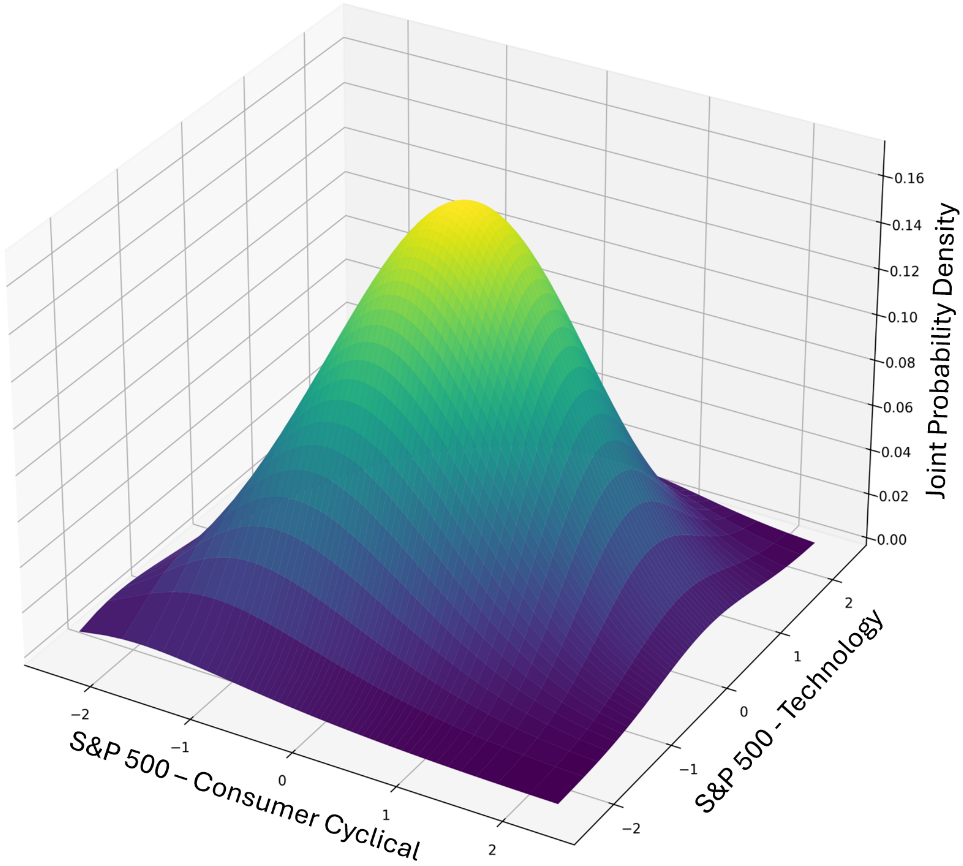
CHALMERS
UNIVERSITY OF TECHNOLOGY



Dynamic Dependence Among Economic Sectors in Equity Markets

Capturing lower-tail dependence using a Markov regime-switching copula model with transitions conditioned on a market-sentiment indicator

Master's thesis in Mathematics Engineering and Computational Science

MARCUS MÖLLER

DEPARTMENT OF MATHEMATICS
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2026
www.chalmers.se

MASTER'S THESIS 2026

Dynamic Dependence Among Economic Sectors in Equity Markets

Capturing lower-tail dependence using a Markov regime-switching copula model with transitions conditioned on a market-sentiment indicator

MARCUS MÖLLER

Department of Mathematics
Division of Applied Mathematics and Statistics
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2026

Dynamic Dependence Among Economic Sectors in Equity Markets
Capturing lower-tail dependence using a Markov regime-switching copula model with transitions conditioned on a market-sentiment indicator
MARCUS MÖLLER

Supervisor: Carl Lindberg, Department of Mathematics
Examiner: Peter Helgesson, Department of Mathematics

Master's Thesis 2026
Department of Mathematics
Division of Applied Mathematics and Statistics
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Illustrative graphic of a joint probability distribution between S&P 500 sector indices

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2026

Dynamic Dependence Among Economic Sectors in Equity Markets
Capturing lower-tail dependence using a Markov regime-switching copula model with transitions conditioned on a market-sentiment indicator
Marcus Möller
Department of Mathematics
Chalmers University of Technology

Abstract

This study investigates the dynamic dependence structure between equity market sectors using a Markov regime-switching copula framework. The analysis focuses on return distributions across economic sectors and their lower-tail dependence, particularly across different market regimes. The proposed model builds on empirical evidence that market drawdowns can be contagious and that pairwise dependence between assets tends to increase during periods of market stress. Modeling dynamic dependence can provide a more robust risk framework for portfolio evaluation and asset allocation. In this setting, dependence is allowed to vary over time through regime shifts and is linked to market sentiment.

This thesis extends *Bubbles and dependence between international equity markets* by Wuyi Ye, Lingbo Gao and Xiaoquan Liu (2024) by applying a similar framework to a different data sample. Specifically, the model is applied to equity market sectors (S&P 500 sub-indices: Industrials, Materials, Energy, Healthcare, Financials, Information Technology, and Consumer Non-Cyclical) rather than a geographic cross section, using daily returns over the period 1989-2026.

Empirically, the conditional regime-switching copula provides a better fit for many index pairs than an unconditional regime-switching copula and a static copula benchmark. However, in an out-of-sample asset allocation exercise, we are unable to replicate the economic gains reported in Wuyi Ye, Lingbo Gao and Xiaoquan Liu (2024): portfolios based on the conditional model do not consistently achieve higher risk-adjusted returns, evaluated by their Sharpe ratio, than equally weighted portfolios or portfolios constructed using a static copula model. Nonetheless, the model-based portfolios often exhibit lower maximum drawdowns, indicating that the framework can capture and mitigate some tail risk.

Keywords: Markov regime-switching, copula, tail dependence, equity sector indices, market contagion, bubble index, market regimes

Contents

1
Introduction

Understanding the interdependence between financial markets and assets has long been a central concern in both academia and practice. This is especially relevant during periods of market distress, where co-movements between markets tend to increase, amplifying systematic risk and challenging the notion of diversification, as shown in Cappiello, Engle and Sheppard (2006), Forbes and Chinn (2004), and Ang and Bekaert (2002). Correlation is a commonly used measure of dependence between assets, however, it has important limitations. In particular, it assumes stationarity over time and an identical dependence structure across the return distribution, assumptions that are often violated in financial data. Global market shocks following the 2008 financial crisis, the COVID-19 pandemic, and Trump's announced tariffs in April 2025, highlight the inadequacy of linear correlation measures and static models in capturing the complex and time-varying nature of financial markets.

Copula functions offer a flexible mathematical framework to model non-linear dependence between variables, given their marginal distributions. Copulas are widely used in financial econometrics to capture asymmetric dependence (e.g. Jondeau and Rockinger (2006), Christoffersen et al. (2012), Lucas et al. (2017) and Koopman et al. (2018)). Yet, a standard copula function assumes a static dependence structure over time, which is inconsistent with empirical evidence of time variation in dependence.

In response, dynamic models have been proposed, such as Bauwens and Otranto (2016), who explore several dynamic models with conditional correlation between stock returns and show that market volatility is a central factor in determining the dependence structure. A subset of the tested models is based on Markov regime-switching models, which perform well empirically and have a clear economic interpretation. This study builds upon the Markov regime-switching approach by assuming regime-specific copula parameters.

Moreover, Longin and Solnik (2001) provide evidence that dependence and correlation in financial markets are largely driven by market trends. Motivated by this, regime transition probabilities are conditioned on an exogenous proxy for market sentiment (the bubble index). Further, Brunnermeier and Oehmke (2013) recount that financial bubbles and asset price booms tend to form before periods of crisis, typically characterized by low volatility and eased monetary policy. To capture such dynamics, the bubble index defined by Wong (2011) is utilized. This index is constructed to measure deviations of asset prices from their long-term trend, acting as a proxy for speculative excess. By linking regime transitions to an observable variable, the model enhances economic interpretability and can be evaluated directly on observed data.

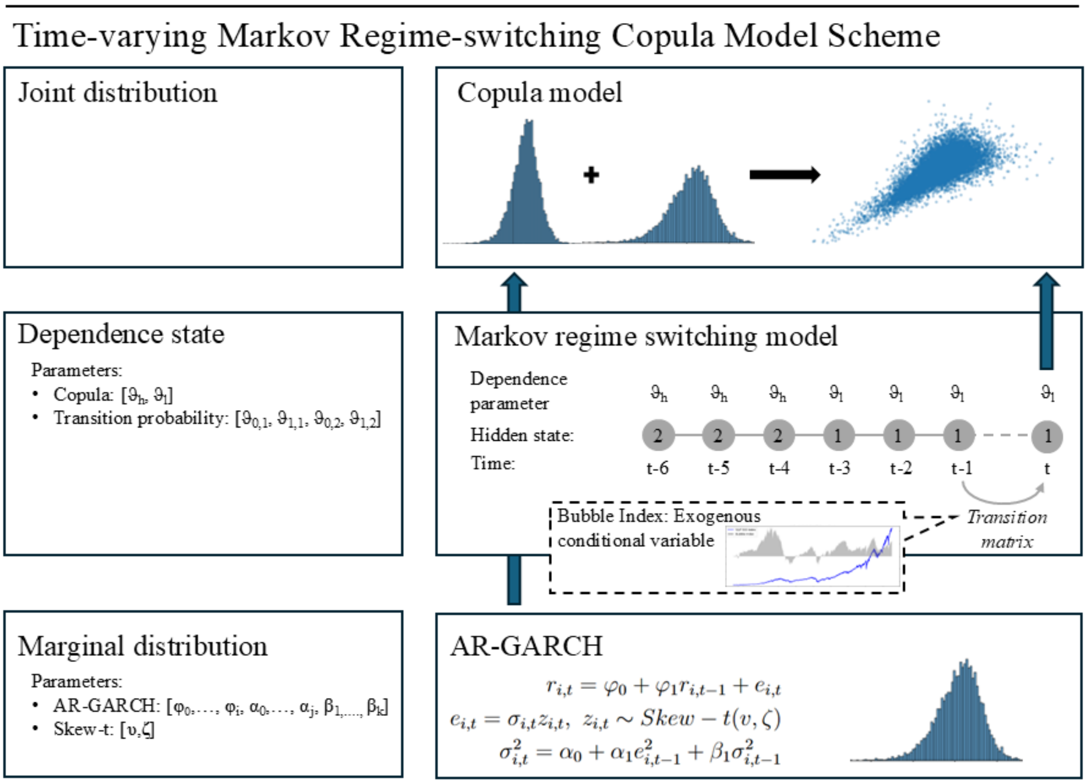


Figure 1.1: Schematic of the proposed time-varying Markov regime-switching copula model with transitions conditioned on an exogenous sentiment proxy (the bubble index). AR-GARCH models specify the marginal return distributions, while a Hidden Markov model governs regime dynamics and the associated copula parameter.

The analysis is carried out on sub-indices of the S&P 500, segmented by GICS sectors: Industrials, Energy, Materials, Healthcare, Financials, Information Technology, and Consumer Non-Cyclical. The sample consists of daily total returns in USD from 1989-09-11 to 2026-01-16. The indices are weighted by the constituents' float-adjusted market capitalization and rebalanced quarterly.

The empirical implementation proceeds in two steps. First, marginal return dynamics are estimated using AR-GARCH models, and the implied cumulative probabilities are used to construct copula observations. Second, dependence is modeled using a two-state Markov regime-switching copula, where regime-specific copula parameters capture changes in lower-tail dependence and transition probabilities vary with the bubble index. Model performance is assessed both statistically, through likelihood-based fit comparisons, and economically, through an out-of-sample portfolio allocation exercise comparing model-based allocations to simpler benchmark strategies.

Contributions. This thesis contributes by applying a sentiment-conditioned Markov regime-switching copula framework to S&P 500 sector indices, extending earlier applications on international equity markets. Evaluating whether conditioning regime transitions on an observable sentiment proxy improves empirical fit. Finally, quantifying the economic value of the resulting dependence forecasts in an out-of-sample portfolio allocation setting.

2

Theory

The following section describes the model framework, and the following statistics that are used to evaluate the model's goodness-of-fit.

2.1 The Bubble index

A price bubble is a market phenomenon in which asset prices rise substantially above levels justified by fundamentals, driven primarily by investor behaviour and expectations of further price appreciation. Such episodes are often followed by sharp drawdowns when the bubble collapses, Blasques et al. (2022). In practice, identifying bubbles and distinguishing them from rationally high valuations is difficult. A range of indicators has therefore been proposed to proxy market sentiment or valuation pressure, including the Buffett indicator, the Fear and Greed index, the CAPE ratio, and statistical approaches such as LPPLS and GSADF.

The sentiment proxy used in this thesis aims to measure deviations of the observed price from a time-varying equilibrium level. Wong (2011) proposes estimating this equilibrium level using a filtered moving-average benchmark computed over a dynamic window size. The window length is defined as

$$m = \min\left[\frac{\operatorname{stdev}(S_t, S_{t-1}, \ldots, S_{t-1000})}{\operatorname{stdev}(S_t, S_{t-1}, \ldots, S_{t-2000})} \cdot 2000,\ 2000\right]. \tag{2.1}$$

A window of 2000 trading days corresponds to roughly eight years, which is shorter than the time between major crises such as 1987, 1997, and 2007. The window length is scaled by the ratio of a four-year standard deviation to an eight-year standard deviation. Intuitively, if recent volatility is low relative to longer-run volatility, the ratio decreases and the window shrinks, making the benchmark more local. Conversely, if recent volatility is elevated, the window expands toward the cap of 2000 days.

