International Statistical Review

International Statistical Review (2024) doi: 10.1111/insr.12562

Maximum Likelihood Estimation of Multivariate Regime Switching Student-*t* Copula Models

Federico P. Cortese[1], Fulvia Pennoni[2] and Francesco Bartolucci[3]

[1]*Department of Economics, Management and Statistics, University of Milano-Bicocca, Via Bicocca degli Arcimboldi, 8, Milan, 20128, Italy*
[2]*Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Via Bicocca degli Arcimboldi, 8, Milan, 20128, Italy*
[3]*Department of Economics, University of Perugia, Via A. Pascoli, 20, Perugia, 60123, Italy*
Corresponding to: Federico P. Cortese, University of Milano-Bicocca, Department of Economics, Management and Statistics, Italy. Email: federico.cortese@unimib.it

Summary

We propose a multivariate regime switching model based on a Student-*t* copula function with parameters controlling the strength of correlation between variables and that are governed by a latent Markov process. To estimate model parameters by maximum likelihood, we consider a two-step procedure carried out through the Expectation–Maximisation algorithm. To address the main computational burden related to the estimation of the matrix of dependence parameters and the number of degrees of freedom of the Student-*t* copula, we show a novel use of the Lagrange multipliers, which simplifies the estimation process. The simulation study shows that the estimators have good finite sample properties and the estimation procedure is computationally efficient. An application concerning log-returns of five cryptocurrencies shows that the model permits identifying bull and bear market periods based on the intensity of the correlations between crypto assets.

Key words: copula models; cryptocurrencies; daily log-returns; expectation–maximisation algorithm; latent variable models.

1 INTRODUCTION

Financial analysis and risk management research shows that the dependence structure of financial time-series changes during crises, with interdependence among assets increasing compared with stable periods (Das & Uppal, 2004; Patton, 2004). This phenomenon, known as *asymmetric dependence* (Ang & Bekaert, 2002; Ang & Chen, 2002; Longin & Solnik, 2002), is particularly relevant in cryptocurrency markets due to their vulnerability to changes in economic developments and news (Garcia & Ghysels, 1998; Kristoufek, 2013; Telli & Chen, 2020). For this reason, it is important to consider suitable specifications for the joint distribution of log-returns to capture possible sudden changes in market dynamics.

Extensive research supports the existence of regime switches in cryptocurrency returns, volatilities, and cross-correlation structure; see Ardia *et al.* (2019), Shen *et al.* (2020) and Cremaschini *et al.* (2023). Hamilton (1989) first argued that the switching dynamics of financial returns may be easily modelled through a Markovian process, see also the monograph by Frühwirth-Schnatter (2006), for further references and more details. In particular, regime switching (RS) copula models, also known as Markov-switching copulas (Jondeau & Rockinger, 2006; Rodriguez, 2007; Okimoto, 2008; Chollete *et al.*, 2009), accurately describe persistent correlation dynamics (Ang & Timmermann, 2012) in log-returns by modelling the joint distribution as a copula function that changes according to latent states. These models involve two stochastic processes: the first corresponds to the observed series, and the other to an underlying (latent) process describing the evolution of the hidden states over time. RS copula models are employed for exchange rates data, for the analysis of the correlation between S&P 500 and NASDAQ indexes, for the study of gold-oil dependence structure, and to describe momentum shifts in football matches (Stöber & Czado, 2014; Härdle *et al.*, 2015; Nasri & Rémillard, 2019; Tiwari *et al.*, 2020; Ötting *et al.*, 2021).

In the present paper, we propose a new model named RS Student-*t* copula (RS*St*C) model tailored to account for stylised facts of cryptocurrency returns, such as heavy-tailed distributions and non-linear dependencies. The model is based on a Student-*t* copula (Demarta & McNeil, 2005), which is parametrised by the number of degrees of freedom and the matrix of dependence parameters. Student-*t* copula is generally preferred to Gaussian copula for financial time-series because it allows the modelling of tail dependence and kurtosis (Breymann *et al.*, 2003; Fischer *et al.*, 2009; Huang *et al.*, 2009). Compared with Archimedean copulas (Genest *et al.*, 2011), which rely on a single dependence parameter for all variables, the proposed RS*St*C formulation offers superior accuracy in modelling the correlation structure. Additionally, while Vine copulas (Joe & Kurowicka, 2011; Czado & Nagler, 2022) are quite flexible to model different distributions, they have a more complex analytical form.

Maximum likelihood estimation of the RS copula parameters is typically performed through the Expectation–Maximisation (EM) algorithm (Dempster *et al.*, 1977). However, even in the case of a simple multidimensional Student-*t* copula model, Hernández *et al.* (2014) show that maximum likelihood estimation can be highly computationally inefficient. We provide a new approximation method for the EM algorithm tailored for estimating RS*St*C models. The proposal consists in an iterative procedure for estimating the matrix of dependence parameters and the number of degrees of freedom of the multivariate RS*St*C model. Following the approach of Trede (2020), developed for estimating a simple Student-*t* copula model, we maximise the log-likelihood function corresponding to the Student-*t* copula density in two steps. At the first step, we estimate the matrix of dependence parameters for a fixed number of degrees of freedom through Lagrange multipliers relying on a closed form solution; then, we numerically optimise the log-likelihood with respect to the number of degrees of freedom, keeping fixed the estimated matrix of dependence parameters. This procedure is simple, computationally feasible, and fast, even for long series with many assets. To evaluate the proposal, we rely on a simulation study assessing the good finite sample properties of the estimates and the computational efficiency of the procedure. An important feature of the proposal is that it can account for persistence in market regimes. This is a relevant aspect because, as suggested in Nystrup *et al.* (2020), when the state sequence contains several jumps, the RS model tends to a finite mixture model (McLachlan & Peel, 2000).

We apply the proposed approach to analyse log-returns of the five cryptocurrencies, Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), and Bitcoin Cash (BCH), for 5 years, from 17 September 2017 to 2 October 2022. At least to our knowledge, these data have never been analysed with RS copula models; for a recent review of the methods proposed in the literature for the analysis of multiple cryptoassets, see, among others, Koki *et al.* (2022). We select

the optimal number of latent states relying on the Integrated Completed Likelihood (ICL) criterion (Biernacki *et al.*, 2000) and, following Pennoni *et al.* (2021), we predict the latent regimes considering global decoding (Viterbi, 1967; Juang & Rabiner, 1991). Additionally, we compare the forecasting performance of the proposed model with that of a more common hidden Markov model (HMM, Zucchini *et al.*, 2017) and a multivariate random walk (MRW) process considered as benchmark.

To summarise, we provide three main contributions to the existing literature. First, we propose a multivariate RSS*t*C model, whereas most previous works focus only on the bivariate case. Second, we implement a novel and computationally efficient method for estimating the matrix of dependence parameters and the number of degrees of freedom. Finally, we show the applicability of the proposal analysing cryptocurrency log-returns in a novel way.

We implemented the code developed to carry out the estimation of the RSS*t*C model and to perform simulations in the `C++` language, through the `R` (R Core Team, 2023) package **Rcpp** (Eddelbuettel & François, 2011). The code is freely available at the following link: https://github.com/FedericoCortese/RSstcopula/find/main.

The remainder of the paper is organised as follows. In Section 2, we introduce the RSS*t*C model. In Section 3, we show the proposed procedure for maximum likelihood estimation of the model parameters, the initialisation strategy, and the convergence criterion chosen for the EM algorithm. In Section 4, we illustrate the simulation study aimed at assessing the validity of the proposed estimation procedure, and we comment on the results. In Section 5, we apply the proposal to analyse the daily log-returns of the five cryptocurrencies and present the results along with comparative analysis. In Section 6, we discuss the obtained results. In Appendix A of the paper, we show additional details of the E- and M-steps of the EM algorithm, and we describe the model selection criterion. Appendix B contains more details on the simulation results. In the supporting information, we show an application in which the model is estimated using a semi-parametric approach (Section A), and we provide further insights into the simulation results (Section B).

2 REGIME SWITCHING STUDENT-*t* COPULA MODEL

We consider an r-dimensional copula function C, which is a multivariate cumulative distribution function on the hypercube $[0,1]^r$ with marginal uniform distributions in $[0,1]$. Estimating such a model through inferential procedures becomes challenging due to limitations in numerical optimisation methods when dealing with high-dimensional parameter vectors. Additionally, the joint likelihood often involves multidimensional integrals, posing difficulties in numerical computations. To solve this problem, we rely on Sklar's theorem (Sklar, 1959), suggesting that it is possible to separately estimate each marginal cumulative distribution function and the copula function. The inference for margins approach of Joe & Xu (1996) allows us to split the estimation into two steps: first, we fit the marginal distribution of each univariate time-series; second, we estimate the joint distribution of integral transforms of these series using a RS copula model.

In the following, for the sake of clarity, we explicitly refer to the practical context of financial data. Let $\boldsymbol{y}_t = (y_{t1}, \ldots, y_{tr})'$ denote the vector of log-returns of the r time-series at time $t = 1, \ldots, T$. Following a parametric approach (Joe, 1997; Nasri & Rémillard, 2019), we assume a generalised error model (Du, 2016) for each of the r univariate time-series. This model postulates a cumulative distribution function denoted as $G_{\boldsymbol{\beta}j}$, and characterises the integral transforms $z_{tj} = G_{\hat{\boldsymbol{\beta}}j}(y_{tj})$ as independent and identically distributed random variables, each with continuous distribution function F_j, $j = 1, \ldots, r$. To eliminate the dependence of the estimated copula parameters on the marginal distributions, as suggested by Nasri & Rémillard (2019), we

initially estimate the parameters $\boldsymbol{\beta}_j$ through a consistent estimator $\hat{\boldsymbol{\beta}}_j$; then we compute the uniform pseudo-observations z_{tj}, and finally, we calculate normalised ranks, denoted by $\hat{e}_{tj} = \text{rank}(z_{tj})/(T+1)$, for $t = 1, \ldots, T$, $j = 1, \ldots, r$. Following a semi-parametric approach, normalised ranks can be directly calculated from the observed log-returns, thus obtaining $\hat{e}_{tj} = \text{rank}(y_{tj})/(T+1)$. This is a valid alternative when there is no interest in estimating a parametric model for the marginal univariate time-series, which might be the case when the focus is only on the association between a set of random variables and not on their marginal distributions. We consider the first method in the application presented in Section 5 to analyse cryptocurrency log-returns, and we also offer the results with the semi-parametric approach in the supporting information.

With a slight abuse of notation, let us denote with $\boldsymbol{y}_t = (y_{t1}, \ldots, y_{tr})'$, $t = 1, \ldots, T$, the r-dimensional vector of pseudo-observations following an RSS*t*C model, and let u_t denote the latent variable following a time homogeneous Markov process of first order with k latent states. We postulate that:

- The latent process is characterised by a vector of initial probabilities $\boldsymbol{\lambda}$ with elements $\lambda_u = P(u_1 = u)$, $u = 1, \ldots, k$, and a transition matrix denoted as $\boldsymbol{\Pi}$, with elements $\pi_{v|u} = P(u_t = v|u_{t-1} = u)$, $u, v = 1, \ldots, k$.
- The vectors of pseudo-observations $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T$ are conditionally independent given the latent regimes $u_1, \ldots, u_T$, each with copula density $c(\cdot\,; \boldsymbol{R}_{u_1}, \nu_{u_1}), \ldots, c(\cdot\,; \boldsymbol{R}_{u_T}, \nu_{u_T})$, where $\boldsymbol{R}_u$ denotes the matrix of dependence parameters with entries $\rho_u^{(ij)}$, $i, j = 1, \ldots, r$, $i \neq j$, each measuring the correlation between asset i and j, and ν_u is the number of degrees of freedom of the Student-t copula.

