, Gumbel and Joe copulas in Bladt and McNeil [2022]. Additionally, the paper points out the models based on sequences of tail-dependent copulas, such as Gumbel, Joe and Clayton, which present slower convergence.

Furthermore, if the S-vine process has a finite order k and k is smaller than the order required for convergence, then the dependence level will be truncated when the lag is bigger than k, which means that the pacf is supposed to be 0 and variables are independent at lags greater or equal to $k + 1$.

3.3 Estimation of the S-vine models

The S-vine models have a hierarchical structure in the copulas part. Dependence between any pair of variables conditional on variables in between is described by a bivariate copula. The bivariate copula $(C_k)_{k\in\mathbb{N}}$ can be the Gaussian, Frank, Gumbel, Joe, Clayton or any other continuous bivariate copulas. In our research, we follow the method referred to in Bladt and McNeil [2022] to select the bivariate copulas with only one parameter to reduce the amount of parameters in the model. Furthermore, in order to improve the efficiency of the fitting process, especially when $n \to \infty$, the copula sequence includes only one or two types of copulas to capture both the positive and negative Kendall's tau values in time series. The details have been discussed in Section 2.4. The methodology of estimation is introduced in the next section.

There are two main kinds of S-vine models estimation methods, parametric or semi-parametric. The parametric S-vine estimation method consists of a parametric margin and an S-vine copula, which can be estimated as a whole model by maximum likelihood estimation (MLE). The semi-parametric S-vine estimation method consists of a non-parametric margin and a parametric S-vine copula, which can be estimated by pseudo maximum likelihood estimation (pseudo-MLE), where the empirical distributions are used as margins to transform data to uniform distribution, details may be found in Genest et al. [1995] and Shih and Louis [1995]. The semi-parametric S-vine can reduce the effect of error in the parametric margins and help us to observe the prediction performance of S-vine copulas independently. Chen and Fan [2006] investigate the semi-parametric estimation and prediction procedure for first-order Markov copula models.

The estimation of S-vine models includes two parts: to find the suitable marginal distribution and to fit the S-vine copula model to approximated uniform distributed data transformed by the chosen marginal distribution.

3.3.1 Estimation of marginal distribution

Parametric estimation of the marginal distribution

The estimation of parameters of marginal distribution is carried out using the MLE method (details may be found in [Rossi, 2018, p. 227]). The comparison between different marginal distributions is made by using the Akaike Information Criterion (AIC), which is a measure of relative forecasting quality. The smaller the AIC value is, the better the model is. Hence, we choose the margins with smaller AIC values to ensure that the transformed data are well approximated by a uniform distribution, and satisfy the assumptions of an S-vine copula process.

According to Proposition 1, the marginal distribution is the true distribution functions of random variables $X_1, ..., X_n$. Given data $x_1, ..., x_d$, the $u_1, ..., u_n$ should be the uniform distribution, where $U_i = F(X_i), i = 1, ..., n$. The transformed uniform data can

be called as probability integral transformation (PIT) values; see details in the Proposition 7.2 of McNeil et al. [2015]. Hence, we can check the feasibility of the marginal distribution through testing if the PIT values are uniformly distributed.

The parametric distribution can be chosen from a wide range of distribution functions. This section introduces the mainly used margins in this thesis only, including normal (norm), skewed Laplace (slap), skewed double Weibull (sdwe), skewed student (sst), hyperbolic (hyp), normal inverse Gaussian (NIG) and generalized hyperbolic (ghyp) distributions. The skewed distributions are chosen according to the family of skewed distributions proposed by Fernández and Steel [1998]. These parametric margins are the functions we have in the software; theoretically, we can choose any continuous distributions.

Non-parametric estimation of marginal distribution

Assume we have data vectors $(X_1, ..., X_n)$, the scaling empirical distribution function is one type of non-parametric margins, where the expression is

$$F_n(x) = \frac{1}{n+1} \sum_{t=1}^{n} I_{\{X_t \leq x\}}, \tag{3.8}$$

see Equation (7.54) in McNeil et al. [2015]. We use the scaling empirical distribution function to ensure the $F_n(x)$ is strictly in the range of $(0, 1)$.

The margins develop the PIT value of the original data as the input for the S-vine copulas. Hence, the selection of marginal distributions is very important. The good margins can yield uniform PIT values. Many kinds of data, particularly in finance, have heavier tails than normal distribution. In practice, the distributions that can describe tails better and transform data close to uniform distribution include sdwe, sst, hyp, NIG and ghyp distributions.

3.3.2 Estimation of S-vine copula processes

After transforming the original time series $(x_1, ..., x_n)$ on to an approximately uniform scale, we fit the S-vine copula process to the transformed data $(\hat{u}_1, ..., \hat{u}_n)$ to estimate the parameters. The parameters of S-vine copula are the parameters of ARMA(p,q) models that are used to express the kpacf of each conditional bivariate copulas. Then, the non-Gaussian copula sequences share the same kpacf values of Gaussian copulas, which are expressed by parameters in ARMA(p,q). Following this, we apply the MLE method estimating the parameters $\boldsymbol{\theta}$ in S-vine copulas via formulas in Table 2.2, where the parameters $\boldsymbol{\theta}$ are expressed by the parameters of an ARMA(p,q) process with the same Kendall's tau values as the S-vine copula process. Hence, there are $p+q$ parameters estimated by MLE.

In order to compare different combination of copulas in S-vines, we calculate AIC values of different combinations and select the ones with smaller AIC values.

The parameters of marginal distribution and S-vine copula can be estimated jointly in one step in the parametric S-vine process. Then, we calculate the AIC values of each estimation process and select the combination of margins and copulas with the smaller AIC values. Furthermore, we can apply the chosen copula sequences and parameters of S-vine copula obtained via the pseudo-MLE method to the MLE method. Because the pseudo-MLE method uses scaling empirical distribution, this can help to find the appropriate bivariate copula sequences and estimate their parameters with the reduction effect of marginal distribution. The results from the pseudo-MLE method can be used as the start values for the MLE approach. This can improve the efficiency of MLE methods.

3.4 Prediction of S-vine models

3.4.1 Methodology of one-step prediction

In this section, we consider one-step predictions. Assume we have $n+m$ data and intend to use the first n data to make m one-step predictions which can be compared with the final m observations. The Figure 3.1 shows the procedures of forecasting using moving window. The forecasting by moving window is presented in the following description. In this study, we decide to predict applying the moving window, since the increasing window does not improve the result and it requires more data which costs time in prediction.

Step 1 Divide the time series $X_1, X_2, ..., X_{n+m}$, into two parts, $X_1, X_2, ..., X_n$ used as training data, $X_{n+1}, X_{n+2}, ..., X_{n+m}$ used as test data.

Step 2 Fit S-vine models to the training data to estimate the parameters in the model.

Step 3 Use the fitted model and training data to predict the next value $\hat{X}_{n+1}$.

Step 4 Add the first test data X_{n+1} to the training data and use the new time series $X_1, X_2, ..., X_{n+1}$ and remove the first value X_1 of the time series. Then, use the new time series and the fitted model in the Step 3 to predict the next one-step predicting value.

Step 5 Similar to Step 4, add test data X_{n+i} to time series $X_{i-1}, X_2, ..., X_{n+i-1}$ and remove the X_{i-1} in the ith step and predict the next value $\hat{X}_{n+i+1}$. Finally, obtain a sequence of one-step predictions $\hat{X}_{n+1}, \hat{X}_{n+2}, ..., \hat{X}_{n+m}$.

The training data are used to estimate parameters of the applied model and the test data are used to evaluate its accuracy. The size of the test data is typically about

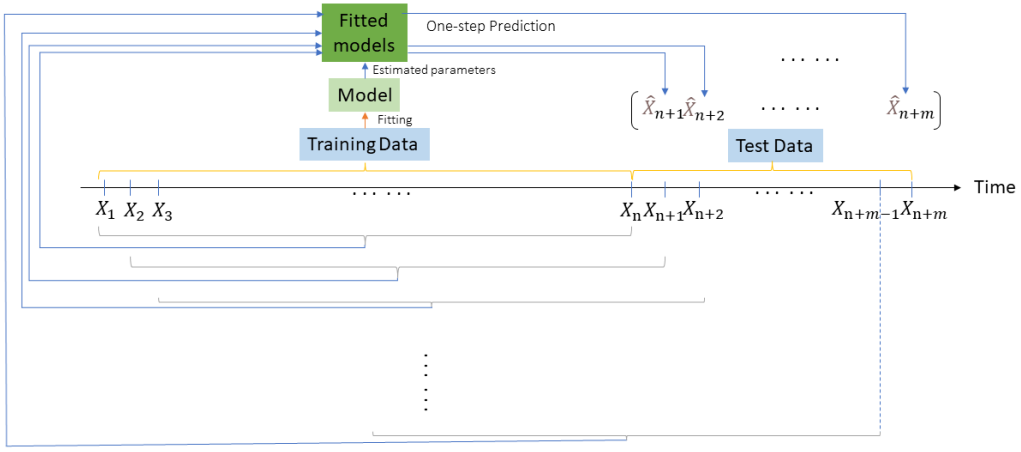


Figure 3.1: Diagram of One step prediction using moving window

20 percent of the total sample, although this value depends on how long the sample is and how far ahead you want to forecast, as discussed in Chapter 3 of Hyndman and Athanasopoulos [2018]. The length of the training data used for prediction in the moving window is fixed.

Moreover, the forecasting approach used above keeps the parameter estimates fixed. It keeps using the fixed models with the fitted parameters in Step 2. In order to observe the effect of using refitted parameters, we can also attempt to refit the testing data either at each step or after a fixed number of steps. For example, if we have quarterly data, refitting every four steps will be a reasonable choice. Details of refitting of the models will be discussed in the following section.

3.4.2 Predictive distributions

We study one-step predictive distributions, which can be described by their cumulative distribution functions (cdf), quantile functions (qf) and probability density functions (pdf). The prediction process uses the Rosenblatt transformation (Chapter 2 in Hofert et al. [2018]), which can transform the dependent uniform data to independent data. The Rosenblatt function used in the prediction is discussed in Bladt and McNeil [2022]. The expression is in Equation 3.4. In empirical study, if the size of training data n is larger than the lag of S-vine k, then the previous value used in the Equation 3.4 is equal to k, where the order of the S-vine model is $k-1$. However, if the size of training data n is smaller than k, then the lag of the S-vine will be n, since the input of the predictive function will be n each step. In our study, we choose $n > k$, so the notation of lag in

the following equation will be k.

The function $R_k(\mathbf{u}, x)$ is applied in the calculation of the cdf of the one-step prediction. The x can take any values between 0 and 1. Given an S-vine process $(X_t)_{t\in\mathbb{Z}}$, if we want to forecast X_{t+1} conditional on the previous k values, we use $F_{t+1|k}(x)$ as a predictive cdf. The predictive distribution function and its density and quantile function can be written as

$$\begin{aligned} F_{t+1|k}(x) &= R_k(F_X(\boldsymbol{x}_{[t-k+1:t]}), F_X(x)) \\ f_{t+1|k}(x) &= r_k(F_X(\boldsymbol{x}_{[t-k+1:t]}), F_X(x)) f_X(x) \\ F^{-1}_{t+1|k}(u) &= F_X^{-1}(R_X^{-1}(F_X(\boldsymbol{x}_{[t-k+1:t]})), u) \end{aligned} \tag{3.9}$$

where F_X and f_X are the marginal distribution function and density of the S-vine process and R_k and r_k are defined in Equation 3.2 and 3.5.

The PIT values behaving like samples of a predictive cdf should be iid uniform distribution if the prediction function is accurate; the inference may be found in Diebold et al. [1997].

In order to compare prediction results with the testing data, we apply the score function introduced by Gneiting and Raftery [2007] to evaluate forecasts of quantiles. The scores can quantify the discrepancy between the prediction and the observation from the distribution. The expected scores are calculated using score functions. Quantiles can minimize the expected scores. The risk measures that can minimize the scoring functions are said to be elicitable (Gneiting and Raftery [2007]). The formal definition of score function is as follows; details may be found in [McNeil et al., 2015, p. 356].

Definition 3.4.1. *A score function is a function $S : \mathbb{R} \times \mathbb{R} \to [0, \infty)$ that satisfies the following conditions for any $y, l \in \mathbb{R}$:*
(i) $S(y, l) \geq 0$ and $S(y, l) = 0$ if and only if $y = l$;
(ii) $S(y, l)$ is increasing for $y > l$ and decreasing for $y < l$;
(iii) $S(y, l)$ is continuous in y.

A real-valued statistical functional T defined on a space of distributions is elicitable if there exists a scoring function S satisfying two conditions for every F on the same space of T, (1)$\int_{\mathbb{R}} S(y, l)dF(l) < \infty, \forall y \in \mathbb{R}$,
(2)$T(F) = \arg\min_{y\in\mathbb{R}} \int_{\mathbb{R}} S(y, l)dF(l)$.
In this case, the scoring function S is said to be strictly consistent for T.

The quantile risk measure is elicitable for strictly increasing distribution functions, so we apply the scoring function for elicitable risk measures introduced in Proposition 9.8 of McNeil et al. [2015].

Proposition 4. *For any $0<\alpha<1$ the statistical functional $T(F_L) = F_L^{-1}(\alpha)$ is elicitable on the set of strictly increasing distribution functions with finite mean. The scoring*

function

$$S_\alpha^q(y,l) = |1_{\{l\leq y\}} - \alpha||l-y| \tag{3.10}$$

is strictly consistent for T.

According to Chapter 9 in McNeil et al. [2015], the approach of computing the empirical quantile score $\mathcal{Q}$ at α quantile is

$$\sum_{t=1}^{m} S_\alpha^q(\widehat{\mathcal{Q}}_\alpha^t, X_t),\ t = 1, ...m \tag{3.11}$$

where $\widehat{\mathcal{Q}}_\alpha^t$ is the t^{th} quantile prediction at quantile α based on information of the previous n realizations. X_t is the t^{th} realization of the testing data.

We use the Equation 3.11 and 3.10 to calculate the quantile scores at 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95-quantiles for the two models. However, there are seven quantile scores at seven quantiles, so it is necessary to find a scoring rules to consider the seven scenarios in one calculation. According to Gneiting and Raftery [2007], the scoring rule could be calculated by

$$S(F,x) = \int_0^1 S_\alpha(F^{-1}(\alpha); x)\nu(d\alpha), \tag{3.12}$$

where S_α denotes a proper scoring function for the quantile α and ν is a Borel measure on $(0,1)$. According to the Equation 3.12, if we assume that each score function has the same weight, the score rule can be computed as the average quantile score function at different quantiles. The method used to calculate the average quantile score is expressed in Equation 3.13.

$$E(S_\alpha(F^{-1}(\alpha); x)) = \frac{1}{m}\sum_{i=1}^{m} S_{\alpha_i}(F^{-1}(\alpha_i); x), \tag{3.13}$$

where m is the amount of chosen quantiles.

3.4.3 Simulations

Before applying the prediction methodology to real data, it is necessary to ensure that the evaluating method is proper and effective. Therefore, we conduct a simulation study to find an appropriate comparison method to distinguish between "good" and "bad" models in this section. Another question we explore in this section is the size of sample required to obtain meaningful comparisons between forecast models. The simulation is developed to ensure that the method used performs as we can expect and can thus find the effective tests to compare prediction results. In the simulation study, we use the non-parametric distribution to avoid the effect from margins. Hence,

the first step is transforming time series into approximately uniform distribution by empirical distribution. The two S-vine copula models used in this section are Gumbel with Gaussian copula sequences and Gaussian S-vine models. The Gumbel with a Gaussian copula sequence replaces the Gumbel copula in the sequence with Gaussian copula when the Kendall's tau becomes negative. We select the Gumbel with Gaussian copulas, because this combination in the class of the S-vine model usually present more accurate fitting result than other copulas.

We use the S-vine with Gumbel and Gaussian copula sequences with kpacf from ARMA(4,3) to generate 191 simulations, which corresponds to the amount of real data we will have in the application of Chapter 4. We choose the ARMA(4,3) that will be used to model UK inflation rates in Chapter 4 and the parameters are exactly the same as the ones used in UK inflation models. The simulated models are applied to mimic real data set. The dependence is strong which can be seen in the kpacf plots of the simulations (see Figure 3.2). Then, we use the first 151 simulations to fit the two models that we have intended to compare; one is the original model we used to generate data, the other is the Gaussian S-vine model. The AIC values of the two models are calculated. The AIC of Gumbel with Gaussian copula sequences is smaller than the benchmark model, because it is the true model that has been used to simulate the data (although the parameters are estimated).

Then, we use the fitted model and the other 40 simulations to do the one step predictions. The following one-step predictions are all forecast using the moving windows and without the refitting process. The PIT values are computed by predictive cdf via Equation 3.9, so as to be applied in the independence and uniformity tests in the next step. Meanwhile, the quantiles at 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95 are calculated via the quantile function of the predictive distributions. Also, the real quantiles at these quantiles are computed using the true models.

We found that the PIT values from the benchmark models do not fail tests of uniformity and independence, despite the fact that the benchmark model is not a "good" model in this case. Hence, a method that is used to distinguish between "good" and "bad" model is required. Therefore, we attempt the quantile score function.

In order to reveal the improved accuracy of the average quantile score, the similar simulation procedures are repeated 1000 times and the quantile scores at equal divided quantiles $0.05, 0.10, ..., 0.95$ and the average quantile scores computed by Equation 3.13 of the "good" and "bad" model are compared. If the "good" model has a lower score, it means the approach to comparing models is reasonable. Hence, we record the percentage of accuracy of the quantile score at all the 19 quantiles and the average quantile scores (AQS). By accuracy we mean the probability that the "good" model presents smaller quantile scores or average quantile scores than the "bad" model. The results of $\alpha = 0.05, 0.10, 0.25, 0.5, 0.75, 0.9, 0.95$ and AQS at all 19 quantiles are presented in Table 3.1.

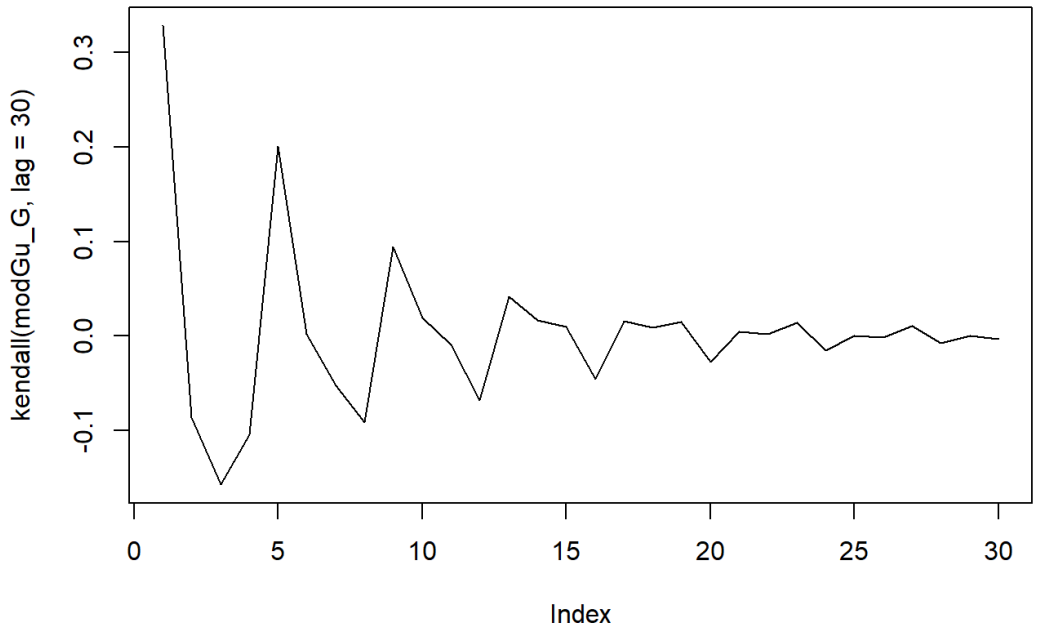


Figure 3.2: Kpacf plots of simulation by S-vine with Gumbel with Gaussian copula sequences and ARMA(4,3) order, whose parameters are $\phi = \{-0.178, -0.100, -0.423, -0.130\}$, and $\psi = \{0.728, 0.598, 0.691\}$.

Accuracy of quantile scores								
Quantiles	0.05	0.10	0.25	0.5	0.75	0.90	0.95	AQS
Accuracy	62.8%	66.3%	72.0%	77.0%	77.3%	77.0%	72.8%	79.9%
Quantile scores								
"Good" model	0.6785	1.1148	1.7713	2.5390	2.0709	1.0996	0.6268	1.4144
Benchmark	0.6771	1.1361	1.9713	2.5443	2.3321	1.4245	0.8321	1.5596

Table 3.1: Accuracy of the quantile scoring function and quantile score values of the benchmark model(Gaussian copula sequences in S-vine) and the "Good" model (Gumbel with Gaussian copulas sequences in the S-vine). The parameters in ARMA(4,3) used for kpacf estimation are $\phi = \{-0.178, -0.100, -0.423, -0.130\}$, and $\psi = \{0.728, 0.598, 0.691\}$.

Absolute values of differences to the real quantile							
Quantiles	0.05	0.10	0.25	0.5	0.75	0.90	0.95
"Good" model	0.6536	0.7997	0.9770	1.0640	1.1080	1.1628	1.1800
Benchmark	1.4134	1.7939	2.4525	3.0215	3.4427	3.5746	3.3973

Table 3.2: Absolute values of differences to the real quantile of the benchmark model(Gaussian copula sequences in the S-vine) and the "good" model (Gumbel with Gaussian copulas sequences in the S-vine). The parameters in ARMA(4,3) used for kpacf estimation are $\phi = \{-0.178, -0.100, -0.423, -0.130\}$, and $\psi = \{0.728, 0.598, 0.691\}$.

The quantile scores of the benchmark model should be greater than the ones of the true model at all the quantiles, but the $\mathcal{Q}_{0.05}$ of the true model is a little greater than the benchmark model. In this case, the AQS can be compared as a reference as well. The conclusion is almost in accordance with our expectation. The true models should have smaller AQS and quantile scores at most quantiles. Even though the sample size is not very big, the advantage of the true model in the quantile score function is evident still. Hence, the quantile score function and AQS is a reliable method to distinguish "good" and "bad" models.

Meanwhile, in order to present the difference between the real model and estimated models more clearly, we calculate the sum of the absolute values of differences between the real quantile and the estimated quantile. The real quantile in one-step prediction could be computed directly in the simulation. The estimated quantile at the six quantiles is predicted by the two models. The results are in Table 3.2.

The advantage of the true model is remarkable compared to the benchmark model in Table 3.2. Moreover, the plots of the one-step predicted quantile at 0.95 and 0.05 are presented in the Figure 3.3.

In Figure 3.3, the green line is close to the black line. Meanwhile, the trend of the red line is generally in accordance with the trend of the real quantile, but the changing range of the prediction from the benchmark model is smaller than the one from the real model. Hence, the S-vine model with Gumbel with Gaussian copula sequences forecasts better than the benchmark model. Moreover, the advantages of the "good" model are more significant at quantile 0.95 than at 0.05, which might be caused by the asymmetry of the Gumbel copula. Also, the benchmark model does better at quantile 0.05 than at 0.95. In practice, the real quantile is unknown, so it is impossible to use this method. The quantile score function might be an advisable option. Furthermore, if the practical time series have strong dependence, the tests we refer to in this section can also be applied as a method to exclude "bad" models. Besides this, practical data will use margins to transform the data into uniform scale data. The prediction results of parametric S-vines have relatively greater variations. These tests can be more useful

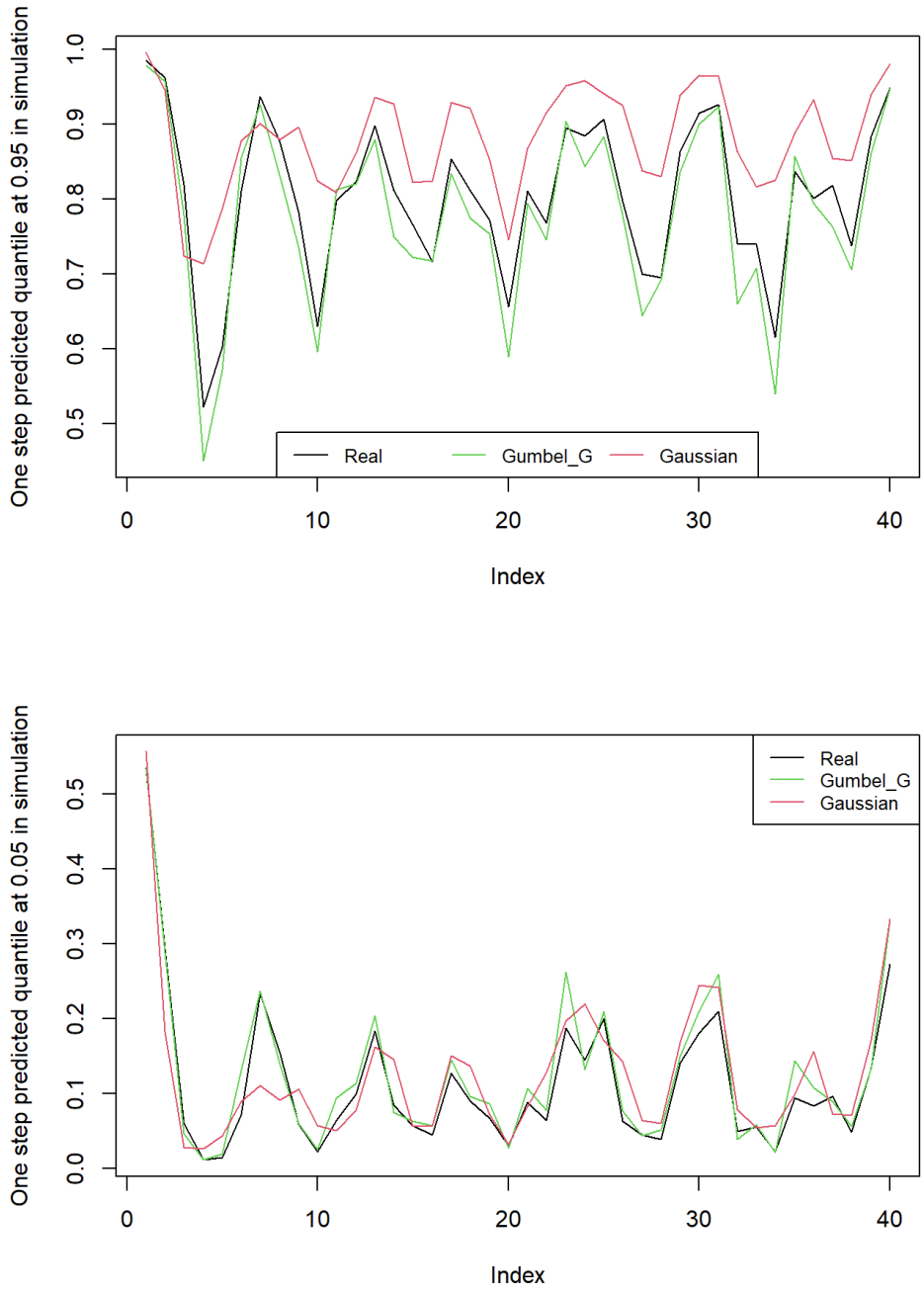


Figure 3.3: One step predicted quantiles at 0.95 and 0.05 by S-vine with Gumbel with Gaussian copula sequences and benchmark model.The computed real quantile is in black line. The red line is the quantile forecasted by benchmark model. The green one is the predicted quantile from Gumbel with Gaussian copula sequences.

to exclude bad cases in parametric S-vines.

The prediction with refitting processes every four steps is also attempted in the simulation data as well. Surprisingly, the refitting does not improve the results. The results are similar but slightly worse than the ones without refitting. The reason might be that there are noises in the refitting process with simulation data sets. In the simulation, the model we fit to the simulations is the model used to simulate these data. Therefore, fitting once may be sufficient to find the accurate parameters. It may generate a lot of noise if we repeat the fitting process. Hence, this phenomenon may only happen in the simulation. The refitting process still has a possibility to improve the prediction results in real data. Therefore, it is attempted, and the comparison with prediction without refitting is demonstrated in Section 4.6.

Chapter 4

Empirical Study of S-vine Models for Inflation Modelling

This chapter describes the empirical study of S-vine models for modelling and predicting inflation rates. The data we used are introduced in Section 4.1. The data processing of the consumer price index is discussed in Section 4.2. There are three methods we used to adjust the data, so as to make it satisfy the stationary condition of the S-vine model. The three approaches are presented in Section 4.3. Section 4.4 lists the choice of marginal distribution and copulas. The estimation results of semi-parametric and parametric S-vine models are shown in Section 4.5. The predictions by S-vine models for inflation rates are demonstrated in Section 4.6. Besides this, the method used to compare predictions from different models is discussed in Section 4.6. Finally, Section 4.7 summarizes the results of the empirical study and develops the possible further study topics.

4.1 Consumer price index

The consumer price index (CPI) measures the overall level in consumer prices over time based on a representative basket of goods and services. It can reflect the trend and the extent of changes of the price level of consumer goods and services of residents in a certain period. The CPI is the most widely used measure of inflation. Modelling and forecasting the inflation rate is important to households, businesses and policy-holders. Central banks aim to forecast and keep inflation stable and near a target level. Therefore, an accurate model of describing and predicting inflation is required.

Autoregressive moving average (ARMA) models are widely used for stationary macroeconomic time series. The linear structure of ARMA models is believed to describe the dependencies well. According to the work of Box et al. [1970], ARMA models have become a standard tool for modelling and forecasting univariate time series. There

are a lot of studies of inflation modelling and forecasting following the work in Box et al. [1970], including Stoviček [2007], Krukovets et al. [2019], Moriyama and Naseer [2009] and Kelikume and Salami [2014]. Marcellino et al. [2006] compares the direct and iterated multistep autoregressive (AR) methods for macroeconomic time series.

More recently, the coefficient instability in CPI inflation is investigated in Stock and Watson [2007], which introduces a time-varying integrated moving average (IMA) model for inflation. Huwiler and Kaufmann [2013] apply the disaggregate forecasts method for inflation using the autoregressive integrated moving average (ARIMA) model. Hassler and Wolters [1995] investigate the long-range dependence in the inflation rates and apply the fractionally integrated ARMA (ARFIMA) models to describe the fluctuations in data. There is a lot of research on variants of ARMA models, such as Zhang et al. [2020], which refers to applications of the stochastic volatility model with ARMA innovations for inflation forecasts. These studies discuss the non-stationary features in the inflation data. In order to simplify the modelling process, Huwiler and Kaufmann [2013] also take seasonal difference to make the inflation rates more stable and model the differenced inflation by ARMA model.

Furthermore, the seasonal ARIMA (SARIMA) models for inflation modelling are discussed in certain studies, such as Arteche [2007], Arlt [2021], Klutse [2020] and Ospina and Padilla Ospina [2019]. These researchers use the SARMA model to model and predict inflation rates to describe the seasonality in inflation more accurately.

Our study investigates the empirical application of S-vine models for CPI inflation rates and compares the results with ARMA model in order to improve both the modelling and prediction results. As noted, the classical ARMA model with Gaussian innovation is a special case of S-vine processes, called Gaussian processes. For this reason we take it as our benchmark model.

In this section, the empirical studies of S-vines with different kinds of bivariate copulas sequences and margins will be presented and the results will be compared to the benchmark model.

The bivariate copulas chosen in this section are Gaussian, Joe, Gumbel, Frank and Clayton. We also consider 180-degree rotations of copulas. For negative dependence, when the copula family is not comprehensive, we take the rotations through 90 or 270 degree, and we take Gaussian and Frank substitutions into consideration as well. According to Table 2.2, the Joe and Gumbel copulas both have upper tail dependence. In contrast, the Clayton copula has lower tail dependence. Hence, if we rotate the Clayton copula, we obtain a model with upper tail dependence as well.

There are nine kinds of margins compared in the empirical study: empirical, normal, skewed Laplace, skewed double Weibull, skewed student t, skewed hyperbolic student t, hyperbolic, Normal inverse Gaussian (NIG) and generalized hyperbolic distributions. The best margins are chosen according to the AIC values and Q-Q plots of margins in

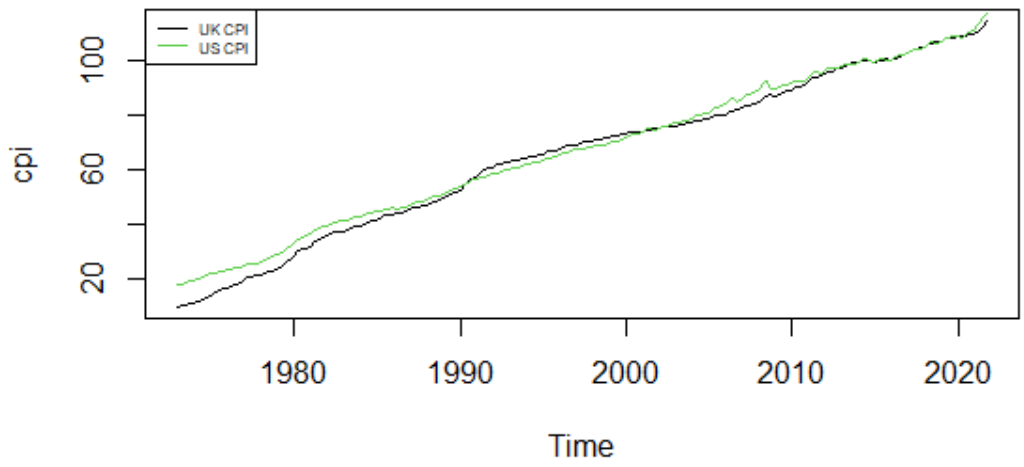


Figure 4.1: Quarterly UK and US CPI between 1973 Q1 and 2021 Q4

the first step of the empirical fitting process. Then, the stationary time series are transformed to uniform in the second step. The transformed times series should generally be uniformly distributed if the marginal distribution is a good model for the data.

4.2 Data processing for consumer price index

We use the Consumer Price Index (CPI) from two countries-the UK and the US. The UK and US data used to estimate the parameters are from the first quarter of 1973 to the fourth quarter of 2011. The other ten years of data (from the first quarter of 2012 to the fourth quarter of 2021) are used to validate the one-step predictions. The data are grouped according to the procedure described in Chapter 3.4 of Hyndman and Athanasopoulos [2018], which says the size of prediction data sets is typically 20 percent of the total sample. The reason we choose the data in this period is that the data excludes the volatility caused by the first oil crisis in 1973; see Hamid and Dhakar [2008] for details. Both of the two data sets are taken from the OECD website. The time series plots of the two sets of data are presented in Figure 4.1.

There are increasing trends in the two time series. Hence, the two time series are not stationary. In order to remove trend and seasonality, the first-order or second-order differences or seasonal differences can be taken. Also, the time series decomposition methods can be applied to remove the trend and seasonality and make the time series more stationary (details can be found in Hyndman and Athanasopoulos [2018]). Three methods are developed in this section.

We decided to take logarithm for each dataset in order to obtain the growth rate of CPI when we take the first order difference at lag one, which can be taken as the inflation rate. Meanwhile, a multiplier, 400 is used to scale the logarithm of CPI, in order to obtain an annualized measurement that can be interpreted (approximately) as a percentage. The explanation of taking logarithm and difference of macroeconomics is referred to in Chapter 14.6 in Stock et al. [2003]. We use y_t to denote the CPI and analyse x_t, the annualized inflation rate being expressed as a percentage, where

$$x_t = 400(\log y_t - \log y_{t-1}). \tag{4.1}$$

The annualized inflation from the UK and the US are shown in Figure 4.2.

The trend and seasonality may still exist in the annualized inflation rate in both the UK and the US according to Figure 4.2. The stationarity tests, Augmented Dickey-Fuller (ADF) (Dickey and Fuller [1979]) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests (Kwiatkowski et al. [1992]) are applied to confirm whether a stationary model is tenable. The null hypothesis of the ADF test is that the data require a model with a unit root, in other words, a non-stationary model. In contrast, the KPSS test null hypothesis is that the data can be explained by a model without unit roots, which is a stationary model. If the null hypothesis is rejected in the ADF test, while being accepted in the KPSS test, then we will have sufficient evidence to deem that the data are suitable to be modelled by stationary models. The p-values of the ADF test and the KPSS test are in Table 4.1. We assume that when the p-value is higher than 0.05, the null hypothesis cannot be rejected. Hence, the null hypothesis for UK inflation does not have enough evidence to be rejected by the ADF test and accepted by the KPSS test, which means that the UK inflation can be explained by a non-stationary model. Similarly, the null hypothesis of US inflation is almost rejected by both ADF and KPSS tests. The results of US inflation are conflicted in the two tests. The time series suitable for the stationary model should generate the same results in the two tests. Hence, US inflation may be adapted for the non-stationary model as well.

It is widely reported that there is seasonality in the inflation rate; some examples can be found in Osborn and Sensier [2009], Bataa et al. [2014] and Arteche [2007]. In our study, the data we used show the seasonal pattern in quarters; see box plots in Figure 4.3. Therefore, we attempt three approaches to remove the seasonality and make the inflation rates suitable for the stationary model in the next section.

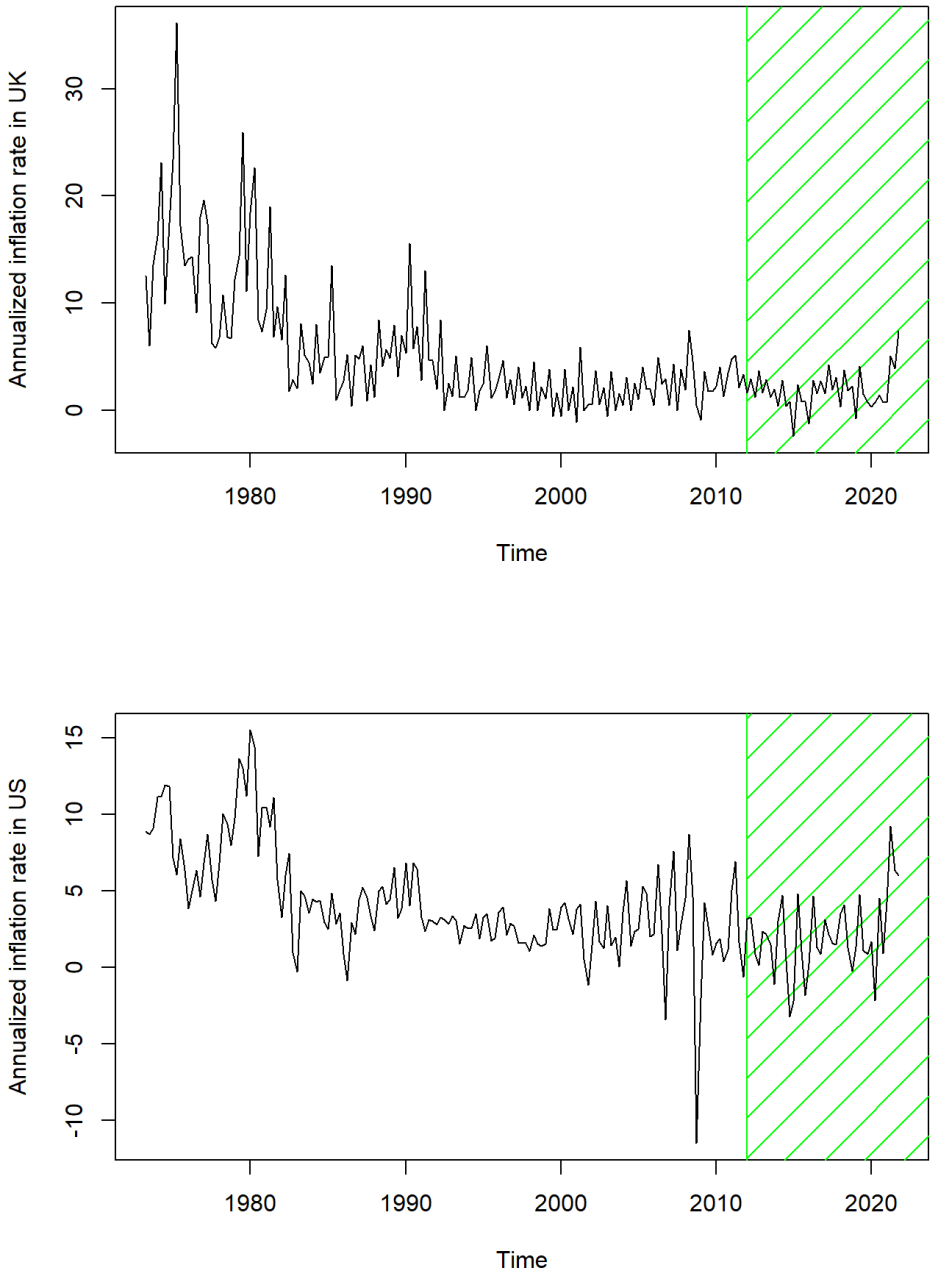


Figure 4.2: Annualized CPI inflation rate in UK and US between 1973 Q1 and 2021 Q4 (in percentage). The data from 1973 Q1 to 2011 Q4 are used as training data. The other data (in the green shaded area) are applied as the testing data.

p-values	ADF	KPSS
UK inflation	0.1220	0.01
US inflation	0.0296	0.01

Table 4.1: P-values of the ADF and KPSS test for UK and US inflation. The null hypothesis of ADF tests is that the data have a unit root, which is non-stationary. The null hypothesis of KPSS tests is that the data have no unit root, which is stationary.

4.3 Three methods to remove seasonality

The S-vine model is a stationary D-vine model, so the time series fitted by it should admit a stationary model. However, the original CPI data or the annualized inflation rate are not consistent with a stationary model. Therefore, it is necessary to try certain methods to transform the time series into stationary form. There are three approaches introduced in this section.

4.3.1 Method 1 : Model seasonal differenced inflation rates with ARMA

The seasonality of inflation rates in the UK and the US are not obvious in the time series plot. In order to present the seasonal features clearly, boxplots are applied to the two datasets; see Figure 4.3. The quarterly seasonality is demonstrated in the box plots, so we decide to take the seasonal difference at lag four to remove the seasonality. The seasonal difference is referred to by Hyndman and Athanasopoulos [2018] in Chapter 8, which says that the seasonal difference is the difference between an observation and the previous observation from the same season and can be written as

$$y'_t = y_t - y_{t-m}, \tag{4.2}$$

where m is the number of seasons and is referred to as the "lag-m differences".

Since the data we use is quarterly CPI, we pick four as the lag in the seasonal difference. The p-values of the UK and US inflation after being taken the seasonal difference are presented in Table 4.2. The null hypothesis of both the two time series are rejected by the ADF test and not rejected by the KPSS test, which means the two datasets are relatively suitable for the stationary models. The box plots of the seasonal differenced UK and US inflation rates are presented in Figure 4.4. The time series plots are demonstrated in the Figure 4.5. The seasonal differenced annualized quarterly inflation rates are without obvious trends and the seasonality is removed, which can be seen from the box plots. Hence, the adjusted time series can generally satisfy the stationary condition of S-vine models.

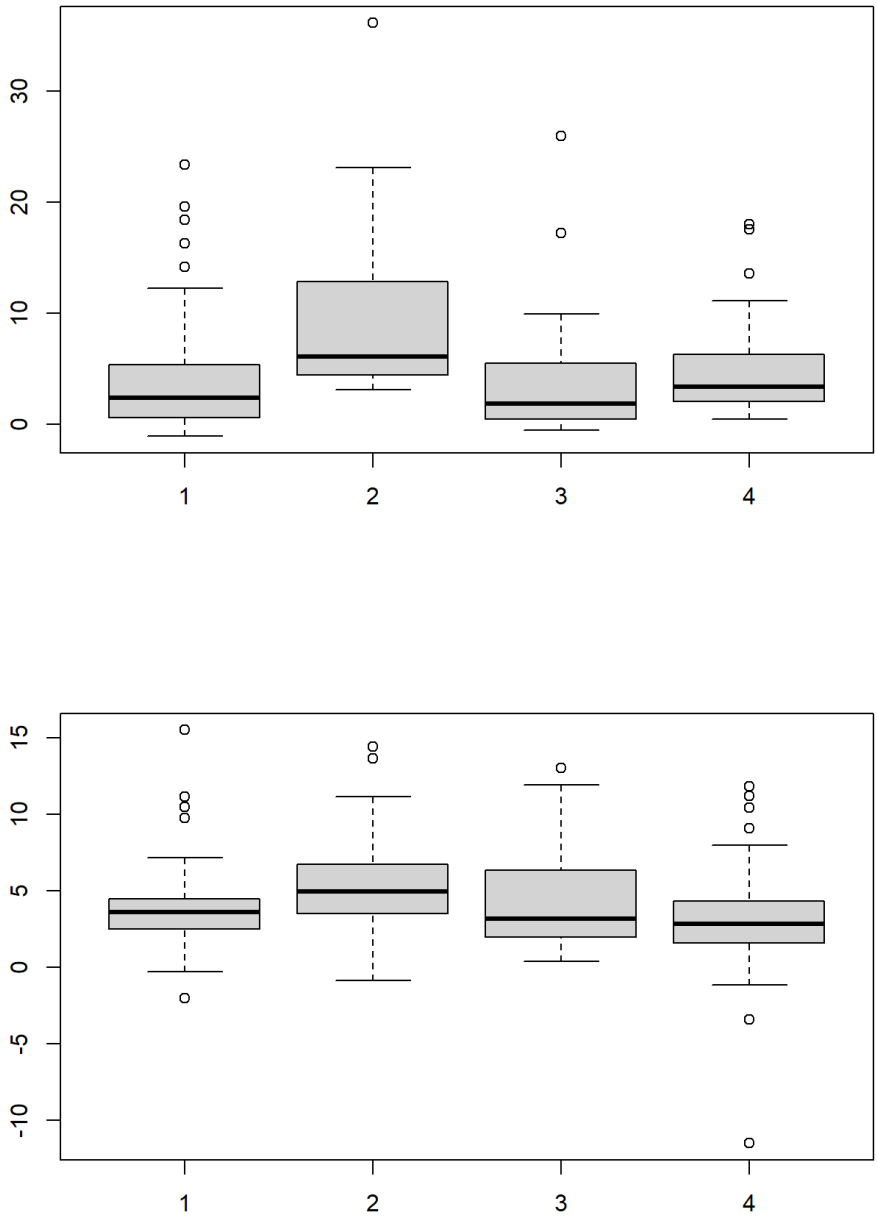


Figure 4.3: Box plots of the training data of annualized quarterly inflation rates in UK (top panel) and US (bottom panel) from 1973 Q1 to 2011 Q4

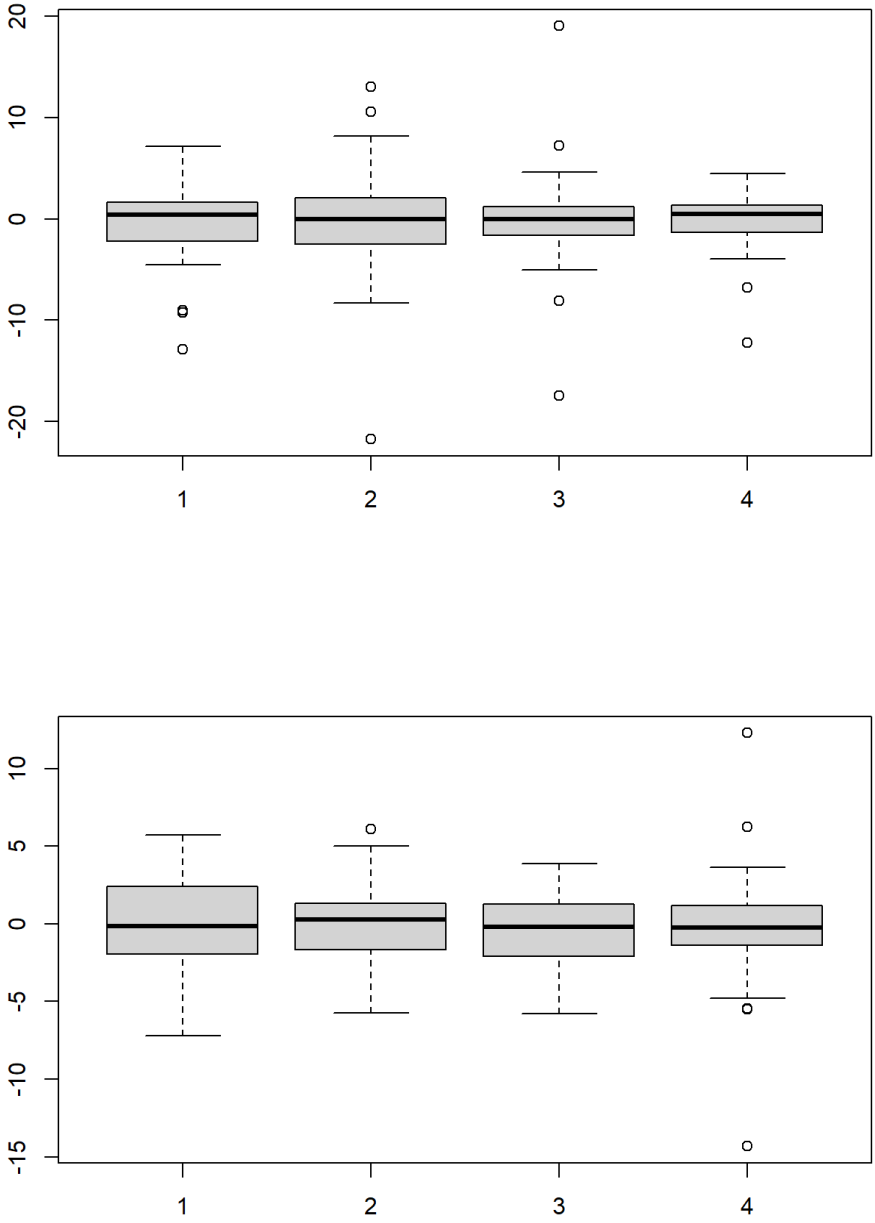


Figure 4.4: Box plots of the training data of seasonal differenced annualized quarterly inflation rates in UK (top panel) and US (bottom panel) from 1973 Q1 to 2011 Q4

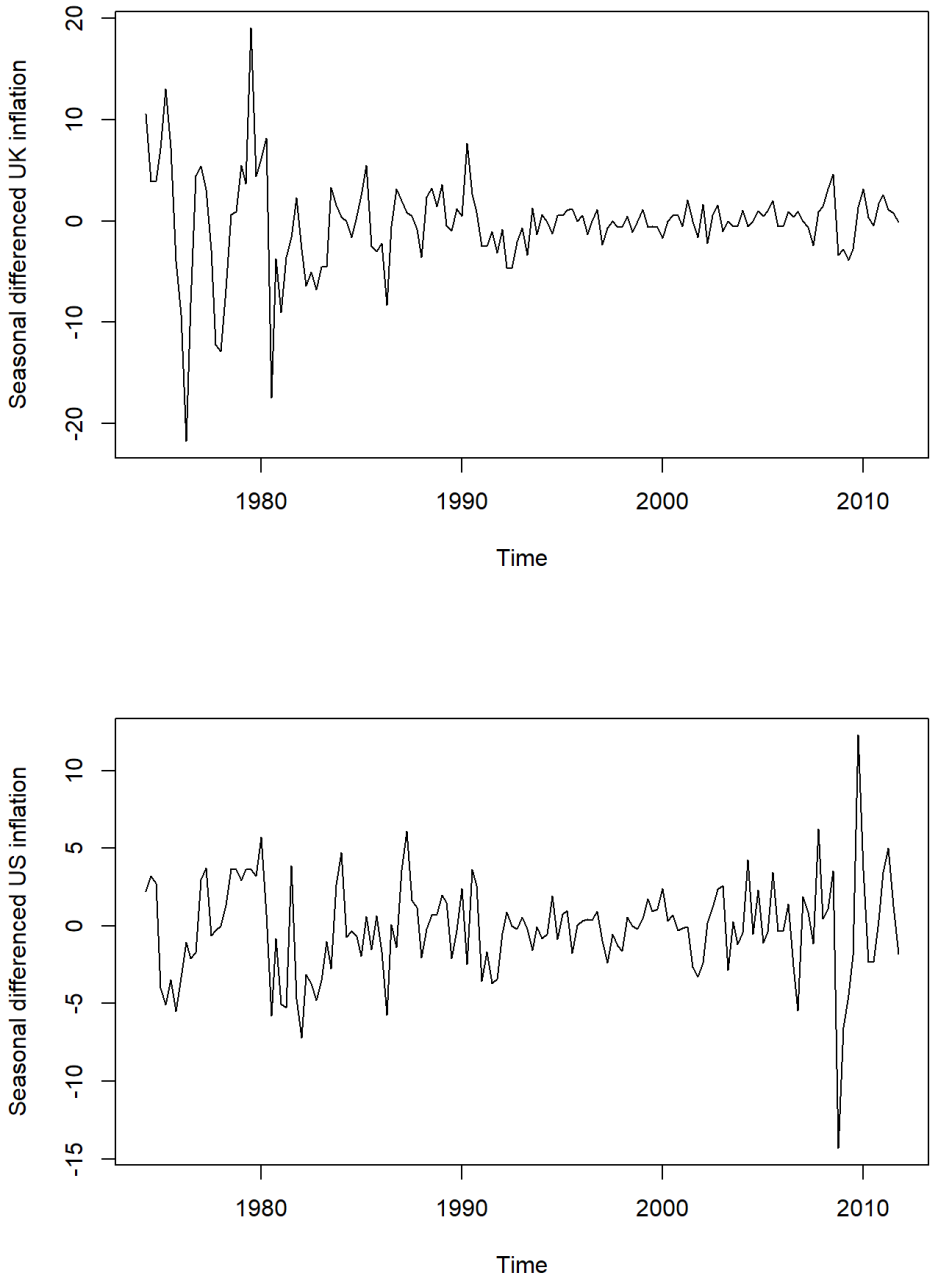


Figure 4.5: Time series plots of the training data of the seasonal differenced annualized quarterly inflation rates in UK (top panel) and US (bottom panel) from 1973 Q1 to 2011 Q4

p-values	ADF	KPSS
Seasonal diff UK inflation	0.01	0.1
Seasonal diff US inflation	0.01	0.1

Table 4.2: P-values of the ADF and KPSS test for seasonal differenced UK and US inflation. The null hypothesis of ADF tests is that the data have a unit root, which is non-stationary, while the null hypothesis of KPSS tests is that the data have no unit root, which is stationary.