The bubble index is constructed as follows:

1. Compute the logarithmic daily return

   $$r_t = \ln\left(\frac{S_t}{S_{t-1}}\right),$$

   where S_t denotes the index level at day t.
2. Set the window length m according to (2.1).

3. Within the window, filter out extreme returns by truncating observations below the 5th percentile and above the 95th percentile:

$$\hat{r}_t = \begin{cases} r_t, & \text{if } P_5(\{r_{t-i}\}_{i=0}^m) < r_t < P_{95}(\{r_{t-i}\}_{i=0}^m)\,, \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

4. Construct the "filtered" price

$$P_{t-1} = \frac{S_t}{\exp(\hat{r}_t)}.$$

5. Define the equilibrium benchmark price as the moving average of the filtered price over the window:

$$\mu_t = \frac{1}{m}\sum_{i=1}^{m} P_{t-i+1}.$$

6. Define the bubble index as the proportional deviation from the benchmark,

$$B_t = \frac{S_t}{\mu_t} - 1.$$

2.2 Independent Asset Dynamics

In the proposed framework, a copula function combines marginal distributions into a joint multivariate distribution. The first step is therefore to estimate the marginal dynamics of sector index returns. Each marginal is modeled using an AR-GARCH specification for the conditional mean and variance, where autoregressive and GARCH orders up to five lags are considered. The final model is selected using the Bayesian Information Criterion (BIC). Empirically, the majority of indices are best described by an AR(0)-GARCH(1,1) specification. The general AR(m)-GARCH(p, q) model is given by

$$\begin{aligned} r_{i,t} &= \varphi_0 + \sum_{j=1}^{m} \varphi_j r_{i,t-j} + e_{i,t}, \\ e_{i,t} &= \sigma_{i,t} z_{i,t}, \qquad z_{i,t} \sim \text{Skew-}t(\nu, \zeta) \\ \sigma_{i,t}^2 &= \alpha_0 + \sum_{k=1}^{p} \alpha_k e_{i,t-k}^2 + \sum_{l=1}^{q} \beta_l \sigma_{i,t-l}^2 \end{aligned} \tag{2.3}$$

where $\{\varphi_j\}_{j=0}^m$ are autoregressive parameters and $\{\alpha_k\}_{k=0}^p$ and $\{\beta_l\}_{l=1}^q$ are GARCH parameters. Here, $r_{i,t}$ denotes the daily log return of sector i at time t.

To allow for skewness and heavy tails in returns, the innovations $z_{i,t}$ are assumed to follow a skewed Student-t distribution. Using the parametrization in (2.4), its density is

$$
\begin{aligned}
q(x \mid \nu, \zeta) &= \xi\kappa \left(1 + \frac{1}{\nu - 2}\left(\frac{\xi x + \alpha}{1 + \zeta \operatorname{sign}\left(x + \frac{\alpha}{\xi}\right)}\right)^2\right)^{-\frac{\nu+1}{2}}, \\
\alpha &= 4\zeta\frac{\nu - 2}{\nu - 1}, \\
\xi &= 1 + 3\zeta^2 - \alpha^2, \\
\kappa &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu - 2)}\,\Gamma\left(\frac{\nu}{2}\right)}.
\end{aligned}
\tag{2.4}
$$

A key characteristic of the skew-t distribution is that it allows for both skewness and heavier tails relative to the normal distribution, while being standardized to have mean zero and unit variance. When $\zeta = 0$, it reduces to the Student-t distribution, and as $\nu \to \infty$ it converges to the normal distribution.

Skewness and kurtosis summarize asymmetry and tail heaviness, respectively. For a random variable X with mean $\mu = \mathbf{E}[X]$ and variance $\sigma^2 = \mathbf{E}[(X - \mu)^2]$, they are defined as

$$
\begin{aligned}
S &= \frac{\mathbf{E}[(X - \mu)^3]}{\sigma^3}, \\
K &= \frac{\mathbf{E}[(X - \mu)^4]}{\sigma^4}.
\end{aligned}
\tag{2.5}
$$

Skewness measures the asymmetry of a distribution around its mean, a normal distribution is symmetric and therefore has skewness equal to zero. Kurtosis measures tail heaviness, higher kurtosis indicates a higher prevalence of extreme outcomes. The normal distribution has kurtosis equal to three.

2.3 Copulas

A copula is a multivariate cumulative distribution function (CDF) with uniform marginal distributions on $[0, 1]$. Copulas are used to model dependence between random variables separately from their marginal behaviour. By Sklar's theorem (Sklar, 1959), any multivariate joint distribution can be written in terms of its univariate marginals and a copula that captures the dependence structure. This representation allows the modeling problem to be separated into estimating the marginals and estimating the copula.

In the bivariate case, let X and Y have marginal CDFs F and G, and define $U = F(X)$ and $V = G(Y)$. Then $U, V \sim \mathcal{U}(0, 1)$ and there exists a copula C such that

$$C(u, v) = \mathbb{P}(U \leq u, V \leq v) = \mathbb{P}(X \leq F^{-1}(u),\, Y \leq G^{-1}(v)).$$

Equivalently, the joint CDF H of (X, Y) can be written as $H(x, y) = C(F(x), G(y))$.

Definition 1. A function $C : [0,1]^d \to [0,1]$ is a d-dimensional copula if C is a CDF on $[0,1]^d$ satisfying:

1. $C(u_1, \ldots, 0, \ldots, u_d) = 0$ whenever any entry equals 0.
2. $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$ for each $i = 1, \ldots, d$.
3. C is d-increasing: for every rectangle $B = \prod_{i=1}^d [x_i, y_i] \subseteq [0,1]^d$, the C-volume is non-negative, i.e. $V_C(B) \geq 0$.

Applications often focus on tail dependence, i.e. the tendency of variables to experience joint extreme events. The (upper and lower) tail dependence coefficients are defined by

$$\begin{aligned} \lambda_u &= \lim_{u \to 1^-} \mathbb{P}\Big(Y > G^{-1}(u) \,\Big|\, X > F^{-1}(u)\Big)\,, \\ \lambda_l &= \lim_{u \to 0^+} \mathbb{P}\Big(Y \leq G^{-1}(u) \,\Big|\, X \leq F^{-1}(u)\Big)\,. \end{aligned} \tag{2.6}$$

Hence, $\lambda_u, \lambda_l \in [0,1]$ measure limiting probabilities of joint upper- and lower-tail events, respectively.

2.3.1 Sklar's Theorem

To represent a multivariate joint distribution using its univariate marginals and a copula, the copula must capture all cross-sectional dependence, and the joint distribution must be recoverable from the decomposition. Sklar's theorem (Sklar (1959)) establishes these results.

Theorem 1. *Sklar (1959): Let H be an d-dimensional joint distribution function with univariate marginal distribution functions $F_1, \ldots, F_d$. Then there exists an d-copula $C : [0,1]^d \to [0,1]$ such that*

$$H(x_1, \ldots, x_d) = C\Big(F_1(x_1), \ldots, F_d(x_d)\Big), \qquad \forall (x_1, \ldots, x_d) \in \mathbb{R}^d. \tag{2.7}$$

Furthermore, C is uniquely determined on $\mathrm{Range}(F_1) \times \cdots \times \mathrm{Range}(F_d)$, and in particular is unique if all marginals are continuous.

Consequently, any multivariate CDF H can be decomposed into its marginals and a copula. Under the assumption of continuous marginal return distributions, the copula describing the dependence structure is also unique.

Moreover, if the joint distribution admits a density h and each marginal admits a density f_i, then the joint density factorizes by the chain rule

$$h(x_1, \ldots, x_d) = \frac{\partial^d H(x_1, ..., x_d)}{\partial x_1 ... \partial x_d} = c\Big(F_1(x_1), \ldots, F_n(x_d)\Big) \prod_{i=1}^d f_i(x_i), \tag{2.8}$$

where c denotes the copula density.

Proof. (continuous marginals), Durante, Fernandez-Sanchez and Sempi (2013):

Let $X = (X_1, \dots, X_d)$ have joint CDF H and continuous marginals $F_1, \dots, F_d$. Define the generalized inverse

$$F_i^{-1}(u) \equiv \inf\{x \in \mathbb{R} : F_i(x) \geq u\}, \qquad u \in [0,1],$$

and define $C : [0,1]^d \to [0,1]$ by

$$C(u_1, \dots, u_d) = H\Big(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\Big).$$

We verify that C is a copula:

1. If some coordinate is zero $u_j = 0$, then $F_j^{-1}(0) = -\infty$, and hence
$$C(u_1, \dots, 0, \dots, u_d) = H(\dots, -\infty, \dots) = 0.$$
2. For any i, since $F_k^{-1}(1) = +\infty$ for $k \neq i$,
$$C(1, \dots, 1, u_i, 1, \dots, 1) = H(+\infty, \dots, F_i^{-1}(u_i), \dots, +\infty) = \mathbb{P}(X_i \leq F_i^{-1}(u_i)) = F_i(F_i^{-1}(u_i))$$
By continuity of F_i, we have $F_i(F_i^{-1}(u_i)) = u_i$, so $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$.
3. *d-increasing.* For any rectangle $[a,b] = \prod_{i=1}^d [a_i, b_i] \subset [0,1]^d$, let $x_i^- := F_i^{-1}(a_i)$ and $x_i^+ := F_i^{-1}(b_i)$. Then the C-volume of the rectangle is
$$V_C([a,b]) = \int_{[a,b]} dC = \mathbb{P}\Big(\bigcap_{i=1}^{d} \{x_i^- < X_i \leq x_i^+\}\Big) \geq 0,$$
so C is d-increasing.

Finally, for any $x = (x_1, \dots, x_d) \in \mathbb{R}^d$,

$$C(F_1(x_1), \dots, F_n(x_d)) = H\Big(F_1^{-1}(F_1(x_1)), \dots, F_d^{-1}(F_d(x_d))\Big) = H(x_1, \dots, x_d),$$

since continuity implies $F_i^{-1}(F_i(x_i)) = x_i$. Thus (2.7) holds. Uniqueness follows because C is determined on $\mathrm{Range}(F_1) \times \cdots \times \mathrm{Range}(F_d)$, and this set equals $[0,1]^d$ when the marginals are continuous. □

2.3.2 Clayton Copula

The Clayton copula is widely used in financial applications because it captures *lower-tail dependence*, i.e. the tendency for joint extreme losses. By construction, the Clayton copula exhibits zero upper-tail dependence. This feature is often considered suitable in equity applications, where dependence is frequently stronger during market downturns than during upturns. For example, Fortin and Kuzmics (2002) document relatively stronger lower-tail dependence and weaker upper-tail dependence for stock index returns.

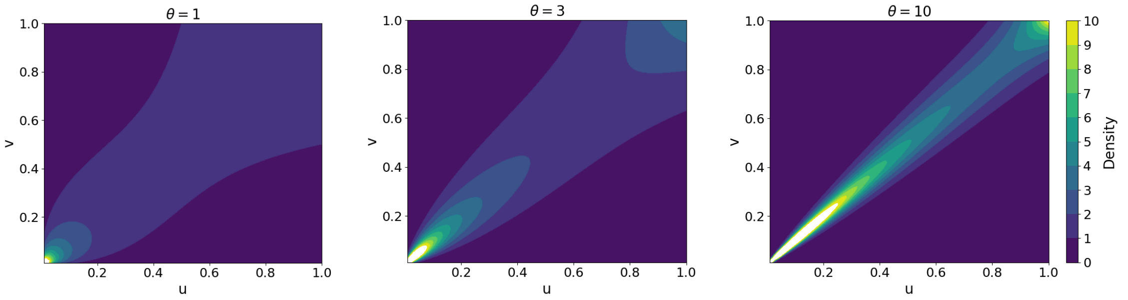
The bivariate Clayton copula with parameter $\theta > 0$ is defined by

$$C(u, v; \theta) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta}, \qquad u, v \in (0,1),\ \theta > 0. \tag{2.9}$$

Its tail dependence coefficients satisfy $\lambda_u = 0$ and $\lambda_l = 2^{-1/\theta}$. The independence copula is obtained in the limit $\theta \to 0$, in which case $C(u, v; \theta) \to uv$.

The corresponding copula density is obtained by differentiating the copula in respect to both arguments

$$c(u, v; \theta) = \frac{\partial^2}{\partial u\,\partial v} C(u, v; \theta) = (1 + \theta)(uv)^{-1-\theta} \left(u^{-\theta} + v^{-\theta} - 1\right)^{-2-\frac{1}{\theta}}, \qquad \theta > 0. \tag{2.10}$$



(a) $\theta = 1$ **(b)** $\theta = 3$ **(c)** $\theta = 10$

Figure 2.1: Clayton copula density for different values of θ. Larger θ implies stronger dependence, particularly in the lower tail.