The joint density of the pseudo-observations is given by

$$f(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T) = \sum_{u_1=1}^{k} \pi_{u_1} c(\boldsymbol{y}_1; \boldsymbol{R}_{u_1}, \nu_{u_1}) \sum_{u_2=1}^{k} \pi_{u_2|u_1} c(\boldsymbol{y}_2; \boldsymbol{R}_{u_2}, \nu_{u_2}) \ldots \sum_{u_T=1}^{k} \pi_{u_T|u_{T-1}} c(\boldsymbol{y}_T; \boldsymbol{R}_{u_T}, \nu_{u_T}). \tag{1}$$

More specifically, following Joe (2014), $c(\boldsymbol{y}_t; \boldsymbol{R}_{u_t}, \nu_{u_t})$, $t = 1, \ldots, T$, is given by

$$c(\boldsymbol{y}_t; \boldsymbol{R}_u, \nu_u) = \frac{t_{r,\nu_u}(\boldsymbol{x}_t; \boldsymbol{R}_u)}{\prod_{j=1}^{r} t_{1,\nu_u}(x_{tj})}, \quad u = 1, \ldots, k,$$

where $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{rt})'$ is the vector with components $x_{tj} = T_{1,\nu_u}^{-1}(y_{tj})$, $j = 1, \ldots, r$, $t = 1, \ldots, T$, and T_{1,ν_u}^{-1} is the inverse cumulative distribution function of a one-dimensional Student-t random variable with ν_u degrees of freedom. The univariate and r-variate Student-t densities, denoted as t_{1,ν_u} and t_{r,ν_u}, are defined as

$$t_{1,\nu_u}(x_{tj}) = \frac{\Gamma((\nu_u+1)/2)}{\sqrt{\pi\nu_u}\,\Gamma(\nu_u/2)} \left(1 + \frac{x_{tj}^2}{\nu_u}\right)^{-(\nu_u+1)/2},$$

$$t_{r,\nu_u}(\boldsymbol{x}_t; \boldsymbol{R}_u) = \frac{\Gamma\left(\frac{\nu_u + r}{2}\right)}{\Gamma\left(\frac{\nu_u}{2}\right) \nu_u^{\frac{r}{2}} \pi^{\frac{r}{2}} |\boldsymbol{R}_u|^{\frac{1}{2}}} \left(1 + \frac{1}{\nu_u} \boldsymbol{x}_t^T \boldsymbol{R}_u^{-1} \boldsymbol{x}_t\right)^{-\frac{\nu_u + r}{2}},$$

where $\Gamma(\cdot)$ is the gamma function.

To measure correlation between variables, we employ the Kendall's tau (Kendall, 1938), which offers advantages over the linear correlation coefficient, particularly in its ability to consider non-linear dependencies. Starting from the dependence parameters $\rho_u^{(ij)}$ of the Student-t copula, Kendall's tau can be computed through the following formula:

$$\tau_u^{(ij)} = \frac{2}{\pi} \arcsin \rho_u^{(ij)}, \qquad i, j = 1, \ldots, r, \; i \neq j. \tag{2}$$

3 MAXIMUM LIKELIHOOD ESTIMATION

Let $\ell(\boldsymbol{\theta}|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T)$ denote the log-likelihood of the proposed model, corresponding to the logarithm of (1), with $\boldsymbol{\theta}$ being the column vector of parameters including the non-redundant elements of $\boldsymbol{R}_u$, together with ν_u and λ_u, for $u = 1, \ldots, k$, and $\pi_{v|u}$, for $u, v = 1, \ldots, k$. Maximum likelihood estimation of the model parameters is performed through the EM algorithm; for the Student-t copula parameters, we use a two-step procedure where we first estimate the matrix of dependence parameters for a fixed number of degrees of freedom, through Lagrange multipliers, and then we estimate the number of degrees of freedom through numerical optimisation of the complete log-likelihood, keeping fixed the previously estimated matrix. In the following section, we show the steps of the EM algorithm, details of which are provided in Appendix A, and in Section 3.2 we provide additional information on the initialisation of the algorithm and its convergence.

3.1 Expectation–Maximisation Algorithm

The complete-data log-likelihood, denoted as $\ell^*(\boldsymbol{\theta}|(\boldsymbol{y}_1, u_1), \ldots, (\boldsymbol{y}_T, u_T))$, is the log-likelihood computed assuming the knowledge of the hidden states $u_1, \ldots, u_t$, and expressed as

$$\begin{aligned} \ell^*(\boldsymbol{\theta}|(\boldsymbol{y}_1, u_1), \ldots, (\boldsymbol{y}_T, u_T)) \quad = \quad & \sum_{t=1}^{T} \sum_{u=1}^{k} w_{tu} \log c(\boldsymbol{y}_t; \boldsymbol{R}_u, \nu_u) + \sum_{u=1}^{k} w_{1u} \log \lambda_u \\ & + \sum_{t=2}^{T} \sum_{u=1}^{k} \sum_{v=1}^{k} z_{tuv} \log \pi_{v|u}, \end{aligned} \tag{3}$$

being $w_{tu} = I(u_t = u)$ an indicator variable equal to 1 when the latent process is in state u at time t (0 otherwise), and $z_{tuv} = I(u_{t-1} = u, u_t = v)$ equal to 1 if the latent process switches from state u at time $t - 1$ to state v a time t (0 otherwise). Note that this log-likelihood is the sum of three components that may be maximised separately.

Starting from some initial values for the parameters collected into the vector $\boldsymbol{\theta}^{(0)}$, the EM-algorithm maximises $\ell(\boldsymbol{\theta}|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T)$, by alternating, at each iteration m, the following two steps until convergence:

- **E-step**. Compute the conditional expected value of $\ell^*(\boldsymbol{\theta}|(\boldsymbol{y}_1, u_1), \ldots, (\boldsymbol{y}_T, u_T))$, given the values of the parameters at the previous iteration and the pseudo-observations. At this step, we rely on the posterior expected values of the previous indicator variables, denoted by $\hat{w}_{tu}$ and $\hat{z}_{tuv}$, whose formulas are provided in Appendix A.
- **M-step**. Maximise the expected value of $\ell^*(\boldsymbol{\theta}|(\boldsymbol{y}_1, u_1), \ldots, (\boldsymbol{y}_T, u_T))$ and update the model parameters. In particular, parameters λ_u and $\pi_{v|u}$ are updated by using the following explicit rules:

$$\lambda_u^{(m)} = \frac{\widehat{w}_{1u}}{\sum_{v=1}^{k} \widehat{w}_{1v}}, \quad u = 1, \ldots, k, \tag{4}$$

$$\pi_{v|u}^{(m)} = \frac{\sum_{t=2}^{T} \widehat{z}_{tuv}}{\sum_{t=2}^{T} \widehat{w}_{t-1u}}, \qquad u, v = 1, \ldots, k. \tag{5}$$

The updated values of the remaining parameters, that is, $\boldsymbol{R}_u^{(m)}$ and $\nu_u^{(m)}$, are obtained by solving the following optimisation problem:

$$\max_{\boldsymbol{R}_u,\, \nu_u} \sum_{t=1}^{T} \widehat{w}_{tu} \log c(\boldsymbol{y}_t;\, \boldsymbol{R}_u,\, \nu_u), \quad u = 1, \ldots, k; \tag{6}$$

direct numerical maximisation of (6) may be performed. However, it results computationally inefficient, especially when the number of available assets is large. Following Trede (2020), we maximise (6) with respect to $\boldsymbol{R}_u$, $u = 1, \ldots, k$, given $\nu_u^{(m-1)}$ using Lagrange multipliers, to obtain $\boldsymbol{R}_u^{(m)}$, and then with respect to ν_u given $\boldsymbol{R}_u^{(m)}$, obtaining $\nu_u^{(m)}$. In particular, we obtain an estimated approximation of the matrix of dependence parameters via the following rule:

$$\boldsymbol{R}_u^{(m)} = \boldsymbol{A}^{(m-1)} + \boldsymbol{R}_u^{(m-1)} \operatorname{diag}\left[\left(\boldsymbol{R}_u^{(m-1)} \circ \boldsymbol{R}_u^{(m-1)}\right)^{-1}\left(\mathbf{1} - \boldsymbol{a}^{(m-1)}\right)\right] \boldsymbol{R}_u^{(m-1)}, \tag{7}$$

with

$$\boldsymbol{A}^{(m-1)} = \frac{\nu_u^{(m-1)} + r}{\nu_u^{(m-1)} \sum_{t=1}^{T} \widehat{w}_{tu}} \sum_{t=1}^{T} \widehat{w}_{tu} \boldsymbol{x}_t \boldsymbol{x}_t' \left[1 + \frac{\boldsymbol{x}_t' \left(\boldsymbol{R}_u^{(m-1)}\right)^{-1} \boldsymbol{x}_t}{\nu_u^{(m-1)}}\right]^{-1},$$

where $\circ$ in (7) denotes the element-wise product, $\mathbf{1}$ is a vector of 1s, $\boldsymbol{a}^{(m-1)}$ denotes the vector of diagonal elements of $\boldsymbol{A}^{(m-1)}$, and $\boldsymbol{x}_t$ is the vector with components $T_{1,\nu_u}^{-1}\left(y_{tj}\right)$, with $j = 1, \ldots, r$. In this way, we do not require numerical optimisation methods for such an estimate, thus reducing the computational effort. See Appendix A for additional details on the derivation of the above formulas.

It is simple to numerically maximise Equation (6) with respect to ν_u once we set $\boldsymbol{R}_u = \boldsymbol{R}_u^{(m)}$, because the number of degrees of freedom ν_u of the Student-t copula is a scalar parameter. The estimates for ν_u, $u = 1, \ldots, k$, are obtained as

$$\nu_u^{(m)} = \operatorname*{argmax}_{\nu_u} \sum_{t=1}^{T} \widehat{w}_{tu} \log c(\boldsymbol{y}_t;\, \boldsymbol{R}_u^{(m)},\, \nu_u), \quad u = 1, \ldots, k. \tag{8}$$

We employ a heuristic approach with specific bounds to ensure successful computation of the objective function for estimating the parameter ν_u. We set the lower bound at 2 for practical purposes, and the upper bound is chosen as 25 to prevent numerical instability of the algorithm at higher values. As also reported in Trede (2020), larger values of ν_u imply a significantly higher computational time needed to achieve convergence to the maximum of the log-likelihood function. Additionally, as the number of degrees of freedom increases, the Student-t copula gradually

approximates the Gaussian copula, and it may result less effective to model extreme returns. In the simulation study of Section 4 and in the application presented in Section 5, we arrange these regimes in ascending order of determinants, from 1 to 0, corresponding to decreasing overall correlation values.

Standard errors for the parameter estimates are computed by parametric bootstrap (Davison & Hinkley, 1997; Chernick, 2011). In particular, we make use of the stationary block bootstrap (Politis & Romano, 1994) to preserve time-series dependence of the data: it consists in resampling blocks of consecutive observations assuming that the length of each block is distributed as a geometric random variable with average size proportional to $\mathcal{O}(T^{2/3})$.

3.2 Initialisation and Convergence of the Algorithm

In the literature, there is currently no consensus on the most appropriate approach for initialising the values in $\boldsymbol{\theta}^{(0)}$ within the context of the EM algorithm. We follow the proposal in Bartolucci *et al.* (2013) using a deterministic rule as an initialisation strategy for $\boldsymbol{R}_u$, $u = 1, \ldots, k$, such that initial values are defined on the basis of the descriptive statistics computed for the observed time-series. To determine the initial values for the dependence parameters, we begin by computing the matrix of sample Kendall's tau. Subsequently, we invert the formula in Equation (2) to obtain initial estimates of $\rho_u^{(ij)}$. Other possible choices for initialization are illustrated in Maruotti & Punzo (2021). The starting values for the initial probabilities λ_u are set equal to $1/k$, and those of the transition probabilities $\pi_{v|u}$ are set equal to $1/(\gamma + k)$ for $v \neq u$ and equal to $(\gamma + 1)/(\gamma + k)$ for $v = u$, where γ is a suitable constant (we use $\gamma = 0$ in the application of Section 5). We note that a moderate initial value of v_u is the best practical choice: it should not be too large or too small because we might encounter convergence issues. For this reason, based on a heuristic strategy, the number of degrees of freedom of the Student-*t* copula is initialised with 4.