4.3.2 Method 2: Model inflation rates with seasonal ARIMA

In Method 1, the seasonal difference is taken to remove the seasonality of inflation rates. In Method 2, the seasonality of the time series is described through the seasonal ARIMA in the S-vine structure. Therefore, the order of the difference depends on the order of the seasonal ARIMA.

In the second method, after taking seasonal difference, we use a stationary seasonal ARIMA(p,d,q)(P,D,Q)[m] ($d = 0, D = 0$), which can be termed as a seasonal ARMA. This is equivalent to a higher order ARMA with certain parameter constraints.

Example 1. *A seasonal ARIMA(1,0,0) (1,0,0)[4] is the ARIMA model with quarterly seasonality. Assume B represents the backward shift operator, which means*

$$BX_t = X_{t-1}.$$

In this example, the mathematical expression of ARIMA(1,0,0)(1,0,0)[4] is

$$(1 - \phi_1 B)(1 - \Phi_1 B^4)X_t = \varepsilon_t.$$

The bracket can be expanded to obtain the polynomial in the left hand side,

$$X_t - \phi_1 B X_t - \Phi_1 B^4 X_t + \phi_1 \Phi_1 B^5 X_t = \varepsilon_t.$$

According to the backward shift expression, the equation could be written as

$$X_t - \phi_1 X_{t-1} - \Phi_1 X_{t-4} + \phi_1 \Phi_1 X_{t-5} = \varepsilon_t.$$

The expression above is a mathematical expression of ARIMA(5,0,0) without X_{t-2} and X_{t-3}, and the parameter for X_{t-5} is the product of parameters for X_{t-1} and X_{t-4}.

Therefore, the seasonal ARMA models can be transformed to higher-order ARMA models, some details can be found in the Chapter 8 of the Hyndman and Athanasopoulos [2018]. We use this property to simplify our model. Moreover, in order to reduce the calculations of the transformation process and avoid generating an extremely high order ARMA model caused by integrated parts in the seasonal ARIMA, we expect that the

p-values	ADF	KPSS
Adjusted UK inflation (Method 2)	0.01	0.1
Adjusted US inflation (Method 2)	0.01	0.1

Table 4.3: P-values of the ADF and KPSS test for adjusted UK and US inflation via Method 2. The null hypothesis of ADF tests is that the data have a unit root, which is non-stationary, while the null hypothesis of KPSS tests is that the data have no unit root, which is stationary.

degree of differencing of the main and seasonal parts of ARIMA are zero ($d = 0, D = 0$), which can be realized by taking differences to data.

The first step is applying the auto.arima function in R, an automatic fitting procedure of seasonal ARIMA referred to in Hyndman and Athanasopoulos [2018] to find the degree of differencing in both the main part (d) and seasonal part (D). Then, we take the difference or seasonal difference to the data according to the suggested differencing order from auto.arima. Finally, the adjusted data can be modelled by the seasonal ARMA model. This step can reduce the amount of parameters produced by transformation of seasonal ARIMA to higher order ARMA, so as to avoid overfitting and increase the efficiency of the fitting process in S-vine models.

The next step is to find the reasonable order (p, q) of ARMA model via the auto.arima function in R. However, the auto.arima function using the stepwise and approximation method and the options within this function may not be comprehensive (details may be found in Chapter 8 of Hyndman and Athanasopoulos [2018]). In order to take more options in consideration, we also write our own functions to find the best order of ARMA or seasonal ARMA according to AIC and AICc. The AICc is similar to AIC, which is the AIC with a correction of small sample sizes. The comparison between AIC and AICc can be found in Burnham et al. [2011].

According to the results of fitting ARIMA(p,d,q)(P,D,Q)[4] process to annualized quarterly inflation rates (X_t), the selected differencing orders are $d = 0, D = 1$ in UK and $d = 1, D = 0$ in the US, respectively. Therefore, seasonal differences are applied to the UK annualized quarterly inflation rates, which using Equation 4.2 with $m = 4$. For US inflation rates (X_t), we take difference at lag one via Equation 4.2, where $m = 1$. The adjusted time series are in Figure 4.6, which can be modelled by the seasonal ARMA structure in the following study. The ADF and KPSS test results are presented in Table 4.3. The adjusted data are suitable for the stationary model according to the two test results.

4.3.3 Method 3: Time series decomposition of inflation rates

In the Chapter 6 in Dickey and Fuller [1979], it is stated that the time series (y_t) at time t can be thought of as comprising three components: a trend-cycle component (T_t), a

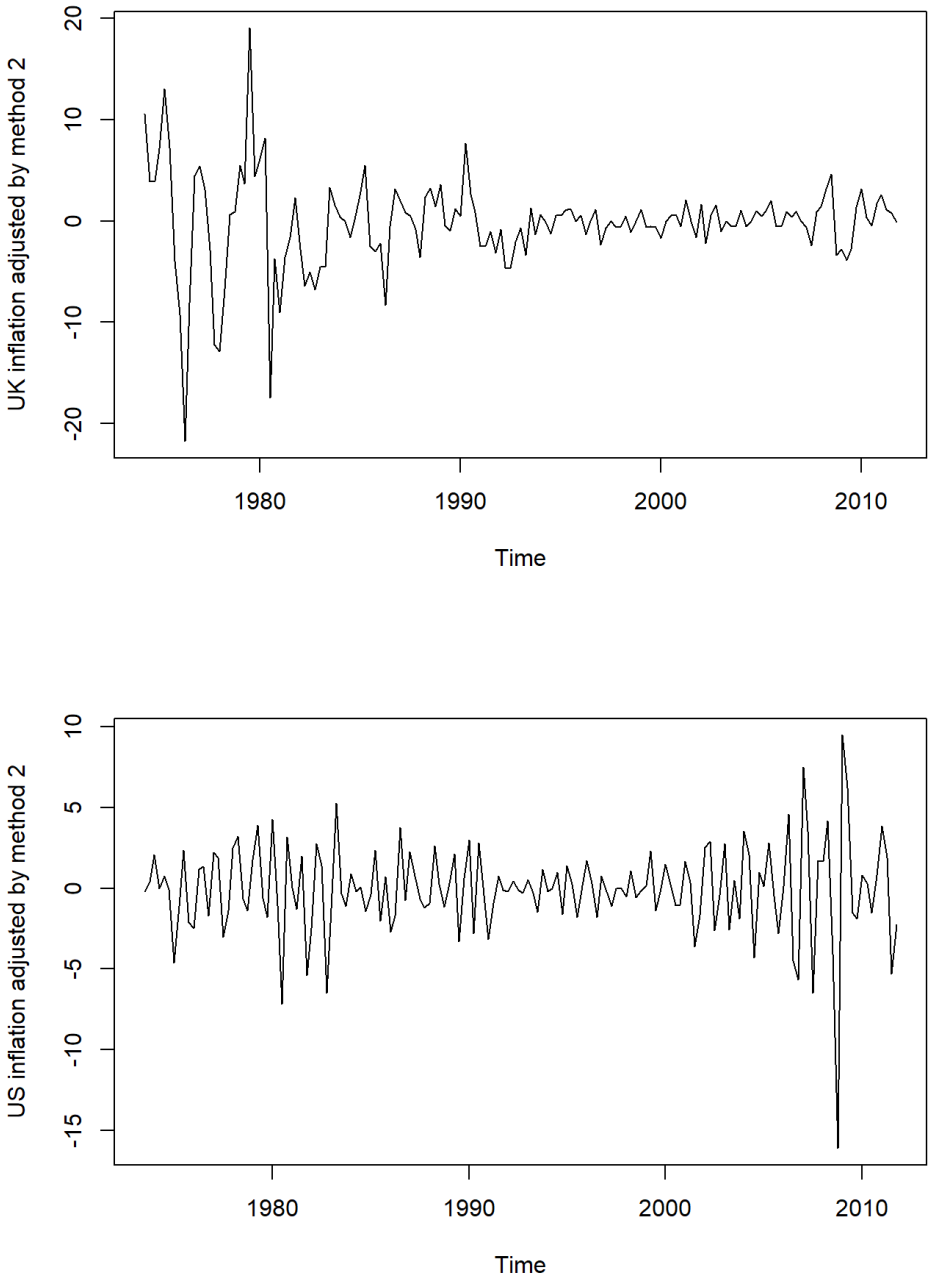


Figure 4.6: Time series plots of annualized quarterly UK (top panel) and US (bottom panel) inflation rates between 1973 Q1 and 2011 Q4 adjusted via Method 2

seasonal component (S_t) and a remainder component (R_t), which includes anything else in the time series.

An additive decomposition can be written as

$$y_t = T_t + S_t + R_t. \tag{4.3}$$

We apply the additive decomposition to $x_t = \log y_t - \log y_{t-1}$. In the CPI inflation rates, we assume that the time series (X_t) can be additively decomposed. Hence, we apply Equation 4.3. In the additive decomposition, the deseasonalized time series can be obtained by the original time series minus the seasonal component, which means

$$y_t - S_t = T_t + R_t.$$

The reason for removing the seasonal component instead of the trend component is that seasonality is usually easier to predict. There are many approaches applied to forecast seasonality, such as the seasonal naïve method, which sets each forecast to be equal to the last observed value from the same season (Hyndman and Athanasopoulos [2018], Section 3.1). However, the trend component is difficult to predict, so is the remaining component. It is unreasonable to decompose the trend out and predict it separately. Therefore, it is advisable to forecast the trend and remaining component as a whole via a time series analysis model and add back the predicted seasonality via a simpler method, such as the naïve method.

There are three main approaches for time series decomposition introduced in Chapter 6 of Hyndman and Athanasopoulos [2018], the X11 decomposition (described in Dagum and Bianconcini [2016]), the seasonal extraction in ARIMA Time Series (SEATS) decomposition (for details see Dagum and Bianconcini [2016] and the seasonal and trend decomposition using Loess (STL) decomposition (developed by Cleveland et al. [1990]). The X11 and SEATS decomposition are commonly used by banks (Dagum and Bianconcini [2016]). They are suitable for analyzing quarterly and monthly data. Alternatively, STL decomposition can handle any kind of seasonality, not only the quarterly and monthly data. Also, the STL decomposition is robust to outliers. The disadvantages of STL decomposition are that it does not handle trading day or calendar variations automatically, they need to be set by the users. The contents of the three decomposition methods are beyond the scope of this thesis; details can be found in the Dagum and Bianconcini [2016].

In the following study, we choose the decomposition methods according to the recommended order from the automatic fitting process of ARMA (auto.arima) for the deseasonalized annualized inflation rates, the robustness of decomposition and the efficiency of calculation. In practice, it is important to have a strong dependence structure in ARMA part, so that the relatively big Kendall's rank autocorrelation helps to forecast

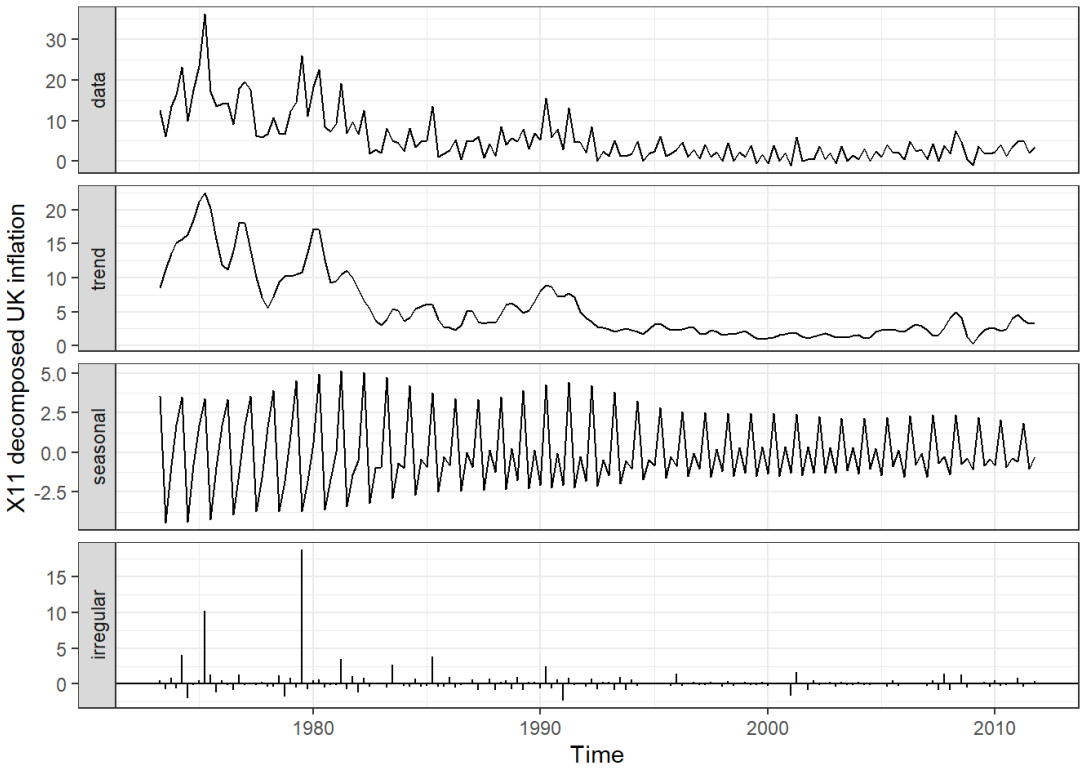


Figure 4.7: The X11 decomposition of annualized quarterly UK inflation rates from 1973 Q1 to 2011 Q4

data, which can embody the differences between distinctive types of bivariate copulas more clearly. Furthermore, a robust decomposition can improve the accuracy of predictions. In addition, the faster and easier operating methods will be preferred.

According to the conditions above, X11 decomposition is applied to the inflation rates of the US and the UK, because X11 decomposition provides better order and stronger dependence of ARMA structure. The decomposition results are presented in Figure 4.7 and Figure 4.8. The inflation rates (X_t) are divided into three components. Both of the UK and the US results have remarkable seasonality, which changes regularly with time. Hence, the seasonality can be easily forecasted by the naïve method. However, the other two components- trend and irregular parts-do not have obvious patterns, which are difficult to predict according to the regularity. Therefore, we extract trend and remainder and model their summation, which are termed as deseasonalized time series, by ARMA process.

The deseasonalized time series of the UK and the US are shown in Figure 4.9, which demonstrates trends contained in the decomposed time series. Therefore, it is necessary to remove the trends from the deseasonalized data by taking differences from them. Similarly, the order of difference is set according to the automatic fitting procedure of ARIMA and the function we write in R for finding the best order of ARMA. Both of

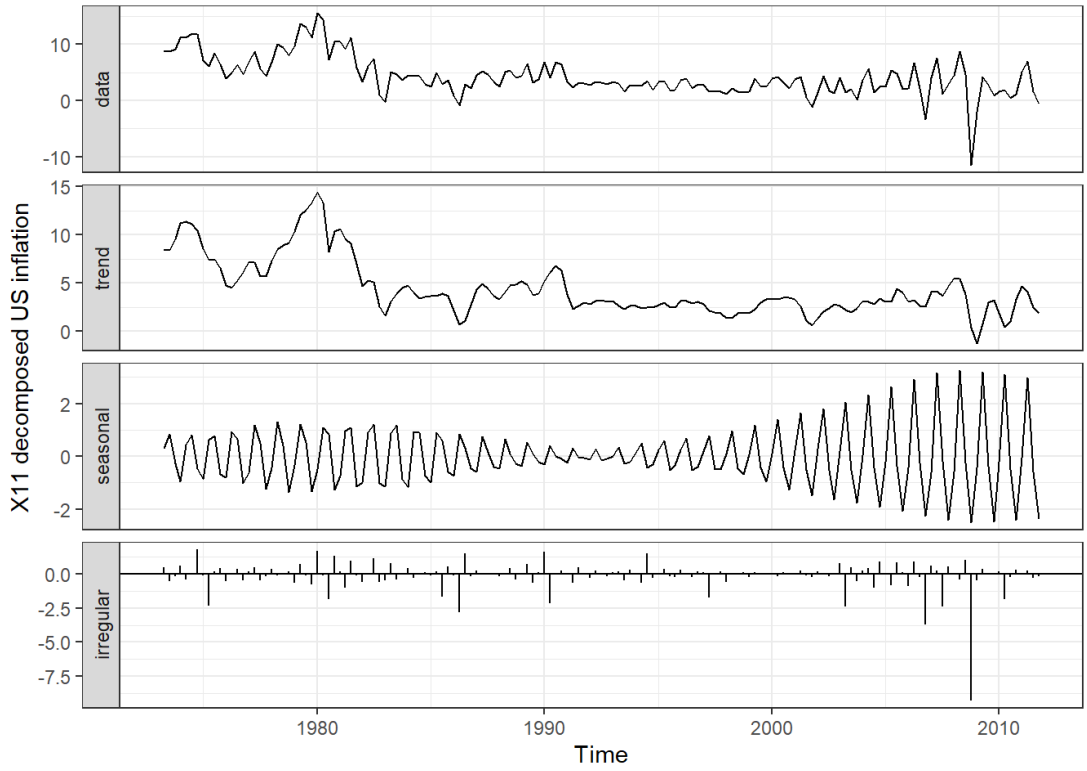


Figure 4.8: The X11 decomposition of annualized quarterly US inflation rates from 1973 Q1 to 2011 Q4

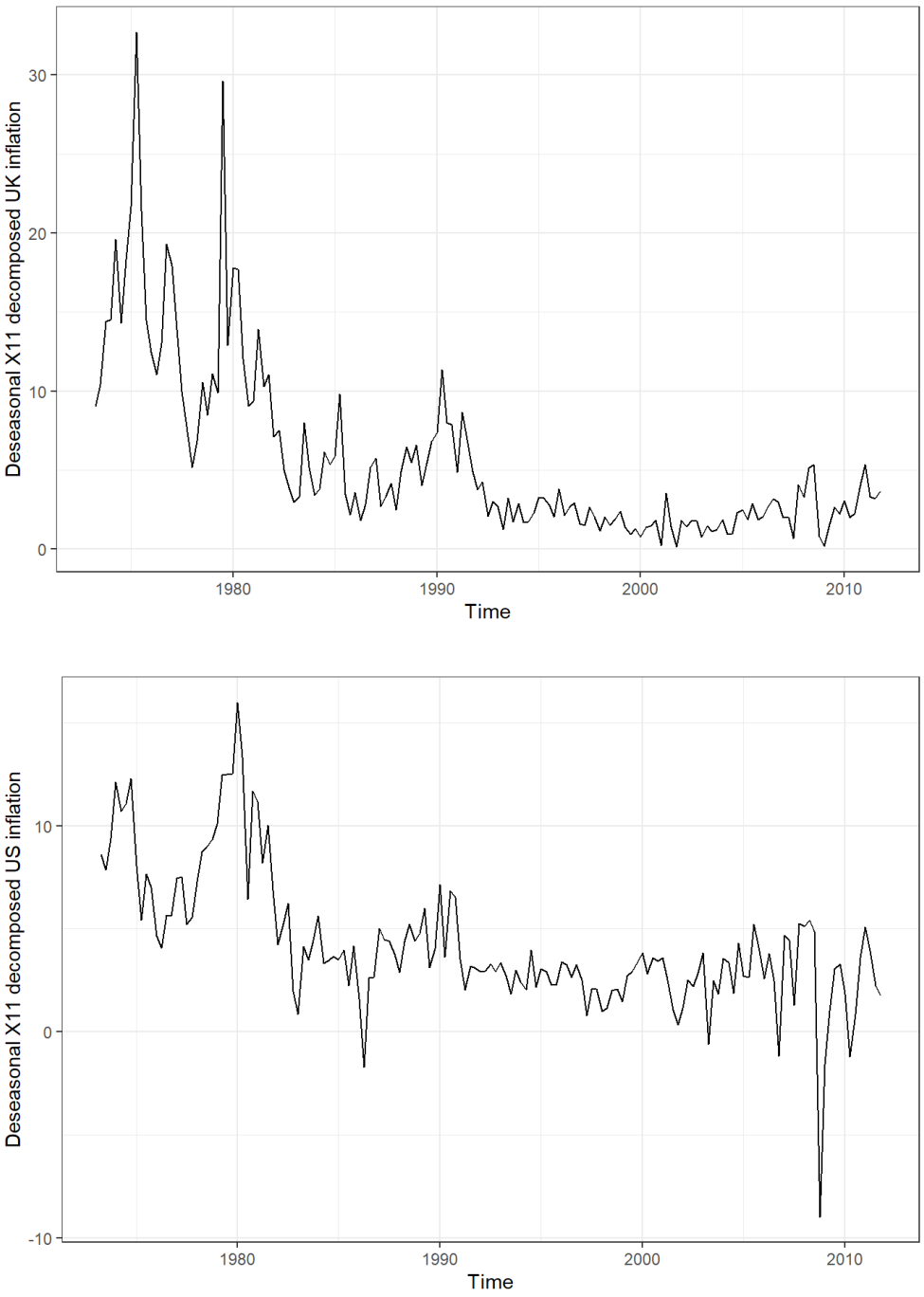


Figure 4.9: Plots of deseasonalized annualized quarterly UK and US inflation rates from 1973 Q1 to 2011 Q4 by X11 methods

p-values	ADF	KPSS
Deseasonalized UK inflation	0.3245	0.01
Deseasonalized US inflation	0.1869	0.01
Diff deseasonalized UK inflation	0.01	0.1
Diff deseasonalized US inflation	0.01	0.1

Table 4.4: P-values of the ADF and KPSS test for deseasonalized UK and US inflation decomposed by X11 method, the third and fourth row are p-values of differenced deseasonalized inflation rates. The null hypothesis of ADF tests is that the data have a unit root, which is non-stationary, while the null hypothesis of KPSS tests is that the data have no unit root, which is stationary.

$ARMA(p,q)(P,Q)$	UK	US
Method one	$ARMA(4,3)$	$ARMA(1,5)$
Method two	$ARMA(3,3)(0,2)$	$ARMA(4,1)(1,1)$
Method three	$ARMA(2,1)$	$ARMA(4,0)$

Table 4.5: The chosen ARMA model for UK and US inflation adjusted by the three methods

the two data sets are taken once difference at lag one to remove the trends. In order to ensure that the adjusted time series are stationary, the ADF and KPSS tests are applied and the results are demonstrated in Table 4.4. According to the Table 4.4, the deseasonalized inflation rates are not rejected by the ADF test and rejected by the KPSS test. Hence, they are hardly suitable for stationary models. On the contrary, the differenced deseasonalized inflation rates are relatively suitable for stationary models and their plots are shown in Figure 4.10, which does not contain obvious trends for both the UK and the US.

The three methods can transform the time series into stationary form, so as to satisfy the conditions to be applied in the S-vine models. The next step is to find the order of ARMA, which is used to estimated the kpacf of the time series.

The order in ARMA fitting

In order to fit the S-vine models to the adjusted time series, we need to set the order of the ARMA process, which is used to estimate the kpacf of the datasets. The order is set according to the automatic fitting process of ARMA. In the second method, the order of the seasonal part should be set on the basis of the automatic fitting procedure with seasonal ARMA. The order of selected ARMA models is presented in the Table 4.5. Furthermore, we input start values of the parameters in different orders of ARMA process for S-vine models. The parameters will be estimated by the maximum likelihood estimation method in the next step. This approach is taken for simplicity, since it takes too long to try S-vine processes with lots of different orders.

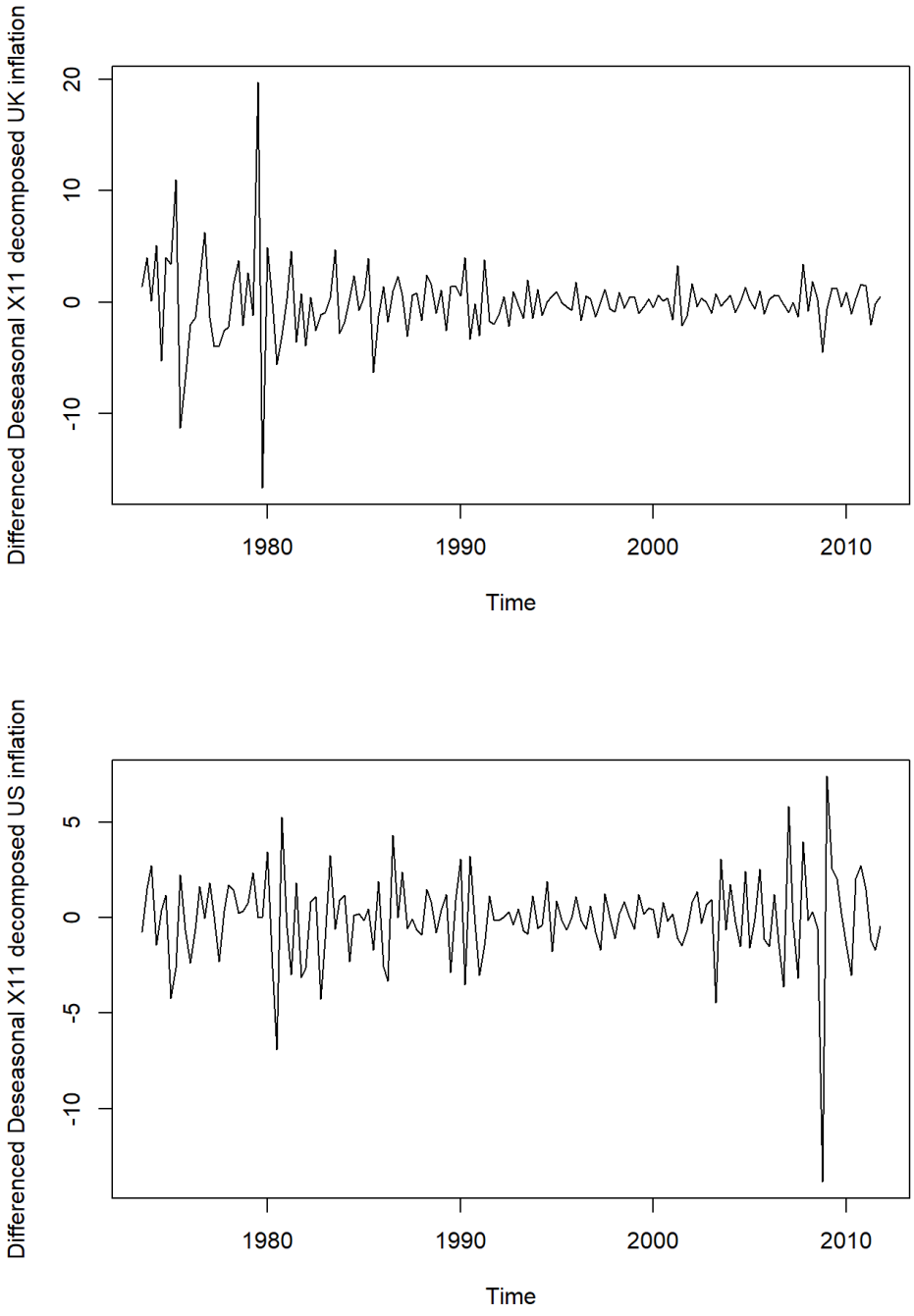


Figure 4.10: Time series plot of annualized quarterly UK (top panel) and US (bottom panel) inflation rates from 1973 Q1 to 2011 Q4 adjusted by Method 3. These data are the differenced time series of Figure 4.9.

Parametric margins used in S-vine models		
Country	UK	US
Method one	norm (889.1283)	norm (770.9587)
	sdwe (811.2870)	slap (758.4566)
	sst (828.9800)	sst (759.9314)
	NIG (826.0457)	hyp (760.1776)
Method two	norm (889.1283)	norm (762.0686)
	sdwe (812.1338)	slap (733.6209)
	sst (828.9800)	sst (735.4019)
	NIG (826.0457)	hyp (735.6752)
Method three	norm (804.3684)	norm (693.8083)
	sdwe (720.0326)	sdwe (655.9986)
	sst (722.8885)	sst (659.2513)
	NIG (723.6248)	NIG (659.6610)

Table 4.6: The parametric marginal distributions chosen to transform the stationary data from the UK and the US into uniform distribution and their AIC values (in brackets).

4.4 Choice of marginal distribution and copulas

After setting the order of ARIMA part, we start by finding out the best margins and sequences of bivariate copulas, separately. The margins can transform the adjusted time series to dataset with range between 0 and 1, which should be uniform distribution if the margin is good enough.

There are two types of margins, non-parametric and parametric. The scaling empirical distributions (expressed in Equation 3.8) are used as non-parametric margins in our study.

There are eight kinds of parametric margins in our study, including normal, skewed Laplace (slap), skewed double Weibull (sdwe), skewed student t (sst), skewed hyperbolic student t (shyt), hyperbolic (hyp), normal inverse Gaussian (NIG) and generalized hyperbolic distributions (Ghyp). We choose the top four margins used in the following research according to the Akaike information criterion (AIC) and Q-Q plots of the marginal distributions. The margins with the smallest AIC and linear qqplots are the one we choose to use in the following research.

There are three approaches to acquire stationary time series, which are distinctive from each other. Hence, we need to find the proper margins and bivariate copulas for them, separately. The Gaussian S-vine process with normal margin is the benchmark model. Thus, the normal distribution is also attempted in the following study. The chosen parametric marginal distributions are summarized in the Table 4.6.

In order to show the improvement of the margins, the Q-Q plots of the chosen margins for the data obtained from the first method are shown in Figure 4.11 as an

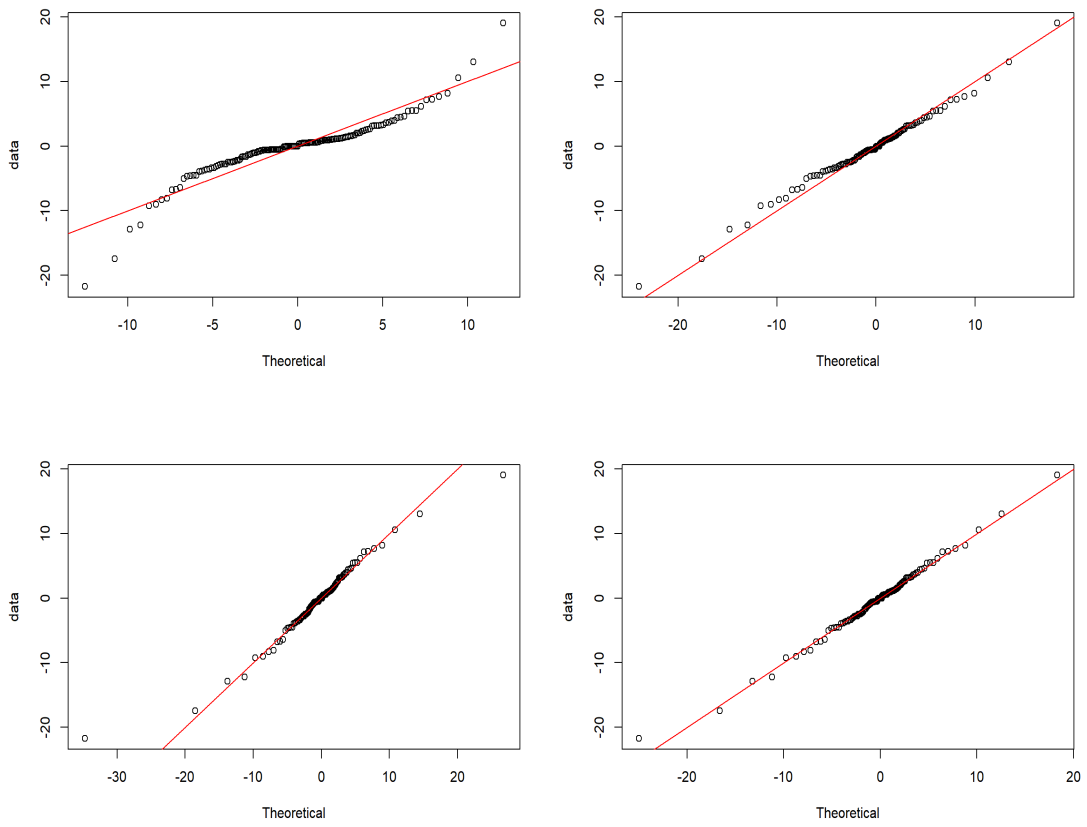


Figure 4.11: Q-Q plots of chosen margins of differenced UK inflation rate in the first method, the one at the top left is the normal distribution, the one at top right is the skewed double Weibull margin. The skewed student t distribution is at the left bottom. And the NIG margin is at the right bottom.

example. Compared to the Q-Q plots of the other three margins, the normal distribution has obvious disadvantages. Nevertheless, the Q-Q plots of the other three marginal distributions do not have significant differences. We decided to apply all the four margins in each case in the combination of margins and copulas in S-vine structures and compare them with the benchmark models.

After choosing the proper margins, the next step is to find the sequences of bivariate copulas, which can describe the data in the S-vine structures. The first step is to find the appropriate options of bivariate copulas and their combinations. In order to avoid the influences of marginal distribution, we use the non-parametric distribution, to transform the stationary time series into uniform distribution.

The options for the sequences of copulas in the S-vine are mentioned in Section 3.2.2. The radially symmetric and comprehensive copulas are Gaussian and Frank copulas. And the non-radially symmetric copulas are Gumbel, Joe and Clayton copulas. Non-radially symmetric copulas can be rotated 180 degrees; to denote the rotated copulas

we write, for example, $Clayton180$. Besides, the 90- and 270-degree rotation of copula are written as R and L, respectively, in order to treat the negative dependence. The treatment of negative dependence can also be denoted as G and F where the Gaussian and Frank copula replace the copula when the kpacf has negative values. For example, ($Clayton180\_F$) represents the 180-degree rotated Clayton with Frank copulas when kpacf has negative values and ($Gumbel\_L$) means Gumbel with left-rotated (270-degree rotation) Gumbel copulas when kpacf has negative values. It is worth pointing out that the order of the S-vine was set to 30, so as to save the calculation time. The orders equal to 30 are mostly sufficient for the convergence of acf or pacf. We only impose the order equal to 30 when the models are with moving-average terms.

Besides, we choose the survival Clayton copula, because when using the bivariate copula to fit the UK and US inflation rate data, the survival Clayton (also written as $Clayton180$) shows better results than Clayton, since there is evidence of the upper tail dependence in the inflation rates.

In order to model data with both upper and lower tail dependence, the two-parameter t copula and BB1 copula are applied in the s-vine models. The BB1 copula can be used to model asymmetric tail dependence (details can be found in Joe et al. [2010] and Nikoloulopoulos et al. [2012]). However, the extra parameters in the two types of copula will increase the difficulty of estimating the parameters of S-vine processes and cause the increase of AIC. Furthermore, Joe et al. [2010] pointed out that if the bivariate copulas at the first tree have upper or lower tail dependence, then all bivariate margins of a vine copula have the corresponding tail dependence. Hence, we substitute the t copula, BB1 copula or rotation of BB1 copula at the first tree and leave the copulas at the other trees as Gaussian copula. The S-vine copula is truncated at lag 30, so the two combinations for the two-parameter copula is written as t(1)_Ga(29), BB1(1)_Ga(29) or sBB1(1)_Ga(29), where sBB1 represents the survival BB1 copula.

4.5 Estimation results for S-vine models

4.5.1 Estimation for semi-parametric S-vine models

The semi-parametric S-vine models have non-parametric margins and parametric S-vine copulas. After we transform the data to uniform scale via scaling empirical distribution, we concentrate on modelling approximated uniformly distributed data by parametric S-vine copulas, referred to as semi-parametric estimations.

In order to decrease the volume of work, we decided to pick five types of sequences of copulas in the following study according to the AIC values, residual plots and the kpacf plots. Meanwhile, in order to test the normality of residuals, the Shapiro-Wilk tests are applied to in the residuals. The Gaussian copulas are used as the benchmark in each dataset, so as to make a comparison between different copulas and present

AIC	Method 1		Method 2		Method 3	
Country	UK	US	UK	US	UK	US
Gaussian	-71.5805	-62.0898	-79.9213	-77.8253	-3.5245	-31.7128
Joe	-2.3212	5.3124	-0.3212	6.6748	2.5733	-0.2405
Joe_G	-57.7207	-35.1237	-60.0648	-73.1653	-4.6685	-29.5735
Joe_F	-50.7055	-32.1583	-53.1106	-74.7198	-5.3209	-26.7659
Joe_L	-40.3367	-29.2928	-41.4346	-57.8568	-1.6985	-38.5731
Joe_R	-73.0530	-26.1545	-72.4664	-40.8520	-25.0859	-20.1309
Gumbel	-5.7180	3.8497	-0.2001	6.8250	3.9172	-2.7763
Gumbel_G	-84.8357	-57.1118	-87.2472	-68.3043	-5.5867	-29.6744
Gumbel_F	-77.5947	-57.4089	-74.6097	-70.6466	-5.5860	-27.2218
Gumbel_L	-76.0798	-52.8068	-78.5224	-74.1481	-11.1290	-42.3491
Gumbel_R	-75.6880	-53.1009	-92.1431	-67.8641	-23.5814	-26.9929
Frank	-57.8343	-60.5853	-70.9216	-70.3200	-4.612572	-29.8039
*Clayton*180	-0.3507	4.9779	1.6493	6.7681	4.655113	-1.6634
*Clayton*180*_G*	-57.7267	-39.4112	-61.0333	-75.4336	-3.524537	-29.9202
*Clayton*180*_F*	-41.2799	-35.8599	-54.0985	-72.2573	-4.612572	-27.4338
*Clayton*180*_L*	-42.4032	-30.3299	-42.8254	-61.4604	3.421432	-36.0535
*Clayton*180*_R*	-70.1552	-35.3284	-68.1050	-55.7069	-25.2234	-21.5908

Table 4.7: AIC values for the S-vine structures with seventeen types of combinations of copulas and the empirical distributions fitted to UK and US differenced inflation rate adjusted in the three methods. The values in lime colour are the AIC values of benchmark model. The ones in yellow colour are the AIC values of best model in each case.

the improvement from the Archimedean copulas. The AIC values of fitting the S-vine with the seventeen types of combinations to datasets transformed via non-parametric methods are summarized in Table 4.7. The datasets are acquired from the three methods used to remove the seasonality.

Table 4.7 presents that the combination of copulas can give superior results than the sequence of Gaussian copulas for UK data. Generally, mixed copula sequences perform better than the sequence of single copulas. However, for the US data, the sequences of Gaussian copulas demonstrate the smallest AIC values in the first and second method. The combinations of copulas show improvemens in the two respective methods compared to the single copula sequences. The advantage of the combination is embodied in the third method in the US data set.

According to Table 4.7, the combination of copulas has a significant improvement compared to the sequences of one type of copula, especially in the Joe, Gumbel or 180 degree rotated Clayton copulas, which can only describe the positive Kendall's tau. Table 4.8 demonstrates that the left-rotated Joe, Gumbel or 180 degree rotated Clayton are chosen when the kpacf has negative values in most data sets. This might be because the left-rotated Archimedean copulas can describe the negative Kendall's tau of inflation

Sequences of copulas used in S-vine models for the empirical distribution		
Country	UK	US
Method one	$Gaussian$	$Gaussian$
	Joe_R	Joe_G
	$Gumbel\_G$	$Gumbel\_G$
	$Frank$	$Frank$
	$Clayton-180\_R$	$Clayton-180\_G$
Method two	$Gaussian$	$Gaussian$
	Joe_R	Joe_F
	$Gumbel\_R$	$Gumbel\_L$
	$Frank$	$Frank$
	$Clayton180\_R$	$Clayton180\_G$
Method three	$Gaussian$	$Gaussian$
	Joe_R	Joe_L
	$Gumbel\_R$	$Gumbel\_L$
	$Frank$	$Frank$
	$Clayton180\_R$	$Clayton180\_L$

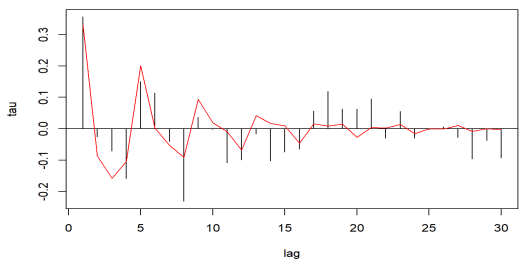
Table 4.8: The sequences of copulas chosen for the empirical distribution of data from the UK and the US.

rates better. There are no rotated combinations in the Gaussian and Frank copulas, because they are capable not only of expressing the positive correlations but also the negative ones.

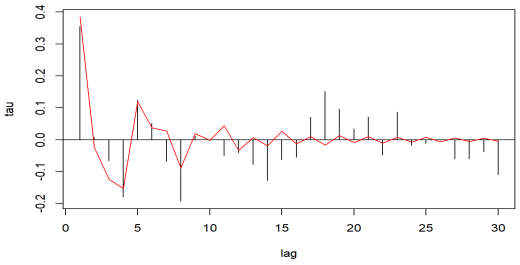
In order to present the comparison between different sequences of copulas and highlight the effects of rotation Archimedean copulas in the negative Kendall's tau expression, we show the kpacf plots of the selected sequences of copulas and the plots of Joe, Gumbel and 180 degree rotated Clayton copulas. There is quite a large amount of kpacf plots in different data sets. Thus, we select two data sets to present the comparison from different types of copula sequences. For the first method of data adjustment, the kpacf plots of UK data are shown in the Figure 4.12. Kendall's tau of UK inflation rates by the second method are presented in Figure 4.13. For the third approach, the US data are selected for kpacf plots in Figure 4.14.

The combinations of copulas improve the kpacf fitting significantly in all the three figures. The second method uses the seasonal ARMA in the estimated kpacf plots, so there are some seasonal features in the kpacf plots in the second method; for details see Figure 4.13. In Figure 4.14, most of the kpacf of US stationary inflation rates are negative, so the Joe, Gumbel and 180 degree rotated Clayton demonstrate poor fitting results compared to the combinations with Gaussian, Frank or their rotations. Nevertheless, the strengths of combinations compared to the Gaussian and Frank copulas are difficult to tell in the figures, so we take the AIC values as a reference as well.

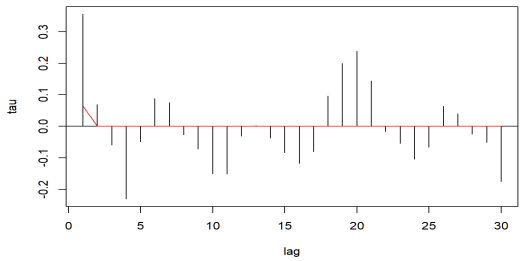
We attempt a Q-Q plot of residuals as well. If the model describes the dataset



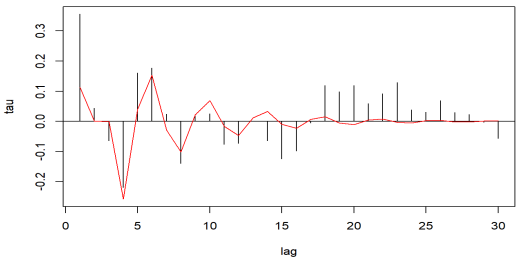
(a) Gaussian



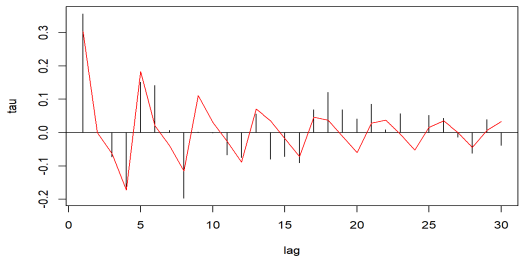
(b) Frank



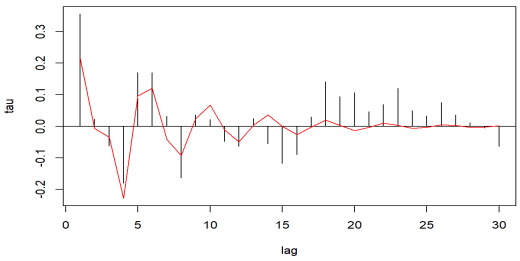
(c) Joe



(d) Joe with right-rotated Joe

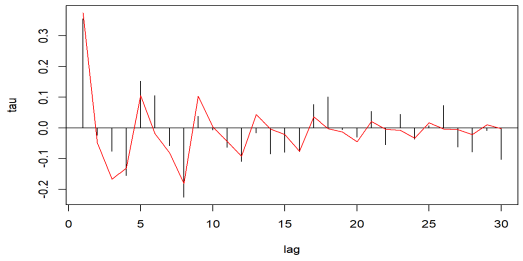


(e) Gumbel with Gaussian

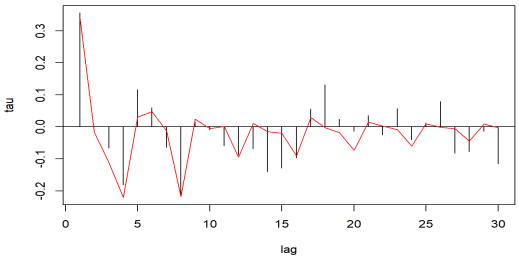


(f) Clayton180 with right-rotated Clayton180

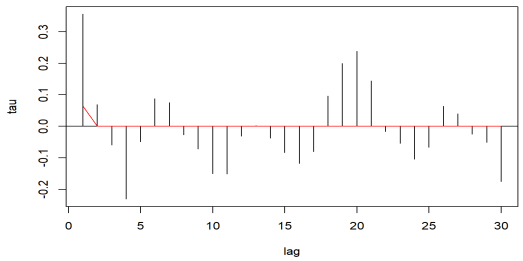
Figure 4.12: Kpacf plots of adjusted UK inflation rates in the first method. From 4.12a to 4.12f, they are Gaussian, Frank, Joe, Joe with right-rotated Joe, Gumbel with Gaussian, survival Clayton with its right-rotation copulas. The first black bar is kpacf calculated from real data and the previous value on the red line. The red line is the kpacf calculated from the S-vine models, details in Bladt and McNeil [2021].



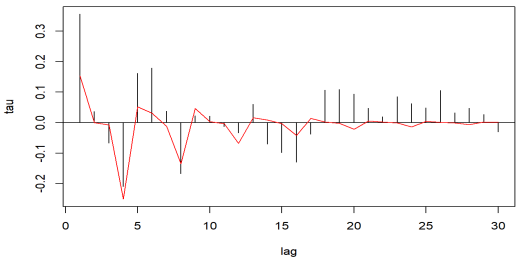
(a) Gaussian



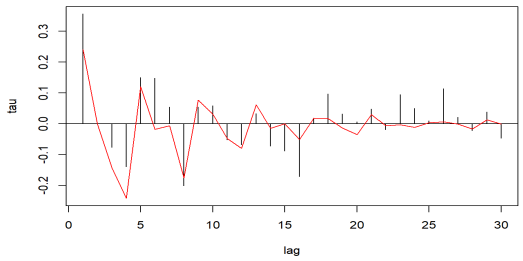
(b) Frank



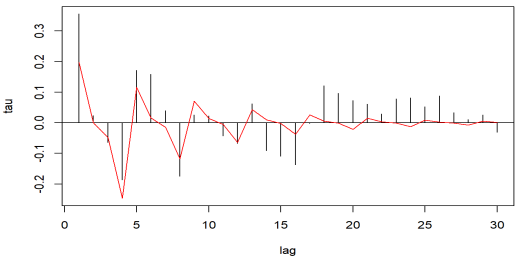
(c) Joe



(d) Joe with right-rotated Joe

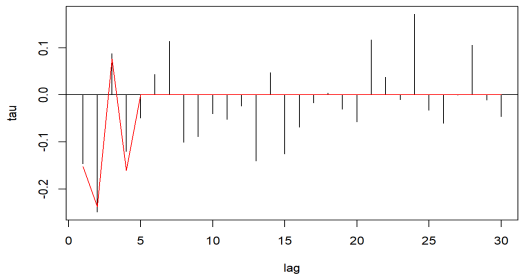


(e) Gumbel with right-rotated Gumbel

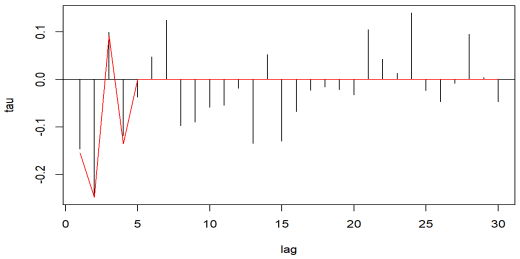


(f) Clayton180 with right-rotated Clayton180

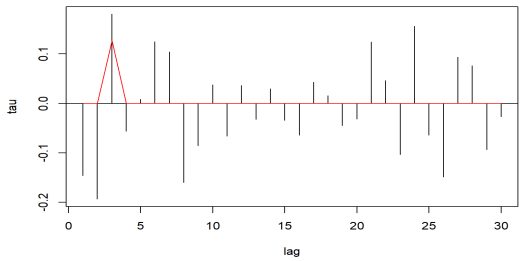
Figure 4.13: Kpacf plots of adjusted UK inflation rates in the second method. From 4.13a to 4.13f, they are Gaussian, Frank, Joe, Joe with right-rotated Joe, Gumbel with right-rotated Gumbel, survival Clayton with right-rotated survival Clayton copulas. The first black bar is kpacf calculated from real data and the previous value on the red line. The red line is the kpacf calculated from the S-vine models, details in Bladt and McNeil [2021].