2.4 Markov Chain

A Markov chain is a stochastic process $\{X_t\}_{t \geq 0}$ that satisfies the *Markov property*: conditional on the present state, the future is independent of the past. Since both the return dynamics and the regime process in this thesis are modeled in discrete time (daily or weekly), the regime indicator is naturally represented as a discrete-time Markov chain:

$$\mathbb{P}(X_{t+1} = x_{t+1} \mid X_t = x_t, X_{t-1} = x_{t-1}, \ldots, X_0 = x_0) = \mathbb{P}(X_{t+1} = x_{t+1} \mid X_t = x_t). \tag{2.11}$$

The set of possible values of a Markov chain is its *state space* $\mathcal{S}$. Depending on the application, $\mathcal{S}$ may be finite, countably infinite, or continuous. For the process $\{X_t\}$, we have:

- $X_t \in \mathcal{S}$ for all t, where $\mathcal{S}$ contains all possible states.
- **Finite:** e.g. $\mathcal{S} = \{1, \ldots, N\}$, a finite set of states.
- **Countable:** e.g. $\mathcal{S} = \mathbb{N}$, a countably infinite set of discrete states.
- **Continuous:** e.g. $\mathcal{S} = \mathbb{R}$, an uncountable state space; the process is then typically referred to as a (continuous-state) Markov process.

For a finite state space $\mathcal{S} = \{1, \ldots, N\}$, transition probabilities are summarized by the transition matrix $\mathbf{P}$, defined by

$$\mathbf{P} = (p_{ij})_{i,j=1}^{N}, \qquad p_{ij} = \mathbb{P}(X_{t+1} = j \mid X_t = i). \tag{2.12}$$

The matrix satisfies the stochastic constraints

$$p_{ij} \geq 0, \qquad \sum_{j=1}^{N} p_{ij} = 1 \quad \text{for each } i = 1, \dots, N. \tag{2.13}$$

2.4.1 Regime switching state model

Originally proposed in statistics by Lindgren (1978) and introduced to econometrics by Hamilton (1989), regime-switching models are widely used to capture structural changes in time series. The underlying idea is that the distributional properties of an observed process may depend on an unobserved *state* (or *regime*). In financial markets, "bull" and "bear" markets are common examples of regimes, characterized by persistently rising and falling prices, respectively. Under the model assumptions, the regime process can be represented as a Markov chain, where the probability of switching regimes depends only on the most recent regime. Since the regime is typically not directly observed, it is treated as a latent state and the resulting model is a Hidden Markov model (HMM).

Following Ye, Gao and Liu (2024), the dependence between indices is captured through a copula parameter that is allowed to be regime-specific. Econometrically, one regime is interpreted as a *high-dependence* regime and the other as a *low-dependence* regime. This is consistent with evidence that dependence tends to increase in turbulent market conditions (see, e.g., Ang and Bekaert (2002)). Let s_t denote the latent regime indicator at time t:

$$\begin{aligned} &s_t \in \{1, 2\}, \\ &\theta_t = \begin{cases} \theta_h, & s_t = 1, \\ \theta_l, & s_t = 2, \end{cases} \end{aligned} \tag{2.14}$$

where θ_h and θ_l denote the copula parameters in the high- and low-dependence regimes, respectively. Hence, the copula parameter is time-varying through the latent regime process $\{s_t\}$, which is assumed to follow an HMM.

2.4.2 Conditional transition probability

In the two-state model, let $p_{ii,t}$ denote the probability of remaining in state i from $t-1$ to t, and let $p_{ij,t}$ denote the probability of switching from state i to state j ($i \neq j$). For a two-state chain, $p_{ij,t} = 1 - p_{ii,t}$. Transition probabilities are modeled as a logistic function of the lagged bubble index B_{t-1}:

$$p_{ii,t} = \mathbb{P}(s_t = i \mid s_{t-1} = i, B_{t-1}) = \frac{\exp(\theta_{0,i} + \theta_{1,i} B_{t-1})}{1 + \exp(\theta_{0,i} + \theta_{1,i} B_{t-1})}, \qquad i \in \{1, 2\}. \tag{2.15}$$

The parameters $(\theta_{0,i}, \theta_{1,i})$ determine the baseline persistence in regime i and the sensitivity of persistence to the bubble index, respectively. In this way, regime transitions are conditioned on an observable proxy for market sentiment.

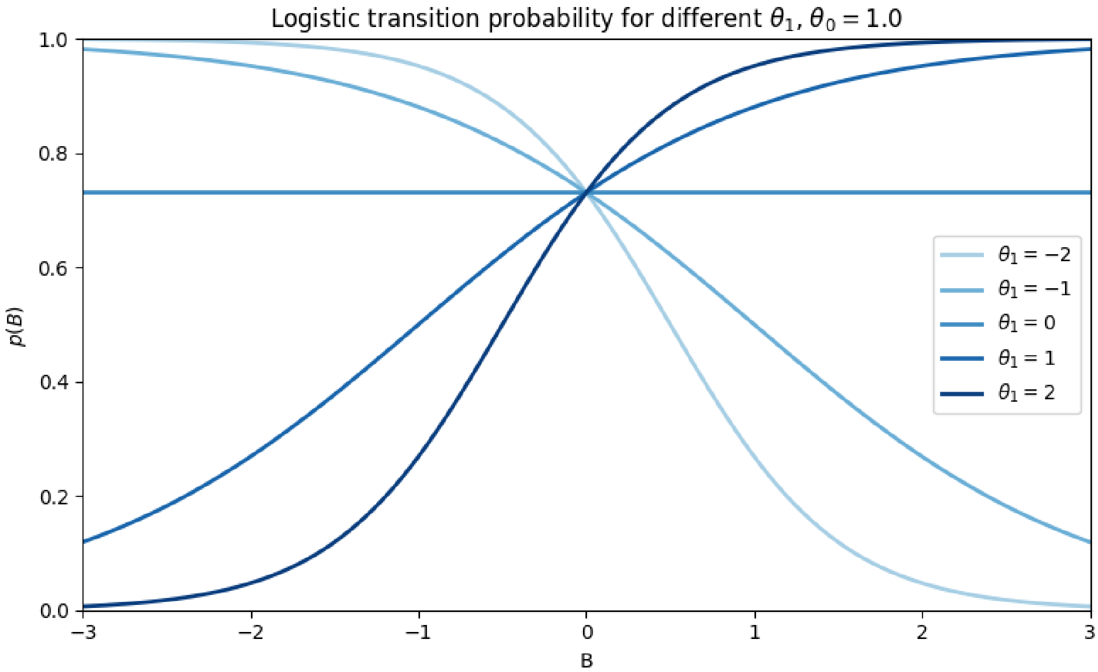


Figure 2.2: Transition probability $p_{ii,t}$ as a function of B_{t-1} for $\theta_{0,i} = 1$ and different values of $\theta_{1,i}$.

As illustrated in Figure 2.2, $\theta_{0,i}$ determines the baseline probability of remaining in state i when $B_{t-1} = 0$:

$$p_{ii,t}\Big|_{B_{t-1}=0} = \frac{1}{1 + \exp(-\theta_{0,i})}.$$

The parameter $\theta_{1,i}$ governs how the bubble index affects the transition probabilities: if $\theta_{1,i} > 0$, higher values of B_{t-1} increase $p_{ii,t}$ (making regime i more persistent), whereas if $\theta_{1,i} < 0$, higher values of B_{t-1} decrease $p_{ii,t}$ (making switches away from regime i more likely).

2.5 Maximum likelihood estimation

Maximum likelihood estimation (MLE) is a method for estimating parameters of a parametric model from observed data. Given a statistical model parametrized by a vector θ, the MLE chooses the parameter value that maximizes the probability density of the observed sample under the model.

Let x denote an observation with density $f(x \mid \theta)$. For an observed sample $\{x_t\}_{t=1}^{n}$, the likelihood function is

$$L(\theta) = \prod_{t=1}^{n} f(x_t \mid \theta), \tag{2.16}$$

and the MLE is defined by

$$\hat{\theta} = \arg\max_{\theta} \; L(\theta). \tag{2.17}$$

In practice, estimation is carried out using the log-likelihood,

$$\ell(\theta) = \log L(\theta) = \sum_{t=1}^{n} \log f(x_t \mid \theta), \tag{2.18}$$

which is numerically more stable. The maximizer of $\ell(\theta)$ equals the maximizer of $L(\theta)$ because the logarithm is strictly increasing.

2.5.1 Likelihood under a copula model

Consider a bivariate vector (X_t, Y_t) with marginal CDFs $F(\cdot \mid \theta_x)$ and $G(\cdot \mid \theta_y)$ and corresponding densities $f(\cdot \mid \theta_x)$ and $g(\cdot \mid \theta_y)$. Let the copula density be $c(\cdot, \cdot \mid \theta_c)$. By Sklar's theorem and the factorization in (2.8), the joint density of (X_t, Y_t) can be written as

$$h(x_t, y_t \mid \theta_x, \theta_y, \theta_c) = c(F(x_t \mid \theta_x), G(y_t \mid \theta_y) \mid \theta_c)\, f(x_t \mid \theta_x)\, g(y_t \mid \theta_y). \tag{2.19}$$

Define pseudo-observations

$$u_t \equiv F(x_t \mid \theta_x), \qquad v_t \equiv G(y_t \mid \theta_y).$$

Then the sample log-likelihood for the joint model becomes

$$\begin{aligned} \ell(\theta_x, \theta_y, \theta_c) &= \sum_{t=1}^{n} \log h(x_t, y_t \mid \theta_x, \theta_y, \theta_c) \\ &= \sum_{t=1}^{n} \log c(u_t, v_t \mid \theta_c) + \sum_{t=1}^{n} \log f(x_t \mid \theta_x) + \sum_{t=1}^{n} \log g(y_t \mid \theta_y). \end{aligned} \tag{2.20}$$

This decomposition motivates a two-step estimation procedure. First estimate the marginal parameters (θ_x, θ_y), then estimate the copula parameters θ_c conditional on the estimated marginals.

Step 1: Marginal estimation: The marginal parameters can be estimated by maximizing the marginal log-likelihoods:

$$\begin{aligned} \hat{\theta}_x &= \arg\max_{\theta_x} \sum_{t=1}^{n} \log f(x_t \mid \theta_x), \\ \hat{\theta}_y &= \arg\max_{\theta_y} \sum_{t=1}^{n} \log g(y_t \mid \theta_y). \end{aligned} \tag{2.21}$$

In this thesis, each marginal is modeled using an AR(m)-GARCH(p, q) specification with skew-t innovations. For an AR(m)-GARCH(p,q) series the recursion yielding the pseudo-observations u_t is

$$\begin{aligned} e_t &= x_t - \left(\varphi_0 + \sum_{j=1}^{m} \varphi_j x_{t-j} \right), \\ \sigma_t^2 &= \alpha_0 + \sum_{k=1}^{p} \alpha_k e_{t-k}^2 + \sum_{l=1}^{q} \beta_l \sigma_{t-l}^2, \\ z_t &= \frac{e_t}{\sigma_t}, \\ u_t &= F(x_t \mid \hat{\theta}_x) = F_z(z_t), \end{aligned} \tag{2.22}$$

where F_z denotes the CDF of the standardized skew-t innovation distribution. The recursion is initialized using the unconditional variance,

$$\mathbb{E}[\sigma_t^2] = \frac{\alpha_0}{1 - \sum_{k=1}^{p} \alpha_k - \sum_{l=1}^{q} \beta_l}, \tag{2.23}$$

which requires $\sum_{k=1}^{p} \alpha_k + \sum_{l=1}^{q} \beta_l < 1$.

Step 2: Copula estimation: Let θ_c denote the parameters of the copula model. In the regime-switching setting considered,

$$\theta_c = \{\theta_h, \theta_l, \theta_{0,1}, \theta_{1,1}, \theta_{0,2}, \theta_{1,2}\},$$

where θ_h and θ_l are the regime-specific copula parameters and $(\theta_{0,i}, \theta_{1,i})$ govern the conditional transition probabilities for $i \in \{1, 2\}$. The copula parameters are then estimated by maximizing the copula log-likelihood,

$$\hat{\theta}_c = \arg\max_{\theta_c} \sum_{t=1}^{n} \log c(u_t, v_t \mid \theta_c), \tag{2.24}$$

where (u_t, v_t) are the pseudo-observations constructed from the estimated marginal models.

2.6 Evaluation measures

To evaluate the model, we use several statistics to assess the goodness-of-fit of AR-GARCH model for the marginal sector index returns and the full copula-based dependence model.

2.6.1 Marginal distributions

- **Jarque-Bera test (JB)**: Presented by Jarque and Bera (1987), the Jarque-Bera test is a hypothesis test for normality. It assesses whether a sample has skewness and kurtosis consistent with a normal distribution (skewness: $S = 0$ and kurtosis $K = 3$). The JB statistic is defined by:

$$JB = \frac{n}{6}(S^2 + \frac{(K-3)^2}{4}) \tag{2.25}$$

Under the null hypothesis of normality, the JB statistic follows a chi-squared distribution with 2 degrees of freedom. A low p-value then indicates rejection of the null hypothesis, and normality is rejected.

- **Ljung-Box test (LB)**: The Ljung-Box test is a hypothesis test used to detect autocorrelation in a time series, up to a chosen lag. The null hypothesis is that the h first autocorrelations are 0 ($H_0 : \rho_0 = \rho_1 = ... = \rho_h = 0$).

$$LB = n(n+2)\sum_{k=1}^{h}\frac{\hat{\rho}_k^2}{n-k} \tag{2.26}$$

where n is the sample size, h number of considered lags, and $\hat{\rho}_k$ the sample autocorrelation at lag k. Shown by Ljung and Box (1978) the LB statistic follows a chi-squared distribution with v degrees of freedom, where v is the number of lags h - estimated paramters in the AR/MA model. A high LB value implies rejection of the null hypothesis and there is autocorrelation in the sample.