Regarding algorithm convergence, we employ two common approaches: monitoring the distance between estimated parameter vectors at consecutive steps and tracking the increase in the log-likelihood function at each step. Specifically, the E- and M-steps iterate until either or both of the following conditions are met

$$\max_h \left| \theta_h^{(m+1)} - \theta_h^{(m)} \right| < \epsilon_1,$$

$$\left| \ell(\boldsymbol{\theta}^{(m+1)}) - \ell(\boldsymbol{\theta}^{(m)}) \right| < \epsilon_2,$$

being $\theta_h^{(m)}$ the h-th element of the vector $\boldsymbol{\theta}^{(m)}$ at the m-th iteration of the algorithm and $\epsilon_1, \epsilon_2 > 0$ suitable tolerance levels. In the simulation study presented in Section 4 and in the empirical analysis of Section 5, both tolerance levels are set equal to 10^{-8}.

4 SIMULATION STUDY

We validate the proposed RSS*t*C modelling approach through a simulation study, examining the properties of the estimators of the dependence parameters, number of degrees of freedom, initial and transition probabilities. In particular, we present the simulation results for a 3-state RSS*t*C model.

We conduct experiments on a Standard NC6 Promo virtual machine with 6 cores and 56 GB of memory. As mentioned in the introduction, we implement the `R` code for the EM algorithm through the package **Rcpp**, which allows the user to easily integrate `C++` into the `R`

environment. The code is available at the following link: https://github.com/FedericoCortese/RSstcopula/find/main.

The estimation procedure is remarkably efficient, as it takes only around 50 seconds to estimate a 3-state model with data consisting of 1,500 observations and 5 marginals. Moreover, the computational time increases linearly with both the length of the data (T) the number of assets (r).

4.1 Three State Regime Switching Student-t Copula Model

We generate data from a 3-state RSStC model drawing $B = 1{,}000$ samples of dimension $r = 5$, each with a total number of observations $T = 1{,}500$, with vector of initial probabilities $\boldsymbol{\lambda} = (1/3, 1/3, 1/3)'$ and transition probability matrix given by

$$\boldsymbol{\Pi} = \begin{bmatrix} 0.700 & 0.200 & 0.100 \\ 0.300 & 0.600 & 0.100 \\ 0.100 & 0.100 & 0.800 \end{bmatrix}.$$

The dependence matrices are defined as

$$\boldsymbol{R}_1 = \begin{bmatrix} 1.000 & - & - & - & - \\ 0.900 & 1.000 & - & - & - \\ 0.700 & 0.750 & 1.000 & - & - \\ 0.800 & 0.900 & 0.700 & 1.000 & - \\ 0.800 & 0.800 & 0.800 & 0.800 & 1.000 \end{bmatrix},$$

$$\boldsymbol{R}_2 = \begin{bmatrix} 1.000 & - & - & - & - \\ 0.500 & 1.000 & - & - & - \\ 0.300 & 0.400 & 1.000 & - & - \\ 0.500 & 0.400 & 0.400 & 1.000 & - \\ 0.400 & 0.500 & 0.500 & 0.300 & 1.000 \end{bmatrix},$$

$$\boldsymbol{R}_3 = \begin{bmatrix} 1.000 & - & - & - & - \\ 0.100 & 1.000 & - & - & - \\ 0.150 & -0.100 & 1.000 & - & - \\ 0.050 & 0.100 & 0.050 & 1.000 & - \\ 0.050 & -0.050 & 0.100 & -0.010 & 1.000 \end{bmatrix},$$

with state-specific numbers of degrees of freedom equal to $\nu_1 = 3,\ \nu_2 = 6$ and $\nu_3 = 10$, respectively.

We evaluate the estimators in terms of the average bias and root mean squared error (RMSE) across $B = 1{,}000$ samples computed for the h-th parameter θ_h as

$$\text{Bias} = \text{E}\left(\widetilde{\theta}_h - \theta_h\right),$$

$$\text{RMSE} = \sqrt{\text{E}\left(\widetilde{\theta}_h - \theta_h\right)^2 + \text{Var}\left(\widetilde{\theta}_h\right)},$$

where $\tilde{\theta}_h$ denotes the h-th component of the vector of parameters $\tilde{\boldsymbol{\theta}}$ estimated on the simulated sample, and θ_h the corresponding component of the vector $\boldsymbol{\theta}$ of true parameters. We also compute bootstrap percentiles confidence intervals (CIs) at 95% confidence level.

Table 1 shows results for the parameters of the proposed model under the above scenarios. The bias is always low and RMSEs of the initial probabilities are around 0.5, while they are below 0.168 for the transition probabilities. CIs of the transition probabilities are narrower for higher probabilities.

Table 2 reports results referred to the dependence parameters $\rho_u^{(ij)}$ and the number of degrees of freedom ν_u. The maximum absolute bias is 0.037 for the pair of marginals $(2, 3)$ in the second state. RMSE values are all below 0.135. CIs for the first state are narrower, indicating more accurate estimates when the correlation is high, while CIs for the second state show higher uncertainty in the estimated parameters. As ν_u decreases, bias tends to decrease, thus using distributions with fat tails provides better results.

We also examined a 2-state RSStC model in a separate simulation study, and the outcomes closely resemble those of the 3-state model. Furthermore, we varied the number of observations, T, and the number of assets, r. Our findings indicate that the proposed approach has good finite sample properties, as evidenced by a decline in RMSE for increasing values of T. Similarly, as r increases, the RMSE decreases for initial and transition probabilities, as well as for the number of degrees of freedom, while it increases with respect to dependence parameters. Comprehensive information on these simulation results can be found in Appendix B.

5 EMPIRICAL STUDY

Data used for the application are multidimensional time-series of the daily log-returns considered at closing prices of BTC, ETH, XRP, LTC, and BCH, which are, in terms of market capitalisation, the less manipulated and more liquid crypto assets. Data are provided by the Crypto Asset Lab (https://www.diseade.unimib.it/it/ricerca/osservatori/crypto-asset-lab), which is an independent academic lab established at the University of Milano-Bicocca. We recall that BTC is the first cryptocurrency that has operated digitally since 2009 with a decentralised ledger system known as blockchain. ETH, released in 2015, has a semi-decentralised network that allows creating and running smart contracts, whereby it differs from other cryptocurrencies with its unlimited supply. LTC is a clone of BTC, created in 2011. Meanwhile, XRP was created in 2012 with a different design from BTC as it has a centralised network and an un-mineable coin. Finally, BCH is an altcoin created in 2007. In particular, recently, they got increasing public attention because they differentiate quite a lot from other more common assets due to their extraordinary return potential in phases of extreme price growth. We consider 1,842 daily closing prices observed over a 5 years period from 17 September 2017 to 2 October 2022. Log-returns of the daily closing prices are given by

$$y_{tj} = \log \frac{p_{t+1,j}}{p_{tj}}, \qquad j = 1, \ldots, r, \quad t = 1, \ldots, T,$$

where p_{tj} denotes the closing price for asset j at time t. Similar data have been analysed in Pennoni *et al.* (2021) through a Gaussian HMM based on discrete latent variables, to which we refer the reader for more details.

Table 3 presents the sample unconditional means and standard deviations of the log-returns for the five cryptocurrencies. Volatilities exhibit remarkably high values, while the average log-returns are approximately 0. Table 4 shows the observed linear correlations: a positive

Table 1. *Simulation results for the 3-state RSStC model: true parameter value, bias,* RMSE, *lower and upper bounds of the* CIs (CI_L *and* CI_U, *respectively) of the initial and transition probabilities.*

	λ_1	λ_2	λ_3	$\pi_{1\|1}$	$\pi_{1\|2}$	$\pi_{1\|3}$	$\pi_{2\|1}$	$\pi_{2\|2}$	$\pi_{2\|3}$	$\pi_{3\|1}$	$\pi_{3\|2}$	$\pi_{3\|3}$
True	0.333	0.333	0.333	0.700	0.300	0.100	0.200	0.600	0.100	0.100	0.100	0.800
Bias	−0.014	0.079	−0.065	0.019	−0.012	−0.007	−0.027	0.085	−0.058	0.000	−0.025	0.026
RMSE	0.475	0.440	0.492	0.080	0.074	0.048	0.099	0.168	0.112	0.040	0.080	0.074
CI_L	0.000	0.000	0.000	0.461	0.100	0.024	0.167	0.225	0.019	0.027	0.015	0.603
CI_U	1.000	1.000	1.000	0.774	0.354	0.220	0.517	0.730	0.384	0.180	0.299	0.859

Table 2. *Simulation results for the 3-state RSStC model: the true parameter value, bias,* RMSE*, lower and upper bounds of the* CIs *(*CI_L *and* CI_U*, respectively) for the dependence parameters, and the number of degrees of freedom.*

State 1	$\rho_1^{(12)}$	$\rho_1^{(13)}$	$\rho_1^{(14)}$	$\rho_1^{(15)}$	$\rho_1^{(23)}$	$\rho_1^{(24)}$	$\rho_1^{(25)}$	$\rho_1^{(34)}$	$\rho_1^{(35)}$	$\rho_1^{(45)}$	ν_1
True	0.900	0.700	0.800	0.800	0.750	0.900	0.800	0.700	0.800	0.800	3.000
Bias	0.001	0.003	0.002	0.002	0.003	0.001	0.003	0.003	0.002	0.002	−0.031
RMSE	0.015	0.034	0.025	0.026	0.030	0.015	0.028	0.034	0.025	0.026	0.441
CI_L	0.869	0.628	0.750	0.746	0.688	0.867	0.742	0.635	0.747	0.745	2.285
CI_U	0.924	0.763	0.842	0.843	0.802	0.923	0.843	0.759	0.842	0.843	4.054
State 2	$\rho_2^{(12)}$	$\rho_2^{(13)}$	$\rho_2^{(14)}$	$\rho_2^{(15)}$	$\rho_2^{(23)}$	$\rho_2^{(24)}$	$\rho_2^{(25)}$	$\rho_2^{(34)}$	$\rho_2^{(35)}$	$\rho_2^{(45)}$	ν_2
True	0.500	0.300	0.500	0.400	0.400	0.400	0.500	0.400	0.500	0.300	6.000
Bias	0.032	0.032	0.031	0.032	0.037	0.035	0.032	0.026	0.029	0.036	0.611
RMSE	0.110	0.117	0.110	0.119	0.124	0.128	0.115	0.109	0.110	0.135	1.844
CI_L	0.302	0.108	0.310	0.197	0.206	0.175	0.302	0.228	0.320	0.194	3.848
CI_U	0.759	0.581	0.741	0.660	0.691	0.719	0.760	0.656	0.736	0.621	10.395
State 3	$\rho_3^{(12)}$	$\rho_3^{(13)}$	$\rho_3^{(14)}$	$\rho_3^{(15)}$	$\rho_3^{(23)}$	$\rho_3^{(24)}$	$\rho_3^{(25)}$	$\rho_3^{(34)}$	$\rho_3^{(35)}$	$\rho_3^{(45)}$	ν_3
True	0.100	0.150	0.050	0.050	−0.100	0.100	−0.050	0.050	0.100	−0.010	10.000
Bias	0.008	0.000	0.011	0.007	0.012	0.007	0.014	0.011	0.011	0.009	1.416
RMSE	0.073	0.062	0.078	0.070	0.078	0.064	0.081	0.073	0.072	0.068	3.725
CI_L	−0.038	0.024	−0.097	−0.081	−0.249	−0.018	−0.194	−0.092	−0.039	−0.145	7.002
CI_U	0.239	0.275	0.199	0.191	0.061	0.225	0.125	0.192	0.239	0.122	21.602

International Statistical Review (2024)

Table 3. *Sample means and standard deviations (SD) of BTC, ETH, XRP, LTC, and BCH log-returns referred to the period from 17 September 2017 to 2 October 2022.*

	Cryptocurrency				
	BTC	*ETH*	*XRP*	*LTC*	*BCH*
Mean (%)	0.090	0.087	0.049	0.002	−0.073
SD (%)	4.166	5.271	6.451	5.659	6.585

Table 4. *Observed linear correlations between log-returns of BTC, ETH, XRP, LTC, and BCH.*

	BTC	*ETH*	*XRP*	*LTC*	*BCH*
BTC	1.000	-	-	-	-
ETH	0.787	1.000	-	-	-
XRP	0.560	0.653	1.000	-	-
LTC	0.765	0.823	0.644	1.000	-
BCH	0.678	0.742	0.583	0.732	1.000

association is present for each pair, with a maximum value of 0.787 for the pair of cryptos BTC-ETH.