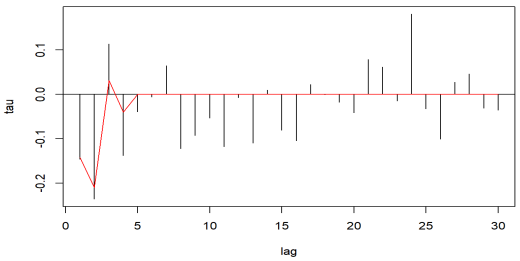
(a) Gaussian



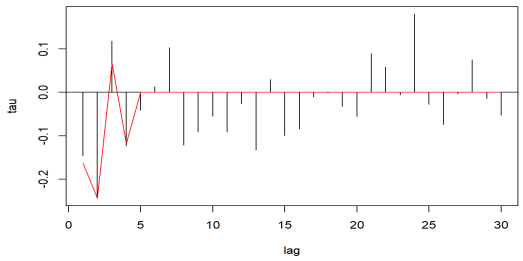
(b) Frank



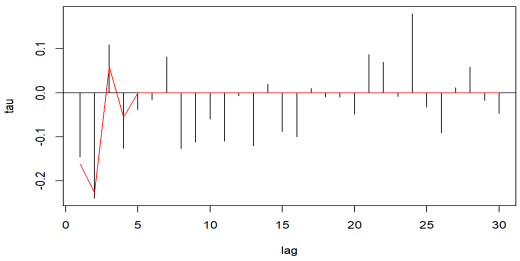
(c) Joe



(d) Joe with left-rotated Joe



(e) Gumbel with left-rotated Gumbel



(f) Clayton180 with left-rotated Clayton180

Figure 4.14: Kpacf plots of adjusted US inflation rates in the third method. From 4.14a to 4.14f, they are Gaussian, Frank, Joe, Joe with left-rotated Joe, Gumbel with left-rotated Gumbel, survival Clayton with its left-rotation copulas. The first black bar is kpacf calculated from real data and the previous value on the red line. The red line is the kpacf calculated from the S-vine models, details in Bladt and McNeil [2021].

AIC	Method 1		Method 2		Method 3	
Country	UK	US	UK	US	UK	US
$t(1)\_Ga(29)$	-89.5103	-80.6217	-92.6245	-83.1878	-23.6984	-42.8055
$sBB1/BB1(1)\_Ga(29)$	-83.8371	-69.6765	-87.3849	-46.4979	7.9494	-18.1245

Table 4.9: AIC values for S-vine models based on empirical margins with two-parameter copulas at first level of tree fitted to UK and US differenced inflation rate in the three methods. Note:only the US inflation rates adjusted by the second method use BB1 copula. The other dataset use the survival BB1 copula.

accurately, the residuals that are the differences between estimated values and observed values should be normally distributed (Pagan and Hall [1983]). Hence, the Q-Q plot of the residuals should be close to the straight red line if the model is reasonable. However, the advantages of S-vine copula models are not presented clearly in these plots, so we have decided not to show details of these plots in this section.

Two-parameter copulas in S-vine models

In order to capture both the upper and lower tail dependence in inflation rates, we apply S-vine models with two-parameter copulas, such as t copula or survival BB1 copula, at level one and Gaussian copula for the other levels. The survival BB1 copula performs better in inflation rates than the BB1 copula, so we decide to use survival BB1 copula in the following steps. The AIC values of the two combinations in S-vine copulas are presented in Table 4.9.

According to Table 4.9, the $t(1)\_Ga(29)$ copulas obtain the smaller AIC values than any other types of S-vine copulas in Table 4.7 in both UK and US datasets in the first and second method. The advantage of $t(1)\_Ga(29)$ model in UK inflation rates in the third method is not apparent. The $sBB1(1)\_Ga(29)$ model improves the estimation results in UK and US inflation rates in the first method, but it can not perform better than the best models in Table 4.7. The $sBB1(1)\_Ga(29)$ model is more suitable for the asymmetric tail dependence. Therefore, the reason for $t(1)\_Ga(29)$ copulas yielding smaller AIC values may be that there is both upper and lower tail dependence existing in the two datasets and the tail dependence of the inflation rates is relatively symmetric.

4.5.2 Estimation for parametric S-vine models

After selecting proper parametric margins for data and finding feasible copula sequences via a semi-parametric estimation approach, we fit the parametric S-vine models to the detrended and deseasonalized time series. The models combining parametric margins and S-vine copulas are called parametric S-vine models. As we discussed in Section 3.3.2, the estimation process of parametric S-vine models is based on the results from the estimation of semi-parametric models. The AIC values, marginal plots, kpacf plots

AIC of the first method of adjusted UK inflation rates				
	norm	sdwe	sst	NIG
Gaussian	811.4814	734.7820	750.2338	744.5911
Joe_R	821.9290	756.9242	746.9493	744.7967
Gumbel_G	841.3379	733.4545	735.9072	730.9872
Frank	785.2184	750.9871	759.5937	757.9247
*Clayton*180_*R*	830.9615	740.0448	753.6860	751.8919
AIC of the second method of adjusted UK inflation rates				
	norm	sdwe	sst	NIG
Gaussian	794.7407	726.0176	739.2647	740.0537
Joe_R	822.1728	727.4579	744.1916	735.3891
Gumbel_R	791.4435	702.2344	728.2249	717.3876
Frank	779.7509	734.3166	746.9644	746.2894
*Clayton*180_*R*	829.7111	736.1130	755.592	751.909
AIC of the third method of adjusted UK inflation rates				
	norm	sdwe	sst	NIG
Gaussian	780.4248	712.6835	717.1248	716.9361
Joe_R	706.5202	689.1821	681.1636	692.2709
Gumbel_R	727.5703	685.4403	681.6743	681.0495
Frank	785.5359	714.9143	717.2708	717.8501
*Clayton*180_*R*	725.7483	694.2959	687.1514	684.954

Table 4.10: AIC values for the S-vine models with selected copula sequences and marginal distributions fitted to stationary UK inflation rates adjusted by the three methods. The AIC in the green cell is the one from the benchmark model. The yellow one is the best full model. The best margins and copula sequences are marked in blue.

and Q-Q plots of residuals are produced. The Shapiro-Wilk tests are applied to the residuals of the full models to check their normality (Shapiro and Wilk [1965]).

We develop the fitting process in each combination of copulas and margins. There are two datasets and each dataset is adjusted by three approaches. Hence, in total, we divide the datasets into six groups. In order to make it brief, we choose the adjusted UK inflation from all the three methods and US inflation from the second method and represent their AIC results. The reason for choosing the three groups of data from the UK is that they contain distinctive types of combinations of copulas, and we can illustrate the advantages of non-Gaussian copulas and margins. Then, we can observe the influences from S-vine copulas in the full models. The AIC values of adjusted UK inflation in the three approaches are presented in Table 4.10. Table 4.11 shows the AIC values of adjusted US inflation in the second approach. The non-Gaussian copulas do not demonstrate improvement in the US data set compared to the other four copulas sequences, but the margins can improve the fitting process.

It is worth pointing out that the combinations of best margins and best copula sequences in the S-vine process are not always the best full models. The UK data set

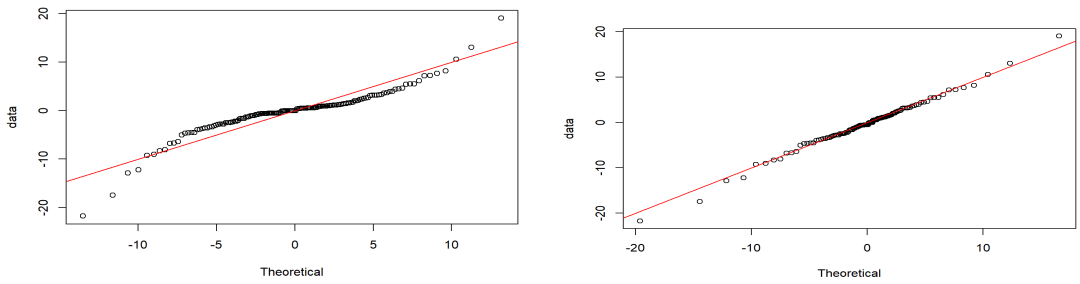
AIC of the second method of adjusted US inflation rates				
	norm	slap	sst	hyp
Gaussian	677.1613	653.3312	664.0922	668.8172
Joe_F	669.0866	657.5672	659.9301	659.5339
Gumbel_L	727.3565	674.0208	665.6345	675.9005
Frank	687.6807	661.5585	666.7535	662.9571
*Clayton*180_*G*	680.5335	653.4326	655.6574	665.1901

Table 4.11: AIC values for the S-vine models with selected copula sequences and marginal distributions fitted to stationary US inflation rates adjusted by the second method. The AIC in the green cell is the one from the benchmark model. The yellow ones are the best full models.The best margins and copula sequences are marked by blue colour.

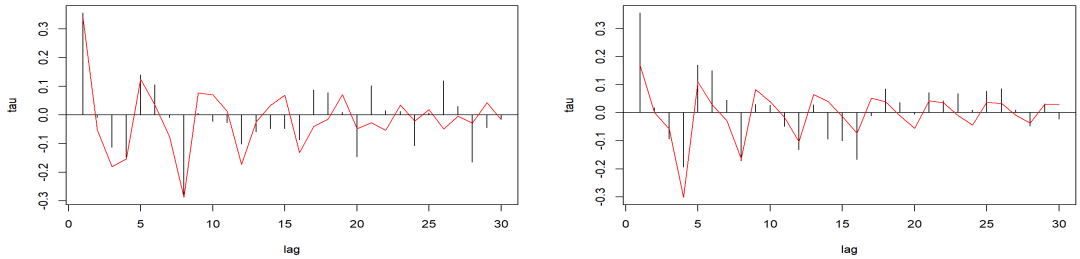
adjusted by the second method in Table 4.10 demonstrates that the S-vine consisting of Gumbel with its right rotation sequences and NIG margins is the best model, which is the combination with neither the best copula sequence nor the best margin. There is a compromise between imperfect margins and imperfect copulas. According to Bladt and McNeil [2022], it may be because one of the copulas has much stronger tail dependence than the others. Hence, the whole S-vine process tends to generate data that are not perfectly uniform in a small sample. A relatively "bad" margin transforming the data into not perfectly uniform may be a compromising choice in this case.

The S-vine with non-Gaussian copulas and marginal distributions shows evident advantages compared to Gaussian processes for both UK and US inflation rates. Joe et al. [2010] provide the theoretical justification and derivations for the method to capture the strong tail dependence at first tree in vine copula and why this model can surpass Gaussian models. In order to shed light on the performance of the non-linear and non-Gaussian models in the class of S-vine processes, we put the marginal distributions, kpacf and residual plots of the best model and benchmark model together. We use the results for UK inflation rates adjusted by the second and third approaches as examples. Figure 4.15 demonstrates the plots of the benchmark and Gumbel with right-rotated Gumbel copulas with skewed double Weibull margins for the UK data set in the second method. The Gumbel_R copula sequence with NIG margins for UK inflation used in the third approach is shown in Figure 4.16. The benchmark model is shown as a comparison.

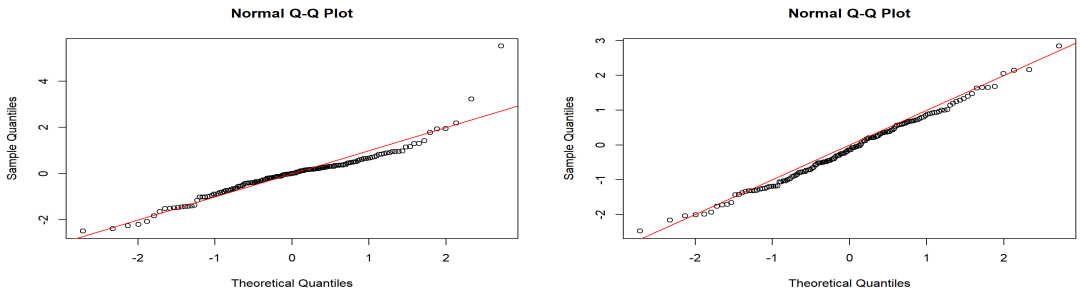
Figure 4.15a and 4.16a show the obvious improvement in the margins of the non-linear and non-Gaussian models compared to the benchmark models. The advantages of kpacf fittings are presented in Figure 4.15b and 4.16b. There is an obvious seasonal pattern in Figure 4.15b which embodies the features of seasonal ARIMA model. The residuals of the non-Gaussian and non-linear models are more close to the red line and are more linear, which means that their residuals are more in accordance with normal



(a) Plots of marginal distributions

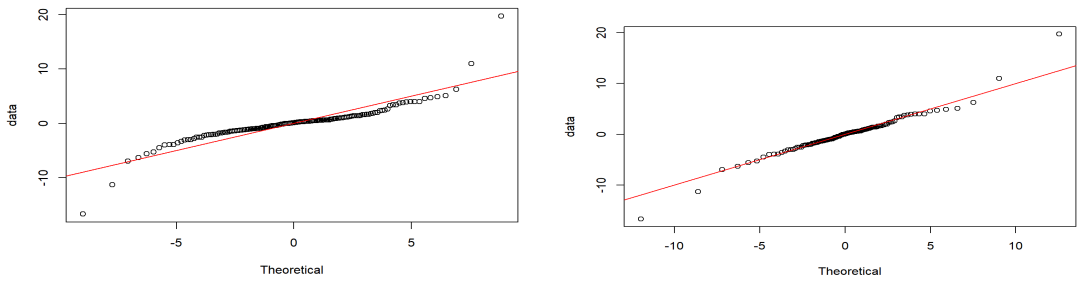


(b) Kpacf plots

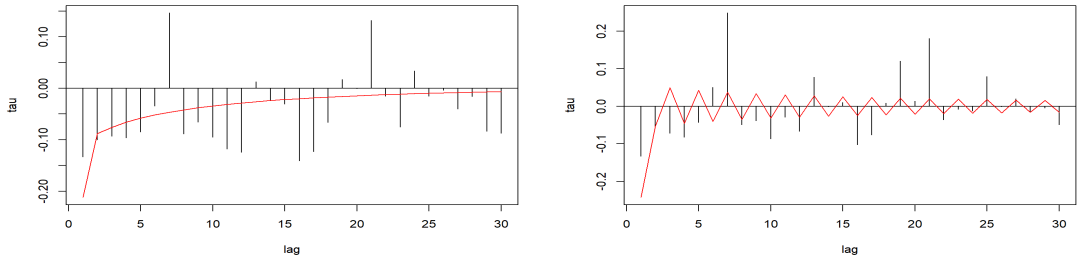


(c) QQ-plots of Residuals

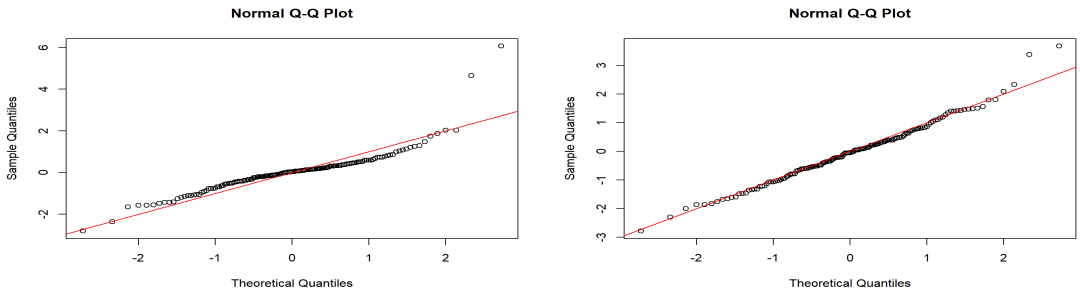
Figure 4.15: The three plots on the left are the marginal, kpacf and residuals plots of benchmark models of the UK inflation rates adjusted by the second approach. The ones on the right are the plots of Gumbel with its right-rotation copulas with skewed double Weibull margins models of the UK inflation rates adjusted by the second approach.



(a) Plots of marginal distributions



(b) Kpacf plots



(c) QQ-plots of Residuals

Figure 4.16: The three plots on the left are the marginal, kpacf and residuals plots of benchmark models of the UK inflation rates adjusted by the third approach. The ones on the right are the plots of Gumbel with right rotated Gumbel copulas with NIG margins models of the UK inflation rates adjusted by the third approach.

Method	One	Two	Three
	Benchmark model		
UK	1.21×10^{-9}	1.218×10^{-7}	1.495×10^{-11}
US	4.759×10^{-7}	9.266×10^{-6}	1.084×10^{-8}
	Best model for each dataset		
UK	0.8584	0.8245	0.1139
	$(Gumbel\_G + NIG)$	$(Gumbel\_R + sdwe)$	$(Gumbel\_R + NIG)$
US	0.04245	0.9804	0.5271
	$(Gaussian + sst)$	$(Gaussian + slap)$	$(Gumbel\_L + NIG)$

Table 4.12: P-values of the Shapiro-Wilk test for adjusted UK and US inflation rates. The null hypothesis of the Shapiro-Wilk test is that the data come from a normally distributed population.

AIC of the first method of adjusted UK inflation rates				
	norm	sdwe	sst	NIG
$t(1)\_Ga(29)$	816.6525	728.9716	735.4622	737.6350
$sBB1(1)\_Ga(29)$	829.8066	731.0628	740.8192	740.1331

Table 4.13: AIC values for the S-vine models with selected copula sequences and marginal distributions fitted to stationary UK inflation rates adjusted by the first method.

distribution; see Figure, 4.15c and 4.16c.

Moreover, the Shapiro-Wilk tests are applied to the residuals of the S-vine processes. The null hypothesis of the Shapiro-Wilk test is that the data are normally distributed. The results are shown in Table 4.12. All the null hypotheses of benchmark models are rejected by the tests, which means their residuals are hardly in accordance with normal distribution. Thus, the Gaussian with normal margin is not a suitable model for inflation rates. Conversely, the null hypotheses of non-Gaussian and non-linear models are not rejected by the normality tests, which indicates that there is not enough evidence against these models. The p-value of US inflation rates adjusted by the first method is slightly smaller than 0.05, which means the null hypothesis has a relatively higher possibility to be rejected, but the value is close to 0.05. Hence, the model is still acceptable.

Two-parameter copulas in S-vine processes

The AIC values of the S-vine with two-parameter copula sequences fitted to estimating UK and US inflation rates adjusted by the first method and the second method, separately, are presented in Table 4.13 and Table 4.14. The reason we add AIC results for US CPI adjusted by the second method is to make it correspond to the results in Table 4.11. Then, we can compare it with the S-vine with one-parameter copulas.

AIC of the second method of adjusted US inflation rates				
	norm	slap	sst	hyp
$t(1)\_Ga(29)$	678.5489	642.1164	645.2675	656.2833
$BB1(1)\_Ga(29)$	715.1267	683.7448	669.7348	697.434

Table 4.14: AIC values of the S-vine models with selected copula sequences and marginal distributions fitted to stationary US inflation rates adjusted by the second method.

Table 4.13 shows that the S-vine processes with skewed double Weibull margin and t copula at level one and Gaussian copula for the other level estimate the UK inflation rates better than the best S-vine processes in Table 4.10. Similarly, the S-vine processes with skewed Laplace or skewed t margins and t copula at level one and Gaussian copula at other levels in Table 4.14 outperform the S-vine processes with one-parameter copula sequences.

4.6 Forecasting inflation rates via S-vine models

The best S-vine models for fitting each data sets are found in the previous section. The goal of this section is to use the selected models to predict the future inflation rates in different methods. The methods and theories developed in the one-step prediction are discussed in Section 3.4.1. The one-step forecasting for real time series is produced in this section.

In order to compare different results, we have decided to calculate the mean, median, quantiles of 0.05, 0.10, 0.25, 0.75, 0.90, 0.95, and the probability integral transform (PIT) values. If the model is good enough, the PIT value in S-vine models should be iid uniform distribution. Hence, the Ljung-Box test and Kolmogorov-Smirnov tests are applied to test the independence and uniformity of PIT values, separately. The null hypothesis of the Ljung-Box test is that the tested data are independent. The null hypothesis of the Kolmogorov-Smirnov tests is that the tested data are uniform distribution. However, this does not mean that the model not rejected by all the tests is a good model. In contrast, if a model is rejected by one of these tests, it means the model is not a reasonable choice for the data. Furthermore, in order to test the accuracy of predictions distributed between quantiles following an arithmetic sequence, such as quantiles equal to $0.05, 0.10, ..., 0.95$, the chi-square test is used as well. In addition, the quantile score function is used to distinguish between the "good" model and "bad" models.

The next step is to use the model to forecast the next quarter inflation rate. Before the fitting and prediction processes, the inflation rates are differenced following the procedure in Section 4.3. There are three approaches to make the data stationary. In prediction, we decide to use the first and second methods in Section 4.3, because the

Quantile scores of UK dataset–Semi-parametric								
Quantiles	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
UK inflation adjusted by the first method								
Gumbel_G	0.777	1.323	2.213	2.590	2.142	1.116	0.615	2.051
Gaussian	0.738	1.249	2.045	2.715	2.199	1.340	0.793	2.017
$t(1)\_Ga(29)$	0.684	1.125	1.865	2.634	2.083	1.229	0.741	1.900
Gumbel_G_4	0.765	1.295	2.274	2.617	2.206	1.114	0.626	2.017
Gaussian_4	0.693	1.210	2.018	2.748	2.320	1.377	0.799	2.041
UK inflation adjusted by the second method								
Gumbel_R	0.662	1.140	2.052	2.481	2.129	1.251	0.745	1.911
Gaussian	0.687	1.148	1.886	2.395	2.173	1.275	0.758	1.869
Gumbel_R_4	0.653	1.110	2.068	2.670	2.219	1.256	0.747	1.983
Gaussian_4	0.713	1.185	1.912	2.693	2.306	1.241	0.719	1.984

Table 4.15: Quantile score values and average quantile scores of one-step prediction for UK inflation adjusted by the first and second approach in semi-parametric S-vine models. *Gumbel_G* is the Gumbel with Gaussian copula sequences without refitting. *Gaussian* is the Gaussian S-vine copulas without refitting. $t(1)\_Ga(29)$ is the S-vine copula with t copula at level one and Gaussian copulas for the other levels. *Gumbel_G_4* is the Gumbel with Gaussian copula sequences with refitting each 4 steps. *Gaussian_4* is the Gaussian S-vine copulas with refitting each 4 steps. *Gumbel_R* is the Gumbel with right-rotated Gumbel sequences in S-vine copula without refitting. *Gumbel_R_4* is the Gumbel with right-rotated Gumbel sequences in S-vine copula with refitting each 4 steps.

third method, the time decomposition method is a little more complex to forecast and is beyond the scope of this thesis.

The parametric S-vine models fit the whole models with adjusted time series and estimated the parameters both in S-vine copulas and margins. The semi-parametric first transforms the data to approximated uniform distribution by empirical distribution. Then, the S-vine models are fitted to the data on uniform scale to estimate the parameters in the S-vine. The semi-parametric method can analyze the S-vine copulas and margins, separately. We only include the model with two-parameter copula at lag one if it gives an improvement in each dataset.

4.6.1 Semi-parametric S-vine process

We choose the best copula sequences according to the AIC values in the fitting process. In the UK inflation dataset adjusted by the first approach that deals with seasonality, the model is the *Gumbel_G* copula and the Gaussian copulas are used to compare with the best model. For seasonal ARMA kpacf in the second approach, the *Gumbel_R* copula sequence is selected to predict the next data. The VaR score values and the average VaR score (Equation 3.11 and Equation 3.13) of the one step prediction are demonstrated in Table 4.15. Meanwhile, the prediction results from refitting each four steps are presented in the Table 4.15.

Quantile scores of US dataset–Semi-parametric								
Quantiles	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
US inflation adjusted by the first method								
$Frank$	0.827	1.482	2.777	3.588	2.764	1.328	0.675	2.583
$Gaussian$	0.823	1.483	2.710	3.580	2.864	1.514	0.899	2.601
$Frank\_4$	0.828	1.459	2.795	3.745	3.047	1.661	0.890	2.730
$Gaussian\_4$	0.791	1.394	2.778	3.932	3.315	2.014	1.237	2.868
US inflation adjusted by the second method								
$Clayton180\_G$	0.749	1.347	2.351	3.317	2.701	1.441	0.846	2.396
$Gaussian$	0.687	1.162	2.223	3.030	2.438	1.318	0.777	2.346
$t(1)\_Ga(29)$	0.720	1.243	2.188	2.912	2.315	1.305	0.756	2.130
$Clayton180\_G\_4$	0.831	1.405	2.368	3.217	2.486	1.296	0.730	2.318
$Gaussian\_4$	0.769	1.216	2.136	2.843	2.251	1.152	0.663	2.076

Table 4.16: Quantile score values and average quantile scores of one-step prediction for US inflation adjusted by the first and second approach in semi-parametric S-vine models. $Frank$ is the Frank copula sequences without refitting. $Gaussian$ is the Gaussian S-vine copulas without refitting. $t(1)\_Ga(29)$ is the S-vine copula with t copula at level one and Gaussian copulas for the other levels. $Frank\_4$ is the Frank copula sequences with refitting each 4 steps. $Gaussian\_4$ is the Gaussian S-vine copulas with refitting each 4 steps. $Clayton180\_G$ is the survival Clayton with Gaussian sequences in S-vine copula without refitting. $Clayton180\_G\_4$ is the survival Clayton with Gaussian sequences in S-vine copula with refitting each 4 steps.

According to the Table 4.15, the $Gumbel\_G$ copula in the first method removing seasonality shows smaller average quantile scores than the Gaussian S-vine process, which means the non-Gaussian S-vine copula can improve the one-step prediction. However, the $Gumbel\_R$ copula in the second method has smaller AIC in the fitting process, but the average quantile score is greater than the Gaussian copulas in prediction. This means that the model fitted best may not predict accurately. In contrast, a model not fitting precisely is possible to forecast well. It has to be remembered that the number of out of sample forecasts is small, so it is possible that this can happen, as we have discussed in the simulation study. The quantile score of US inflation (Table 4.16) is an example for this conclusion. Moreover, the quantile scores of refitting each four steps are greater than the ones without refitting. Hence, the refitting process does not improve the one-step prediction.

Furthermore, the S-vine with t copula at level one improves the prediction results according to the mean of quantile scores in the UK inflation rates adjusted by the first method. However, refitting for the S-vine copulas with two-parameter copula will create difficulties in calculation and the fitting process will be slower, so we do not do refit for the $t(1)\_Ga(29)$ model.

In the US inflation dataset, the Gaussian S-vine is the best model in the one-parameter copulas, but in the first method dealing with seasonality of data, the average quantile scores of the Gaussian are greater than the Frank copula sequences, which

means it is possible that a good model does not predict well or this result may be caused by the small sample size as we discussed before. Furthermore, the quantile scores at different quantiles manifest distinctive comparison results. This may be because of the asymmetry of non-Gaussian copulas.

Similar to UK dataset, the S-vine with t copula at level one improves the prediction results according to the mean of quantile scores in the US inflation rates adjusted by the second method. However, refitting for the S-vine copulas with two-parameter copula will increase the difficulty of calculation and make the process slower, so we do not do refit for the $t(1)\_Ga(29)$ model in this dataset as well.

4.6.2 Parametric S-vine process

The advantages of non-Gaussian copulas and margins S-vine models are pronounced in parametric S-vine processes as well. Table 4.17 has the quantile scores of a "good" model and the benchmark model in UK inflation rates adjusted by the first and second method.

Most of the quantile scores of the benchmark model are greater than non-Gaussian S-vine models, according to Table 4.17. However, the scores of refitting each four steps are smaller than the ones without refitting in the parametric models. This may be because the refitting could improve the estimation of margins. Moreover, the differences between the non-Gaussian S-vine and benchmark models are much greater than the semi-parametric S-vine models in both the first and second method. The t copula does not show advantage in parametric case for UK inflation in the first method. The seasonal ARMA presents better results in the parametric S-vine models. In order to present quantile values in each prediction, the predicted median and quantile values at quantiles $0.10, 0.25, 0.5, 0.75, 0.90$ of the $Gumbel\_G$ copulas with *sdwe* margin in the S-vine model for UK inflation adjusted by the first approach are demonstrated and the one of the benchmark model for UK inflation adjusted by the first approach is presented in Figure 4.17.

The black lines in Figure 4.17 present the observations, which should have exceedance points beyond the lines describing the predicted quantiles α for 40α times when $\alpha < 0.5$ and $40(1-\alpha)$ times when $\alpha > 0.5$. The exceedances of the benchmark model are barely beyond the quantile lines. In contrast, the line of observations crosses the quantile lines many times, which means the exceedances of the predicted quantiles of the S-vine model with the class of $Gumbel\_G$ copula and *sdwe* margin are closer to our expectation. Therefore, the quantile prediction from the chosen model is more reasonable than the benchmark model in this dataset.

Quantile scores of UK dataset–Parametric								
Quantiles	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
UK inflation adjusted by the first method								
Gumbel_G_sdwe	11.973	15.904	19.492	20.380	16.619	8.983	5.879	16.765
Benchmark	10.843	16.985	23.608	21.473	24.293	16.981	10.443	18.797
t(1)_*Ga*(29)_*sdwe*	10.078	14.525	21.403	23.014	19.675	12.693	7.581	18.797
Gumbel_G_sdwe_4	9.950	14.013	17.443	18.911	19.091	12.862	9.008	16.632
Benchmark_4	8.831	14.563	21.943	21.619	21.818	16.607	11.452	17.084
UK inflation adjusted by the second method								
Gumbel_R_sdwe	7.228	11.032	17.585	22.806	21.060	13.861	8.909	18.060
Benchmark	11.439	18.218	26.152	22.081	20.645	15.338	9.658	20.865
Gumbel_R_sdwe_4	6.959	10.990	16.807	22.103	21.369	14.879	10.152	17.967
Benchmark_4	8.789	14.323	22.064	22.218	20.224	14.643	9.819	18.589

Table 4.17: Quantile score values and their means of one-step prediction for UK inflation adjusted by the first and second approach in parametric S-vine models. *Gumbel_G_sdwe* is the Gumbel with Gaussian copula sequences with skewed double Weibull margins without refitting. *Benchmark* is the Gaussian S-vine copulas and margins in S-vine without refitting. *t*(1)_*Ga*(29)_*sdwe* is the S-vine with skewed double Weibull margin and t copula at level one. *Gumbel_G_sdwe_4* is the Gumbel with Gaussian copula sequences with skewed double Weibull margins refitting each 4 step. *Benchmark_4* is the Gaussian S-vine copulas and margins in S-vine refitting each 4 step. *Gumbel_R_sdwe* is the Gumbel with its right-rotation copula sequences with skewed double Weibull margins without refitting. *Gumbel_R_sdwe_4* is the Gumbel with its right-rotation copula sequences with skewed double Weibull margins refitting each 4 step.

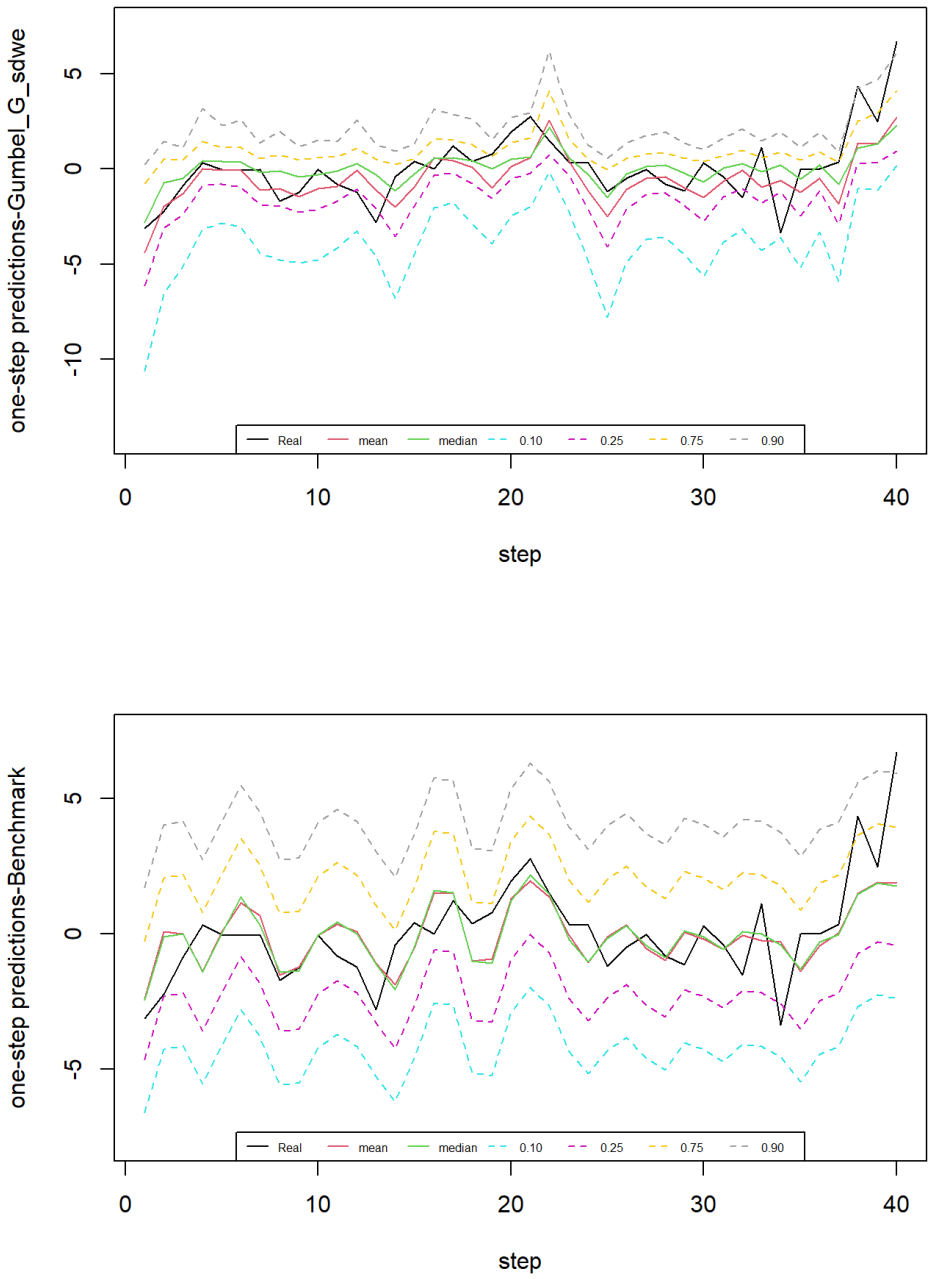


Figure 4.17: One step predicted median and quantile at quantiles 0.1,0.25,0.5,0.75,0.9 of S-vine with Gumbel_G copula sequences and skewed double weibull margin model (top panel) and benchmark model (bottom panel) in UK inflation adjusted by the first approach.

4.7 Discussion

S-vines with non-Gaussian copulas and non-linear structures can improve both the fitting and prediction processes compared with ARMA processes, which are equal to the Gaussian copula sequences and normal margins in the S-vine with infinite orders. This section finds the quantile score and average quantile scores constitute an effective method to compare "good" and "bad" predictions if the sample size is large enough. Moreover, we assume that the weights of quantile scores at different quantiles are the same. Different weights could be attempted at different quantiles.

In addition, it is necessary to find more sensitive statistical tests for the small sample. In fitting processes, the AICcs can be a proper approach to select the best model for a small sample size.

Furthermore, the one-step prediction for the data adjusted by the third method could be realized in further research. Other types of asymmetric copulas with two parameters, such as the BB7 copula, could be considered to model features of data more precisely in the future. The reason why we did not try BB7 in this thesis is that the Kendall's tau formula is more complicated in BB7, which makes it less suitable than BB1 in our methodology.

Chapter 5

Vt-S-vine Models for Asset Return Data

In previous chapters, we discussed the time series that can be modelled by stationary vine copulas processes. However, many financial data exhibit volatility, so models that can capture volatile patterns are required. Loaiza-Maya et al. [2018] pointed out that the heteroskedastic data, which are the data with volatility, are usually with a cross shape of copula densities. Few bivariate copulas can meet these requirements in practice. Only the t copula with a correlation parameters and degrees of freedom both close to zero can do so (Demarta and McNeil [2005]). However, the t copula with the degrees of freedom approaching to zero is numerically unstable and difficult to conduct in practice. Therefore, Loaiza-Maya et al. [2018] developed the method to use mixtures of copulas and their rotations to model financial data. Moreover, they proposed a symmetric transformation to combine with mix copulas, so as to facilitate measuring persistence in the volatility. Similarly, in order to model volatile financial time series more accurately, McNeil [2021] proposes combining S-vine processes with a class of uniformity-preserving transformations called v-transform. The theory of v-transforms in this chapter is based on McNeil [2021].

Let $x_1, ..., x_n$ be a time series of assets return that can be modelled by a strictly stationary stochastic process $(X_t)_{t\in\mathbb{N}}$, which has a marginal distribution F_X. Generally, the returns on financial assets X_t have limited serial correlation. However, the squared (X_t^2) or absolute $(|X_t|)$ processes have prominent and persistent positive serial correlation, which is caused by volatility clustering; Cuthbertson [2004] and Cont [2001] discussed this stylized fact of financial return data. The volatility clustering is defined in Cont [2001] that the large changes in price tend to be followed by large changes and small changes tend to be followed by small changes, which also exhibits the fact that high-volatility events tend to cluster in time.

Transformed series like $|X_t|$ are a typical example of symmetric volatility proxy

series; see McNeil [2021] for details. In these series, the volatility is exhibited via serial correlation. We denote the volatility proxy series as $(T(X_t))$, where $T : \mathbb{R} \to \mathbb{R}$ is a transformation for this absolute process that satisfies the conditions: (i) depends on a change point μ_T that may be zero, (ii) the transformation is increasing in $X_t - \mu_T$ for $X_t \geq \mu_T$ and (iii) increasing in $\mu_T - X_t$ for $X_t \leq \mu_T$.

The $(T(X_t))$ for $|X_t|$ process is symmetric and linear on both sides of the change point μ_T. Nevertheless, one can also construct volatility proxies that can be asymmetric and non-linear. The expression of $T(X_t)$ is diversified and flexible. Moreover, the volatility proxy series has its own distribution as well, where $F_{T(X)}$ denotes the cdf of $T(X_t)$. In this chapter, we concentrate on modelling the relationship between PIT series of a volatility proxy series (V_t) and PIT series of stationary stochastic process (U_t). We follow the definition developed by McNeil [2021] and use the same notation to define $V_t = F_{T(X)}(T(X_t))$ and $U_t = F_X(X_t)$ for all t. Then, a v-transform is a function to describe their relationship. The theory of v-transforms will be developed later.

Given (V_t) and letting v-transform be written as $\mathcal{V}$, we can have the following chain of transformations:

$$X_t \xrightarrow{F_X} U_t \xrightarrow{\mathcal{V}} V_t \xrightarrow{F_{T(X)}(v)^{-1}} T(X_t), \tag{5.1}$$

where $F_{T(X)}$ is the marginal distribution of the volatility proxy series $T(X_t)$. Under the chain of transformations in 5.1, V_t can be modelled by a stochastic process of correlated uniform variables, such as S-vine processes. Hence, by selecting different types of marginal distribution or v-transforms, we can produce distinctive processes via the transforming chain. The relationship between v-transforms and volatility proxy is embodied in the this chain.

5.1 V-transforms for stochastic volatility

In practice, there is stochastic volatility existing in financial and economic data. In order to capture the volatility, the volatility proxy transformation is used. V-transforms describe the relationship between quantiles of the stationary distribution of time series and quantiles of the distribution of a predictable volatile proxy variable. The linear v-transform is the simplest v-transformation. Hence, we start from the linear v-transform.

Example 2 (Linear v-transform)**.** *We can construct a symmetric volatility proxy transformation $T(X_t) = |X_t|$. Assume the marginal distribution of X_t is F_X and the marginal distribution of $|X_t|$ is $F_{|X|}$, where F_X and $F_{|X|}$ are absolutely continuous and the density f_X satisfies $f_X(x) = f_X(-x)$ for all $x > 0$. The probability integral transformation (PIT) can be written as $U_t = F_X(X_t)$ and $V_t = F_{|X|}(|X_t|)$.*

Since the density f_X is symmetric, the V_t can be expressed as

$$V_t = F_{|X|}(|X_t|) = \begin{cases} F_{|X|}(-X_t) & = 1 - 2F_X(X_t) = 1 - 2U_t, if X_t < 0 \\ F_{|X|}(X_t) & = 2F_X(X_t) - 1 = 2U_t - 1, if X_t \geq 0 \end{cases} \tag{5.2}$$

where we could obtain the relationship between V_t and U_t that

$$V_t = \mathcal{V}(U_t) = |2U_t - 1|. \tag{5.3}$$

The v-transform $\mathcal{V}(u) = |2u - 1|$ is symmetric and linear v-shaped function.

The linear v-transform is a special case of v-transforms. There are other types of v-transforms that are asymmetric and non-linear.

5.1.1 Asymmetric and non-linear v-transforms

In order to generalize non-linear v-transforms, we introduce the definition of volatility proxy and profile proposed in McNeil [2021] first.

Definition 5.1.1 (Volatility proxy transformation and profile.)**.** *Let T_1 and T_2 be strictly increasing, continuous and differentiable functions on $\mathbb{R}^+ = [0, \infty)$ where $T_1(0) = T_2(0)$. Let $\mu_T \in \mathbb{R}$, then any transformation $T : \mathbb{R} \to \mathbb{R}$ of the form*

$$T(x) = \begin{cases} T_1(\mu_T - x) & x \leq \mu_T \\ T_2(x - \mu_T) & x > \mu_T \end{cases} \tag{5.4}$$

is a volatility proxy transformation. The parameter μ_T is the change point of T and the associated function $g_T : \mathbb{R}^+ \to \mathbb{R}^+$, $g_T(x) = T_2^{-1}(T_1(x))$ is the profile function of T.

From Equation 5.4, there is a change point μ_t in which the different functions T_1 and T_2 for returns may contribute distinctive influence to the volatility proxy depending on the sign of the return. This is a stylized fact of asset returns, called leverage effect, which is suggested in economic theory that market information should have an asymmetric effect on volatility, which means that the bad news leading to a fall in the equity value of a company tends to increase the volatility. The phenomenon is similar to the leverage effects in GARCH type models referred to by Ding et al. [1993] and Cont [2001].

In the case that a volatility proxy transformation is symmetric about μ_T, where $T_1 = T_2$, the profile function $g_T(x) = x$. Besides, the profile function of a volatility proxy transformation is a strictly increasing, continuous and differentiable function in $\mathbb{R}^+$. Generally, if the change point μ_T and profile function g_T are known, the information of constructing a v-transform is fully obtained. Moreover, by applying different continuous marginal distribution F_X and volatility proxy transformation T, we can construct different v-transforms $\mathcal{V}(U)$.

Proposition 5. *Suppose F_X is an absolutely continuous and strictly increasing marginal distribution of a random variable X on $\mathbb{R}$ and a volatility proxy transformation expressed by T. Let $U = F_X(x)$ and $V = F_{T(X)}(T(X))$ be the PIT values of X and $T(X)$. Then, the v-transform $V = \mathcal{V}(U)$ can be written as*

$$\mathcal{V}(u) = \begin{cases} F_X(\mu_t + g_T(\mu_T - F_X^{-1}(u))) - u, & u \leq F_X(\mu_T) \\ u - F_X(\mu_T - g_T^{-1}(F_X^{-1}(u) - \mu_T)), & u > F_X(\mu_T). \end{cases} \tag{5.5}$$

The Proposition 5 is given by McNeil [2021]. In Example 2, the volatility proxy transformation $T_X(X_t) = |X_t|$ with change point 0.5 and profile function $g_T(x) = x$ produces the symmetric v-transform in the form of Equation 5.3. Nevertheless, if we change the volatility proxy transformation to $T_X(X_t) = X_t^2$, the change point and profile function will be exactly the same as the ones in Example 2. Therefore, we obtain the same v-transform as Equation 5.3. The result reveals that any two volatility transformations T_X and $\tilde{T}_X$ sharing the same change point μ_T and profile function g_T lead to the same resulting v-transform. The formal definition of the v-transform and the fulcrum proposed by McNeil [2021] is in Definition 5.1.2.

Definition 5.1.2 (V-transform and fulcrum)**.** *Any transformation $\mathcal{V}$ that can be obtained from Equation 5.5 by choosing an absolutely continuous and strictly increasing marginal distribution F_X on $\mathbb{R}$ and a volatility proxy transformation T is a v-transform. The value $\delta = F_X(\mu_T)$ is the fulcrum of the v-transform.*

According to Definition 5.1.2, the fulcrum δ is an important parameter for v-transforms. If we set the fulcrum δ to zero, then $\mathcal{V}(u) = u$ for $u \in [0, 1]$. For a linear v-transform, the fulcrum is the only parameter, so the linear v-transform can be called one parameter v-transform as well. In practice, the v-transform can be flexible and constructed by applying an asymmetric and parametric F_X, for example, a distribution of the type introduced by Fernández and Steel [1998]. We follow the functions and assumptions suggested by McNeil [2021] to set $\mu_T = 0$ and $g_T(x) = kx^\xi$ for $k > 0$ and $\xi > 0$. Clearly, the profile function $g_T(x) = x$ of a symmetric volatility proxy transformation is a special case of this profile, where $k = \xi = 1$.

Let f_0 be a symmetric density and set $\gamma > 0$ to be a scalar parameter. The model suggested by Fernández and Steel [1998] is

$$f_X(x; \gamma) = \begin{cases} \frac{2\gamma}{1+\gamma^2} f_0(\gamma x) & x \leq 0 \\ \frac{2\gamma}{1+\gamma^2} f_0(\frac{x}{\gamma}) & x > 0. \end{cases} \tag{5.6}$$

We can develop the parametric v-transform by using Equation 5.6 and setting the f_0 to be a special case of a Laplace or double exponential distribution $f_0(x) = 0.5 \exp(-|x|)$, in order to produce a tractable expression; see McNeil [2021] for details.

Proposition 6 (Three-parameter v-transform)**.** *Let $F_X(x;\gamma)$ be the cumulative distribution function (cdf) of the density in Equation 5.6 when $f_0(x) = 0.5\exp(-|x|)$. Set $\mu_T = 0$ and let $g_T(x) = kx^\xi$ for $k, \xi > 0$. The v-transform in Equation 5.5 is given by*

$$\mathcal{V}_{\delta,\kappa,\xi}(u) = \begin{cases} 1 - u - (1-\delta)\exp(-\kappa(-\ln(\frac{u}{\delta}))^\xi) & u \le \delta, \\ u - \delta\exp(-\kappa^{-1/\xi}(-\ln(\frac{1-u}{1-\delta}))^{1/\xi}) & u > \delta, \end{cases} \tag{5.7}$$

where $\delta = F_X(0) = (1+\gamma^2)^{-1} \in (0,1)$ and $\kappa = k/\gamma^{\xi+1} > 0$.

The expression in Equation 5.7 is a three-parameter v-transform, which is a uniformity-preserving transformation. In the case $\xi = 1$, we obtain the two-parameter v-transform. Another special case is linear v-transform when $\xi = 1$ and $\kappa = 1$, which can be written as

$$\mathcal{V}_\delta(u) = \begin{cases} (\delta - u)/\delta & u \le \delta, \\ (u-\delta)/(1-\delta) & u > \delta. \end{cases} \tag{5.8}$$

Equation 5.8 contains the special case of Example 2 when $\delta = 0.5$. The linear v-transform is the simplest case of v-transform. In the empirical study, the linear v-transform are widely used to be combined with other models, in order to simplify the entire models and reduce the amount of parameters.

In order to present the features of v-transforms intuitively, we demonstrate one example for each of the three types of v-transform, linear, two-parameter and three-parameter in Figure 5.1. The fulcrum δ is set to 0.4 in the first two transforms. In the v-transform with two parameters, $\kappa = 0.6$. In the case of three parameters, $delta = 0.35$, $\kappa = 1.2$ and $\xi = 0.8$.

According to Figure 5.1, the v-shape is skewed to the right when $\kappa > 1$ and skewed to the left when $\kappa < 1$, which can be proved by taking the second order derivatives of Equation 5.7. When $\kappa > 1$, the second order derivative is negative on the left side and positive on the right side. Hence, the function decrease faster on the left and increase faster on the right, which leads to the v-shape being skewed to the right and vice versa.

In applications, the v-transform is usually constructed by Equation 5.7 and the cdf F_X is set at outset. The profile function is generated via Equation 5.9, proposed in Proposition 3 by McNeil [2021].

$$g_T(x) = F_X^{-1}(F_X(\mu_T - x) + \mathcal{V}(F_X(\mu_T - x))) - \mu_T, \quad x \ge 0, \tag{5.9}$$

where the change point $\mu_T = F_X^{-1}(\delta)$.

5.1.2 Properties of v-transforms

The v-transforms plots in Figure 5.1 exhibit the main properties of v-transforms, which are summarized in Definition 5.1.3; the details are discussed by McNeil [2021].

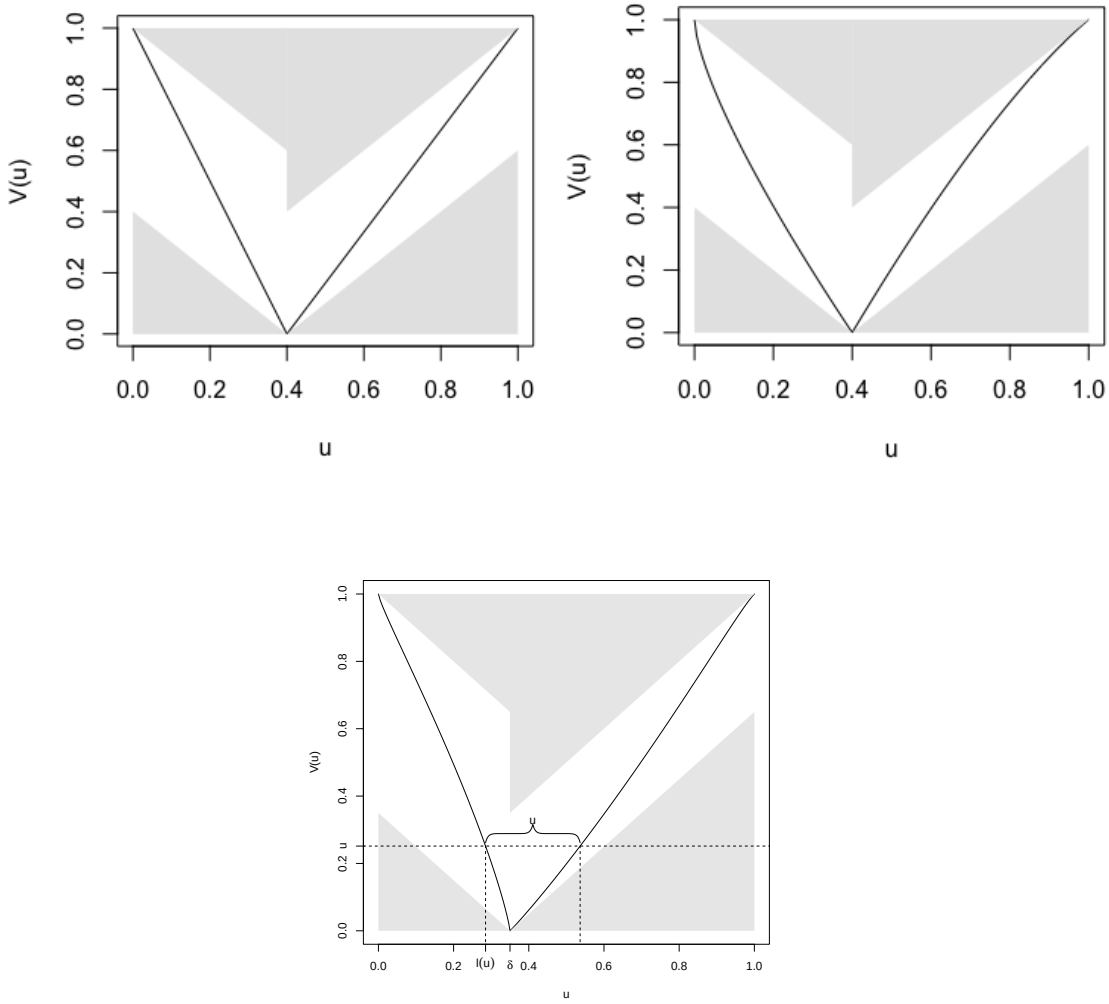


Figure 5.1: Three types of v-transform plots. Linear v-transform with parameters $\delta = 0.4$ (topleft). Two parameters v-transform where $\delta = 0.4$ and $\kappa = 0.6$ (topright). V-transform with three parameters, which are $\delta = 0.35$, $\kappa = 1.2$ and $\xi = 0.8$ (bottom).The white space is the space that the black line can lie in. For given fulcrum δ, a v-transform can never enter the gray shaded area of the plot.