- **ARCH Lagrange Multiplier test (ARCH LM)**: The ARCH Lagrange multiplier test, introduced by Engle (1982), assesses the presence of autoregressive conditional heteroscedasticity in the residuals of a time series model. Under the null hypothesis there are no ARCH effects, and the GARCH-model captures the conditional heteroscedasticity.

$$\begin{aligned} \hat{\epsilon}_t^2 &= \phi_0 + \sum_{i=1}^{q}\phi_i\hat{\epsilon}_{t-i}^2 + \eta_t \\ LM &= n*(1 - \frac{\sum_{t=1}^{n}\eta_t^2}{\sum_{t=1}^{n}(\hat{u}_t - \bar{u})}) \end{aligned} \tag{2.27}$$

The test statistic LM is chi-squared distributed with q (lags) degrees of freedom. A high LM implies rejected null hypothesis and further ARCH behaviour in the residuals.

- **Augmented Dickey-Fuller test (ADF)**: The ADF test, formulated by Dickey and Fuller (1979), is a hypothesis test, assessing whether a time series is stationary. The null hypothesis H_0 is that there is a unit root present in the time series sample. Consider an AR(1) process ($r_t = \varphi_o + \varphi_1 r_{t-1} + e_{i,t}$) where a unit root corresponds to $\varphi_1 = 1$, this implies the series is non-stationary, i.e. time dependent mean and variance. Stationarity of the time series follows from the alternative hypothesis H_1. The ADF regression is:

$$\begin{aligned} \Delta y_t &= \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p}\delta_i\Delta y_{t-i} + \varepsilon_t, \\ \text{ADF} &= \frac{\hat{\gamma}}{\text{SE}(\hat{\gamma})}. \end{aligned} \tag{2.28}$$

We test $H_0 : \gamma = 0$ against $H_1 : \gamma < 0$, where the ADF test statistic is the t-statistic for $\hat{\gamma}$.

- **Kolmogorov-Smirnov test (KS)**: Kolmogorov (1933) and Smirnov (1939) formulated the KS hypothesis test which compares two cumulative distribution. Under the null hypothesis the two samples are drawn from the same sample and the Kolmogorov-Smirnov statistic is defined as the maximal vertical distance between the distributions:

$$D_{n,m} = \sup_x |F_{1,n}(x) - F_{2,m}(x)| \tag{2.29}$$

where $F_{1,n}$ and $F_{2,m}$ are the empirical distributions with n and m samples. The null hypothesis can then be rejected at a level α if:

$$D_{n,m} > \sqrt{-\log(\frac{\alpha}{2})\frac{1 + m/n}{2m}} \tag{2.30}$$

The test is applied to the standard residuals obtained in the AR-GARCH model, and compared to the Skew-t distribution.

2.6.2 Copula model

To evaluate the fit of the conditional Markov regime-switching copula model, comparative criteria are used, including the Akaike information criterion (AIC), the Bayesian information criterion (BIC), and the likelihood-ratio (LR) test. The conditional model is compared to an unconditional Markov regime-switching specification. These measures assess in-sample fit while penalizing model complexity through the number of estimated parameters.

- **Akaike information criterion (AIC)**: As presented by Akaike (1974) the AIC is a comparative estimator of the predictive performance of a model. The selection rule is motivated by information theory, and selects the model expected to lose the least information from the actual distribution. Consider the data $\{x\}_{i=0}^{n}$ generated by the true distribution g, with the model $f(x|\theta)$ fitted by maximizing the likelihood $\theta^* = \arg\max_\theta \; L(\theta)$. The measure of discrepancy between the truth g and fitted model f is the Kullback-Leibler information loss, Kullback and Leibler (1951):

$$D_{KL}(g||f_\theta) = \int g(x) \log \frac{g(x)}{f(x|\theta)} dx \tag{2.31}$$

While not being possible to compute, with unknown g Akaike showed for large n minimizing the expected KL divergence of a fitted model, is equivalent to minimizing $2k - 2\log(L(\theta))$ where k is the number of parameters. Thus, the AIC estimate for a model is:

$$AIC = 2k - 2\log(L(\theta)) \tag{2.32}$$

Where the model with the lowest AIC therefore indicates the preferred model.

- **Bayesian information criterion (BIC):** Similar to AIC, the Bayesian information criterion (BIC) can be used to compare models, with smaller values indicating a preferred trade-off between fit and complexity. Presented in Schwarz (1978), BIC is derived as a large-sample (Laplace) approximation to $-2\log p(x \mid M)$, where $p(x \mid M)$ is the marginal likelihood under model M:

$$p(x \mid M) = \int p(x \mid \theta, M)\, \pi(\theta \mid M)\, d\theta, \tag{2.33}$$

where $\pi(\theta \mid M)$ denotes the prior under model M. Let $L_n(\theta) = p(x \mid \theta, M)$ and $\ell_n(\theta) = \log L_n(\theta)$, and let $\hat{\theta}$ denote the MLE that maximizes $\ell_n(\theta)$. A second-order Taylor expansion of $\ell_n(\theta)$ around $\hat{\theta}$ yields

$$\ell_n(\theta) \approx \ell_n(\hat{\theta}) - \frac{1}{2}(\theta - \hat{\theta})^\top J_n(\hat{\theta})(\theta - \hat{\theta}), \tag{2.34}$$

where

$$J_n(\hat{\theta}) = -\frac{\partial^2 \ell_n(\theta)}{\partial\theta\, \partial\theta^\top}\bigg|_{\theta=\hat{\theta}} \tag{2.35}$$

is the observed information matrix. Using a Laplace approximation and assuming $\pi(\theta \mid M)$ is smooth and positive in a neighborhood of $\hat{\theta}$,

$$\begin{aligned} p(x \mid M) &\approx \exp\Big(\ell_n(\hat{\theta})\Big)\, \pi(\hat{\theta} \mid M) \int \exp\Big(-\frac{1}{2}(\theta - \hat{\theta})^\top J_n(\hat{\theta})(\theta - \hat{\theta})\Big)\, d\theta \\ &= \exp\Big(\ell_n(\hat{\theta})\Big)\, \pi(\hat{\theta} \mid M)\, (2\pi)^{k/2}\, |J_n(\hat{\theta})|^{-1/2}, \end{aligned} \tag{2.36}$$

and therefore

$$\log p(x \mid M) \approx \ell_n(\hat{\theta}) + \log \pi(\hat{\theta} \mid M) + \frac{k}{2}\log(2\pi) - \frac{1}{2}\log|J_n(\hat{\theta})|. \tag{2.37}$$

Under standard regularity conditions, $\log|J_n(\hat{\theta})| = k\log n + O(1)$ as $n \to \infty$, so

$$-2\log p(x \mid M) \approx -2\ell_n(\hat{\theta}) + k\log n + O(1). \tag{2.38}$$

Dropping terms that are $O(1)$ in n, the BIC is

$$\mathrm{BIC} = -2\ell_n(\hat{\theta}) + k\log n = -2\log L_n(\hat{\theta}) + k\log n, \tag{2.39}$$

and the preferred model is the one with the smallest BIC. Compared to AIC, BIC applies a larger penalty for additional parameters through the $\log n$ term.

- **Likelihood ratio test**: The likelihood-ratio test is a hypothesis test comparing the goodness-of-fit of two models. Typically the comparison is made between a simpler model and a more complex model, where the null hypothesis states that the simple model is sufficient $H_0 : \theta \in \Theta_0, H_1 : \theta \in \Theta_1$.

$$\begin{aligned} \Lambda(x) &= \frac{\sup_{\theta\in\Theta_0} L(\theta)}{\sup_{\theta\in\Theta_1} L(\theta)} = \frac{L(\hat{\theta}_0)}{L(\hat{\theta}_1)} \\ LR &= -2\log\Lambda = 2(\log L_c - \log L_s) \end{aligned} \tag{2.40}$$

Where L_s is the likelihood of the simpler model, L_c the likelihood of the more complex model. For large sample sizes Wilks (1938) showed, by Wilks' theorem, that the LR statistic asymptotically approaches χ^2_d distribution if the null hypothesis is true, where the degrees of freedom d are the number of additional parameters in the complex model. The likelihood ratio test requires the models to be nested, i.e. the more complex model can be transformed to the simple by setting constraints on some parameters, often setting them to 0.

3 Methods

This section outlines the calibration procedure and explains how the estimated model is used to generate predictive return distributions and to evaluate its usefulness in an investor portfolio allocation setting.

3.1 Calibration

The model specified by Ye, Gao, Liu (2024) and described in the *Theory* is implemented and solved numerically. The AR-GARCH parameters are solved using the Python package **arch**, where the order is selected by the optimal BIC. The dynamic Markov regime-switching is, however solved recursively because the latent regime follows a Hidden Markov model and depends on the previous filtered state. We implement the quasi-maximum likelihood method and forward algorithm described by Kim and Nelson (1999).

1. Initialize the unconditional state probability

$$P(s_0 = 1) = 0.5 \tag{3.1}$$

2. Formulate the prediction probability

$$P(s_t = i, s_{t-1} = j|I_{t-1}; \Theta_c) = P(s_{t-1} = j|I_{t-1}; \Theta_c) * P(s_t = i|s_{t-1} = j, I_{t-1}; \Theta_c) \tag{3.2}$$

This is used to evaluate the dynamic Markov regime copula as:

$$E[c(u_t, v_t|I_{t-1}; \Theta_c)] = \sum_i \sum_j c(u_t, v_t|s_t = i, I_{t-1}; \Theta_c) * P(s_t = i, s_{t-1} = j|I_{t-1}; \Theta_c) \tag{3.3}$$

The conditional switching probability $P(s_t = i|s_{t-1} = j, I_{t-1}; \Theta_c)$ is simply the transition probabilities defined by the logistic function in eq. 2.15. The filtered probability $P(s_{t-1} = j|I_{t-1}; \Theta_c)$, i.e. the expected state probability given observed variables, is obtained via Bayes' rule using the copula density and calculated by:

$$\begin{gathered} P(s_{t-1} = j|I_{t-1}; \Theta_c) = \sum_{i=1}^{2} P(s_{t-1} = j, s_{t-2} = 1|I_{t-1}; \Theta_c) = \\ \sum_{i=1}^{2} \frac{c(u_{t-1}, v_{t-1}|s_{t-1} = j) * P(s_{t-2} = i|I_{t-2}, \Theta_c) * p_{i,j,t-1}}{E[c(u_t, v_t|I_{t-1}; \Theta_c)]} \end{gathered} \tag{3.4}$$

3. 1) and 2) are used to formulate the log likelihood of the copula as a function of the parameters Θ_c. Implemented as Algorithm 1, which is then maximized under Θ_c:

Algorithm 1 Copula log likelihood

$\alpha_0 = [0.5, 0.5]$ ▷ Initial state probability s_0
$\log L = 0$
for $t = 1 \to T$ **do**
  $\bar{c}_t = [c(F(x_t), G(y_t), \theta_1),\ c(F(x_t), G(y_t), \theta_2)]$
  $c_t = \alpha_t \circ \bar{c}_t$ ▷ Expected copula
  $logL \mathrel{+}= \ln(c_t)$
  $\alpha_t = \frac{\alpha_t \circ \bar{c}_t}{c_t}$ ▷ Filtered state probability s_t given I_t
  $p_{11} = \text{logistic}(B_t, \theta_{1,0}, \theta_{1,1})$
  $p_{22} = \text{logistic}(B_t, \theta_{2,0}, \theta_{2,1})$
  $P = \begin{bmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{bmatrix}$ ▷ Transition matrix
  $\alpha_{t+1} = P^T \alpha_t$ ▷ State probability s_{t+1} given I_t
end for
return $logL$

In the regime-switching model, the likelihood function exhibits an inherent symmetry due to the unobservable regimes. Specifically, the regimes labelled as $s = 1$ (high) and $s = 2$ (low) are interchangeable, equivalently, the model can be described by swapping these labels as $s = 1$ (low) and $s = 2$ (high).

Let us partition the parameters into two sets: $\Theta_1 = \{\theta_1, \theta_{0,1}, \theta_{1,1}\}$ and $\Theta_2 = \{\theta_2, \theta_{0,2}, \theta_{1,2}\}$. The liklihood function $L(\Theta_1, \Theta_2)$ then satisfies the symmetry property

$$L(\Theta_1, \Theta_2) = L(\Theta_2, \Theta_1) \tag{3.5}$$

This symmetry implies that if $(\hat{\Theta}_1, \hat{\Theta}_2)$ is a global maximizer, then so is $(\hat{\Theta}_2, \hat{\Theta}_1)$. Consequently, the model does not admit a unique solution.

Assuming that the trivial solution corresponding to a static copula—where $\hat{\Theta}_2 = \hat{\Theta}_1$ is not a global optimum then $\hat{\Theta}_2 \neq \hat{\Theta}_1$. Under this assumption, the set of global optima is non-convex, which further implies that the likelihood function L is non-convex.

To effectively address this non-convexity, the optimization is carried out using both a global constrained optimizer, namely **SHGO**, and an unconstrained gradient-based method, **BFGS**. By comparing solutions from these two distinct approaches, and with symmetric or identical results, we gain relative confidence that the identified solution corresponds to a global optimum.

3.2 Application and Asset Allocation

This section describes how the model is applied in practice and how an investor could use it for asset allocation. We also outline the backtesting scheme, which is subsequently used to quantify the economic value of the model.