First, we assume the well-known ARMA(1,1)-GARCH(1,1) model (Engle & Bollerslev, 1986) for the marginals. Its efficacy, when combined with copula models, has been demonstrated in previous studies such as Bauwens *et al.* (2006) and Patton (2012). It postulates the following autoregressive equations for the conditional mean and variance of each log-return series:

$$\begin{aligned} y_{tj} &= \alpha_{1j} y_{t-1,j} + \alpha_{2j}\zeta_{t-1,j} + \sqrt{\sigma^2_{tj}}\xi_{tj}, \\ \sigma^2_{tj} &= \omega_{0j} + \omega_{1j}\zeta^2_{t-1,j} + \omega_{2j}\sigma^2_{t-1,j}, \end{aligned} \tag{9}$$

where $\alpha_{1j},\ \alpha_{2j},\ \omega_{0j},\ \omega_{1j}$, and ω_{2j}, are the ARMA(1,1)-GARCH(1,1) parameters for time-series j, $j = 1, \ldots, r$, and $\zeta_{tj} = y_{tj} - \overline{y}_{tj}$, with $\overline{y}_{tj} = \alpha_{1j} y_{t-1,j} + \alpha_{2j}\zeta_{t-1,j}$. Second, we assume that the innovations ξ_{tj} follow a skewed generalised error distribution (SGED, Theodossiou, 2015), whose density is given by

$$f(x;\ \phi,\ \kappa) = \frac{\kappa \exp\left[-\frac{1}{\kappa}\left|\frac{x+\delta_1}{\delta_2(1+\phi\,\mathrm{sign}(x+\delta_1))}\right|^{\kappa}\right]}{2\delta_2\Gamma\left(\frac{1}{\kappa}\right)}, \tag{10}$$

where ϕ is the skewness parameter, κ the shape parameter and

$$\delta_1 = \frac{2^{\frac{2}{\kappa}}\delta_2\phi\Gamma\left(\frac{1}{2}+\frac{1}{\kappa}\right)}{\sqrt{\pi}},$$

$$\delta_2 = \frac{\pi\left(1+3\phi^2\right)\Gamma\left(\frac{3}{\kappa}\right) - 16^{\frac{1}{\kappa}}\phi^2\Gamma\left(\frac{1}{2}+\frac{1}{\kappa}\right)\Gamma\left(\frac{1}{\kappa}\right)}{\pi\Gamma\left(\frac{1}{\kappa}\right)}.$$

Table 5. *Estimated parameters of the ARMA(1,1)-GARCH(1,1) model as in Equation (9). The coefficients ϕ_j and κ_j, $j = 1, \ldots, 5$, refer to the skewness and shape parameters of the SGED. Standard errors (in brackets) are obtained through the nonparametric block bootstrap.*

	Cryptocurrency				
Parameter	*BTC*	*ETH*	*XRP*	*LTC*	*BCH*
α_{1j}	−0.122	−0.214	−0.021	−0.045	−0.124
	(0.009)	(0.015)	(0.003)	(0.012)	(0.014)
α_{2j}	0.038	0.113	−0.149	−0.060	0.018
	(0.003)	(0.010)	(0.008)	(0.014)	(0.009)
ω_{0j}	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ω_{1j}	0.067	0.085	0.126	0.079	0.062
	(0.006)	(0.017)	(0.021)	(0.019)	(0.013)
ω_{2j}	0.923	0.854	0.854	0.864	0.911
	(0.005)	(0.025)	(0.022)	(0.033)	(0.018)
ϕ_j	0.963	0.972	0.990	0.986	1.016
	(0.006)	(0.016)	(0.015)	(0.018)	(0.041)
κ_j	0.910	1.040	0.874	1.054	0.908
	(0.047)	(0.043)	(0.033)	(0.042)	(0.035)

Table 6. *P-values of the parametric bootstrap (*PB*) and Dickey–Fuller (*DF*) tests.*

	Cryptocurrency				
Test	*BTC*	*ETH*	*XRP*	*LTC*	*BCH*
PB	0.106	0.433	0.369	0.894	0.146
DF	<0.01	<0.01	<0.01	<0.01	<0.01

SGED is particularly flexible because it reduces to the standard Gaussian distribution when $\phi = 0$ and $\kappa = 2$ and to the Laplace distribution when $\phi = 0$ and $\kappa = 1$. It has been used previously to model univariate cryptocurrency time-series in Cerquetti *et al.* (2020).

Table 5 reports the estimated coefficients of the marginals models. Standard errors are obtained with the nonparametric block bootstrap as detailed in Section 3.1, considering an average block length of 127 and $B = 1{,}000$ bootstrap samples.

Following Nasri & Rémillard (2019), we perform a parametric bootstrap (PB) test to evaluate the adequacy of the marginal models (Rémillard, 2011). The implementation proceeds in two steps: firstly, we generate simulated data based on the estimated marginal model; secondly, we compute the Cramér-Von Mises test statistic for the bootstraped data and compare this value with the value of the test statistic computed for the observed data in order to assess model adequacy. Results reported in Table 6 show that the null hypothesis of a correct specification for the marginal distribution is never rejected at each statistical significance level. In the same table, results from the Dickey-Fuller (DF) test suggest that the null hypothesis of non-stationarity of the innovations is rejected at each significance level.

We also investigate the presence of change-points in the residuals, employing the wild binary search procedure proposed by Fryzlewicz (2014). Results indicate the existence of a minimum of six change-points for all cryptocurrencies, with some displaying an even higher number of changes.

Once we have computed the marginal pseudo-observations through the normalised ranks of the integral transformation of the innovations from the previous models, we can estimate the RSS*t*C model, and perform model selection. The Bayesian information criterion (BIC,

Table 7. *Integrated completed likelihood (*ICL*) and Bayesian information criteria (*BIC*) computed for increasing values of the number of hidden regimes k. The minimum values are indicated in bold.*

	k			
Information Criterion	1	2	3	4
ICL	−9,469.583	**−9,616.781**	−9,449.637	−9,601.758
BIC	−9,469.583	−9,887.301	−9,935.296	**−9,988.286**

Table 8. *Estimated number of degrees of freedom ν_u, and determinant of the estimated matrices of dependence parameters under the 2-state RSStC model. Standard errors (in brackets) are obtained with nonparametric block bootstrap.*

State	$u = 1$	$u = 2$
ν_u	6.231	9.416
	(1.275)	(3.627)
$\det(\boldsymbol{R}_u)$	0.001	0.065

Schwarz, 1978) is commonly employed to choose a suitable number of latent states, although, it may overestimate this number. Alternatively, as demonstrated in Pohle *et al.* (2017), ICL provides more parsimonious results. In Table 7, we show the values of BIC and ICL under the RSS*t*C model with k ranging from 1 to 4. While the BIC decreases as k increases, ICL leads us to select a model with two latent states. The maximum log-likelihood for the 2-state RSS*t*C model is $\hat{\ell}(\boldsymbol{\theta}|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T) = 4{,}902.366$ with $K = 25$ estimated parameters.

Table 8 reports the estimated number of degrees of freedom and the determinant of the fitted matrices of dependence parameters under the RSS*t*C model with 2 states. Notably, regimes with strong dependence exhibit a lower estimated number of degrees of freedom, indicating distributions with fat tails for states showing high correlations. This suggests that when the correlation among crypto assets is high, joint high losses (or earnings) occur more frequently.

Table 9 presents the matrix of the estimated dependence parameters and Table 10 displays the computed Kendall's tau values using Equation (2). These estimates allow us to characterise each regime based on pair-specific correlations. The first regime exhibits the highest Kendall's tau values, indicating a highly correlated market state. In contrast, the second regime displays correlation values ranging from 0.324 to 0.519, suggesting a market regime with lower interdependence.

Table 11 reports the estimated transition probability matrices under the 2-state RSS*t*C model. We notice a general persistence in each regime: the maximum off-diagonal entry is observed from regime 2 to regime 1 (0.121). The estimated stationary distribution has probabilities (0.579, 0.421).

In Table 12 we present the estimated state-conditional means and standard deviations for the five cryptocurrency log-returns, with the state prediction performed through the Viterbi algorithm (Viterbi, 1967). Findings reveal that the 1st state, which represents a regime characterised by high correlations among cryptocurrencies, is associated with negative daily log-returns. Conversely, the 2nd state demonstrates positive average returns. Based on these observations, we can characterise the two states as bearish and bullish market regimes. Moreover, the state-conditional standard deviations indicate high volatility in both regimes.

Figure 1 displays the decoded state sequence alongside the prices of BTC, ETH, XRP, LTC, and BCH. The analysis reveals distinct periods characterised by different market regimes. Initially, a bullish market regime dominates, followed by a significant presence of a bearish market

Table 9. *Estimated dependence parameters $\rho_u^{(ij)}$ under the 2-state RSStC model. Standard errors (in brackets) are obtained with nonparametric block bootstrap.*

State $u = 1$	*BTC*	*ETH*	*XRP*	*LTC*	*BCH*
BTC	1.000	-	-	-	-
	(0.000)				
ETH	0.911	1.000	-	-	-
	(0.016)	(0.000)			
XRP	0.902	0.910	1.000	-	-
	(0.025)	(0.020)	(0.000)		
LTC	0.875	0.902	0.910	1.000	-
	(0.034)	(0.026)	(0.034)	(0.000)	
BCH	0.902	0.907	0.927	0.901	1.000
	(0.023)	(0.022)	(0.027)	(0.035)	(0.000)
State $u = 2$	*BTC*	*ETH*	*XRP*	*LTC*	*BCH*
BTC	1.000	-	-	-	-
	(0.000)				
ETH	0.652	1.000	-	-	-
	(0.128)	(0.000)			
XRP	0.667	0.728	1.000	-	-
	(0.067)	(0.039)	(0.000)		
LTC	0.487	0.613	0.592	1.000	-
	(0.122)	(0.087)	(0.066)	(0.000)	
BCH	0.569	0.645	0.672	0.517	1.000
	(0.156)	(0.090)	(0.123)	(0.145)	(0.000)

Table 10. *Kendall's tau as in Equation (2) computed with the estimated dependence parameters $\rho_u^{(ij)}$ under the 2-state RSStC model.*

State $u = 1$	*BTC*	*ETH*	*XRP*	*LTC*	*BCH*
BTC	1.000	-	-	-	-
ETH	0.730	1.000	-	-	-
XRP	0.715	0.728	1.000	-	-
LTC	0.678	0.715	0.728	1.000	-
BCH	0.716	0.724	0.756	0.714	1.000
State $u = 2$	*BTC*	*ETH*	*XRP*	*LTC*	*BCH*
BTC	1.000	-	-	-	-
ETH	0.452	1.000	-	-	-
XRP	0.465	0.519	1.000	-	-
LTC	0.324	0.420	0.403	1.000	-
BCH	0.385	0.446	0.469	0.346	1.000

regime until mid-2018. After a brief period of price increases, bearish periods become prominent until early 2020. Subsequently, all cryptocurrencies exhibit positive returns until late 2021, and from this period a prevailing bearish trend reemerges.