Definition 5.1.3. *A v-transform is a mapping* $\mathcal{V} : [0,1] \to [0,1]$ *which satisfies the following properties:*

1. $\mathcal{V}(0) = \mathcal{V}(1) = 1$;
2. *There is a fulcrum point* δ *existing such that* $0 < \delta < 1$ *and* $\mathcal{V}(\delta) = 0$;
3. $\mathcal{V}$ *is continuous;*
4. $\mathcal{V}$ *is strictly decreasing on* $[0, \delta]$ *and strictly increasing on* $[\delta, 1]$;
5. *Every point* $u \in [0,1]$ *except* δ *has a dual point* u^* *on the opposite side of the fulcrum* δ *that satisfies* $\mathcal{V}(u) = \mathcal{V}(u^*)$ *and* $|u - u^*| = \mathcal{V}(u)$. *This property is called square property.*

The square property in Definition 5.1.3 is proved by McNeil [2021]. This property means that the events $\{V \leq v\}$ and $\{\min(u, u^*) \leq U \leq \max(u, u^*)\}$ are the same. Therefore, the probabilities of the two events happening are equal, which is

$$v = \mathbb{P}(V \leq v) = \mathbb{P}(\min(u, u^*) \leq U \leq \max(u, u^*)) = |u^* - u|. \tag{5.10}$$

The square property can be used to derive the inverse function of the v-transform. Inverting v-transform is complicated because each v corresponds to two u when $v \neq \delta$. The two dual points that yield the same value are distributed on two sides of the v-shape with different probabilities when in asymmetric v-transforms. The probability of the inverse value on the left side of the v-transform is denoted as $\Delta(v)$. We denote the inverse function by $\mathcal{V}^{-1}$ when $u \in [0, \delta]$. The partial inverse function is given by $\mathcal{V}^{-1} : [0,1] \to [0, \delta], \mathcal{V}^{-1}(v) = \inf\{u : \mathcal{V}(u) = v\}$ and let $\mathcal{V}'$ denote the gradient of $\mathcal{V}$. The gradient of v-transform is defined for all points $u \in [0,1]$, except for the change point δ. The expression of the inverse of the v-transform will be discussed in Section 5.1.3.

5.1.3 Modelling v-transforms by copulas

V-transforms $V = \mathcal{V}(V)$ describes the relationship between two uniform random variables U and V. The joint distribution function of (U, V) is a copula. Therefore, this section focuses on modelling the dependence of U and V by copulas.

Proposition 7. *Let* V *and* U *be uniform preserving random variables linked to the v-transform* $V = \mathcal{V}(U)$. *The copula* C *describing the dependence between* U *and* V *is*

$$C(u,v) = \mathbb{P}(U \leq u, V \leq v) = \begin{cases} 0 & u < \mathcal{V}^{-1}(v). \\ u - \mathcal{V}^{-1}(v) & \mathcal{V}^{-1}(v) \leq u < \mathcal{V}^{-1}(v) + v \\ v & u \geq \mathcal{V}^{-1}(v) + v. \end{cases} \tag{5.11}$$

McNeil [2021] proves that the inverse of v-transform can be obtained conditional on $V = v$, where the distribution of U is given by

$$U = \begin{cases} \mathcal{V}^{-1}(v) & \text{with probability } \Delta(v) \text{ if } v \neq 0 \\ \mathcal{V}^{-1}(v) + v & \text{with probability } 1 - \Delta(v) \text{ if } v \neq 0 \\ \delta & if v = 0 \end{cases} \tag{5.12}$$

where

$$\Delta(v) = -\frac{1}{\mathcal{V}'(\mathcal{V}^{-1}(v))}. \tag{5.13}$$

The probability $\Delta(v)$ is referred to as the conditional down probability of the v-transform and satisfies $E(\Delta(V)) = \delta$. It is the probability of a binomial distribution, which describe the probability that the value is on the left side of the fulcrum under the condition of $V = v$. The stochastic inversion function can be defined according to these features. This function can be constructed by a Bernoulli event, by which the value of V can be randomly distributed to one of the dual point U and U^* where $\mathcal{V}(U) = \mathcal{V}(U^*) = V$.

Definition 5.1.4 (Stochastic inversion function of a v-transform.)**.** *Let $\mathcal{V}$ be a v-transform with conditional down probability $\Delta(.)$. The piece-wise function $\boldsymbol{\mathcal{V}}^{-1} : [0,1] \times [0,1] \to [0,1]$ is*

$$\boldsymbol{\mathcal{V}}^{-1}(v, w) = \begin{cases} \mathcal{V}^{-1}(v) & if \quad w \leq \Delta(v) \\ v + \mathcal{V}^{-1}(v) & if \quad w > \Delta(v) \end{cases} \tag{5.14}$$

which is the stochastic inversion function of $\mathcal{V}$.

Suppose the random variables V and W are iid uniform distributed variables and a v-transform $\mathcal{V}$ with stochastic inversion function $\boldsymbol{\mathcal{V}}$ defined in Definition 5.1.4. Clearly, $\mathcal{V}(\boldsymbol{\mathcal{V}}^{-1}(v, w)) = v$ for any w. Given $U = \boldsymbol{\mathcal{V}}^{-1}(V, W)$, then $\mathcal{V}(U) = V$ and $U \sim U(0, 1)$, which means U has the conditional distribution expressed in Equation 5.12 and must be uniformly distributed.

Let $\mathcal{V}$ be a v-transform and $\boldsymbol{U} = (U_1, ..., U_d)^\top$ and $\boldsymbol{V} = (\mathcal{V}(U_1), ..., \mathcal{V}(U_d))^\top$ be vectors of uniform random variables with copula densities $c_{\boldsymbol{u}}$ and $c_{\boldsymbol{v}}$, separately. McNeil [2021] sets the assumption that $W_1, .., W_d$ are iid uniformly distributed random variables to the stochastic inversion v-transform, where

$$\boldsymbol{U} = (\boldsymbol{\mathcal{V}}^{-1}(V_1, W_1), ..., \boldsymbol{\mathcal{V}}^{-1}(V_d, W_d))^\top.$$

In this case,

$$c_{\boldsymbol{U}}(u_1, ..., u_d) = c_{\boldsymbol{V}}(\mathcal{V}(u_1), ..., \mathcal{V}(u_d)). \tag{5.15}$$

The derivation of Equation 5.15 can be found in Theorem 3 of Bladt and McNeil [2021].

5.2 Vt-ARMA models

After introducing the v-transforms and their properties, we construct a simple combination of v-transform and stationary model, the VT-ARMA model, which consists of a v-transform $\mathcal{V}$ and a Gaussian S-vine process. Following the procedure in Equation 5.1, we can generate a stationary time series from the Gaussian S-vine process.

- Firstly, we generate a causal and invertible Gaussian $ARMA(p,q)$ process (Z_t) with mean zero and variance one.
- Then, we transform the $ARMA(p,q)$ process Z_t to uniformly distributed process V_t by $V_t = \Phi(Z_t)$ for all t.
- Then, we simulate iid random variables $(W_t) \sim U(0,1)$. The series PIT process (U_t) is obtained by taking the stochastic inverses $U_t = \boldsymbol{\mathcal{V}}^{-1}(V_t, W_t)$.
- Finally, the process (X_t) can be produced by using $X_t = F_X^{-1}(U_t)$, where F_X is a continuous cdf.

Any stochastic process (X_t) generated by this process is a VT-ARMA process. Therein, the process (U_t) is a VT-ARMA copula process (McNeil [2021]). The VT-ARMA copula U_t is a strictly stationary process since the underlying ARMA process Z_t is strictly stationary. According to Equation 5.15, the process (U_t) is invariant under time shift concordance with PIT process of ARMA (V_t). Generally, the VT-ARMA is a special case of vt-S-vine, which corresponds to S-vine copula processes of a Gaussian ARMA model.

VT-ARMA copula process is a process on variants of ARMA copula. Therefore, certain properties of VT-ARMA are related to ARMA model. Assume U_t is a VT-ARMA copula process with v-transform $\mathcal{V}$ and an underlying ARMA model with autocorrelation function $\rho(k)$. Then, the random vectors $(U_{t_1}, ..., U_{t_k})$ for $k \in \mathbb{N}$ has joint density $c^{Ga}_{P(t_1,...,t_k)}(\mathcal{V}(u_1), ..., \mathcal{V}(u_k))$, where $c^{Ga}_{P(t_1,...,t_k)}$ is the density of the Gaussian copula $C^{Ga}_{P(t_1,...,t_k)}$ and $P(t_1, ..., t_k)$ is a correlation matrix with (i,j) element given by $\rho(|t_j - t_i|)$ (McNeil [2021]).

The expression of VT-ARMA copula between two variables U_t and U_{t+k} can be written as

$$\begin{aligned} C(u_t, u_{t+k}) &= \int_0^{u_t} \int_0^{u_{t+k}} c_{\boldsymbol{u}}(u_t, u_{t+k}) du_t du_{t+k}. \\ &= \int_0^{u_t} \int_0^{u_{t+k}} c_{\boldsymbol{v}}(\mathcal{V}(u_t), \mathcal{V}(u_{t+k})) du_t du_{t+k}. \\ &= \int_0^{u_t} \int_0^{u_{t+k}} c^{Ga}_{\rho(k)}(\mathcal{V}(u_t), \mathcal{V}(u_{t+k})) du_t du_{t+k}. \end{aligned} \tag{5.16}$$

The Kendall rank correlation of VT-ARMA process can be calculated by applying Equation 2.10 and 5.16.

5.3 Vt-S-vine models

VT-S-vine model consists of v-transform and S-vine models. The S-vine model is introduced in Chapter 3. The definition of vt-S-vine copula processes is given by Bladt and McNeil [2021].

Definition 5.3.1. *Let $\mathcal{V}$ be a v-transform and (V_t) be an S-vine(k) copula process. Suppose (W_t) is arbitrary strictly stationary copula process that is independent of (V_t). Set $U_t = \boldsymbol{\mathcal{V}}^{-1}(V_t, W_t)$. If (W_t) is an iid process, then the (U_t) is called a vt-S-vine(k) copula process. Otherwise, (U_t) is a generalised vt-S-vine(k) copula process, denoted as gvt-S-vine(k) process.*

Assuming a time series (V_t) is an S-vine copula process, according to Equation 3.1, the d-dimensional marginal density of the random vector $(V_t, ..., V_{t+d-1})$ is

$$c_{\boldsymbol{V}}(v_1, ..., v_d) = \prod_{k=1}^{d-1} \prod_{j=k+1}^{d} c_k(R^*_{k-1}(\boldsymbol{v}_{[j-k+1:j-1]}, v_{j-k}), R_{k-1}(\boldsymbol{v}_{[j-1:j-k+1]}, v_j)) \tag{5.17}$$

where $R^*_{k-1}(\boldsymbol{v}, x))$ and $R_{k-1}(\boldsymbol{v}, x))$ are given in Equation 3.2.

The joint density of a gvt-S-vine process is

$$\begin{aligned} c_{\boldsymbol{u}}(u_1, ..., u_d) = c_{\boldsymbol{W}}(u_1, ..., u_d) \prod_{k=1}^{d-1} \prod_{j=k+1}^{d} c_k(R^*_{k-1}(\boldsymbol{v}_{[j-k+1:j-1]}, v_{j-k}), \\ R_{k-1}(\boldsymbol{v}_{[j-1:j-k+1]}, v_j))|_{v_1=\mathcal{V}(u_1),...,v_d=\mathcal{V}(u_d)}. \end{aligned} \tag{5.18}$$

Under the condition that (W_t) is an iid process, the vt-S-vine can be written as

$$\begin{aligned} c_{\boldsymbol{u}}(u_1, ..., u_d) = \prod_{k=1}^{d-1} \prod_{j=k+1}^{d} c_k(R^*_{k-1}(\boldsymbol{v}_{[j-k+1:j-1]}, v_{j-k}), \\ R_{k-1}(\boldsymbol{v}_{[j-1:j-k+1]}, v_j))|_{v_1=\mathcal{V}(u_1),...,v_d=\mathcal{V}(u_d)}. \end{aligned} \tag{5.19}$$

In the case $\delta = 0$, the expression of Equation 5.19 is reduced to

$$c_{\boldsymbol{u}}(u_1, ..., u_d) = c_{\boldsymbol{v}}(u_1, ..., u_d), \tag{5.20}$$

where the joint density function of the vt-S-vine copula process is simplified to an S-vine copula process. In other words, the S-vine copula can be identified as a boundary or a special case of a vt-S-vine(k) copula.

The conditional density of a gvt-S-vine copula is

$$\begin{aligned} f_{U_t|U_{t-1},...,U_{t-k}}(u_t|u_{t-1},...,u_{t-k}) &= \frac{c_{\boldsymbol{U}}(u_{t-k},...,u_t)}{c_{\boldsymbol{U}}(u_{t-k},...,u_{t-1})} \\ &= c_W(u_{t-1},u_t)\prod_{i=1}^{k} c_i(R^*_{k-1}(\boldsymbol{v}_{[t-k+1:t-1]},v_{t-k}), \\ & \quad R_{k-1}(\boldsymbol{v}_{[t-1:t-k+1]},v_t))|_{v_{t-k}=\mathcal{V}(ut-k),...,v_t=\mathcal{V}(u_t)} \end{aligned} \tag{5.21}$$

For the iid process (W_t), $c_W(u_{t-1},u_t)=1$, so the conditional density of the vt-S-vine copula can be obtained from Equation 5.21.

Chapter 6

Can We Replicate GARCH Processes by Vt-S-vine Models?

Autoregressive conditional heteroskedasticity (ARCH) processes were originally proposed by Engle [1982]. Bollerslev [1986] gave the condition for covariance stationarity and introduced GARCH processes where the squared volatility is dependent on the previous squared volatilities and squared values of the process.

GARCH processes are widely applied in modelling financial time series. A GARCH option pricing model with filtered historical simulation is proposed by Barone-Adesi et al. [2008]. Franses and Van Dijk [1996] predict the stock market volatility via non-linear GARCH models. The copula based GARCH method is applied by Huang et al. [2009] to estimate VaR of portfolios.

Furthermore, variants and extensions of GARCH methods are developed in a large amount of studies. For example, Nelson [1991] proposes the exponential GARCH (EGARCH) processes that takes the leverage effect of financial time series into consideration. Similarly, the GJR-GARCH process is introduced by Glosten et al. [1993], who pointed out that the process can model the leverage effect in the nominal excess returns on stocks. Other types of GARCH models are produced to model the stylized facts of financial data in recent years, such as censored-GARCH model for asset returns with price limit (Wei [2002]), power GARCH investigating the gold price (Tully and Lucey [2007]).

GARCH models have attracted lots of researchers to study their structure and properties since they fit financial time series accurately but it is very difficult to explain why they work so well. Hence, there are some works that try to explore the structure of GARCH models by comparing the effect of modelling empirical financial time series by GARCH processes and the models we are familiar with. Loaiza-Maya et al. [2018] apply S-vine copulas to model the temporal dependence in stationary heteroscedastic time series and compare the result with GARCH models. Bladt and McNeil [2022] show that

the VT-ARMA and vt-S-vine model can rival to GARCH model in modelling Bitcoin data. Moreover, applying other kinds of models to mimic GARCH processes is also a reasonable approach to reveal the reason why GARCH models work so well. Zhao et al. [2022] mimic GARCH and GJR-GARCH processes by first-order semi-parametric S-vine process.

In this section, we attempt to replicate the GARCH processes with the vt-S-vine introduced in Chapter 5. Section 6.1 is an introduction of three types of GARCH processes. Section 6.2 demonstrates the methods and results of modelling GARCH type processes by VT-S-vine models. The procedure and results of one-step prediction by VT-S-vine for GARCH type processes are shown in Section 6.3.

6.1 GARCH process

Definition 6.1.1. *Let $(Z_t)_{t\in\mathbb{Z}}$ be a strictly white noise with mean zero and variance one. The process $(X_t)_{t\in\mathbb{Z}}$ is a GARCH(p,q) process if it is strictly stationary and satisfies for all $t \in \mathbb{Z}$ and a strictly positive-valued process $(\sigma_t)_{t\in\mathbb{Z}}$, the equations*

$$\begin{aligned} X_t &= \sigma_t Z_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \end{aligned} \tag{6.1}$$

where $\alpha_0 > 0, \alpha_i \geq 0, i = 1, ..., p$, and $\beta_j \geq 0, j = 1, ..., q$.

The process $(Z_t)_{t\in\mathbb{Z}}$ is the innovation process that consists of the innovations, which can be normal, t and any other types of distributions. The GARCH(1,1) model is the most widely applied GARCH model, since in this low-order GARCH model periods of high volatility tend to be persistent (Nelson [1990]). Meanwhile, the expression of GARCH(1,1) is simpler for analysis. Therefore, we concentrate on GARCH(1,1) model in our study.

If $(\sigma_t^2)_{t\in\mathbb{Z}}$ is a strictly stationary process, then the GARCH process $(X_t)_{t\in\mathbb{Z}}$ is stationary as well, since $X_t = \sigma_t Z_t$ and $(Z_t)_{t\in\mathbb{Z}}$ is strictly white noise. The GARCH(1,1) process is a covariance-stationary white noise process if and only if $\alpha_1 + \beta_1 < 1$. The variance of the covariance-stationary process is equal to $\alpha_0/(1 - \alpha_1 - \beta_1)$ (Proposition 4.21 in McNeil et al. [2015]). For GARCH(1.1) process, if it satisfies the condition of covariance-stationary, then it is also a strictly stationary process. In other words, the covariance-stationary has the more strict condition in GARCH(1,1).

The covariance-stationary GARCH(1,1) can be expressed as an ARMA(1,1) model. Let V_t be a martingale difference and $V_t = \sigma^2(Z_t^2 - 1)$, using Equation 6.1, the

GARCH(1,1) process can be written as

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + V_t. \tag{6.2}$$

Since $\sigma_{t-1}^2 = X_{t-1}^2 - V_{t-1}$, the X_t^2 can be expressed as

$$X_t^2 = \alpha_0 + (\alpha_1 + \beta_1) X_{t-1}^2 - \beta_1 V_{t-1} + V_t. \tag{6.3}$$

We can rebuild Equation 6.3 to write it in a mode of ARMA(1,1) process.

$$\left(X_t^2 - \frac{\alpha_0}{1-\alpha_1-\beta_1}\right) = (\alpha_1+\beta_1)\left(X_{t-1}^2 - \frac{\alpha_0}{1-\alpha_1-\beta_1}\right) - \beta_1 V_{t-1} + V_t. \tag{6.4}$$

Equation 6.4 presents an ARMA(1,1) process under the condition that $E(X_t^4) < \infty$ and $\alpha_1 + \beta_1 < 1$; details can be found in [McNeil et al., 2015, p. 120].

The standard GARCH model, describes the time series in a rigidly symmetric way where the volatility reacts to recent returns without concern about their sign. This feature of GARCH model violates the economic theory of leverage effect, where the market information should have an asymmetric effect on volatility. Therefore, we introduce two GARCH models that add an additional parameter to describe the leverage effect.

Nelson [1991] proposes the EGARCH(1,1) model

$$\ln \sigma_t^2 = \alpha_0 + \alpha_1 Z_{t-1} + \gamma(|Z_{t-1}| - E(|Z_{t-1}|)) + \beta_1 \ln \sigma_{t-1}^2, \tag{6.5}$$

where the coefficient α_1 manifests the sign effect and γ describes the size effect. The expression $\alpha_1 Z_{t-1} + \gamma(|Z_{t-1}| - E(|Z_{t-1}|)$ allows the volatility σ_t^2 to be affected asymmetrically by rises and falls of financial time series. In EGARCH process, the $\ln \sigma_t^2$ is a linear process, so its covariance-stationarity and ergodicity are easy to check.

Let $\epsilon_t = \alpha_1 Z_{t-1} + \gamma(|Z_{t-1}| - E(|Z_{t-1}|))$, then $(\epsilon_t)_{t\in\mathbb{Z}}$ is strictly stationary, since ϵ only depends on Z_t and Z_t is a strictly white noise. If we let $Y_t = \ln \sigma_t^2$, then $Y_t = \alpha_0 + \epsilon_t + \beta_1 Y_{t-1}$. Therefore, $(Y_t)_{t\in Z}$ is an AR(1) process, which will be strictly stationary process under the condition $|\beta_1| < 1$. For EGARCH(1,1), the condition of covariance stationary is the same as strictly stationary, which is $|\beta_1| < 1$; proof can be found in Exercise 4.20 in Hofert et al. [2020].

Furthermore, the GJR-GARCH(1,1) model with an additional parameter γ can also describe the asymmetric effect of good and bad news on volatility. Glosten et al. [1993] defined the GJR-GARCH process(1,1) by

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \gamma I_{\{X_{t-1}<0\}} X_{t-1}^2, \tag{6.6}$$

where $I_{t-1} = 0$ if $X_{t-1} \geq 0$ and $I_{t-1} = 1$ if $X_{t-1} < 0$. γ represents the leverage term.

In the GJR-GARCH model, the factor I_{t-1} can add the influence of negative values in financial time series to GARCH model and parameter γ facilitates the description of asymmetry effect on volatility.

6.2 Modelling GARCH processes by vt-S-vine model

In order to study the structure of GARCH processes, we use vt-S-vine models to fit the simulated data from standard GARCH, EGARCH and GJR-GARCH processes. Zhao et al. [2022] have also investigated whether S-vine models can mimic GARCH, but their study is restricted to models of first and second order with t copulas. We extend the model to higher-order S-vine copulas. Meanwhile, the v-transform is combined with the S-vine copulas to describe the cross-shaped copula density in GARCH type process.

6.2.1 Methods of replicating GARCH type processes

There are two types of estimation methods for S-vine models, semi-parametric and parametric. Similarly, the vt-S-vine has the two modes as well. For semi-parametric vt-S-vine models, we use empirical distribution to transform the simulations from GARCH processes to uniform scale and fit them to vt-S-vine copulas. The parametric vt-S-vine model is selecting a suitable parametric marginal distribution first and then model the data with the entire model consisting of the vt-S-vine copula and margins.

We try three types of v-transforms, linear, two-parameter, and three-parameter v-transforms. The choices of copulas we used is the same as the ones used in Section 4.4. The estimation approach in S-vine copulas is also the same as in Section 4.4, where applying kpacf from ARMA process to approximate the kpacf values of S-vine process to reduce the amount of the parameters. The order of vt-S-vine copula is 40. The order of ARMA process used to estimate the kpacf values is $(1, 1)$, because the GARCH(1,1) process can be written as the form of ARMA (1,1); the expression is in Equation 6.4. The process of mimicking GARCH type processes to semi-parametric vt-S-vine model is:

- Simulate 5000 data X_t from standard GARCH(1,1), EGARCH(1,1) and GJR-GARCH(1,1) processes with normal innovation.
- Transform simulations to uniform scale U_t by empirical distribution.
- Find the best fulcrum by comparing maximum likelihood values (we input values of fulcrum evenly distributed between 0 and 1, and calculate the likelihood values in each case to select the fulcrum with the largest likelihood value) and plot profile likelihood at a regularly spaced grid of values for δ.

- Fix the fulcrum δ to v-transform and fit the model with v-transform and S-vine copulas to the uniformly distributed U_t to estimate the parameters and calculate the AIC values.

- Compare the AIC values with the first and second-order S-vine process with t copula.

The process of replicating GARCH type processes by the parametric vt-S-vine model is similar to the one by the semi-parametric vt-S-vine model. We add one step to find the suitable marginal distribution to describe the simulations from GARCH processes before transforming data to uniformly distributed. The choice of parametric marginal distributions are same as the ones used in Section 4.4. Then, the simulations are transformed to uniform scale by the chosen parametric margin in the second step. After finding the best fulcrum, we fit the full model with margins, v-transforms and S-vine copulas to the simulations and estimate the parameters.

The parameters of vt-S-vine models are estimated by the maximum likelihood estimation (MLE) approach. Let $\mathbf{x} = \{x_1, ..., x_n\}$ be a realization from a strictly stationary process with parametric marginal distribution $F_X(x; \boldsymbol{\theta}_m)$ and joint copula density $c_{\mathbf{U}}(u_1, ..., u_n; \boldsymbol{\theta}_c)$, where $\boldsymbol{\theta}_m$ is the parameters in marginal distribution and $\boldsymbol{\theta}_c$ and $\boldsymbol{\theta}_v$ are the parameters of the S-vine copula density and v-transform, separately. The full log-likelihood is

$$\begin{aligned} L(\boldsymbol{\theta}_c, \boldsymbol{\theta}_v, \boldsymbol{\theta}_m; \mathbf{x}) = \sum_{k=1}^{d-1} \sum_{j=k+1}^{d} \log c_k(R^*_{k-1}(\boldsymbol{v}_{[j-k+1:j-1]}, v_{j-k}), \\ R_{k-1}(\boldsymbol{v}_{[j-1:j-k+1]}, v_j))|_{v_1=\mathcal{V}(u_1;\boldsymbol{\theta}_v),...,v_d=\mathcal{V}(u_d;\boldsymbol{\theta}_v)} \\ + \sum_{i=1}^{n} \log\left(f_X(x_i; \boldsymbol{\theta}_m); \boldsymbol{\theta}_c\right), \end{aligned} \tag{6.7}$$

where d is the order of the vt-S-vine copula and $R^*_{k-1}(\boldsymbol{v}, x)$ and $R_{k-1}(\boldsymbol{v}, x)$ are given in Equation 3.2.

In our copula-based estimation method, we follow the inference-functions-for-margins (IFM) approach introduced in Joe [1997] and estimate the marginal distribution and copula model in two steps. The suitable margins $\hat{F}_X$ are selected first by using the second part of Equation 6.7 and then by using the pseudo-copula data $\{\hat{u}_i = \hat{F}_X(x_i; \hat{\boldsymbol{\theta}}_m), i = 1, ..., n\}$ to estimate fulcrum δ in v-transform. Finally, we fit the full model using Equation 6.7. Before estimating the parameters in copula sequences, the ARMA(1,1) is used to approximate the kpacf in S-vine and the parameters in bivariate copula sequences $\boldsymbol{\theta}_c$ are expressed by the parameters $\{\phi_1, \theta_1\}$ in ARMA(1,1). Hence, only two parameters are left in the MLE process for S-vine copulas. The details of the approach are exactly the same as the one discussed in Section 3.2.2. For semi-

parametric vt-S-vine model, the log-likelihood is equal to the first part of the Equation 6.7.

Note that for some original data point x_t, $\delta = u_t = F_X(x_t)$, where F_X is empirical distribution, then $\mathcal{V}(u_t) = 0$ and the log-likelihood for the copula takes the value $-\infty$, which implies that the profile likelihood of δ is not differentiable at such points and has multiple local maxima. Therefore, we apply a grid search to finding the optimal estimation of δ, instead of continuous optimization in the interval $[0, 1]$; details are discussed by Bladt and McNeil [2021].

6.2.2 Simulations

The GARCH type data are simulated from GARCH(1,1), EGARCH(1,1) and GJR-GARCH(1,1) processes. In order to compare the results with Zhao et al. [2022], we use the GARCH and GJR-GARCH processes with the same parameters and innovation in this paper. The innovation distribution we use here is normal. We set the parameters to be $[\alpha_0, \alpha_1, \beta_1] = [0.05, 0.1, 0.85]$ for the GARCH process and $[\alpha_0, \alpha_1, \beta_1, \gamma] = [0.05, 0.1, 0.85, 0.05]$ for the GJR-GARCH process. These parameters match the real world financial data broadly (Oh and Patton [2013]). Furthermore, the parameters of EGARCH process are chosen corresponding to GARCH and GJR-GARCH. Also, we make slight adjustment to the parameters to make the dependence between data stronger, which can illustrate the differences between vt-S-vine models more clearly. The parameters of EGARCH with normal innovation are set to be $[\alpha_0, \alpha_1, \beta_1, \gamma] = [0.05, -0.03, 0.95, 0.25]$.

We simulate 5000 data for each model. The simulated time series are plot in Figure 6.1. The plots present volatility clustering features of GARCH type simulations.

In order to show the conditional dependence of the data, we plot the partial autocorrelation (pacf) of the data. Moreover, the pacf of absolute values of the data are plotted to demonstrate the conditional dependence of the volatility proxy series of data as an example. The pacf plots are shown in Figure 6.2.

Figure 6.2 demonstrates that the absolute values of GARCH type processes have strong dependence, which embodies the positive persistence serial correlations.

6.2.3 Modelling GARCH type processes by vt-S-vine

After simulating the GARCH type time series, the vt-S-vine models are fitted to the time series. We also attempt the first-order and second-order S-vine model with t copulas and compare the AIC values with vt-S-vine models. The results of semi-parametric and parametric models are discussed in the following sections.

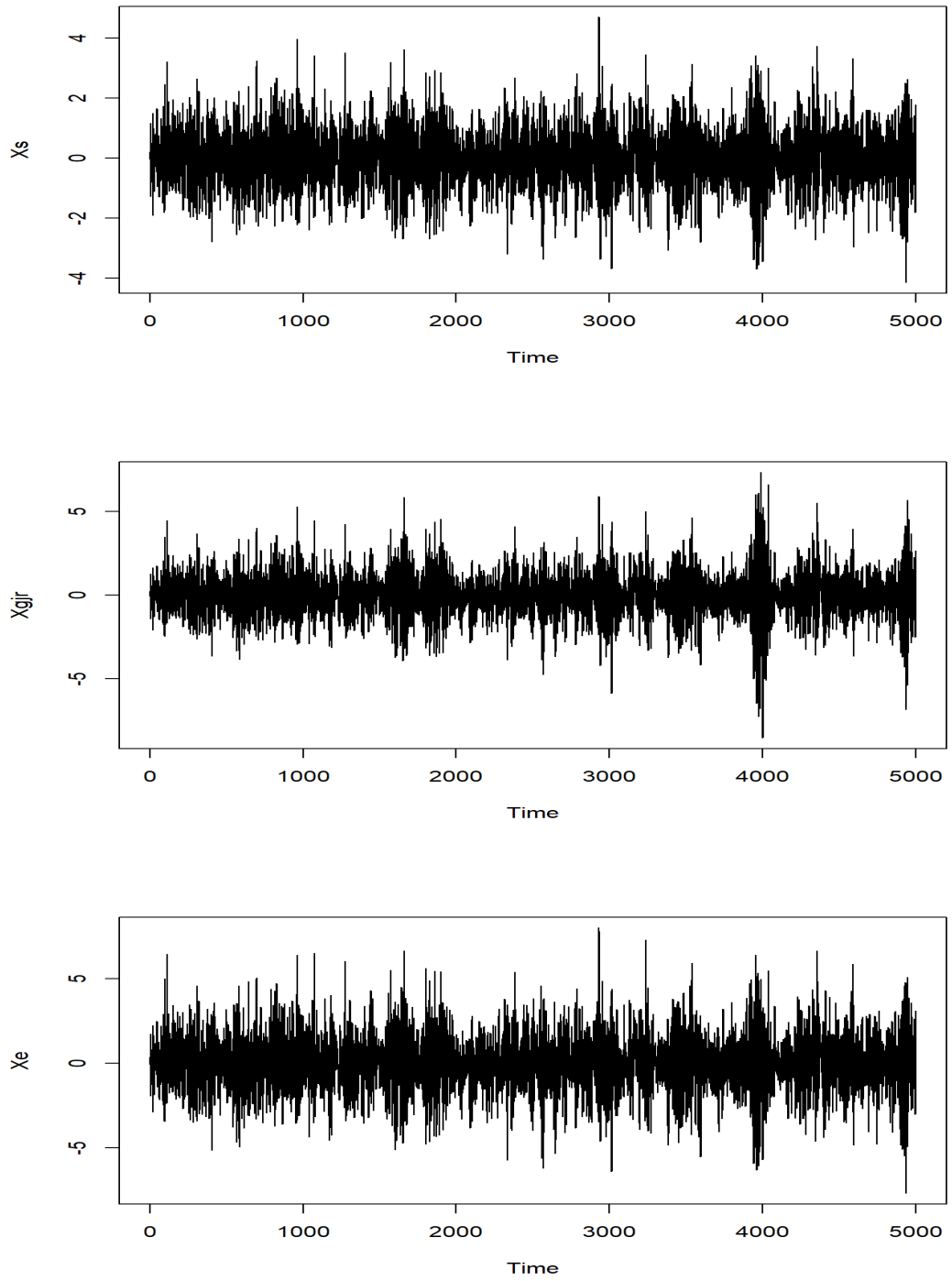


Figure 6.1: Three simulations from GARCH processes (top), GJR-GARCH (middle) and EGARCH (bottom). The parameters are presented in Section 6.2.2.

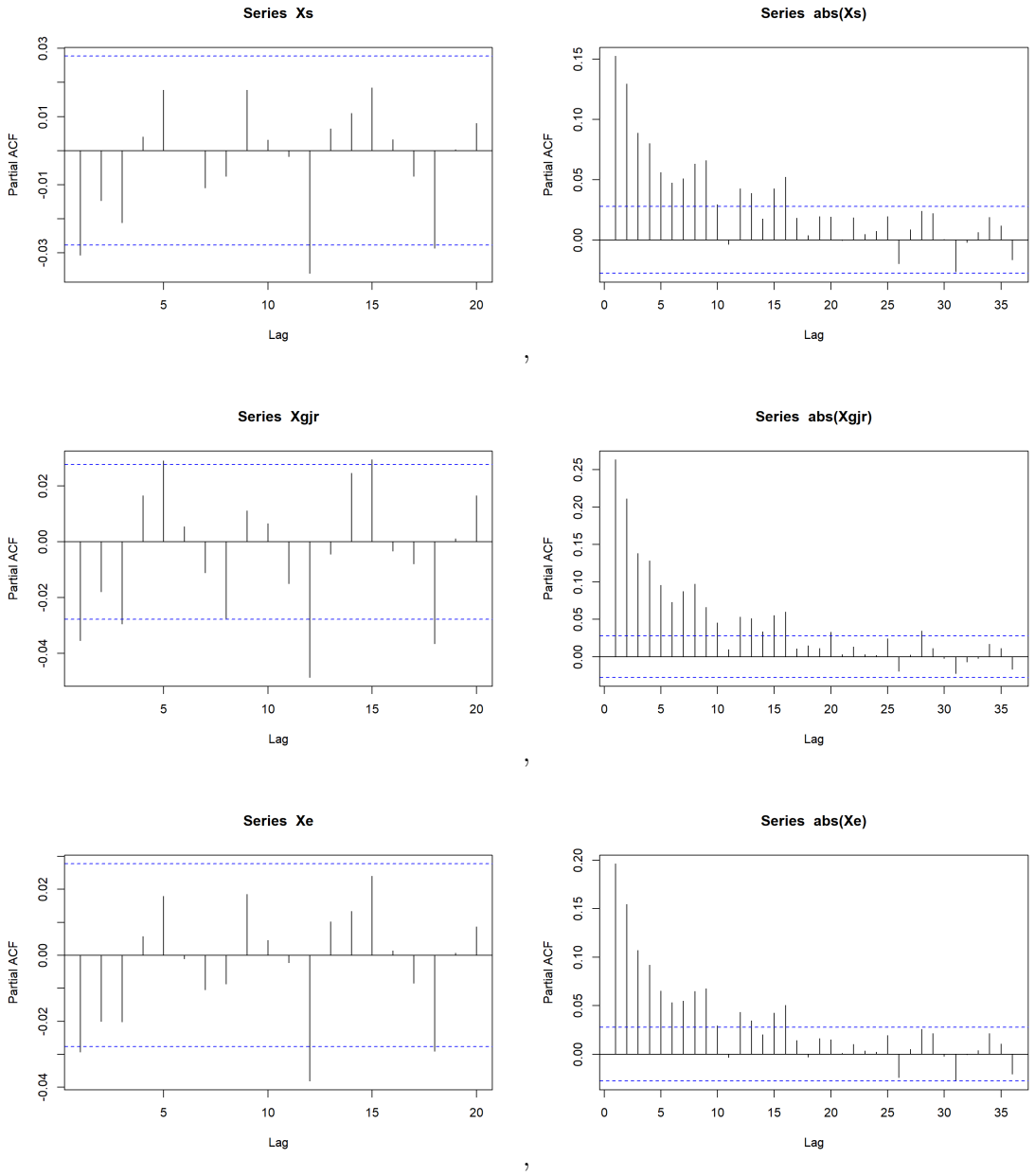


Figure 6.2: Pacf plots of GARCH (left-top), absolute values of GARCH (right-top), GJR-GARCH (left-middle), absolute values of GJR-GARCH (right-middle), EGARCH (left-bottom) and absolute values of EGARCH (right-bottom).

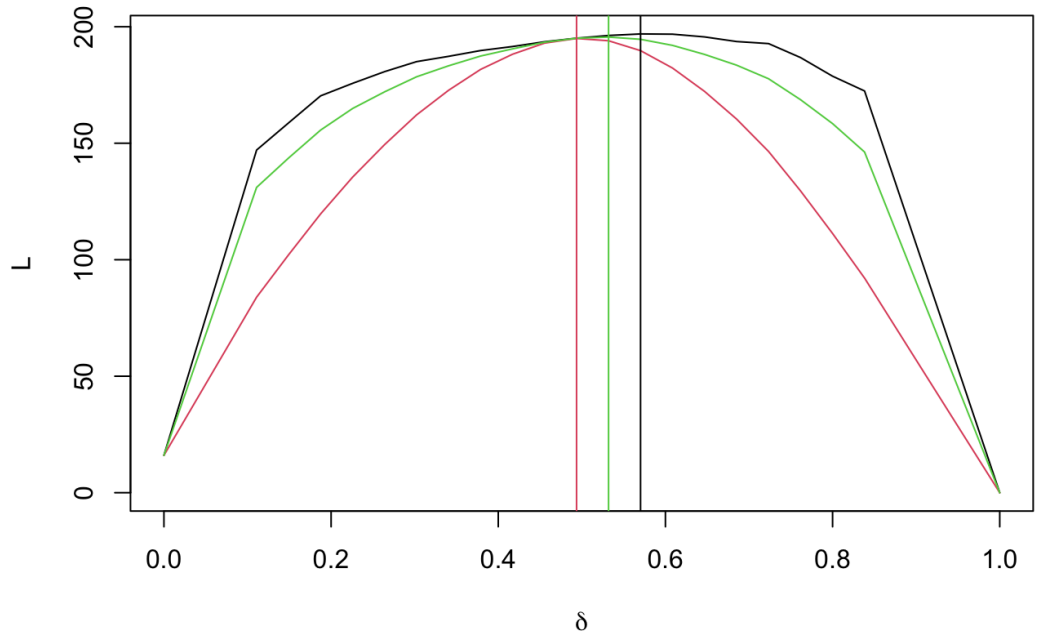


Figure 6.3: The profile log-likelihood function describing the relationship between fulcrum values δ and the corresponding log likelihood values of the vt-S-vine models with survival Clayton copulas fitted to the simulations from GARCH processes. The x-axis represents the fulcrum values and y-axis shows the corresponding log likelihood values. The red, green and black lines are the vt-S-vine copulas with linear, two-parameter and three-parameter v-transforms, respectively. The vertical lines give the best fulcrum corresponding to the maximum log likelihood values in the three types of v-transforms.

Semi-parametric vt-S-vine models

In the semi-parametric vt-S-vine models, the simulations are transformed to the uniformly distributed time series U_t by empirical distribution. For the fulcrum search, we draw the profile likelihood of the fulcrum values for different copula sequences, including Gaussian, Gumbel, Joe, survival Clayton and Frank copulas. In this section, we also apply the models based on Gaussian copula sequences with non-Gaussian substitutes at the first few lags, including the model with Gumbel for the first ten orders and Gaussian for the following thirty orders denoted as $Gu(10)\_Ga(30)$ and the model with survival Clayton for the first ten orders and Gaussian copulas thereafter denoted as $SC(10)\_Ga(30)$. Figure 6.3 presents the relationship between fulcrum values δ and the corresponding log likelihood values of the vt-S-vine models with survival Clayton copulas fitted to the simulations from GARCH processes.

According to Figure 6.3, the three-parameter v-transform always has the largest log likelihood and the two-parameters v-transform is the second choice for each fulcrum.

However, the fulcrums leading to the maximum log likelihood values concentrate on the range of $[0.4, 0.6]$. Besides, the log likelihood values of the three types of v-transforms are closer when the fulcrum is taken in this range. Therefore, the selection of v-transform does not affect the estimation results significantly when the optimal fulcrum is chosen. In other words, we can use linear v-transform as a representative of v-transforms to reduce the amount of parameters and simplify the vt-S-vine model in some cases.

After finding the best fulcrum, we fit the vt-S-vine copula with fixed fulcrum to the transformed uniformly distributed time series U_t and calculate the AIC values. Furthermore, the first-order and second-order S-vine copulas with the t copula are used to model U_t as a comparison. These results are shown in Table 6.1.

Table 6.1 demonstrates that the higher-order vt-S-vine processes capture the stylized facts of GARCH type simulations much better than the first and second-order S-vine processes with t copulas. The vt-S-vine process with the survival Clayton copula yields the best results in the three data groups. Moreover, the vt-S-vine with mixture copula sequences does not improve the results of copula sequences with single copula. The linear v-transform shows smaller AIC values than the v-transforms with two and three parameters in many cases and the AIC values from the three types of v-transform are closer to each other. This observation is in accordance with the discussion in Figure 6.3. Another method to assess the estimation results is to do the residual plot. The residuals of semi-parametric vt-S-vine processes with survival Clayton copula sequences and linear v-transform fitted to EGARCH process are shown in Figure 6.4.

According to Figure 6.4, the upper tail of residuals slightly deviates from the standard line. This may be caused by the asymmetry of the simulated time series. The situations of the other two datasets are similar. The survival Clayton sequences may not capture the upper tail dependence perfectly. It is difficult to assess if the semi-parametric vt-S-vine models GARCH type process well enough. However, we can explore the improvement of estimating GARCH type processes via comparing the AIC values with the existing models, such as S-vine processes with t copula. Another approach is to fit the simulated time series to the GARCH process and compare the AIC values with parametric vt-S-vine to see how much difference the AIC with the real models. The parametric vt-S-vine model is presented in the next section.

Semi-parametric vt-S-vine models with two-parameter copulas

This subsection investigates the replication of GARCH-type processes by vt-S-vine copulas with linear v-transform and two-parameter copulas, such as t and BB1 copulas, which are used by Nikoloulopoulos et al. [2012] to model the asymmetric tail dependence in time series. In the previous section, we found that vt-S-vine copulas with survival Clayton copula sequences can provide the best replication for GARCH-type processes. Hence, we apply survival Clayton copulas at higher level in vt-S-vine models and sub-

AIC values for GARCH simulations			
$S-vine(1)-t$	-112.8375		
$S-vine(2)-t$	-194.4746		
	Vlinear (fulcrum)	V2p (fulcrum)	V3p (fulcrum)
$Gaussian$	-262.20 (0.4937)	-264.48 (0.5320)	-263.29 (0.5702)
$Gumbel$	-354.78 (0.4937)	-355.94 (0.5320)	-355.77 (0.5702)
Joe	-351.71 (0.4554)	-351.33 (0.4937)	-350.09 (0.5702)
$Clayton180$	-385.92 (0.4937)	-385.14 (0.5320)	-385.82 (0.5702)
$Frank$	-224.36 (0.4937)	-224.49 (0.5320)	-224.90 (0.5702)
$Gu(10)\_Ga(30)$	-350.90 (0.4554)	-349.99 (0.4937)	-348.97 (0.5702)
$SC(10)\_Ga(30)$	-374.00 (0.4937)	-372.14 (0.4937)	-374.52 (0.6085)
AIC values for GJR-GARCH simulations			
$S-vine(1)-t$	-279.6331		
$S-vine(2)-t$	-478.6405		
	Vlinear (fulcrum)	V2p (fulcrum)	V3p (fulcrum)
$Gaussian$	-600.14 (0.4937)	-600.96 (0.5320)	-600.13 (0.5320)
$Gumbel$	-818.78 (0.4937)	-818.28 (0.4937)	-818.59 (0.5320)
Joe	-830.67 (0.4937)	-831.23 (0.4937)	-831.65 (0.5320)
$Clayton180$	-849.42 (0.4937)	-849.89 (0.4937)	-849.52 (0.5320)
$Frank$	-530.15 (0.4937)	-529.14 (0.5320)	-527.93 (0.5320)
$Gu(10)\_Ga(30)$	-807.68 (0.4937)	-806.69 (0.4937)	-806.13 (0.4937)
$SC(10)\_Ga(30)$	-830.85 (0.4937)	-830.47 (0.4937)	-829.60 (0.6085)
AIC values for EGARCH simulations			
$S-vine(1)-t$	-168.0197		
$S-vine(2)-t$	-278.5629		
	Vlinear (fulcrum)	V2p (fulcrum)	V3p (fulcrum)
$Gaussian$	-410.02 (0.5320)	-409.09 (0.5320)	-407.57 (0.5320)
$Gumbel$	-484.29 (0.4937)	-482.80 (0.5320)	-481.94 (0.5320)
Joe	-457.97 (0.4937)	-456.07 (0.4937)	-454.82 (0.5320)
$Clayton180$	-538.66 (0.5320)	-537.46 (0.5320)	-537.60 (0.5320)
$Frank$	-366.47 (0.5320)	-366.50 (0.5320)	-364.64 (0.5320)
$Gu(10)\_Ga(30)$	-483.68 (0.4937)	-481.69 (0.4937)	-479.75 (0.4937)
$SC(10)\_Ga(30)$	-527.56 (0.4937)	-526.28 (0.4937)	-525.22 (0.5320)

Table 6.1: The AIC values of the semi-parametric S-vine and vt-S-vine copulas fitted to GARCH, GJR-GARCH and EGARCH processes. $S-vine(1)-t$ and $S-vine(2)-t$ are the first-order and second-order S-vine processes with t copulas. Vlinear, V2p and V3p are the linear, two-parameter and three-parameter v-transforms, separately. The values in the brackets are the fulcrum chosen according to the profile fulcrum functions. The order of the vt-S-vine process is 40. $Clayton180$ represents the survival Clayton copulas. $Gu(10)\_Ga(30)$ is the S-vine copula with the Gumbel copula in the first ten order and Gaussian copula in the following 30 orders. $SC(10)\_Ga(30)$ is the one with survival Clayton replacing the first ten Gumbel copulas in $Gu(10)\_Ga(30)$. The numbers in yellow are the AIC values of the best models in each dataset. The blue colour highlights the best copula sequences in each dataset.

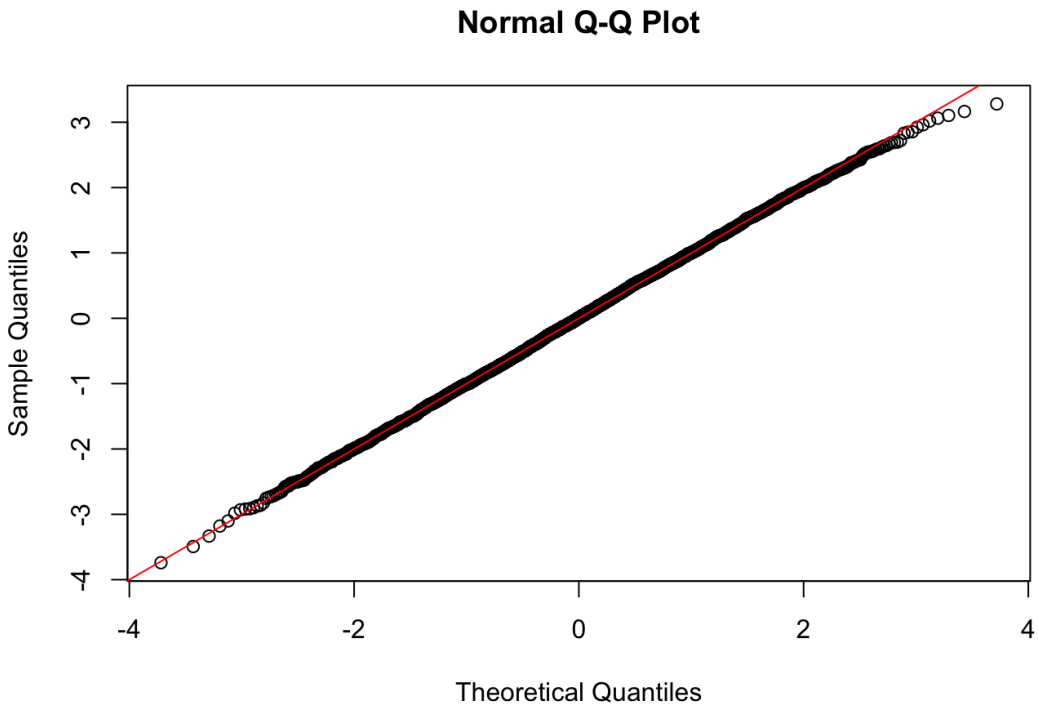


Figure 6.4: The residuals plot of semi-parametric vt-S-vine process with survival Clayton copula sequences and linear v-transform fitted to EGARCH process.

	GARCH	GJR-GARCH	EGARCH
$Vlinear + t(1)\_sclay(39)$	-255.63	-503.82	-332.43
$Vlinear + sBB1(1)\_sclay(39)$	-230.53	-523.98	-344.79

Table 6.2: The AIC values of the semi-parametric S-vine and vt-S-vine copulas fitted to GARCH, GJR-GARCH and EGARCH processes.

stitute the copula at fist level to t or BB1 copula. Because the vt-S-vine is truncated at lag 40, the models are denoted as $t(1)\_sclay(39)$ and $sBB1(1)\_sclay(39)$. The survival BB1 copula obtain better estimation results in GARCH replication, so we decide to use survival BB1 copula. In addition, the linear v-transform is used to simplify the calculation. The estimation results of semi-parametric vt-S-vine models with two-parameter copulas are shown in Table 6.2.

According to Table 6.2, the two types of two-parameter copula do not improve the replication of GARCH-type processes, which means the feature of tail dependence of GARCH-type processes may not be similar to the one of t or BB1 copulas. In v-transforms, the large changes in either positive or negative sides will be transformed to large changes; small changes will still be small changes. It means that only large movements in price has obvious tail dependence. That's why the one-tail dependence copula is better for the financial data.

Parametric vt-S-vine models

The first step for the parametric vt-S-vine model is to choose the suitable marginal distributions. The candidate margins includes normal (norm), skewed Laplace (slap), double Weibull (dwe), skewed double Weibull (sdwe), student t (t), skewed student t (st), normal inverse Gaussian (NIG), hyperbolic (hyp) and generalised hyperbolic (ghyp) distributions. The candidate margins are estimated by the GARCH type time series. We select the margins with the smallest AIC values and the most accurate QQ-plot. The AIC values of the three datasets are demonstrated in Table 6.3.

According to Table 6.3, the best margins for GARCH, GJR-GARCH and EGARCH simulation processes are student t, student t and hyperbolic distributions. In order to demonstrate the margins intuitively, the QQ-plots of the best marginal distributions of the three datasets are shown in Figure 6.5.

The QQ-plots of marginal distributions are close to the standard red line in GARCH and EGARCH simulations. The estimation GJR-GARCH process deviates slightly from the red line, but the student t distribution is the best margin among the nine candidates for GJR-GARCH. Hence, we use the chosen margins to transform data to an approximately uniform and use the same approach to find the fulcrums. Following this, the full models with vt-S-vine and chosen margins are fitted to GARCH type processes. The AIC values are shown in Table 6.4.

