Journal of Data Science 24 (2), 296–318
April 2026
DOI: 10.6339/26-JDS1229
Statistical Data Science

Semiparametric Dynamic Copula Models using Rolling-window Portfolio Optimization

Savita Pareek[1,*] and Sujit K. Ghosh[2]

[1]*Department of Mathematics and Statistics, Auburn University, Auburn, United States*
[2]*Department of Statistics, NC State University, Raleigh, United States*

Abstract

The mean-variance portfolio model, based on the risk-return trade-off for optimal asset allocation, remains fundamental in portfolio optimization. However, its reliance on restrictive assumptions about asset return distributions limits its applicability to real-world data. Parametric copula structures provide a novel way to overcome these limitations by accounting for asymmetry, heavy tails, and time-varying dependencies. Existing methods have been shown to rely on fixed or static dependence structures, thus overlooking the dynamic nature of the financial market. In this study, a semiparametric model is proposed that combines nonparametrically estimated copulas with parametrically estimated marginals to allow all parameters to dynamically evolve over time. A novel framework was developed that integrates time-varying dependence modeling with flexible empirical beta-copula structures. Marginal distributions were modeled using the skewed generalized t-family. This effectively captures asymmetry and heavy tails and makes the model suitable for predictive inferences in real-world scenarios. Furthermore, the model was applied to rolling windows of financial returns from the USA, India, and Hong Kong economies to understand the influence of dynamic market conditions. The approach addresses the limitations of models that rely on parametric assumptions. By accounting for asymmetry, heavy tails, and cross-correlated asset prices, the proposed method offers a robust solution to optimize diverse portfolios in an interconnected financial market. Through adaptive modeling, it allows for better management of risk and return across varying economic conditions, leading to more efficient asset allocation and improved portfolio performance.

Keywords *empirical beta copula; Markowitz portfolio optimization; skewed generalized t-distribution*

1 Introduction

Portfolio optimization aims to construct investment strategies by selecting an optimal mix of assets to achieve objectives such as maximizing returns, minimizing risk, and ensuring stability. Markowitz's mean-variance portfolio model (Markowitz, 1952) remains foundational, enabling optimal asset allocation based on the risk-return trade-off. However, its reliance on restrictive assumptions about asset return distributions limits its applicability to real-world data, which often exhibit asymmetry, heavy tails, and time-varying dependencies.

To address these limitations, advanced strategies have been developed to be used alongside parametric copula structures. These include maximum Sharpe Ratio (MSR), Global Minimum

*Corresponding author. Email: savy.pareek@gmail.com.

Received September 16, 2025; Accepted March 21, 2026

Variance (GMV), Conditional Value-at-Risk (CVaR) and hierarchical risk parity based strategies. For example, Huang et al. (2015) constructed a time-varying copula with varying window lengths to examine the impact of parametric copula families (e.g., Gaussian, Gumbel, Clayton) on portfolio performance during economic cycles, utilizing Markowitz's variance minimization method. Sahamkhadam and Stephan (2023) extended portfolio optimization through out-of-sample predictive models, employing AR-GARCH-type processes for marginal asset returns and Vine copulas for joint distributions, focusing on GMV, CVaR, and MSR strategies. Valeyre (2023) proposed an optimal trend-following portfolio integrating Markowitz, risk parity, agnostic risk parity, and trend-following strategies. Similarly, Avella (2024) evaluated the performance of Markowitz optimization, highlighting the superiority of global minimum variance and mean-variance portfolios over randomly weighted approaches.

The application of copula models in portfolio optimization has evolved significantly, with numerous studies exploring both parametric and nonparametric approaches. While parametric copulas such as Gaussian, Student-t, and Archimedean families have been extensively studied (McNeil and Nešlehová, 2009; Patton, 2012), they impose restrictive assumptions on dependence structures that may not capture the complex relationships observed in financial markets. Nonparametric copula estimation, including kernel-based methods (Chen and Huang, 2007), Bernstein copulas (Sancetta and Satchell, 2004), and empirical copulas (Deheuvels, 1979), offers greater flexibility but often suffers from the curse of dimensionality and requires large sample sizes for reliable estimation. More recently, *B-spline copulas* (Shen et al., 2008; Dou et al., 2025) have emerged as an important addition to this class; they are valid copulas with uniform margins and include the Bernstein copula as a special case, providing a smooth and computationally efficient framework for flexible dependence modeling.

Semiparametric copula models, which combine parametric marginals with nonparametric copulas or vice versa, have emerged as a compromise between flexibility and efficiency. Fermanian and Scaillet (2002) developed semiparametric copula estimation with parametric margins, while Chen and Fan (2006) proposed nonparametric margins with parametric copulas. However, the choice between parametric margins with nonparametric copulas versus the reverse depends critically on the specific characteristics of the data and the dimensionality of the problem. Our approach of using parametric Skewed Generalized t (SGT)-margins with nonparametric empirical beta copulas is motivated by three key considerations: (1) the empirical beta copula provides a genuine copula for finite samples unlike many nonparametric alternatives, (2) the SGT-distribution effectively captures the well-documented stylized facts of individual asset returns, and (3) the computational efficiency of the approach enables practical implementation in rolling window frameworks. Figure 1 presents an overview of the proposed semi-parametric framework, which integrates SGT marginals, empirical beta copulas, and simulation-based optimization for dynamic portfolio allocation.

Specifically, modeling the univariate marginals using the flexible SGT family allows us to capture heavy tails and asymmetry that are commonly observed in financial return data. The SGT-distribution has been extensively validated across a wide range of asset classes, markets, and time periods, providing a robust parametric framework for marginal behavior. Moreover, univariate parametric models such as SGT facilitate the use of well-established goodness-of-fit tests, which enable rigorous model validation, an advantage not easily available for fully nonparametric marginal models. In contrast, nonparametric density estimators (e.g., kernel methods) can be sensitive to bandwidth selection and may not perform well in capturing extreme tail behavior in large datasets of financial returns.

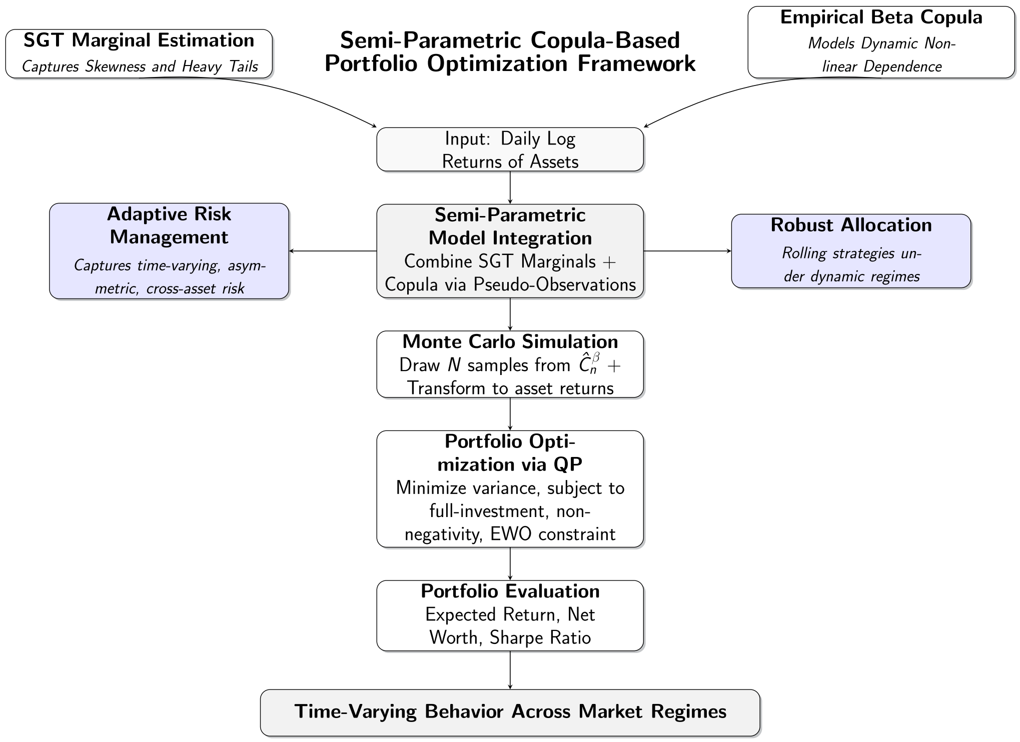

Figure 1: Overview of the semi-parametric copula-based portfolio optimization framework.

On the other hand, specifying a parametric copula model in higher dimensions can impose restrictive assumptions on the form of dependence (e.g., as in vine or elliptical copulas). To mitigate this, we adopt a nonparametric empirical beta copula, which provides greater flexibility in capturing complex and potentially nonlinear dependence structures without requiring strong parametric assumptions.

In summary, the semiparametric framework is designed to balance flexibility and interpretability: the parametric SGT marginals ensure accurate modeling of heavy-tailed and asymmetric univariate behaviors with testable fit, while the nonparametric beta copula allows for data-driven inference on the dependence structure. This pairing leverages the strengths of both approaches while mitigating their individual limitations. Additional figures, tables, and implementation details are provided in the Supplementary Material.

2 Methodology

This section outlines the proposed semiparametric dynamic copula framework for portfolio optimization. The modeling pipeline integrates three main components. First, the marginal distributions of asset returns are modeled using the SGT-distribution, which captures heavy tails and asymmetry that characterize financial data. Second, the cross-asset dependence structure is modeled nonparametrically via the empirical beta copula, providing flexibility to represent nonlinear and tail dependence without restrictive functional forms. Third, the estimated joint distribution is employed within a mean–variance optimization framework that incorporates three practical constraints: full investment, long-only positions, and an equal-weighted outperformance

(EWO) condition, which requires the optimized portfolio to achieve a mean return no lower than the average return of its constituent assets. All components: marginal estimation, copula construction, and constrained optimization, are implemented in a rolling-window scheme, enabling the model to evolve adaptively with changing market conditions.