Notably, a consistent increase or descrese in prices corresponds to bullish or bearish market regimes, which can be identified solely by examining correlations. This implies that the presence of a specific market regime can be detected without relying on the first-order and second-order moments of cryptocurrency log-returns. Bullish and bearish regimes are visited 59.1%, 40.9% of the time, respectively, and the average sojourn times are equal to 23 days for the first state and 16 days for the second state.

Table 11. *Estimated transition probabilities $\pi_{u|v}$ under the 2-state RSStC model. Standard errors (in brackets) are obtained with nonparametric block bootstrap.*

State	$u = 1$	$u = 2$
$v = 1$	0.912	0.088
	(0.046)	(0.041)
$v = 2$	0.121	0.879
	(0.046)	(0.041)

Table 12. *Estimated state-conditional means and standard deviations of the five cryptocurrencies log-returns with state allocation obtained through global decoding under the 2-state RSStC model.*

State 1	*Mean (%)*	*SD (%)*	*State 2*	*Mean (%)*	*SD (%)*
BTC	−0.363	4.121	BTC	0.744	4.147
ETH	−0.509	5.400	ETH	0.948	4.957
XRP	−0.699	5.377	XRP	1.131	7.619
LTC	−0.725	5.454	LTC	1.052	5.788
BCH	−0.902	6.013	BCH	1.125	7.169

Our findings corroborate the results from the previous study by Ardia *et al.* (2019), in which the authors use a 2-state Markov-switching GARCH model fitted on univariate BTC time-series. Their model employs a fat-tailed distribution across two regimes identified by low and high unconditional volatilities which exhibit strong persistence. Koki *et al.* (2022) estimate different HMMs to the log-returns of three cryptocurrencies, namely, BTC, ETH, and XRP, with different numbers of states. They determine that a 4-state model provides the most accurate forecasting results among all considered model specifications. Nonetheless, the statistical properties of the hidden states exhibit differences among the three cryptocurrencies, making the interpretation of these latent states as distinct economic regimes a more challenging task.

5.1 Comparative Analysis

We compare the results of our proposal with those obtained with the basic HMM (Zucchini *et al.*, 2017). We also show the forecast performance of both models and a benchmark model, based on the MRW process.

The basic HMM assumes a conditional Gaussian distribution and it is generally not robust for analysing extreme events usually observed in financial data. Cryptocurrency markets are characterised by heavy-tailed distributions that lead to frequent extreme price movements, whether positive or negative, more than in traditional financial markets. In this regard, the Student-*t* copula accommodates heavy-tailed distributions, allowing us to model the idiosyncrasies of cryptocurrency log-returns. Thus, the proposed RSS*t*C model may appropriately represents the observed underlying trends of the cryptocurrency log-returns.

Estimation of the parameters of the HMM is performed with the routines provided within the R package **RcppHMM** (Ardenas-Ovando *et al.*, 2017); for an alternative HMM formulation the **LMest** package (Bartolucci *et al.*, 2017) can be used. Model selection performed with both BIC and ICL criteria suggests a HMM with 6 regimes. In the following, we show some results obtained with both 2- and 6-states HMMs, denoted as HMM-2, and HMM-6, respectively.

We note that the self-transition probabilities estimated under the HMM-2 are 0.633 and 0.873, respectively, and those of the HMM-6 range in the interval (0.425, 0.673). Notably, as illustrated in the previous section and shown in Table 11, the RSS*t*C model exhibits higher

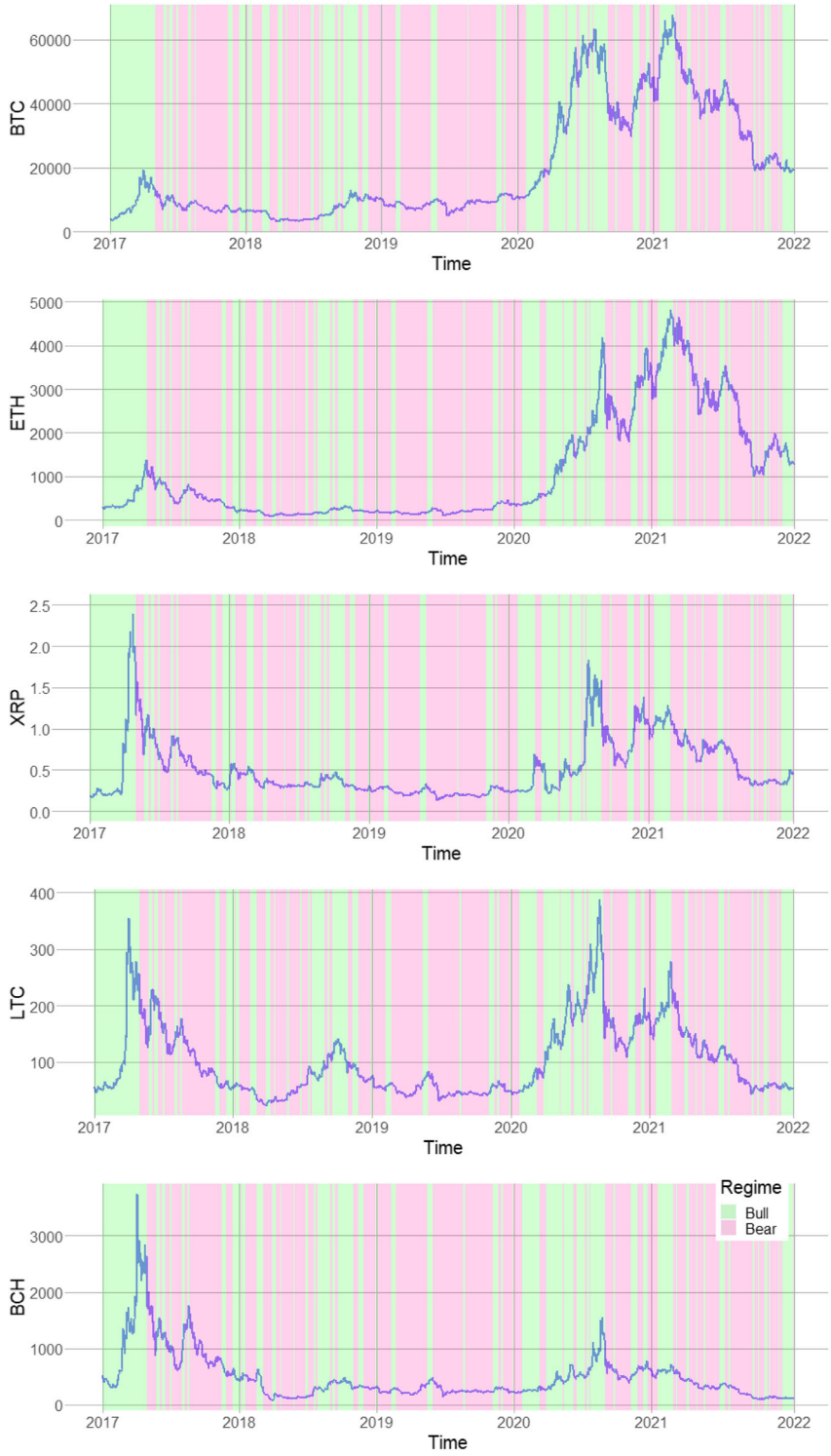

Figure 1. *Observed prices of BTC, ETH, XRP, LTC, and BCH (17 September 2017 to 2 October 2022) with the global decoding state sequence highlighted in red for state 1 (bearish market) and green for state 2 (bullish market).*

values. Such a model is sometimes preferable to make profitable investment decisions because asset allocation is more stable and the portfolio turnover is low. In fact, disposing of a model with more stable regimes allows us to avoid frequent reallocation or trading in response to short-term market fluctuations. This strategy implies reduced transaction costs and tax-efficient investments strategies, as also noticed in Nystrup *et al.* (2020).

Forecasts are implemented through a rolling window approach for each model: it consists in dividing the available data into overlapping windows, each consisting of 1,500 observations. We estimate parameters using each window and employ the estimates to generate one-step-ahead forecasts for the log-returns of the five cryptocurrencies. In more details, we obtain forecasts for the RSS*t*C model according to the following steps, as suggested in Simard & Rémillard (2015):

1 Estimate the RSS*t*C model as explained in Sections 2 and 3.
2 Simulate pseudo-marginals from the fitted copula model. This involves generating random samples from the copula function corresponding to the estimated dependence structure. In particular, we consider 1,000 observations.
3 Transform the generated uniform samples into the target marginal distributions using the inverse of the estimated marginal cumulative distribution functions.
4 Use data obtained at the previous step to forecast future values of each variable in the time-series using the sample mean, and estimate prediction intervals using the sample quantiles.

To estimate HMM and MRW forecasts, we adopt an approach similar to the previous one. We utilise the estimated parameters for the HMM to simulate portfolio realisations for the one-step-ahead observation. This process involves generating latent states based on the HMM's predictive distribution. These latent states are then used to create corresponding portfolio returns for the next period. Similarly, for the MRW process, we use the estimated parameters to simulate future portfolio returns.

We evaluate forecast quality through two different metrics. We employ the RMSE and the percentage of correct sign predictions (CSP). RMSE quantifies the accuracy between true and forecasted values by calculating the average of squared differences. CSP, on the other hand, measures the frequency with which we accurately forecast the sign of returns. We present the results in Table 13. We additionally test models forecast accuracy using Diebold and Mariano (2002) test, finding no significant difference in squared error estimates. We also consider the model confidence set (MCS) procedure of Hansen et al. (2011), identifying RSS*t*C as the top model for LTC, XRP, and BCH log-returns, and MRW for BTC and ETH. We also evaluate the accuracy of Value-at-Risk forecasts using Christoffersen's (1998) conditional coverage test. Our findings show that, at a 1% significance level, the hypothesis that the models accurately predict losses exceeding the Value-at-Risk threshold cannot be rejected for RSStC and HMMs.

As Timmermann (2018) wisely noted: '*detecting breaks in financial forecasting models is a formidable task, and transforming such evidence into more accurate forecasts is even more challenging*'. Our approach acknowledges this inherent complexity, positioning it as a practical

Table 13. *RMSE between true and forecasted values of the five cryptocurrencies and percentage CSP obtained under the 2-state RSStC, HMM-2, HMM-6 and MRW models.*

	RSStC	*HMM-2*	*HMM-6*	*MRW*
RMSE	**0.061**	0.065	0.067	0.093
CSP (%)	**53.26**	49.03	50.73	50.67

solution for decision-making processes, even without aiming for extraordinary predictive accuracy. In fact, in terms of forecasting performance, the RSS*t*C model achieves lower RMSE compared with traditional models, thus providing enhanced accuracy in predicting cryptocurrency returns which can results particularly useful in case of future financial crisis events.

6 DISCUSSION

In this paper, we provide three main contributions to the existing literature on regime switching copula models. First, we generalise the regime switching Student-*t* copula model to the multivariate case. Second, we propose a novel maximum likelihood estimation procedure for multivariate regime switching Student-*t* copula models through a two-step method, initially estimating the dependence matrix and then the number of degrees of freedom. Third, we analyse the joint distribution of the log-returns of five cryptocurrencies during the period 2017–2022 with the proposed model.

Our simulation studies demonstrate the good finite sample properties of the proposed estimator highlighting its ability to accurately detect the true number of degrees of freedom, especially in scenarios where this number is small, denoting a process with fat tails.

By analysing 5 years of time-series data encompassing the log-returns of Bitcoin, Ethereum, Ripple, Litecoin, and Bitcoin Cash, we show the suitability of a 2-state regime switching Student-*t* copula model for detecting market trends. Through the application of the Viterbi algorithm using the decoded state sequences, we can effectively distinguish between bullish and bearish market phases. Notably, bearish periods in financial markets correspond to increasing correlations among assets. In comparison to two commonly used models, our approach is more robust and yields slightly improved forecasting results. These advantages translate into the potential for making more profitable investment decisions and implementing portfolio trading strategies. The estimated allocation into each regime, as determined through global decoding, plays a crucial role in achieving these benefits. Given the ongoing growth of these cryptocurrencies and the presence of co-integration and dynamic interdependencies between them, the ability to detect signals of market dynamics is of primary importance, not only for optimising investment strategies but also for identifying early warnings of potential financial crises.

As lines for future research, we highlight that it would be of interest to model marginals distributions with a Markov process along with the joint distribution. Additionally, it could be interesting to investigate joint distributions that incorporate skewness, such as a skewed Student-*t* copula, to gain an even greater understanding of the underlying dynamics.

ACKNOWLEDGEMENTS

We acknowledge the University of Milano-Bicocca Data Science Lab (datalab) for supporting this work by providing some computational resources and the Crypto Asset Lab for providing the data employed to illustrate the feasibility of the proposed model. F. Pennoni and F. Bartolucci acknowledge the financial support from the grant 'Hidden Markov Models for Early Warning Systems' of Ministero dell'Università e della Ricerca (PRIN 2022TZEXKF) funded by European Union - Next Genereation EU.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the authors with the permission of the Crypto Asset Lab of the University of Milano-Bicocca.

REFERENCES

Ang, A. & Bekaert, G. (2002). International asset allocation with regime shifts. *The Rev. Finan. Stud.*, **15**, 1137–1187. https://doi.org/10.1093/rfs/15.4.1137

Ang, A. & Chen, J. (2002). Asymmetric correlations of equity portfolios. *J. Finan. Econ.*, **63**, 443–494.

Ang, A. & Timmermann, A. (2012). Regime changes and financial markets. *Ann. Rev. Finan. Econ.*, **4**, 313–337.

Ardenas-Ovando, R., Noguez, J. & Rangel-Escareno, C. 2017. RcppHMM: Rcpp Hidden Markov Model. R package version 1.2.2.

Ardia, D., Bluteau, K. & Rüede, M. (2019). Regime changes in bitcoin GARCH volatility dynamics. *Finance Res. Lett.*, **29**, 266–271.

Bartolucci, F., Farcomeni, A. & Pennoni, F. (2013). *Latent Markov Models for Longitudinal Data*. Boca Raton, FL: Chapman & Hall/CRC Press.

Bartolucci, F., Pandolfi, S. & Pennoni, F. (2017). LMest: An R package for latent Markov models for longitudinal categorical data. *J. Stat. Softw.*, **81**, 1–38.

Baum, L.E. & Petrie, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *The Ann. Math. Stat.*, **37**, 1554–1563.

Bauwens, L., Laurent, S. & Rombouts, J.V.K. (2006). Multivariate GARCH models: A survey. *J. Appl. Econometr.*, **21**, 79–109.

Biernacki, C., Celeux, G. & Govaert, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Trans. Pattern Anal. Mach. Intell.*, **22**, 719–725.

Breymann, W., Dias, A. & Embrechts, P. (2003). Dependence structures for multivariate high-frequency data in finance. *Quant. Finance*, **3**, 1–14.

Cerquetti, R., Giacalone, M. & Mattera, R. (2020). Skewed non-Gaussian GARCH models for cryptocurrencies volatility modelling. *Inform. Sci.*, **527**, 1–26.

Chernick, M.R. (2011). *Bootstrap Methods: A Guide for Practitioners and Researchers*. Newtown, PA: John Wiley & Sons.

Chollete, L., Heinen, A. & Valdesogo, A. (2009). Modeling international financial returns with a multivariate regime-switching copula. *J. Finan. Econometr.*, **7**, 437–480.

Christoffersen, P.F. (1998). Evaluating interval forecasts. *Int. Econ. Rev.*, 841–862.

Cremaschini, A., Punzo, A., Martellucci, E. & Maruotti, A. (2023). On stylized facts of cryptocurrencies returns and their relationship with other assets, with a focus on the impact of COVID-19. *Appl. Econ.*, **55**, 3675–3688.

Czado, C. & Nagler, T. (2022). Vine copula based modeling. *Ann. Rev. Stat. Appl.*, **9**, 453–477. https://doi.org/10.1146/annurev-statistics-040220-101153

Das, S.R. & Uppal, R. (2004). Systemic risk and international portfolio choice. *The J. Finance*, **59**, 2809–2834.

Davison, A.C. & Hinkley, D.V. (1997). *Bootstrap Methods and Their Application*. Cambridge, MA: Cambridge University Press.

Demarta, S. & McNeil, A.J. (2005). The *t* copula and related copulas. *Int. Stat. Rev.*, **73**, 111–129.

Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc.: Ser. B*, **39**, 1–22.

Diebold, F.X. & Mariano, R.S. (2002). Comparing predictive accuracy. *J. Bus. Econ. Stat.*, **20**, 134–144.

Du, Z. (2016). Nonparametric bootstrap tests for independence of generalized errors. *The Econometr. J.*, **19**, 55–83.

Eddelbuettel, D. & François, R. (2011). Rcpp: Seamless R and C++ integration. *J. Stat. Softw.*, **40**, 1–18.

Engle, R.F. & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Economet. Rev.*, **5**, 1–50.

Fischer, M., Köck, C., Schlüter, S. & Weigert, F. (2009). An empirical analysis of multivariate copula models. *Quant. Finance*, **9**, 839–854.

Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *The Ann. Stat.*, **42**, 2243–2281.

Frühwirth-Schnatter, S. (2006) Finite mixture and Markov switching models. Springer, New York.

Garcia, R. & Ghysels, E. (1998). Structural change and asset pricing in emerging markets. *J. Int. Money Finance*, **17**, 455–473.

Genest, C., Nešlehová, J. & Ziegel, J. (2011). Inference in multivariate Archimedean copula models. *Test*, **20**, 223–256.

Härdle, W.K., Okhrin, O. & Wang, W. (2015). Hidden Markov structures for dynamic copulae. *Economet. Theory*, **31**, 981–1015.

Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, **57**, 357–384.

Hansen, P.R., Lunde, A. & Nason, J.M. (2011). The model confidence set. *Econometrica*, **79**, 453–497.

Hernández, L., Tejero, J. & Vinuesa, J. 2014. Maximum likelihood estimation of the correlation parameters for elliptical copulas. arXiv:1412.6316, 1–13.

Huang, J.-J., Lee, K.J., Liang, H. & Lin, W.F. (2009). Estimating value at risk of portfolio by conditional copula-GARCH method. *Insur.: Math. Econ.*, **45**, 315–324.
Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. London: Chapman & Hall.
Joe, H. (2014). *Dependence Modeling With Copulas*. Boca Raton, FL: CRC Press.
Joe, H. & Kurowicka, D. (2011). *Dependence Modeling: Vine Copula Handbook*. Singapore: World Scientific.
Joe, H. & Xu, J.J. (1996). The estimation method of inference functions for margins for multivariate models. In 166, University of British Columbia, Department of Statistics.
Jondeau, E. & Rockinger, M. (2006). The copula-GARCH model of conditional dependencies: An international stock market application. *J. Int. Money Finance*, **25**, 827–853.
Juang, B.H. & Rabiner, L.R. (1991). Hidden Markov models for speech recognition. *Technometrics*, **33**, 251–272.
Kendall, M.G. (1938). A new measure of rank correlation. *Biometrika*, **30**, 81–93.
Koki, C., Leonardos, S. & Piliouras, G. (2022). Exploring the predictability of cryptocurrencies via Bayesian hidden Markov models. *Res. Int. Bus. Finance*, **59**, 101554.
Kristoufek, L. (2013). Bitcoin meets Google trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Sci. Rep.*, **3**, 1–7.
Longin, F. & Solnik, B. (2002). Extreme correlation of international equity markets. *The J. Finance*, **56**, 649–676.
Maruotti, A. & Punzo, A. (2021). Initialization of hidden Markov and semi-Markov models: A critical evaluation of several strategies. *Int. Stat. Rev.*, **89**, 447–480.
McLachlan, G.J. & Peel, D. (2000). *Finite Mixture Models*. New York: Wiley.
Nasri, B.R. & Rémillard, B.N. (2019). Copula-based dynamic models for multivariate time series. *J. Multivariate Anal.*, **172**, 107–121.
Nasri, B.R., Rémillard, B.N. & Thioub, M.Y. (2020). Goodness-of-fit for regime-switching copula models with application to option pricing. *The Can. J. Stat.*, **48**, 79–96.
Nystrup, P., Lindström, E. & Madsen, H. (2020). Learning hidden Markov models with persistent states by penalizing jumps. *Expert Syst. Appl.*, **150**, 113307.
Okimoto, T. (2008). New evidence of asymmetric dependence structures in international equity markets. *J. Financial Quant. Anal.*, **43**, 787–816.
Ötting, M., Langrock, R. & Maruotti, A. (2021). A copula-based multivariate hidden Markov model for modelling momentum in football. *AStA Adv. Stat. Anal.*, 1–19. https://doi.org/10.1007/s10182-021-00395-8
Patton, A.J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *J. Financial Economet.*, **2**, 130–168.
Patton, A.J. (2012). A review of copula models for economic time series. *J. Multivariate Anal.*, **110**, 4–18.
Pennoni, F., Bartolucci, F., Forte, G. & Ametrano, F. (2021). Exploring the dependencies among main cryptocurrency log-returns: A hidden Markov model. *Econ. Notes*, **51**, e12193. https://doi.org/10.1111/ecno.12193
Pohle, J., Langrock, R., van Beest, F.M. & Schmidt, N.M. (2017). Selecting the number of states in hidden Markov models: Pragmatic solutions illustrated using animal movement. *J. Agricul., Biol. Environ. Stat.*, **22**, 270–293.
Politis, D.N. & Romano, J.P. (1994). The stationary bootstrap. *J. Am. Stat. Assoc.*, **89**, 1303–1313.
R Core Team 2023. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
Rémillard, B. 2011. Validity of the parametric bootstrap for goodness-of-fit testing in dynamic models. Available at SSRN 1966476, 1–43.
Remillard, B. (2013). *Statistical Methods for Financial Engineering*. CRC Press: Boca Raton, FL.
Rodriguez, J.C. (2007). Measuring financial contagion: A copula approach. *J. Empir. Finance*, **14**, 401–423.
Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.*, **6**, 461–464.
Shen, D., Urquhart, A. & Wang, P. (2020). Forecasting the volatility of Bitcoin: The importance of jumps and structural breaks. *Eur. Finan. Manag.*, **26**, 1294–1323.
Simard, C. & Rémillard, B. (2015). Forecasting time series with multivariate copulas. *Depend. Model.*, **3**, 59–82.
Sklar, A. (1959). Fonctions de répartition à *n* dimensions et leurs marges. *Publications de l'Institut Statistique de l'Université de Paris*, **8**, 229–231.
Stöber, J. & Czado, C. (2014). Regime switches in the dependence structure of multidimensional financial data. *Comput. Stat. Data Anal.*, **76**, 672–686.
Telli, S. & Chen, H. (2020). Structural breaks and trend awareness-based interaction in crypto markets. *Phys. A: Stat. Mech. Appl.*, **558**, 124913.
Theodossiou, P. (2015). Skewed generalized error distribution of financial assets and option pricing. *Multinatl. Finance J.*, **19**, 223–266.
Timmermann, A. (2018). Forecasting methods in finance. *Ann. Rev. Financial Econ.*, **10**, 449–479.
Tiwari, A.K., Aye, G.C., Gupta, R. & Gkillas, K. (2020). Gold-oil dependence dynamics and the role of geopolitical risks: Evidence from a Markov-switching time-varying copula model. *Energy Econ.*, **88**, 104748.

Trede, M. (2020). Maximum likelihood estimation of high-dimensional Student-*t* copulas. *Stat. Probab. Lett.*, **159**, 108678.
Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans. Inform. Theory*, **13**, 260–269.
Welch, L.R. (2003). Hidden Markov models and the Baum-Welch algorithm. *IEEE Inform. Theory Soc. Newslett.*, **53**, 1–13.
Zucchini, W., MacDonald, I.L. & Langrock, R. (2017). *Hidden Markov Models for Time Series: An Introduction Using R*. Boca Raton, FL: CRC Press.