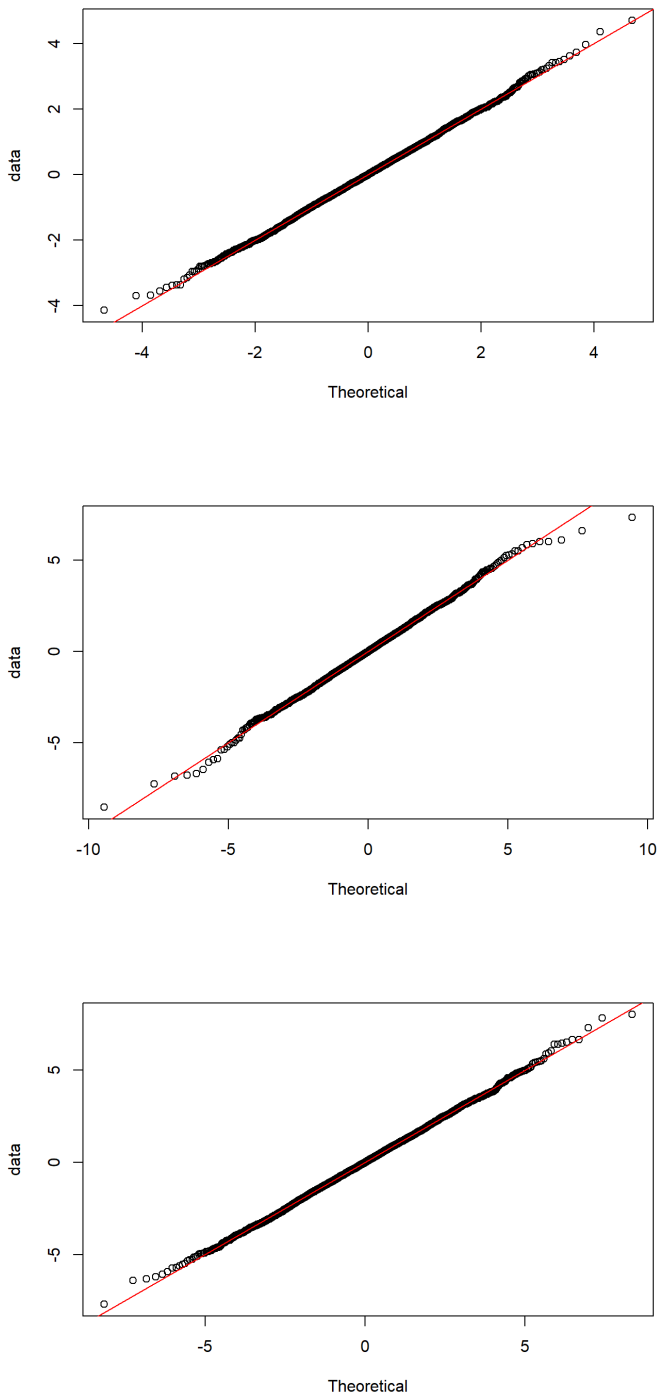


Figure 6.5: QQ-plots of best margins in simulations from GARCH (top), GJR-GARCH (middle) and EGARCH (bottom).

		GARCH	GJR-GARCH	EGARCH
Margins	norm	14096.22	17404.29	19665.19
	slap	14343.52	17243.92	19817.38
	dwe	14065.97	17107.61	19589.53
	sdwe	14067.79	17109.52	19591.07
	t	14048.80	17098.08	19584.17
	st	14050.77	17100.03	19585.99
	NIG	14049.79	17102.09	19582.96
	hyp	14049.68	17107.65	19582.40
	ghyp	14051.94	17101.36	19585.62

Table 6.3: The AIC values of marginal distributions fitted to the simulations from GARCH, GJR-GARCH and EGARCH processes. The numbers in yellow are the AIC values of the best models in each dataset.

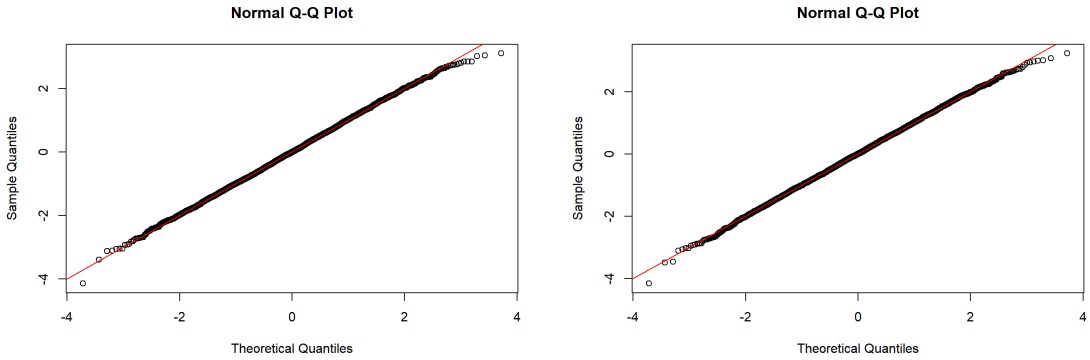
Table 6.4 shows that the parametric vt-S-vine processes perform much better than S-vine process with t copulas. The best copula sequences are survival Clayton copulas. Three-parameter v-transforms present better results than the other two v-transforms in most cases, but the differences are small. Hence, it is still reasonable to take the linear v-transform as representative v-transform to reduce the number of parameters and simplify the full models. The GARCH type simulations are also modelled by the true model and calculate the AIC values as a benchmark to analyse the estimation results from vt-S-vine and S-vine processes.

The best models are vt-S-vine processes with three-parameter v-transform and survival Clayton copula sequences in all these three datasets. Although there still is a small gap between their AIC values and true model's AIC, the improvement is significant. The full models consist of the marginal distributions and vt-S-vine copulas. Hence, the gap may be caused by the marginal distribution, since the margins we attempt are not perfect, especially the tail estimation, which can be shown in QQ-plot in Figure 6.5. Furthermore, the mixed copula sequences in vt-S-vine process do not present advantages in estimating the GARCH type processes compared to the vt-S-vine with single types of copulas.

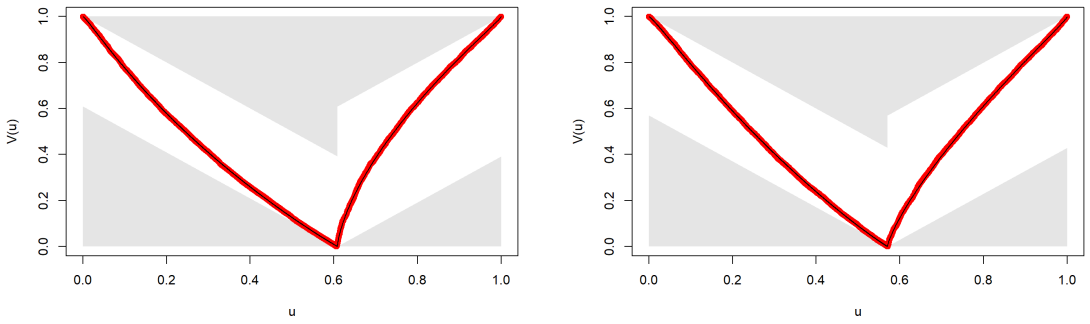
In order to present the estimation results, we put the residuals, v-transforms and kpacf plots of fitting parametric vt-S-vine models with three-parameters v-transform and survival Clayton copula sequences to GARCH and EGARCH processes in Figure 6.6. The residuals of the two datasets have bias in the upper tail, which is in accordance with the conclusion in semi-parametric vt-S-vine estimation. Therefore, it may be caused by the copulas sequences. The three-parameter v-transform of the two datasets have right-twist shape and fit the uniformly distributed time series accurately. The kpacf plots demonstrate that it is reasonable to use ARMA process to approximate the kpacf of GARCH type processes in vt-S-vine copula estimation.

AIC values for GARCH simulations [13644.67]			
$S-vine(1)-t$	13935.91		
$S-vine(2)-t$	13854.34		
	Vlinear (fulcrum)	V2p (fulcrum)	V3p (fulcrum)
$Gaussian$	13787.63 (0.4937)	13780.95 (0.5320)	13782.75 (0.5320)
$Gumbel$	13694.65 (0.4937)	13695.50 (0.5320)	13690.52 (0.5702)
Joe	13698.16 (0.4937)	13699.95 (0.4937)	13696.71 (0.5702)
$Clayton180$	13664.55 (0.4937)	13665.69 (0.5320)	13663.62 (0.6085)
$Frank$	13822.68 (0.4937)	13822.41 (0.5320)	13824.33 (0.5702)
$Gu(10)\_Ga(30)$	13697.75 (0.4554)	13699.36 (0.4554)	13696.73 (0.5702)
$SC(10)\_Ga(30)$	13675.78 (0.4937)	13677.75 (0.4937)	13674.61 (0.6085)
AIC values for GJR-GARCH simulations [16202.66]			
$S-vine(1)-t$	16815.19		
$S-vine(2)-t$	16617.27		
	Vlinear (fulcrum)	V2p (fulcrum)	V3p (fulcrum)
$Gaussian$	16495.45 (0.4937)	16490.34 (0.5320)	16491.61 (0.5320)
$Gumbel$	16276.58 (0.4937)	16277.31 (0.4937)	16276.11 (0.5320)
Joe	16263.66 (0.4937)	16263.34 (0.4937)	16263.32 (0.5320)
$Clayton180$	16246.37 (0.5320)	16246.90 (0.4937)	16245.95 (0.5320)
$Frank$	16565.89 (0.4937)	16566.44 (0.5320)	16567.80 (0.5320)
$Gu(10)\_Ga(30)$	16285.72 (0.4937)	16286.79 (0.4937)	16287.95 (0.4937)
$SC(10)\_Ga(30)$	16264.51 (0.4937)	16264.43 (0.4937)	16264.65 (0.6085)
AIC values for EGARCH simulations [19008.73]			
$S-vine(1)-t$	19414.06		
$S-vine(2)-t$	19303.95		
	Vlinear (fulcrum)	V2p (fulcrum)	V3p (fulcrum)
$Gaussian$	19178.76 (0.4937)	19172.07 (0.5320)	19173.65 (0.5320)
$Gumbel$	19098.96 (0.4937)	19100.46 (0.4937)	19097.40 (0.5702)
Joe	19124.66 (0.4937)	19125.70 (0.4937)	19121.04 (0.5702)
$Clayton180$	19046.06 (0.5320)	19047.48 (0.4937)	19045.01 (0.5702)
$Frank$	19214.72 (0.5320)	19214.72 (0.5320)	19216.64 (0.5320)
$Gu(10)\_Ga(30)$	19098.55 (0.4937)	19100.20 (0.4937)	19100.12 (0.5320)
$SC(10)\_Ga(30)$	19056.41 (0.4937)	19057.34 (0.4937)	19057.42 (0.6085)

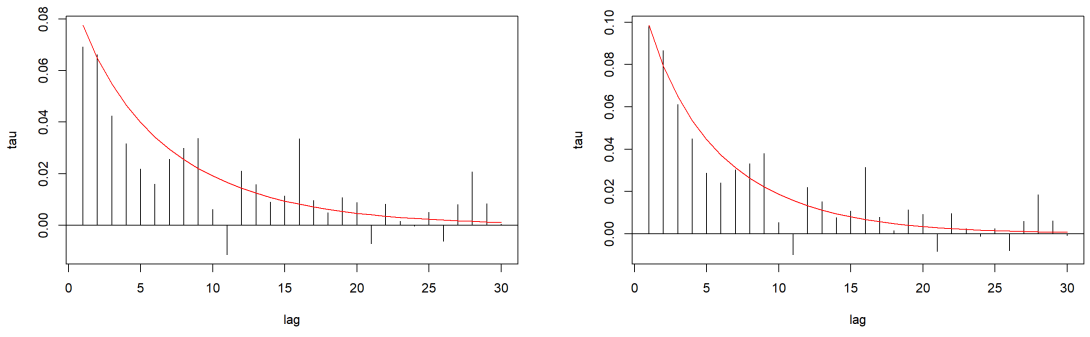
Table 6.4: The AIC values of parametric S-vine and vt-S-vine processes fitted to GARCH, GJR-GARCH and EGARCH processes. The margins for GARCH, GJR-GARCH and EGARCH simulation processes are student t, student t and hyperbolic distributions. $S-vine(1)-t$ and $S-vine(2)-t$ are the first-order and second-order S-vine process with t copulas. Vlinear, V2p and V3p are the linear, two-parameter and three-parameter v-transforms, separately. The values in the brackets are the fulcrum chosen according to the profile fulcrum functions. The order of vt-S-vine process is 40. $Clayton180$ represents the survival Clayton copulas. $Gu(10)\_Ga(30)$ is S-vine copula with Gumbel copula in the first ten order and Gaussian copula in the following 30 orders. $SC(10)\_Ga(30)$ is the one with survival Clayton replacing the first ten Gumbel copulas in $Gu(10)\_Ga(30)$. The values in the square brackets are the AIC values of GARCH models fitted to the simulations from the GARCH model. The numbers in yellow are the AIC values of the best models in each dataset. The blue colour highlights the best copula sequences in each dataset.



(a) Residuals in GARCH (left) and EGARCH (right) datasets



(b) V-transforms in GARCH (left) and EGARCH (right) datasets



(c) Kpacf plots in GARCH (left) and EGARCH (right) datasets

Figure 6.6: The residuals plot (top), v-transform plots (middle) and kpacf plots (bottom) of the estimating simulations from GARCH (left) and EGARCH (right) processes by parametric vt-S-vine process with three-parameter v-transform and survival Clayton copula sequences.

	GARCH (t)	GJR-GARCH (t)	EGARCH (hyp)
Gumbel	125.82	168.31	220.69
Gaussian	128.39	177.65	221.22
Joe	126.04	168.42	221.41
*Clayton*180	124.81	166.90	216.10
Frank	131.21	191.04	227.32
$Gu(10)\_Ga(30)$	126.42	169.41	221.25
$SC(10)\_Ga(30)$	125.27	168.79	217.56
$t(1)\_sclay(39)$	131.46	194.12	229.81
$sBB1(1)\_sclay(39)$	132.02	195.99	230.16

Table 6.5: The quantile scores of estimating the simulations from GARCH, GJR-GARCH and EGARCH processes by marginal distributions student t, student t and hyperbolic distributions, separately.The quantiles in this table is 0.01. The numbers in yellow are the quantile scores of the best model in each dataset.

	GARCH (t)	GJR-GARCH (t)	EGARCH (hyp)
Gumbel	132.99	178.20	233.67
Gaussian	137.32	186.54	238.74
Joe	132.94	178.21	234.10
*Clayton*180	132.47	175.50	228.83
Frank	139.54	196.20	243.78
$Gu(10)\_Ga(30)$	133.03	177.58	232.64
$SC(10)\_Ga(30)$	131.20	175.55	228.61
$t(1)\_sclay(39)$	136.34	192.22	239.45
$sBB1(1)\_sclay(39)$	138.88	195.34	243.92

Table 6.6: The quantile scores of estimating the simulations from GARCH, GJR-GARCH and EGARCH processes by marginal distributions student t, student t and hyperbolic distributions, separately.The quantiles in this table is 0.99. The numbers in yellow are the quantile scores of the best model in each dataset.

Extreme quantiles estimation of parametric vt-S-vine models

The best fitting model should not only have relatively small AIC values, but also have a accurate match to extreme quantiles. Nikoloulopoulos et al. [2012] propose to use the log-likelihood ratio test to test the quantile violation rates. We use the quantile score methodology instead. In order to compare different replication results of vt-S-vine models, we calculate in-sample quantile scores at quantile 0.01 and 0.99 of the vt-S-vine models. Similar to out-of-sample quantile score, the vt-S-vine models are fit to the simulations from GARCH-type processes and the parameters are estimated. The vt-S-vine models with estimated parameters are used to compute next quantiles at 0.01 or 0.99. Then, the formula 3.10 and 3.11 are applied to calculate quantile scores. The quantiles scores of estimation at quantile 0.01 and 0.99 are shown in Table 6.5 and Table 6.6, respectively.

According to Table 6.5, the survival Clayton has the smallest quantile scores in the three datasets at quantile 0.01. Meanwhile, the vt-S-vine models with survival Clayton copula sequences perform very well in extreme quantiles (0.99) estimations as well. The Gaussian copula with the fist ten levels replaced by survival Clayton copulas ($SC(10\_Ga(30))$ outperform the other models in GARCH and EGARCH processes. These results correspond to the observations of AIC values. The best models are the best models for both methods.

6.3 Predict GARCH by vt-S-vine models

After fitting vt-S-vine processes to the simulated data from GARCH type models, we perform one-step prediction using the estimated vt-S-vine models. There are two types of prediction, semi-parametric and parametric predictions. The vt-S-vine process consists of v-transform and S-vine copulas. In order to simplify the model and reduce the parameters in vt-S-vine models, we use linear v-transforms for all vt-S-vine processes. As discussed in Section 6.2.3, the differences between the three types of v-transforms are not very significant. Moreover, to predict the process, we need to calculate the inverse of the v-transform and the inverse of a linear v-transform is simple to obtain. We carry out a simulation study using the following steps:

1. Simulate 7000 data from GARCH(1,1), GJR-GARCH(1,1) and EGARCH(1,1) processes, separately. The first 5000 data are used to estimate the parameters in vt-S-vine copulas and the remaining 2000 data are applied for one step prediction.

2. Follow the routine of modelling GARCH type time series by vt-S-vine models in Section 6.2.3 and estimate the parameters in vt-S-vine models.

3. Predict using the estimated vt-S-vine process via the approach used to predict S-vine process in Section 3.4. By adding the linear v-transform, the predictions can be transformed back to U_t by inverse linear v-transform.

4. For quantile predictions, we predict at even quantiles $\{0.05, 0.10, 0.15, ..., 0.95\}$ and calculate the quantile scores at these quantiles. The mean of these quantile scores is also calculated to compare prediction results for different vt-S-vine models.

5. The simulated time series and predictions at different quantiles are plotted.

We develop the one-step prediction in three sample size, 500, 1000 and 2000. However, the prediction results are inconsistent with the expectations when sample size are 500 and 1000. The best models for predictions are vt-S-vine with copula sequences that do not fit well when the sample size is 500 or 1000. However, when the sample size increased to 2000, we found that the survival Clayton copulas, which are the best

models in fitting process, usually predict accurately, which is in accordance with our expectations. Hence, we decide to simulate 2000 data for prediction. The results will be presented and discussed in this section.

The prediction procedure for the parametric vt-S-vine is very similar to the semi-parametric one. The only difference is the choice of margins. The margins used for estimation and prediction in parametric vt-S-vine models are the parametric ones as discussed in Section 6.2.3. For semi-parametric vt-S-vine models, we use empirical margins instead. Furthermore, the first-order and second-order S-vine with t copulas are applied to predict the GARCH type processes; details can be found in Zhao et al. [2022]. The results are in the following sections.

6.3.1 Predicting semi-parametric vt-S-vine processes

The simulations from GARCH(1,1), GJR-GARCH(1,1) and EGARCH(1,1) processes are transformed to uniformly distributed time series by empirical distribution. Then, we fit the vt-S-vine copulas with linear v-transform and S-vine process with copula sequences, including Gaussian, Gumbel, Joe, survival Clayton, Frank and the mixture of Gumbel and Gaussian ($Gu(10)\_Ga(30)$) or mixture of survival Clayton and Gaussian copulas ($SC(10)\_Ga(30)$). Predictions of first and second-order of S-vine with t copulas are compared with the higher-order vt-S-vine processes. The quantile score values are presented in Table 6.7. Although quantile scores at $\{0.05, 0.10, 0.15, ..., 0.95\}$ quantiles are calculated, we only present the results at quantiles $0.05, 0.10, 0.25, 0.5, 0.75, 0.9, 0.95$ and the mean of the quantile scores at all quantiles being predicted. Table 6.7 demonstrates the quantile scores of the best two copula sequences in vt-S-vine models and the two S-vine copulas with t copula sequences for each GARCH type time series.

The higher-order semi-parametric vt-S-vine processes exhibit smaller quantile scores at lower and upper tail. The mean of the quantile scores presents the advantages of vt-S-vine in one-step prediction. Among the copula sequences we attempt, the Gumbel, Joe and survival Clayton copulas show the better predictions according to the mean of quantile scores in both the GARCH and the GJR-GARCH datasets. The combination of Gumbel and Gaussian copula sequences show the best results in the EGARCH dataset. The vt-S-vines with survival Clayton copula sequences, which are the best models in the fitting process for all three datasets, are usually one of the top two models in prediction.

It is reasonable that the models fitting GARCH type time series accurately can also predict well. In order to show the advantage of the higher-order semi-parametric vt-S-vine process in replication of GARCH type time series, the time series plot of the predictions for EGARCH process are shown in Figure 6.7. In order to exhibit the time series clearly, we only show the plot of the first 500 predictions.

According to Figure 6.7, the quantile predictions of semi-parametric vt-S-vine processes describe the GARCH type time series better. The predicted lines from the vt-

Quantile scores of GARCH simulations–Semi-parametric								
Quantiles	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
$S-vine(1)-t$	47.411	89.856	186.969	249.709	187.371	89.498	47.198	174.719
$S-vine(2)-t$	47.150	89.432	187.443	249.972	186.986	89.458	47.021	174.738
$Gumbel$	46.423	87.874	185.728	249.868	187.030	88.772	46.765	174.093
$Clayton180$	46.327	87.799	185.585	249.847	187.233	88.824	46.751	174.112
Quantile scores of GJR-GARCH simulations–Semi-parametric								
$S-vine(1)-t$	46.875	88.809	185.838	249.784	186.083	88.487	46.704	173.978
$S-vine(2)-t$	46.227	87.716	185.502	250.029	184.870	87.967	46.238	173.498
Joe	45.017	85.330	182.826	249.857	184.157	86.254	45.460	172.190
$Clayton180$	44.863	85.122	182.477	249.833	184.271	86.275	45.376	172.154
Quantile scores of EGARCH simulations–Semi-parametric								
$S-vine(1)-t$	47.279	89.547	186.804	249.742	186.755	89.088	47.012	174.488
$S-vine(2)-t$	46.888	88.974	187.197	250.091	186.126	88.815	46.781	174.431
$Gu(10)\_Ga(30)$	45.954	87.117	184.800	249.859	185.726	87.599	46.251	173.387
$Gumbel$	45.939	87.072	184.902	249.856	185.715	87.601	46.271	173.406
$Clayton180$	45.737	86.827	185.015	250.377	185.780	87.456	46.093	173.448

Table 6.7: Chosen quantile score values and the means of one-step prediction at evenly distributed quantiles for semi-parametric first-order and second-order S-vine with t copulas and vt-S-vine process with the best two copula sequences in GARCH, GJR-GARCH and EGARCH processes. $S-vine(1)-t$ and $S-vine(2)-t$ are the first-order and second-order S-vine processes with t copulas, respectively. The v-transform used in this table is linear v-transform. The blue colour highlights the best copula sequences in fitting process.

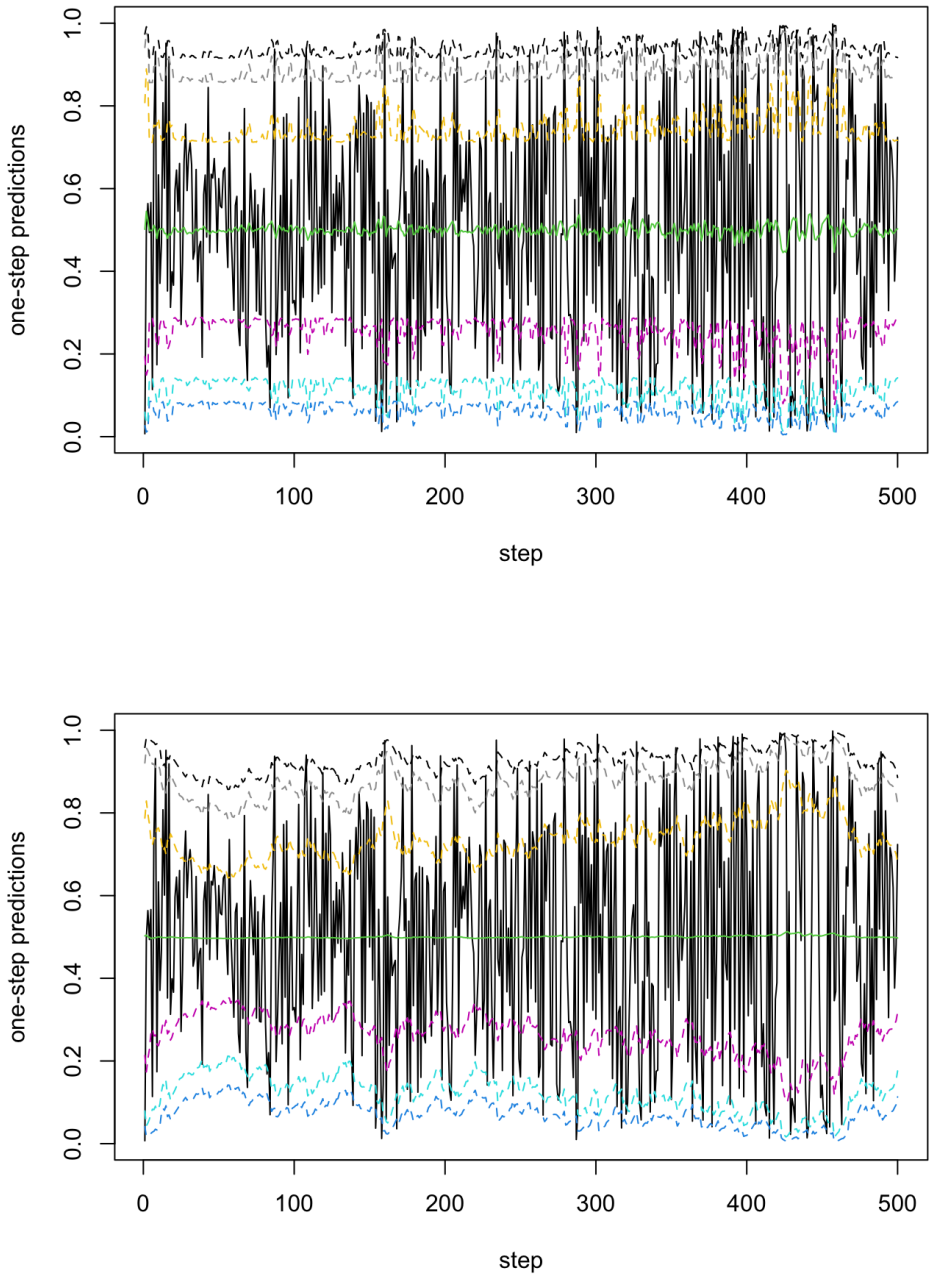


Figure 6.7: The one-step prediction of simulations from EGARCH(1,1) process by semi-parametric second-order S-vine (top) and fortieth-order vt-S-vine with linear v-transform and $Gu(10)\_Ga(30)$ copula sequences (bottom). The black line is the simulations from EGARCH(1,1) process, the 0.05,0.10,0.25,0.5,0.75,0.9,0.9-quantiles are plot in blue, shallow blue, pink, green, yellow, grey lines and black dash lines, separately.

S-vine follow the change of EGARCH processes closely. However, the second-order S-vines with t copulas do not reflect the fluctuation well, where the quantile predictions are almost a straight line with very small volatility. The second-order S-vine process does not capture the trend of EGARCH processes properly. However, the higher-order vt-S-vine with $Gu(10)\_Ga(30)$ copula sequences can realise that, which reveals that the v-transforms are effective at improving the modelling of GARCH type processes.

The advantages of vt-S-vine are more pronounced in parametric vt-S-vine models. We show the results in the next section.

6.3.2 Predicting parametric vt-S-vine processes

For the parametric one step predictions, the margins applied are student t distribution for GARCH and GJR-GARCH processes and hyperbolic distribution for EGARCH processes. The quantile scores are presented in Table 6.8. We only show the results of the best two vt-S-vine models for each dataset. Also, the quantile scores of prediction by first and second-order S-vine models are presented in Table 6.8. The quantiles presented are in accordance with Table 6.7.

The best copula sequences for parametric vt-S-vine models are $SC(10)\_Ga(30)$, Gumbel and survival Clayton copulas in GARCH and GJR-GARCH processes predictions. In EGARCH process, the vt-S-vine with Gaussian and survival Clayton copulas are the best two models for prediction. The vt-S-vines with Clayton copula sequences usually model the in-sample data quite well, so it is reasonable that it also predict well. However, the Gaussian copula sequences perform poorly in the fitting process, but it predicts even better than survival Clayton copula sequences. This may be caused by errors in the prediction process or the 2000 data are not enough for distinguishing between good and bad predictive models.

The advantage for parametric higher-order vt-S-vine models is remarkable compared to first or second-order S-vine processes with t copulas. Similar to the conclusion in semi-parametric predictions, the quantile predictions are improved significantly for the tails of each dataset, especially for GJR-GARCH and EGARCH datasets. The time series plot of predictions at selected quantiles are shown in Figure 6.8. The quantile predictions are predicted from second-order S-vine model with t copula and fortieth-order vt-S-vine with linear v-transform and Gaussian copula sequences, where the margins are hyperbolic margins in Figure 6.8.

According to Figure 6.8, the forecasting time series from parametric vt-S-vine model capture the trend and changes of the EGARCH process accurately and show great improvement compared to the S-vine with t copulas. This observation is in concordance with semi-parametric predictions. The models without v-transforms can not reflect the fluctuations in predicted data. Another finding in both semi-parametric and parametric vt-S-vine models is that the quantile scores calculated from vt-S-vine processes with

Quantile scores of GARCH simulations–Parametric								
Quantiles	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
$S-vine(1)-t$	214.901	355.143	622.562	762.138	606.550	333.195	199.686	570.939
$S-vine(2)-t$	211.032	351.468	623.889	762.700	605.820	332.179	196.799	570.535
$SC(10)\_Ga(30)$	202.782	342.070	617.709	762.431	606.406	328.506	193.337	567.560
$Clayton180$	201.919	341.136	617.611	762.420	606.641	328.984	193.907	567.516
Quantile scores of GJR-GARCH simulations–Parametric								
$S-vine(1)-t$	342.213	535.285	887.013	1061.921	851.976	486.496	301.763	810.369
$S-vine(2)-t$	313.806	511.015	881.443	1062.388	845.684	473.013	283.997	802.618
$Gumbel$	284.815	479.305	864.681	1062.149	842.457	458.069	271.544	791.790
$Clayton180$	284.698	480.433	863.986	1064.583	841.868	456.561	267.422	791.607
Quantile scores of EGARCH simulations–Parametric								
$S-vine(1)-t$	380.417	624.272	1082.716	1319.469	1054.608	586.352	354.592	993.304
$S-vine(2)-t$	370.991	616.575	1085.623	1320.782	1052.583	583.210	348.511	992.150
$Gaussian$	355.074	596.472	1068.272	1319.920	1049.901	572.480	340.501	983.813
$Clayton180$	351.206	593.322	1071.922	1322.104	1050.641	570.558	338.349	984.089

Table 6.8: Chosen quantile score values and the means of one-step prediction at evenly distributed quantiles for parametric first-order and second-order S-vine with t copulas and vt-S-vine process with the best two copula sequences in GARCH, GJR-GARCH and EGARCH processes. $S-vine(1)-t$ and $S-vine(2)-t$ are the first-order and second-order S-vine processes with t copulas, respectively. The v-transform used in this table is the linear v-transform. The blue colour highlights the best copula sequences in fitting process.

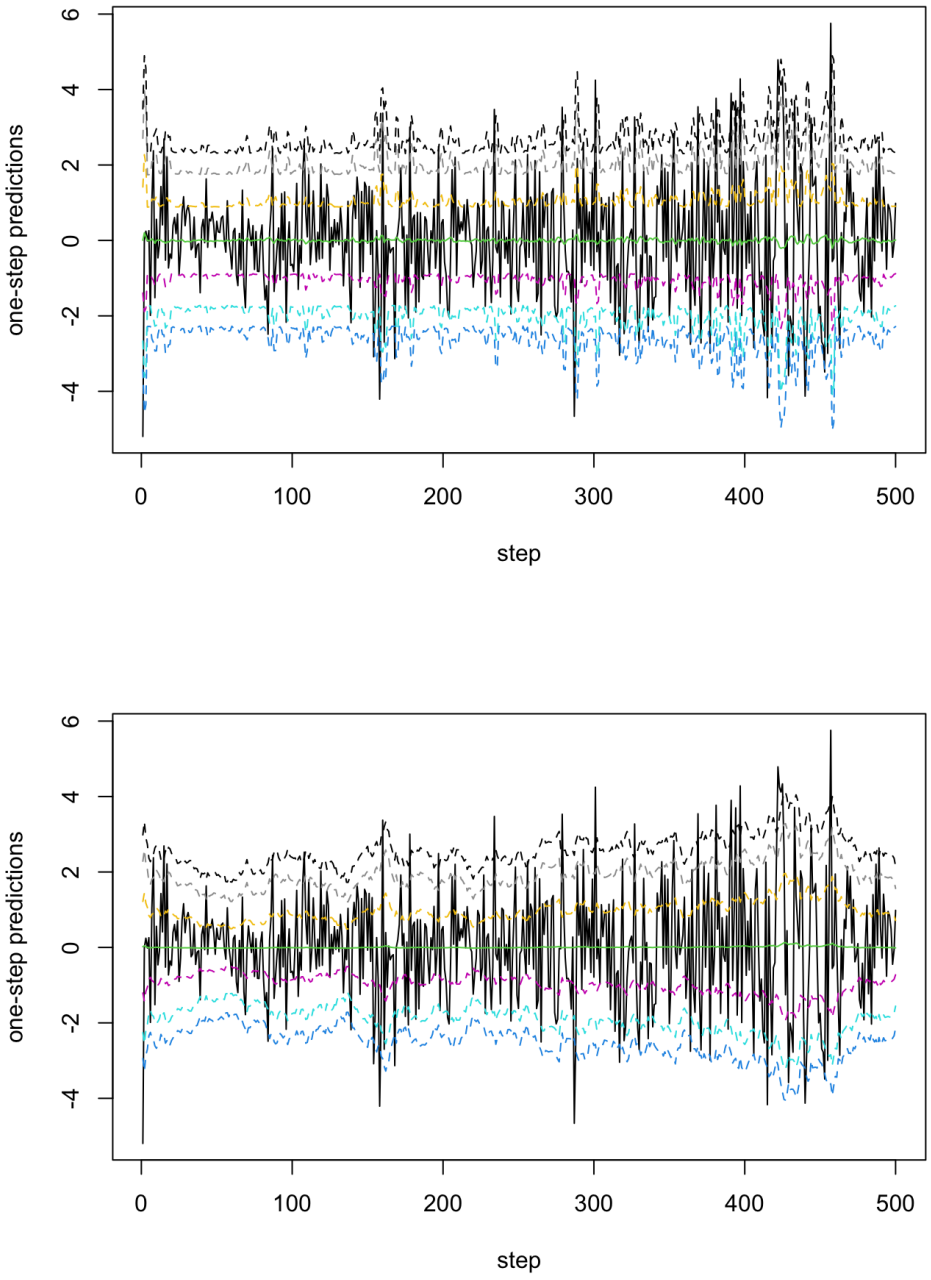


Figure 6.8: The one-step prediction of simulations from EGARCH(1,1) process by parametric second-order S-vine (top) and fortieth-order vt-S-vine with linear v-transform and Gaussian copula sequences (bottom). The black line is the simulations from EGARCH(1,1) process, the 0.05,0.10,0.25,0.5,0.75,0.9,0.9-quantiles are plot in blue, shallow blue, pink, green, yellow, grey lines and black dash lines, separately.

different types of copula sequences are not highly variable, so the choice of copulas may not have great influence on the results.

Chapter 7

Modelling Volatile Return Series with Vt-S-vine Models

This chapter investigates the application of vt-S-vine processes to the kinds of time series that are usually modelled well by GARCH processes. The results are compared to the ones from GARCH processes. The empirical data are introduced in Section 7.1. Section 7.2 elaborates the method used to model the chosen time series by vt-S-vines and presents the estimation results. The methodology and results of one-step predictions using vt-S-vine models are shown in Section 7.3.

7.1 Volatile data

Modelling and predicting volatility of financial data has been an important topic of many empirical and theoretical studies. Volatility, as measured by the standard deviation or variance of returns, is often used as a crude measure of the total risk of financial assets ([Brooks, 2014, p.420]). The volatility of financial time series appears to varying over time. It is often formally modelled by the processes with the changing standard deviation of financial returns given historical information in practice.

In GARCH type processes, the volatility is a conditional standard deviation, which is a time-varying function of the previous squared values of the processes. Hence, GARCH models can describe the changing of volatility effectively. Furthermore, there is a phenomenon known as volatility clustering in the financial time series. Volatility clustering is the tendency for extreme returns to be followed by other extreme returns. GARCH type processes can match the volatility clustering features of financial time series very well. Therefore, GARCH processes are widely used in modelling the volatile financial time series, such as returns on stocks, stock market index, oil or volatility index. Besides, other stylised facts of financial returns can be captured accurately via the variation of GARCH models, such as the leverage effect modelled by EGARCH.

Therefore, leverage effects reflect the asymmetric effect of the market information on volatility, whereby bad news leading to a fall in financial returns tends to increase the volatility ([McNeil et al., 2015, p.122]).

In this section, we investigate whether vt-S-vines are realistic competitors to GARCH and related models. In order to realise this objective, we use the daily price of Bitcoin in USD and the daily price of Apple Inc. stock as two examples to show the fitting and predicting results from GARCH type processes and vt-S-vine models. The time period of daily prices of Bitcoin is from 1st January 2016 to 2nd September 2022 and time period of daily prices of Apple stock is from 1st January 2016 to 1st September 2022. Figure 7.1 presents the prices of Bitcoin and their log-returns values, which are scaled by multiplying 100 in percentage scale. The prices of Apple stocks and scaled log-returns are shown in Figure 7.2. The log-returns of Apple stock is scaled by multiplying 100 to make the data in percentage scale. We take log returns to make the data stationary and be in accordance with the method other studies used to deal with data before modelling them using GARCH type processes.

In Figure 7.1 and Figure 7.2, there are volatility clustering stylized facts existing in the log-returns of the two datasets. For example, we can see periods in early 2020 when the pandemic happened: they are marked by large negative moves in both Bitcoin and Apple stock returns, according to Figure 7.1 and Figure 7.2. Meanwhile, there are some extreme values appearing in 2020 in both of the two log-returns datasets, which may cause the tail risk in the fitting process. In order to show the dependence of the financial returns, we present the acf of the two datasets and the acf of their absolute values in Figure 7.3.

In Figure 7.3, the correlograms of the raw data and their absolute values for the two data sets are presented. There is very little evidence of serial correlation in the raw data for both data sets, while their absolute values appear to show evidence of serial dependence. The serial dependence in the absolute values of returns is equally apparent to the one in squared returns. Hence, the acf plots in both data sets can be seen as an evidence of the presence of volatility clustering.

7.2 In-sample estimation of vt-S-vines

7.2.1 Estimation of parametric vt-S-vines

The Bitcoin returns and stock returns are widely modelled and predicted by GARCH type processes. Katsiampa [2017] applied several GARCH models to explain the Bitcoin price volatility. Lamoureux and Lastrapes [1990] investigated twenty traded stock returns modelled by GARCH and ARCH processes. There is a large amount of studies about using variations of GARCH processes to model and predict the financial return series. For example, Harvey and Sucarrat [2014] investigate financial returns, includ-

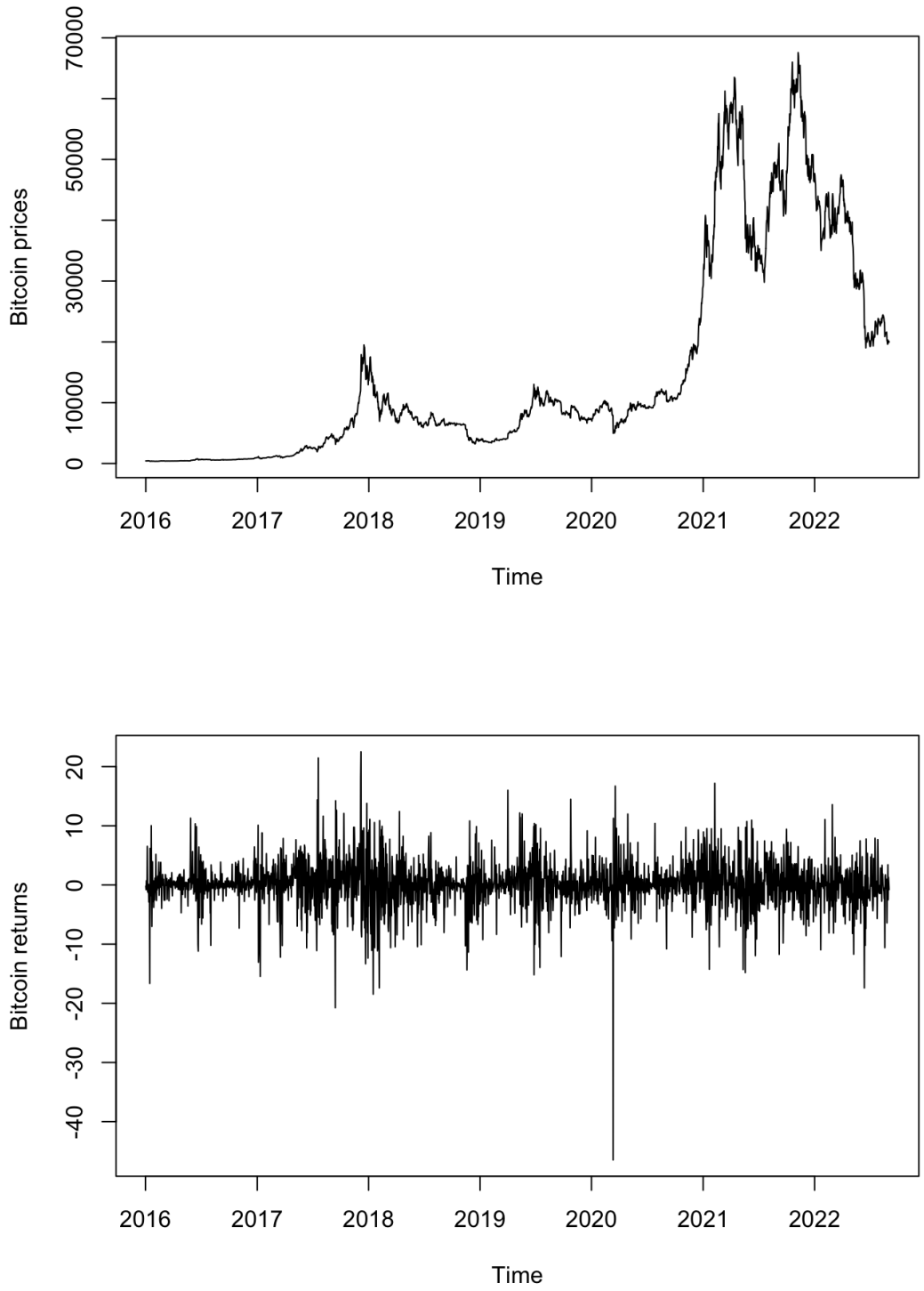


Figure 7.1: The time series plot of prices of Bitcoin (top) and scaled log-returns on Bitcoin prices (bottom).

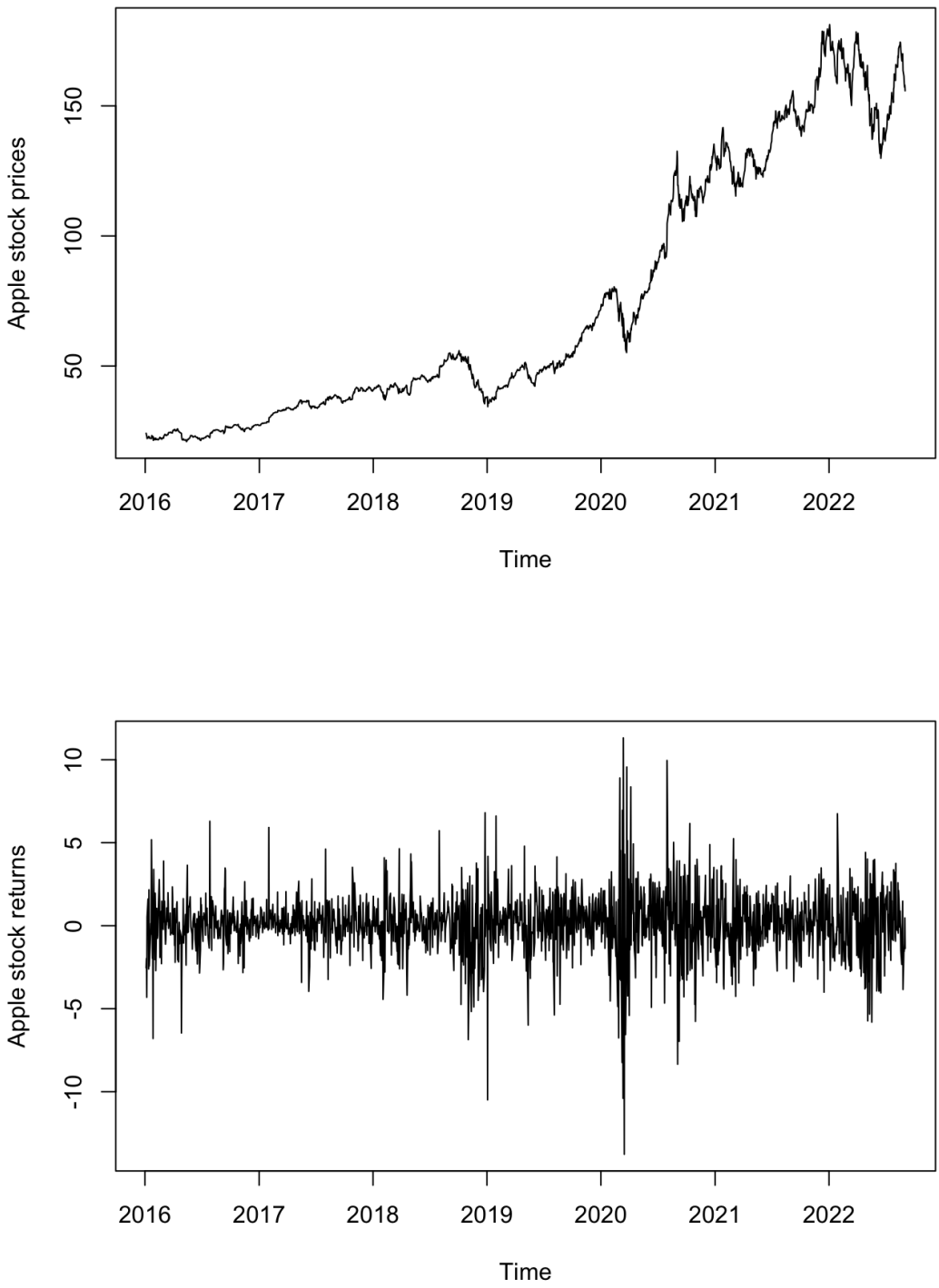
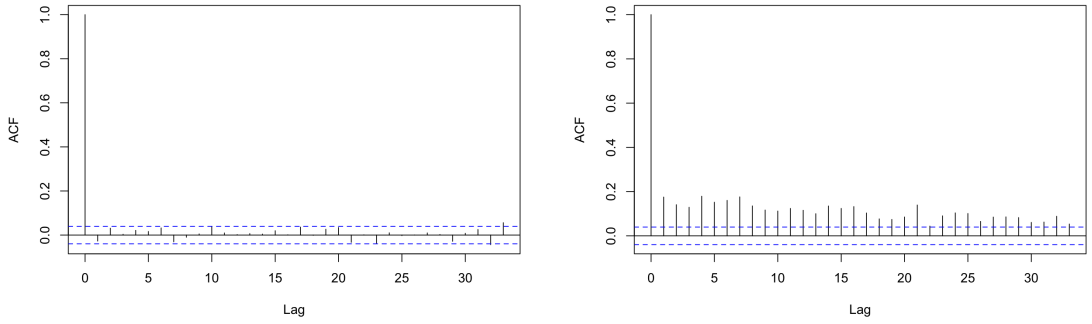
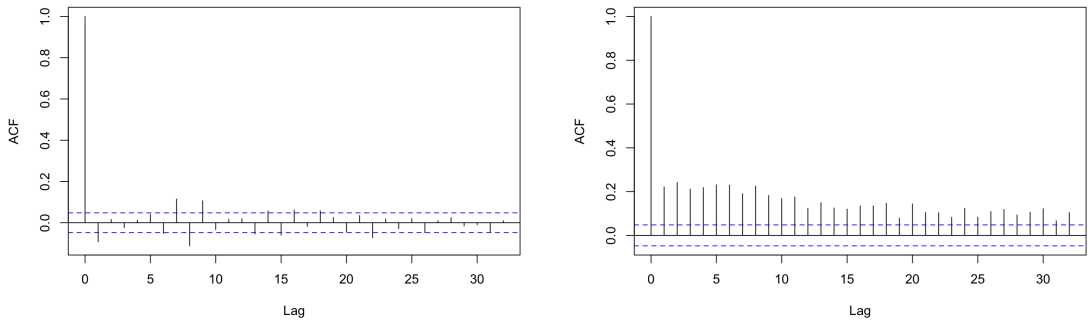


Figure 7.2: The time series plot of prices of Apple stock (top) and scaled log-returns on Apple stock prices (bottom).



(a) Bitcoin returns



(b) Apple stock returns

Figure 7.3: The acf of Bitcoin returns (top-left) and the acf of absolute values of Bitcoin returns (top-right). The acf of Apple stock returns (bottom-left) and the acf of absolute values of Apple stock returns (bottom-right). Dotted lines mark the standard 95 % confidence intervals for autocorrelations of a process of iid finite-variance random variables.

	P-values	
Data	kpss test	arch test
Bitcoin	0.1	6.361e-05
Apple	0.1	2.2e-16

Table 7.1: The p-values of the kpss test and arch test in Bitcoin returns and Apple stock returns. The null hypothesis of kpss test is that the time series is stationary. The null hypothesis of arch test is that the time series is homoscedastic.

AIC values of GARCH type processes fitted to log-returns of Bitcoin			
	GARCH(1,1)	EGARCH(1,1)	GJR-GARCH(1,1)
norm	9233.45	9224.76	9223.48
t	8649.80	8605.24	8649.96
st	8641.91	8593.02	8640.42
NIG	8618.85	8593.13	8618.95

Table 7.2: The AIC values of estimating the log-returns of Bitcoin from GARCH, GJR-GARCH and EGARCH processes with normal (norm), student t (t), skewed student t (st) and normal-inverse Gaussian (NIG) distribution. The AIC value in yellow colour represents the best process for describing log-returns of Bitcoin.

ing the Apple stock returns modelled by GARCH, EGARCH, GJR-GARCH and their variations and compare the results with standard GARCH.

In this section, we separate each dataset into two parts, one as the training data, the other as the testing data. For Bitcoin price, we used the first 1707 as the training data and the recent two years data, which are 707 as testing data. For the Apple stock price, the first 1175 data are used in fitting process and the last two years data, 504, are applied as an input for one-step prediction. The data amount of Bitcoin is larger than the one of Apple stock, so we decide to leave more data for prediction.

The candidate GARCH type processes we applied are GARCH(1,1), EGARCH(1,1) and GJR-GARCH(1,1) with normal (norm), student t (t), skewed student t (st) and normal-inverse Gaussian (NIG) distribution as the innovation distributions. Before we fit the candidate models to Bitcoin and Apple stock returns, it is necessary to test stationarity and heteroscedasticity of the two time series via kpss test and arch test. The p-values of the two tests are shown in Table 7.1.