To emphasize tail risk, we consider an investor with *constant relative risk aversion* (CRRA), following Patton (2006). Let S_0 denote initial wealth and let X_{t+1} and Y_{t+1} denote one-period returns on the two risky assets. For portfolio weights ω_x and ω_y, next-period gross portfolio wealth is

$$S_{t+1}(\omega) = S_0\left(1 + \omega_x X_{t+1} + \omega_y Y_{t+1}\right).$$

The CRRA utility function is then

$$U_\eta\Big(S_{t+1}(\omega)\Big) = \begin{cases} \frac{S_{t+1}(\omega)^{1-\eta}}{1-\eta}, & \eta \neq 1, \\ \log S_{t+1}(\omega), & \eta = 1, \end{cases} \tag{3.6}$$

where η is the coefficient of relative risk aversion. Since S_0 is a multiplicative constant, it does not affect the optimal weights; we therefore set $S_0 = 1$. Figure 3.1 illustrates how larger values of η increase the penalty on negative outcomes and reduce the marginal utility of large gains.

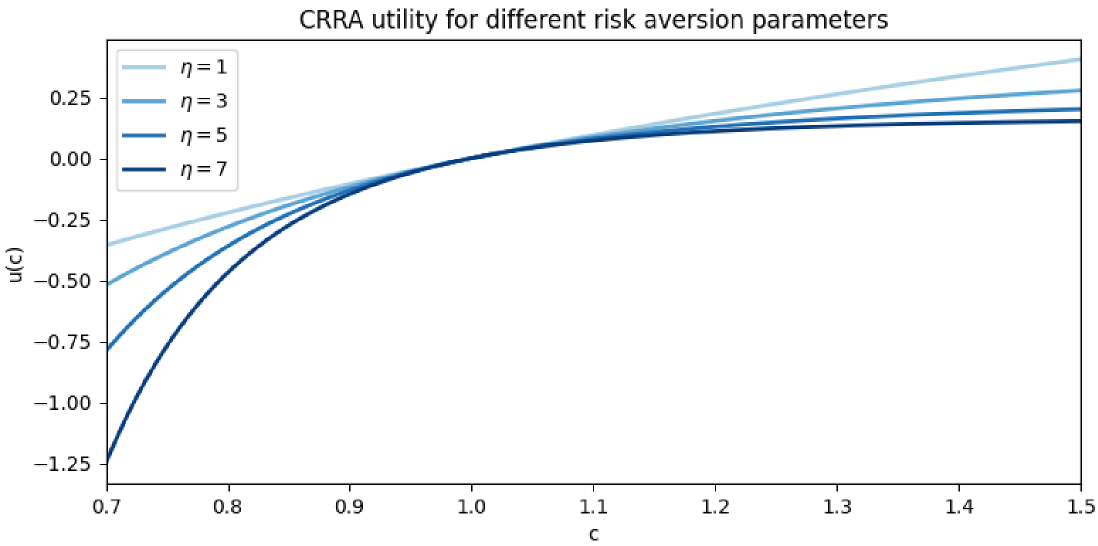


Figure 3.1: CRRA utility $U_\eta(c)$, where $c = S_0(1 + \omega_x X + \omega_y Y)$. The parameter η determines the level of risk aversion.

To obtain a one-step-ahead predictive return distribution and compute optimal weights, we use the following Monte Carlo scheme:

1. **Re-estimate the model.** Using data available up to time t, estimate the marginal AR-GARCH models and the regime-switching copula model, and store the resulting parameter estimates in $\hat{\Theta}_t$. Using the filtered state probabilities, compute the one-step-ahead predicted regime probabilities α_{t+1}.

2. **Simulate from the predictive copula.** For $i = 1, \ldots, N$ (with $N = 10{,}000$), draw a regime $s_i \in \{1, 2\}$ from α_{t+1} and then sample pseudo-observations from the corresponding copula:

$$(U_i, V_i) \mid s_i \sim C(\cdot, \cdot \mid \hat{\theta}_{s_i,t}), \qquad i = 1, \ldots, N. \tag{3.7}$$

3. **Transform to return space.** Using the fitted marginal models, map pseudo-observations to returns via the inverse CDF functions:

$$X_i = F^{-1}(U_i \mid \hat{\theta}_{x,t}), \qquad Y_i = G^{-1}(V_i \mid \hat{\theta}_{y,t}), \qquad i = 1, \ldots, N. \tag{3.8}$$

   This yields a Monte Carlo approximation of the predictive joint return distribution for (X_{t+1}, Y_{t+1}).

4. **Optimize expected utility.** Portfolio weights are chosen to maximize simulated expected utility:

$$\begin{aligned} \omega_t^* &= \arg\max_{\omega \in \Omega} \ \frac{1}{N} \sum_{i=1}^{N} U_\eta(1 + \omega_x X_i + \omega_y Y_i)\,, \\ \Omega &= \big\{(\omega_x, \omega_y) \in [0,1]^2 : \ \omega_x + \omega_y = 1\big\}. \end{aligned} \tag{3.9}$$

   The constraint set Ω implies no short positions and full investment in the two risky assets in each period.

5. **Backtesting.** Steps (1)-(4) are repeated sequentially over the out-of-sample period. At each re-estimation, the model is fitted using an expanding window of data up to time t. For numerical efficiency, the previous parameter estimates $\hat{\Theta}_{t-1}$ are used as starting values in the optimization at time t.

4

Results

This chapter presents the empirical results. Section 4.1 describes the bubble index constructed from the S&P 500 and relates it to broad market phases. Section 4.2 reports estimation results and diagnostics for the marginal AR-GARCH models. Section 4.3 presents the copula estimation results, comparing the conditional regime-switching specification with an unconditional regime-switching model and a static copula benchmark. Finally, Section 4.4 evaluates the models in an asset-allocation exercise based on out-of-sample portfolio performance.

4.1 Bubble index

Figure 4.1 displays the bubble index constructed from the S&P 500 over the full sample (1989-09-11 to 2026-01-16). High (low) values of the bubble index tend to coincide with expanding (contracting) market conditions, consistent with its interpretation as a proxy for deviations from a long-run benchmark and, by extension, speculative excess. Notably, elevated values are observed around well-known episodes of market exuberance, while sharp declines coincide with major drawdowns of the dot-com crash in early 2000s, great financial crisis in 2008, and the COVID-19 pandemic in 2020.

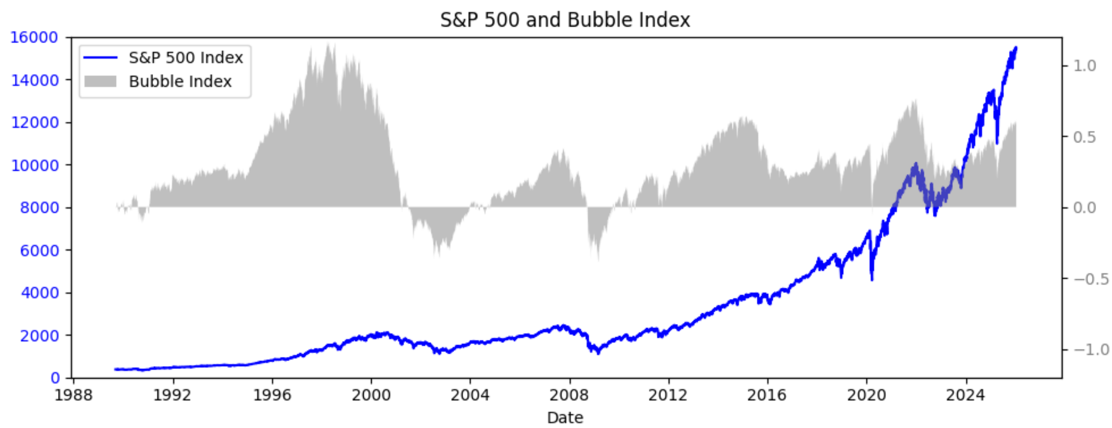


Figure 4.1: S&P 500 index level (blue, left axis) and corresponding bubble index (grey, right axis), evaluated daily from 1989-09-11 to 2026-01-16.

4.2 Marginal distributions

To estimate marginal return dynamics for each sector index, we fit AR-GARCH models with skew-t innovations as specified in (2.3). This specification captures time-varying conditional volatility and allows for heavy tails and skewness in the standardized innovations.

Table 4.1 reports descriptive statistics and diagnostic test results for the daily log returns. Mean daily returns range from 0.031% to 0.051%, while standard deviations range from 0.95% to 1.68%. Information Technology has the highest average return and among the highest volatility, whereas Consumer Non-Cyclical exhibits the lowest volatility (0.95%), consistent with common sector-level risk characteristics. The return distributions are generally negatively skewed and strongly leptokurtic (kurtosis > 3), indicating heavy tails. The Jarque-Bera test rejects normality for all sector indices.

Serial dependence is assessed using the Ljung-Box test at lag 10. For most sectors the p-values are below 5%, implying rejection of the null of no autocorrelation in returns. (Materials is a partial exception, with a p-value of 0.0616 at lag 10.) The ARCH LM test rejects the null of no ARCH effects for all sectors, consistent with pronounced conditional heteroskedasticity and motivating a GARCH-type volatility model. Finally, the ADF test rejects the unit-root null for all indices, providing no evidence of non-stationarity in the return series.

Table 4.1: Descriptive statistics and diagnostic tests for sector indices, applied to daily log returns from 1989-09-12 to 2026-01-16.

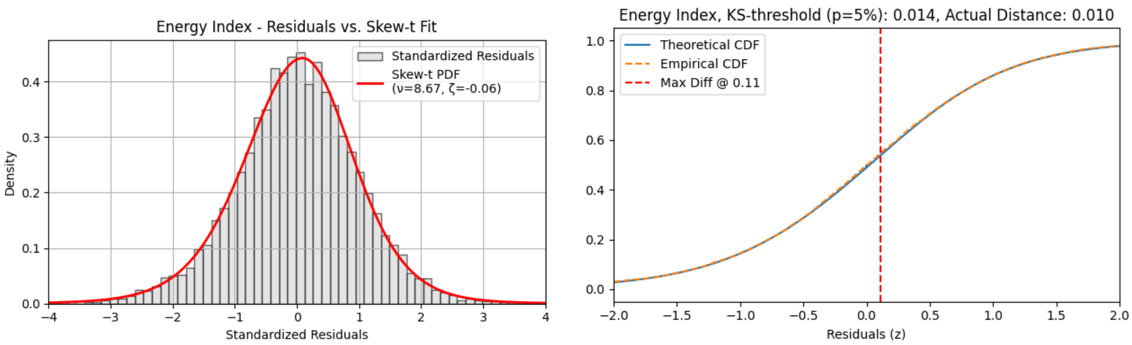
Sector	Mean (%)	Std Dev (%)	Skew	Kurt	JB (p)	LB(10) stat (p)	ARCH (p)	ADF t (p)
S&P 500	0.041%	1.14%	-0.38	10.92	45,637 (0e-03)	57.65 (0e-03)	800 (0e-03)	-17.53 (0e-03)
Industrials	0.040%	1.25%	-0.40	8.46	27,491 (0e-03)	30.88 (0e-03)	665 (0e-03)	-17.45 (0e-03)
Energy	0.036%	1.61%	-0.56	14.47	80,276 (0e-03)	40.11 (0e-03)	444 (0e-03)	-15.63 (0e-03)
Materials	0.031%	1.38%	-0.29	7.91	23,981 (0e-03)	16.26 (0.0616)	603 (0e-03)	-22.01 (0e-03)
Healthcare	0.043%	1.15%	-0.20	6.43	15,808 (0e-03)	52.59 (0e-03)	764 (0e-03)	-17.30 (0e-03)
Financials	0.035%	1.68%	-0.18	17.19	112,638 (0e-03)	101.83 (0e-03)	976 (0e-03)	-16.68 (0e-03)
Information Technology	0.051%	1.67%	0.03	6.03	13,868 (0e-03)	54.89 (0e-03)	643 (0e-03)	-23.27 (0e-03)
Consumer, non-cyclical	0.040%	0.95%	-0.23	9.62	35,372 (0e-03)	31.56 (0e-03)	1,053 (0e-03)	-21.69 (0e-03)

Table 4.2 reports the estimated AR-GARCH and skew-t parameters. For all indices except Consumer Non-Cyclical, the preferred specification under BIC is AR(0)-GARCH(1,1), indicating that the conditional mean is adequately described by a constant while volatility is time-varying. For Consumer Non-Cyclical, BIC selects AR(2)-GARCH(1,1), suggesting mild additional predictability in the conditional mean.

As a distributional diagnostic, the fitted skew-t distribution is compared to the empirical distribution of standardized residuals using the Kolmogorov-Smirnov (KS) statistic. Since the skew-t parameters are estimated on the same sample used to compute the KS statistic, the resulting p-values should be interpreted cautiously. For Industrials, Healthcare, Financials, and Information Technology, the KS p-values fall below 5%, suggesting remaining discrepancies between the fitted skew-t and the empirical residual distribution. In contrast, Energy, Materials, and Consumer Non-Cyclical exhibit larger KS p-values, indicating closer agreement with the fitted innovation distribution.

Table 4.2: Parameter estimations for the marignal distributions of the sector indices. Calibrated on daily log-returns from 1989-09-12 to 16-01-2026.