APPENDIX A

Let us denote with $\boldsymbol{\theta}$ the full vector of parameters and with $\ell^*(\boldsymbol{\theta}|(\boldsymbol{y}_1, u_1), \ldots, (\boldsymbol{y}_T, u_T))$ the complete-data log-likelihood as in Equation (3). Starting from an initial parameter vector $\boldsymbol{\theta}^{(0)}$, at the m-iteration the algorithm performs the following steps:

- **E-step**: compute the posterior expected value of the indicator variables w_{tu} and z_{tuv} given by the quantities

$$\hat{w}_{tu} = P(u_t = u|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T),$$

$$\hat{z}_{tuv} = P(u_{t-1} = u, u_t = v|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T,),$$

for $t = 1, \ldots, T$, and for all $u, v = 1, \ldots, k$. We also define (see also Remillard, 2013; Nasri *et al.*, 2020)

$$\overline{\eta}_t(u) = P(u_t = u|\boldsymbol{y}_{t+1}, \ldots, \boldsymbol{y}_T), \quad t = 1, \ldots, T,$$

$$\eta_t(u) = P(u_t = u|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_t), \quad t = 2, \ldots, T,$$

where the conditioning argument disappears from the first expression for $t = T$. The above quantities are initialised as

$$\overline{\eta}_T(u) = 1/k, \; \eta_1(u) = \frac{\lambda_u c(\boldsymbol{y}_1; \boldsymbol{R}_u, \nu_u)}{\sum_{v=1}^k \pi_v c(\boldsymbol{y}_1; \boldsymbol{R}_v, \nu_v)}, \quad u = 1, \ldots, k,$$

and are computed recursively (Baum & Petrie, 1966; Welch, 2003; Nasri *et al.*, 2020) through

$$\eta_t(u) = \frac{c(\boldsymbol{y}_t; \boldsymbol{R}_u, \nu_u)\sum_{v=1}^k \eta_{t-1}(v)\pi_{u|v}}{\sum_{a=1}^k c(\boldsymbol{y}_t; \boldsymbol{R}_a, \nu_a)\sum_{v=1}^k \eta_{t-1}(v)\pi_{a|v}}, \quad t = 2, \ldots, T,$$

$$\overline{\eta}_t(u) = \frac{\sum_{v=1}^k \overline{\eta}_{t+1}(v)\pi_{v|u} c(\boldsymbol{y}_{t+1}; \boldsymbol{R}_v, \nu_v)}{\sum_{a=1}^k \sum_{v=1}^k \overline{\eta}_{t+1}(v)\pi_{v|a} c(\boldsymbol{y}_{t+1}; \boldsymbol{R}_v, \nu_v)}, \quad t = 1, \ldots, T-1,$$

to be evaluated in reverse order. From the previous expressions, we obtain $\hat{w}_{tu}$ and $\hat{z}_{tvu}$ as

$$\hat{w}_{tu} = \frac{\eta_t(u)\overline{\eta}_t(u)}{\sum_{a=1}^k \eta_t(a)\overline{\eta}_t(a)}, \quad t = 1, \ldots, T, \quad u = 1, \ldots, k,$$

$$\hat{z}_{tuv} = \frac{\pi_{v|u}\eta_{t-1}(u)\overline{\eta}_t(v)c(\boldsymbol{y}_t; \boldsymbol{R}_u, v_u)}{\sum_{a=1}^{k}\sum_{b=1}^{k}\pi_{b|a}\eta_{t-1}(a)\overline{\eta}_t(b)c(\boldsymbol{y}_t; \boldsymbol{R}_b, v_b)}, \quad t = 2, \ldots, T, \ u, v = 1, \ldots, k.$$

The expected value of the complete-data log-likelihood is then obtained by substituting $\hat{w}_{tu}$ and $\hat{z}_{tuv}$ to w_{tu} and z_{tuv}, in Equation (3). This expected value is denoted by $\mathcal{Q}(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m-1)})$ where $\boldsymbol{\theta}^{(m-1)}$ is the vector of parameters provided by the previous M-step, on the basis of which $\hat{w}_{tu}$ and $\hat{z}_{tuv}$ are computed.

- **M-step**. The new parameter vector $\boldsymbol{\theta}^{(m)}$ is obtained as $\text{argmax}_{\boldsymbol{\theta}}\mathcal{Q}(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m-1)})$. Parameters λ_u and $\pi_{v|u}$ are updated by using formulas in (4) and (5). Following Trede (2020), the updated values of the remaining parameters, $\boldsymbol{R}_u^{(m)}$ and $v_u^{(m)}$, are obtained as follows: for a given u, $u = 1, \ldots, k$, we maximise

$$\sum_{t=1}^{T}\hat{w}_{tu}\log c(\boldsymbol{y}_t; \boldsymbol{R}_u, v_u),$$

subject to the restriction that $\boldsymbol{R}_u$ is symmetric, positive definite, and with all diagonal elements equal to 1. The Lagrangians are the following

$$\mathcal{L}(\boldsymbol{R}_u|v_u) = \sum_{t=1}^{T}\hat{w}_{tu}\log c(\boldsymbol{y}_t; \boldsymbol{R}_u, v_u) + \sum_{j=1}^{r}\mu_j\left(\rho_u^{(jj)} - 1\right),$$

with $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_r)'$ being the Lagrange multipliers. Setting the first derivative with respect to $\boldsymbol{R}_u$ equal to 0 turns to

$$\frac{\partial\mathcal{L}(\boldsymbol{R}_u|v_u)}{\partial\boldsymbol{R}_u} = -\frac{\sum_t\hat{w}_{tu}}{2}\boldsymbol{R}_u^{-1} + \frac{v_u + r}{2v_u}\sum_{t=1}^{T}\hat{w}_{tu}\boldsymbol{R}_u^{-1}\boldsymbol{x}_t\boldsymbol{x}_t'\boldsymbol{R}_u^{-1}\left(1 + \frac{1}{v}\boldsymbol{x}_t'\boldsymbol{R}_u^{-1}\boldsymbol{x}_t\right)^{-1} + \boldsymbol{M} = \boldsymbol{0},$$

where we denote with $\boldsymbol{M} = \text{diag}(\boldsymbol{\mu})$ the matrix with diagonal elements equal to $\mu_1, \ldots, \mu_r$, and zero in all other positions, and with $\boldsymbol{x}_t$ the vector with components $T_{1,v_u}^{-1}\left(y_{tj}\right)$, with $j = 1, \ldots, r$. Multiplying both sides by $\boldsymbol{R}_u$ gives

$$\boldsymbol{R}_u = \frac{v_u + r}{v_u\sum_{t=1}^{T}\hat{w}_{tu}}\sum_{t=1}^{T}\hat{w}_{tu}\boldsymbol{x}_t\boldsymbol{x}_t'\left(1 + \frac{1}{v_u}\boldsymbol{x}_t'\boldsymbol{R}_u^{-1}\boldsymbol{x}_t\right)^{-1} + \frac{2}{\sum_t\hat{w}_{tu}}\boldsymbol{R}_u\boldsymbol{M}\boldsymbol{R}_u.$$

Let denote the first term of the previous expression with

$$\boldsymbol{A} = \frac{v_u + r}{v_u\sum_t\hat{w}_{tu}}\sum_{t=1}^{T}\hat{w}_{tu}\boldsymbol{x}_t\boldsymbol{x}_t'\left(1 + \frac{1}{v_u}\boldsymbol{x}_t'\boldsymbol{R}_u^{-1}\boldsymbol{x}_t\right)^{-1},$$

and let $\boldsymbol{a}$ be the vector of diagonal elements of $\boldsymbol{A}$. The Lagrange multipliers $\boldsymbol{\mu}$ satisfy the equation

$$\frac{2}{\sum_{t=1}^{T}\hat{w}_{tu}}(\boldsymbol{R}_u \circ \boldsymbol{R}_u)\boldsymbol{\mu} = \mathbf{1} - \boldsymbol{a},$$

with $\boldsymbol{R}_u \circ \boldsymbol{R}_u$ being the matrix of squared elements of $\boldsymbol{R}_u$. In order to meet the restrictions, the Lagrange multipliers are

$$\boldsymbol{\mu} = \sum_{t=1}^{T} \hat{w}_{tu}(\boldsymbol{R}_u \circ \boldsymbol{R}_u)^{-1}(\mathbf{1} - \boldsymbol{a})/2,$$

which, substituted in the equation for $\boldsymbol{R}_u$, yields to

$$\boldsymbol{R}_u = \boldsymbol{A} + \boldsymbol{R}_u \text{diag}\left[(\boldsymbol{R}_u \circ \boldsymbol{R}_u)^{-1}(\mathbf{1} - \boldsymbol{a})\right]\boldsymbol{R}_u.$$

The above equation cannot be solved analytically, and we consider the iterative solution as in Equation (7). The estimate for v_u is obtained by solving the optimisation problem in Equation (8).

In the application presented in Section 5, the number of latent states k is selected according to the ICL criterion (Biernacki *et al.*, 2000), defined as

$$\text{ICL} = -2\log \ell^*(\hat{\boldsymbol{\theta}}|(\boldsymbol{y}_1, u_1), \ldots, (\boldsymbol{y}_T, u_T)) + K\log(T), \tag{A1}$$

being $\hat{\boldsymbol{\theta}}$ the vector of estimated RSStC parameters and K the number of free parameters, computed as $K = (k - 1) + k(k - 1) + k(r(r + 1)/2) + k$. As suggested in Pohle *et al.* (2017), the unknown sequence $u_1, \ldots, u_T$ may be replaced with the decoded time-series $\hat{u}_1, \ldots, \hat{u}_T$ obtained applying the Viterbi algorithm to the posterior probabilities estimated with the selected RSStC model.

APPENDIX B

We report the simulation results regarding the 2-state RSStC model and the complete results for the 2- and for the 3-state model when varying the number of observations T or the number of assets r.

B.1 Two State Regime Switching Student-t Copula Model

We generate data from a 2-state RSStC model, and we simulate $B = 1{,}000$ samples of dimension $r = 5$, each with a total number of observations $T = 1{,}500$. We set the vector of initial probabilities $\boldsymbol{\lambda} = (1/2, 1/2)'$ so that each state is equally likely, and we fix the transition matrix as follows

$$\boldsymbol{\Pi} = \begin{bmatrix} 0.800 & 0.200 \\ 0.200 & 0.800 \end{bmatrix},$$

the matrices of dependence parameters as

$$\boldsymbol{R}_1 = \begin{bmatrix} 1.000 & - & - & - & - \\ 0.900 & 1.000 & - & - & - \\ 0.900 & 0.900 & 1.000 & - & - \\ 0.900 & 0.900 & 0.900 & 1.000 & - \\ 0.900 & 0.900 & 0.900 & 0.900 & 1.000 \end{bmatrix},$$

$$\boldsymbol{R}_2 = \begin{bmatrix} 1.000 & - & - & - & - \\ 0.200 & 1.000 & - & - & - \\ 0.000 & 0.000 & 1.000 & - & - \\ 0.100 & 0.200 & 0.100 & 1.000 & - \\ 0.000 & 0.200 & 0.200 & 0.000 & 1.000 \end{bmatrix},$$

and the state-specific numbers of degrees of freedom as $\nu_1 = 5$ and $\nu_2 = 15$, respectively. We simulate a turbulent market scenario with high asset correlation and fat tails, and a more stable market scenario with low dependencies and higher degrees of freedom. We evaluate the results using the criteria outlined in Section 4.

Table B1 presents the true value, bias, RMSE, and CI for each parameter. The maximum absolute bias for the initial probabilities is 0.033, and for the transition probabilities is approaching zero. The maximum RMSE is 0.499 for the initial probabilities and 0.018 for the transition probabilities. CIs for the transition probabilities are narrow and centered around the true values, and those of the initial probabilities reflect the unitary nature of the maximum likelihood estimator for these parameters (Zucchini *et al.*, 2017).

Table B2 presents simulation results for the dependence parameters $\rho_u^{(ij)}$ and the number of degrees of freedom ν_u. Regarding the dependence parameters, the maximum absolute bias is 0.002, and the highest RMSE occurs for $\rho_2^{(15)}$ and $\rho_2^{(45)}$, with a value of 0.041. The CIs are narrower in the first state, indicating more accurate estimation of high correlations. In terms of number of degrees of freedom, we observe that as ν_u decreases, the absolute bias and RMSE decrease while the CI narrows.

B.2 Increasing the Series Length and the Number of Assets

In this simulated scenario, we investigate the RMSE by varying the series length (T) and the number of assets (r). We simulate data from 2- and 3-state RSS*t*C models using the values of the

Table B1. *Simulation results for the 2-state RSStC model: true parameter value, bias, RMSE, lower and upper bounds of the* CIs *(*CI_L *and* CI_U*, respectively) of the initial and transition probabilities.*

	λ_1	λ_2	$\pi_{1\|1}$	$\pi_{1\|2}$	$\pi_{2\|1}$	$\pi_{2\|2}$
True	0.500	0.500	0.800	0.200	0.200	0.800
Bias	−0.033	0.033	0.000	0.000	0.000	0.000
RMSE	0.499	0.499	0.018	0.018	0.017	0.017
CI_L	0.000	0.000	0.764	0.168	0.168	0.764
CI_U	1.000	1.000	0.832	0.236	0.236	0.832

Table B2. *Simulation results for the 2-state RSStC model: true parameter value, bias, RMSE, lower and upper bounds of the* CIs *(*CI_L *and* CI_U*, respectively) of the dependence parameters, and of the number of degrees of freedom.*

State $u = 1$	$\rho_1^{(12)}$	$\rho_1^{(13)}$	$\rho_1^{(14)}$	$\rho_1^{(15)}$	$\rho_1^{(23)}$	$\rho_1^{(24)}$	$\rho_1^{(25)}$	$\rho_1^{(34)}$	$\rho_1^{(35)}$	$\rho_1^{(45)}$	ν_1
True	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900	5.000
Bias	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	−0.001	0.000	−0.157
RMSE	0.008	0.009	0.008	0.008	0.009	0.008	0.009	0.009	0.009	0.008	0.916
CI_L	0.882	0.881	0.883	0.882	0.880	0.882	0.881	0.882	0.882	0.882	3.724
CI_U	0.915	0.915	0.915	0.915	0.915	0.915	0.915	0.916	0.916	0.915	7.159
State $u = 2$	$\rho_2^{(12)}$	$\rho_2^{(13)}$	$\rho_2^{(14)}$	$\rho_2^{(15)}$	$\rho_2^{(23)}$	$\rho_2^{(24)}$	$\rho_2^{(25)}$	$\rho_2^{(34)}$	$\rho_2^{(35)}$	$\rho_2^{(45)}$	ν_2
True	0.200	0.000	0.100	0.000	0.000	0.200	0.200	0.100	0.200	0.000	15.000
Bias	0.002	0.000	0.001	0.001	0.002	0.001	0.001	0.001	−0.001	0.002	0.668
RMSE	0.038	0.040	0.040	0.041	0.040	0.037	0.037	0.039	0.038	0.041	3.476
CI_L	0.128	−0.077	0.024	−0.080	−0.077	0.132	0.125	0.022	0.123	−0.080	10.480
CI_U	0.274	0.083	0.175	0.077	0.081	0.270	0.274	0.180	0.275	0.075	23.502

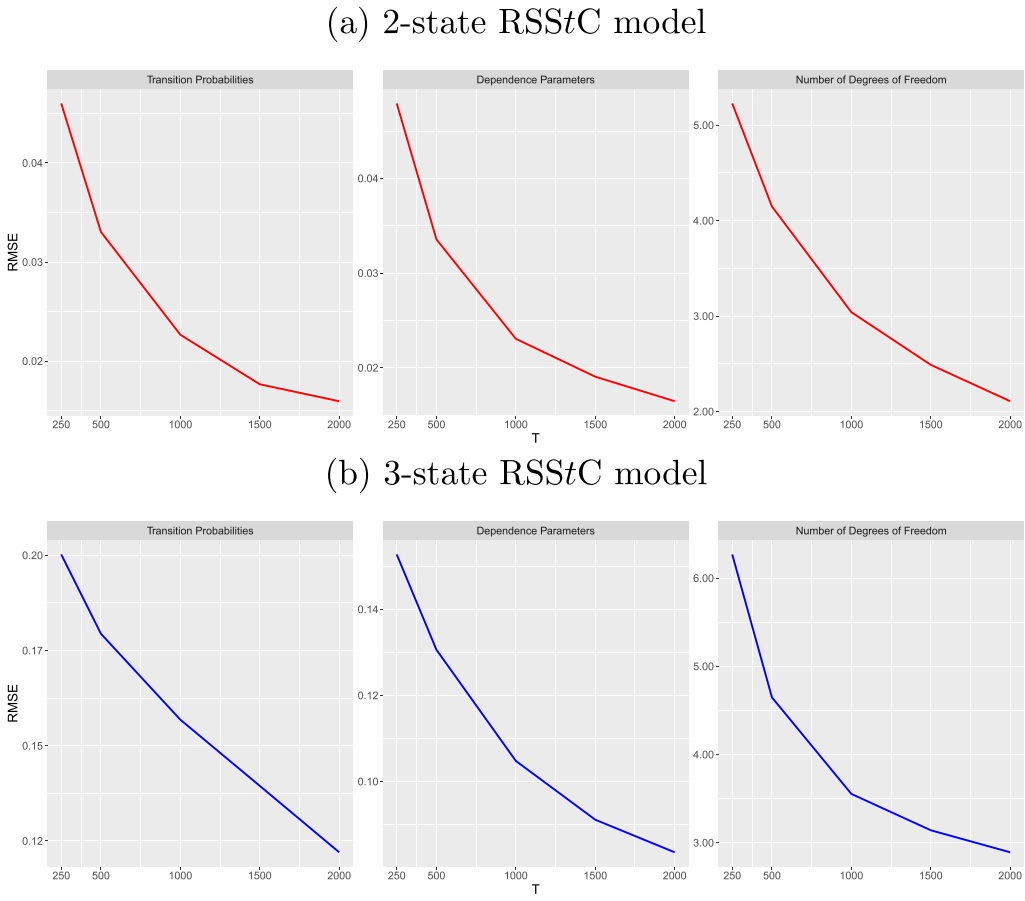

Figure B.1. *Average RMSE for transition probabilities, dependence parameters, and number of degrees of freedom in the 2-state (a) and 3-state (b) RSStC models, with series length (T) varying from 250 to 2,000 and* $r = 5$ *assets.*

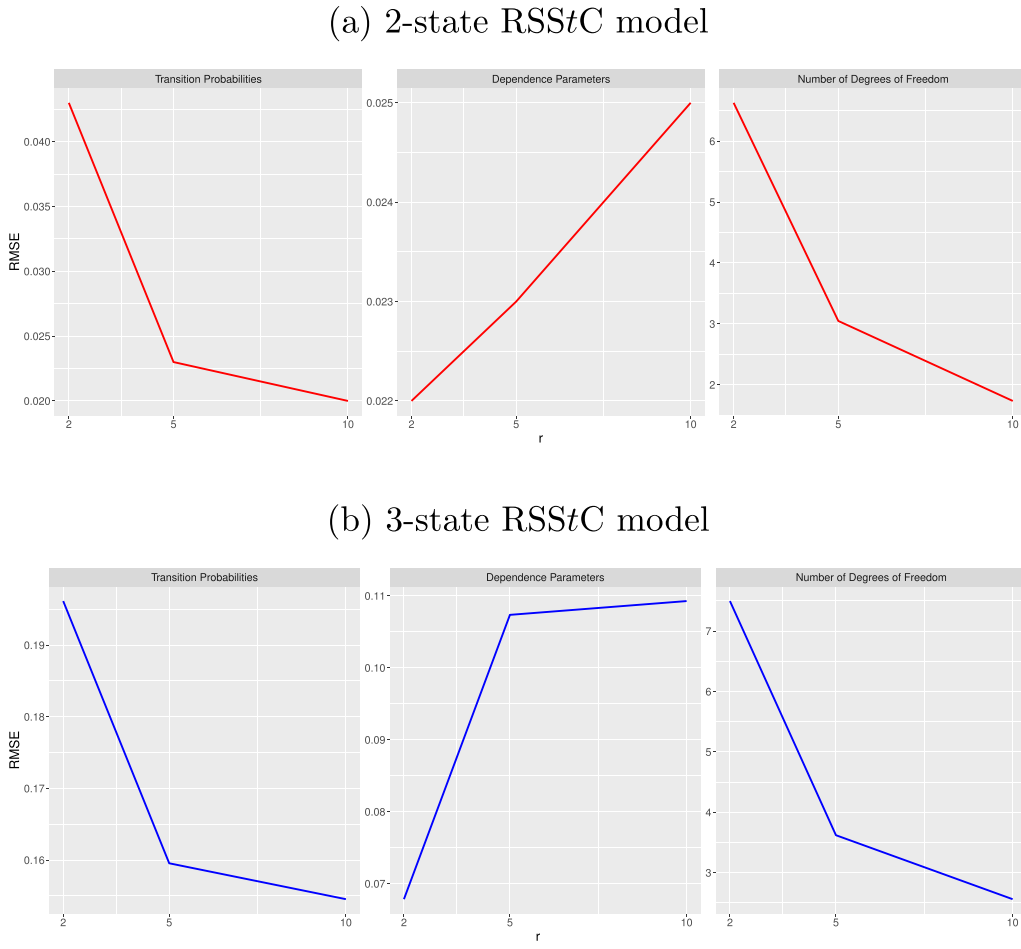

Figure B.2. *Average RMSE for transition probabilities, dependence parameters, and number of degrees of freedom in the 2-state (a) and 3-state (b) RSStC models, with number of assets (r) varying from 2 to 10 and* $T = 1,500$.

parameters employed for the simulated scenarios presented in Sections B.1 and 4.1. For each simulated sample, we calculate the average RMSE over $B = 1{,}000$ samples for each type of parameters, concerning the initial probabilities, the transition probabilities, the number of degrees of freedom, and the dependence parameters. We consider five different series of lengths $T = 250$, 500, 1,000, 1,500, 2,000, referred to $r = 5$ assets.

We summarise the results in Figure B.1, illustrating that the average RMSE decreases rapidly as the series length increases for all parameters. We omit the average RMSE for initial probabilities as it remains around 0.5 up to the fourth decimal digit.

Then, we increase the number of assets while keeping the series length fixed at $T = 1{,}000$. We consider three plausible values for r, namely, 2, 5, and 10, and we simulate $B = 1{,}000$ samples from the 2- and 3-state RSStC models. In the 2-state model, the initial and transition probabilities, as well as the vector of degrees of freedom, remain unchanged from the first simulation study. However, the dependence parameters are kept identical for all pairs of observations. Specifically, we use $\rho_1^{(ij)} = 0.9$ for the first regime and $\rho_2^{(ij)} = 0.1$ for the second regime, for $i \neq j$. For the 3-state RSStC model, we follow a similar procedure. The initial probabilities,

transition probabilities, and number of degrees of freedom remain the same as in the second simulation study. The dependence parameters are set equal to $\rho_1^{(ij)} = 0.9$, $\rho_2^{(ij)} = 0.5$, and $\rho_3^{(ij)} = 0.1$, for $i \neq j$. In Figure B.2, we observe that the average RMSE for the dependence parameters increases when the number of assets is higher than 2. However, the RMSE for the number of degrees of freedom and the transition probabilities decreases as the number of marginals increases.

[*Received December 2022; accepted December 2023*]