According to Table 7.1, the null hypothesis is accepted in the kpss test and rejected in the arch test for both data sets. Hence, the two data sets are highly possible stationary and heteroscedastic, which means they can be modelled by GARCH type processes.

Then, we apply the two financial return series to estimate the parameters of candidate models. The AIC values of fitting GARCH type processes to log-returns of Bitcoin are presented in Table 7.2. The ones of Apple stocks are shown in Table 7.3.

Table 7.2 reveals that the EGARCH(1,1) with skewed student t innovation distribution is the best model describing the log-returns of Bitcoin. Similarly, the best process

AIC values of GARCH type processes fitted to log-returns of Apple stock			
	GARCH(1,1)	EGARCH(1,1)	GJR-GARCH(1,1)
norm	4440.66	4387.49	4387.37
t	4273.26	4241.80	4246.11
st	4267.52	4237.88	4241.29
NIG	4265.02	4237.61	4239.31

Table 7.3: The AIC values of estimating the log-returns of Apple stock from GARCH, GJR-GARCH and EGARCH processes with normal (norm), student t (t), skewed student t (st) and normal-inverse Gaussian (NIG) distribution. The AIC value in yellow colour represents the best process for describing log-returns of Apple stock.

	Bitcoin (EGARCH(1,1)-st)	Apple(EGARCH(1,1)-NIG)
α_0	0.0273 (0.0092)	0.0612 (0.0175)
α_1	0.0590 (0.0316)	-0.1432 (0.0304)
β_1	0.9941 (0.0000)	0.9518 (0.0145)
γ	0.4085 (0.1293)	0.2385 (0.0430)
shape	2.2372 (0.0182)	0.1584 (0.0500)
skew	0.9285 (0.1606)	0.0500 (0.1584)

Table 7.4: The estimated parameters in the best model (in bracket) in each data set. EGARCH(1,1)-st denotes the EGARCH(1,1) model with skewed t innovation. EGARCH(1,1)-NIG denotes the EGARCH(1,1) model with NIG innovation. The numbers in the bracket is are the standard errors for each parameters estimation.

for modelling log-returns of Apple stock is EGARCH(1,1) with normal-inverse Gaussian distribution, according to Table 7.3. The estimated parameters of the best models in the two data sets are shown in Table 7.4. In order to exhibit the estimation results clearly, we draw the QQ-plot of standardized residuals of the best models in each data set in Figure 7.4. Besides, the acf of the standardized residuals and acf of their absolute values are presented in Figure 7.5.

The estimated β_1 of EGARCH(1,1) is close to 1, which is the boundary of condition for strictly stationary EGARCH(1,1) processes, in both Bitcoin and Apple stock returns, according to Table 7.4. The shape parameter of Bitcoin is the degree of freedom in t distribution. The degree of freedom is small, which reflects the heavy tail of the Bitcion returns. Similarly, the parameters of NIG distribution also exhibit the heavy tails existing in Apple stock returns.

According to Figure 7.5, the QQ-plot of in-sample estimation of EGARCH(1,1) with skewed t innovation distribution seems not quite well in Bitcoin returns. The standardized residuals of Apple stock returns is almost consistent with the NIG innovation, which reflects the EGARCH(1,1) with NIG innovation distribution process describes the Apple stock returns relatively accurately. There is very little evidence in Figure 7.5 to show the serial dependence in the acf of the standard residuals and acf of their

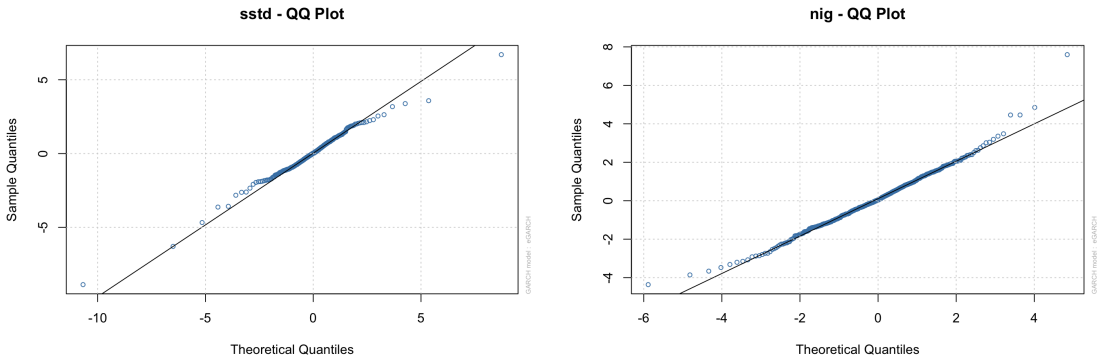
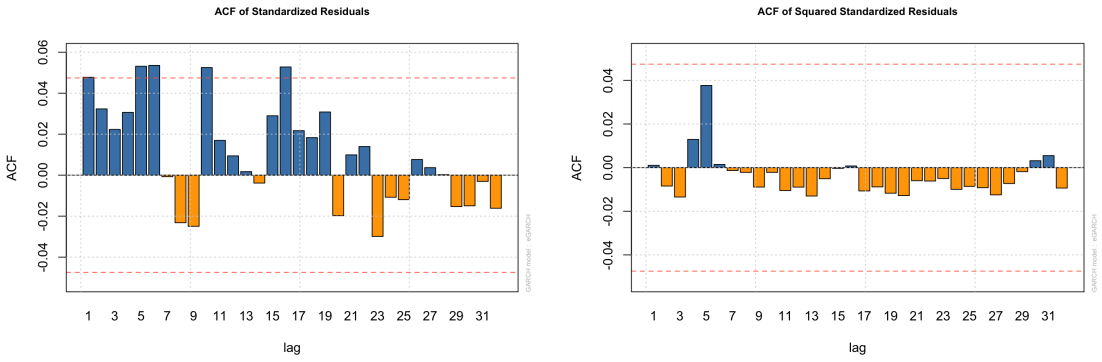
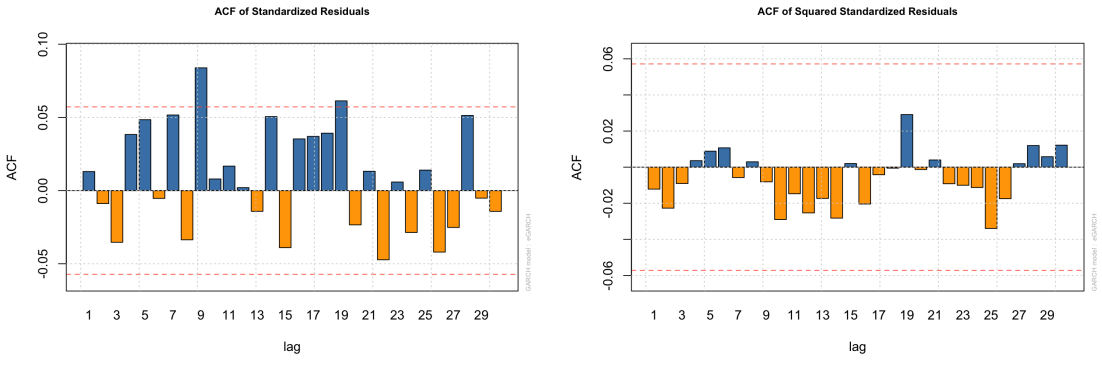


Figure 7.4: The QQ-plots of standardized residuals of Bitcoin returns modelled by EGARCH(1,1) with skewed t innovation distribution (left) and the QQ-plots of standardized residuals of Apple stock returns modelled by EGARCH(1,1) with normal-inverse Gaussian innovation distribution (right).



(a) Bitcoin returns



(b) Apple stock returns

Figure 7.5: The acf of standardized residuals of Bitoin returns modelled by EGARCH(1,1) with skewed t innovation distribution (left-top) and the acf of their absolute values (top-right). The acf of standardized residuals of Apple stock returns modelled by EGARCH(1,1) with normal-inverse Gaussian innovation distribution (bottom-left) and the acf of their absolute values (bottom-right). Dotted lines mark the standard 95 % confidence intervals for autocorrelations of a process of iid finite-variance random variables.

AIC values of candidate margins fitted to Bitcoin and Apple stock		
Margins	Bitcoin	Apple
normal	9593.90	4784.93
slap	8966.83	4439.24
dwe	8890.07	4434.83
sdwe	8891.95	4433.92
t	8976.76	4413.21
st	8978.76	4415.20
NIG	9078.57	4411.98

Table 7.5: The AIC values of candidate margins fitted to log-returns of Bitcoin and Apple stock. The AIC value in yellow colour represent the best margins for each data set. *slap*, *dwe*, *sdwe*, *t*, *st* and *NIG* denote skewed Laplace, double Weibull, skewed double Weibull, student t, skewed student t and normal-inverse Gaussian distributions, respectively.

absolute values. Hence, the standardized residuals in the two data sets conform to the condition for a reasonable model.

In order to compare the results with GARCH processes, we model the two data sets by parametric vt-S-vines and calculate the AIC values as well. The first step is to select the best margins. The margin we choose for Bitcoin is double Weibull distribution (dwe). Meanwhile, Normal-inverse Gaussian distribution (NIG) is the most suitable margin for log-returns of Apple stock. The AIC values of the potential margins are shown in Table 7.5.

After finding the best marginal distribution, we combine it with vt-S-vine copulas and estimate all the parameters jointly by maximum likelihood estimation as an entirety. The method used in this section is the same as the one in Section 6.2.3. The v-transform in the vt-S-vine copulas is the linear v-transform, which can simplify the model and reduce the number of parameters. The candidates of copulas are stated in Section 6.2.3 as well. The AIC values of fitting vt-S-vine models to log-returns of Bitcoin and Apple stock are shown in Table 7.6.

Table 7.6 shows that the vt-S-vine with Frank copula sequences is the best model for Bitcoin. Comparing to the AIC values in Table 7.2, the best vt-S-vine model can describe log-returns of Bitcoin better than all the GARCH type processes in the table. However, for the log-returns of Apple stock, the EGARCH(1,1) with t,st,and NIG innovation distributions yield better result than the best vt-S-vine, so as the GJR-GARCH(1,1) with st and NIG distributions. Moreover, the best vt-S-vine can outperform all the GARCH(1,1) processes, according to Table 7.3.

The findings can be extended to other stock price or index datasets we attempted. In most of the datasets, the GARCH(1,1) can be surpassed by the best vt-S-vine and so as GJR-GARCH(1,1). However, EGARCH(1,1) processes with proper innovation

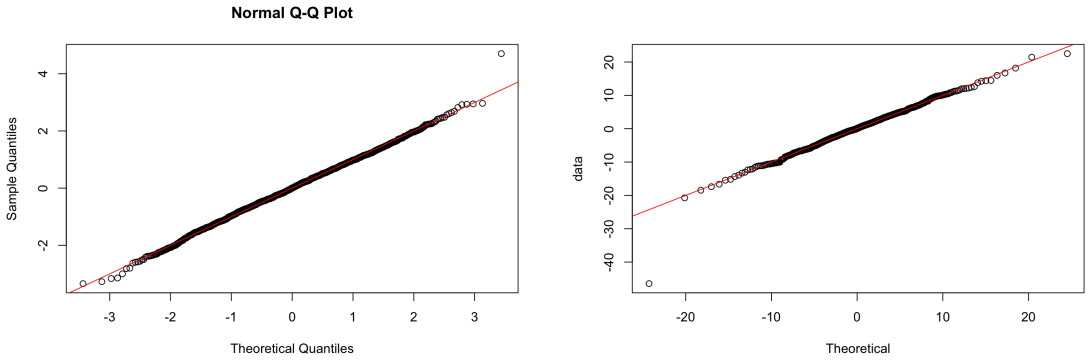
AIC values of VT-S-vine fitted to Bitcoin and Apple stock		
Copulas	Bitcoin (dwe)	Apple (NIG)
$Gumbel$	8625.96	4261.27
$Gaussian$	8589.72	4262.06
Joe	8684.97	4271.29
$Clayton180$	8619.09	4245.19
$Frank$	8579.84	4282.31
$Gu(10)\_Ga(30)$	8610.47	4259.08
$SC(10)\_Ga(30)$	8613.78	4243.48

Table 7.6: The AIC values of estimating the log-returns of Bitcoin and Apple stock by vt-S-vine models with $Gumbel$, $Gaussian$, Joe, $Clayton180$, $Frank$, $Gu(10)\_Ga(30)$ and $SC(10)\_Ga(30)$ copula sequences. The AIC value in yellow colour represent the best models for each dataset. The words in brackets of data sets represent the margins selected for each data set.

distributions are very competitive and hardly be surpassed. That is an open question for our study now. It may be caused by the limited options of margins and copulas. In order to show the estimation results of the two data sets clearly, we present the QQ-plots of residuals, acf of residuals, acf of their absolute values, QQ-plot of margins, v-transforms plots and the kpacf plots of the best vt-S-vine models in Bitcoin and Apple stock returns in Figure 7.6 and Figure 7.7.

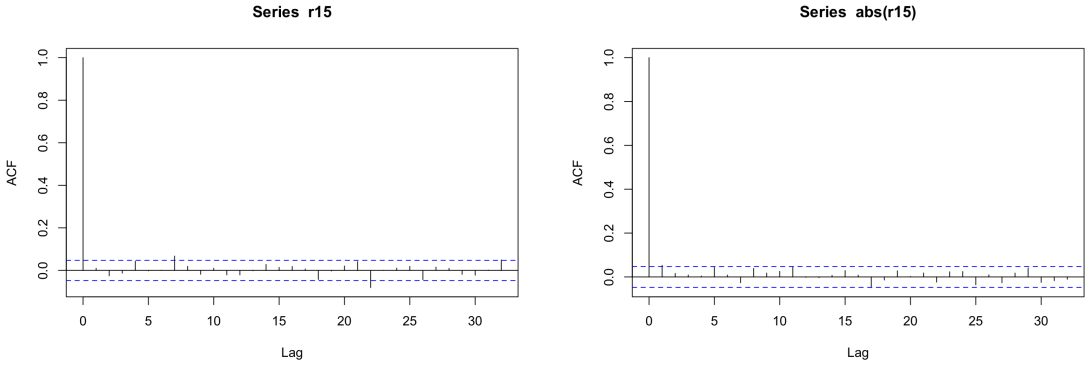
According to Figure 7.6, the QQ-plot of residuals and margins are very accurate. Meanwhile, there is very little evidence against the independence of residuals according to acf plots of residuals and absolute residuals. In addition, we apply the Shapiro-Wilk test to test the normality of the residuals. The p-value of Bitcoin returns is 0.09272. Hence, the null hypothesis is accepted that the residuals of Bitcoin returns are relatively normally distributed. The kpacf plot is plausible in Bitcoin as well, which means the Frank copulas capture the serial dependence in Bitcoin returns adequately. Moreover, the v-transform plots of Bitcoin has an apparent twist against the linear v-transform. It can be improved by using v-transforms with two or three-parameters v-transforms. Generally, the vt-S-vine model presented in Figure 7.6 describes the Bitcoin returns very well.

Figure 7.7 exhibits that the QQ-plots of residuals and margins are reasonable and acceptable as well for Apple stock returns. Similarly, there is little evidence of serial dependence between residuals and absolute residuals. The p-value of the Shapiro-Wilk test is 0.1216. There is little evidence against the null hypothesis of normally distributed. The kpacf and v-transforms plots of Apple stock returns are accurate. However, the AIC values of the vt-S-vine models fitted to Apple stock returns are larger than the one of EGARCH(1,1) with NIG innovation process, which means the in-sample estimation of vt-S-vines can be improved. Superior and advanced margins and copulas can be searched in the further study to realize the object, since the imperfect QQ-plots of



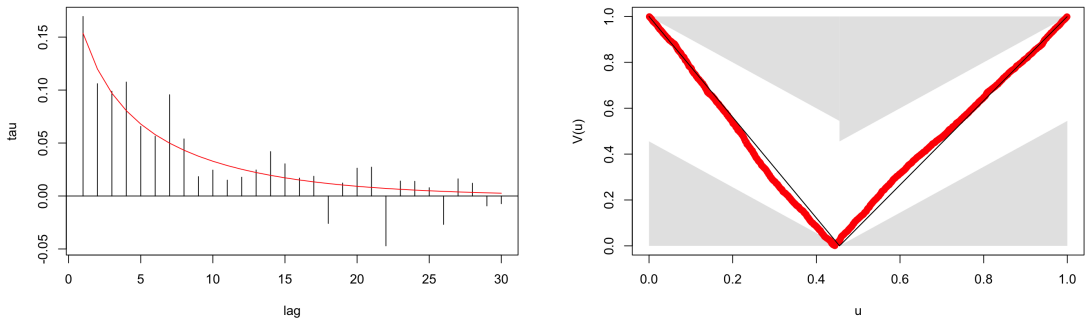
,

(a) QQ-plot of residuals (left) and margins (right) in Bitcoin



,

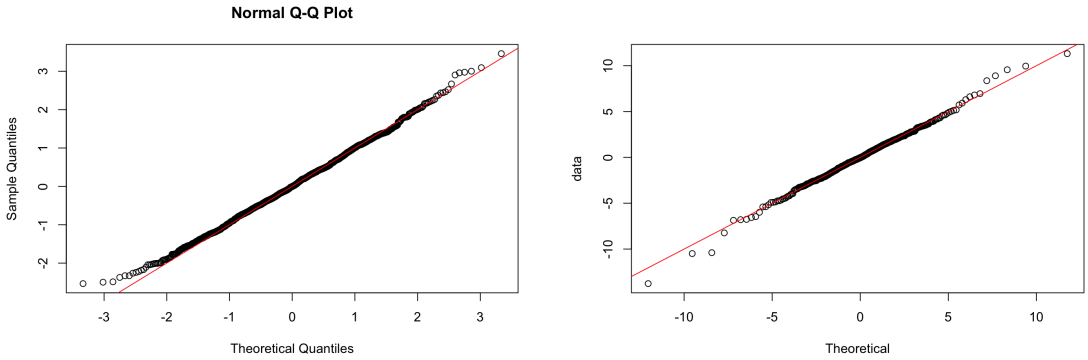
(b) Acf of residuals (left) and acf of their absolute values (right) in Bitcoin



,

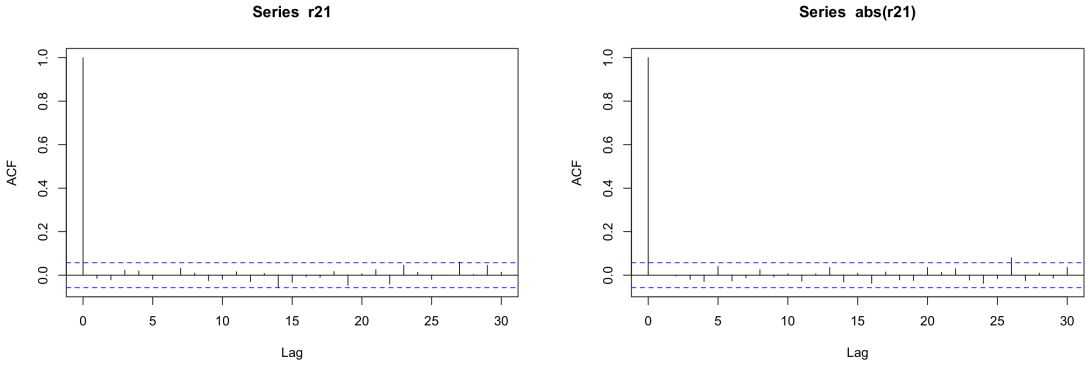
(c) Kpacf of copulas (left) and v-transforms estimations (right) in Bitcoin

Figure 7.6: The QQ-plots of residuals, margins, acf of residuals, acf of their absolute values, kpacf plots and v-transforms plots of the vt-S-vine models with double Weibull margins, linear v-transforms and Frank copula sequences estimated by Bitcoin returns.



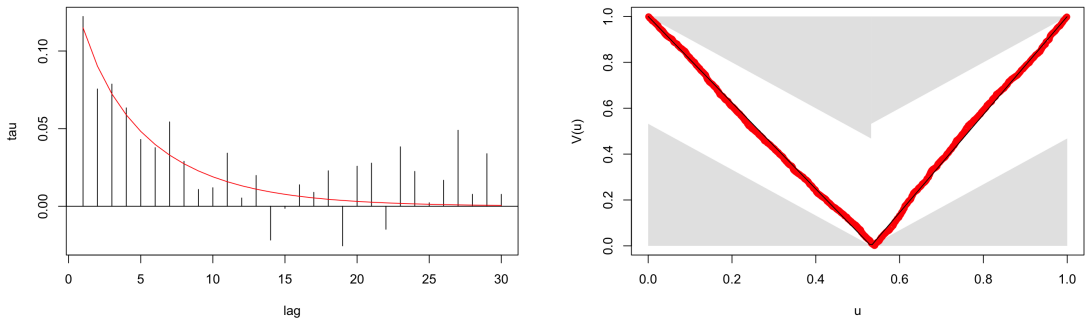
,

(a) QQ-plot of residuals (left) and margins (right) in Apple



,

(b) Acf of residuals (left) and acf of their absolute values (right) in Apple



,

(c) Kpacf of copulas (left) and v-transforms estimations (right) in Apple

Figure 7.7: The QQ-plots of residuals, margins, acf of residuals, acf of their absolute values, kpacf plots and v-transforms plots of the vt-S-vine models with NIG margins, linear v-transforms and $SC(10)\_Ga(30)$ copula sequences estimated by Apple stock returns.

AIC values of VT-S-vine fitted to Bitcoin and Apple stock		
Copulas	Bitcoin (dwe)	Apple (NIG)
$t(1)\_sclay(39)$	8641.73	4301.17
$sBB1(1)\_sclay(39)$	8647.84	4298.38
$t(1)\_Ga(39)$	8577.55	4262.19
$sBB1(1)\_Ga(39)$	8584.49	4259.30

Table 7.7: The AIC values of vt-S-vine models with $t(1)\_sclay(39)$, $sBB1(1)\_sclay(39)$, $t(1)\_Ga(39)$ and $sBB1(1)\_Ga(39)$ copula sequences fitted to the log-returns of Bitcoin and Apple stock.

margins and kpacf plots imply that there are defects existing in the two parts.

In-sample estimation with two-parameter copulas in vt-S-vine

The vine copulas with BB1 and t copulas are applied to financial returns, especially for modelling their tail dependence; referred to by Nikoloulopoulos et al. [2012]. Therefore, we fit the vt-S-vine with linear v-transform and copula sequences including two-parameter copulas at level one, such as t and survival BB1 copula to Bitcoin and Apple stock returns. The combinations include t or BB1 copula at level one and survival Clayton for the other 39 levels denoted as $t(1)\_sclay(39)$ or $sBB1(1)\_sclay(39)$. Besides, to compare with the vt-S-vine model with Gaussian copula at the higher order, the model with t or BB1 copulas at level one and Gaussian copula at higher level are applied, written as $t(1)\_Ga(39)$ or $sBB1(1)\_Ga(39)$. The marginal distributions are double Weibull distribution in Bitcoin returns and NIG distribution in Apple stock returns. The AIC values of the four vt-S-vine models are shown in Table 7.7.

According to Table 7.7, the vt-S-vine model with t copula at level one and Gaussian copula at the other 39 levels surpasses the best model in Bitcoin returns in Table 7.6, which is vt-S-vine with Frank copula sequences. However, the vt-S-vine models with copula sequences including two-parameter copulas cannot outperform the models only with one-parameter copulas in Apple stock returns.

Extreme quantile estimations

In order to compare the estimations of extreme quantiles, we calculate the quantile scores for the models in Table 7.6 and 7.7. The results are demonstrated in Table 7.8 and Table 7.9.

According to Table 7.8 and 7.9, the vt-S-vine with $t(1)\_Ga(39)$ copula sequences obtain the most precise quantile estimation at both 0.01 and 0.99 quantile in Bitcoin returns. This results are corresponding to the observations from AIC values. The vt-S-vine with $SC(10)\_Ga(30)$ sequences is the best model in Apple stock dataset, according to AIC values. The $SC(10)\_Ga(30)$ yields the smallest quantile scores at quantile 0.01

Quantile scores at 0.01 of VT-S-vine fitted to Bitcoin and Apple stock		
Copulas	Bitcoin (dwe)	Apple (NIG)
$Gumbel$	263.66	67.27
$Gaussian$	260.19	66.50
Joe	263.47	68.14
$Clayton180$	258.19	65.57
$Frank$	261.38	73.37
$Gu(10)\_Ga(30)$	259.48	66.25
$SC(10)\_Ga(30)$	256.38	64.98
$t(1)\_sclay(39)$	262.48	70.92
$sBB1(1)\_sclay(39)$	263.19	72.71
$t(1)\_Ga(39)$	256.20	65.74
$sBB1(1)\_Ga(39)$	256.67	67.45

Table 7.8: Quantile scores of estimating the log-returns of Bitcoin and Apple stock by vt-S-vine models with chosen copula sequences at quantile 0.01. The numbers in yellow are the quantile scores of the best model in each dataset.

Quantile scores at 0.99 of VT-S-vine fitted to Bitcoin and Apple stock		
Copulas	Bitcoin (dwe)	Apple (NIG)
$Gumbel$	227.89	68.10
$Gaussian$	227.94	69.58
Joe	227.10	67.98
$Clayton180$	226.42	68.64
$Frank$	230.42	72.49
$Gu(10)\_Ga(30)$	228.16	68.72
$SC(10)\_Ga(30)$	228.47	68.58
$t(1)\_sclay(39)$	226.16	74.21
$sBB1(1)\_sclay(39)$	225.90	73.33
$t(1)\_Ga(39)$	225.03	69.62
$sBB1(1)\_Ga(39)$	225.50	69.93

Table 7.9: Quantile scores of estimating the log-returns of Bitcoin and Apple stock by vt-S-vine models with chosen copula sequences at quantile 0.99. The numbers in yellow are the quantile scores of the best model in each dataset.

AIC values of GARCH type processes fitted to ARMA(0,1) residuals of Bitcoin			
	GARCH(1,1)	EGARCH(1,1)	GJR-GARCH(1,1)
norm	9224.87	9216.14	9219.24
t	8630.79	8584.11	8627.63
st	8632.17	8585.62	8629.08
NIG	8609.74	8585.98	8608.83

Table 7.10: The AIC values of estimating the ARMA(0,1) residuals of Bitcoin from GARCH, GJR-GARCH and EGARCH processes with normal (norm), student t (t), skewed student t (st) and normal-inverse Gaussian (NIG) distribution. The AIC value in yellow colour represents the best process for describing the ARMA(0,1) residuals of Bitcoin.

and a quantile scores very close to the smallest values from Joe copulas sequences. The vt-S-vine with Joe copula sequences does not improve the quantile estimation at 0.99 significantly, compared to $SC(10)\_Ga(30)$.

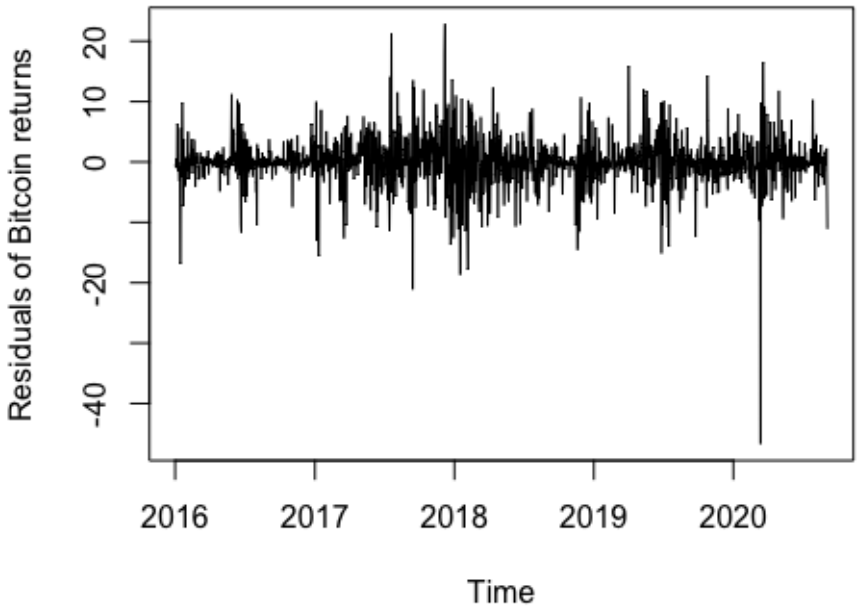
7.2.2 Modelling residuals of ARMA filter

In order to figure out if there is an ARMA influence in modelling the returns, we apply the ARMA filters to remove the ARMA component from the data and fit the GARCH type and vt-S-vine models to the residuals. The automatic function is used to find the appropriate order of ARMA in each dataset. The ARMA(0,1) is chosen for Bitcoin returns and ARMA(3,0) is used for Apple stock returns. After fitting the chosen ARMA models to the two data sets, the residuals are extracted and estimated by the GARCH type models and vt-S-vine models in the next step. The time series plots of the two data sets are shown in Figure 7.8.

The time series plots are not very different to the ones in Figure 7.1 and 7.2. The ARMA filters do not change the data too much. We also plot the acf and acf of the absolute values of the residuals to check if the ARMA filters change the serial dependence in the time series. The plots are presented in Figure 7.9.

Figure 7.9 demonstrates the acf of the two ARMA residuals are not remarkable, while the acf of their absolute values are slowly decreased. The serial dependence in the residuals in the two data sets are not significantly changed by ARMA filters. The next step is to fit the GARCH type models and vt-S-vines to the residuals and compare the results to the ones without ARMA filters. The AIC values of GARCH type models fitted to ARMA residuals of Bitcoin returns are shown in Table 7.10 and the ones for Apple stock returns are demonstrated in Table 7.11.

Table 7.10 and 7.11 exhibit that the best GARCH model is EGARCH(1,1) with student t innovation distribution for the two data sets. Meanwhile, we can draw the same conclusion that the EGARCH(1,1) is always the best model no matter which



,

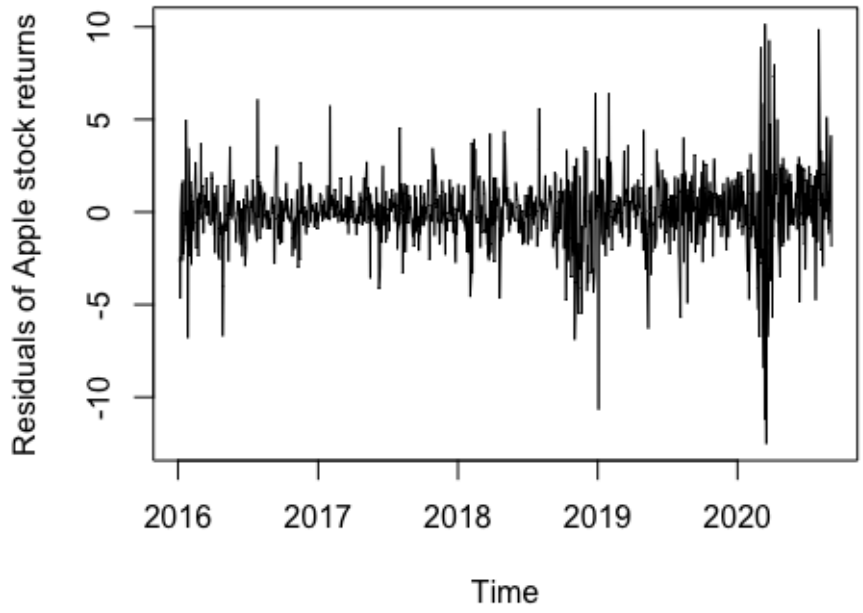
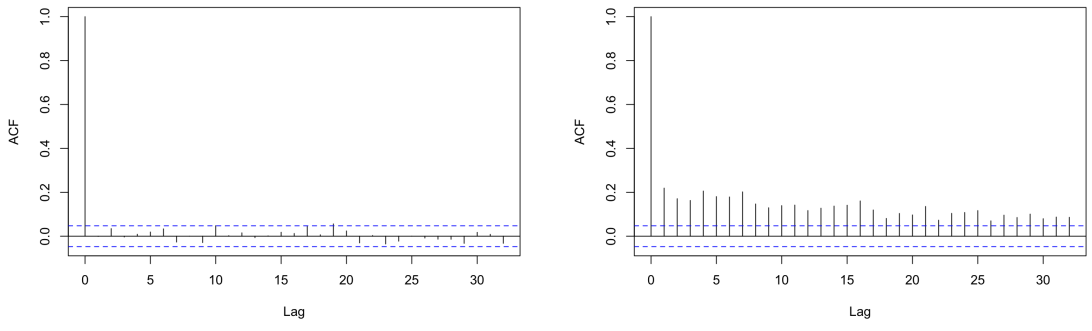
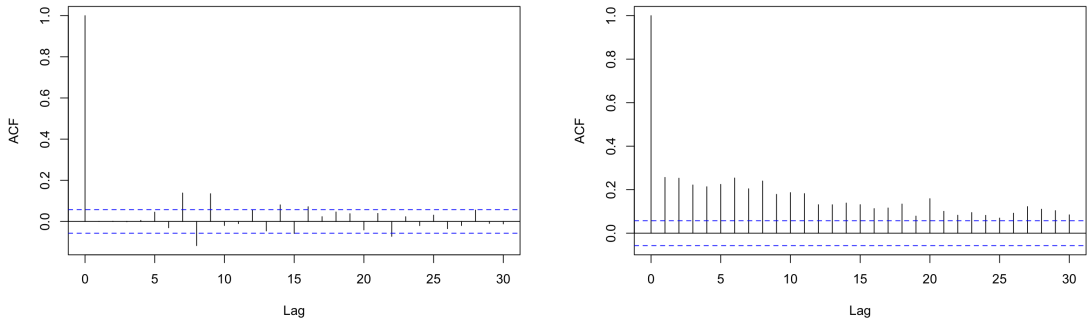


Figure 7.8: The time series plots of ARMA residuals of Bitcoin returns (top) and Apple stock returns (bottom).



(a) Bitcoin returns.



(b) Apple stock returns.

Figure 7.9: The acf of ARMA residuals of Bitcoin returns (left-top) and Apple stock returns (left-bottom). The acf of their absolute values, the one in right-top is Bitcoin returns and the one in right-bottom is the Apple stock returns.

AIC values of GARCH type processes fitted to ARMA(3,0) residuals of Apple stock			
	GARCH(1,1)	EGARCH(1,1)	GJR-GARCH(1,1)
norm	4423.28	4378.49	4381.02
t	4259.01	4233.91	4237.39
st	4260.41	4235.78	4239.13
NIG	4257.51	4234.64	4237.08

Table 7.11: The AIC values of estimating the ARMA(3,0) residuals of Apple stock from GARCH, GJR-GARCH and EGARCH processes with normal (norm), student t (t), skewed student t (st) and normal-inverse Gaussian (NIG) distribution. The AIC value in yellow colour represents the best process for describing the ARMA(3,0) residuals of Apple stock.

	Bitcoin (EGARCH(1,1)-t)	Apple(EGARCH(1,1)-t)
α_0	0.0239 (0.0071)	0.0398 (0.0158)
α_1	0.0611 (0.0273)	-0.1244 (0.0274)
β_1	0.9945 (0.0001)	0.9589 (0.0139)
γ	0.3587 (0.0524)	0.2575 (0.0453)
shape	2.2818 (0.1367)	3.9501 (0.4775)

Table 7.12: The estimated parameters in the best model (in bracket) in each data set. EGARCH(1,1)-t denotes the EGARCH(1,1) model with student t innovation. The numbers in the bracket is are the standard errors for each parameters estimation.

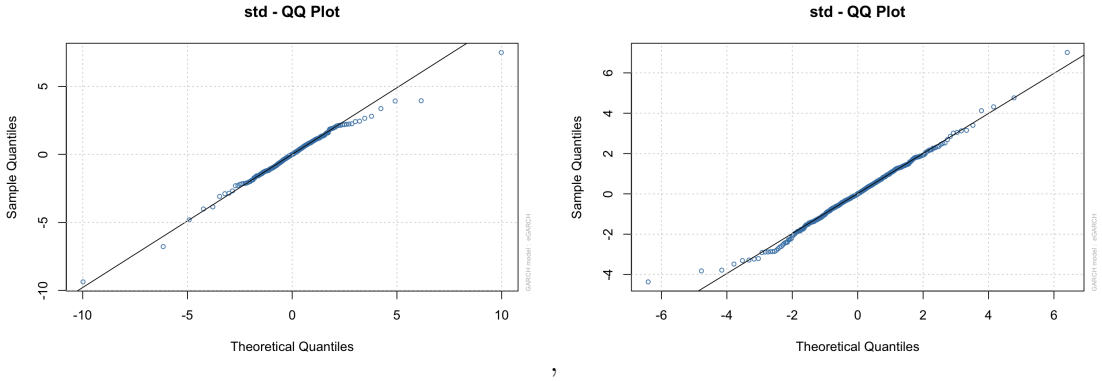


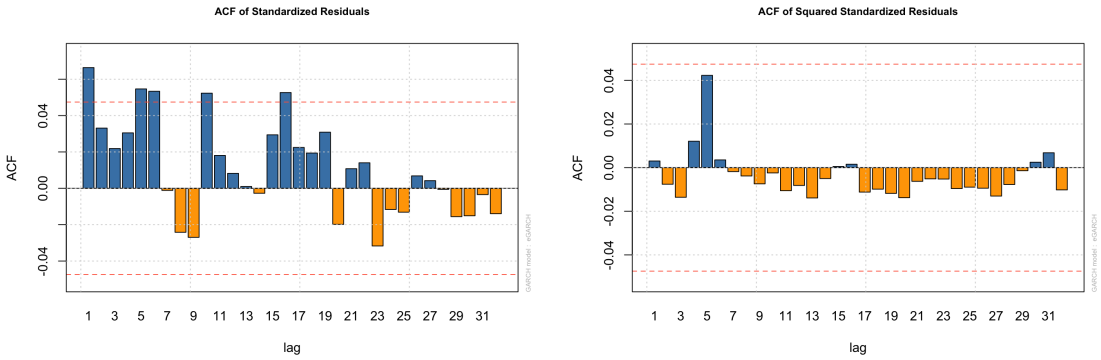
Figure 7.10: The QQ-plots of standardized residuals of Bitcoin returns with ARMA filter modelled by EGARCH(1,1) with student t innovation distribution (left) and the QQ-plots of standardized residuals of Apple stock returns with ARMA filter modelled by EGARCH(1,1) with student t innovation distribution (right).

innovation distribution are used. To present the result further, we show the estimated parameters of EGARCH(1,1) with t innovation in the two data sets in Table 7.12.

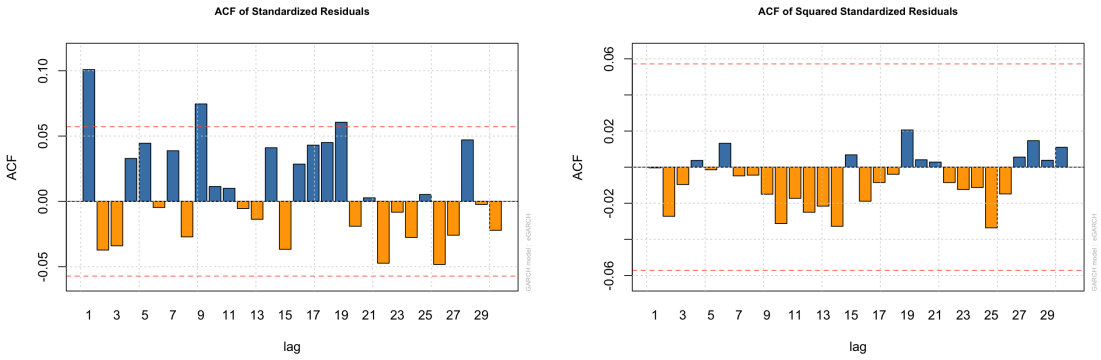
The parameters of EGARCH(1,1) in Table 7.12 are similar to the ones in Table 7.4. The ARMA filters do not exert remarkable influence on the parameters estimation of GARCH part. The QQ-plots of the residuals are presented in Figure 7.10. The QQ-plots of the two data sets with ARMA filter reveals that the tail of the residuals of the two data sets are not estimated very well by student t distribution. Therefore, it is possible to improve the accuracy of estimation by finding the more suitable innovation distribution for GARCH type models. The acf of the residuals are shown in Figure 7.11.

According to Figure 7.11, there are acf values at lag 1, 5, 6, 9 and 15 beyond the standard 95 % confident interval in Bitcoin data set. Similarly, the acf values at lag 1, 9 and 18 go over the dotted line in Apple stock returns. The residuals of the two data sets seem have slight serial dependence, which exhibits that the model can be improved.

In order to compare the estimation results with vt-S-vine models, we fit the vt-S-vines to the Bitcoin and Apple stock returns residuals of ARMA filter. The first step is the marginal distribution selection. We find that the best margins are the same as the



(a) Bitcoin returns



(b) Apple stock returns

Figure 7.11: The acf of standardized residuals of Bitoin returns with ARMA filter modelled by EGARCH(1,1) with student t innovation distribution (left-top) and the acf of their squared values (top-right). The acf of standardized residuals of Apple stock returns with ARMA filter modelled by EGARCH(1,1) with student t innovation distribution (bottom-left) and the acf of their squared values (bottom-right). Dotted lines mark the standard 95 % confidence intervals for autocorrelations of a process of iid finite-variance random variables.

ones in returns without ARMA filters, which is double Weibull distribution for Bitcoin and normal-inverse Gaussian distribution for Apple stock. The v-transform is still the linear v-transform with the suitable fulcrum. The copulas we used for Bitcoin include Gaussian and Frank copulas. The Gaussian copulas are used as the benchmark copulas, while the Frank copulas are the best copulas for Bitcoin returns in vt-S-vine models. The AIC value of vt-S-vine with Gaussian copulas fitted to Bitcoin returns with ARMA filter is 8572.896, the one of vt-S-vine with Frank copulas is 8568.113. The AIC values of vt-S-vines are both smaller than the best GARCH models.

For Apple stock returns, we model the ARMA residuals with vt-S-vine with Gaussian (as benchmark), survival Clayton and the combination of survival Clayton and Gaussian copulas. The combination copula sequences are the best choice in Apple stock returns, so we fit it in the returns with ARMA filter as well. The AIC values of the vt-S-vine with Gaussian, survival Clayton copulas and combination of the two copulas are 4281.210, 4254.989 and 4257.001, respectively. The best copulas are survival Clayton copulas, which are still larger than the best GARCH model. Therefore, the ARMA filters do not improve the in-sample estimation process of vt-S-vines. In order to explore this question further, we develop the estimation of vt-S-vine models of the normalized returns.

7.2.3 Modelling the normalized returns

The distribution of Bitcoin and Apple stock returns have heavier tail than normal distribution. Generally, the GARCH type models with t or skewed t innovation distributions describe the heavy tail of data precisely. If we normalize the returns, it may spoil the power tail structure, which is modelled well by GARCH type models. However, the normalization will not change the serial dependence of time series, so it may not affect the estimation of vt-S-vine models significantly. In order to present it, we take empirical distribution of the returns to transform the data to uniform distribution, then calculate the corresponding quantile values of the standard normal distribution to normalize the returns. The time series plot of the normalized Bitcoin and Apple stock returns are shown in Figure 7.12.

According to Figure 7.12, the normalization brings the data into a range that is compatible with a standard normal distribution; the values all lie between approximate -3 and 3. The fluctuations of the two data sets are decreased and the tails of the normalized returns will be smaller. Then, we fit the GARCH type models to the normalized returns and calculate the AIC values. The AIC values of Bitcoin and Apple stock are shown in Table 7.13 and Table 7.14, respectively. The innovation distributions of GARCH type models we applied for the two data sets are normal, student t, skewed student t, normal-inverse Gaussian (NIG) and generalized error (ged) distributions. The ged distribution is very competitive when with GARCH processes in normalized Apple stock returns, but it does not estimate accurately in normalized Bitcoin returns. Hence, we present

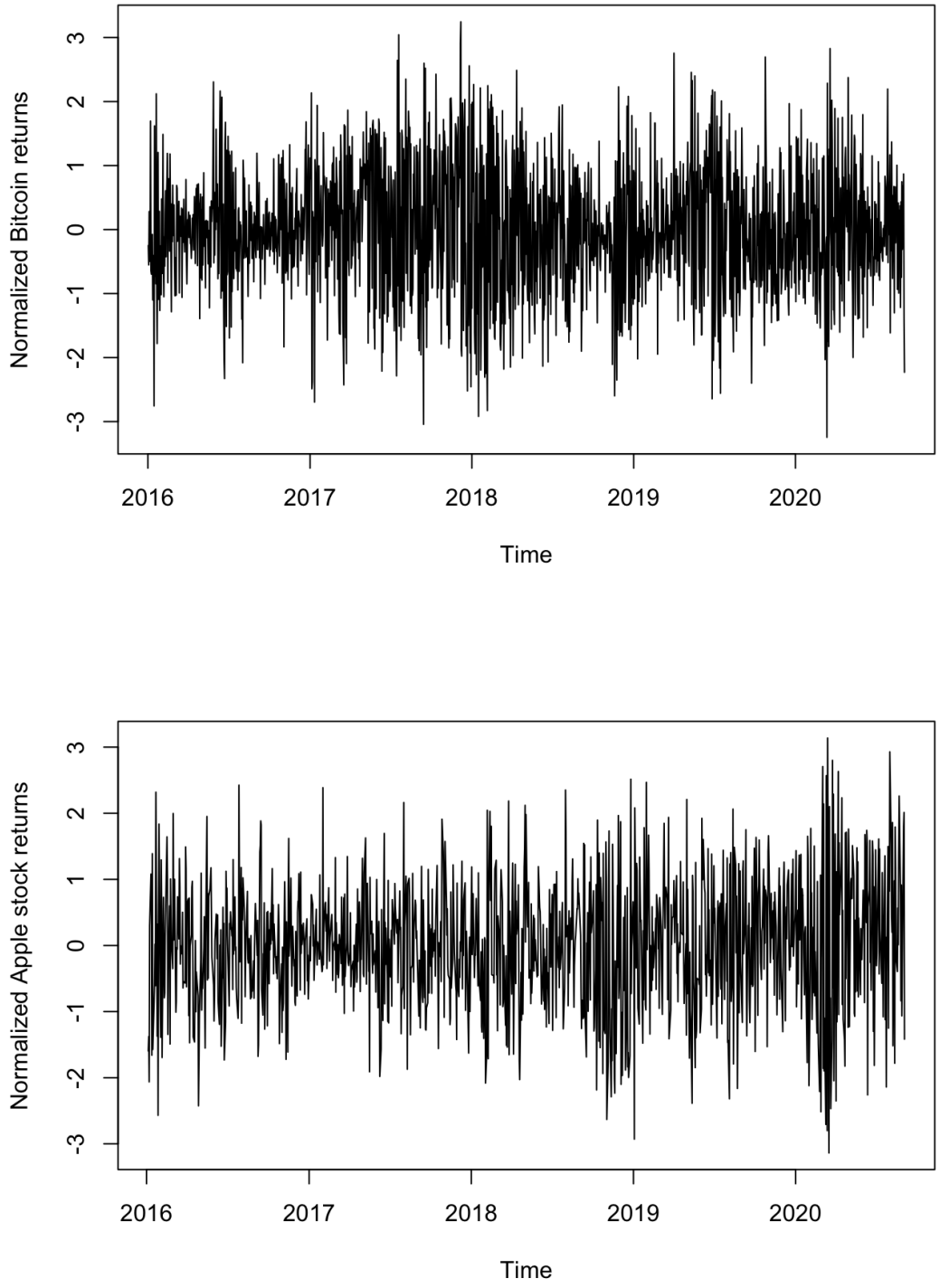


Figure 7.12: The time series plots of normalized Bitcoin returns (top) and Apple stock returns (bottom).

AIC values of GARCH type processes fitted to normalized Bitcoin returns				
	Process	GARCH(1,1)	EGARCH(1,1)	GJR-GARCH(1,1)
Innovation	norm	4611.63	4616.75	4613.28
	t	4612.46	4617.24	4613.93
	st	4614.44	4619.23	4615.93
	NIG	4614.69	4619.51	4616.22

Table 7.13: The AIC values of estimating the normalized Bitcoin returns from GARCH, GJR-GARCH and EGARCH processes with normal (norm), student t (t), skewed student t (st) and normal-inverse Gaussian (NIG) distribution. The AIC value in yellow colour represents the best process for describing the normalized Bitcoin returns.

AIC values of GARCH type processes fitted to normalized Apple stock returns				
	Process	GARCH(1,1)	EGARCH(1,1)	GJR-GARCH(1,1)
Innovation	norm	3197.43	3183.10	3180.27
	t	3201.41	3187.22	3184.49
	st	3204.75	3190.31	3187.92
	ged	3191.15	3170.05	3170.91

Table 7.14: The AIC values of estimating the normalized Apple stock returns from GARCH, GJR-GARCH and EGARCH processes with normal (norm), student t (t), skewed student t (st) and generalized error distribution (ged). The AIC value in yellow colour represents the best process for describing the normalized Apple stock returns.

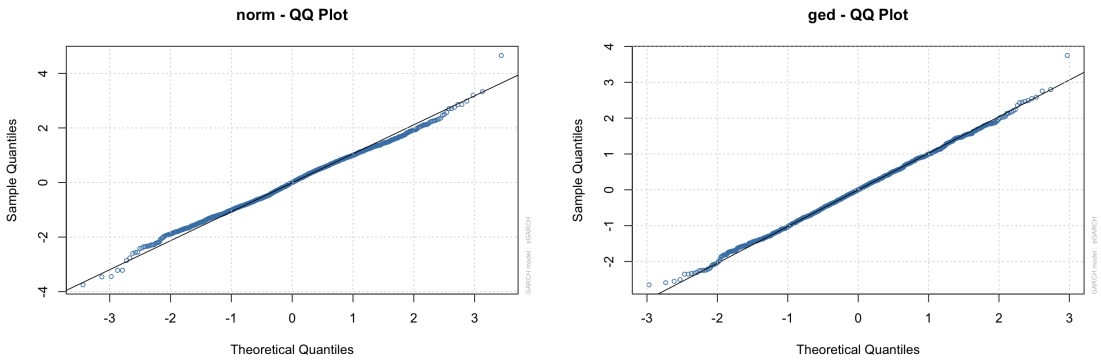
the AIC of GARCH type processes with ged innovation instead of NIG distribution in Apple stock returns, and does not show results of ged innovation in Bitcoin returns.

Table 7.13 embodies that the best GARCH type model is standard GARCH(1,1) with normal innovation distribution in Bitcoin, which is different from the results in Table 7.2. Moreover, the EGARCH(1,1) develops the worst results with all the four innovation distributions compared to GARCH(1,1) and GJR-GARCH(1,1). Therefore, the normalization may break the patterns in the time series that are captured well by EGARCH processes. The best GARCH type model for normalized Apple stock returns is EGARCH(1,1) with ged innovation. The EGARCH(1,1) and GJR-GARCH(1,1) still perform beter than standard GARCH process. The normalization does not affect the estimation results of GARCH type model significantly. For further analysis, we present the estimated parameters and the QQ-plots of the residuals of the best GARCH type models for the two data sets in Table 7.15 and Figure 7.13, separately.

The parameters of GARCH(1,1) process $\alpha_1 + \beta_1$ are close to 1 for Bitcoin returns in Table 7.15, which is close to the boundary of condition for strictly stationary of GARCH(1,1) processes. Similarly, the β_1 is also a big value for the boundary of EGARCH(1,1) in Apple stock returns. The shape parameter for Apple stock is small, which means the tail of the data is heavy. It seems that the normalization does not spoil the tail pattern of the Apple stocks. Figure 7.13 exhibits the QQ-plots of the stan-

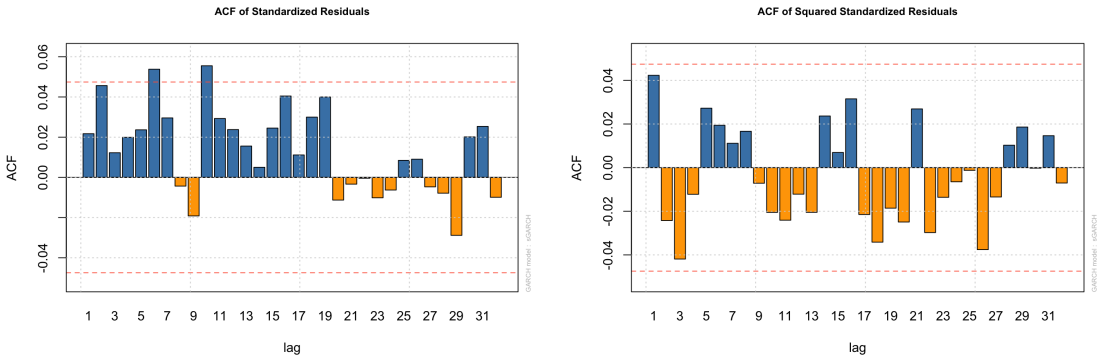
	Bitcoin (GARCH(1,1)-norm)	Apple(EGARCH(1,1)-ged)
α_0	0.0261 (0.0081)	-0.0093 (0.0057)
α_1	0.1112 (0.0193)	-0.0863 (0.0194)
β_1	0.8641 (0.0232)	0.9355 (0.0168)
γ	-	0.1768 (0.0348)
shape	-	2.5624 (0.1767)

Table 7.15: The estimated parameters in the best model (in bracket) in each data set. GARCH(1,1)-norm is the GARCH(1,1) process with normal distribution innovation. EGARCH(1,1)-ged denotes the EGARCH(1,1) model with generalized error distribution innovation. The numbers in the bracket is are the standard errors for each parameters estimation.

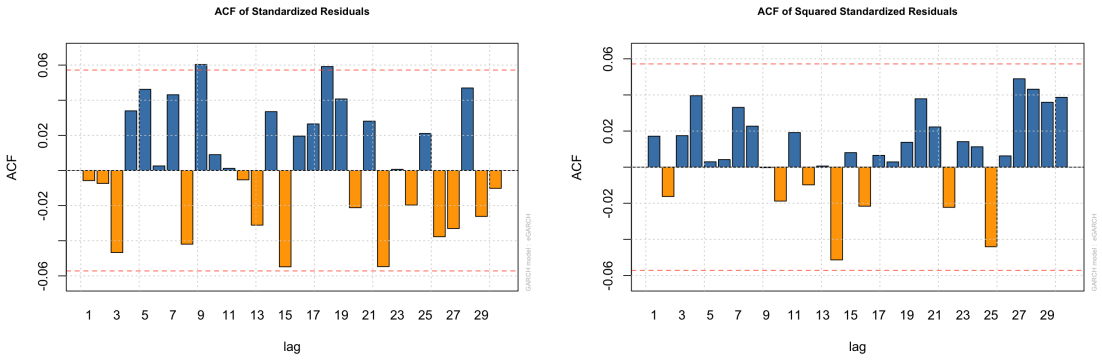


,

Figure 7.13: The QQ-plots of standardized residuals of normalized Bitcoin returns modelled by GARCH(1,1) with normal innovation distribution (left) and the QQ-plots of standardized residuals of normalized Apple stock returns modelled by EGARCH(1,1) with generalized error distribution innovation (right).



(a) Bitcoin returns



(b) Apple stock returns

Figure 7.14: The acf of standardized residuals of normalized Bitoin returns modelled by GARCH(1,1) with normal innovation distribution (left-top) and the acf of their squared values (top-right). The acf of standardized residuals of normalized Apple stock returns modelled by EGARCH(1,1) with generalized error distribution innovation (bottom-left) and the acf of their squared values (bottom-right). Dotted lines mark the standard 95 % confidence intervals for autocorrelations of a process of iid finite-variance random variables.

dardized residuals of estimations from the two data sets. The QQ-plots of standardized residuals are almost straight lines for both of the data sets. The tail of residuals of the normalized Bitcoin and Apple stock returns is not perfectly straight, which means the more suitable innovation distributions can be explored.

Figure 7.14 exhibits the acf of the standardized residuals and their squared values in both Bitcoin and Apple stock returns. The acf of Bitcoin has two violations at lag 6 and 10, while most of the acf values are consistent with the standard 95 % confidence intervals values. The situation of acf in Apple stock returns is similar to the one of Bitcoin. Most of the acf are in the range of the 95 % confidence intervals, so as the acf of their squared values.

In order to analyze the influence of the normalization to the returns, we fit the vt-S-vine models to the normalized data. The data are normally distributed, so the

AIC values of VT-S-vine fitted to Bitcoin and Apple stock			
Model	Bitcoin	Model	Apple
Gaussian	4533.59	*Gaussian*	3187.96
Frank	4526.49	*Clayton*180	3161.20
		$SC(10)\_Ga(30)$	3159.48

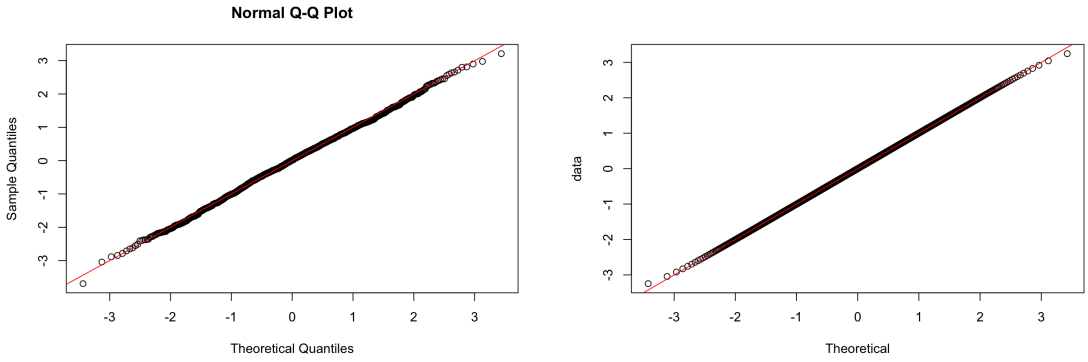
Table 7.16: The AIC values of estimating the normalized Bitcoin by vt-S-vine models with *Gaussian* and *Frank* copulas and estimating the normalized Apple stock returns by vt-S-vines with *Gaussian*, *Clayton*180 and $SC(10)\_Ga(30)$ copula sequences. The AIC value in yellow colour represent the best models for each data set. The margin is the normal distribution and v-transform is the linear v-transform.

margin of the vt-S-vine is normal distribution. The linear v-transform is used in the normalized data as well. The copulas we used for Bitcoin are Gaussian copula (as a benchmark) and Frank copula (the best copula for Bitcoin returns). In Apple stock returns, the copulas selected are Gaussian copula (as a benchmark), survival Clayton copula, and combination of the two copulas (the best copula in Apple stock returns). The AIC values of the fitting results are shown in Table 7.16.

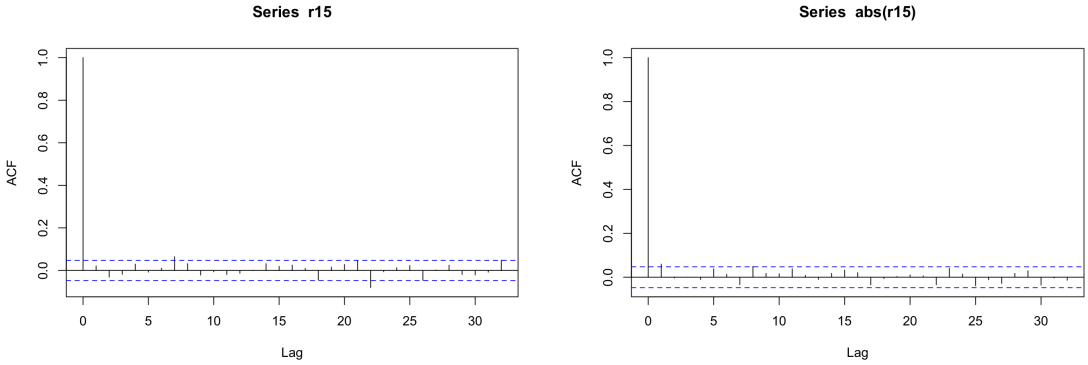
According to Table 7.16, the Frank copula is still the best copula for normalized Bitcoin returns and the combination of the survival Clayton and Gaussian copulas is the best choice for normalized Apple stock returns. It is worth noting that the vt-S-vine models with survival Clayton and the combination copulas can outperform the best GARCH type models in normalized Apple stock returns. This may be because the normalization spoils the features can be modelled well by GARCH, but it does not change the serial dependence in the time series. Another possible reason is that the marginal distribution is known here, so the estimation of vt-S-vine models will not be affected remarkably by the marginal component. Hence, the estimation process is improved by normalized data.

The vt-S-vine models with Gaussian and Frank copulas yield smaller AIC values than any GARCH type models in Bitcoin. The advantage of the vt-S-vines is overwhelming compared to GARCH type processes in both the Bitcoin returns and normalized returns. The advantage may be resulting from both margin and copula components. Also, the serial dependence of Bitcoin may be very suitable for the vine structure and symmetric copulas, like Gaussian and Frank copulas. In order to analyse the results further, we present the QQ-plots of residuals, acf of residuals, acf of their absolute values, QQ-plot of margins, v-transforms plots and the kpacf plots of the best vt-S-vine models in normalized Bitcoin and Apple stock returns in Figure 7.15 and Figure 7.16.

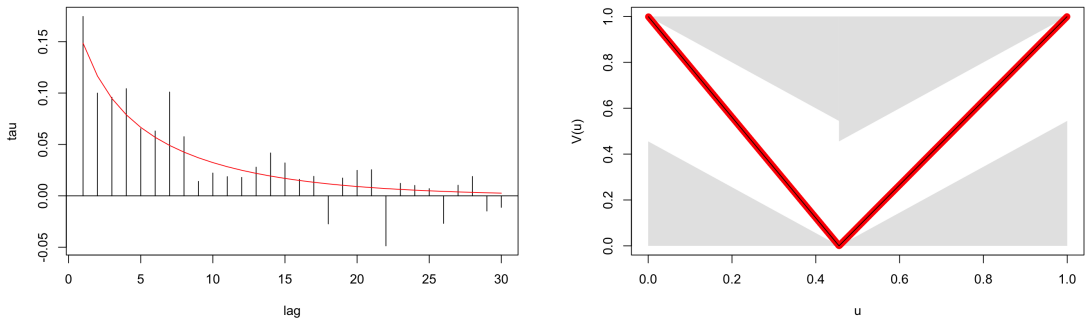
The QQ-plot of residuals in Figure 7.15 is almost straight except for some point at tails. Meanwhile, the acf plots of residuals and absolute value of residuals are all in the range the confidence level at 95 %, which means the in-sample estimation by vt-S-vine for normalized Bitcoin returns is acceptable. The margin and v-transform estimation



(a) QQ-plot of residuals (left) and margins (right) in Bitcoin

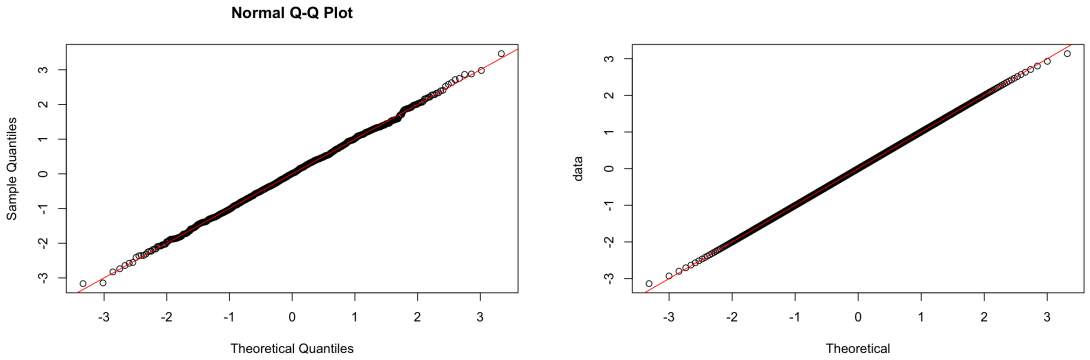


(b) Acf of residuals (left) and acf of their absolute values (right) in Bitcoin

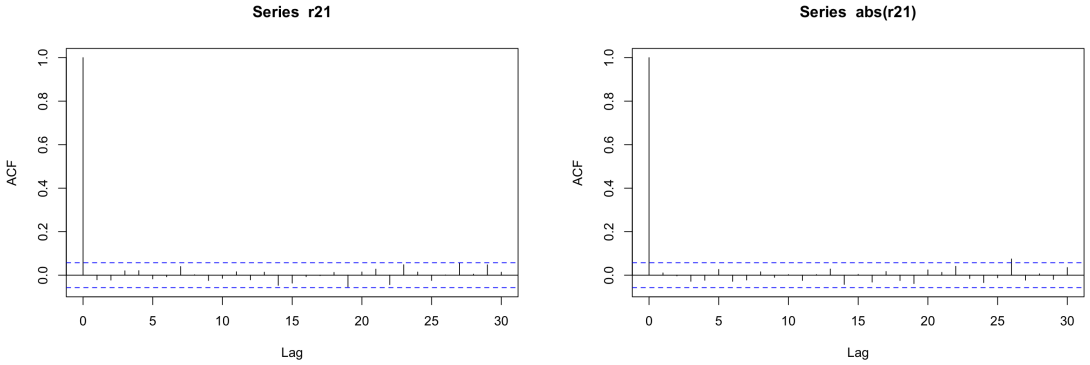


(c) Kpacf of copulas (left) and v-transforms estimations (right) in Bitcoin

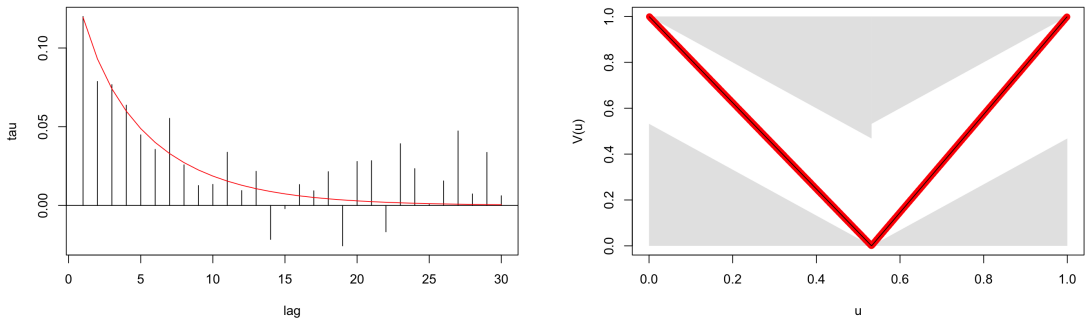
Figure 7.15: The QQ-plots of residuals, margins, acf of residuals, acf of their absolute values, kpacf plots and v-transforms plots of the vt-S-vine models with normal margins, linear v-transforms and Frank copula sequences estimated by normalized Bitcoin returns.



(a) QQ-plot of residuals (left) and margins (right) in Apple



(b) Acf of residuals (left) and acf of their absolute values (right) in Apple



(c) Kpacf of copulas (left) and v-transforms estimations (right) in Apple

Figure 7.16: The QQ-plots of residuals, margins, acf of residuals, acf of their absolute values, kpacf plots and v-transforms plots of the vt-S-vine models with normal margins, linear v-transforms and $SC(10)\_Ga(30)$ copula sequences estimated by normalized Apple stock returns.

are all very accurate. Because the true margin is normal distribution, the estimation with the true margin should be accurate. Also, the v-transform is linear, so if we find the proper fulcrum, the v-transform will be fitted precisely. Besides, the kpacf fitting is also accurate and similar to the Bitcoin returns, which proves the normalization does not change the serial dependence significantly.

Similar conclusion can be drawn for the Apple stock returns, according to Figure 7.16. There are an acceptable QQ-plots and acf plots of residuals. Also, the margin and v-transform are quite accurate. There are some defects of kpacf plots, but the kpacf estimation is reasonable. Compare the Figure 7.15 and Figure 7.16 to Figure 7.6 and Figure 7.7, the residuals and v-transform estimations are improved in the two data sets. The improvement may be caused by the proper margin selection.

In conclusion, vt-S-vine models are very competitive for modelling in-sample financial return series. Usually, they can surpass GARCH, even GJR-GARCH processes, while the EGARCH processes are difficult to be surpassed. However, the normalization can change certain features of financial return series and make vt-S-vine models more suitable than EGARCH processes. It is possibly caused by the distribution changed, which spoils the structures that can be captured by GARCH type processes, while the influence on serial dependence is hardly to be observed, which exerts less influence on vt-S-vine estimations. Besides, after normalizing the financial returns, we know the true marginal distribution is normal, so the errors of the estimation of margins will be reduced and the more uniformly distributed PIT values fitted by vt-S-vine copulas are produced, which can improve the estimation results of vt-S-vines.

7.3 Prediction using vt-S-vines

After estimating the GARCH type processes and vt-S-vine models, we do the one-step prediction with the two types of models and compare their quantile scores. Among all the datasets we tried, vt-S-vine models show better estimation results in Bitcoin. Hence, we use the Bitcoin data as an example in this section.

The quantile scores of the best two vt-S-vine and the best two GARCH type processes are shown in Table 7.17. The best GARCH type processes are GARCH with NIG innovation distribution and EGARCH with skewed t innovation. The quantiles are taken from 0.05 to 0.95, evenly distributed. The average quantile scores are calculated between all the intervals. Table 7.17 presents quantiles at $\{0.05, 0.10, 0.25, 0.5, 0.75, 0.9, 0.95\}$ as examples.

According to Table 7.17, the best two vt-S-vine models yield smaller average quantile scores compared to the GARCH type processes. The vt-S-vines with $Gu(10)\_Ga(30)$ and $SC(10)\_Ga(30)$ copula sequences have relatively smaller quantile scores at 0.05 and 0.10, which suggests that vt-S-vine predict better in the lower tails. The $t(1)\_Ga(39)$

Quantile scores of Bitcoin dataset								
Quantiles	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
$GARCH(1,1)-NIG$	337.39	515.70	847.03	997.74	829.84	507.04	311.36	778.21
$EGARCH(1,1)-st$	341.48	519.73	846.51	1001.86	824.52	494.70	302.43	776.49
$Gu(10)\_Ga(30)$	333.53	513.44	847.21	998.10	824.12	494.67	309.34	775.41
$SC(10)\_Ga(30)$	333.36	514.35	846.11	998.53	824.22	492.33	303.80	774.92
$Frank$	334.31	523.72	852.00	997.99	826.81	494.52	302.57	777.26
$t(1)\_Ga(39)$	335.67	522.18	851.61	998.08	824.28	492.41	300.51	776.42

Table 7.17: Chosen quantile score values and the means of one-step prediction at evenly distributed quantiles for vt-S-vine process with the best two copula sequences and the best two GARCH type processes in modelling log-returns of Bitcoin. $GARCH(1,1)-NIG$ and $EGARCH(1,1)-st$ are GARCH with NIG innovation and EGARCH processes with skewed t innovation, respectively. The v-transform used in this table is linear v-transform. The best two copula sequences are $Gu(10)\_Ga(30)$ and $SC(10)\_Ga(30)$. The copula sequences in blue colour is the best two copula sequences in fitting process.

yeilds smaller quantile scores at upper tail. In general, the vt-S-vine with $t(1)\_Ga(39)$ copula sequences can surpass GARCH processes in predicting, but hardly outperform EGARCH processes.

The Frank copula sequences in vt-S-vine have remarkable advantage over GARCH type processes in estimation procedure. Also, they improve the prediction results of GARCH processes. However, they do not improve the prediction compared to EGARCH processes. The difference between the prediction from vt-S-vine and EGARCH processes are not significant according to the average quantile scores in Table 7.17. It is possible that the size of data used for prediction is not large enough, so the results are not very stable and representative. When we use 500 simulations for prediction, it is difficult to distinguish bad and good models. However, if we increase it to 2000 simulations, the advantage of good model is demonstrated. In Bitcoin datasets, there are 707 testing data, which is a small sample. The replication and prediction of GARCH processes are improved more significantly than EGARCH processes, especially in vt-S-vines with $Gu(10)\_Ga(30)$ and $SC(10)\_Ga(30)$ copula sequences.

7.4 Conclusion

In this chapter, we tried three approaches to fit the GARCH type models and vt-S-vines to the financial returns. We found the ARMA filters exert slight influence on the in-sample estimations. However, the normalization can improve the effects of estimation of vt-S-vine. The main reason may be the margin is known, so we obtain a relatively perfect marginal estimation and uniformly distributed PIT values. This also reflects the importance of the marginal distribution selection. Therefore, the estimation can be elevated by finding more suitable margins. Also, the normalization breaks the features of financial returns that can be modelled by GARCH type models. Combine with the normalization, the vt-S-vine models can give superior estimations of GARCH, EGARCH and GJR-GARCH processes.

For the one-step prediction, we find the average quantile scores are more effective when the sample size is 2000 or more. However, in practice, the sample size is smaller than 2000, the results may contains certain bias. Besides, the differences between different models are not large. Even so, there are still small advantages appearing in the best vt-S-vines over GARCH type models. The estimation and prediction of vt-S-vines can be improved further when the more suitable margins and copulas are found.

Chapter 8

On the Distribution of VaR Exeedances in a Vt-S-vine Model

Techniques for measuring risk are important to the process of managing risk in financial institutions. In the banking and insurance industries, probability distributions are used to model and quantify risk. Given a profit and loss (P&L) distribution, risk measure provides a point estimate of the amount of capital required as a cushion against insolvency. The main risk measure used by financial institutions and regulators is value-at-risk (VaR), which is defined as a quantile of the P&L distribution.

The regulations usually focus on the tail of the P&L, so the quantiles for VaR are mainly chosen to be 97.5% or 99%. The Basel regulations for market risk at banks assume that if a bank has an adequate risk model, the number of VaR exceedances is binomially distributed with n being the number of days backtesting and p is the VaR percentile (1% typically for 99% VaR). Under this assumption, the Basel Committee on Banking Supervision [2019] applies the traffic light system as a regulation based on a sample with $n = 250$ observations. The system defines the backtesting green, amber and red zones according to the critical number of exceedances of levels based on 99% VaR.

The quantile exceedance indicator variables for the α-quantile are iid Bernoulli variables with success probability $1 - \alpha$ in the traffic light system. However, the quantile exceedance indicators variables tend to be positively dependent in practice. Hence, the distribution of exceedances in the P&L with serial dependence caused by stochastic volatility is not a binomial distribution. The distribution of the number of exceedances in a sequence of n values will have an identical mean to the independent case ($n \times (1-\alpha)$) but it will tend to have a higher variance and a heavier tail. The insight that dependence caused by stochastic volatility produces a distribution of VaR exceedances that is heavier tailed than binomial should provide an incentive for banks to use conditional or dynamic VaR estimation procedures.

The vt-S-vines we have referred to in the previous chapters can describe the serial dependence in the P&Ls caused by stochastic volatility. Hence, in order to explore the influence of the dependence on distribution of quantile exceedances in P&Ls, we use simulations of vt-S-vines to mimic the P&L of its trading book and study the exceedances of the determined VaR in this chapter.

Another question for risk managers is: What is the best empirical quantile estimator for VaR? Empirical quantile estimation by far is the most popular method of estimating a quantile from a sample of losses. All empirical quantile estimators work by interpolating the order statistics of the sample but there are a number of different ways of doing this and they each have subtly different properties. We employ six widely used empirical quantile estimator approaches in the simulations of vt-S-vines to investigate which method banks should prefer in the sample with dependence.

Section 8.1 introduces the methods used to estimate VaR. The approaches applied to calculate exceedances and VaR estimation errors in iid case are shown in Section 8.2. Section 8.2 also investigates the exceedances of simulations with serial dependence and volatility, including simulations of vt-S-vine models. In Section 8.3, we study the distribution of exceedances in vt-S-vine models and the beta-binomial distribution is applied to model the exceedances. The conclusions of this chapter are found in Section 8.4.

8.1 Estimation of VaR

8.1.1 Historical simulation

The historical simulation method is based on re-sampling of historical risk-factor changes or returns. Supposing the loss L_t incurred on a portfolio of traded assets in time period t. Assume there is a window of n historical datapoints (can also be called lookback period of length n), the procedure obtains a sample of losses $S = \{L_1, ..., L_n\}$ at a fixed time.

Let $\alpha \in [0, 1]$ and let $\widehat{VaR}_\alpha$ denote the estimated value of VaR_α at fixed time t. The $\widehat{VaR}_\alpha$ can be calculated as an empirical α-quantile of the sample S. There are many approaches to compute the empirical quantiles and all of them involve the order statistics, which are denoted by $L_{(1)} < \cdots < L_{(n)}$. We assume no ties which is true of continuous data with probability 1.

8.1.2 Empirical quantiles

In this section, we calculate the empirical quantiles following the procedure referred to by Hyndman and Fan [1996]. Let

$$K(\alpha, n) = \lfloor n\alpha + \eta(\alpha) \rfloor, \tag{8.1}$$

where $\lfloor \cdot \rfloor$ denotes the floor-function which represents the largest integer not exceeding the observation. $\eta(\alpha)$ is a linear function of α taking values in $[0, 1]$. The function $\eta(\alpha)$ has varied forms to define different expressions of empirical quantiles, which we referred to as the adjustment function. The $K(\alpha, n)$ identify the order statistics we were used to estimate the α-quantile in a sample of size n for a particular choice for the adjustment function. Let

$$\lambda(\alpha, n) = n\alpha + \eta(\alpha) - K(\alpha, n), \tag{8.2}$$

where $\lambda(\alpha, n) \in [0, 1)$.

The estimated α-quantiles from a sample $\{L_1, ..., L_n\}$ can be expressed as

$$\begin{cases} (1 - \lambda(\alpha, n))L_{(K(\alpha,n))} + \lambda(\alpha, n)L_{(K(\alpha,n)+1)} & \text{if} \quad K(\alpha, n) \in \{1, ..., n-1\}, \\ L_{(1)} & \text{if} \quad K(\alpha, n) = 0, \\ L_{(n)} & \text{if} \quad K(\alpha, n) = n. \end{cases} \tag{8.3}$$

For every $K(\alpha, n) \in \{1, ..., n\}$ there is a value $\alpha(k, n)$ satisfying

$$K(\alpha(k, n), n) = k \quad \text{and} \quad \lambda(\alpha(k, n), n) = 0, \tag{8.4}$$

where we can also obtain $\widehat{VaR}_{\alpha(k,n)} = L_{(k)}$ and k represents the kth order statistic of the observations. For values of α such that $\alpha(k, n) < \alpha < \alpha(k + 1, n)$ the quantile estimation is given by linear interpolation between $L_{(k)}$ and $L_{(k+1)}$. Hyndman and Fan [1996] summarize nine types of definitions for $\eta(\alpha)$. In this section, we follow them, considering 6 popular choices for $\eta(\alpha)$.

1. $\eta(\alpha) = 0$. In this case the $\alpha(k, n) = k/n$ so that $L_{(k)}$ estimates the (k/n)-quantile. In this definition $L_{(1)}$ is the α-quantile estimate for $\alpha \leq 1/n$; $L_{(n)}$ is the 100%-quantile estimate.

2. $\eta(\alpha) = 1/2$. In this definition $L_{(1)}$ is the α-quantile estimate for $\alpha \leq 1/(2n)$; $L_{(n)}$ is the α-quantile estimate for $\alpha \geq 1 - 1/(2n)$. The $L_{(1)}$ and $L(n)$ have more symmetric properties in this case.

3. $\eta(\alpha) = \alpha$. In this case $\alpha(k, n) = k/(n+1)$. In this definition $L_{(1)}$ is the α-quantile estimate for $\alpha \leq 1/(n + 1)$; $L_{(n)}$ is the α-quantile estimate for $\alpha \geq n/(n + 1)$.

4. $\eta(\alpha) = 1 - \alpha$. In this case $\alpha(k, n) = (k-1)/(n-1)$. In this definition $L_{(1)}$ and $L_{(n)}$ is the 0-quantile and 100%-quantile estimate, respectively.

5. $\eta(\alpha) = (\alpha + 1)/3$. In this case $\alpha(k, n) = (k - 1/3)/(n + 1/3)$. In this definition $L_{(1)}$ is the α-quantile estimate for $\alpha \leq 1/(3n+1)$; $L_{(n)}$ is the α-quantile estimate for $\alpha \geq (3n-1)/(3n+1)$.

6. $\eta(\alpha) = (2\alpha + 3)/8$. In this case $\alpha(k, n) = (8k - 3)/(8n + 2)$. In this definition $L_{(1)}$ is the α-quantile estimate for $\alpha \leq 5/(8n+2)$; $L_{(n)}$ is the α-quantile estimate for $\alpha \geq (8n-3)/(8n+2)$.

8.2 Exceedances of simulations from forecasting models

The exceedance probability is the probability that a loss L exceeds the estimated or calculated $\widehat{VaR}_\alpha$, which can be expressed as $\mathbb{P}(L > \widehat{VaR}_\alpha)$. If we assume the sample of losses $S = \{L_1, ..., L_n\}$ are independent and identically-distributed, then the exceedance probabilities can be calculated. We show the calculation and results of exceedance probabilities in the case of iid losses in Section 8.2.1.

Practically, there is serial dependence in the losses in most cases. Therefore, the exceedance probabilities of simulations from models with serial dependence are studied as well. We investigate the exceedances of vt-S-vine models in Section 8.2.2. In order to present the results clearly, we also calculate the expected exceedances, mean absolute error (MAE) and the bias in each model.

8.2.1 Independent and identically distributed case

When the losses are iid, the general calculation of the exceedance probability can be computed by using an integral over the joint density of the order statistics. Let $\lambda = \lambda(\alpha, n)$ to simplify the notation and k is the order of the observation, the exceedance probabilities can be calculated as

$$\begin{aligned}
&\mathbb{P}(L > \widehat{VaR}_\alpha) \\
&= \int_{-\infty}^{\infty} \int_{x}^{\infty} \mathbb{P}(L > (1-\lambda)x + \lambda y) f_{L_{(k)}, L_{(k+1)}}(x, y) dy dx \\
&= \int_{-\infty}^{\infty} \int_{x}^{\infty} \mathbb{P}(L > (1-\lambda)x + \lambda y) \frac{n! F(x)^{k-1} (1 - F(y))^{n-k-1}}{(k-1)!(n-k-1)!} f(x) f(y) dy dx,
\end{aligned} \tag{8.5}$$

where F is the cdf of L. Then, make the substitutions that $u = F(x)$, $v = F(y)$ and $U = F(L)$,

$$\begin{aligned}
&\mathbb{P}(L > \widehat{VaR}_\alpha) \\
&= \int_0^1 \int_u^1 \mathbb{P}(U > F((1-\lambda)F^{-1}(u) + \lambda F^{-1}(v))) \frac{n! u^{k-1}(1-v)^{n-k-1}}{(k-1)!(n-k-1)!} dv du \\
&= \int_0^1 \int_u^1 (1 - F((1-\lambda)F^{-1}(u) + \lambda F^{-1}(v))) \frac{n! u^{k-1}(1-v)^{n-k-1}}{(k-1)!(n-k-1)!} dv du.
\end{aligned} \tag{8.6}$$

Equation 8.6 can be computed numerically. The underlying distribution F in Equation 8.6 is the key factor for calculation of exceedance probabilities in the iid case. We select three commonly used distributions in the calculation, which are normal, student t and skewed student t distributions. The sample size n is 250 and 500, since there are around 250 trading days in the financial equity markets each year and we decide to study the historical simulations for one or two years. The target quantiles in the simulations are 0.975 and 0.99, which are widely used values in regulation and banking. The quantile estimation methods used are the six referred to in Section 8.1.2. The calculated exceedance probabilities are shown in Table 8.1. Table 8.2 presents the expected exceedances corresponding to Table 8.1.

			methods						
dist	α	n	1	2	3	4	5	6	true
norm	0.975	250	0.0288	0.0268	0.0249	0.0287	0.0261	0.0263	0.0250
		500	0.0269	0.0259	0.0250	0.0269	0.0256	0.0257	0.0250
	0.99	250	0.0137	0.0120	0.0096	0.0137	0.0111	0.0113	0.0100
		500	0.0120	0.0109	0.0100	0.0120	0.0106	0.0107	0.0100
t	0.975	250	0.0288	0.0268	0.0249	0.0287	0.0261	0.0263	0.0250
		500	0.0269	0.0259	0.0250	0.0269	0.0256	0.0257	0.0250
	0.99	250	0.0136	0.0120	0.0096	0.0136	0.0110	0.0112	0.0100
		500	0.0120	0.0109	0.0100	0.0120	0.0106	0.0106	0.0100
skt	0.975	250	0.0288	0.0268	0.0249	0.0287	0.0261	0.0263	0.0250
		500	0.0269	0.0259	0.0250	0.0269	0.0256	0.0257	0.0250
	0.99	250	0.0136	0.0120	0.0096	0.0136	0.0110	0.0112	0.0100
		500	0.0120	0.0109	0.0100	0.0120	0.0106	0.0106	0.0100

Table 8.1: Exceedance probabilities in the iid case (calculated). *norm* is the standard normal distribution. *t* represents student t distribution with degree of freedom, $\nu = 6$. *skt* is the skewed student t distribution with degree of freedom 6 and skew parameter equals to 0.8. *true* is the true values of exceedance probabilities.

According to Table 8.1 and 8.2, the exceedance probabilities calculated by the third method are the closest to the true values, which should be 0.025 in the 0.975-quantile

			methods						
dist	α	n	1	2	3	4	5	6	true
norm	0.975	250	7.20	6.70	6.22	7.18	6.53	6.57	6.25
		500	13.46	12.97	12.48	13.43	12.81	12.85	12.50
	0.99	250	3.42	2.99	2.41	3.41	2.77	2.82	2.50
		500	5.99	5.45	5.00	5.98	5.29	5.33	5.00
t	0.975	250	7.20	6.69	6.22	7.17	6.53	6.57	6.25
		500	13.45	12.97	12.48	13.43	12.80	12.85	12.50
	0.99	250	3.41	2.99	2.39	3.40	2.75	2.81	2.50
		500	5.99	5.44	5.00	5.98	5.28	5.32	5.00
skt	0.975	250	7.20	6.69	6.22	7.17	6.53	6.57	6.25
		500	13.45	12.97	12.48	13.43	12.80	12.85	12.50
	0.99	250	3.41	2.99	2.39	3.40	2.76	2.81	2.50
		500	5.99	5.44	5.00	5.98	5.28	5.32	5.00

Table 8.2: Expected exceedances in iid case (calculated). *norm* is the standard normal distribution. *t* represents student t distribution with degree of freedom, $\nu = 6$. *skt* is the skewed student t distribution with degree of freedom 6 and skew parameter equals to 0.8. *true* is the true values of expected exceedances.

and 0.010 in the 0.99-quantile. For example, the true exceedance probability is 0.025 when $\alpha = 0.975$, the probability calculated by the third method from t distribution is 0.0249 and 0.025 in sample size 250 and 500, respectively. The two calculated values are quite similar to the true value. Furthermore, the exceedance probabilities and expected exceedances are closer to the true values when the sample size increased to 500 from 250. It reveals there is an influence of the sample size in the accuracy of the quantile estimation methods. Moreover, the exceedance probabilities and expected exceedances are not very different among the three types of distributions, which implies that the distribution does not affect the exceedances probabilities and expected exceedances too much in iid case.

Furthermore, it is worth noting that only the third method understates the exceedances probabilities and expected exceedances. The other five approaches yields the overstated results. The regulations of bank are based on numbers of exceedances. Banks that consistently experience too many VaR exceedances in a time period may be subject to higher capital requirements and may lose their option to use the internal-models-based approach to capital that most sophisticated institutions prefer to use. Hence, the other five approaches which always overstate the exceedances may not be reasonable choices for banks.

In order to exhibit the results further, we calculate the MAE and bias of the esti-

			methods					
dist	α	n	1	2	3	4	5	6
norm	0.975	250	0.1329	0.1317	0.1365	0.1328	0.1322	0.1320
		500	0.0943	0.0949	0.0956	0.0943	0.0947	0.0947
	0.99	250	0.1801	0.1843	0.1973	0.1799	0.1830	0.1828
		500	0.1325	0.1290	0.1379	0.1323	0.1308	0.1302
t	0.975	250	0.2262	0.2281	0.2407	0.2263	0.2304	0.2296
		500	0.1614	0.1638	0.1665	0.1614	0.1640	0.1639
	0.99	250	0.3688	0.3918	0.4495	0.3687	0.3975	0.3946
		500	0.2742	0.2733	0.2981	0.2739	0.2792	0.2775
skt	0.975	250	0.1792	0.1805	0.1902	0.1793	0.1822	0.1816
		500	0.1278	0.1296	0.1316	0.1278	0.1297	0.1296
	0.99	250	0.2891	0.3067	0.3511	0.2891	0.3110	0.3088
		500	0.2148	0.2139	0.2332	0.2146	0.2185	0.2172

Table 8.3: Mean absolute error of the VaR estimator in the iid case (calculated). *norm* is the standard normal distribution. t represents student t distribution with degree of freedom, $\nu = 6$. *skt* is the skewed student t distribution with degree of freedom 6 and skew parameter equals to 0.8.

mator $\widehat{VaR}_\alpha$. The MAE can be calculated as

$$\mathbb{E}(|\widehat{VaR}_\alpha - F^{-1}(\alpha)|) = \int_0^1 \int_u^1 |(1-\lambda)F^{-1}(u) + \lambda F^{-1}(v) - F^{-1}(\alpha)| \frac{n!u^{k-1}(1-v)^{n-k-1}}{(k-1)!(n-k-1)!} dvdu. \tag{8.7}$$

The results for MAE are presented in Table 8.3 and for the bias in Table 8.4.

The MAE and bias reflect the difference between the true VaR and the estimated VaR. According to Table 8.3, the second and sixth approaches give the smallest MAE values in iid case with normal distribution, which means the two approaches develop VaR estimation closer to the true VaR than other methods in normal distributions. For student t and skewed student t distributions, the first and forth methods produce more accurate VaR estimations, since they yield smaller MAE values. The best methods for VaR estimations are distinctive in different distributions and at different quantiles according to MAE values. However, the third method for quantile estimation yields the biggest MAE values, which means the third method yields the least accurate quantile estimates.

Combining the bias in Table 8.4, the method 3,5 and 6 have always given the positive bias, which means that these methods overestimate the true values. They are conservative in regard to VaR estimation. In contrast, methods 1 and 4 underestimate the true values. It is apparent that larger sample size can improve the quantile estimation,

			methods					
dist	α	n	1	2	3	4	5	6
norm	0.975	250	-0.0371	-0.0040	0.0294	-0.0355	0.0071	0.0043
		500	-0.0189	-0.0024	0.0144	-0.0180	0.0032	0.0018
	0.99	250	-0.0739	-0.0108	0.0768	-0.0726	0.0184	0.0111
		500	-0.0405	-0.0033	0.0332	-0.0398	0.0089	0.0058
t	0.975	250	-0.0516	0.0058	0.0648	-0.0490	0.0255	0.0205
		500	-0.0263	0.0021	0.0314	-0.0249	0.0118	0.0094
	0.99	250	-0.1211	0.0112	0.2161	-0.1184	0.0795	0.0624
		500	-0.0687	0.0111	0.0894	-0.0671	0.0372	0.0307
skt	0.975	250	-0.0415	0.0039	0.0505	-0.0394	0.0194	0.0156
		500	-0.0212	0.0013	0.0245	-0.0200	0.0090	0.0071
	0.99	250	-0.0959	0.0077	0.1677	-0.0938	0.0610	0.0477
		500	-0.0543	0.0082	0.0694	-0.0531	0.0286	0.0235

Table 8.4: Bias of the VaR estimator in the iid case (calculated). *norm* is the standard normal distribution. t represents student t distribution with degree of freedom, $\nu = 6$. *skt* is the skewed student t distribution with degree of freedom 6 and skew parameter equals to 0.8.

according to Table 8.3. Moreover, the signs of bias are not influenced remarkably by the distributions, but highly depend on the methods used.

Furthermore, one interesting finding is that the quantile estimator is that closest on average to the true value of the quantile of a distribution may not be the one that tends to give the most accurate value for the exceedance probability. For example, the third method yields the exceedance probabilities closest to the true values according to Table 8.1 and 8.2. However, the MAE values of Method 3 are larger than the ones estimated by any other approaches, which means the VaR estimations from Method 3 are the least accurate among the six approaches. Therefore, the exceedance probabilities and MAE or bias give conflicting recommendations of the best method, which gives an indicator as to the method of empirical quantile estimation that banks should prefer.

Traffic light system of iid losses

Validation of models is largely based on the concepts of VaR exceptions, which is the exceedances of VaR estimates by the realized loss. In a trading year of 250 days, too many exceptions to the 99% VaR can have undesired consequences for a bank. The traffic light system in trading book stipulate:

- 5 or more VaR exceedances lead to an amber traffic light and also lead to an increase in the multiplier. The number 5 corresponds roughly to the 95% quantile of a binomial distribution $B(250, 0.01)$.

- 10 or more VaR exceedances for the overall bank portfolio leads to a red traffic light and a maximum multiplier applied to a bank's capital calculations. The idea is that 10 is approximately the 99.99% quantile of the distribution of a $B(250, 0.01)$ distribution ([on Banking Supervision, 2019, p. 128]).

- More than 12 exceedances for a particular desk means that capital for that desk's positions must be determined using the standardised approach, which is likely to be more capital intensive. The supervisor has the option of disallowing use of the internal model when there are 13 or more VaR exceedances ([on Banking Supervision, 2019, p. 83]).

Hence, there are some critical levels in the number of exceedances of the 99% VaR in 250 days: 5+, 10+ and 13+. The number of exceedances of the true α-quantile in a sequence of n values from an iid process follows a binomial distribution $B(n, 1-\alpha)$. The α used here is the exceedance probability calculated by each distribution at selected quantile. Hence, we calculate the probabilities at the three levels $\mathbb{P}(N \geq 5)$, $\mathbb{P}(N \geq 10)$ and $\mathbb{P}(N \geq 13)$, where N is the number of exceedances. The results are shown in Table 8.5.

		Methods					
dist	levels	1	2	3	4	5	6
normal	5	0.259606	0.181709	0.096410	0.257768	0.146966	0.154977
	10	0.002630	0.000978	0.000189	0.002577	0.000556	0.000640
	13	0.000051	0.000013	0.000001	0.000050	0.000006	0.000007
t	5	0.256763	0.181709	0.093762	0.254954	0.144279	0.152689
	10	0.002548	0.000978	0.000176	0.002497	0.000530	0.000615
	13	0.000049	0.000013	0.000001	0.000048	0.000005	0.000007
skt	5	0.256892	0.181709	0.093858	0.255082	0.144386	0.152782
	10	0.002552	0.000978	0.000176	0.002501	0.000531	0.000616
	13	0.000049	0.000013	0.000001	0.000048	0.000005	0.000007

Table 8.5: The calculated probabilities in binomial distribution at 5+, 10+, 13+ exceedances in iid case. *norm* is the standard normal distribution. *t* represents student t distribution with degree of freedom, $\nu = 6$. *skt* is the skewed student t distribution with degree of freedom 6 and skew parameter equals to 0.8.

According to Table 8.5, the third method always gives the smallest probabilities at all three levels and distributions, which implies that the bank applies the third method will have the smallest probabilities of getting the amber, red zone or disallowing use of the internal model in traffic light system. The first and fourth methods yielding probabilities that exceedances over 13 have different orders of magnitude compared to Method 3. The probabilities of disallowing the internal models in the two approaches are much higher than the third method. Moreover, the probabilities are not affected

by the choice of distributions. For level 5 and 10, all the approaches, except for the third one are bigger than the standard $5\%(= 1 - 95\%)$ and $0.01\%(= 1 - 99.99\%)$ on Banking Supervision [2019], which means they tend to overstate the probabilities of the amber and red zones in the traffic light system. For level 13, the results are very close to 0 among distinctive approaches, which means the probability of the internal model being refused by traffic light system is not influenced largely by different quantile estimators.

8.2.2 The case of vt-S-vine models

The previous section investigates the VaR estimation in the case of iid losses. In this section, we consider the losses with both serial dependence caused by stochastic volatility. Therefore, we use the vt-S-vine models to simulate data and study the exceedances estimation in the simulations. We apply the linear v-transform with fulcrum 0.5 in vt-S-vine copulas. The parameters in ARMA(1,1) which are used to estimate partial autocorrelations are $\phi_1 = 0.95$ and $\theta_1 = -0.85$. The values of ϕ_1 and θ_1 are chosen according to the results of empirical study in the previous chapter. The copula sequences used in simulation are Gaussian and survival Clayton copulas. The vt-S-vine with Gaussian copula sequences are equal to the VT-ARMA models. Besides, survival Clayton copulas usually perform well in modelling empirical time series, which is shown in Chapter 6.

Assume the uniform distributed simulations generated from vt-S-vine copulas are $u_1, ..., u_n$ and the marginal distribution we select is F. Then, the simulated losses are denoted as $\{L_1, ..., L_n\}$, where $L_i = F^{-1}(u_i)$, $i = 1, ..., n$. Following this, the estimate $\widehat{VaR}_\alpha$ of the α-quantile from the sample $\{L_1, ..., L_n\}$ is obtained using the six approaches introduced in Section 8.2.1. The exceedance probabilities in vt-S-vines can be calculated as

$$\mathbb{P}(L > L_{(k)}) = 1 - F(\widehat{VaR}_\alpha). \tag{8.8}$$

The expected exceedances are written as $(1 - F(\widehat{VaR}_\alpha)) \times n$, where n is the sample size.

Besides, the MAE and bias can be calculated using the estimated $\widehat{VaR}_\alpha$ of the α-quantile from the simulated sample $\{L_1, ..., L_n\}$ as

$$\begin{aligned} MAE &= \mathbb{E}(|\widehat{VaR}_\alpha - F^{-1}(\alpha)|), \\ Bias &= \mathbb{E}(\widehat{VaR}_\alpha - F^{-1}(\alpha)). \end{aligned} \tag{8.9}$$

The α is set to 0.975 and 0.99, which is the same as the ones in iid case. The sample sizes n used are 250 and 500. The simulations are repeated 1000 times in each case. The MAE and bias are calculated from the 1000 repetitions. The exceedance probabilities are estimated using Equation 8.8. The results are put in Table 8.6. The expected

exceedances are shown in Table 8.7.

According to Table 8.6, the method 3 gives the exceedance probabilities that are closest to true values, which are 0.025 for $\alpha = 0.975$ and 0.010 for $\alpha = 0.99$. Comparing to Table 8.1, the vt-S-vine models yield bigger exceedance probabilities in all the cases than in the iid case. The exceedances in Table 8.7 are all bigger than the true values, which means the empirical quantile estimation always results in an exceedance probability larger than the true values in a vt-S-vine. Moreover, the selection of distribution does not affect the exceedances estimation, which is the same observation in Table 8.6 and 8.7. It is noteworthy that the vt-S-vine with survival Clayton copulas yields exceedances apparently further than true values than Gaussian copulas. The choice of copulas in vt-S-vine models may crucially effect the exceedances estimations.

In order to assess the accuracy of the VaR estimates, we exhibit the MAE and Bias of estimator $\widehat{VaR}_\alpha$ in Table 8.8 and Table 8.9. The MAE values of vt-S-vine with survival Clayton copulas are larger than the ones of vt-S-vine with Gaussian copulas, according to Table 8.8, which means the VaR estimated by quantile estimation methods in the survival Clayton copulas are less precisely than the ones in Gaussian copulas. Furthermore, the normal distribution yields the smallest MAE values in all the six methods for both Gaussian and survival Clayton copulas in vt-S-vine models. Moreover, it is difficult to find a method that estimate VaR most or least accurately in all cases. It may be advisable to choose different methods according to distributions, quantiles and sample sizes. Furthermore, compared to Table 8.2, the vt-S-vine models produce higher MAE values than the iid losses, which reveals that the empirical quantile estimates less accurately in the presence of serial dependence.

Table 8.9 shows that the first, second and forth approaches have negative bias, which reveals that the three approaches understate VaR. Interestingly, method 3 has positive bias in vt-S-vine with Gaussian copulas, while negative bias in vt-S-vine with survival Clayton copulas, which embodies that vt-S-vine can overstate VaR with Gaussian copulas, but understate VaR with survival Clayton copulas using the third method. Therefore, the nature of the serial dependence that is generated from copulas affects the accuracy of quantile estimation very much in vt-S-vine models.

Traffic light system in vt-S-vines

In order to show the application of vt-S-vine models, we generate the traffic light system of the simulations from vt-S-vines as well. The critical levels in the number of exceedances of the 99% VaR in 250 days are 5,10 and 13. However, in the vt-S-vine simulated losses, the quantile exceedance indicator variables for the α-quantile are identically, but not independent distributed Bernoulli variables with success probability $1 - \alpha$. Instead they will tend to be positively dependent.

The distribution of the number of exceedances in a sequence of n values will have

			Methods						
dist	α	n	1	2	3	4	5	6	true
					Gaussian copula				
norm	0.975	250	0.0306	0.0286	0.0267	0.0306	0.0279	0.0281	0.0250
		500	0.0280	0.0270	0.0260	0.0279	0.0266	0.0267	0.0250
	0.99	250	0.0149	0.0130	0.0107	0.0148	0.0121	0.0123	0.0100
		500	0.0127	0.0116	0.0107	0.0126	0.0113	0.0114	0.0100
t	0.975	250	0.0306	0.0286	0.0266	0.0305	0.0279	0.0281	0.0250
		500	0.0279	0.0270	0.0260	0.0279	0.0266	0.0267	0.0250
	0.99	250	0.0148	0.0130	0.0106	0.0148	0.0121	0.0123	0.0100
		500	0.0127	0.0116	0.0107	0.0126	0.0113	0.0114	0.0100
skt	0.975	250	0.0306	0.0286	0.0266	0.0305	0.0279	0.0281	0.0250
		500	0.0279	0.0270	0.0260	0.0279	0.0266	0.0267	0.0250
	0.99	250	0.0148	0.0130	0.0106	0.0148	0.0121	0.0123	0.0100
		500	0.0127	0.0116	0.0107	0.0126	0.0113	0.0114	0.0100
					Survival Clayton copula				
norm	0.975	250	0.0366	0.0347	0.0329	0.0365	0.0341	0.0342	0.0250
		500	0.0312	0.0302	0.0292	0.0311	0.0299	0.0299	0.0250
	0.99	250	0.0204	0.0183	0.0157	0.0204	0.0174	0.0176	0.0100
		500	0.0158	0.0147	0.0137	0.0158	0.0144	0.0144	0.0100
t	0.975	250	0.0366	0.0347	0.0328	0.0365	0.0340	0.0342	0.0250
		500	0.0312	0.0302	0.0292	0.0311	0.0299	0.0299	0.0250
	0.99	250	0.0204	0.0183	0.0156	0.0203	0.0173	0.0175	0.0100
		500	0.0158	0.0147	0.0137	0.0158	0.0143	0.0144	0.0100
skt	0.975	250	0.0366	0.0347	0.0328	0.0365	0.0340	0.0342	0.0250
		500	0.0312	0.0302	0.0292	0.0311	0.0299	0.0299	0.0250
	0.99	250	0.0204	0.0183	0.0156	0.0203	0.0173	0.0176	0.0100
		500	0.0158	0.0147	0.0137	0.0158	0.0143	0.0144	0.0100

Table 8.6: exceedance probabilities of vt-S-vine with symmetric linear v-transformation with Gaussian copula and survival Clayton copulas. The parameters in ARMA(1,1) estimation part are $\phi_1 = 0.95$ and $\theta_1 = -0.85$. *norm* is the standard normal distribution. t represents student t distribution with degree of freedom, $\nu = 6$. *skt* is the skewed student t distribution with degree of freedom 6 and skew parameter equals to 0.8.