Portfolio theory provides a quantitative framework for building models of volatile assets. It involves three key steps: identifying the joint distribution and dependence structure of returns, selecting a mathematical model to represent the risk-return trade-off, and solving the model to achieve optimal portfolio outcomes. Consider a portfolio consisting of m asset prices, where the return of each asset is observed over n time points. For each asset j, the return vector is denoted as $\mathbf{R}_j = (R_{2j}, \ldots, R_{nj}, R_{(n+1)j})^T$, representing the returns across n time periods. The mean return of the j^{th} asset is given by $\mu_{\boldsymbol{R}_j}$. The steps to construct the portfolio are outlined below:

1. **Definition of Returns:** The returns are defined as:

$$R_{tj} = \log\left(\frac{p_{tj}}{p_{(t-1)j}}\right),$$

where p_{tj} denotes the adjusted closing price of the j^{th} asset on the t^{th} day, with $t = 2, \ldots, n+1$ and $j = 1, 2, \ldots, m$.

2. **Marginal Distribution Estimation and Pseudo-Uniform Transformation:** Estimate the marginal distribution of the returns R_{tj} and transform each marginal to pseudo-uniform variates:

$$U_{tj} = \hat{F}_j(R_{tj}), \quad t = 2, \ldots n+1; \; j = 1, 2, \ldots, m,$$

where $\hat{F}_j$ is the estimated marginal cumulative distribution function using SGT model. More details are provided in Section 2.1. This step generates an $n \times m$ matrix of pseudo-uniform marginals.

3. **Copula Estimation:** Estimate the empirical beta copula denoted by $\hat{C}_n^{\beta}(u_1, u_2, \ldots, u_m)$ using the pseudo-uniform variates $\{(U_{t1}, \ldots, U_{tm}) : t = 2, \ldots, n+1\}$. Further details are provided in Section 2.2. Supplementary Figure S1 summarizes the data preprocessing and dependence modeling framework, encompassing Steps 1 through 3 of the proposed methodology.

4. **Generation of Random Variates and Reverse Transformation:** Generate random variates from the estimated genuine copula and transform them back to the marginal returns using the quantile function or by reversing the method in Step 2:

$$\text{Generate } (\tilde{U}_{i1}, \tilde{U}_{i2}, \ldots, \tilde{U}_{im}) \overset{iid}{\sim} \hat{C}_n^{\beta}(u_1, u_2, \ldots, u_m);$$
$$\tilde{R}_{ij} = \hat{F}_j^{-1}(\tilde{U}_{ij}), \; i = 1, \ldots, N; \; j = 1, \ldots, m,$$

where N is the Monte Carlo sample size chosen to achieve a desired precision. This step gives a $N \times m$ matrix of $\tilde{\boldsymbol{R}}$'s which can be used to estimate covariance matrix $\boldsymbol{\Sigma}_{\tilde{\boldsymbol{R}}}$ based on the multivariate sample $\{(\tilde{R}_{i1}, \ldots, \tilde{R}_{im}) : i = 1, 2, \ldots, N\}$

5. **Portfolio Optimization:** Let $\mathbf{w} = (w_1, \ldots, w_m)^\top$ denote the m-dimensional vector of portfolio weights, where w_j represents the proportion of total wealth invested in asset j. The portfolio is optimized using Markowitz's mean–variance framework:

$$\min_{\mathbf{w}} \mathbf{w}^T \boldsymbol{\Sigma}_{\tilde{\boldsymbol{R}}} \mathbf{w}, \quad \text{subject to:}$$

- $\sum_{j=1}^{m} w_j = 1$ (Full investment, weights sum to 1)

- $w_j \geqslant 0,\ j = 1, 2, \dots m$ (Long-only, non-negativity constraint)
- $\sum_{j=1}^{m} w_j \mu_{\boldsymbol{R}_j} \geqslant \frac{1}{m} \sum_{j=1}^{m} \mu_{\boldsymbol{R}_j}$ (EWO, optimal weighted return is greater or equal the equal-weighted return).

In this step, alternative optimization strategies such as MSR, CVaR, or VaR can also be employed. The portfolio weights $\hat{w}_j$, are estimated by solving quadratic programming optimization problem described above.

6. **Portfolio Net Worth Value and Sharpe Ratio Computation:** The portfolio's net worth and Sharpe ratio are computed using the next day's expected returns, covariance matrix, and the current optimal portfolio weights. Specifically:

$$\text{Net Worth} = \hat{\boldsymbol{w}}^\top E[\boldsymbol{r}_{(t+f)}], \qquad \text{Sharpe Ratio} = \frac{\hat{\boldsymbol{w}}^\top E[\boldsymbol{r}_{(t+f)}]}{\sqrt{\hat{\boldsymbol{w}}^\top \mathrm{Cov}[\boldsymbol{r}_{(t+f)}]\hat{\boldsymbol{w}}}},$$

where $\hat{\boldsymbol{w}}$ denotes the vector of optimal portfolio weights estimated from the rolling window $\boldsymbol{R}_{(t)}$. Here, $\boldsymbol{R}_{(t)} \in \mathbb{R}^{L \times m}$ represents the matrix of daily log-returns for m assets over a window of length L, spanning the period from the t^{th} to the $(t + L - 1)^{\text{th}}$ trading day. The vector $\boldsymbol{r}_{(t+f)} = (r_{t+f,1}, r_{t+f,2}, \dots, r_{t+f,m})^\top \in \mathbb{R}^m$ represents the asset returns observed at the next rebalancing point $(t + f)$. Assuming a rolling window of length L and a rebalancing frequency f, the window $\boldsymbol{R}_{(t)}$ is updated sequentially by including new observations and discarding the oldest ones. We consider $f = 1$ (daily) and $f = 5$ (weekly) rebalancing frequencies, although the framework generalizes naturally to any desired horizon, such as fortnightly ($f = 10$) or monthly ($f = 21$). Supplementary Figure S2 illustrates the simulation, optimization, and evaluation pipeline corresponding to Steps 4–6 of the proposed framework.

7. **Comparison of Portfolio Net Worth and Sharpe Ratio:** Compare the portfolio's next day net worth, Sharpe ratio with results obtained using equal-weighted portfolios or alternative optimization strategies.
8. **Performance Plots:** Generate plots for: (a) Empirical rolling average returns in a portfolio, (b) Empirical rolling standard deviations in a portfolio (c) Rolling optimal weights using a specific optimization criterion, and (d) Compare next day portfolio performance with other methods.

Next, we describe the details of each of the key steps above in our models.

2.1 Marginal Density Modeling using Skewed Generalized *t*-Distribution

It is well-known that financial log-returns exhibit high kurtosis, pronounced skewness, and heavy tails. While kernel density estimators (KDEs) are useful for exploratory analysis of return distributions, they are often inadequate for modeling log-returns of stock prices that display heavy tails, skewness, and high kurtosis. In particular, KDEs suffer from large relative bias and variance in the tails, are highly sensitive to bandwidth choice, and are effectively restricted to the observed data support, thereby failing to extrapolate plausible extreme values needed for risk assessment (Silverman, 1986; Sheather, 2004). These shortcomings make KDEs unreliable for risk measures and tail-dependent inference (VaR, expected shortfall) unless combined with tail-specific transforms or extreme-value methods. To address these issues, parametric families such as the SGT-distribution provide a flexible alternative. The five-parameter SGT-distribution (with parameters controlling location, scale, skewness, and two tail-shape exponents) allows simultaneous modeling of asymmetry and heavy tails, making it particularly suitable for financial returns. Empirical evidence shows that the SGT can capture the stylized facts of financial time series far more effectively than KDEs (Theodossiou, 1998; Fernández and Steel, 1995; Davis, 2015).

Due to its ability to account for skewness and excess kurtosis, the SGT is particularly effective at capturing the true distribution of asset returns, exchange rates, and commodity prices. Theodossiou (1998) provides empirical evidence that this model represents financial risks with a better fit than symmetrical models. In addition, Theodossiou and Savva (2016) demonstrate that skewness plays a key role in explaining inconsistencies in the risk-return relationship, demonstrating that models that assume symmetry underestimate risk exposure.

Many well-known heavy-tailed and skewed distributions are special cases of the SGT-distribution, including the Skewed Generalized Error, Generalized t-distribution, Skewed t-distribution, Skewed Laplace, Generalized Error, Skewed Normal, Student's t-distribution, Skewed Cauchy, Laplace, Uniform, Normal, and Cauchy distributions. The probability density function (PDF) of the SGT-distribution is defined as follows.

A random variable X is said to follow the SGT-distribution, denoted $X \sim \mathrm{SGT}(\mu, \sigma, \lambda, p, q)$, if its PDF is given by:

$$f_{\mathrm{SGT}}(x; \mu, \sigma, \lambda, p, q) = \frac{p}{2v\sigma q^{1/p} B(\frac{1}{p}, q)} \left(\frac{|x - \mu + m|^p}{q(v\sigma)^p(\lambda\,\mathrm{sign}(x - \mu + m) + 1)^p} + 1 \right)^{-\frac{1}{p} - q},$$

where $x \in \mathbb{R}$, and the parameters satisfy $\mu \in \mathbb{R}$, $\sigma > 0$, $-1 < \lambda < 1$, and $p, q > 0$. The parameter μ defines the central location of the distribution, while σ determines its scale. The skewness of the distribution is controlled by λ, whereas p and q jointly regulate its kurtosis.

When $\lambda = 0$, the distribution is symmetric. For $-1 < \lambda < 0$, it is negatively skewed, while $0 < \lambda < 1$ results in positive skewness. Smaller values of p and q produce a leptokurtic distribution with heavy tails, whereas larger values yield a platykurtic distribution with lighter tails. The h-th moment of the SGT-distribution exists only if $pq > h$, indicating that the finiteness of higher-order moments depends on the choice of these parameters.