Sector	φ_0	φ_1	φ_2	α_0	α_1	β_1	ζ	υ	**KS test (p)**
Industrials	6.250 e-04			1.556 e-06	0.0859	0.9042	7.5256	-0.0860	0.0195 (0.0018)
Energy	4.783 e-04			1.252 e-06	0.0641	0.9317	8.6899	-0.0574	0.0098 (0.3465)
Materials	4.936 e-04			1.967 e-06	0.0842	0.9057	8.3577	-0.0663	0.0120 (0.1394)
Healthcare	5.943 e-04			1.938 e-06	0.0876	0.8986	7.1815	-0.0628	0.0163 (0.0156)
Financials	6.689 e-04			2.149 e-06	0.0952	0.8967	6.7231	-0.0470	0.0165 (0.0135)
Information Technology	8.729 e-04			1.806 e-06	0.0868	0.9093	7.9823	-0.0964	0.0178 (0.0060)
Consumer N.C.	5.390 e-04	-0.0307	-0.0272	1.455 e-06	0.0834	0.8994	7.1922	-0.0578	0.0125 (0.1129)



Standardized residuals and calibrated Skew-t distribution

Empirical and fitted cumulative density function

Figure 4.2: Standardised residuals from the fitted AR(0)-GARCH(1,1) model compared to the fitted skew-t distribution. Both the pdf and cdf indicate that the model captures the main distributional features of the data

4.3 Dynamic copula

We estimate time-varying dependence between sector indices using the conditional Markov regime-switching copula model (TV-MRSB) described in Section 2. The conditional specification is evaluated against an unconditional Markov regime-switching copula (TV-MRS) and a static copula benchmark. Model comparison is conducted using AIC, BIC, and likelihood-ratio (LR) tests.

Across several sector pairs, the conditional specification provides a materially better in-sample fit than the unconditional TV-MRS model. In particular, the pairs Industrials-Information Technology, Industrials-Materials, Financials-Information Technology, Energy-Materials, and Materials-Information Technology exhibit the largest

LR statistics and lower AIC/BIC under TV-MRSB, indicating that conditioning transition probabilities on the bubble index improves explanatory power for these dependence dynamics.

The dependence dynamics is shown in table 4.3 by the calibrated θ and λ_l -parameters. There is a notable difference between λ_l^h and λ_l^l, implied by the regime-specific Clayton parameters. A consistent pattern is that the high-dependence regime features substantially larger lower-tail dependence than the low-dependence regime, supporting the interpretation of the latent states as “stress” and “normal” dependence regimes. For example looking at the TV-MRSB Copula for Information Technology - Consumer, non-cyclical, the low dependence regime has very little dependence between the indices with only $\lambda_l^l = 0.04$ while significant co-movements in the higher regime with $\lambda_l^h = 0.58$. The highest dependence in both high and low regimes we see between Industrials - Materials ($\lambda_l^l = 0.54, \lambda_l^h = 0.81$) and Industrials - Financials ($\lambda_l^l = 0.59, \lambda_l^h = 0.84$).

The strong dependence between Industrials - Materials is consistent with a shared exposure to the global economic cycle, and coupling of their value-chains. Materials companies supply key materials (metals, chemicals, construction materials) that feed into industrial production, so both sectors respond to common drivers such as global growth, capex cycle, and commodity prices. These common shocks tend to be the main subject of sector-specific news implying high co-movements in both regimes.

The high dependence between Industrials - Financials is more naturally interpreted through macro-economics and the credit-cycle. Both sectors are strongly pro-cyclical and sensitive to tightening financial conditions: higher risk-free rates, and wider credit spreads raise the cost of capital constraining industrial investments, while financials are affected by higher expected credit losses, deteriorating asset quality, and lower risk appetite.

The transition parameters $\theta_{0,i}$ are positive on a range [2.10, 5.58] indicating a persistence of regimes, i.e. the probability is higher to remain in the current regime than switching. The transition parameters $\theta_{1,i}$ links the Bubble Index to the transition matrix and obtains both positive and negative values, which implies a mixed dependence on the variable. The impact of the Bubble Index (B) to the transition probabilities can be described by the partial derivative:

$$\frac{\partial p_{ii}}{\partial B} = p_{ii}(B)(1 - p_{ii}(B))\theta_{1,i} \tag{4.1}$$

We see a negative $\theta_{1,i}$ implies a higher switching probability with increasing B. Being in the low dependence state this can be congruent with the interpretation: an increasing bubble index indicates the creation of a bubble, which precedes a downturn and high dependency regime, thus motivating increased probability of switching regime. However, it is mostly the high regime parameter $\theta_{1,1}$ that is negative, while the low regime parameter $\theta_{1,2}$ is positive. This indicates a higher Bubble index increases the probability of switching from high dependency regime to low, while decreasing the probability of switching from low to high dependency regime. This challenges the result of introducing the bubble index in the model as defined by Ye, Gao and Liu (2024), that *with a rise in the bubble index, markets exhibit stronger tendency to move to the high correlation regime* (Note: Ye, Gao, Liu (2024) constructs the Bubble index with MSCI World index while this study base it of the S&P 500). The interpretation is rather that periods of market expansion, and a high Bubble index, are persistent and coincide with a low dependency regime. A low Bubble index coincides with turbulent times and a high dependency regime, which can be observed in figure 4.1.

Table 4.3: Parameter estimations for time-varying dynamic Markov switching copula and unconditional Markov switching copula. Calibrated on daily log-returns from 1989-09-12 to 16-01-2026.

	TV-MRSB Copula								TV-MRS-Copula						Statistics
Sector	λ_l^h	λ_l^l	$\theta_{0,1}$	$\theta_{1,1}$	$\theta_{0,2}$	$\theta_{1,2}$	**AIC**	**BIC**	λ_l^h	λ_l^l	$\theta_{0,1}$	$\theta_{0,2}$	**AIC**	**BIC**	**LR (P-value)**
Industrials - Energy	0.71	0.29	4.67	-2.73	4.97	-1.59	53,337	53,379	0.71	0.30	4.05	4.69	53,342	53,371	9.59 (0.008)
Industrials - Materials	0.81	0.54	3.68	-1.68	2.79	0.65	46,380	46,423	0.80	0.54	3.54	3.32	46,405	46,433	28.43 (6.7e-07)
Industrials - Healthcare	0.73	0.23	3.43	-1.12	2.86	-0.55	46,946	46,989	0.71	0.21	3.40	2.87	46,946	46,975	4.28 (0.118)
Industrials - Financials	0.84	0.59	2.10	0.05	1.99	1.46	47,079	47,121	0.84	0.59	2.17	2.47	47,091	47,120	16.72 (2.3e-04)
Industrials - Information Technology	0.76	0.39	3.36	0.08	2.61	2.67	52,240	52,282	0.75	0.38	3.78	3.73	52,265	52,293	29.24 (4.5e-07)
Industrials - Consumer, n.c.	0.71	0.32	3.94	1.10	3.44	2.77	43,241	43,283	0.72	0.32	4.11	4.06	43,242	43,271	5.77 (0.056)
Energy - Materials	0.76	0.36	4.04	-3.19	4.26	-1.26	54,980	55,023	0.74	0.36	3.69	4.39	54,995	55,024	19.09 (7.2e-05)
Energy - Healthcare	0.64	0.11	4.09	-1.75	4.41	-0.47	54,318	54,361	0.64	0.12	3.77	4.50	54,324	54,352	9.75 (0.008)
Energy - Financials	0.65	0.20	4.32	0.67	4.16	2.68	57,045	57,087	0.64	0.20	4.71	5.06	57,046	57,075	5.55 (0.062)
Energy - Information Technology	0.67	0.10	4.83	-3.10	5.22	-0.88	60,495	60,538	0.66	0.10	4.12	5.12	60,505	60,534	14.14 (8.5e-04)
Energy - Consumer, n.c.	0.66	0.13	3.86	-0.16	4.24	1.69	50,615	50,658	0.65	0.12	3.83	4.67	50,616	50,645	5.20 (0.074)
Materials - Healthcare	0.69	0.23	3.79	-1.39	3.60	-0.01	50,520	50,563	0.69	0.24	3.59	3.82	50,526	50,555	9.92 (0.007)
Materials - Financials	0.75	0.40	3.20	-0.76	2.66	0.94	52,183	52,225	0.74	0.40	3.26	3.21	52,195	52,224	16.68 (2.4e-04)
Materials - Information Technology	0.71	0.27	3.74	-1.20	3.23	1.51	56,191	56,234	0.70	0.27	3.76	3.99	56,210	56,239	23.37 (8.4e-06)
Materials - Consumer, n.c.	0.72	0.31	3.31	-0.85	3.58	0.73	46,849	46,891	0.71	0.31	3.31	4.03	46,855	46,884	10.30 (0.006)
Healthcare - Financials	0.66	0.16	3.83	1.48	2.73	2.92	50,553	50,595	0.67	0.17	3.91	3.28	50,554	50,583	5.32 (0.070)
Healthcare - Information Technology	0.67	0.09	3.65	-0.35	3.07	0.62	53,935	53,978	0.67	0.09	3.64	3.34	53,935	53,964	4.16 (0.125)
Healthcare - Consumer, n.c.	0.74	0.37	4.01	-1.47	3.67	-0.28	42,498	42,541	0.74	0.38	3.69	3.75	42,504	42,532	9.73 (0.008)
Financials - Information Technology	0.68	0.22	3.92	1.17	2.60	4.18	56,576	56,619	0.67	0.23	4.85	4.46	56,597	56,625	25.16 (3.4e-06)
Financials - Consumer, n.c.	0.66	0.26	4.18	1.88	3.42	3.68	47,041	47,084	0.66	0.26	4.54	4.24	47,050	47,079	13.31 (0.001)
Information Technology - Consumer, n.c.	0.58	0.04	5.58	-1.45	4.68	0.43	50,853	50,896	0.58	0.05	5.25	4.94	50,854	50,882	4.80 (0.091)

4.4 Application and Asset Allocation

The proposed regime-switching copula model has practical relevance for risk management and asset allocation because it provides a forward-looking approximation to the joint return distribution with time-varying dependence. In this section, the model is applied in a weekly rebalancing backtest, where one-step-ahead predictive return distributions are generated via Monte Carlo simulation and portfolio weights are chosen by maximizing CRRA expected utility. The allocation exercise is conducted on weekly returns over the out-of-sample period 2001-03-16 to 2026-01-16 (see Tables 4.4-4.5).

To reduce model complexity in the allocation exercise, we impose an AR(0)-GARCH(1,1) specification for the marginal return dynamics, which is the BIC-selected model for all but one sector in the full-sample marginal analysis. This restriction provides robust marginal filter while keeping the dependence model as the primary object of interest.

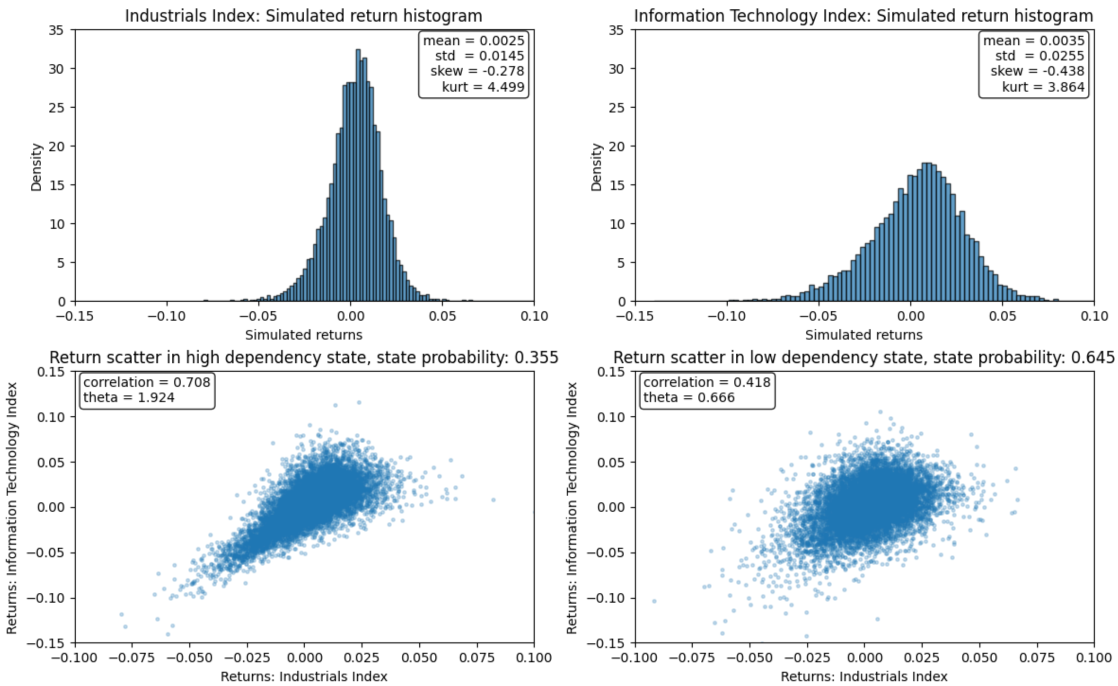


Figure 4.3: Monte Carlo simulation Industrials - Information Technology indices: weekly return distribution from 16-01-2026 to 23-01-2026. With state probabilities $P(s_1) = 0.355$ (high), $P(s_2) = 0.645$ (low) and corresponding copula parameters $\theta_h = 1.924$, $\theta_l = 0.666$

Figure 4.3 illustrates the simulated one-week-ahead return distribution for Industrials and Information Technology over 2026-01-16 to 2026-01-23, using information up to 2026-01-16. The marginal histograms exhibit skewness and excess kurtosis, and Information Technology displays a higher variance and mean than Industrials. The dependence implied by the copula simulation is state-dependent: the predicted regime probabilities place more weight on the low-dependence regime ($P(s_{t+1} = 2) = 0.645$) than the high-dependence regime ($P(s_{t+1} = 1) = 0.355$), implying weaker joint downside clustering than would be expected under the high-dependence regime, while still allowing for non-linear dependence.