			Methods						
dist	α	n	1	2	3	4	5	6	true
Gaussian copula									
norm	0.975	250	7.66	7.15	6.67	7.64	6.98	7.02	6.25
		500	13.98	13.49	13.01	13.95	13.32	13.36	12.50
	0.99	250	3.72	3.25	2.67	3.71	3.04	3.09	2.50
		500	6.33	5.81	5.37	6.32	5.66	5.69	5.00
t	0.975	250	7.66	7.14	6.66	7.63	6.97	7.02	6.25
		500	13.97	13.49	13.01	13.95	13.32	13.36	12.50
	0.99	250	3.70	3.25	2.65	3.69	3.02	3.07	2.50
		500	6.33	5.80	5.37	6.32	5.65	5.69	5.00
skt	0.975	250	7.66	7.15	6.66	7.63	6.98	7.02	6.25
		500	13.97	13.49	13.01	13.95	13.32	13.36	12.50
	0.99	250	3.70	3.25	2.65	3.69	3.02	3.07	2.50
		500	6.33	5.80	5.37	6.32	5.65	5.69	5.00
Survival Clayton copula									
norm	0.975	250	9.16	8.67	8.22	9.13	8.52	8.56	6.25
		500	15.59	15.10	14.60	15.56	14.93	14.97	12.50
	0.99	250	5.11	4.57	3.93	5.10	4.34	4.40	2.50
		500	7.91	7.35	6.86	7.90	7.18	7.22	5.00
t	0.975	250	9.15	8.67	8.21	9.13	8.51	8.55	6.25
		500	15.58	15.10	14.60	15.56	14.93	14.97	12.50
	0.99	250	5.10	4.57	3.91	5.08	4.33	4.39	2.50
		500	7.91	7.34	6.86	7.90	7.17	7.21	5.00
skt	0.975	250	9.15	8.67	8.21	9.13	8.51	8.55	6.25
		500	15.58	15.10	14.60	15.56	14.93	14.97	12.50
	0.99	250	5.10	4.57	3.91	5.08	4.33	4.39	2.50
		500	7.91	7.34	6.86	7.90	7.17	7.21	5.00

Table 8.7: Expected exceedance of vt-S-vine with symmetric linear v-transformation with Gaussian copulas and survival Clayton copulas. The parameters in ARMA(1,1) estimation part are $\phi_1 = 0.95$ and $\theta_1 = -0.85$. *norm* is the standard normal distribution. t represents student t distribution with degree of freedom, $\nu = 6$. *skt* is the skewed student t distribution with degree of freedom 6 and skew parameter equals to 0.8.

				methods					
copula	dist	α	n	1	2	3	4	5	6
Ga	norm	0.975	250	0.1788	0.1767	0.1785	0.1788	0.1766	0.1766
			500	0.1241	0.1237	0.1248	0.1240	0.1238	0.1237
		0.99	250	0.2204	0.2182	0.2268	0.2201	0.2162	0.2163
			500	0.1646	0.1617	0.1677	0.1644	0.1628	0.1624
	t	0.975	250	0.3051	0.3068	0.3160	0.3052	0.3087	0.3081
			500	0.2119	0.2128	0.2166	0.2118	0.2136	0.2133
		0.99	250	0.4510	0.4633	0.5146	0.4508	0.4689	0.4662
			500	0.3403	0.3418	0.3617	0.3401	0.3467	0.3453
	skt	0.975	250	0.2417	0.2428	0.2497	0.2418	0.2442	0.2437
			500	0.1678	0.1684	0.1713	0.1677	0.1690	0.1688
		0.99	250	0.3537	0.3628	0.4020	0.3535	0.3669	0.3649
			500	0.2667	0.2676	0.2830	0.2665	0.2714	0.2703
SC	norm	0.975	250	0.2562	0.2512	0.2480	0.2559	0.2498	0.2501
			500	0.1947	0.1923	0.1912	0.1946	0.1918	0.1919
		0.99	250	0.3154	0.3020	0.2887	0.3150	0.2955	0.2970
			500	0.2426	0.2337	0.2312	0.2424	0.2322	0.2325
	t	0.975	250	0.4376	0.4352	0.4362	0.4374	0.4351	0.4351
			500	0.3357	0.3342	0.3351	0.3356	0.3343	0.3342
		0.99	250	0.6299	0.6208	0.6210	0.6294	0.6161	0.6168
			500	0.4977	0.4885	0.4925	0.4973	0.4884	0.4884
	skt	0.975	250	0.3469	0.3447	0.3451	0.3467	0.3445	0.3445
			500	0.2658	0.2645	0.2650	0.2657	0.2644	0.2644
		0.99	250	0.4947	0.4869	0.4863	0.4944	0.4830	0.4836
			500	0.3902	0.3828	0.3856	0.3900	0.3826	0.3826

Table 8.8: MAE of vt-S-vine with symmetric linear v-transformation with Gaussian copula (Ga) and survival Clayton copulas (SC). The parameters in ARMA(1,1) estimation part are $\phi_1 = 0.95$ and $\theta_1 = -0.85$. *norm* is the standard normal distribution. t represents student t distribution with degree of freedom, $\nu = 6$. *skt* is the skewed student t distribution with degree of freedom 6 and skew parameter equals to 0.8.

				methods					
copula	dist	α	n	1	2	3	4	5	6
Ga	norm	0.975	250	-0.0446	-0.0123	0.0210	-0.0431	-0.0012	-0.0040
			500	-0.0257	-0.0095	0.0065	-0.0249	-0.0042	-0.0055
		0.99	250	-0.0835	-0.0210	0.0621	-0.0822	0.0067	-0.0003
			500	-0.0476	-0.0129	0.0210	-0.0469	-0.0016	-0.0044
	t	0.975	250	-0.0546	0.0020	0.0612	-0.0521	0.0217	0.0168
			500	-0.0338	-0.0059	0.0221	-0.0324	0.0035	0.0011
		0.99	250	-0.1268	0.0035	0.1995	-0.1242	0.0688	0.0525
			500	-0.0750	-0.0003	0.0728	-0.0735	0.0240	0.0179
	skt	0.975	250	-0.0445	0.0003	0.0471	-0.0425	0.0159	0.0120
			500	-0.0273	-0.0052	0.0169	-0.0262	0.0021	0.0003
		0.99	250	-0.1009	0.0013	0.1542	-0.0989	0.0523	0.0395
			500	-0.0595	-0.0011	0.0561	-0.0583	0.0180	0.0132
SC	norm	0.975	250	-0.0946	-0.0670	-0.0393	-0.0933	-0.0578	-0.0601
			500	-0.0475	-0.0322	-0.0165	-0.0468	-0.0270	-0.0283
		0.99	250	-0.1765	-0.1220	-0.0537	-0.1754	-0.0992	-0.1049
			500	-0.1034	-0.0713	-0.0398	-0.1027	-0.0608	-0.0634
	t	0.975	250	-0.1177	-0.0696	-0.0205	-0.1155	-0.0532	-0.0573
			500	-0.0554	-0.0289	-0.0011	-0.0541	-0.0197	-0.0220
		0.99	250	-0.2809	-0.1711	-0.0197	-0.2787	-0.1206	-0.1332
			500	-0.1635	-0.0954	-0.0286	-0.1622	-0.0731	-0.0787
	skt	0.975	250	-0.0956	-0.0575	-0.0187	-0.0939	-0.0446	-0.0478
			500	-0.0453	-0.0243	-0.0023	-0.0442	-0.0170	-0.0188
		0.99	250	-0.2230	-0.1367	-0.0183	-0.2213	-0.0972	-0.1071
			500	-0.1297	-0.0764	-0.0240	-0.1287	-0.0589	-0.0633

Table 8.9: BIAS of vt-S-vine with symmetric linear v-transformation with Gaussian copula (Ga) and survival Clayton copulas (SC). The parameters in ARMA(1,1) estimation part are $\phi_1 = 0.95$ and $\theta_1 = -0.85$. *norm* is the standard normal distribution. t represents student t distribution with degree of freedom, $\nu = 6$. *skt* is the skewed student t distribution with degree of freedom 6 and skew parameter equals to 0.8.

			methods					
DGP	dist	levels	1	2	3	4	5	6
Ga	norm	5	0.2610	0.2110	0.1350	0.2600	0.1785	0.1835
		10	0.0285	0.0170	0.0095	0.0270	0.0140	0.0140
		13	0.0040	0.0030	0.0015	0.0040	0.0030	0.0030
	t	5	0.2600	0.2110	0.1320	0.2580	0.1770	0.1825
		10	0.0265	0.0170	0.0095	0.0265	0.0140	0.0140
		13	0.0040	0.0030	0.0015	0.0040	0.0025	0.0030
	skt	5	0.2600	0.2110	0.1320	0.2580	0.1770	0.1825
		10	0.0265	0.0170	0.0095	0.0265	0.0140	0.0140
		13	0.0040	0.0030	0.0015	0.0040	0.0025	0.0030
SC	norm	5	0.4280	0.3700	0.2895	0.4260	0.3455	0.3520
		10	0.1085	0.0855	0.0525	0.1075	0.0725	0.0765
		13	0.0420	0.0345	0.0175	0.0415	0.0265	0.0295
	t	5	0.4255	0.3700	0.2880	0.4250	0.3435	0.3520
		10	0.1085	0.0855	0.0520	0.1075	0.0710	0.0765
		13	0.0420	0.0345	0.0175	0.0415	0.0265	0.0295
	skt	5	0.4255	0.3700	0.2880	0.4250	0.3435	0.3520
		10	0.1085	0.0855	0.0520	0.1075	0.0720	0.0765
		13	0.0420	0.0345	0.0175	0.0415	0.0265	0.0295

Table 8.10: The probabilities of exceedances over levels 5, 10, 13 in simulations from vt-S-vine with symmetric linear v-transformation with Gaussian copula (Ga) and survival Clayton copulas (SC). *norm* is the standard normal distribution. t represents student t distribution with degree of freedom, $\nu = 6$. *skt* is the skewed student t distribution with degree of freedom 6 and skew parameter equals to 0.8.

an identical mean to the independent case ($n \times (1-\alpha)$) but it will tend to have a higher variance and a heavier tail. Hence, the number of exceedances of the true α-quantile in a sequence of n values from a vt-S-vine process does not follow a binomial distribution. Hence, we can not calculate the cumulative probabilities in traffic light system. Instead we compare the exceedances (N) with the levels and repeat the simulations for 1000 times and count the probabilities $\mathbb{P}(N \geq 5)$, $\mathbb{P}(N \geq 10)$ and $\mathbb{P}(N \geq 13)$. The results are presented in Table 8.10.

The probabilities in Table 8.10 are all larger than the ones in iid case, which means the probabilities of getting the amber, red zone or having internal models disallowed by traffic light system are higher in serial dependent losses than iid losses. Furthermore, the vt-S-vine with survival Clayton copulas yields larger probabilities than the one with Gaussian copulas, which embodies that the quantile estimators have more possibilities to obtain the exceedances in the amber or red zone, even with the internal model being refused by the regulation. The distributions do not exert great influence on the probabilities in vt-S-vine process as well. Besides, the Method 3 gives the smallest

violating probabilities in traffic light system in both Gaussian and survival Clayton copulas, which is good for banks, because they prefer less violations in practice.

8.3 The distributions of exceedances

In this section, we study the distributions of exceedances for vt-S-vine models. It is simple to calculate the expectation of exceedances for α-quantiles in vt-S-vine models. Suppose the sample size is n and let $u = 1-\alpha$, the number of exceedances for the left-tail and right-tail can be expressed by $N_l = \sum_{i=1}^{n} I_{\{L_i \leq F_L^{-1}(u)\}}$ and $N_r = \sum_{i=1}^{n} I_{\{L_i \geq F_L^{-1}(\alpha)\}}$, respectively. Then, the expectations of exceedances of the left and right tail can be written as $\mathbb{E}(N_l) = nu$ and $\mathbb{E}(N_r) = n(1-\alpha)$. The variance of the exceedaneces is calculated in Proposition 8.

Proposition 8. *The variance of the number of exceedances of the α-quantile in a vt-S-vine sample with linear v-transform of length n is given by*

$$var(N_l) = nu - (nu)^2 + 2\sum_{k=1}^{n-1}(n-k)C_{(k)}(u,u) \tag{8.10}$$

and

$$var(N_r) = n(1-\alpha) - (n(1-\alpha))^2 + 2\sum_{k=1}^{n-1}(n-k)\hat{C}_{(k)}(\alpha,\alpha) \tag{8.11}$$

where $u = 1-\alpha$, $C_{(k)}(u,u)$ is the joint distribution that denotes the probability that U_t and U_{t+k} jointly take values in a lower orthant defined by the point (u,u). $\hat{C}_{(k)}$ denotes the survival copula of $C_{(k)}(u,u)$, which the expression can be found in Equation 2.2.

Proof. We argue that

$$\begin{aligned}
\text{var}(N_l) &= \text{var}\left(\sum_{i=1}^{n} I_{\{L_i \leq F_L^{-1}(u)\}}\right) = \text{var}\left(\sum_{i=1}^{n} I_{\{U_i \leq u\}}\right) \\
&= \mathbb{E}\left((\sum_{i=1}^{n} I_{\{U_i \leq u\}})^2\right) - \mathbb{E}\left(\sum_{i=1}^{n} I_{\{U_i \leq u\}})\right)^2 \\
\mathbb{E}\left(\sum_{i=1}^{n} I_{\{U_i \leq u\}}\right) &= nu \\
\mathbb{E}\left((\sum_{i=1}^{n} I_{\{U_i \leq u\}})^2\right) &= \mathbb{E}\left(\sum_{i=1}^{n} I^2_{\{U_i \leq u\}} + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} I_{\{U_i \leq u\}} I_{\{U_j \leq u\}}\right) \\
&= \sum_{i=1}^{n} \mathbb{E}\left((I_{\{U_i \leq u\}})^2\right) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathbb{E}\left(I_{\{U_i \leq u\}} I_{\{U_j \leq u\}}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \mathbb{P}\left(U_i \leq u\right) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathbb{P}\left(U_i \leq u, U_j \leq u\right) \\
&\texttt{let } j = i+k \\
&= nu + 2\sum_{i=1}^{n-k}\sum_{k=1}^{n-1} \mathbb{P}\left(U_i \leq u, U_{i+k} \leq u\right) \\
&= nu + 2\sum_{k=1}^{n-1}(n-k)\mathbb{P}\left(U_i \leq u, U_{i+k} \leq u\right) \\
&= nu + 2\sum_{k=1}^{n-1}(n-k)C_{(k)}(u,u) \\
\mathrm{var}(N_l) &= nu + 2\sum_{k=1}^{n-1}(n-k)C_{(k)}(u,u) - (nu)^2 \\
&= nu - (nu)^2 + 2\sum_{k=1}^{n-1}(n-k)C_{(k)}(u,u).
\end{aligned}
$$

Similarly, the expectation and variance of right tail exceedances N_r is

$$
\begin{aligned}
\mathrm{var}(N_r) &= \mathrm{var}\left(\sum_{i=1}^{n} I_{\{L_i \geq F_L^{-1}(\alpha)\}}\right) = \mathrm{var}\left(\sum_{i=1}^{n} I_{\{U_i \geq \alpha\}}\right) \\
&= \mathbb{E}\left((\sum_{i=1}^{n} I_{\{U_i \geq \alpha\}})^2\right) - \mathbb{E}\left(\sum_{i=1}^{n} I_{\{U_i \geq \alpha\}})\right)^2 \\
\mathbb{E}\left(\sum_{i=1}^{n} I_{\{U_i \geq \alpha\}}\right) &= n(1-\alpha) \\
\mathbb{E}\left((\sum_{i=1}^{n} I_{\{U_i \geq \alpha\}})^2\right) &= \mathbb{E}\left(\sum_{i=1}^{n} I^2_{\{U_i \geq \alpha\}} + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} I_{\{U_i \geq \alpha\}} I_{\{U_j \geq \alpha\}}\right) \\
&= \sum_{i=1}^{n} \mathbb{P}\left(U_i \geq \alpha\right) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathbb{P}\left(U_i \geq \alpha, U_j \geq \alpha\right) \\
&\texttt{let } j = i+k \\
&= n(1-\alpha) + 2\sum_{i=1}^{n-k}\sum_{k=1}^{n-1} \mathbb{P}\left(U_i \geq \alpha, U_{i+k} \geq \alpha\right) \\
&= n(1-\alpha) + 2\sum_{i=1}^{n-k}\sum_{k=1}^{n-1} \mathbb{P}\left(1-U_i \leq 1-\alpha, 1-U_{i+k} \leq 1-\alpha\right) \\
&= n(1-\alpha) + 2\sum_{k=1}^{n-1}(n-k)\hat{C}_{(k)}(1-\alpha, 1-\alpha)
\end{aligned}
$$

$$\begin{aligned}\text{var}(N_r) &= n(1-\alpha) + 2\sum_{k=1}^{n-1}(n-k)\hat{C}_{(k)}(1-\alpha, 1-\alpha) - (n(1-\alpha))^2 \\ &= n(1-\alpha) - (n(1-\alpha))^2 + 2\sum_{k=1}^{n-1}(n-k)\hat{C}_{(k)}(1-\alpha, 1-\alpha) \\ &= nu - (nu)^2 + 2\sum_{k=1}^{n-1}(n-k)\hat{C}_{(k)}(u,u).\end{aligned}$$

□

Two types of vt-S-vine are studied in previous section, one with Gaussian copulas, the other with survival Clayton copulas. We keep using the linear v-transform in this section to simplify the calculation. When the copula sequences are Gaussian copulas, the expression $C_{(k)}(u,u)$ in Equation 8.10 can be computed using the method proposed by Bladt and McNeil [2020], which is

$$C_{(k)}(u,u) = \delta^2 C^{Ga}_{\rho(k)}\left(\frac{u}{\delta}, \frac{u}{\delta}\right), u \leq \delta, \tag{8.12}$$

where δ is the fulcrum of the linear v-transformation and $\rho(k)$ is the autocorrelation of Gaussian process with copula sequences $C_{(k)}$. The proof of Equation 8.12 can be found in McNeil [2020]. The left and right tail of the exceedance distribution in vt-S-vine with Gaussian copulas are exactly the same, since Gaussian copula is symmetric. However, for survival Clayton copula, they will be distinctive.

In order to understand the distribution of exceedances in vt-S-vine, we model the exceedances by different distributions. In stationary vt-S-vine processes, the quantile exceedance indicator variables for the α-quantile are identically distributed Bernoulli variables. If these Bernoulli variables are independent, the sum of them will follow a binomial distribution. The beta-binomial distribution is a common model for the sum of dependent Bernoulli distributions (see Yu and Zelterman [2002]). Therefore, we inference that the distribution of exceedances in vt-S-vine may be estimated properly by the beta-binomial distribution.

Therefore, we fit beta-binomial distribution to the number of exceedances of quantiles in vt-S-vine models with Gaussian and survival Clayton copulas and symmetric linear v-transform. We apply the normal distribution as margin for the vt-S-vines. Therefore, the VaR of time series, expectations and variance of the number of exceedances in vt-S-vine are calculated.

The method used to fit beta-binomial distributions to the number of exceedances at different quantile levels in a vt-S-vine model with normal margin and Gaussian copulas is:

Step 1 : Calculate the variance and expectation of the number of exceedances of 0.025-

quantile and 0.01-quantile in vt-S-vine models. The variance is computed by the Equation 8.10 and the copula $C_{(k)}(u,u)$ is calculated by the Gaussian copula in Equation 8.12. The expectation is equal to nu.

Step 2 : Compute the parameters of beta-binomial distribution by method of moments, setting the number of trials equal to n, where n is 250 or 500.
$a = \frac{(\mu^2(n-\mu)-\mu\sigma^2)}{(n\sigma^2-\mu n+\mu^2)}$, $b = (\frac{n}{\mu}-1)a$
where μ and σ^2 are the expectation and variance calculated in Step 1.

Step 3 : Simulate n data from vt-S-vine models, where n is equal to 250 and 500, respectively.

Step 4 : Calculate the VaR in 0.025 and 0.01 levels where $VaR_u = \Phi^{-1}(u)$, because the data generated from vt-S-vine models are normal distribution.

Step 5 : Compare the simulated time series in Step 2 with the VaR and count the number of exceedances of different quantiles. Repeat the Step 2 and Step 4 for 1000 times and we will obtain 1000 exceedances..

Step 6 : Plot the density of the number of exceedances of quantiles and density function of beta-binomial distribution with computed parameters in every circumstance.

The steps above are fitting the beta-binomial distribution to the number of exceedances for the left tail in vt-S-vine with Gaussian copulas. The distribution of exceedances for the left tail is exactly the same as right tail in Gaussian copula. Hence, we only present the fitting process of left-tail exceedances in Gaussian copula. The approach used for vt-S-vine with survival Clayton copulas is very similar to the one with Gaussian copulas. However, $C_{(k)}(u,u)$ in Equation 8.10 can not be calculated, since the copula between U_t and U_{t+k} is unknown. The vt-S-vine with survival Clayton copulas have different properties to Gaussian copulas, the copula for two variables that are not adjacent may not be survival Clayton. Hence, we use the empirical variance instead. The variance is calculated from the 1000 exceedances from Step 5. The other steps are exactly the same as vt-S-vine with Gaussian copulas. The left tail and right tail of the vt-S-vine with survival Clayton copulas are distinctive, so we study them separately.

The density plots of exceedances in the vt-S-vine with symmetric linear v-transform and Gaussian copulas are shown in Figure 8.1. Figure 8.2 and Figure 8.3 exhibit the density plots of exceedances in the left and right tail of vt-S-vine with symmetric linear v-transform and survival Clayton copulas, respectively. The red line is the estimated beta-binomial distributions, which are very close to the density of exceedances of simulations from vt-S-vine in both figures. The exceedances in models with Gaussian copulas are estimated more accurately by beta-binomial distribution, compared to survival Clayton

p-values of KS test				
α	n	Gaussian	Clayton180 (left)	Clayton180 (right)
0.975	250	0.9937	0.4324	0.1995
	500	0.7591	0.0149	0.0225
0.99	250	0.8879	0.0112	0.0012
	500	0.9883	0.0692	0.0869

Table 8.11: P-values of Kolmogorov-Smirnov test for exceedances of simulations from vt-S-vine models with Gaussian and survival Clayton copulas. The "left" and "right" in bracket means the left and right tail of vt-S-vine with survival Clayton copulas, respectively. The null hypothesis of the test is that the exceedances and the data simulated from beta-binomial distribution are drawn from the same continuous distribution.

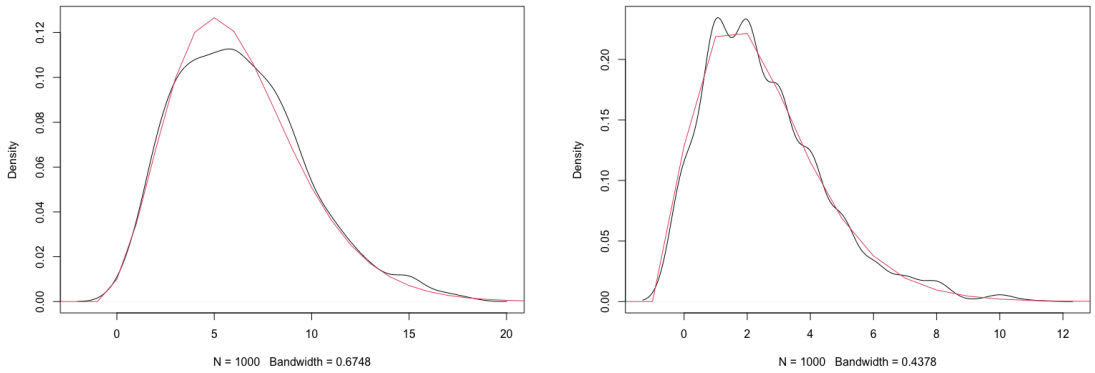
copulas. The peak of the density in both left and right tail of vt-S-vine with survival Clayton copulas are not modelled well by beta-binomial distribution.

In order to ensure the exceedances of vt-S-vines can be approximated by beta-binomial distributions, the Kolmogorov-Smirnov test is applied to test the exceedances. The p-values of the test for both vt-S-vine with Gaussian and survival Clayton copulas are presented in Table 8.11.

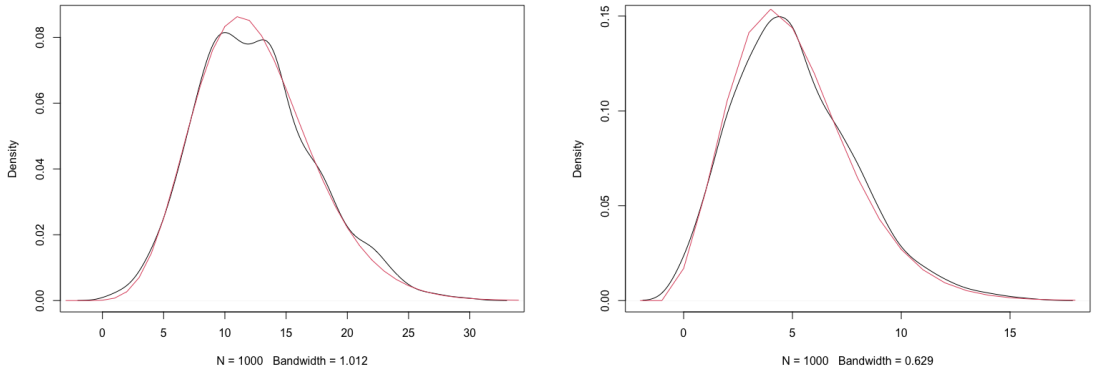
According to Table 8.11, all the p-values in Gaussian copulas accord with the standard if we set the significance level to 0.01. Hence, there is very little evidence against the null hypothesis in Gaussian copulas. The exceedances of vt-S-vine with Gaussian copulas can be approximated by beta-binomial distributions at both 0.975 and 0.99 quantiles in both 250 and 500 sample size. However, the p-value of the right tail of vt-S-vine with survival Clayton copulas in 250 sample size at 0.99 quantile is smaller than 0.01. Thus, the beta-binomial distribution may not be suitable for the right tail of vt-S-vine with survival Clayton copulas at 0.99 quantile in 250 sample size. Furthermore, the p-values of both tails in survival Clayton copulas are much smaller than the ones of Gaussian copulas, which exemplifies that the beta-binomial distribution may estimate the exceedances better in vt-S-vine with Gaussian copula sequences.

8.4 Conclusions

In this chapter, we investigated the application of vt-S-vine models in trading book of bank. The first point we find is that the accuracy of empirical quantile estimates of VaR decreases when the losses have serial dependence. Moreover, the six quantile estimate methods yield exceedances closer to the true values 0.025 and 0.01 in the vt-S-vine with Gaussian copulas than in the vt-S-vine with survival Clayton copulas. Meanwhile, the empirical quantile estimates of VaR by the six approaches are worse for non-Gaussian models, compared to the Gaussian models. Furthermore, we find the distributions do not exert remarkably influence on the exceedance probabilities, while the methods of estimating VaR affect the estimation of exceedances and VaR remarkably. Generally,

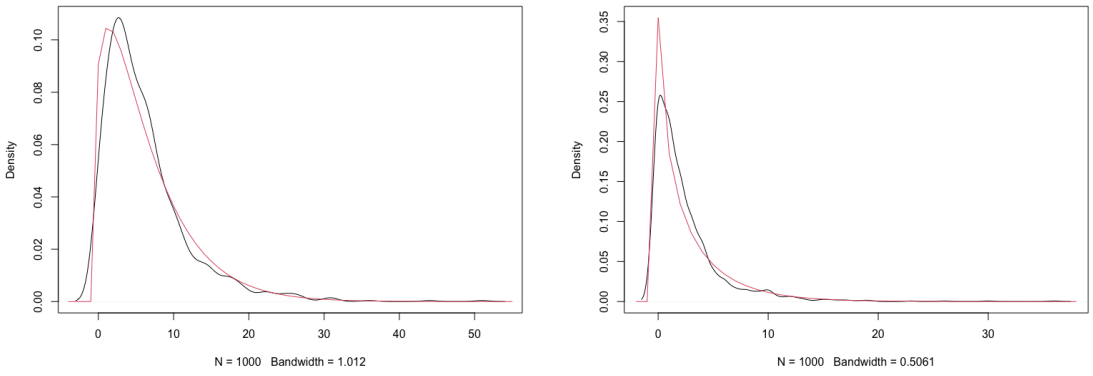


(a) Sample size is 250.



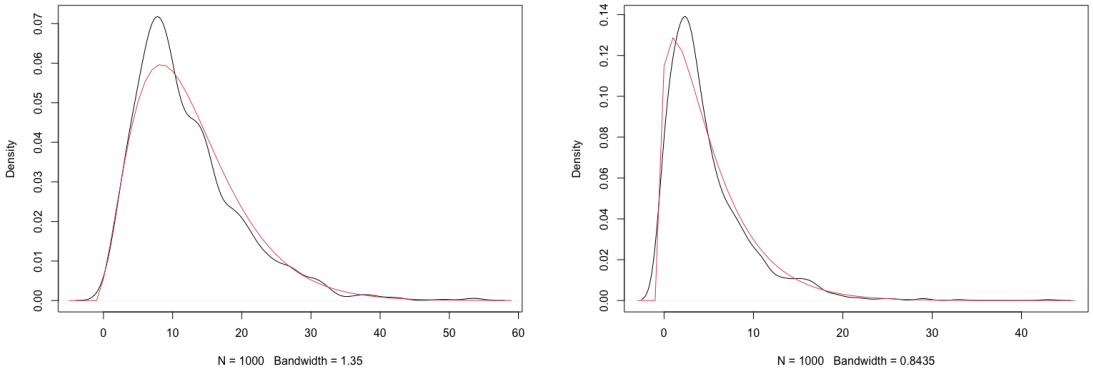
(b) Sample size is 500.

Figure 8.1: The density plot of exceedances in vt-S-vine models with Gaussian copulas and symmetric linear v-transform. The number of simulation in each experiment is 1000 and the number of trail in beta-binomial distribution is equal to sample size, 250 or 500. The two figures on the left are the ones at quantile 0.975. The two on the right are the ones at quantile 0.99. The black line is the density of exceedances in vt-S-vine simulations. The red line is the beta-binomial distributions estimated by the method of moments.



,

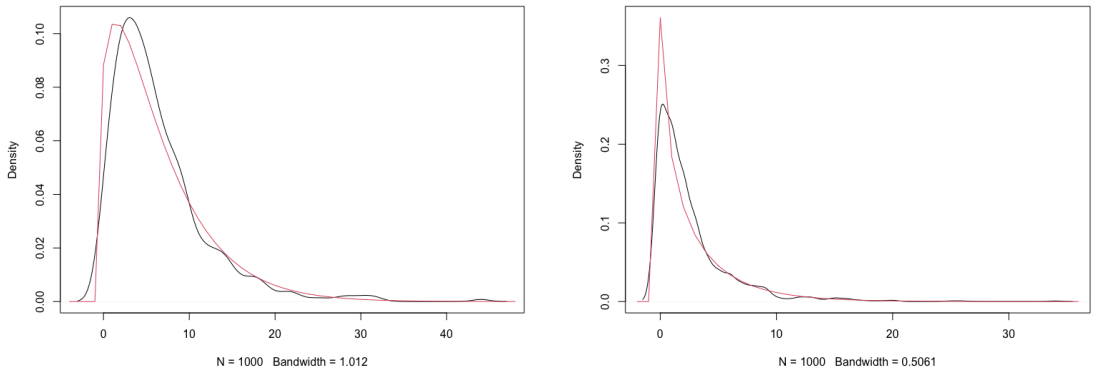
(a) Sample size is 250.



,

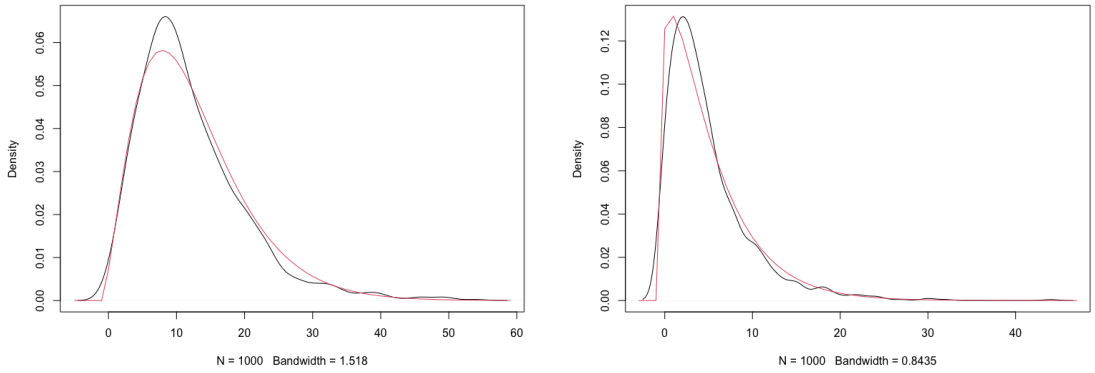
(b) Sample size is 500.

Figure 8.2: The density plot of exceedances in the left tail of vt-S-vine models with survival Clayton copulas and symmetric linear v-transform. The number of simulation in each experiment is 1000 and the number of trail in beta-binomial distribution is equal to sample size, 250 or 500. The two figures on the left are the ones at quantile 0.975. The two on the right are the ones at quantile 0.99. The black line is the density of exceedances in vt-S-vine simulations. The red line is the beta-binomial distributions estimated by the method of moments.



,

(a) Sample size is 250.



,

(b) Sample size is 500.

Figure 8.3: The density plot of exceedances in the right tail of vt-S-vine models with survival Clayton copulas and symmetric linear v-transform. The number of simulation in each experiment is 1000 and the number of trail in beta-binomial distribution is equal to sample size, 250 or 500. The two figures on the left are the ones at quantile 0.975. The two on the right are the ones at quantile 0.99. The black line is the density of exceedances in vt-S-vine simulations. The red line is the beta-binomial distributions estimated by the method of moments.

the third method yields the best exceedances estimation in both iid and vt-S-vine cases. In most cases, method 3 overstates VaR, which means it is a conservative approach.

Another counterintuitive finding is that the quantile estimator is closest on average to the true value of the quantile of a distribution may not be the one gives the most accurate value for the exceedance probability. Since regulation is based on numbers of exceedances and not closeness to the unknown true value of VaR, this gives an indicator as to the method of empirical quantile estimation that banks should prefer.

The exceedances of vt-S-vine models are estimated by beta-binomial distributions. It is surprising to find that the exceedances from vt-S-vine with Gaussian copulas can be modelled so accurately by beta-binomial distributions. The exceedances of processes with survival Clayton copulas can also be modelled by beta-binomial distributions, but not for all quantiles and sample sizes. Moreover, the distribution of exceedances for both left and right tail of the vt-S-vine with survival Clayton copulas have similar densities, despite the asymmetry in survival Clayton copula. The fit of beta-binomial distributions to exceedances in survival Clayton copulas is not as good as Gaussian copulas, especially in the peak of the density.

Summary and Conclusions

This thesis is devoted to the exploration of the properties and applications of S-vine models and vt-S-vine models in macroeconomics, finance and banking regulation. There are three topics in this thesis. The first one is modelling and predicting the inflation rates by S-vine models. We found that the S-vine models with non-Gaussian copula sequences, marginal distributions and non-linear structures can outperform ARMA models in most cases. Rotating Gumbel, Joe or Clayton copulas that can only describe the positive dependence in time series can optimize the estimation and prediction results of S-vine models. Besides, the simplification of S-vine can reduce the number of parameters and lead to smaller AIC values, which improve the S-vine models in estimations. For prediction, the average quantile scores may not find the best models for forecasting in the data with small sample size. Therefore, it is possible to attempt various weights at different quantiles.

The second topic of this thesis is the vt-S-vine models applied to modelling and predicting volatile time series. We use vt-S-vine models to mimic GARCH type processes. Generally, the vt-S-vine processes with survival Clayton copula sequences obtain the smallest AIC values among the candidate copulas in both parametric and semi-parametric methods. Moreover, the vt-S-vine used in this thesis can compete with the lower-order S-vine models with t copulas, which are proposed by Zhao et al. [2022]. The higher-order vt-S-vine models with proper copula sequences has an advantage over the lower-order S-vine models with t copulas in forecasting GARCH type processes. In the empirical study of the second topic, the vt-S-vines can surpass GARCH or GJR-GARCH processes in most cases. However, the EGARCH processes yield better estimation results than vt-S-vine sometimes, while the transformation of the marginal distributions of the time series can reverse the results. This phenomenon implies that the better marginal distribution may enhance the estimations of vt-S-vine models.

The last topic seeks to understand the distribution of quantile exceedances in volatile data generated from vt-S-vines. This has an application to the "traffic light system" used in the regulation of banks' trading books. We found the best approach of quantile estimation in vt-S-vine models in different cases according to VaR exceedances and VaR estimations. Typically, the copula selection of vt-S-vine models exerts great influence

on the VaR exceedances and estimations. However, we find the marginal distributions do not exert remarkably influence on the exceedance probabilities. Moreover, a counterintuitive finding is that the quantile estimator is closest on average to the true value of the quantile of a distribution may not be the one that tends to give the most accurate value for the exceedance probability. Since regulation is based on numbers of exceedances and not closeness to the unknown true value of VaR, this gives an indicator as to the method of empirical quantile estimation that banks should prefer.

In conclusion, S-vine and vt-S-vine models discussed in this thesis are flexible, efficient and competitive models for many kinds of macroeconomic and financial time series. They can compete with and often surpass certain classical stochastic models, such as the family of ARMA and GARCH processes. The vt-S-vines can even mimic GARCH type processes to some extent. The estimation and prediction by using S-vines or vt-S-vines may be improved by exploring more detailed marginal models and more flexible pair copula choices. More advanced approaches of backtesting prediction results from S-vine or vt-S-vine models will be developed in the further studies.

Bibliography

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

D. Acemoglu, A. Ozdaglar, and A. Tahbaz-Salehi. Microeconomic origins of macroeconomic tail risks. *American Economic Review*, 107(1):54–108, 2017.

T. W. Anderson. An introduction to multivariate statistical analysis, john wiley & sons. *Inc., New York*, 54, 1958.

J. Arlt. The problem of annual inflation rate indicator. *International Journal of Finance & Economics*, 2021.

J. Arteche. The analysis of seasonal long memory: The case of spanish inflation. *Oxford Bulletin of Economics and Statistics*, 69(6):749–772, 2007.

O. Barndorff-Nielsen and G. Schou. On the parametrization of autoregressive models by partial autocorrelations. *Journal of multivariate Analysis*, 3(4):408–419, 1973.

G. Barone-Adesi, R. F. Engle, and L. Mancini. A garch option pricing model with filtered historical simulation. *The review of financial studies*, 21(3):1223–1258, 2008.

E. Bataa, D. R. Osborn, M. Sensier, and D. v. Dijk. Identifying changes in mean, seasonality, persistence and volatility for g7 and euro area inflation. *Oxford Bulletin of Economics and Statistics*, 76(3):360–388, 2014.

B. K. Beare. Copulas and temporal dependence. *Econometrica*, 78(1):395–410, 2010.

B. K. Beare and J. Seo. Vine copula specifications for stationary multivariate markov chains. *Journal of Time Series Analysis*, 36(2):228–246, 2015.

T. Bedford and R. M. Cooke. Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial intelligence*, 32(1):245–268, 2001.

T. Bedford and R. M. Cooke. Vines–a new graphical model for dependent random variables. *The Annals of Statistics*, 30(4):1031–1068, 2002.

W. P. Bergsma. *Testing conditional independence for continuous random variables*. Citeseer, 2004.

M. Bladt and A. J. McNeil. Time series copula models using d-vines and v-transforms: an alternative to garch modelling. *arXiv preprint arXiv:2006.11088*, 2020.

M. Bladt and A. J. McNeil. Time series copula models using d-vines and v-transforms. *Econometrics and Statistics*, 2021.

M. Bladt and A. J. McNeil. Time series with infinite-order partial copula dependence. *Dependence Modeling*, 10(1):87–107, 2022.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

G. E. Box, G. M. Jenkins, and G. Reinsel. Time series analysis: forecasting and control holden-day san francisco. *BoxTime Series Analysis: Forecasting and Control Holden Day1970*, 1970.

E. C. Brechmann and C. Czado. Copar—multivariate time series modeling using the copula autoregressive model. *Applied Stochastic Models in Business and Industry*, 31 (4):495–514, 2015.

P. J. Brockwell and R. A. Davis. *Time series: theory and methods*. Springer, New York, 2nd edition, 1991.

C. Brooks. *Introductory Econometrics for Finance*. Cambridge University Press, 3 edition, 2014.

K. P. Burnham, D. R. Anderson, and K. P. Huyvaert. Aic model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral ecology and sociobiology*, 65(1):23–35, 2011.

X. Chen and Y. Fan. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of econometrics*, 135(1-2):125–154, 2006.

R. B. Cleveland, W. S. Cleveland, J. E. McRae, and I. Terpenning. Stl: A seasonal-trend decomposition. *J. Off. Stat*, 6(1):3–73, 1990.

R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative finance*, 1(2):223, 2001.

R. M. Cooke. Markov and entropy properties of tree-and vine-dependent variables. In *Proceedings of the ASA section of Bayesian statistical science*, volume 27, 1997.

I. P. Cornfeld, S. V. Fomin, and Y. G. Sinai. *Ergodic theory*, volume 245. Springer Science & Business Media, 2012.

K. Cuthbertson. Introductory econometrics for finance, chris brooks, cambridge university press, cambridge, 2002. *International Journal of Finance & Economics*, 9(1):82, 2004.

C. Czado. Analyzing dependent data with vine copulas. *Lecture Notes in Statistics, Springer*, 2019.

E. B. Dagum and S. Bianconcini. *Seasonal adjustment methods and real time trend-cycle estimation*. Springer, 2016.

W. F. Darsow, B. Nguyen, E. T. Olsen, et al. Copulas and markov processes. *Illinois journal of mathematics*, 36(4):600–642, 1992.

S. Demarta and A. J. McNeil. The t copula and related copulas. *International statistical review*, 73(1):111–129, 2005.

D. A. Dickey and W. A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a): 427–431, 1979.

F. X. Diebold, T. A. Gunther, and A. Tay. Evaluating density forecasts, 1997.

Z. Ding, C. W. Granger, and R. F. Engle. A long memory property of stock market returns and a new model. *Journal of empirical finance*, 1(1):83–106, 1993.

F. Domma, S. Giordano, and P. F. Perri. Statistical modeling of temporal dependence in financial data via a copula function. *Communications in Statistics-Simulation and Computation*, 38(4):703–728, 2009.

R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, pages 987–1007, 1982.

G. Fagiolo, M. Napoletano, and A. Roventini. Are output growth-rate distributions fat-tailed? some evidence from oecd countries. *Journal of Applied Econometrics*, 23 (5):639–669, 2008.

C. Fernández and M. F. Steel. On bayesian modeling of fat tails and skewness. *Journal of the american statistical association*, 93(441):359–371, 1998.

M. J. Frank. On the simultaneous associativity off (x, y) andx+ y- f (x, y). *Aequationes mathematicae*, 19(1):194–226, 1979.

P. H. Franses and D. Van Dijk. Forecasting stock market volatility using (non-linear) garch models. *Journal of forecasting*, 15(3):229–235, 1996.

G. A. Fredricks and R. B. Nelsen. On the relationship between spearman's rho and kendall's tau for pairs of continuous random variables. *Journal of statistical planning and inference*, 137(7):2143–2150, 2007.

C. Genest, K. Ghoudi, and L.-P. Rivest. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3): 543–552, 1995.

I. Gijbels, M. Omelka, and N. Veraverbeke. Partial and average copulas and association measures. *Electronic Journal of Statistics*, 9(2):2420–2474, 2015.

L. R. Glosten, R. Jagannathan, and D. E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801, 1993.

T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, 102(477):359–378, 2007.

I. H. Haff, K. Aas, and A. Frigessi. On the simplified pair-copula construction—simply useful or too simplistic? *Journal of Multivariate Analysis*, 101(5):1296–1310, 2010.

S. A. Hamid and T. S. Dhakar. The behaviour of the us consumer price index 1913–2003: a study of seasonality in the monthly us cpi. *Applied Economics*, 40(13):1637–1650, 2008.

A. Harvey and G. Sucarrat. Egarch models with fat tails, skewness and leverage. *Computational Statistics & Data Analysis*, 76:320–338, 2014.

U. Hassler and J. Wolters. Long memory in inflation rates: International evidence. *Journal of Business & Economic Statistics*, 13(1):37–45, 1995.

M. Hofert, I. Kojadinovic, M. Mächler, and J. Yan. *Elements of copula modeling with R*. Springer, 2018.

M. Hofert, R. Frey, and A. J. McNeil. The quantitative risk management exercise book, 2020.

J.-J. Huang, K.-J. Lee, H. Liang, and W.-F. Lin. Estimating value at risk of portfolio by conditional copula-garch method. *Insurance: Mathematics and economics*, 45(3): 315–324, 2009.

M. Huwiler and D. Kaufmann. Combining disaggregate forecasts for inflation: The snb's arima model: the snb's arima model. *Swiss National Bank Economic Study*, 7, 2013.

R. J. Hyndman and G. Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2018.

R. J. Hyndman and Y. Fan. Sample quantiles in statistical packages. *The American Statistician*, 50(4):361–365, 1996.

R. Ibragimov. On the robustness of economic models to heavy-tailedness assumptions. Technical report, Mimeo, Yale University. Available at http://post. economics. harvard. edu . . . , 2004.

R. Ibragimov. Copula-based characterizations for higher order markov processes. *Econometric Theory*, 25(3):819–846, 2009.

H. Joe. Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters. *Lecture Notes-Monograph Series*, pages 120–141, 1996.

H. Joe. *Multivariate models and multivariate dependence concepts*. CRC press, 1997.

H. Joe. Generating random correlation matrices based on partial correlations. *Journal of multivariate analysis*, 97(10):2177–2189, 2006.

H. Joe. *Dependence modeling with copulas*. CRC press, 2014.

H. Joe and T. Hu. Multivariate distributions from mixtures of max-infinitely divisible distributions. *Journal of multivariate analysis*, 57(2):240–265, 1996.

H. Joe, H. Li, and A. K. Nikoloulopoulos. Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101(1):252–270, 2010.

P. Katsiampa. Volatility estimation for bitcoin: A comparison of garch models. *Economics letters*, 158:3–6, 2017.

I. Kelikume and A. Salami. Time series modeling and forecasting inflation: Evidence from nigeria. *The International Journal of Business and Finance Research*, 8(2): 41–51, 2014.

S. K. Klutse. Inflation forecasting in developing economies using sarma models: The case of ghana. In *Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged*, pages 286–302, 2020.

D. Krukovets, O. Verchenko, et al. Short-run forecasting of core inflation in ukraine: a combined arma approach. *Visnyk of the National Bank of Ukraine*, 248:12–22, 2019.

D. Kurowicka and R. M. Cooke. *Uncertainty analysis with high dimensional dependence modelling*. John Wiley & Sons, 2006.

D. Kwiatkowski, P. C. Phillips, P. Schmidt, and Y. Shin. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1-3):159–178, 1992.

C. G. Lamoureux and W. D. Lastrapes. Heteroskedasticity in stock return data: Volume versus garch effects. *The journal of finance*, 45(1):221–229, 1990.

R. Loaiza-Maya, M. S. Smith, and W. Maneesoonthorn. Time series copulas for heteroskedastic data. *Journal of Applied Econometrics*, 33(3):332–354, 2018.

M. Marcellino, J. H. Stock, and M. W. Watson. A comparison of direct and iterated multistep ar methods for forecasting macroeconomic time series. *Journal of econometrics*, 135(1-2):499–526, 2006.

G. Maruyama. Infinitely divisible processes. *Theory of Probability & Its Applications*, 15(1):1–22, 1970.

A. J. McNeil. Modelling volatility with v-transforms. *arXiv preprint arXiv:2002.10135*, 2020.

A. J. McNeil. Modelling volatile time series with v-transforms and copulas. *Risks*, 9(1): 14, 2021.

A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: concepts, techniques and tools-revised edition*. Princeton university press, 2015.

K. Moriyama and A. Naseer. Forecasting inflation in sudan. 2009.

T. Nagler, D. Krüger, and A. Min. Stationary vine copula models for multivariate time series. *Journal of Econometrics*, 2022.

R. B. Nelsen. *An introduction to copulas*. Springer Science & Business Media, 2007.

D. B. Nelson. Stationarity and persistence in the garch (1, 1) model. *Econometric theory*, 6(3):318–334, 1990.

D. B. Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the econometric society*, pages 347–370, 1991.

A. K. Nikoloulopoulos, H. Joe, and H. Li. Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics & Data Analysis*, 56(11):3659–3673, 2012.

D. H. Oh and A. J. Patton. Simulated method of moments estimation for copula-based multivariate models. *Journal of the American Statistical Association*, 108(502):689–700, 2013.

B. C. on Banking Supervision. Minimum capital requirements for market risk. *Bank for international settlements*, 136, 2019.

D. R. Osborn and M. Sensier. Uk inflation: persistence, seasonality and monetary policy. *Scottish Journal of Political Economy*, 56(1):24–44, 2009.

J. H. Ospina and A. M. Padilla Ospina. Penalised regressions vs. autoregressive moving average models for forecasting inflation. *ECONÓMICAS CUC*, 2019.

A. R. Pagan and A. D. Hall. Diagnostic tests as residual analysis. *Econometric Reviews*, 2(2):159–218, 1983.

A. J. Patton. A review of copula models for economic time series. *Journal of Multivariate Analysis*, 110:4–18, 2012.

F. L. Ramsey. Characterization of the partial autocorrelation function. *The Annals of Statistics*, pages 1296–1301, 1974.

R. J. Rossi. *Mathematical statistics: an introduction to likelihood based inference*. John Wiley & Sons, 2018.

U. Schepsmeier. Maximum likelihood estimation of c-vine pair-copula constructions based on bivariate copulas from different families. 2010.

S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611, 1965.

P. Shi and L. Yang. Pair copula constructions for insurance experience rating. *Journal of the American Statistical Association*, 113(521):122–133, 2018.

J. H. Shih and T. A. Louis. Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, pages 1384–1399, 1995.

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

M. Smith, A. Min, C. Almeida, and C. Czado. Modeling longitudinal data using a pair-copula decomposition of serial dependence. *Journal of the American Statistical Association*, 105(492):1467–1479, 2010.

J. H. Stock and M. W. Watson. Why has us inflation become harder to forecast? *Journal of Money, Credit and banking*, 39:3–33, 2007.

J. H. Stock, M. W. Watson, et al. *Introduction to econometrics*, volume 104. Addison Wesley Boston, 2003.

K. Stoviček. Forecasting with arma models: The case of slovenian inflation. *Bank of Slovenia*, 2007.

E. Tully and B. M. Lucey. A power garch examination of the gold market. *Research in International Business and Finance*, 21(2):316–325, 2007.

S. X. Wei. A censored–garch model of asset returns with price limits. *Journal of Empirical Finance*, 9(2):197–223, 2002.

C. Yu and D. Zelterman. Sums of dependent bernoulli random variables and disease clustering. *Statistics & probability letters*, 57(4):363–373, 2002.

G. U. Yule and M. G. Kendall. An introduction to the theory of statistics. 14t ed., revised and enlarged. London: Charles Griffin & Co., Ltd. XXIV, 701 p. (1950)., 1950.

B. Zhang, J. C. Chan, and J. L. Cross. Stochastic volatility models with arma innovations: An application to g7 inflation forecasts. *International Journal of Forecasting*, 36(4):1318–1328, 2020.

Z. Zhao, P. Shi, and Z. Zhang. Modeling multivariate time series with copula-linked univariate d-vines. *Journal of Business & Economic Statistics*, 40(2):690–704, 2022.