The terms m and v are defined as:

$$m = \frac{2v\sigma\lambda q^{1/p} B(\frac{2}{p}, q - \frac{1}{p})}{B(\frac{1}{p}, q)},$$

$$v = q^{-\frac{1}{p}} \left[(3\lambda^2 + 1) \frac{B(\frac{3}{p}, q - \frac{2}{p})}{B(\frac{1}{p}, q)} - 4\lambda^2 \left(\frac{B(\frac{2}{p}, q - \frac{1}{p})}{B(\frac{1}{p}, q)} \right)^2 \right]^{-\frac{1}{2}},$$

where the beta function is defined as $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$, with $\Gamma(\cdot)$ denoting the gamma function. This study utilizes the SGT-distribution to model financial return data with asymmetry and fat tails. Its flexibility in capturing skewness and excess kurtosis makes it ideal for such data. By applying a rolling window approach, we fit the SGT-distribution dynamically, allowing for time-varying analysis and improved modeling of financial market dynamics.

2.2 Nonparametric Copula Model Using Smoothed Beta Copula

Copula models are increasingly used in financial time series analysis to capture the evolving dependence structure in asset returns and stock price movements. Their ability to model complex dependencies makes them a fundamental tool for analyzing multivariate data, as they link univariate distribution functions into a multivariate framework. By Sklar's theorem (Sklar, 1959), any joint distribution can be uniquely expressed in terms of its marginal distributions and a copula, which encapsulates the dependence structure among multiple random variables. This decomposition enables flexible estimation, allowing marginals and the copula to be modeled separately using distinct techniques (Jaworski et al., 2010; Patton, 2012; Joe, 2014). For a portfolio return vector of d assets $(X_1, \ldots, X_d)$ with a joint cumulative distribution function (CDF) F and continuous marginal CDFs $F_j,\ j = 1, \ldots, d$, Sklar's theorem asserts that:

$$F(x_1, \ldots, x_m) = C\big(F_1(x_1), \ldots, F_d(x_d)\big),$$

where $C(\cdot)$ is the copula function. A copula is the joint CDF of a transformed random vector $(U_1 = F_1(X_1), \ldots, U_d = F_d(X_d))$, where the marginal distributions are uniform on $[0, 1]$. While Sklar's theorem applies to discrete variables, this study focuses on continuous multivariate random vectors (e.g., stock returns), assuming all marginal CDFs F_j are absolutely continuous.

Copulas can be estimated through parametric or nonparametric methods. Parametric models, such as Gaussian, *t*, or Archimedean copulas (Nelsen, 2007; McNeil and Nešlehová, 2009), are computationally efficient and interpretable but may fail to capture complex dependencies like asymmetry or tail dependence, especially under model misspecification. Vine copulas address some of these challenges by constructing dependence structures using bivariate copulas at each node of a vine tree. However, their estimation often involves high-dimensional integrals, posing computational challenges. Nonparametric approaches, including empirical, Bernstein, and beta copulas, provide greater flexibility in capturing arbitrary dependencies (Sancetta and Satchell, 2004; Genest et al., 2017; Segers et al., 2017). However, many nonparametric estimators are valid copulas only asymptotically, limiting their applicability to finite samples. Furthermore, dependence measures like Spearman's rho and Kendall's tau, based on these estimators, can fall outside their natural range, making them impractical in some scenarios. To overcome these limitations, Segers et al. (2017) introduced the empirical beta copula (EBC), a valid, smooth copula constructed as a special case of the empirical Bernstein copula with polynomial degrees equal to the sample size. It improves bias and mean squared error over classical estimators, but may have higher variance than Bernstein copulas with lower degrees. Lu and Ghosh (2023) extended this idea via the empirical checkerboard Bernstein copula (ECBC), allowing for varying polynomial degrees. In this work, we adopt the EBC due to its greater computational efficiency compared to ECBC. The EBC is defined as:

$$C_n^{\beta}(\boldsymbol{u}) = \frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d} F_{n,R_{i,j}^{(n)}}(u_j), \quad \boldsymbol{u} = (u_1, \ldots, u_d) \in [0, 1]^d,$$

where $R_{i,j}^{(n)}$ denotes the rank of X_{ij} among $(X_{1j}, \ldots, X_{nj})$. For $u \in [0, 1]$ and $r \in \{1, \ldots, n\}$,

$$F_{n,r}(u) = \mathbb{P}(U_{r:n} \leqslant u) = \sum_{s=r}^{n}\binom{n}{s}u^s(1-u)^{n-s}$$

is the cumulative distribution function of beta distribution $\mathcal{B}(r, n + 1 - r)$. Here, $U_{1:n} < \cdots < U_{n:n}$ denote the order statistics based on n independent random variables $U_1, \ldots, U_n$ uniformly distributed on $[0, 1]$.

We employ the EBC to estimate time-varying copulas over rolling windows of financial returns. This approach is particularly useful for capturing dynamic dependence structures, as dependencies between assets often evolve during periods of market turbulence, such as financial crises. We implement the EBC using the `empCopula` function from the `copula` package in R, with smoothing set to `beta` (Hofert et al., 2024). This method is further integrated into a semiparametric framework, where the marginals are modeled using the SGT-distribution, enabling us to capture asymmetry, heavy tails, and time-varying dependencies in financial markets.

To capture the time-varying nature of financial returns, the proposed semiparametric copula framework is implemented through a rolling-window estimation scheme. For each time index $t = L, \ldots, T$, where L denotes the rolling window length, both the SGT marginal parameters and the empirical beta-copula dependence structure are re-estimated using the most recent L observations $\{(X_{i,t-L+1}, \ldots, X_{i,t}) : i = 1, \ldots, d\}$. The resulting joint distribution is then used to update the optimal portfolio weights $\mathbf{w}_t$ under full-investment, long-only, and EWO constraints. This sequential re-estimation and re-optimization process enables the model to adapt to evolving market conditions, such as volatility clustering, structural breaks, and regime shifts, while preserving local stationarity within each window.

The rolling window approach is consistent with the notion of *local stationarity* (see, e.g., Dahlhaus, 1997; Politis, 2011), which assumes that the underlying data-generating process is approximately stationary within short time intervals but may evolve over longer horizons. By contrast, expanding-window approaches implicitly assume long-term stationarity and may therefore be less responsive to recent market conditions. Hence, the rolling window estimation provides a practical and flexible mechanism for modeling and optimizing portfolios under locally stationary, time-varying dependence structures.

2.3 Portfolio Optimization using Quadratic Programming

Optimizing portfolios and diversifying investments have been crucial in financial decision making. Markowitz's mean variance optimization model (MVO) (Markowitz, 1952) introduced a quantitative framework to balance risk and return, formulating portfolio selection as an optimization problem. It emphasizes diversification, where risk depends on asset correlations rather than individual asset risk. Using the model, investors select portfolios with the lowest variance and reject inefficient portfolios with higher risk. This approach revolutionized classical financial analysis by shifting the focus from single-asset valuation to portfolio-level risk management. The MVO framework is typically formulated as a quadratic programming problem, where investors maximize expected returns for a given risk level or minimize portfolio variance for a required return. For a detailed exposition and practical considerations in portfolio optimization, refer to Bacon (2008), Kolm et al. (2014), and Palomar (2025). Let m be the number of assets in an investment universe with uncertain future returns denoted by $\mathbf{r} = (r_1, r_2, \ldots, r_m)^T$. A portfolio is represented by the weight vector $\boldsymbol{w} = (w_1, w_2, \ldots, w_m)^T$, where w_i is the proportion of total funds allocated to asset i. The portfolio return is given by:

$$r_p(\boldsymbol{w}) = \boldsymbol{w}^T \mathbf{r}.$$

Expected portfolio returns and standard deviation (risk) of the portfolio return are represented as:

$$\mu(\boldsymbol{w}) = \boldsymbol{\mu}^T \boldsymbol{w}, \quad \sigma(\boldsymbol{w}) = \sqrt{\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}},$$

where:

- $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_m)^T$, $\mu_i = E(r_i)$ is the expected return of asset i.
- $\boldsymbol{\Sigma} = \operatorname{diag}(\boldsymbol{\sigma})\boldsymbol{P}\operatorname{diag}(\boldsymbol{\sigma})$ is the covariance matrix, which is positive semi-definite, ensuring $\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} \geqslant 0$.
- $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_m)^T$ is the vector of standard deviations.

- $\boldsymbol{P} = [\rho_{ij}]$ is the correlation matrix, where ρ_{ij} is the correlation between asset i and j.

Let $\Omega \subset \mathbb{R}^m$ denote the set of permissible portfolios, where $\boldsymbol{w} \in \Omega$ satisfies the portfolio constraints. Using this framework, the MVO problem is formulated as:

$$\max_{\boldsymbol{w}\in\Omega} \boldsymbol{\mu}^T \boldsymbol{w} - \gamma \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}, \quad \text{subject to: } \mathbf{1}^T \boldsymbol{w} = 1, \quad \boldsymbol{w} \geqslant 0,$$

where γ is the investor-specific risk-aversion parameter governing the trade-off between expected portfolio return and portfolio risk. When $\gamma = 0$, the optimization problem focuses solely on maximizing expected return, resulting in the global maximum return portfolio. Conversely, as $\gamma \to \infty$, the problem minimizes risk entirely, producing the GMV portfolio. Alternative formulations of portfolio optimization include the MSR portfolio, as well as approaches based on alternative risk measures such as downside risk, semivariance, VaR, CVaR, and drawdown (Gunjan and Bhattacharyya, 2023).

In this study, we implement the GMV portfolio strategy using a covariance matrix estimated from copula-based simulated returns. To ensure a reasonable level of predictive performance, the optimization is subject to long-only, full investment, and an additional EWO return constraint requiring the portfolio's expected return to exceed the average return of the individual assets. The optimization problem is stated as:

$$\min_{\boldsymbol{w}\in\Omega} \boldsymbol{w}^\top \boldsymbol{\Sigma}_{\tilde{\mathbf{R}}} \boldsymbol{w}, \quad \text{subject to:} \quad \mathbf{1}^\top \boldsymbol{w} = 1, \quad \boldsymbol{w} \geqslant 0, \quad \boldsymbol{\mu}^\top \boldsymbol{w} \geqslant \bar{r},$$

where $\bar{r} = \frac{1}{m} \sum_{i=1}^{m} \mu_i$ denotes the average expected return of the m assets, $\boldsymbol{\mu}$ is the vector of expected asset returns, and $\boldsymbol{\Sigma}_{\tilde{\mathbf{R}}}$ is the covariance matrix estimated from the copula-implied joint return distribution. We solve this optimization problem dynamically over rolling windows using the `solve.QP` function from `quadprog` package in `R` (Turlach and Weingessel, 2019). By obtaining optimal weights for each window, we capture changes in market dynamics and improve portfolio performance over time.

The proposed portfolio optimization framework introduces a flexible structure by combining three components: (1) nonparametric copula-based modeling of dependence, (2) the inclusion of EWO return constraint to ensure practical applicability, and (3) a rolling window of approximately one year of daily returns to dynamically adjust portfolio weights. The first component addresses the limitations of commonly used parametric copulas, such as Gaussian and t-copulas, which can capture certain nonlinear and tail dependencies but are limited by their fixed functional forms. These limitations become more pronounced in high-dimensional settings or when the true dependence structure is complex or asymmetric. In contrast, nonparametric copulas provide greater flexibility by allowing the joint dependence structure to be estimated directly from the data without imposing a specific parametric form (Nelsen, 2007; Patton, 2012). This is particularly advantageous when asset returns exhibit non-Gaussian features, as is commonly observed in financial markets. The second component, an EWO return constraint, ensures that the optimized portfolio achieves an expected return at least as high as that of a naive equal-weighted portfolio. This not only aligns the model with realistic investor expectations, but also safeguards against low-return allocations that may arise in purely risk-driven optimization settings. The third component incorporates a rolling window framework (typically 250 trading days, adjusted for the specific market's calendar) with daily or weekly rebalancing to account for changing market dynamics, producing real-time optimal allocations that respond to structural shifts in the joint return distribution.

A key theoretical motivation for this framework lies in the inadequacy of assuming multivariate normality. When asset marginals are normally distributed, the joint distribution is also multivariate normal, and the covariance matrix fully characterizes the dependence structure. However, empirical evidence consistently shows that financial returns exhibit skewness, leptokurtosis, and heavy tails, often due to volatility clustering, price jumps, and higher-moment dependencies (Bali and Theodossiou, 2007). When marginals follow skewed distributions such as SGT or skew-normal, there is no natural or unique multivariate extension that preserves the given marginals while maintaining analytical tractability (Theodossiou, 1998; Azzalini and Capitanio, 2003). Moreover, the covariance structure cannot be inferred directly from these marginals, making mean-variance optimization unsuitable. In contrast, the copula-based approach allows one to model such joint behaviors by specifying each marginal independently and linking them through a copula, enabling simulation of artificial datasets that reflect the empirical joint behavior. Although this approach has a small computational overhead, it offers significantly improved realism and predictive accuracy. While our approach is anchored in the classical mean-variance paradigm, it can be extended to accommodate higher-order moments or incorporate quantile-based risk measures, such as VaR and CVaR, to better reflect tail-risk behavior in non-Gaussian return distributions.

It is to be noted that the rolling window approach assumes local stationarity within each window, which is a more flexible and realistic assumption than global (long-term) stationarity, especially for financial return series that often exhibit structural breaks or volatility clustering. This local assumption is one of the key motivations for adopting a rolling window framework instead of an expanding window, as it allows model parameters to adapt over time and remain robust to evolving market conditions.

2.4 Visualization of Portfolio Performance Measures

This section analyzes the time-varying performance of two U.S. assets, Netflix (NFLX), a volatile stock, and Costco (COST), a stable stock, using daily log returns from April 2018 to March 2025 with weekly rebalancing. We examine rolling mean returns, standard deviations, Sharpe ratios, correlations, and return densities to capture their dynamic risk-return profiles and regime shifts over time.

Figure 2 shows the dynamic nature of stock returns through rolling density plots. For COST, high peaks suggest that returns are concentrated around the mean, indicating lower variability, while lower peaks signal periods of increased volatility. NFLX, on the contrary, exhibits wider and flatter distributions, reflecting greater variance and market sensitivity. These evolving density patterns suggest that the return dynamics is influenced by market conditions and investor sentiment, although the changes are not drastic. The adaptability of the SGT-distribution effectively captures high kurtosis and skewness, making it a reliable model for financial returns.

Figure 3 provides a comprehensive view of the evolving return dynamics between NFLX and COST. Panel (a) displays the rolling mean returns, highlighting pronounced fluctuations for NFLX (ranging from approximately -0.004 to 0.004), while COST exhibits a narrower band (approximately -0.001 to 0.002), consistent with its lower market sensitivity. Panel (b) shows the rolling standard deviations, where NFLX consistently exhibits higher volatility (0.02 to 0.05), depicting its riskier profile. Panel (c) presents the rolling Sharpe ratios, which remain generally below 0.2 for both stocks, indicating modest risk-adjusted performance over time. Panel (d) illustrates the time-dependent relationship between NFLX and COST through rolling correlations. The correlation strength fluctuates between 0.2 and 0.5, reflecting varying co-movement influenced by market conditions. Pearson correlation indicates a strong yet time-varying linear dependence, while Spearman and Kendall correlations exhibit relative stability, suggesting con-

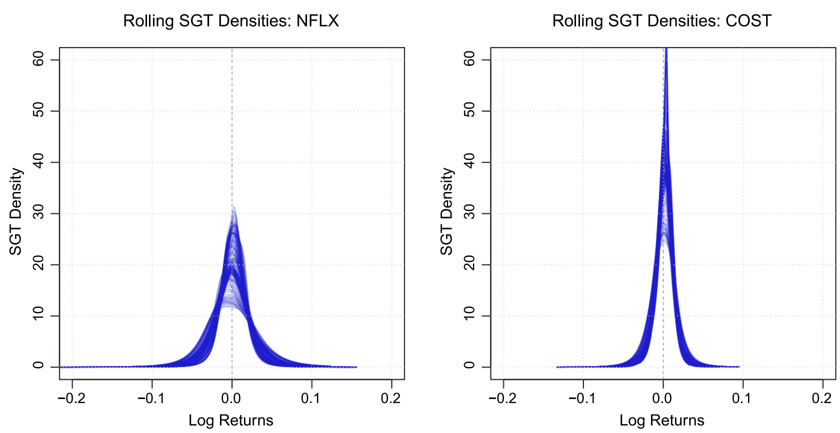

Figure 2: Overlaid SGT-fitted density plots for the log-differenced adjusted returns of NFLX and COST over a 250-day rolling window with weekly rebalancing. The comparison highlights differences in tail behavior and asymmetry between a high-volatility (NFLX) and low-volatility (COST) asset.

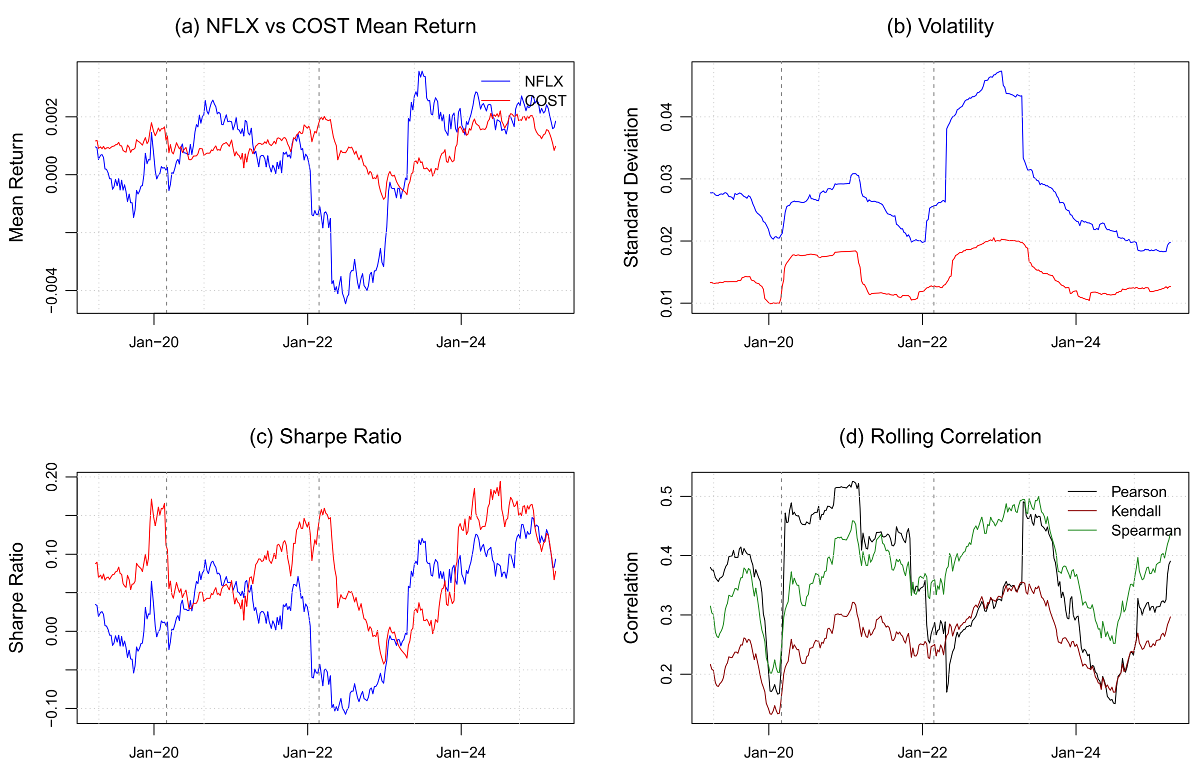

Figure 3: Rolling performance metrics for NFLX and COST based on log Returns (250-Day window, weekly rebalancing): Plot (a) represents the rolling expected portfolio return, (b) shows the rolling standard deviation, and (c) illustrates the rolling Sharpe ratio, defined as the ratio of mean return to standard deviation. Plot (d) depicts the rolling correlation using Spearman's rank, Kendall's tau, and Pearson correlation, capturing the time-varying relationships between NFLX and COST assets.

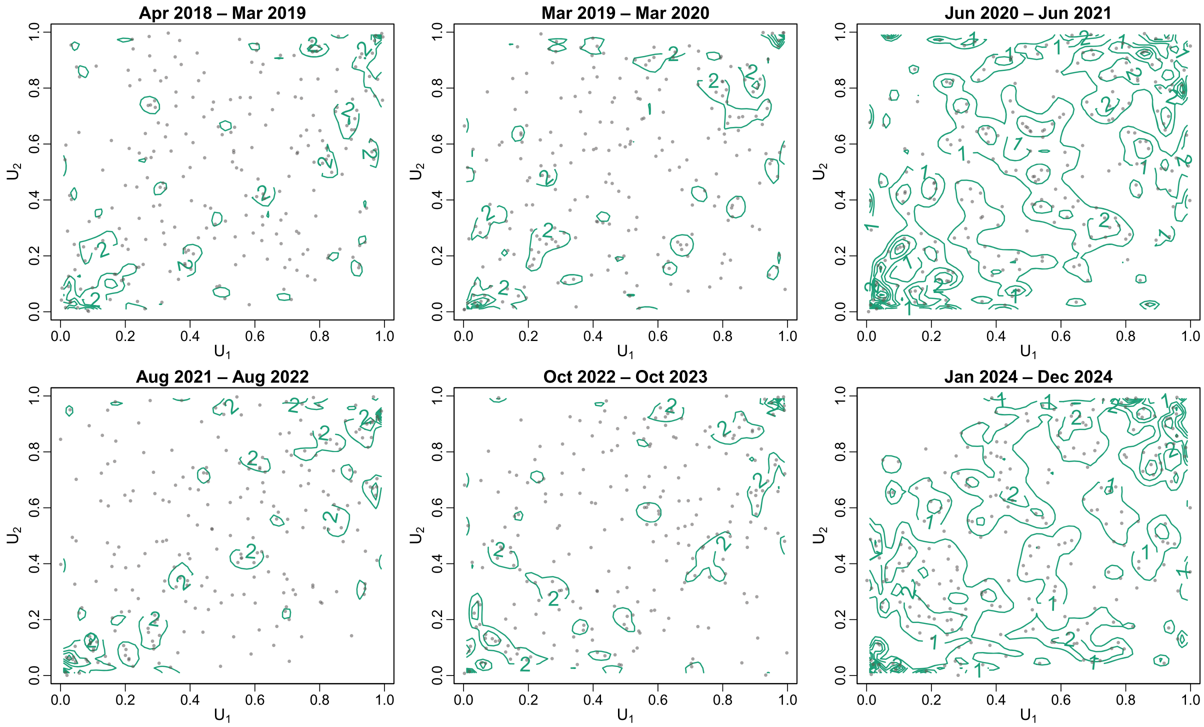

Figure 4: Rolling empirical beta copula contours for log returns of Netflix (U1) and Costco (U2) with SGT-fitted marginals. The six panels correspond to rolling 12-month windows from April 2018 to December 2024, illustrating the evolving dependence structure between the two assets.

sistent rank ordering despite fluctuations in absolute returns. Kendall's τ correlation is lower than Spearman's ρ due to its stricter pairwise ranking comparison, making it more sensitive to minor fluctuations, whereas Spearman captures broader monotonic relationships.

Figure 4 presents contour plots of empirical beta copulas estimated from pseudo-uniform transformed log returns of Netflix and Costco, with marginals modeled using the SGT-distribution. The contours reveal substantial temporal variation in dependence between the two assets. Initially, in the pre-pandemic period (April 2018–March 2019), dependence is relatively symmetric and weak, indicative of limited joint extreme events. During the early stages of the COVID-19 pandemic (March 2019–June 2021), the contours exhibit enhanced upper-tail dependence, reflecting stronger simultaneous positive return shocks. In contrast, the post-pandemic period (August 2021–December 2024) features asymmetry and pronounced nonlinear dependence, characterized by contours that deviate significantly from the diagonal, highlighting more complex co-movements and persistent upper-tail clustering.

In summary, these findings highlight the empirical beta copula's flexibility in capturing dynamic, non-Gaussian dependencies, particularly strong tail dependence during market stress, compared to static parametric models. This underscores the necessity of rolling-window, non-parametric approaches in financial applications.

Table 1: Summary of rolling window (RW) setup and market characteristics.

Economy	Daily returns	RW size	# RW (daily)	# RW (weekly)	Avg volatility (SD, %)
United States	1,758	250	1,508	301	2.01
India	1,728	245	1,483	296	1.80
Hong Kong	1,719	244	1,475	295	1.72

3 Financial Data

We evaluate the proposed portfolio optimization framework using stock price data from three major markets: the United States (developed), India (developing), and Hong Kong (representing the Far East Asian financial hub). The empirical analysis spans April 1 2018 to March 31st 2025, capturing both stable periods and episodes of extreme market stress, including the COVID-19 crisis. Daily equity prices are retrieved from Yahoo Finance via the `quantmod` package in `R` Ryan and Ulrich (2023). For each asset, we extract adjusted closing prices using the `Ad()` function and compute log returns as `diff(log(Ad(get(ticker))))`, ensuring consistency by accounting for corporate actions such as dividends and stock splits. To capture evolving market dynamics and time-varying dependencies, we implement a rolling window approach based on one trading year, with window lengths aligned to local exchange calendars.

Each portfolio comprises 20 large-cap companies selected from the S&P 500 (United States), NIFTY 50 (India), and the top-traded stocks by turnover listed in HKEX (Hong Kong), chosen for their substantial market capitalization and their significant influence on their respective economies, providing a representative and diversified portfolio for analysis. A summary of the rolling window setup and average volatility is provided in Table 1, with asset-level details in Supplementary Tables S1–S3. While the proposed framework is readily extendable to higher-dimensional portfolios, we focus on 20 assets per market to maintain clarity in visualization and facilitate interpretable inference. The seven-year sample period offers a comprehensive setting to evaluate the procedure's adaptability across a range of economic environments, including episodes of structural shifts and increased market volatility.

4 Empirical Analysis

We apply the proposed stepwise procedure to the selected markets to evaluate its empirical performance. Using seven years of daily equity data, we compute optimal portfolio weights within rolling windows of 250 trading days for the U.S., 245 for India, and 244 for Hong Kong. The analysis is conducted under both daily and weekly rebalancing schemes, where portfolio weights are updated accordingly, every trading day in the daily scheme, and every five trading days in the weekly scheme. This setup enables us to assess the adaptability of the procedure across varying market conditions and rebalancing frequencies.

Multivariate normality of asset returns was evaluated using the energy test (Székely and Rizzo, 2005) implemented in the `assetsTest()` function of the `fPortfolio` package (Wuertz et al., 2023). The results uniformly rejected the null hypothesis across all markets, with p-values below the numerical precision threshold ($< 2.2 \times 10^{-16}$). The marginal distributions of the asset returns are modeled using the SGT-distribution, with parameters estimated via the `sgt.mle` function. Goodness-of-fit is assessed using the Anderson-Darling (AD) test, with all p-values exceeding 0.2, indicating adequate fit. Dependence structures across assets were modeled using the empirical beta copula (`empCopula`), which flexibly captures non-Gaussian dependencies and tail behavior. To simulate joint return scenarios from the fitted copula, we employed Monte Carlo sampling with 10^5 replications per market. Monte Carlo sample sizes were chosen to ensure target numerical accuracies of $\varepsilon = 8 \times 10^{-4},\ 7 \times 10^{-4}$, and 5×10^{-4} for the U.S., India, and Hong Kong, respectively. These targets control the simulation error in the empirical copula approximation. The required number of replications was determined using Proposition 1 in Lu and Ghosh (2023), which provides an upper bound on the approximation error based on the largest eigenvalue of the covariance matrix. Portfolio optimization was subsequently performed via quadratic programming using the `solve.QP` function.

The proposed framework, referred to as the *copula-based covariance model with three constraints*, is compared against three benchmark strategies: (i) the *mean–variance model* under a multivariate normality assumption using the sample covariance matrix, (ii) the *copula-based covariance model with two constraints* (restricted to full investment and long-only positions), and (iii) an *equally weighted naïve portfolio*. For clarity and reproducibility, these models are labeled in the figures by their implementation names: `copula_cov_3constraint`, `data_cov_3constraint`, `copula_cov_2constraint`, and `eq_weights`, respectively. All models are evaluated under both daily and weekly rebalancing schemes, and performance is assessed using the average return and average Sharpe ratio over time.

Figure 5 presents rolling weekly estimates of next-day mean returns and Sharpe ratios for the U.S. market from April 2019 to March 2025, with a focused view on the COVID-19 crisis period (March 2020–February 2022). Portfolios optimized using copula-based dependence structures with three constraints achieve relatively large Sharpe ratios, particularly during crisis period (e.g., mid-2020). Interestingly, in this sample, the sample covariance-based strategy performs comparably to the copula-based approach over extended periods. However, this alignment may be sample specific, as sample covariance matrices are known to be unstable under non-Gaussian return distributions or in high-dimensional settings (Ledoit and Wolf, 2004). By contrast, copula-based models explicitly account for tail risk and nonlinear dependence, offering greater flexibility in environments where normality assumptions do not hold.

The evolution of portfolio weights under the *copula-based covariance model with three constraints* exhibits distinct and interpretable shifts across three major market regimes: pre-pandemic, pandemic, and post-pandemic (Figure 6). The terms pre-pandemic, pandemic, and post-pandemic refer to the periods April 2019–February 2020, March 2020–February 2022, and March 2022–March 2025, respectively. During the pre-pandemic phase, the portfolio concentrated in large-cap defensive stocks such as PepsiCo, Procter & Gamble, and Johnson & Johnson, consistent with a stable, low-volatility environment. In the pandemic period, the model responded dynamically to elevated uncertainty and shifting return distributions by reallocating toward resilient consumer and healthcare firms, including Amazon, Netflix, Costco, and JnJ. This period is marked by high volatility, structural breaks, and regime instability, which the copula-based framework accommodates through flexible modeling of tail dependence. In the post-pandemic phase, the portfolio progressively rotated into growth and high-beta stocks, such as Tesla, NVIDIA, and Meta, particularly during the 2023 tech-led rally.

We replicate this analysis for the Indian and Hong Kong markets (Supplementary Figures S7–S10). In both India and Hong Kong, the *copula-based covariance model with three constraints* consistently outperforms equal-weighted and sample covariance benchmarks in terms of Sharpe ratio, particularly during the COVID-19 recovery period, where it delivers higher and more stable risk-adjusted returns. During the pre-pandemic phase, the Indian portfolio emphasized a mix of private financials and defensives, with top allocations to HDFC Bank, HUL,

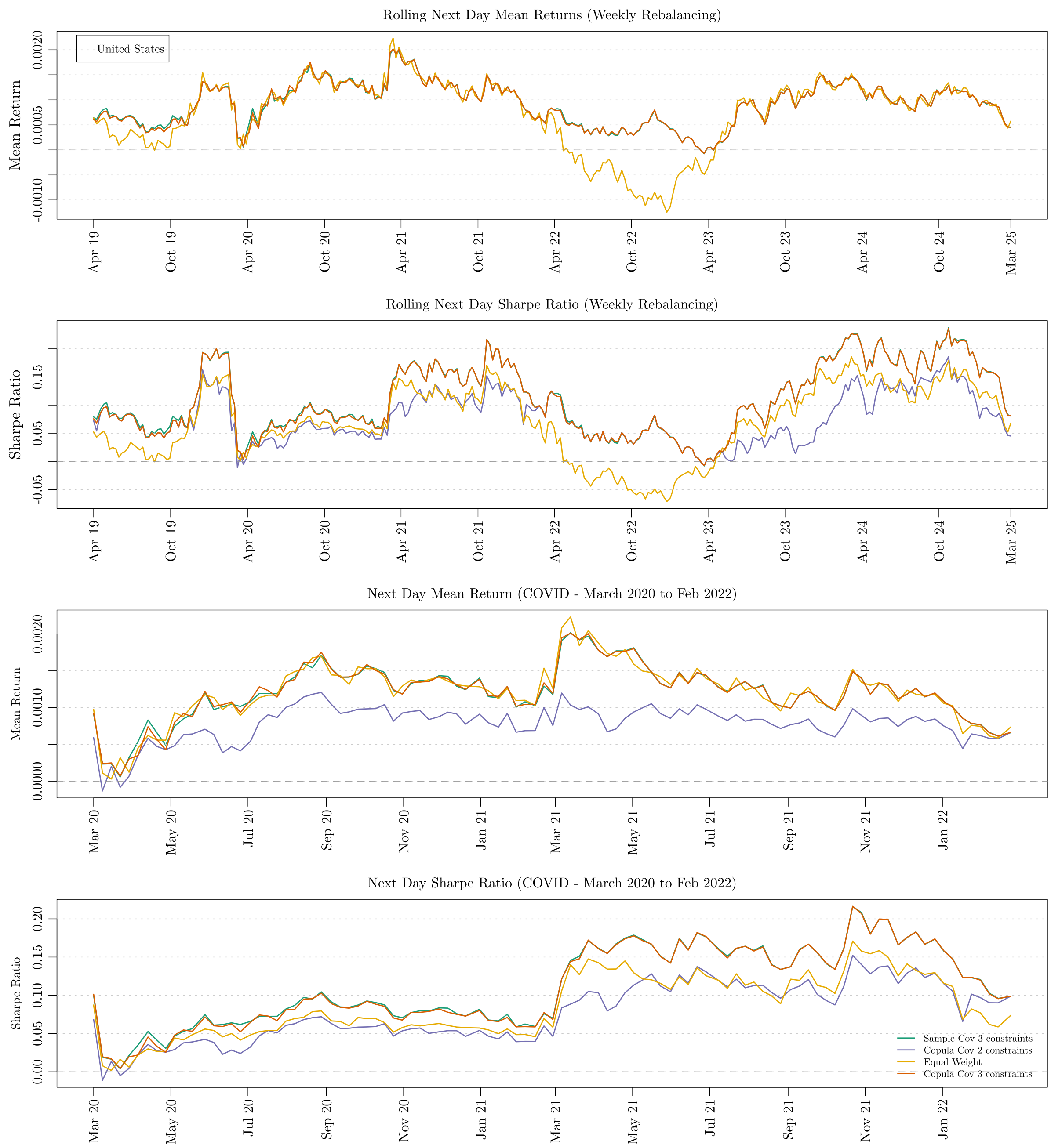

Figure 5: Rolling performance metrics for the U.S. market under weekly rebalancing (April 2019–March 2025) and during the COVID-19 period (March 2020–February 2022). The top two panels display rolling next-day mean returns and Sharpe ratios for the U.S. market based on portfolios constructed using equal weighting, sample covariance, and copula-based covariance matrices with two and three constraints. The bottom two panels focus on the COVID-19 crisis, highlighting sharper contrasts among strategies.

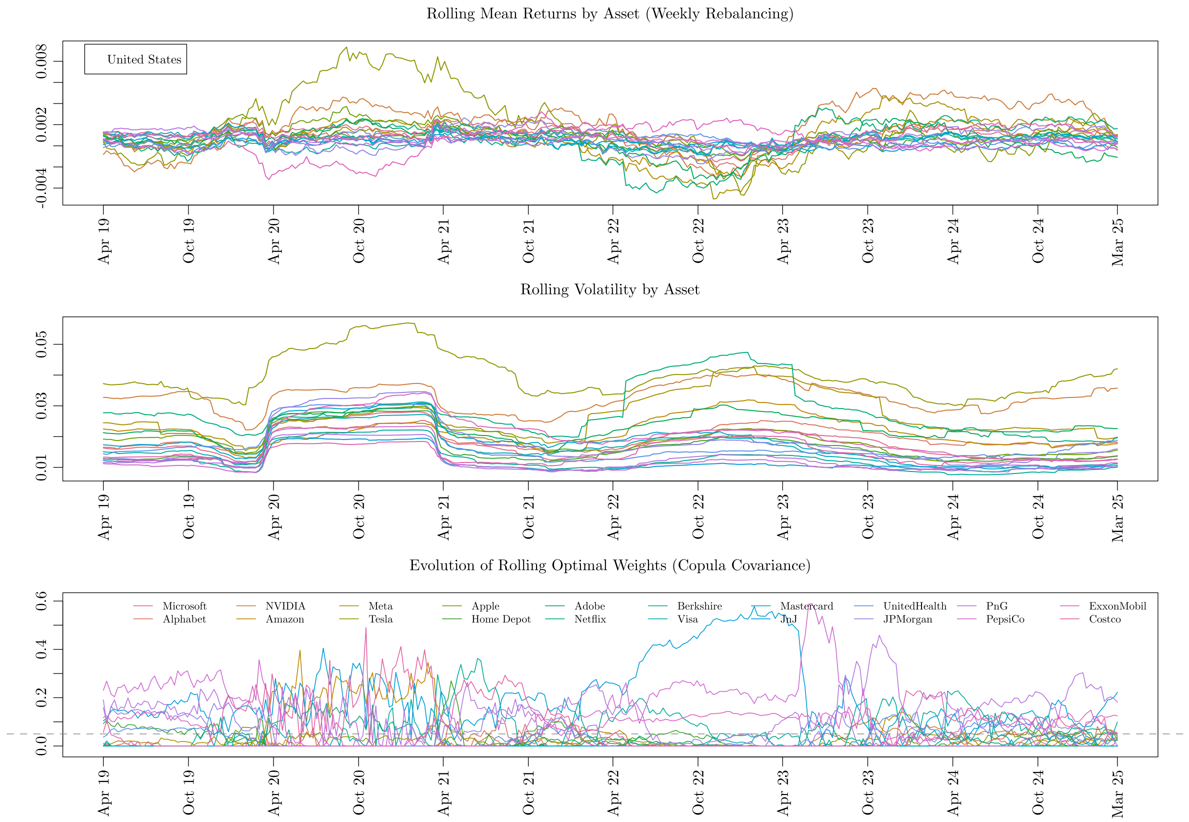

Figure 6: Rolling asset-level statistics and optimal weights for the U.S. market under weekly rebalancing (April 2019–March 2025). The top panels present rolling mean returns and volatilities by asset, revealing heterogeneous risk-return profiles and temporal shifts, particularly during high-volatility episodes (e.g., early 2020 and 2022). The bottom panel shows the evolution of optimal portfolio weights obtained via `copula_cov_3constraint`.

NTPC, and TCS, reflecting a stable, moderately growth-oriented stance. In the pandemic period, allocations shifted more conservatively toward utilities and non-cyclicals, led by NTPC, Asian Paints, HUL, and Sun Pharma, indicating a defensive realignment in response to heightened volatility and uncertainty. In the post-pandemic phase, the portfolio continued favoring defensive allocations: Sun Pharma, HUL, and ITC ranked highest, while partially rotating into telecom and IT names like Bharti Airtel and TCS. In the Hong Kong market, pre-pandemic phase, the portfolio was concentrated in defensives such as HKCG, HSBC, CK Infrastructure, and CLP. During the pandemic, exposure shifted further toward utilities and low-risk financials, led by CLP, HKCG, Ping An, and ICBC, reflecting heightened risk aversion. Post-pandemic, the portfolio partially rotated into telecom and diversified financials, with China Mobile, CLP, and HSBC receiving the largest weights. These dynamics reflect a cautious recovery posture and support the presence of a barbell strategy balancing stability with selective growth exposure. These findings persist under daily rebalancing schemes (Supplementary Figures S3–S4, S11–S12, S15–S16), confirming consistency of the framework in modeling time-varying risk-return trade-offs.

The time-averaged distributions of the rolling optimal weights and the corresponding risk contributions (Supplementary Figures S5, S13, S17) reveal clear market-specific patterns. The normalized risk contribution of asset i is given by: $RC_i = \frac{w_i \cdot (\Sigma w)_i}{w^\top \Sigma w}$, where w_i is the portfolio weight of asset i, Σ is the covariance matrix of asset returns, and $w^\top \Sigma w$ is the total portfolio variance. This formulation ensures that $\sum_{i=1}^{N} RC_i = 1$, allowing interpretation of RC_i as the percentage contribution of asset i to total portfolio risk. In the U.S., weights are broadly distributed, with defensives such as PepsiCo, Procter & Gamble, and Costco contributing consistently to total portfolio risk. In the Indian market, optimal portfolio allocations derived from copula-based covariance structures exhibit broader sectoral diversification, with noticeable weights in large-cap financials and defensives, such as TCS, HUL, ITC, Asian Paints, Sun Pharma, and NTPC. In Hong Kong, the portfolio is dominated by CLP, China Mobile, and HKCG, with these few assets accounting for a substantial share of total risk. These cross-market contrasts highlight how the copula framework dynamically adjusts exposure and risk budgeting in response to varying market structures and dependency patterns.

During the COVID-19 crisis period, we evaluate the relative performance of the *copula-based covariance model with three constraints* strategy against the sample covariance-based approach. Table 2 presents a comparative analysis of portfolio performance under copula-based and sample covariance-based optimization frameworks across the US, India, and Hong Kong markets. The copula-based method demonstrates superior performance in most scenarios, particularly under weekly rebalancing. Notably, the Hong Kong market exhibits the largest relative gains, with improvements of 28.9% in mean return and 5.1% in Sharpe ratio, highlighting the method's advantage in capturing asymmetric and tail-dependent relationships. Across all three markets, the copula approach achieved higher average Sharpe ratios and returns in a substantial proportion of rolling windows, for instance, outperforming in 54.5% (US), 69.7% (India), and 30.3% (Hong Kong) of windows in terms of return, and in 29.7%, 59.6%, and 26.3% of windows in terms of Sharpe ratio, respectively. In the Hong Kong market, although the number of outperformance windows is lower, relative gains in those periods are substantially greater, indicating that when a copula-based strategy outperforms, it does so by a significant margin. These results shows that copula-based allocations, especially during volatile periods such as COVID-19 crisis, by effectively modeling non-linear dependencies and extreme co-movements often missed by sample covariance-based models.

Portfolios were additionally estimated using Gaussian and t-copulas with SGT marginals to enable a direct comparison with those obtained from the nonparametric empirical β-copula. Figure 7 compares the distribution of rolling optimal portfolio weights under the t-, Gaussian-, and β-copulas with SGT marginals at the weekly rebalancing frequency for the U.S. market. Across assets, the t-copula generates the widest range of weights, reflecting stronger tail dependence and larger rebalancing adjustments. The Gaussian copula produces more concentrated allocations, while the β-copula lies in between, maintaining moderate dispersion and non-zero exposure for a broader set of assets. This balance suggests that the empirical β-copula preserves diversification without the extreme weight variability observed under the t-copula. Comparable patterns are observed for the Indian and Hong Kong markets (see Supplementary Figures S19–S20). As shown in the Supplementary Table S4, the empirical β-copula yields consistently higher Sharpe ratios than both Gaussian and t copulas during COVID, despite only minor return differences.

The robustness of portfolio allocations can also be evaluated through rebalancing costs, which offer a practical measure of portfolio stability. Although robustness in this study primarily refers to the flexibility of the copula specification, the rolling-window framework implicitly captures turnover associated with portfolio adjustments. Higher turnover indicates greater trading activity and transaction costs, whereas smoother weight dynamics suggest more stable allocations. In this context, the empirical β-copula achieves a balanced adjustment pattern, im-

Table 2: Portfolio performance: copula vs. sample covariance (weekly rebalancing with daily in parentheses).

Metric	US			India			Hong Kong		
	Copula	Sample	Gain (%)	Copula	Sample	Gain (%)	Copula	Sample	Gain (%)
Average return (%)	0.1204	0.1205	0.0 (−0.5)	0.0911	0.0894	1.9	0.0063	0.0049	28.9
	(0.1206)	(0.1212)		(0.0914)	(0.0893)	(2.4)	(0.0071)	(0.0059)	(20.1)
Average Sharpe ratio (%)	11.1307	11.2479	−1.0	9.4066	9.3228	0.9	1.8521	1.7627	5.1
	(11.1923)	(11.3327)	(−1.2)	(9.4184)	(9.3094)	(1.2)	(1.9273)	(1.8525)	(4.0)
No. of windows		101 (505)			99 (497)			99 (494)	
% Higher return windows		54.5 (49.3)			69.7 (64.6)			30.3 (70.1)	
% Higher Sharpe ratio windows		29.7 (15.8)			59.6 (50.9)			26.3 (23.1)	

Note: Weekly rebalancing results are shown first, with daily rebalancing values in parentheses. ‘Gain’ columns report the relative improvement of copula-based portfolios over sample covariance portfolios during the COVID-19 period (March 2020–February 2022).

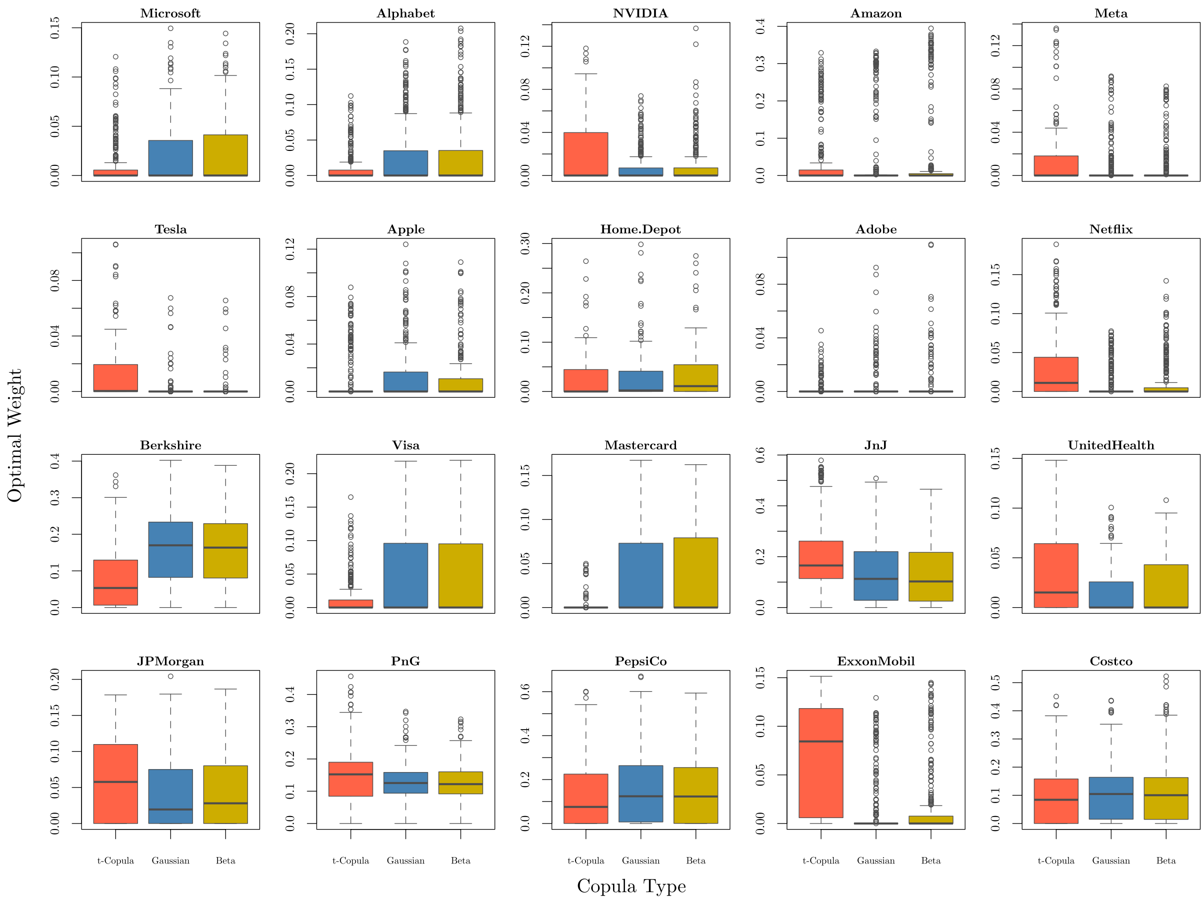

Figure 7: Distribution of rolling optimal portfolio weights under the t-, Gaussian-, and empirical β-copulas with SGT marginals at the weekly rebalancing frequency for the U.S. market. The t-copula produces the widest dispersion of weights, the Gaussian copula yields more concentrated allocations, and the empirical β-copula exhibits an intermediate pattern with broader participation but controlled variability.

plying lower effective rebalancing costs compared with the t- and Gaussian copulas. Hence, the framework can be naturally extended to incorporate cost-adjusted robustness as an additional evaluation criterion.

4.1 Skewed Generalized t-Distribution for Asset Returns

We employed the SGT-distribution to model individual asset returns, which were then input into a copula function to characterize their dependence structure. The SGT-distribution is well-suited to financial returns because it can flexibly capture skewness, heavy tails, and excess kurtosis, thus providing a suitable alternative to conventional symmetric distributions. It features distinct parameters for location (mean), scale (standard deviation), and shape (skewness and tail thickness).

Using the `sgt` package in `R` (Davis, 2015), the parameters of the SGT-distribution can be estimated by maximum likelihood using the function `sgt.mle(x, mu, sigma, lambda, p,` `q, mean.cent = TRUE, var.adj = TRUE)`. For our analysis, we initialized the parameter estimates with the sample mean (μ) and sample standard deviation (σ) of the data, along with small values for the shape parameters ($p = q = 2$) to capture the heavy tails that are typically observed in financial returns. The `mean.cent = TRUE` option ensures that μ corresponds to the actual mean of the SGT-distribution, while `var.adj = TRUE` adjusts σ to represent the standard deviation. Details on the functional form and derivations are in the `sgt` package vignette, which provides a detailed implementation guide. After estimating the parameters, we first plot the fitted density for rolling windows to visualize the evolution of the marginal distributions over time, capturing the dynamic nature of financial returns.

Figure 2 presents the SGT-fitted overlaid rolling density plots for the Netflix and Costco US market stocks. The variations in these plots show the dynamic nature of financial returns over time, potentially reflecting shifts in market conditions, investor sentiment, or external factors. Although densities exhibit changes over time, these changes are not overly drastic. The SGT family appears to effectively account for different distributional shapes characterized by high kurtosis and skewness. Similar behavior is observed in the Indian and Hong Kong markets.

To assess the goodness of fit of the SGT-distribution, we use the p-value from the AD test, which evaluates how closely the estimated SGT-distribution represents the observed stock returns. Since the true distribution of the data is unknown, we consider the fitted SGT-distribution as a proxy for the true distribution. A high p-value suggests that the observed stock returns align well with the theoretical SGT-distribution, validating its suitability for modeling financial return data. The P values were calculated in R using the `ADGofTest` package (Bellosta, 2011), which applies the AD test and an approximate p-value method from Marsaglia and Marsaglia (2004). All p-values exceed 0.2, providing strong evidence that the SGT-distribution effectively models the marginal densities of financial returns. Summary measures, including time-averaged p-values and estimated asymmetry parameters from the SGT fit, are presented as box plots in Supplementary Figures S6, S14, S18.

5 Discussion

The proposed semiparametric model which combines nonparametrically estimated copula with parametrically estimated marginals allows all parameters to dynamically evolve over time making it suitable for predictive inference. Applied to rolling windows of financial returns from the USA, India and Hong Kong economies, this approach addresses the limitations of traditional models that rely on parametric assumptions. The skewed generalized t- distribution captures skewness and kurtosis often observed in financial returns, while the empirical beta copula, as a nonparametric estimator, accommodates arbitrary and multivariate dependence structures. The copula-implied dependence structure enables adaptive, data-driven portfolio decisions across heterogeneous market environments. The model effectively tracks evolving inter-asset dependencies and regime shifts, translating structural changes into rebalanced allocations over time.

By allowing parameters to evolve over time, the proposed framework (*copula-based covariance model with three constraints*) captures dynamic dependencies linked to economic episodes, offering substantial improvements in modeling time-varying, nonlinear, and asymmetric relationships relative to other approaches. It demonstrates consistent outperformance during periods of financial stress, where classical models often fail to reflect joint tail risk and dynamic correlation shifts. Each rolling window, comprising marginal SGT estimation, empirical beta copula construction, and simulation-based optimization with $m = 10^5$ samples, completes in approximately 20–25 seconds on a standard Intel Core i7 (2.90 GHz, 16 GB RAM) machine. The method entails a modest computational burden but provides greater flexibility and reliability in crisis regimes. Its modular design is readily parallelizable, supporting scalable implementation in high-frequency or large-universe settings. From the market analyses, the model performed reasonably well across all three markets by identifying dynamically evolving weights that optimize portfolio returns. The results indicate that optimally weighted portfolios from *copula-based covariance model with three constraints* consistently outperform equal-weighted portfolios, with higher weights allocated to high-return assets, reflecting their importance in the portfolio. Future work will involve exploring other risk measures (e.g., VaR and CVaR) using the proposed semiparametric dynamic models. The R code supporting the results of this study will be made publicly available on the first author's GitHub repository upon acceptance of the paper.

Disclosure Statement: No potential conflict of interest was reported by the author(s).

Supplementary Material

Supplementary Figures and Tables & Software Implementation Details.

Funding

No funding was received.

References

Avella A (2024). Empirical Analysis of Portfolio Optimization Using the Markowitz Model. Available at SSRN.

Azzalini A, Capitanio A (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society, Series B, Statistical Methodology*, 65(2): 367–389. https://doi.org/10.1111/1467-9868.00391

Bacon C (2008). *Practical Portfolio Performance Measurement and Attribution*. John Wiley & Sons, Hoboken, NJ.

Bali TG, Theodossiou P (2007). A conditional-SGT-VaR approach with alternative GARCH models. *Annals of Operations Research*, 151: 241–267. https://doi.org/10.1007/s10479-006-0118-4

Bellosta CJG (2011). *ADGofTest: Anderson-Darling GoF test*. R package version 0.3.

Chen SX, Huang TM (2007). Nonparametric estimation of copula functions for dependence modelling. *Canadian Journal of Statistics*, 35(2): 265–282. https://doi.org/10.1002/cjs.5550350205

Chen X, Fan Y (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics*, 135(1–2): 125–154. https://doi.org/10.1016/j.jeconom.2005.07.027

Dahlhaus R (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics*, 25(1): 1–37.

Davis C (2015). *sgt: Skewed Generalized T Distribution Tree*. R package version 2.0.

Deheuvels P (1979). La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. *Bulletin de l'Académie Royale de Belgique*, 65: 274–292.

Dou X, Kuriki S, Lin GD, Richards D (2025). EM estimation of the B-spline copula with penalized pseudo-likelihood functions. *Statistical Papers*, 66(1): 30. https://doi.org/10.1007/s00362-024-01647-w

Fermanian JD, Scaillet O (2002). Nonparametric Estimation of Copulas for Time Series. *Technical Report 57*, FAME – International Center for Financial Asset Management and Engineering, Geneva, Switzerland.

Fernández C, Steel MFJ (1995). A family of skew distributions with kurtosis and skewness parameters. *Journal of the American Statistical Association*, 90(431): 1331–1341.

Genest C, Nešlehová J, Quessy JF, Remillard B (2017). Asymptotic behavior of the empirical multilinear copula process under broad conditions. *Journal of Multivariate Analysis*, 159: 82–110. https://doi.org/10.1016/j.jmva.2017.04.002

Gunjan A, Bhattacharyya S (2023). A brief review of portfolio optimization techniques. *Artificial Intelligence Review*, 56: 3847–3886. https://doi.org/10.1007/s10462-022-10273-7

Hofert M, Kojadinovic I, Maechler M, Yan J (2024). *copula: Multivariate Dependence with Copulas*. R package version 1.1-4.

Huang CW, Hsu CP, Chiou WJP (2015). *Can Time-Varying Copulas Improve the Mean-Variance Portfolio?* Springer, New York.

Jaworski P, Durante F, Härdle WK, Rychlik T (2010). *Copula Theory and Its Applications*, volume 198 of *Lecture Notes in Statistics*. Springer, Berlin/Heidelberg, Germany.

Joe H (2014). *Dependence Modeling with Copulas*. Chapman and Hall/CRC, New York, NY, USA.

Kolm PN, Tütüncü R, Fabozzi FJ (2014). 60 years of portfolio optimization: practical challenges and current trends. *European Journal of Operational Research*, 234(2): 356–371. https://doi.org/10.1016/j.ejor.2013.10.060

Ledoit O, Wolf M (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2): 365–411. https://doi.org/10.1016/S0047-259X(03)00096-4

Lu L, Ghosh S (2023). Nonparametric estimation of multivariate copula using empirical Bayes methods. *Mathematics*, 11(20): 4383. https://doi.org/10.3390/math11204383

Markowitz H (1952). Portfolio selection. *The Journal of Finance*, 7(1): 77–91. https://doi.org/10.2307/2975974

Marsaglia G, Marsaglia J (2004). Evaluating the Anderson–Darling distribution. *Journal of Statistical Software*, 9(2): 1–5. https://doi.org/10.18637/jss.v009.i02

McNeil AJ, Nešlehová J (2009). Multivariate Archimedean copulas, d-monotone functions, and l_1-norm symmetric distributions. *The Annals of Statistics*, 37: 3059–3097.

Nelsen RB (2007). *An Introduction to Copulas*, 2nd edition, *Springer Series in Statistics*. Springer, Berlin/Heidelberg, Germany.

Palomar DP (2025). *Portfolio Optimization: Theory and Application*. Cambridge University Press, Cambridge, UK.

Patton AJ (2012). A review of copula models for economic time series. *Journal of Multivariate Analysis*, 110: 4–18. https://doi.org/10.1016/j.jmva.2012.02.021

Politis DN (2011). Higher-order accurate, positive semidefinite estimation of large-sample covariance and spectral density matrices. *Econometric Theory*, 27: 703–744. https://doi.org/10.1017/S0266466610000484

Ryan JA, Ulrich JM (2023). *quantmod: Quantitative Financial Modelling Framework*. R package version 0.4.26.

Sahamkhadam M, Stephan A (2023). Portfolio optimization based on forecasting models using vine copulas: An empirical assessment for global financial crises. *Journal of Forecasting*, 42(8). https://doi.org/10.1002/for.3009

Sancetta A, Satchell S (2004). The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory*, 20(3): 535–562. https://doi.org/10.1017/S026646660420305X

Segers J, Sibuya M, Tsukahara H (2017). The empirical beta copula. *Journal of Multivariate Analysis*, 155: 35–51. https://doi.org/10.1016/j.jmva.2016.11.010

Sheather SJ (2004). Density estimation. *Statistical Science*, 19(4): 588–597. https://doi.org/10.1214/088342304000000297

Shen X, Zhu Y, Song L (2008). Linear B-spline copulas with applications to nonparametric estimation of copulas. *Computational Statistics & Data Analysis*, 52(7): 3806–3819. https://doi.org/10.1016/j.csda.2008.01.002

Silverman BW (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall.

Sklar M (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de L'Institut de Statistique de L'Université de Paris*, 8: 229–231.

Székely GJ, Rizzo ML (2005). A new test for multivariate normality. *Journal of Multivariate Analysis*, 93: 58–80. https://doi.org/10.1016/j.jmva.2003.12.002

Theodossiou P (1998). Financial data and the skewed generalized t distribution. *Management Science*, 44(12): 1650–1661. https://doi.org/10.1287/mnsc.44.12.1650

Theodossiou P, Savva CS (2016). Skewness and the relation between risk and return. *Management Science*, 62(6): 1598–1609. https://doi.org/10.1287/mnsc.2015.2201

Turlach BA, Weingessel A (2019). *quadprog: Functions to Solve Quadratic Programming Problems*. R package version 1.5-8.

Valeyre S (2023). Optimal Trend Following Portfolios. Available at SSRN.

Wuertz D, Setz T, Chalabi Y, Chen W, Theussl S (2023). *fPortfolio: Rmetrics – Portfolio Selection and Optimization*. R package version 3042.89.