Similar to Ye, Gao, Liu (2024), we evaluate whether the predictive distribution from the dependence model can generate economic gain in portfolio allocation. Portfolio weights are obtained by maximizing the expected CRRA utility with risk-aversion parameter $\eta = 3$, using 10,000 Monte Carlo draws each period. For comparability across strategies, Tables 4.4 and 4.5 report the mean weekly return, standard deviation, Sharpe ratio, and maximum drawdown for each portfolio. While the optimization criterion is CRRA utility, the Sharpe ratio provides a standard risk-adjusted performance summary commonly used for benchmarking.

Table 4.4 compares fully invested portfolios ($\omega_x + \omega_y = 1$). The TV-MRSB Copula model attains the highest Sharpe ratio in 5 out of 21 sector pairs, while TV-MRS Copula is best in 4 pairs. However, the equally weighted benchmark achieves the highest Sharpe ratio in 10 pairs, and the static copula strategy is best in several additional cases.

Table 4.5 considers an expanded asset set where the investor may allocate to the risk-free asset of US 3-month T-bills, i.e. $\omega_x + \omega_y \leq 1$. Under this specification, TV-MRSB Copula model attains the highest Sharpe ratio in 2 pairs, while the equally weighted benchmark for the two indices is best in 13 pairs.

Overall, the results provide limited evidence that conditioning regime transitions on the bubble index yields a consistent improvement in mean-variance risk-adjusted performance relative to simple benchmarks.

Importantly, portfolio weights are obtained by maximizing CRRA utility rather than directly maximizing the Sharpe ratio. CRRA preferences place relatively greater weight on adverse outcomes, which can lead to allocations that sacrifice some expected return in exchange for reduced downside exposure. Consistent with this, the model-based strategies frequently exhibit smaller maximum drawdowns than the corresponding fully invested benchmark portfolios, indicating some improved downside-risk control.

Figure 4.4 provides examples of the time series of log total wealth and portfolio weights under TV-MRSB, illustrating how allocations shift toward the less risky asset and the risk-free asset, during periods of market stress. For Energy - Information technology the model yields slightly lower returns than the equally weighted portfolio, while with lower volatility from having an average 11.9% allocation to the risk free asset. For Financials - Consumer Non Cyclical the TV-MRSB-Copula yields a similar return to the equally weighted portfolio. This while having lower standard deviation of 1.92%, compared to 2.50% of the equally weighted portfolio. Thus the model has the same return, while at lower risk than equally weighted portfolio. We note the outperformance in particular stems from a decreased exposure to financials, and increased allocation to the risk free asset during the 2008 financial crisis.

A key limitation of the allocation exercise is that it excludes transaction costs. Since the optimal weights can vary substantially from week to week, turnover may be high, and realistic rebalancing costs could materially reduce performance.

Table 4.4: Economic value of alternative models. (Fully invested $w \in \Omega = \{(w_x, w_y) : w_x + w_y = 1, w_x, w_y \geq 0\}$)

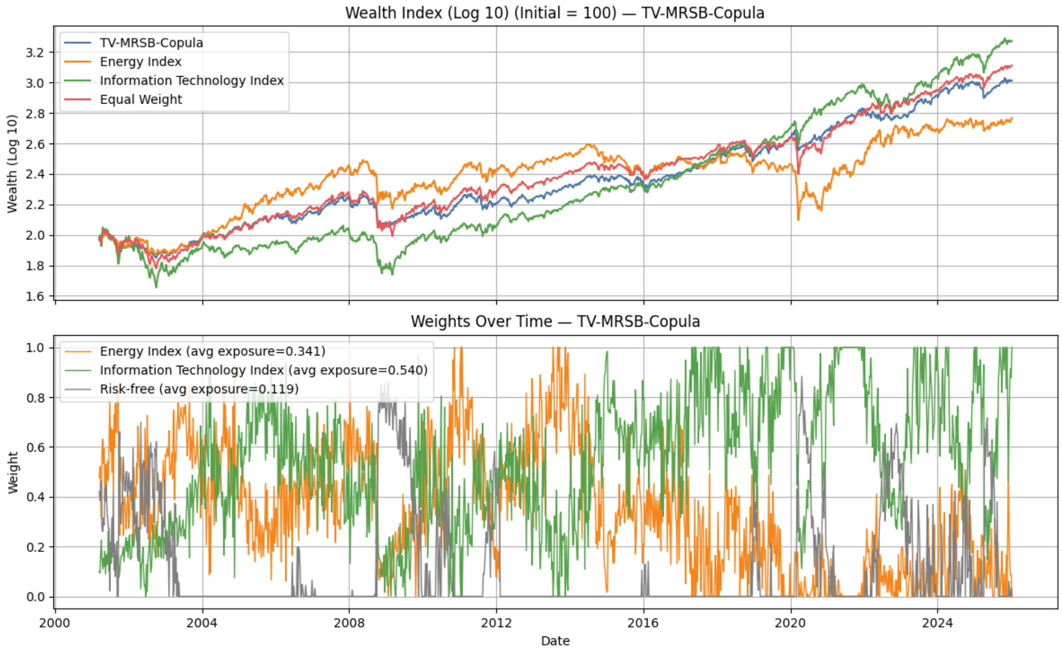
	TV-MRSB-Copula				TV-MRS-Copula				Constant Copula				Equally Weighted			
	Return	s.d.	SR	MDD	Return	s.d.	SR	MDD	Return	s.d.	SR	MDD	Return	s.d.	SR	MDD
Industrials - Energy	0.1749	2.8433	0.0499	−62.92	0.1894	2.8417	0.0550	−61.42	0.1521	2.8359	0.0420	−61.06	0.2036	2.9714	**0.0574**	−56.03
Industrials - Materials	0.2012	2.8305	0.0594	−62.65	0.2031	2.8353	**0.0600**	−62.70	0.1964	2.8413	0.0575	−63.36	0.2046	2.8754	0.0596	−60.44
Industrials - Healthcare	0.1575	2.2995	0.0541	−43.86	0.1585	2.3053	0.0544	−42.91	0.1516	2.2943	0.0516	−45.47	0.1916	2.3772	**0.0667**	−51.40
Industrials - Financials	0.1927	2.9014	0.0550	−65.09	0.1994	2.9085	**0.0572**	−64.42	0.1874	2.9044	0.0531	−66.05	0.1908	3.1180	0.0506	−73.66
Industrials - Information Technology	0.2608	2.7483	**0.0828**	−56.94	0.2605	2.7477	**0.0828**	−54.58	0.2575	2.7525	0.0815	−55.78	0.2416	2.8475	0.0732	−57.26
Industrials - Consumer, n.c.	0.1632	1.9790	0.0657	−40.26	0.1552	1.9623	0.0622	−38.95	0.1546	1.9736	0.0616	−40.68	0.1883	2.1760	**0.0713**	−48.22
Energy - Materials	0.1668	3.0843	0.0433	−60.64	0.1825	3.0866	0.0484	−60.05	0.1737	3.0761	0.0457	−61.00	0.2043	3.1068	**0.0551**	−56.19
Energy - Healthcare	0.1787	2.3358	**0.0623**	−38.22	0.1649	2.3280	0.0566	−40.09	0.1641	2.3285	0.0562	−40.12	0.1914	2.5922	0.0611	−43.61
Energy - Financials	0.1604	3.0101	0.0423	−58.01	0.1676	3.0465	0.0442	−60.68	0.1690	3.0226	0.0449	−60.79	0.1906	3.2522	**0.0484**	−67.41
Energy - Information Technology	0.2655	2.8536	**0.0814**	−50.27	0.2508	2.8635	0.0760	−52.18	0.2451	2.8540	0.0743	−51.80	0.2413	2.9286	0.0711	−49.51
Energy - Consumer, n.c.	0.1490	1.9456	0.0596	−37.04	0.1494	1.9805	0.0587	−39.16	0.1592	1.9454	**0.0648**	−36.00	0.1881	2.3988	0.0646	−41.18
Materials - Healthcare	0.1711	2.3191	0.0595	−41.59	0.1760	2.3258	0.0614	−41.62	0.1650	2.3153	0.0569	−42.32	0.1924	2.4548	**0.0649**	−47.77
Materials - Financials	0.1972	3.0313	0.0541	−66.23	0.1993	3.0338	0.0548	−66.65	0.2055	3.0535	**0.0564**	−65.06	0.1915	3.1549	0.0502	−71.40
Materials - Information Technology	0.2558	2.8657	0.0777	−54.71	0.2665	2.8916	**0.0807**	−54.69	0.2594	2.8956	0.0781	−54.99	0.2423	2.9034	0.0721	−54.39
Materials - Consumer, n.c.	0.1547	1.9170	0.0634	−35.31	0.1562	1.9228	0.0640	−36.22	0.1607	1.9272	0.0662	−35.54	0.1891	2.2521	**0.0693**	−46.54
Healthcare - Financials	0.1614	2.3265	**0.0551**	−42.60	0.1597	2.3179	0.0546	−45.26	0.1535	2.3180	0.0519	−40.97	0.1786	2.7068	0.0537	−65.93
Healthcare - Information Technology	0.1816	2.3120	0.0642	−42.41	0.1958	2.3035	0.0706	−42.09	0.2005	2.3169	0.0722	−43.52	0.2294	2.4618	**0.0797**	−44.35
Healthcare - Consumer, n.c.	0.1578	2.0860	0.0597	−36.36	0.1645	2.0744	0.0633	−34.20	0.1541	2.0789	0.0582	−35.43	0.1761	1.9406	**0.0737**	−34.55
Financials - Information Technology	0.2526	2.7643	0.0794	−56.43	0.2521	2.7919	0.0784	−54.46	0.2627	2.7894	**0.0823**	−57.01	0.2286	3.1048	0.0630	−69.19
Financials - Consumer, n.c.	0.1694	2.0586	**0.0662**	−35.63	0.1633	2.0577	0.0633	−35.13	0.1681	2.0526	0.0658	−34.30	0.1753	2.4997	0.0569	−62.73
Information Technology - Consumer, n.c.	0.1962	2.0438	0.0798	−36.30	0.1961	2.0224	0.0806	−38.36	0.1968	2.0386	0.0803	−36.48	0.2261	2.2543	**0.0856**	−41.10

Notes: Columns report weekly portfolio returns, standard deviation (s.d.), Sharpe ratio (SR), and maximum drawdown (MDD, %) for each model. Boldface indicates the highest SR within each row. Evaluated over 2001-03-16 - 2026-01-16 and based on out-of-sample forecast maximizing the CRRA utility with risk-aversion parameter $\eta = 3$.

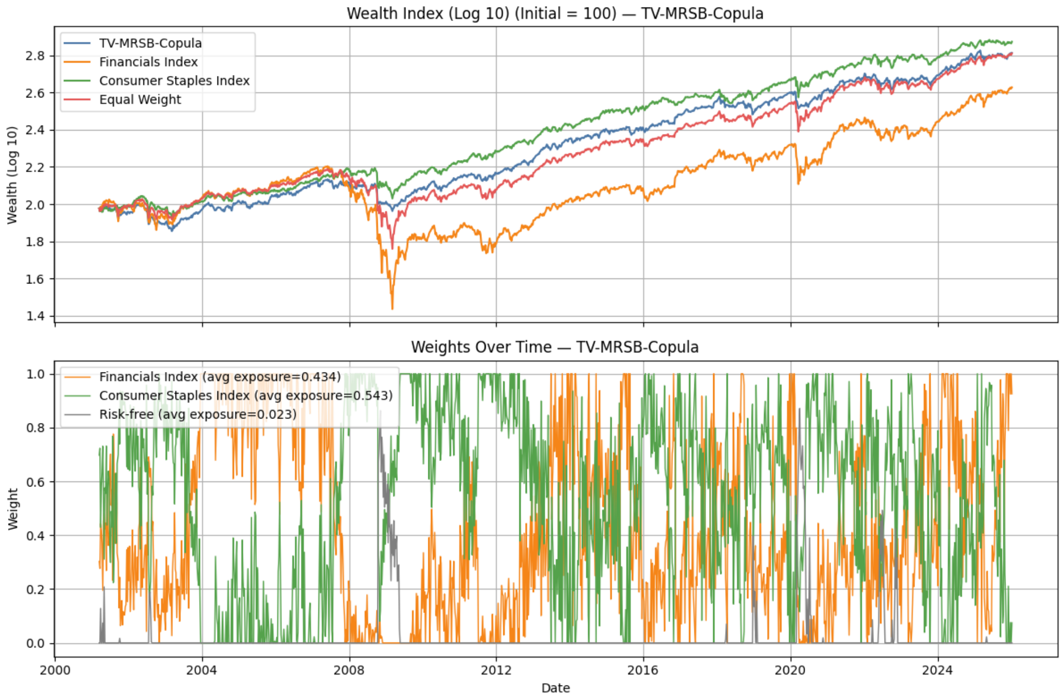
Table 4.5: Economic value of alternative models. (Including investment option in risk free asset $w \in \Omega = \{(w_x, w_y) : w_x + w_y \leq 1, w_x, w_y \geq 0\}$)

	TV-MRSB-Copula				TV-MRS-Copula				Constant Copula				Equally Weighted			
	Return	s.d.	SR	MDD	Return	s.d.	SR	MDD	Return	s.d.	SR	MDD	Return	s.d.	SR	MDD
Industrials - Energy	0.1222	2.1740	0.0410	−43.23	0.1351	2.1795	0.0468	−41.90	0.1521	2.8359	0.0420	−61.06	0.2036	2.9714	**0.0574**	−56.03
Industrials - Materials	0.1362	2.0938	0.0492	−40.30	0.1453	2.0931	0.0536	−39.00	0.1964	2.8413	0.0575	−63.36	0.2046	2.8754	**0.0596**	−60.44
Industrials - Healthcare	0.1450	2.1021	0.0532	−35.28	0.1405	2.1036	0.0510	−35.22	0.1516	2.2943	0.0516	−45.47	0.1916	2.3772	**0.0667**	−51.40
Industrials - Financials	0.1431	2.2510	0.0489	−43.43	0.1520	2.2423	0.0530	−39.94	0.1874	2.9044	**0.0531**	−66.05	0.1908	3.1180	0.0506	−73.66
Industrials - Information Technology	0.1876	2.2104	0.0699	−40.58	0.1886	2.2018	0.0706	−36.74	0.2575	2.7525	**0.0815**	−55.78	0.2416	2.8475	0.0732	−57.26
Industrials - Consumer, n.c.	0.1592	1.8225	0.0692	−34.80	0.1529	1.8089	0.0662	−33.03	0.1546	1.9736	0.0616	−40.68	0.1883	2.1760	**0.0713**	−48.22
Energy - Materials	0.1126	2.2547	0.0353	−38.92	0.1361	2.3209	0.0444	−42.58	0.1737	3.0761	0.0457	−61.00	0.2043	3.1068	**0.0551**	−56.19
Energy - Healthcare	0.1553	2.1566	0.0566	−37.76	0.1444	2.1466	0.0518	−37.59	0.1641	2.3285	0.0562	−40.12	0.1914	2.5922	**0.0611**	−43.61
Energy - Financials	0.1235	2.3032	0.0392	−40.78	0.1294	2.3360	0.0412	−40.10	0.1690	3.0226	0.0449	−60.79	0.1906	3.2522	**0.0484**	−67.41
Energy - Information Technology	0.2073	2.3240	**0.0749**	−36.46	0.1889	2.3450	0.0664	−40.94	0.2451	2.8540	0.0743	−51.80	0.2413	2.9286	0.0711	−49.51
Energy - Consumer, n.c.	0.1377	1.7790	0.0588	−34.54	0.1415	1.8061	0.0600	−35.28	0.1592	1.9454	**0.0648**	−36.00	0.1881	2.3988	0.0646	−41.18
Materials - Healthcare	0.1478	2.1206	0.0541	−36.50	0.1555	2.1121	0.0579	−35.49	0.1650	2.3153	0.0569	−42.32	0.1924	2.4548	**0.0649**	−47.77
Materials - Financials	0.1381	2.2484	0.0467	−43.56	0.1503	2.2561	0.0519	−35.90	0.2055	3.0535	**0.0564**	−65.06	0.1915	3.1549	0.0502	−71.40
Materials - Information Technology	0.1861	2.2354	0.0684	−42.69	0.2045	2.2352	0.0767	−35.38	0.2594	2.8956	**0.0781**	−54.99	0.2423	2.9034	0.0721	−54.39
Materials - Consumer, n.c.	0.1521	1.7469	0.0681	−29.09	0.1466	1.7534	0.0647	−30.97	0.1607	1.9272	0.0662	−35.54	0.1891	2.2521	**0.0693**	−46.54
Healthcare - Financials	0.1344	2.1320	0.0475	−39.66	0.1434	2.1205	0.0520	−39.93	0.1535	2.3180	0.0519	−40.97	0.1786	2.7068	**0.0537**	−65.93
Healthcare - Information Technology	0.1589	2.1449	0.0587	−37.35	0.1732	2.1267	0.0658	−38.18	0.2005	2.3169	0.0722	−43.52	0.2294	2.4618	**0.0797**	−44.35
Healthcare - Consumer, n.c.	0.1481	1.9712	0.0583	−33.24	0.1544	1.9638	0.0617	−33.26	0.1541	2.0789	0.0582	−35.43	0.1761	1.9406	**0.0737**	−34.55
Financials - Information Technology	0.1815	2.2780	0.0651	−43.09	0.1845	2.2945	0.0660	−38.28	0.2627	2.7894	**0.0823**	−57.01	0.2286	3.1048	0.0630	−69.19
Financials - Consumer, n.c.	0.1632	1.9252	**0.0675**	−32.50	0.1556	1.9280	0.0635	−32.56	0.1681	2.0526	0.0658	−34.30	0.1753	2.4997	0.0569	−62.73
Information Technology - Consumer, n.c.	0.1868	1.9055	0.0806	−33.96	0.1846	1.8896	0.0801	−33.11	0.1968	2.0386	0.0803	−36.48	0.2261	2.2543	**0.0856**	−41.10

Notes: Columns report weekly portfolio returns, standard deviation (s.d.), Sharpe ratio (SR), and maximum drawdown (MDD, %) for each model. Boldface indicates the highest SR within each row. Evaluated over 2001-03-16 - 2026-01-16 and based on out-of-sample forecast maximizing the CRRA utility with risk-aversion parameter $\eta = 3$.



(a) Energy - Information Technology



(b) Financials - Consumer n.c.

Figure 4.4: Example evaluations.
Note: (Above) Total wealth (Log 10) for each respective index, equally weighted portfolio and TV-MRSB Copula model, and (below) portfolio weights for TV-MRSB Copula. The evaluation period is from 2001-03-16 to 2026-01-16.

5

Conclusion

This study evaluates the Time-Varying Markov Regime-Switching Copula model, with bubble-index-conditioned transitions (TV-MRSB), proposed in Ye, Gao and Liu (2024) using sector sub-indices of the S&P500. The objective is to assess whether a market-sentiment proxy, the bubble index, helps explain and predict time variation in cross-asset dependence, particularly joint downside risk. Marginal return dynamics are modelled using AR-GARCH specifications with skew-t innovations, while dependence is modelled via a Clayton copula with two regime-specific parameters. The regime states follow a Hidden Markov model with transition probabilities conditioned on the bubble index, yielding interpretable high- and low-dependence regimes.

Based on likelihood ratio tests as well as AIC and BIC, the conditional TV-MRSB Copula model improves in-sample fit relative to the corresponding unconditional regime-switching model (TV-MRS) for 5 sector pairs, to a high degree of certainty. However, compared to the results obtained in Ye, Gao and Liu (2024) some sector model parameters seem to indicate a high bubble index favours a low dependence regime by the transition probabilities. The direction of the impact from the Bubble index on the transition probabilities is not fully consistent as the parameters obtain both positive and negative values.

The model is further evaluated in a portfolio allocation setting. Using one-step-ahead predictive return distributions simulated from the fitted dependence model, portfolio weights are chosen to maximize expected CRRA utility (with $\eta = 3$) in a weekly rebalancing backtest over 2001-03-16 to 2026-01-16. Across sector pairs, the regime-switching copula strategies do not deliver a consistent improvement in risk-adjusted performance, as summarized by Sharpe ratios, relative to an equally weighted benchmark, and the conditional TV-MRSB model does not consistently outperform the unconditional TV-MRS model. However, model-based portfolios frequently exhibit smaller maximum drawdowns, especially when allocation to a risk-free asset is permitted, suggesting that the regime-switching dependence structure can help reduce exposure to lower tail-risk.

Several limitations of the study should be noted. In the portfolio allocation exercise, the model is applied at a weekly frequency and the marginal dynamics are restricted to an AR(0)-GARCH(1,1) model. While this model is selected for daily returns in the marginal analysis, it need not remain optimal for weekly returns. Similarly, the comparative in-sample fit improvements of the conditional copula model are documented for daily returns, whereas the allocation exercise is conducted on weekly returns. Moreover, the asset-allocation backtest excludes transaction costs, which may impose a significant drag on performance given high portfolio turnover. Future work could therefore evaluate the strategy under daily rebalancing, and incorporate transaction costs or turnover constraints. Finally, since the model is primarily motivated by downside dependence and lower tail-risk, additional evaluation metrics focusing on tail risk, such as Value-at-Risk and Expected Shortfall backtests, could provide a more targeted assessment of whether the model improves the prediction and management of downside risk.

Bibliography

[1] Wong, M. (2011) *Market buVaR: A countercyclical risk metric* J. Risk Manag. Financ. Inst. 419-432

[2] Sklar, M. (1959) *Fonctions de répartition à n dimensions et leurs marges* Ann. De L'ISUP, 8, 229–231

[3] Lindgren, G. (1978). *Markov Regime Models for Mixed Distributions and Switching Regressions* Scandinavian Journal of Statistics, 5, p. 81-91.

[4] Hamilton, J.D. (1989). *A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle* Econometrica, 57, 357-384.

[5] Schwarz, G. (1978). *Estimating the Dimension of a Model* The Annals of Statistics, Vol 6., No.2, 461-464

[6] Bauwens, L. and Otranto, E. (2016). *Modeling the dependence of conditional correlations on market volatility* J. Bus. Econ. Stat., 34(2), 254–268.

[7] Wilks, S.S. (1938). *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypothesis*, Ann. Math. Statist. 60-62.

[8] Kullback, S. and Leibler R.A. (1951) *On Information and Sufficiency.* Ann. Math. Statist. Vol. 22 79-86.

[9] Akaike, H. (1974). *A new look at the statistical model identification.* Transactions on Automation Control. Vol 19. 716-723.

[10] Ang, A., and Bekaert, G. (2002). *International Asset Allocation With Regime Shifts*, Review of Financial Studies, 15, 1137–1187.

[11] Fortin, I. and Kuzmics, C. (2002). *Tail-dependence in stock-return pairs* Computational Methods in Econ. and Fin. 89-107.

[12] Ljung, G.M. and Box, G.E.P. (1978) *On a Measure of Lack of Fit in Time Series Models.* Oxford University Press. 297-303.

[13] Jarque , C.M. and Bera, K. (1987) *A Test for Normality of Observations and Regresion Residuals* International Statistical Review Vol.55. 163-172.

[14] Dickey, D.A. and Fuller, W.A. (1976). *Distribution of the Estimator for Autoregressive Time Series with a Unit Root.* Journal of the American Statistical Association Vol.74, 427-431.

[15] Engle, R.F. (1982) *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation.* Econometrica Vol.50 987-1007.

[16] Cappiello, L., Engle, R. F., and Sheppard, K. (2006), *Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns*, Journal of Financial Econometrics, 4, 537–572.

[17] Smirnov, N. (1933) *Table for estimating the goodness of fit of empirical distributions.* Annals of Mathematical Statistics. 19. 279-281

[18] Forbes, C. S., and Chinn, M. D. (2004), *A Decomposition of Global Linkages in Financial Markets Over Time*, The Review of Economics and Statistics,86, 705–722.

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

[20] Kolmogorov, A.N. (1933), *Sulla determinazione empirica di una legge di distribuzione*, Giornale dell'Instituto Italiano degli Attauri. 4. 83-91.

[21] Durante, F. Fernandez-Sanchez, J., Sempi, C. (2013) *How to Prove Sklar's Theorem* Advances in Intelligent Systems and Computing Vol. 228

[22] Brunnermeier, M.K. and Oehmke, M., *Bubbles, financial crises, and systemic risk*, Handbook of the Economics of Finance,Vol.2, pp. 1221–1288, 2013.

[23] Jondeau, E. and Rockinger, M., *The copula-GARCH model of conditional dependencies: An international stock market application*, J. Int. Money. Finance., 2006, 25(5), 827–853.

[24] Christoffersen, P., Errunza, V., Jacobs, K. and Langlois, H., *Is the potential for international diversification disappearing? A dynamic copula approach*, Rev. Financ. Stud., 2012

[25] Lucas, A., Schwaab, B. and Zhang, X., *Modeling financial sector joint tail risk in the euro area*, J. Appl. Econ., 2017, 32(1), 171 191.

[26] Blasques, F., Koopman, SJ. and Nientker, M., *A time-varying paramter model for local explosion.*, J. Econom., 2022, 227, 65-84.

[27] Koopman, S.J., Lit, R., Lucas, A. and Opschoor, A., *Dynamic discrete copula models for high-frequency stock price changes*, J. Appl. Econ., 2018, 33(7), 966–985

[28] Wuyi Ye, Lingbo Gao and Xiaoquan Liu (2024) *Bubbles and dependence between international equity markets*, Quantitative Finance, 24:1, 119-138

[29] Kim,C.J. and Nelson, C.R., *State-Space Models with Regime Switching: Classical And Gibbs-Sampling Approaches with Applications*, Vol. 1, 1999

[30] Andrew J. Patton (2006) *On the Out-of-Sample Importance f Skewness and Asymmetric Dependence for Asset Allocation*, Int. Econ. Rev. 2006, 47(2), 527-556

DEPARTMENT OF MATHEMATICS
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY