Integrating Choquet Portfolios and Machine Learning Interpretability for Robust Cryptocurrency Investment Strategies

João Pedro M. Franco[a,*], Márcio P. Laurini[b]

[a]*Department of Economics - FEARP USP*
[b]*Department of Economics - FEARP USP*

Abstract

This study proposes an alternative approach to portfolio optimization in the cryptocurrency market by applying the Choquet integral portfolio, positioning it within the broader context of Robo-Advisor literature. The Choquet integral formulation offers a flexible framework for risk management, prioritizing subsets of assets that minimize significant losses, especially when assessed using metrics like Conditional Value-at-Risk (CVaR). The performance of the Choquet-based portfolio is compared to traditional portfolio construction methods, including mean-variance optimization, naive diversification, and a BTC-only portfolio. Empirical results show that the Choquet portfolio, which accounts for higher-order moments and captures asset interdependencies, outperforms its counterparts in terms of cumulative returns and tail risk mitigation. Shapley Values are used to analyze the temporal evolution of asset weight allocations in both the Markowitz and Choquet portfolios. Furthermore, the study applies Local Interpretable Model-Agnostic Explanations (LIME) to identify the cryptocurrencies—specifically BTC, BNB, and XRP—that have the greatest impact on portfolio construction.

Keywords: Choquet Portfolio, Cryptocurrency Risk Management, Tail Risk, Shapley Values, Local Interpretable Model-Agnostic Explanations.

⋆Declarations of interest: none

⋆⋆The authors acknowledge funding from Capes, CNPq (310646/2021-9) and FAPESP (2023/02538-0). Declarations of interest: none

*Corresponding author - FEARP-USP - Av. dos Bandeirantes 3900, 14040-905, Ribeirão Preto, SP, Brazil. Tel.: +55-16-33290867 - email - laurini@fearp.usp.br

1. Introduction

The rapid rise of cryptocurrencies has attracted significant investment capital, driven by exponential growth in market capitalization, trading volumes, and the prices of individual assets. This remarkable expansion underscores the potential for swift and substantial changes in the cryptocurrency market. Since its introduction in 2008, Bitcoin has maintained its position as the cryptocurrency with the largest market capitalization (Anyfantaki et al., 2018). As highlighted by Bowala and Singh (2022), the cryptocurrency market is characterized by extreme volatility, a lack of regulatory oversight, and significant challenges in constructing optimal portfolios. Despite the application of advanced portfolio optimization models, the resultant portfolios often remain highly risky, reflecting the inherent unpredictability and speculative nature of the market.

In the context of cryptocurrency markets, studies such as Platanakis et al. (2018) and Brauneis and Mestel (2019) have explored portfolio optimization using a mean-variance framework. For instance, Brauneis and Mestel (2019) found that a naive diversification strategy outperformed more than 75% of mean-variance optimal portfolios in terms of Sharpe ratio, while Platanakis et al. (2018) reported a small and statistically insignificant difference between naive and mean-variance portfolios. Similarly, Arzova and Özdurak (2021) applied the model proposed by Markowitz (1952) to portfolios comprising Bitcoin, Ethereum, Ripple, and the BIST 30 benchmark.

While the mean-variance model is widely used for portfolio optimization due to its simplicity (Xu et al., 2016), it has well-known limitations. These include the reliance on variance as a risk measure, computational and statistical challenges in large portfolios, and the assumption of Gaussian-distributed returns, which does not align with the observed distribution of financial assets (Areewong et al., 2009; De Athayde and Flôres Jr, 2004). The non-normality of cryptocurrency returns has been documented extensively, with Aljinović et al. (2021) highlighting the need for alternative risk measures and models. Addressing the issue of non-normality, Hrytsiuk et al. (2019) demonstrated that Bitcoin, Bitcoin Cash, Litecoin, XRP, Ethereum, and NEM follow a Cauchy distribution with reasonable precision. This finding supports the use of Value-at-Risk (VaR) methods for risk measurement but does not overcome the shortcomings of the mean-variance framework.

To expand beyond mean return and variance, Aljinović et al. (2021) proposed a multi-criteria model that incorporates daily returns, standard deviation, VaR, Conditional Value-at-Risk (CVaR), volume, market capitalization, and asset attractiveness. This approach outperformed naive diversification, mean-variance optimization, maximum Sharpe ratio, and the mean-CVaR efficient frontier. Similarly, Bowala and Singh (2022) included skewness and kurtosis in the portfolio optimization problem, demonstrating superior performance over standard deviation-based approaches.

To address estimation challenges and incorporate higher-order moments, Bassett et al. (2004) introduced a minimum CVaR-based portfolio selection approach grounded in pessimistic decision-making functions. This method, by using the Choquet expected utility theory (Gilboa, 2009), minimizes Expected Shortfall using a concave distortion function ν and a linear utility function, offering a coherent risk measure as defined by Artzner et al.

(1999).

The Choquet integral, as explored by Xu et al. (2016), enables a pessimistic decision criterion by assigning greater weight to unfavorable events. By utilizing a quantile regression framework, this model simplifies numerical optimization and focuses on tail behaviors rather than mean outcomes. Laurini (2007) demonstrated the effectiveness of this approach in outperforming the Ibovespa index and mean-variance portfolios in terms of risk measurement criteria.

Despite its theoretical and empirical advantages (Laurini, 2007; Xu et al., 2016), the Choquet portfolio model has not yet been applied to cryptocurrency markets. The use of Choquet portfolios in the cryptocurrency market is particularly justified by the unique characteristics of this asset class, including large variations in returns, extreme losses, and heightened tail risk. Cryptocurrencies exhibit pronounced volatility and frequent deviations from normality in return distributions, making traditional portfolio optimization models, such as the mean-variance framework, less effective and potentially misleading. The Choquet portfolio model, by incorporating a pessimistic decision criterion, allows for a more robust handling of these variations by placing greater emphasis on unfavorable scenarios, which are prevalent in the cryptocurrency market. This paper addresses this gap by constructing portfolios for leading cryptocurrencies by market capitalization—BTC, ETH, BNB, SOL, XRP, ADA, and LTC (data sourced from `coinmarketcap.com`).

We extend the literature by comparing the Markowitz mean-variance model and the Choquet portfolio, utilizing Machine Learning interpretability tools to explain differences in portfolio weights and returns. Specifically, we apply Shapley Values (Shapley, 1953) and Local Interpretable Model-Agnostic Explanations (LIME; Ribeiro et al., 2016) to enhance model transparency.

Several studies have highlighted the rapid growth of machine learning models and their diverse applications in the financial market (Leow et al., 2021; Phoon and Koh, 2018; Brigo et al., 2021; Leo et al., 2019; Rezaei et al., 2021). A notable area within this literature, as emphasized by Phoon and Koh (2018), pertains to Robo-Advisors (RAs). As defined by Park et al. (2016), Robo-Advisors are artificial intelligence systems that leverage large-scale data and algorithmic decision-making to provide trading and investment advisory services. These systems are designed to assist in security trading and offer advisory services to investors. The increasing adoption of RAs has been well-documented in recent literature (Brenner and Meyll, 2020; Lam, 2016; Rossi and Utkus, 2020), with key factors driving their popularity including the mitigation of conflicts of interest (Leow et al., 2021), low operational costs (Brenner and Meyll, 2020), and the provision of automated financial advice (Rossi and Utkus, 2020).

As noted by Rezaei et al. (2021), automated portfolio allocation powered by artificial intelligence is proving increasingly valuable, with demand for such methods continuing to grow (Park et al., 2016). Portfolio asset allocation occupies a central role in economic decision theory, influencing both the financial sector and policymakers (Babaei et al., 2022; Bassett et al., 2004). Consequently, testing innovative portfolio optimization methods in different markets, while explaining their underlying mechanisms, is of critical importance.

We construct three classes of portfolios with the most influential cryptocurrencies identified by interpretability methods: naive, Markowitz, and Choquet portfolios. Results indicate that the Choquet portfolio consistently achieves higher cumulative returns, reduced tail risk, and superior replicability. By addressing the limitations of classical models and incorporating measures that account for the inherent risks in cryptocurrencies, the Choquet portfolio provides a more comprehensive and tailored approach to portfolio optimization in this domain. Its ability to minimize tail risks and focus on extreme losses aligns well with the risk management needs of investors in this high-volatility market. Additionally, when combined with explainable machine learning methods, such as Shapley Values and LIME, the Choquet portfolio can offer insights into the drivers of risk and return, fostering greater transparency and trust in portfolio decisions.

This paper is structured as follows: Section 2 outlines the mean-variance and Choquet models; Section 3 describes the dataset; Section 4 presents the empirical findings; Sections 5 and 6 detail the interpretability models and their results, respectively; and Section 7 concludes.

2. Portfolio Models

In this section, we provide a concise overview of the Markowitz model, widely regarded as the most classic and commonly used framework for portfolio optimization, and introduce the model proposed by Bassett et al. (2004), which is grounded in pessimistic functions and leverages a quantile regression framework.

For portfolio construction, we adopt the approach outlined in Babaei et al. (2022), assuming the presence of a Robo-advisor (RA) that updates portfolio weights daily based on a rolling h-day time window. It is important to emphasize that the covariance matrix for each portfolio is recalculated daily, incorporating data from the h days preceding the portfolio's rebalancing. This dynamic updating process ensures that the portfolio remains responsive to recent market conditions.

2.1. Markowitz Model

The mean-variance model, which was the seminal model in modern portfolio theory, was introduced by Markowitz (1952). The main objective of this method is to find a set of portfolio weights that, while giving the minimum variance, also gives the targeted expected return; in other words, there is a trade-off between expected return and variance that should be taken into account. Following Constantinides and Malliaris (1995), the investor problem can be formulated as

$$\min_{\mathbf{x}} \quad \sigma^2_{p,t+1} = \mathbf{x}^T \mathbf{V} \mathbf{x} \tag{1}$$

$$\text{s. t.} \quad \mathbf{x}^T \mathbf{1} = 1 \tag{2}$$

$$\mathbf{x}^T \mathbf{R} = R_{p,t+1} \tag{3}$$

where $\sigma^2_{p,t+1}$ denotes the portfolio variance, x is a vector that denotes the weight of the wealth allocated to each asset i and $\mathbf{V}$ denote the covariance matrix, which must be non-singular, symmetric, and positive definite. Therefore, the portfolio variance will be minimized with two restrictions: the weights must have sum one and, also, the portfolio return, $x^T\mathbf{R}$, must be equal to an expected return ($R_{p,t+1}$).

Ravipati (2012) points out that, for mean-variance model works properly, some assumptions must be done for investors and about the market. First, the investors must have the same expected single period investment horizon, aim to maximize the expected return, only accept greater risk if they are compensated with a higher expected return (risk-adverse), and take investment decisions based only on expected return and risk, which in this model is measure using the standard deviation. Also, Ravipati (2012) highlights that markets must be perfectly efficient, in the sense that taxes and transactions costs are not considerate in this model.

Constantinides and Malliaris (1995) point out that the model uses the utility function that indirectly represents the investor's preferences in the problem. The assumption is the existence of a utility function which is defined on the basis of the mean and variance of the portfolio's return, whose investor prefers a higher mean and a lower variance. As we can note, the Markowitz model uses only the two first moments (mean and variance) as parameters for estimated the portfolio weights, leaving aside the rest of the information regarding the distribution. As Laurini (2007) point out, this process can be problematic, in a sense that the minimization led to the same values for the weights if distributions with different tail risks have the same vector of means and same variance-covariance matrix. Therefore, the importance of estimating the portfolio allocation problem using all possible distribution information, i.e., superior moments, has already been highlighted in the literature (Areewong et al., 2009; De Athayde and Flôres Jr, 2004; Harvey et al., 2010; Krokhmal, 2007).

2.2. Choquet portfolio construction

In order to solve the estimation problems of the classic portfolio model and also use higher order moments to calculate the optimal weights for a cryptocurrency portfolio, in this section we will describe the model proposed by Bassett et al. (2004), known as the Choquet portfolio. Unlike the classic mean-variance model, which uses least squares solutions, the Choquet portfolio calculates the optimal weights by solving a quantile regression problem that allows the risk associated with a specific α-quantile to be minimized.

For compare the expected utilities, Bassett et al. (2004) consider two random variables, X and Y, which have, respectively, F and G as distribution functions. Therefore, X will be preferred to Y if

$$\mathbb{E}_F u(X) = \int_{-\infty}^{\infty} u(x)\, dF(x) > \int_{-\infty}^{\infty} u(x)\, dG(x) = \mathbb{E}_G u(Y). \tag{4}$$

Equation (4) is reformulated by Bassett et al. (2004), in the sense that the expected utilities depend on the quantile functions, $F^{-1}(t)$ and $G^{-1}(t)$, for the respective random variables X and Y. This modification is illustrated in the following equation.

$$\mathbb{E}_F u(X) = \int_0^1 u(F^{-1}(t))\,dt > \int_0^1 u(G^{-1}(t))\,dt = \mathbb{E}_G u(Y), \tag{5}$$

where the events are now computed based on a uniform distribution within the range $[0, 1]$.

As highlighted by Bassett et al. (2004), Choquet's theory of expectations is incorporated into the problem through a distortion function, ν, which assigns a differentiated probability to each possible event realization. This approach transforms the expected utility derived from the realizations of the variables F and G, as further emphasized by Laurini (2007).

Thus, the preferences in Equation (5) can be represented by a pair of functions denoted by (μ, ν), where u transforms the outcomes in terms of utility and the distortion function, ν, transforms the probabilities, in the sense that they are inflated or deflated, according to the ranking order of the outcomes. Now, according to Bassett et al. (2004), X will be preferred to Y if

$$\mathbb{E}_{\nu,F} u(X) = \int_0^1 u(F^{-1}(t))\,d\nu(t) > \int_0^1 u(G^{-1}(t))\,d\nu(t) = \mathbb{E}_{\nu,G} u(Y). \tag{6}$$

It is important to note that it is these distortion functions that allow the introduction of pessimistic functions, which occur when, as pointed out by Bassett et al. (2004), instead of giving uniform weighting in the expected utility criterion, more weight is given to less favorable events (extreme negative returns) and less weight is given to more favorable occurrences (positive returns).

For inflate the probabilities of α least favorable events and discount totally the $1 - \alpha$ most favorable events, Bassett et al. (2004) choose $\nu_\alpha(t) = \min\{t/\alpha, 1\}$, where $\alpha \in [0, 1]$, which is a simple one parameter distortion function. Therefore, Equation (6) with pessimistic functions is given by

$$\mathbb{E}_{\nu_\alpha} u(X) = \alpha^{-1} \int_0^\alpha u(F^{-1}(t))\,dt. \tag{7}$$

Bassett et al. (2004) note that Equation (7) can be simply interpreted as the negative of a coherent risk measure, in terms describe by Artzner et al. (1999), known as Expected Shortfall. The following equation gives this formulation.

$$\varrho_{\nu_\alpha}(X) = -\int_0^1 F^{-1}(t)\,d\nu(t) = -\alpha^{-1} \int_0^\alpha F^{-1}(t)\,dt. \tag{8}$$

Therefore, as Bassett et al. (2004) suggests, the minimization of the Expected Shortfall subject to an expected return becomes a natural substitute for the standard deviation in the estimation of weights in the optimal portfolio problem.

Bassett et al. (2004) define a risk measure, ϱ, as pessimistic if, for some probability measure $\varphi \in [0,1]$ we have the following form

$$\varrho(X) = \int_0^1 \varrho_{\nu_\alpha}(X)\, d\varphi(\alpha). \tag{9}$$

Substituting Equation (8) and using the Fubini Theorem, we have that

$$\varrho(X) = -\int_0^1 \alpha^{-1} \int_0^\alpha F^{-1}(t)\, dt\, d\varphi(\alpha) = -\int_0^\alpha F^{-1}(t) \int_0^1 \alpha^{-1}\, d\varphi(\alpha)\, dt. \tag{10}$$

In this way, according to Bassett et al. (2004), φ can be taken as a finite sum of Dirac point masses, which are denote as $d\varphi = \sum_{i=0}^{m} \varphi_i \delta_{\tau_i}$ where $\varphi_i \geq 0$, $\sum \varphi_i = 1$ and $0 = \tau_0 < \tau_1 < \cdots < \tau_m \leq 1$. Therefore, using $\int_t^1 \alpha^{-1}\delta_\tau(\alpha)\, d\alpha = \tau^{-1} I(t < \tau)$, we can rewrite equation (10) as follows

$$\varrho(X) = \int_0^1 \varrho_{\nu_\alpha}(X)\, d\varphi(\alpha) = -\varphi_0 F^{-1}(0) - \int_0^1 F^{-1}(t)\gamma(t)\, dt \tag{11}$$

where $\gamma(t) = \sum_{i=1}^{m} \varphi_i \tau_i^{-1} I(t < \tau_i)$. Note that the concave distortion function, ν, can be approximated using a linear concave function, which here is the weighted sums of Dirac. It is important to note that, with this formulation, the probability in least favorable events is accentuated, while the probability of most favorable events is depressed.

2.2.1. Pessimistic Allocations and Quantile Regressions

Bassett et al. (2004) suggests quantile regressions[1] to solve the optimal portfolio model describe previously. The problem can be formulated as follows

$$\min_{\xi \in \mathbb{R}} E\rho_\alpha(X - \xi), \tag{12}$$

where the quantile loss function is define as $\rho_\alpha(u) = u(\alpha - I(u < 0))$ and the solution, given by the value of ξ that solve this problem, is the α quantile of the random variable X. Bassett et al. (2004) argues that find the α quantile is similar to compute the sum of expected return and the α risk (Expected Shortfall) of X multiplied by α. Therefore, the problem present by equation (12) can be rewrite as

$$\min_{\xi \in \mathbb{R}} E\rho_\alpha(X - \xi) = \alpha(\mu + \varrho_{\nu_\alpha}(X)). \tag{13}$$

[1] For details on quantile regression methods, we point to Koenker and Bassett Jr (1978), Koenker (2005) or Koenker et al. (2017).

Let $\{x_i : i = 1, ..., n\}$ denote a random sample on X, Bassett et al. (2004) formulated an empirical analogue of the α-risk (Expected Shortfall) as follows

$$\hat{\varrho}_{\nu_\alpha}(x) = (n\alpha)^{-1} \min_{\xi \in \mathbb{R}} \sum_{i=1}^{n} \rho_\alpha(x_i - \xi) - \hat{\mu}_n \tag{14}$$

where $\hat{\mu}_n$ denote the sample mean $\bar{x}_n$, which is used as estimator of $E[X] = \mu$.

Due to the nature of the problem and the fact that the solution is given by minimizing the risk subject to a restriction on the mean, Bassett et al. (2004) points out that a Lagrangian can be used to formulate the problem as follows

$$\min_{\pi} \varrho_{\nu_\alpha}(X^T\pi) - \lambda\mu(X^T\pi), \tag{15}$$

where $X^T\pi$ denote a portfolio of assets with vector of weights π. Imposing the constrain that portfolio weights, π, must sum to one we have that

$$\begin{aligned} \min_{\pi} \quad & \varrho_{\nu_\alpha}(X^T\pi) && (16) \\ \text{s. t.} \quad & \mu(X^T\pi) = \mu_0 && (17) \\ & 1^T\pi = 1. && (18) \end{aligned}$$

Following Bassett et al. (2004) and considering the first asset as numeraire, the final problem can be written as follows.

$$\begin{aligned} \min_{(\beta,\xi)\in\mathbb{R}^p} \quad & \sum_{i=1}^{n} \rho_\alpha(x_{i1} - \sum_{j=2}^{p}(x_{i1} - x_{ij})\beta_j - \xi) && (19) \\ \text{s. t.} \quad & \bar{x}^T\pi(\beta) = \mu_0, && (20) \end{aligned}$$

where $\pi(\beta) = (1 - \sum_{j=2}^{p} \beta_j, \beta^T)^T$. Bassett et al. (2004) highlight that any concave distortion function can be approximated by a piecewise linear function, therefore, a general formulation can be written for Equation (19) using the following weighted averages of α-risks,

$$\varrho_\nu(X) = \sum_{k=1}^{m} \nu_k \varrho_{\nu_{\alpha_k}}(X), \tag{21}$$

where $\nu_k > 0$ and $\sum_{k=1}^{m} \nu_k = 1$. Therefore, the general problem of pessimistic preferences is given by the following equation

$$\begin{aligned} \min_{(\beta,\xi)\in\mathbb{R}^{p+m}} \quad & \sum_{k=1}^{m}\sum_{i=1}^{n} \nu_k\rho_\alpha(x_{i1} - \sum_{j=2}^{p}(x_{i1} - x_{ij})\beta_j - \xi_k) && (22) \\ \text{s. t.} \quad & \bar{x}^T\pi(\beta) = \mu_0. && (23) \end{aligned}$$

3. Data

In order to compare these two portfolio methods in the cryptocurrency market, we choose a daily dataset[2] starting on 2020-04-11 and ending on 2024-04-07, of the following cryptocurrencies: Bitcoin (BTC), Ethereum (ETH), Binance Coin (BNB), Solana (SOL), Ripple (XRP), Cardano (ADA) and Litecoin (LTC).

The study period is selected to capture a comprehensive view of the cryptocurrency market, including its recovery after COVID-19 and subsequent growth. This timeframe encompasses various market conditions, such as bull runs (e.g., late 2020 and early 2021) and corrections (e.g., 2022–2023), which are crucial for assessing portfolio strategies under different regimes. The chosen period also includes extreme events and significant price fluctuations, allowing the evaluation of portfolio methods during the high-growth and contraction phases. In addition, the data for the selected cryptocurrencies (Bitcoin, Ethereum, Binance Coin, Solana, Ripple, Cardano, and Litecoin) are reliable and consistent from platforms like CoinMarketCap, making this period relevant for understanding portfolio performance in the evolving cryptocurrency market.

Figure 1, which shows the daily evolution of returns, allows us to highlight the characteristics of the cryptocurrency market, such as the substantial variation in returns and the existence of extreme values in the distribution. The characteristics of kurtosis and skewness, which are higher moments of the distribution of returns, can be seen in Table 1 and also in Figure 2, which plots the density of the returns of each cryptocurrency and compares them with a normal distribution.

In Table 1 and Figure 2, we observe evidence of skewness and excess kurtosis relative to the normal distribution. To formally assess these characteristics, we applied the Jarque-Bera test. The results indicate that, for all cryptocurrencies, the null hypothesis of normality is rejected. This implies that higher moments beyond the mean and variance deviate from those of a Gaussian distribution, confirming the observations in Table 1 and Figure 2.

4. Results

We estimate optimal portfolios using rolling windows of 30, 60, 180, and 365 days to account for time-varying dynamics in asset returns and risk. For each rolling window, the optimal portfolio is calculated using four distinct methodologies:

- **Naive Portfolio**: An equal-weight portfolio where all assets receive the same allocation, providing a simple benchmark for comparison.
- **Markowitz Portfolio**: Based on the mean-variance optimization framework, this approach seeks to maximize the Sharpe ratio by considering the trade-off between expected returns and portfolio risk (variance).

[2]Data on cryptocurrency returns was sourced from coinmarketcap.com

Figure 1: Evolution of daily returns of the analyzed assets (2020-04-11 - 2024-04-07).

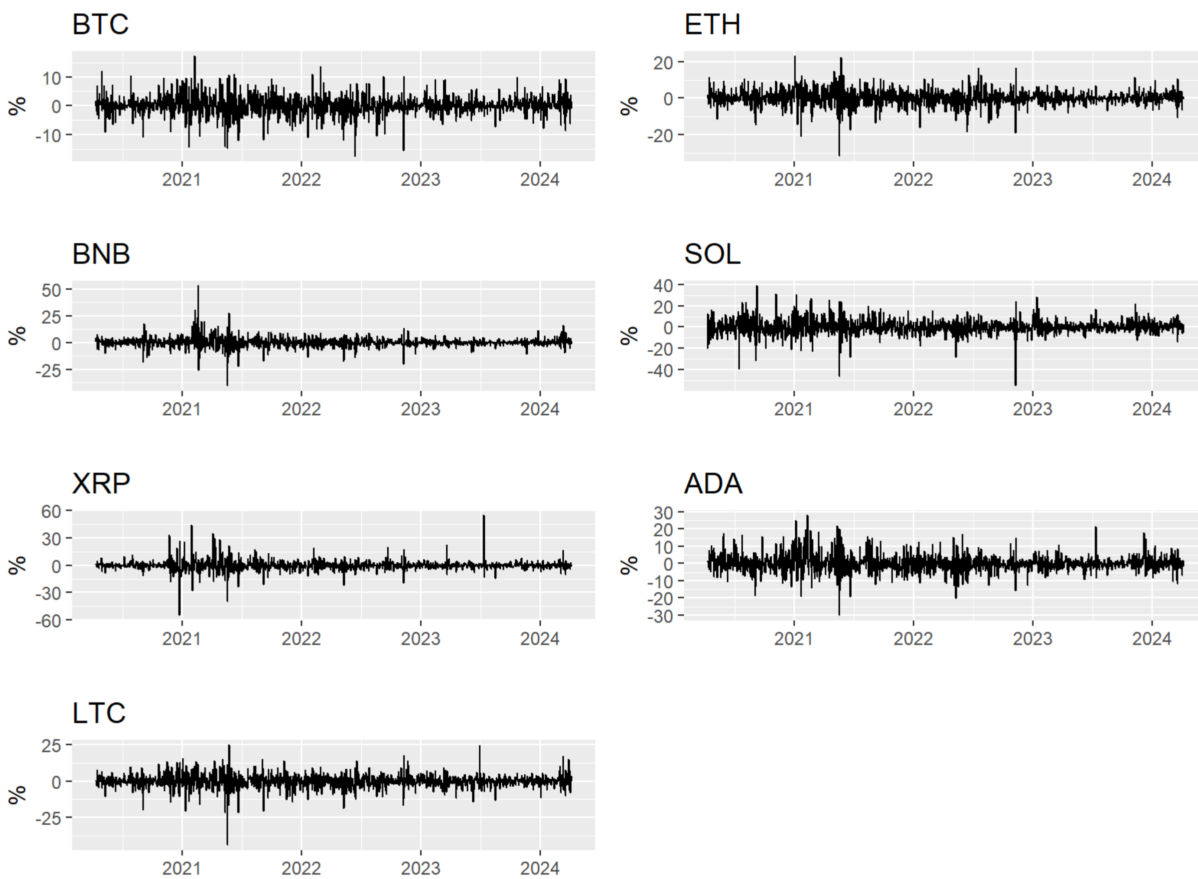
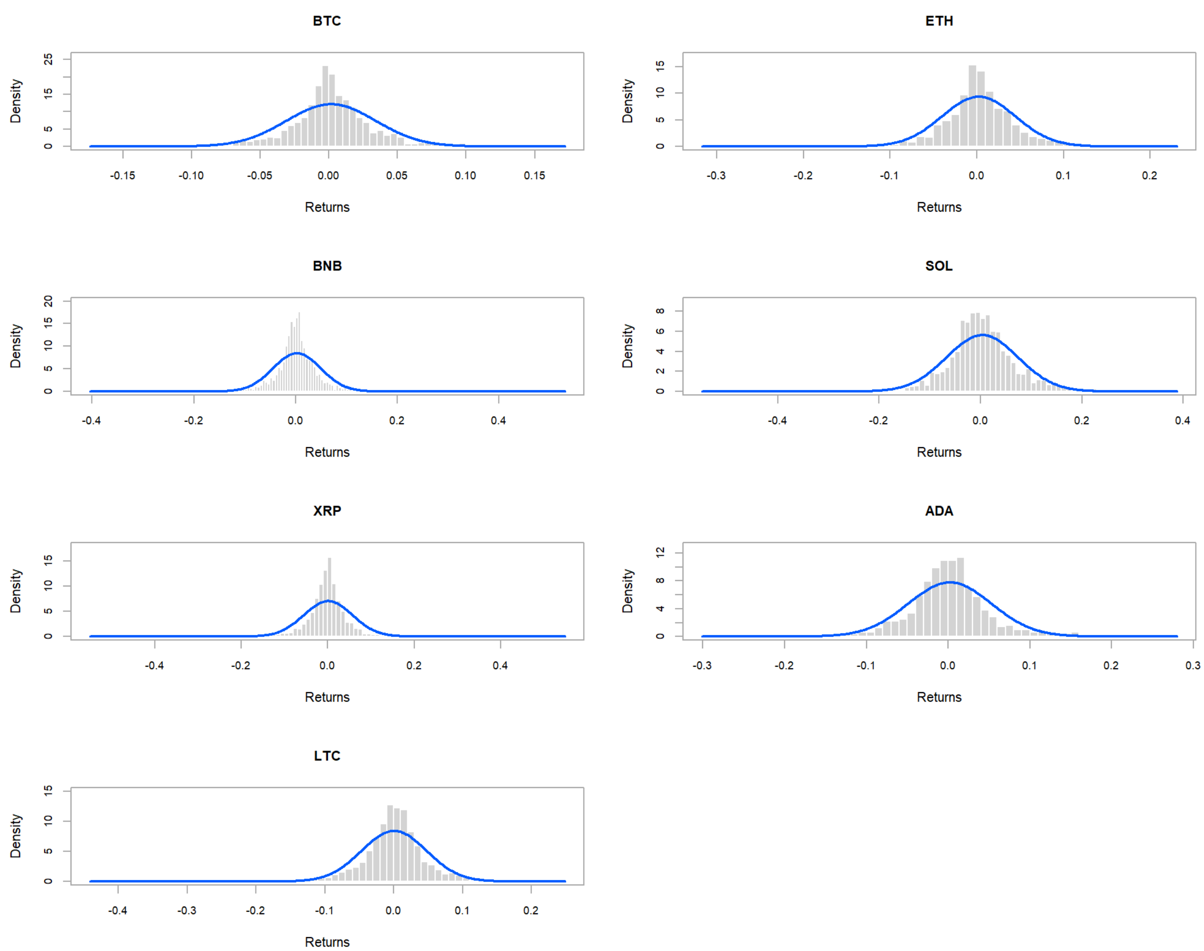


Table 1: Descriptive Statistics

	BTC	ETH	BNB	SOL	XRP	ADA	LTC
Mean (%)	0.159	0.210	0.257	0.359	0.079	0.197	0.060
Std (%)	3.282	4.261	4.699	7.102	5.653	5.127	4.746
Skew	-0.181	-0.395	0.727	-0.283	0.713	0.350	-0.704
Kurt	3.467	5.097	19.144	6.380	21.234	3.643	8.063
Min.	-17.405	-31.746	-40.445	-54.958	-55.050	-30.123	-44.119
Quantile 5%	-5.157	-6.508	-6.196	-9.705	-7.438	-7.630	-7.343
Median	0.079	0.166	0.155	0.053	0.054	0.070	0.144
Quantile 95%	5.367	7.014	6.792	12.051	7.238	8.931	7.395
Max.	17.182	23.070	52.922	38.718	54.855	27.944	24.843
Jarque-Bera test	738.109***	1615.9***	22394.021***	2492.107***	27514.859***	835.929***	4070.066***

Note: This table reports minimum, mean, median, maximum, standard deviation, skewness, excess kurtosis (measured in relation to a normal distribution), quantile of 5% and 95% and the statistics for Jarque-Bera test for BTC, ETH, BNB, SOL, XRP, ADA and LTC. The p-values for Jarque-Bera test are represented for $^{***}p < 0.01$ $^{**}p < 0.05$ and $^{*}p < 0.10$.

Figure 2: Density of the daily assets returns (2020-04-11 - 2024-04-07).



- **Choquet Portfolios**: Four different specifications of portfolios based on Choquet integration, which generalizes the traditional mean-variance framework to account for non-linear preferences or alternative risk measures. These specifications capture distinct risk-return preferences, providing a flexible approach to portfolio optimization.

- **Bitcoin**: The comparison benchmark is a portfolio made up of only Bitcoin.

By comparing these methodologies across different window lengths, we evaluate their performance under varying time horizons and market conditions. This approach enables a comprehensive analysis of portfolio dynamics and the robustness of each strategy.

For the Choquet portfolio, we follow the methodology outlined by Laurini (2007), exploring four cases of distortion functions:

1. A distortion function with two quantile parameters, where we minimize α-risk for quantiles $(0.1, 0.3)$ and assign weights $(0.7, 0.3)$.
2. A distortion function with three quantile parameters, $(0.01, 0.05, 0.1)$, and weights $(0.7, 0.2, 0.1)$.
3. A distortion function with three quantile parameters, $(0.01, 0.05, 0.1)$, but with modified weights $(0.85, 0.1, 0.05)$.
4. A distortion function with three quantile parameters, $(0.01, 0.05, 0.1)$, and weights $(0.85, 0.5, 0.05)$.

By exploring these configurations, we aim to assess the flexibility and performance of the Choquet portfolio under different risk-weighting schemes, effectively capturing both tail risks and variations in returns.

The results presented in Table 2 provide a detailed comparison of the performance and risk characteristics of various portfolio models, including the Bitcoin benchmark (BTC), the Naive portfolio, the Markowitz portfolio, and multiple cases of the Choquet portfolio, across different rolling window sizes. The table includes key metrics such as standard deviation (SD), Sharpe Ratio, Value-at-Risk (VaR), Modified Sharpe Ratio (MSR), and Expected Shortfall (ES), which are crucial for assessing both the risk and return profiles of the portfolios under consideration. Modified Sharpe Ratio replaces the standard deviation as the risk measure in the Sharpe Ratio, using the (negative) VaR(5%) as the risk measure.

The Modified Sharpe Ratio (MSR) (Gregoriou and Gueyie, 2003) extends the traditional Sharpe Ratio by addressing its limitations in capturing the risks associated with non-normal return distributions. While the standard Sharpe Ratio evaluates investment performance based on excess return per unit of standard deviation, the MSR incorporates measures that account for tail risks, such as Value-at-Risk (VaR) or Expected Shortfall (ES). This adjustment is particularly important for portfolios that include assets with heavy-tailed distributions, such as cryptocurrencies, where extreme price movements and higher kurtosis are prevalent.

By replacing standard deviation with a downside risk measure, the MSR provides a more accurate assessment of portfolio efficiency in contexts where returns deviate significantly

Table 2: Comparison of Metrics Across Window Sizes

Window	Model	SD	Sharpe Ratio	VaRq	ModifiedSR	ES
30	BTC	0.03081	0.06847	-0.04356	0.10318	-0.07389
	Naive	0.03729	0.06628	-0.05529	0.07951	-0.09113
	Markowitz	0.02423	0.07967	-0.03535	0.12527	-0.05285
	Choquet - case 1	0.02699	0.06971	-0.02901	0.20786	-0.04620
	Choquet - case 2	0.02911	0.06638	-0.03318	0.15227	-0.05958
	Choquet - case 3	0.02987	0.06729	-0.03460	0.14099	NA
	Choquet - case 4	0.03049	0.06614	-0.03435	0.14554	NA
60	BTC	0.03122	0.06004	-0.04448	0.06226	-0.07625
	Naive	0.03809	0.05624	-0.05766	0.04968	-0.09419
	Markowitz	0.02682	0.06273	-0.04039	0.06790	-0.06084
	Choquet - case 1	0.02828	0.05793	-0.03564	0.08166	-0.05808
	Choquet - case 2	0.02986	0.05574	-0.03966	0.07179	-0.06168
	Choquet - case 3	0.03262	0.05239	-0.04467	0.06346	NA
	Choquet - case 4	0.03330	0.05107	-0.04566	0.06091	NA
180	BTC	0.03212	0.04560	-0.04837	0.03725	-0.07782
	Naive	0.03978	0.03687	-0.06100	0.02946	-0.10052
	Markowitz	0.03029	0.04812	-0.04712	0.03951	-0.07075
	Choquet - case 1	0.03114	0.04661	-0.04427	0.04384	-0.06961
	Choquet - case 2	0.03185	0.04565	-0.04645	0.04090	-0.07983
	Choquet - case 3	0.03384	0.04286	-0.05078	0.03652	-0.09515
	Choquet - case 4	0.03402	0.04244	-0.05119	0.03604	-0.17497
365	BTC	0.03335	0.02974	-0.05212	0.02211	-0.08158
	Naive	0.04134	0.02299	-0.06422	0.01883	-0.10779
	Markowitz	0.03233	0.02956	-0.05074	0.02273	-0.07733
	Choquet - case 1	0.03291	0.02909	-0.04943	0.02221	-0.07646
	Choquet - case 2	0.03343	0.02862	-0.05131	0.02146	-0.09053
	Choquet - case 3	0.03453	0.02766	-0.05433	0.02044	-0.11597
	Choquet - case 4	0.03478	0.02771	-0.05481	0.02041	-0.16256

Note: This table reports, in average terms, the standard deviation, Sharpe Ratio, the Var(5%), Modified Sharpe ratio and Expected Shortfall for BTC, Naive portfolio, Markowitz portfolio and for all cases of Choquet portfolio. NA indicates that there were no VaR violations in the estimation window, preventing the calculation of the Expected Shortfall.

from normality. For instance, when VaR is used in the calculation, the MSR evaluates the excess return per unit of potential maximum loss at a given confidence level, such as 95% or 99%. This focus on downside risk allows the MSR to capture the effects of skewness and kurtosis, offering a perspective on portfolio performance that aligns with the realities of fat-tailed distributions.

The interpretation of the MSR is consistent with its traditional counterpart: higher values indicate a portfolio's ability to generate greater returns per unit of risk, with risk defined in terms of tail losses rather than overall volatility. Consequently, portfolios with superior MSR values are considered more efficient, as they balance return and tail risk more effectively. This makes the MSR particularly valuable for evaluating portfolios in volatile markets, as it provides insights into performance that are not captured by variance-based measures alone.

In terms of standard deviation, the Markowitz portfolio consistently shows the lowest value across all window sizes, reflecting its focus on minimizing overall volatility. In contrast, the Choquet portfolio, particularly in Case 1, exhibits slightly higher standard deviation values. This observation suggests that the Choquet portfolio does not prioritize variance minimization but instead addresses tail risk, which is a critical consideration in markets such as cryptocurrencies, where extreme price movements are common. Although the Choquet portfolio displays a higher standard deviation, it offers more robust protection against large negative returns, as indicated by its lower VaR and ES values compared to the Markowitz portfolio. This characteristic is particularly valuable in environments with high volatility and extreme price movements, such as the cryptocurrency market.

When examining the risk-adjusted performance metrics, the Sharpe Ratio of the Choquet portfolio in Case 1 is comparable to that of the Markowitz portfolio, although slightly lower in most cases. This trade-off can be understood in the context of the Choquet portfolio's focus on minimizing tail risk rather than overall volatility. While this results in a lower Sharpe Ratio, it enhances the portfolio's resilience against extreme losses, making it more suitable for high-risk, high-volatility assets like cryptocurrencies.

The Modified Sharpe Ratio (MSR) provides further evidence of the Choquet portfolio's superior performance in managing risk-adjusted returns. For shorter window sizes, such as 30 and 60 days, the Choquet portfolio (Case 1) exhibits significantly higher MSR values than the Markowitz and Naive portfolios. For instance, with a 30-day window, the MSR for Choquet (Case 1) is 0.20786, which is notably higher than the Markowitz portfolio's 0.12527 and the Naive portfolio's 0.07951. This indicates that, despite a higher standard deviation, the Choquet portfolio performs better when considering the trade-off between return and extreme risk. As the window size increases, the MSR values for all portfolios tend to decline, but the Choquet portfolio remains competitive, particularly in managing tail risk.

The results demonstrate that the Choquet portfolio, by minimizing exposure to assets with higher tail risk, offers superior protection in the face of extreme negative price movements, a common occurrence in the cryptocurrency market. This strategy allows the Choquet portfolio to outperform traditional models like the Markowitz portfolio, particularly in terms of tail risk management, as indicated by its superior MSR and lower ES values.

While the Choquet portfolio may not minimize variance as effectively as the Markowitz model, its superior ability to manage tail risk makes it a more appropriate choice for cryptocurrency investors navigating highly volatile and non-normal return distributions. Over shorter investment horizons, where tail risks are particularly pronounced, the Choquet portfolio offers a more favorable balance between risk and return. However, as the investment horizon lengthens, the disparities in risk measures between the Choquet and

Markowitz portfolios gradually diminish. In such cases, the Markowitz model may become more appealing to investors who prioritize variance minimization over the mitigation of tail risk.

Choquet-case1 stands out as the best-performing portfolio model in this analysis, largely due to its use of a distortion function with two quantile parameters, specifically designed to minimize α-risk for quantiles at 0.1 and 0.3, with assigned weights of 0.7 and 0.3, respectively. This approach allows the Choquet portfolio to effectively manage tail risk, focusing on the lower quantiles of the return distribution, where extreme losses are more likely to occur.

By minimizing the α-risk for these quantiles, Choquet-case1 prioritizes the reduction of potential losses in the left tail of the distribution, a critical consideration in highly volatile markets like cryptocurrencies. The choice of quantiles—0.1 and 0.3—addresses the risk of large negative movements while maintaining a balanced approach to the overall portfolio risk. The weighting of these quantiles (0.7 for 0.1 and 0.3 for 0.3) ensures that the portfolio is more sensitive to the most extreme losses, thus enhancing its ability to mitigate the impact of severe downturns in asset prices.

The success of Choquet-case1 can be attributed to this tailored risk management strategy, which goes beyond traditional risk measures such as variance and Sharpe Ratio, and instead focuses on safeguarding the portfolio from the most damaging losses. This is especially important in cryptocurrency markets, where extreme negative returns can occur with high frequency. By concentrating on tail risk, Choquet-case1 achieves a more robust risk-adjusted performance, demonstrated by its superior Modified Sharpe Ratio and lower Value-at-Risk and Expected Shortfall values compared to other models.

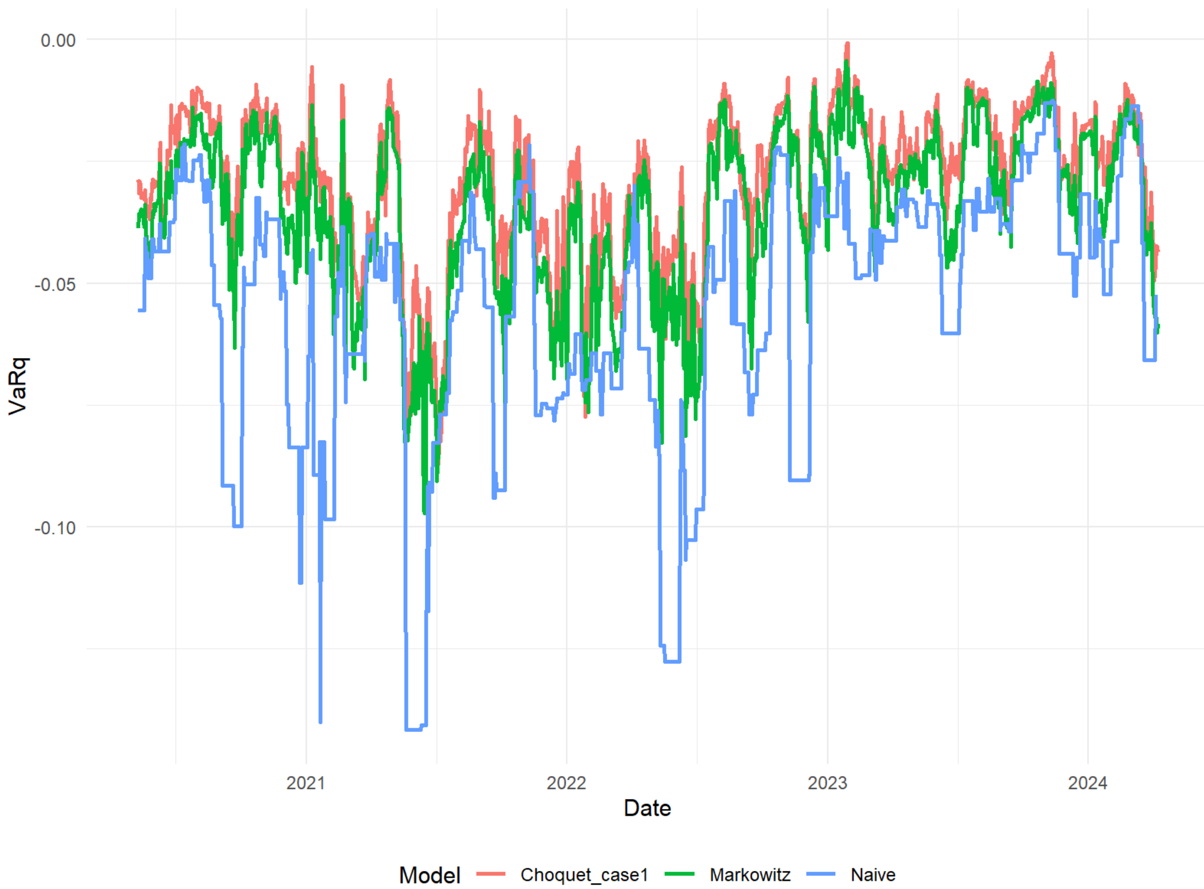
The Choquet portfolio's ability to achieve a high Modified Sharpe Ratio, despite higher standard deviation, demonstrates its robustness in balancing risk and return, making it a valuable alternative to traditional portfolio models like Markowitz, especially in markets characterized by extreme volatility and tail risks.

Figure 3 illustrates the temporal evolution of the 5% Value-at-Risk (VaR) for the optimal portfolios constructed using the Choquet, Markowitz, and Naive allocation methods using the rolling samples of size 30. It is evident that the Choquet method consistently exhibits the lowest tail risk across all periods analyzed, thereby confirming its superior ability to mitigate tail risk in the construction of optimal portfolios.

Figures 4 and 5 illustrate the efficient frontiers and risk measures, respectively, for the Markowitz (1952) portfolio and the Choquet portfolio (focusing on Case 1 due to its superior Sharpe Ratio) using a grid of log-returns in the interval $[0.00170, 0.0022]$ for the full sample. It is evident that, for all expected return values in this grid, the Choquet portfolio exhibits a higher standard deviation compared to the Markowitz portfolio. This observation aligns with the summary presented in Table 2 based on the Sharpe Ratio.

However, as noted by Laurini (2007) and Xu et al. (2016), variance is an inadequate risk measure because it cannot distinguish between extreme negative and positive returns. Unlike the Markowitz portfolio, which explicitly minimizes variance, the Choquet portfolio employs a different approach. Its objective function focuses on minimizing tail risk by incorporating a distortion function that assigns greater weight to less favorable outcomes,

Figure 3: Value-at-Risk - Window Size 30



Note: This figure plots the VaR (5%) for Choquet portfolio (case 1), Markowitz and Naive portfolios.

as outlined by Bassett et al. (2004). This methodology ensures a distinct risk management approach, prioritizing the reduction of extreme negative returns rather than treating all deviations from the mean equally.

As a result, while the Choquet portfolio may exhibit higher variance, this is not indicative of greater risk in the traditional sense. Instead, its optimization framework better accounts for the asymmetries and heavy tails characteristic of financial returns, particularly in volatile markets like cryptocurrencies, offering a distinct advantage over the variance-minimization approach of the Markowitz portfolio.

To explore alternative risk measures, Figure 5 illustrates the Value-at-Risk (VaR) for the mean-variance and the α-risk for Choquet portfolios. The results reveal that the Choquet portfolio exhibits lower risk, as measured by α-risk, compared to the classic portfolio optimization model, even as expected returns increase. This highlights the Choquet portfolio's ability to mitigate extreme downside risks more effectively than traditional mean-variance optimization.

Figure 6 presents the portfolio growth trajectories for BTC (benchmark), the Naive portfolio, the Markowitz portfolio, and the Choquet portfolio (case 1), starting with an initial capital of 100 for the full sample estimation. Although the Markowitz and Choquet portfolios follow similar growth patterns, the Choquet portfolio consistently achieves higher returns throughout the analyzed period, culminating in the largest accumulated portfolio return.

For context, holding only BTC in the portfolio for the entire period would yield an accumulated value of approximately 307.11. In contrast, a Naive portfolio with equal weights ($1/N$, where N is the number of cryptocurrencies) would achieve an accumulated return of about 369.74, the Markowitz a value of 489.71 and the Choquet portfolio a final value of 504.98. These benchmarks underscore the competitive performance of the Choquet portfolio.

Our findings suggest that the model proposed by Bassett et al. (2004) demonstrates significant advantages in the cryptocurrency market, maintaining superior performance characteristics in terms of both returns and risk measures. These results are consistent with the findings of prior studies that applied pessimistic approaches to portfolio optimization, reaffirming the Choquet portfolio's ability to balance return and volatility effectively.

A key advantage of the Choquet portfolio lies in its explicit focus on minimizing portfolio exposure to assets with higher tail risk. Unlike the mean-variance framework, which evaluates risk using variance, a symmetric measure that treats positive and negative deviations equally, the Choquet approach incorporates distortion functions to account for the asymmetry of returns. This allows it to penalize extreme negative outcomes more heavily, making it particularly well-suited for contexts characterized by severe tail risks, such as the cryptocurrency market.

Cryptocurrencies are known for their pronounced volatility and susceptibility to extreme negative price movements, which often occur unexpectedly and can significantly impact portfolio performance. In this environment, minimizing exposure to tail risk becomes a critical component of an effective asset allocation strategy. The Choquet portfolio's superior performance, as evidenced by its higher MSR and lower ES, demonstrates

Figure 4: Efficient frontiers

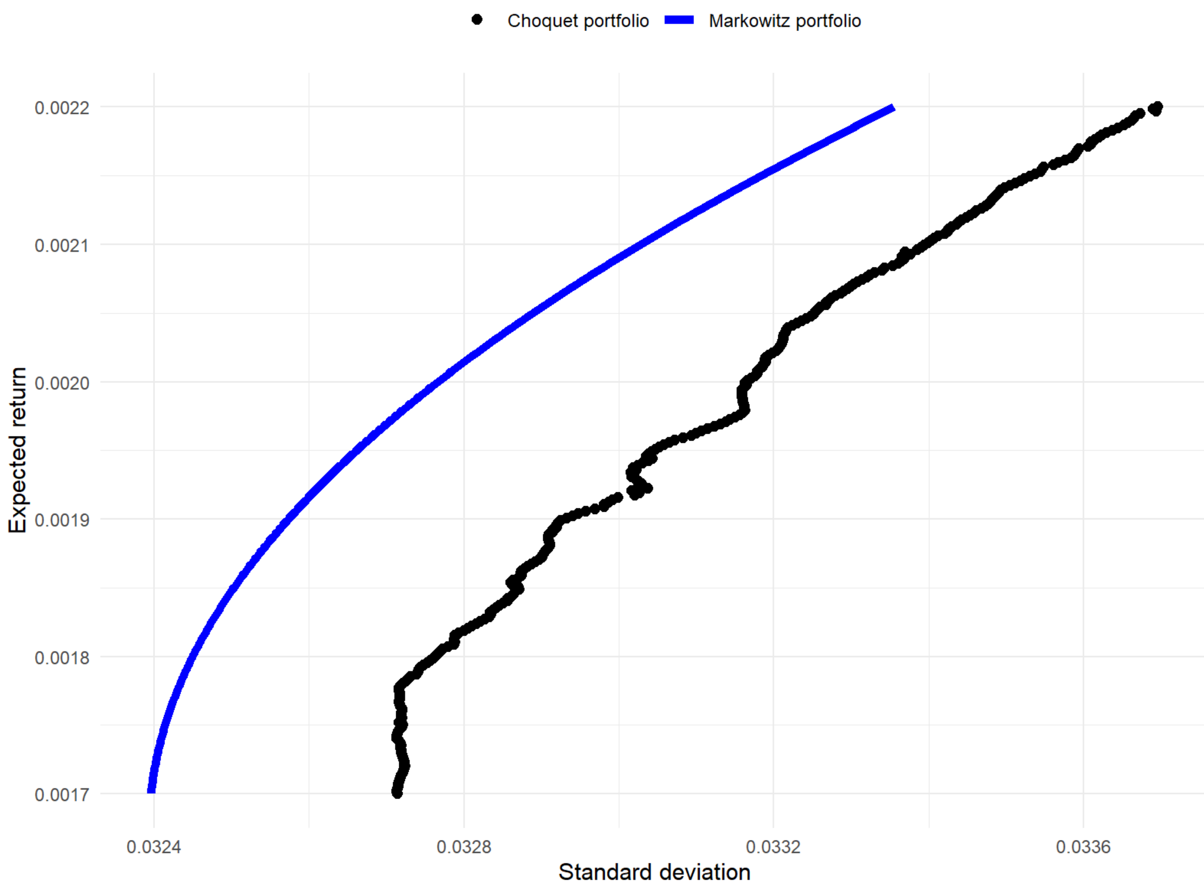


Figure 5: Value-at-Risk and α-risk – Markowitz and Choquet

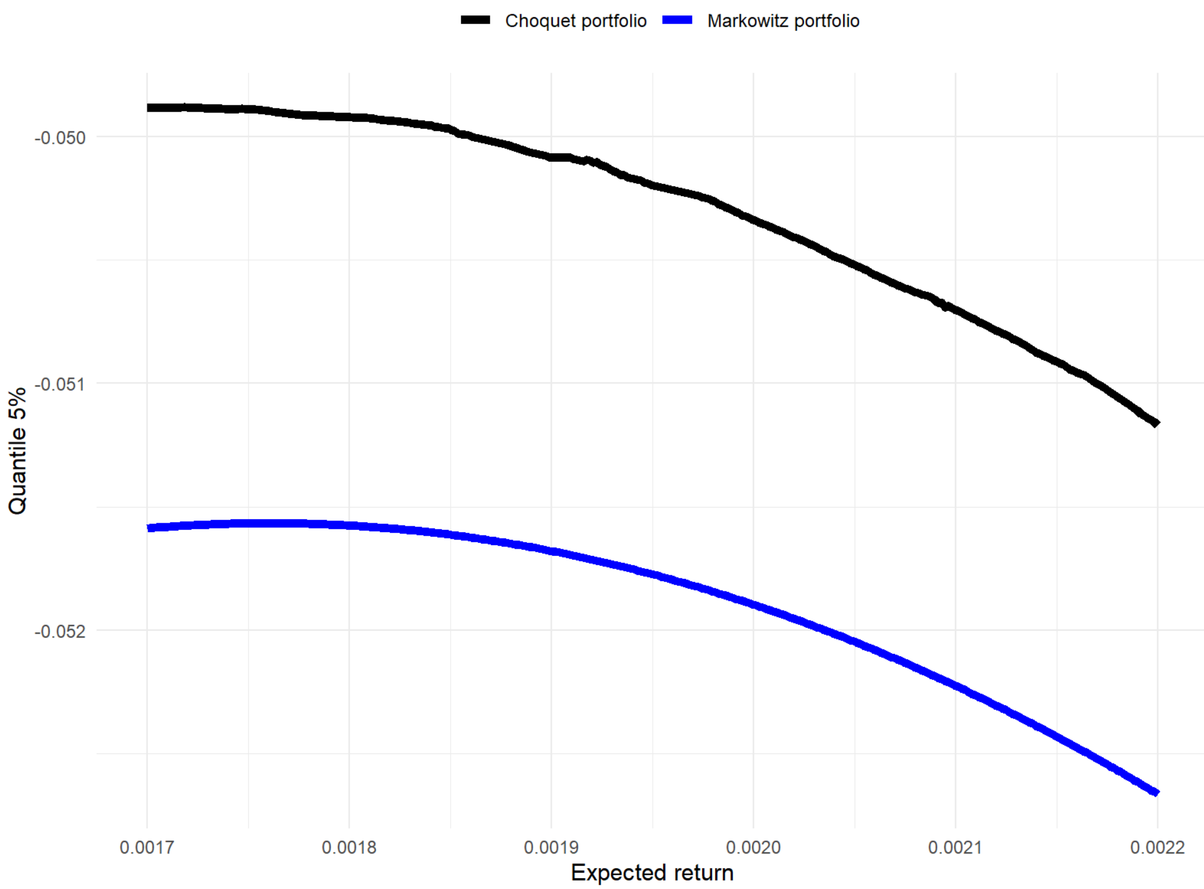
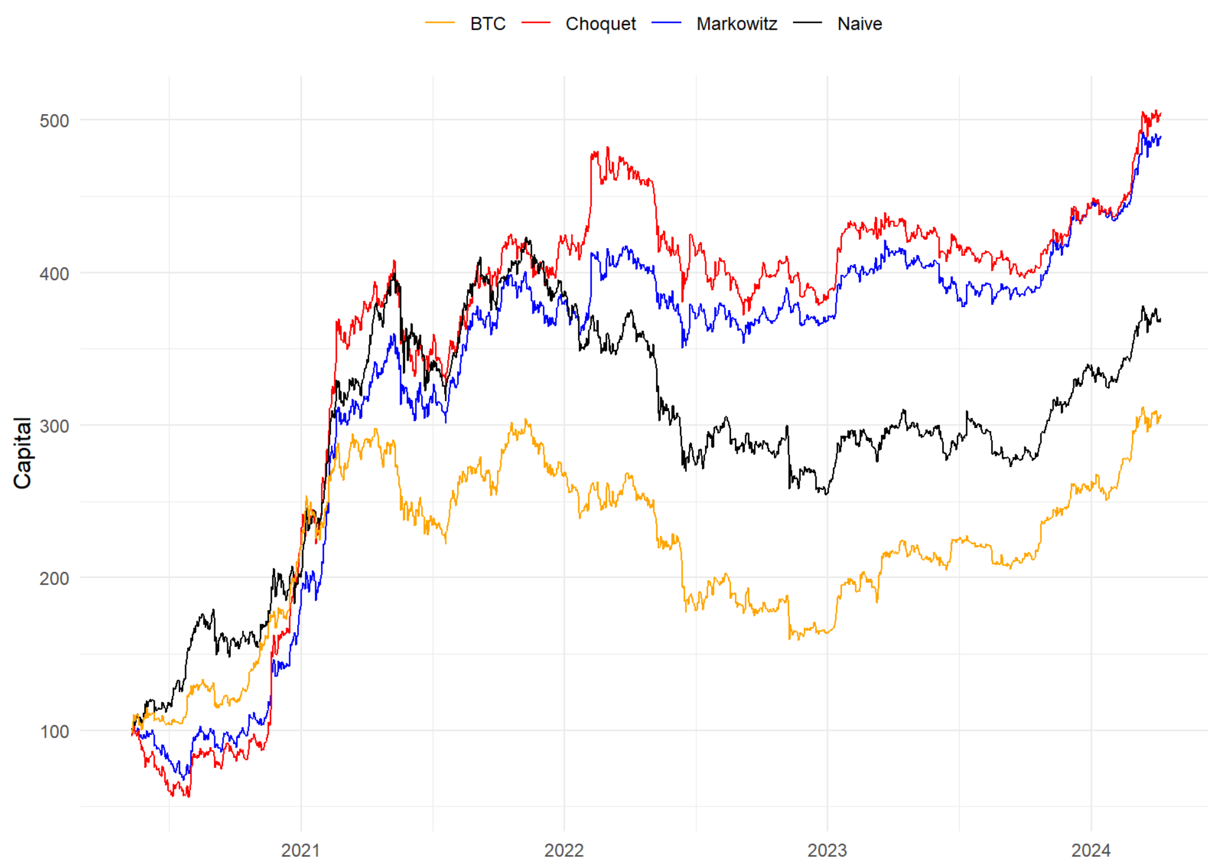


Figure 6: Portfolio Growth - portfolio with all cryptocurrencies



Note: This figure plots the accumulated portfolio return for BTC (benchmark; yellow), Choquet portfolio (case 1; red), Markowitz portfolio (blue) and Naive portfolio (black).

its robustness in addressing these challenges. By systematically reducing the impact of assets with higher downside risk, the Choquet portfolio achieves a more stable risk-return tradeoff, even in highly volatile markets.

Moreover, this performance advantage is not merely theoretical but has practical implications for investors seeking to navigate the unique risk profile of cryptocurrencies. The Choquet portfolio offers a more resilient allocation strategy, better aligning with the needs of risk-averse investors who prioritize downside protection without sacrificing return potential.

Our findings suggest that the model proposed by Bassett et al. (2004) is not only effective but also particularly well-suited for markets characterized by extreme price movements and heightened tail risks, such as cryptocurrencies. This reinforces the broader applicability of pessimistic portfolio optimization approaches in environments where traditional risk measures, like variance, fail to adequately capture the complexity of real-world risks.

5. Model Interpretability

Regarding portfolio allocation methods, an essential question for the users of RAs (Robo-Advisors) is: "How are portfolio weights determined?" In such cases, Babaei et al. (2022) posits that RAs can be conceptualized as models that solve the optimal portfolio problem without providing explicit explanations for the resulting allocations. This perspective aligns RAs with black-box machine learning models, which generate outcomes based on a variety of decision features (Redelmeier et al., 2020). However, these outcomes can be made interpretable through appropriate explainable machine learning (ML) methods (Babaei et al., 2022).

To address the explanatory challenges associated with these "black boxes", the second part of this article seeks to bridge the gap in understanding portfolio models implemented by RAs. Specifically, we compare the Markowitz model and the Choquet model as they relate to portfolio allocation. To align with ML methodologies, we employ Z-scores as standardized metrics that account for both volatility and return. These scores are computed for individual cryptocurrencies and portfolios. The Z-score for each cryptocurrency serves as an explanatory variable to predict the portfolio's overall Z-score. The Z-score for a given cryptocurrency, Z_i, is defined as follows:

$$Z_i = \frac{R_i - E[R_i]}{\sigma_i}, \tag{24}$$

where R_i, $E(R_i)$ and σ_i denotes, respectively, the return of cryptocurrency on the pth day, the average return and the standard deviation of the corresponding cryptocurrency return during the 30-day time window. Meanwhile, the Z score for each portfolio, Z_p, is computed by

$$Z_p = \frac{R_p - E[R_p]}{\sigma_p} \tag{25}$$

where R_p corresponds to the portfolio return on the pth day, $E[R_p] = \sum_{i=1}^{n} w_i E[R_i]$ and σ_p is measured as follows:

$$\sigma_p^2 = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_i w_j cov(R_i, R_j), \tag{26}$$

with $cov(R_i, R_j)$ being the covariance between ith and jth cryptocurrency.

5.1. Interpretable Machine Learning Models

In the following subsections, we outline the interpretable ML methods utilized in this study: the Shapley value and Local Interpretable Model-agnostic Explanations (LIME). First, we consider the Shapley value, which is extensively referenced in the literature (Shapley, 1953; Chen et al., 2018; Fréchette et al., 2016; Brigo et al., 2021; Jaeger et al., 2021). Subsequently, we address the Local Interpretable Model-Agnostic Explanations (LIME) proposed by Ribeiro et al. (2016), with the aim of expanding the analysis presented in Babaei et al. (2022).

For the implementation of interpretable models, we employed a random forest model comprising 100 trees. It is important to note that, consistent with the approach in Babaei et al. (2022), our primary focus is not on evaluating prediction accuracy or comparing the fit of Z-scores across ML models.

5.1.1. Shapley Values

Interpretable models for analyzing results are becoming increasingly critical as the adoption of ML methods expands, both in terms of their application and their penetration into new fields. According to Merrick and Taly (2020), the most prominent models in this domain aim to assign a score to each feature, quantifying its contribution to the prediction.

Among various feature-importance models, the Shapley value model (Shapley, 1953) has garnered significant attention in recent years (Štrumbelj and Kononenko, 2014; Datta et al., 2016; Lundberg and Lee, 2017; Chen et al., 2018; Lundberg et al., 2019; Ancona et al., 2019; Aas et al., 2021). Rooted in cooperative game theory, this model provides a mathematically fair and unique method for distributing the payoff of a cooperative game among its participants, as highlighted by Merrick and Taly (2020).

Following Merrick and Taly (2020), define a set of players as $\mathcal{M} = \{1, \ldots, M\}$ and a set function as $v : 2^{\mathcal{M}} \rightarrow \mathbb{R}$, where the payoff is given by $v(S)$ for each $S \subseteq \mathcal{M}$ and, also $v(\emptyset) = 0$. Through this construction, the Shapley values of player i, $\phi_i(v)$, can be computed as

$$\phi_i(v) = \frac{1}{M} \sum_{S \subseteq \mathcal{M} \setminus \{i\}} \binom{M-1}{|S|}^{-1} (v(S \cup \{i\}) - v(S)), \tag{27}$$

where $v(S)$ denotes the function that maps subsets of players to the contribution of the set S, $|S|$ represents the number of players in the set S. Moreover, according to Redelmeier et al. (2020), this payoff $\phi_i(v)$ can be seem as a weighted sum of the player's marginal contributions to each set S. It is worth noting that the Shapley value satisfies four axioms: symmetry, Dummy (Null player), efficiency and linearity[3].

In a ML context, the players of the game are the input features, and the payoff is the prediction resulting of the ML model (Merrick and Taly, 2020). Therefore, following Babaei et al. (2022), the marginal contribution of each predictor X_k, where $k = 1, \ldots, K$, can be computed as

$$\phi(\hat{f}(X_i)) = \sum_{X' \subseteq \mathcal{C}(X) \setminus X_k} \frac{|X'|!(K - |X'| - 1)!}{K!} [\hat{f}(X' \cup X_k)_i - \hat{f}(X')_i] \tag{28}$$

5.1.2. Local Interpretable Model-agnostic Explanations (LIME)

It is important to note that, in the Shapley Value method, none of the cryptocurrencies are excluded from the explanatory dataset, as illustrated in Figures 7 and 8. According to Karanika et al. (2020), extracting only the features that most significantly explain a model's predictions can introduce challenges related to dimensionality reduction.

To enhance confidence in individual predictions generated by ML models, Ribeiro et al. (2016) propose the Local Interpretable Model-Agnostic Explanations (LIME) method, which relies on local approximations to interpret individual predictions. As highlighted by Brigo et al. (2021), the LIME model constitutes an additive feature attribution approach, as it can be expressed as a linear function of binary variables, as demonstrated in Equation (29).

$$g(x') = \phi_0 + \sum_{i=1}^{M} \phi_i x'_i, \tag{29}$$

Here, $x' \in 0, 1^M$, $\phi_i \in \mathbb{R}$, and M represent the number of simplified inputs x', also referred to as interpretable inputs. These inputs are generated using a mapping function h_x, which translates a binary vector of interpretable inputs, x', back into the original inputs, i.e., $x = h_x(x')$, as noted by Brigo et al. (2021). It is important to highlight that in Equation (29), x' assigns a value of 1 if the input component is present and 0 otherwise.

Let $f(x)$ denote the prediction to be explained. To calculate the interpretability effect on the prediction, Ribeiro et al. (2016) propose minimizing a squared loss function subject to a complexity measure associated with the explanation model $g \in G$. This optimization problem can be expressed as follows:

$$\xi(x) = \arg\min_{g \in G} \mathcal{L}(f, g, \pi_x) + \Omega(g), \tag{30}$$

[3]For more details about the axioms see Merrick and Taly (2020).

where the explanation model g, among all possible models G, is the argument that minimize the square loss function $\mathcal{L}$ for x'. The measure of proximity between x' and x is denoted by π_x and, also $\Omega(g)$ represents a measure of complexity of the explanation model $g \in G$.

As highlighted in Ribeiro et al. (2016), this procedure approximated $g(x')$ using least squares adjusted with Lasso (Tibshirani, 1996), named K-Lasso by the authors. Note that the penalization parameter in Lasso, λ, is chosen so that you have m interpretable inputs x' and, in general, a low value for K is chosen to maintain an easy interpretation of the results (Izbicki and dos Santos, 2020). Here we set $K = 3$, with the aim of highlighting the three cryptocurrency that contribute most to the prediction of the portfolio's Z-score.

6. Interpretability Results

In this section, we present and analyze the results of interpretable ML models applied to the Markowitz portfolio model and the Choquet portfolio approach (Case 1) for the full sample estimation, using the mean return as target return in the optimization procedures. First, the results derived from the Shapley value method are shown in Figures 7 and 8, followed by the results of the LIME model, presented in Figure 9. To examine how each cryptocurrency contributes to the prediction of the portfolio Z-score (Z_p), Figure 7 illustrates the Shapley value decomposition for four portfolios, which were randomly selected from all daily allocations, under the Markowitz and Choquet (Case 1) portfolio models. The Z_p values, computed using Equation (25), are displayed at the top of each portfolio, calculated for each optimal portfolio method.

The Shapley values vary across the four randomly selected portfolios and differ between the mean-variance and Choquet portfolio models. For instance, in Portfolio 1, cryptocurrencies such as ADA, BNB, and LTC, computed using the Markowitz portfolio model (a), appear to have greater importance (in absolute terms) in predicting Z_p. Conversely, the Choquet model for the same portfolio (b) assigns higher importance to BTC, XRP, and LTC in the Z_p prediction. For Portfolio 2, Figures (c) and (d) indicate that BTC and ADA are the two most significant features in predicting the portfolio Z-score. The importance of the remaining cryptocurrencies (SOL, BNB, LTC, ETH, and XRP) is notably reduced compared to BTC and ADA and is also relatively consistent when comparing the two portfolio approaches.

To analyze the overall contribution of each cryptocurrency to the portfolio Z-score predictions, we employed the SHAP Feature Importance technique, which computes the average absolute Shapley values for each feature. The results, shown in Figure 8, indicate that for both portfolio methods analyzed, the cryptocurrencies that most significantly explain the variations in portfolio Z-scores—which are influenced by both return and risk—are, in order, BTC, BNB, and XRP.

Figure 8 also highlights the different order of contribution for the rest of the cryptocurrencies. In the Markowitz (1952) portfolio, the full order of contribution is given by BTC, BNB, XRP, ETH, ADA, LTC and SOL. Meanwhile, the cryptocurrencies that contribute the most to variations in portfolio allocation in the Choquet portfolio (case 1) are BTC, BNB, XRP, LTC, ETH, SOL and ADA. These different allocation weights that the portfolio methods have given to other cryptocurrencies may help explain the difference in portfolio returns shown in Figure 6.

To extend the work of Babaei et al. (2022) by incorporating another interpretable ML method and to gather further evidence on which cryptocurrencies most contribute to explaining portfolio Z-score predictions, we utilized the LIME model. This method employs a LASSO-regularized linear regression, limited to K features, to approximate the predictions. The results of the LIME model are presented in Figure 9, using the same four randomly selected portfolios as in Figure 7. A notable result is that for all portfolios, the K-most influential cryptocurrencies identified by LIME are BTC, BNB, and XRP, consistent with the findings of the SHAP Feature Importance approach.

To further explore the predictive explanatory properties of these three cryptocurrencies in our portfolios, we constructed three additional portfolios using only the K-most significant cryptocurrencies: the Naive K-portfolio, Markowitz K-portfolio, and Choquet K-portfolio. These new portfolios were compared with the original portfolios in Figure 10. The results reveal that a Naive portfolio constructed with only the K-most significant cryptocurrencies (Naive K-portfolio) replicates approximately 88.39% of a portfolio including all cryptocurrencies analyzed (Naive portfolio). Similarly, the Markowitz K-portfolio achieves approximately 90.07% replication of the complete Markowitz portfolio.

However, the Choquet portfolio not only exhibits the highest accumulated portfolio return but also achieves the highest replication rate among the methods studied, approximately 94.42%. These results suggest an additional property of the Bassett et al. (2004) model: the portfolio replication property, particularly when combined with an interpretable ML model and an RA-based portfolio construction technique.

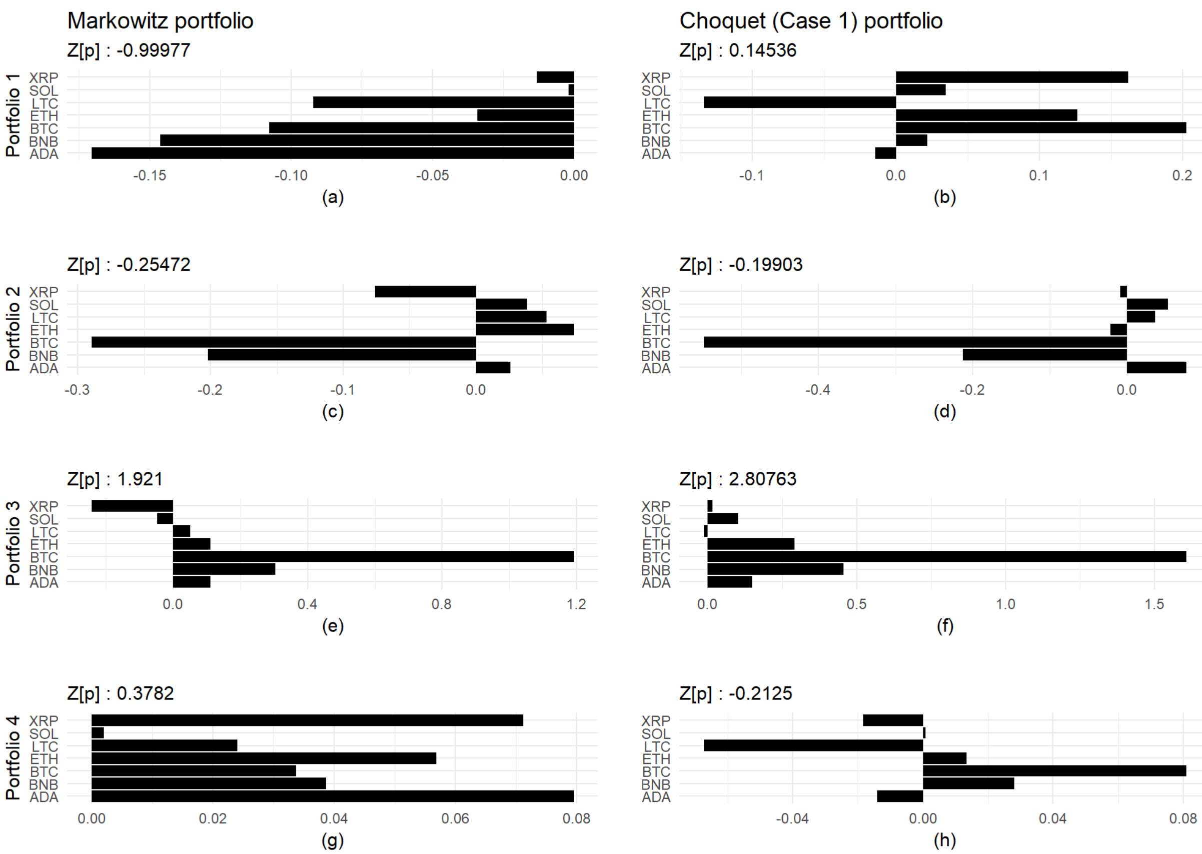
7. Conclusions

The Choquet portfolio model, grounded in pessimistic decision-making and utilizing the Choquet integral, has emerged as a promising approach for addressing extreme risks and tail events. By emphasizing unfavorable outcomes, the model aligns with the unique characteristics of cryptocurrencies, including extreme losses and heightened tail risks.

The Choquet portfolio model, with its reliance on α-risk formulation of Conditional Value-at-Risk (CVaR) as a risk measure, is better suited for this environment. CVaR directly addresses tail risk by focusing on the worst-case outcomes in the return distribution, providing a more realistic measure of potential losses in extreme market conditions. This study advances the literature by applying the Choquet portfolio to cryptocurrency markets for the first time, comparing it to the traditional Markowitz mean-variance model. Furthermore, Machine Learning interpretability tools, such as Shapley Values and LIME, are employed to elucidate differences in portfolio weights and returns, enhancing transparency and understanding of the optimization process.

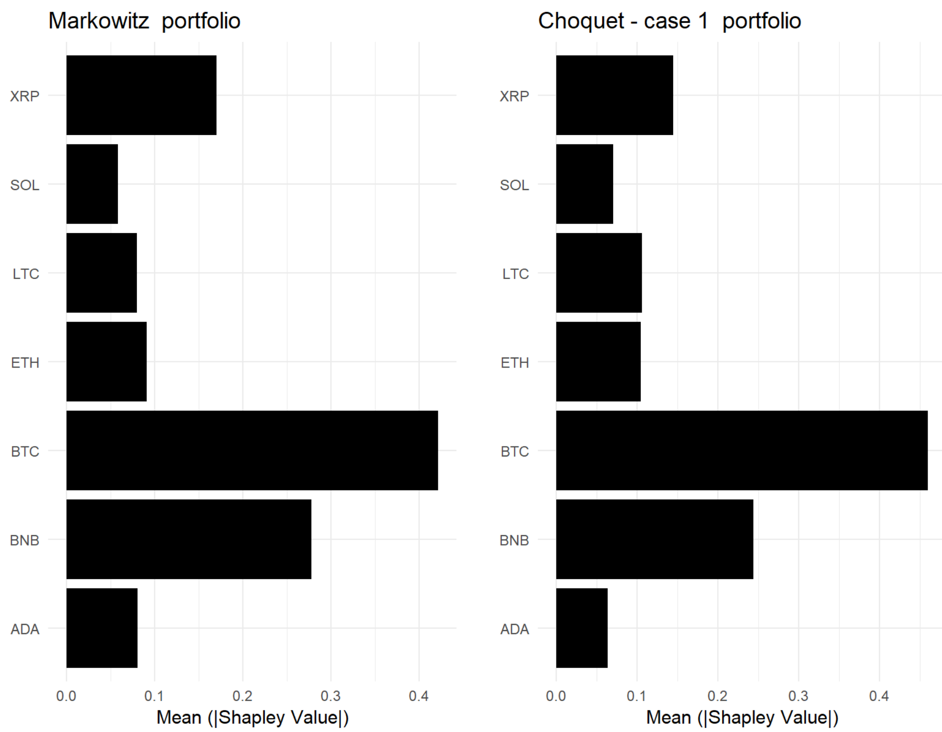
Results demonstrate that the Choquet portfolio achieves superior performance in cumulative returns, tail risk mitigation, and replicability compared to traditional models, offering a more robust framework for managing the unique risks of cryptocurrency investments. These findings underscore the importance of integrating interpretable ML tools

Figure 7: Contribution of each explanatory variable to the Shapley's decomposition of four generated portfolios.



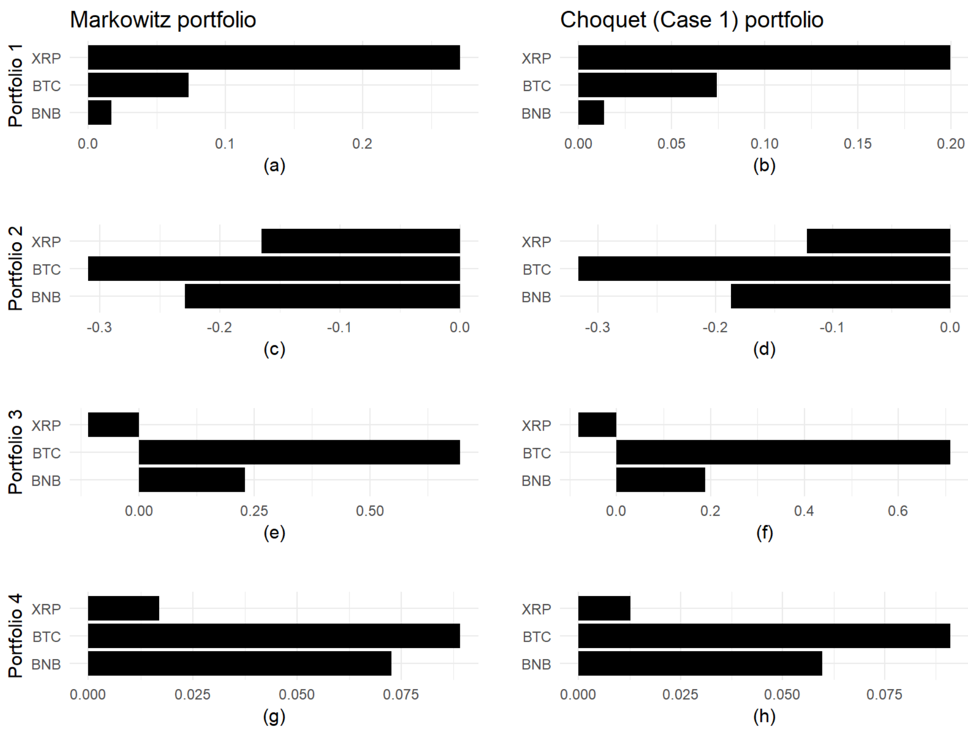
Note: This figure shows the Shapley value for four different portfolios choose randomly.

Figure 8: Overall contribution of cryptocurrencies to the Z scores of portfolios



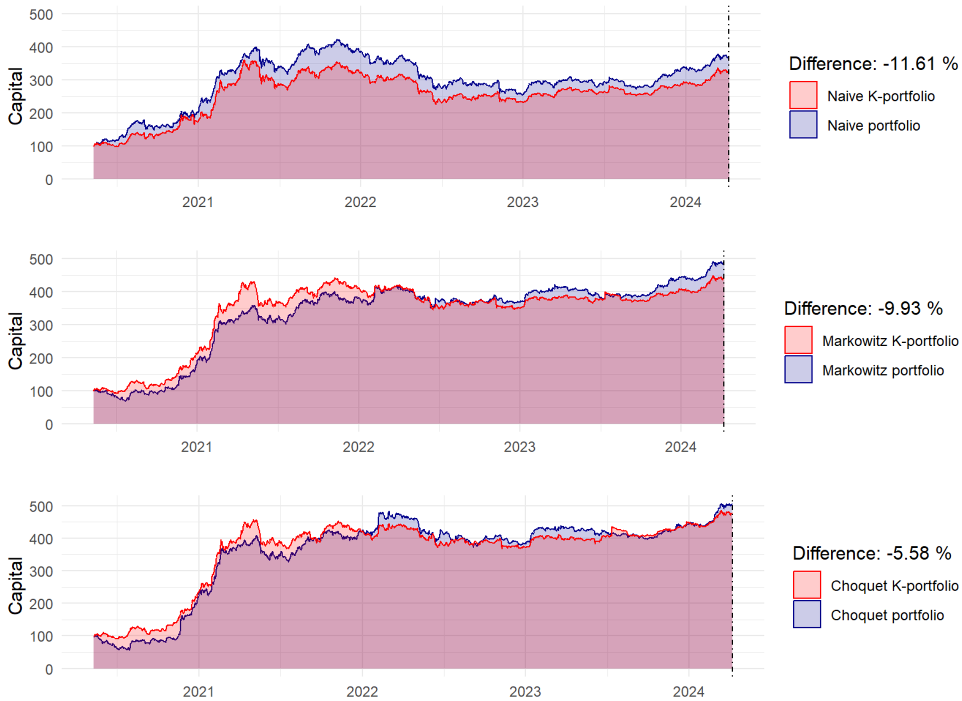
Note: This figure shows the average contribution of each cryptocurrencies in the predictions, i.e. the global importance for each feature.

Figure 9: Contribution of each explanatory variable to the LIME decomposition of three generated portfolios.



Note: This figure shows the contribution measured with LIME model for the same four portfolios as in Figure 7.

Figure 10: Portfolio replication with $K = 3$ (BTC, XRP and BNB)



with advanced portfolio models to provide deeper insights into the drivers of portfolio performance, thereby fostering more informed and transparent investment decisions.

Since that the cryptocurrencies do not follow a Gaussian distribution, the method proposed by Bassett et al. (2004), which incorporates highers moments into portfolio decision, seems to be theoretically and empirical more appropriate among the portfolios optimal methods analysed. Our results indicate that the Choquet portfolio (case 1) present higher accumulated portfolio return and low tail risk when compared to the Markowitz (1952)‘s model.

In addition, the second part of our article aims to explain the weights assigned by a Robo-Advisor for Markowitz (1952)´s model and the Choquet portfolio, in certain periods. For this, we apply the Shapley Values method, as in Babaei et al. (2022), using the z-scores of each cryptocurrency as explanatory variables to predict the z-score of the portfolio. Furthermore, in order to extend the analyses and, find out which are the K-cryptocurrencies that most explain the predictions, the Local Interpretable Model-agnostic Explanations (LIME) model is applied.

Through the LIME model results, we found out that, among all analysed cryptocurrencies (BTC, ETH, BNB, SOL, XRP, ADA and LTC), the most explainable ones are, in order, BTC, BNB and XRP. With the aim of testing the power of these K cryptocurrencies in Robo-Advisor's choice of weights for each portfolio optimization model, three more portfolios were created only with BTC, BNB and XRP: Naive K-portfolio, Markowitz K-portfolio and Choquet K-portfolio. The results point out that the Choquet portfolio (case

1), beyond display the higher portfolio return accumulated, presents the lowest difference rate between the portfolio computed with all cryptocurrencies and the K-portfolio. This indicates that the model proposed by Bassett et al. (2004), when combined with an interpretable ML model and also an RA portfolio construction technique, presents higher performance for portfolio replication compared to conventional portfolio methods.

Our paper dialogues with the literature on dimensionality reduction and, above all, with the literature on replication and the construction of optimal portfolios. In addition, our results help investors understand how the these automatic process (Robo-Advisors) chooses the weights according to an portfolio model in the cryptocurrency market, which is known for its volatility.

References

Aas, K., Jullum, M., and Løland, A. (2021). Explaining individual predictions when features are dependent: More accurate approximations to Shapley values. *Artificial Intelligence*, 298:103502.

Aljinović, Z., Marasović, B., and Šestanović, T. (2021). Cryptocurrency portfolio selection—a multicriteria approach. *Mathematics*, 9(14):1677.

Ancona, M., Oztireli, C., and Gross, M. (2019). Explaining deep neural networks with a polynomial time algorithm for Shapley value approximation. In *International Conference on Machine Learning*, pages 272–281. PMLR.

Anyfantaki, S., Arvanitis, S., and Topaloglou, N. (2018). Diversification, integration and cryptocurrency market. Working paper, Bank of Greece.

Areewong, S. et al. (2009). Higher moments with foreign investment: a thailand', s perspective. Technical report, Thammasat University.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical finance*, 9(3):203–228.

Arzova, S. B. and Özdurak, C. (2021). Optimal cryptocurrency and BIST 30 portfolios with the perspective of markowitz portfolio theory. *J. Financ. Econ*, 9:146–154.

Babaei, G., Giudici, P., and Raffinetti, E. (2022). Explainable artificial intelligence for crypto asset allocation. *Finance Research Letters*, 47:102941.

Bassett, Gilbert W., J., Koenker, R., and Kordas, G. (2004). Pessimistic portfolio allocation and choquet expected utility. *Journal of Financial Econometrics*, 2(4):477–492.

Bowala, S. and Singh, J. (2022). Optimizing portfolio risk of cryptocurrencies using data-driven risk measures. *Journal of Risk and Financial Management*, 15(10):427.

Brauneis, A. and Mestel, R. (2019). Cryptocurrency-portfolios in a mean-variance framework. *Finance Research Letters*, 28:259–264.

Brenner, L. and Meyll, T. (2020). Robo-advisors: a substitute for human financial advice? *Journal of Behavioral and Experimental Finance*, 25:100275.

Brigo, D., Huang, X., Pallavicini, A., and Borde, H. S. d. O. (2021). Interpretability in deep learning for finance: a case study for the Heston model. *arXiv preprint arXiv:2104.09476*.

Chen, J., Song, L., Wainwright, M. J., and Jordan, M. I. (2018). L-Shapley and C-Shapley: Efficient model interpretation for structured data.

Constantinides, G. M. and Malliaris, A. G. (1995). Portfolio theory. *Handbooks in Operations Research and Management Science*, 9:1–30.

Datta, A., Sen, S., and Zick, Y. (2016). Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems. In *2016 IEEE symposium on security and privacy (SP)*, pages 598–617. IEEE.

De Athayde, G. M. and Flôres Jr, R. G. (2004). Finding a maximum skewness portfolio—a general solution to three-moments portfolio choice. *Journal of Economic Dynamics and Control*, 28(7):1335–1352.

Fréchette, A., Kotthoff, L., Michalak, T., Rahwan, T., Hoos, H., and Leyton-Brown, K. (2016). Using the shapley value to analyze algorithm portfolios. *Proceedings of the AAAI Conference on Artificial Intelligence*, 30(1).

Gilboa, I. (2009). *Choquet Expected Utility*, page 147–153. Econometric Society Monographs. Cambridge University Press.

Gregoriou, G. N. and Gueyie, J.-P. (2003). Risk-adjusted performance of funds of hedge funds using a modified sharpe ratio. *The Journal of Wealth Management*, 6(3):77–83.

Harvey, C. R., Liechty, J. C., Liechty, M. W., and Müller, P. (2010). Portfolio selection with higher moments. *Quantitative Finance*, 10(5):469–485.

Hrytsiuk, P., Babych, T., and Bachyshyna, L. (2019). Cryptocurrency portfolio optimization using Value-at-Risk measure. In *6th International Conference on Strategies, Models and Technologies of Economic Systems Management (SMTESM 2019)*, pages 385–389. Atlantis Press.

Izbicki, R. and dos Santos, T. M. (2020). *Aprendizado de máquina: uma abordagem estatística*. Rafael Izbicki.

Jaeger, M., Krügel, S., Marinelli, D., Papenbrock, J., and Schwendner, P. (2021). Interpretable machine learning for diversified portfolio construction. *The Journal of Financial Data Science*, 3(3):31–51.

Karanika, A., Oikonomou, P., Kolomvatsos, K., and Anagnostopoulos, C. (2020). An ensemble interpretable machine learning scheme for securing data quality at the edge. In *Machine Learning and Knowledge Extraction: 4th IFIP TC 5, TC 12, WG 8.4, WG 8.9, WG 12.9 International Cross-Domain Conference, CD-MAKE 2020, Dublin, Ireland, August 25–28, 2020, Proceedings 4*, pages 517–534. Springer.

Koenker, R. (2005). Quantile regression. *Econometric society monographs.*

Koenker, R. and Bassett Jr, G. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society*, pages 33–50.

Koenker, R., Chernozhukov, V., He, X., and Peng, L. (2017). *Handbook of Quantile Regression*. CRC Press.

Krokhmal, P. A. (2007). Higher moment coherent risk measures. *Quantitative Finance*, 7(4):373–387.

Lam, J. W. (2016). Robo-advisors: A portfolio management perspective. *Senior thesis, Yale College*, 20.

Laurini, M. P. (2007). Pessimistic preferences and portfolio allocation-empirical analysis and risk management applications. *Revista de Economia e Administração*, 6(3).

Leo, M., Sharma, S., and Maddulety, K. (2019). Machine learning in banking risk management: A literature review. *Risks*, 7(1):29.

Leow, E. K. W., Nguyen, B. P., and Chua, M. C. H. (2021). Robo-advisor using genetic algorithm and BERT sentiments from tweets for hybrid portfolio optimisation. *Expert Systems with Applications*, 179:115060.

Lundberg, S. M., Erion, G. G., and Lee, S.-I. (2019). Consistent individualized feature attribution for tree ensembles.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 4768–4777, Red Hook, NY, USA. Curran Associates Inc.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1):77–91.

Merrick, L. and Taly, A. (2020). The explanation game: Explaining machine learning models using Shapley values. In *Machine Learning and Knowledge Extraction: 4th IFIP TC 5, TC 12, WG 8.4, WG 8.9, WG 12.9 International Cross-Domain Conference, CD-MAKE 2020, Dublin, Ireland, August 25–28, 2020, Proceedings 4*, pages 17–38. Springer.

Park, J. Y., Ryu, J. P., and Shin, H. J. (2016). Robo advisors for portfolio management. *Advanced Science and Technology Letters*, 141(1):104–108.

Phoon, K. F. and Koh, C. C. F. (2018). Robo-advisors and wealth management. *Journal of Alternative Investments*, 20(3):79.

Platanakis, E., Sutcliffe, C., and Urquhart, A. (2018). Optimal vs naïve diversification in cryptocurrencies. *Economics Letters*, 171:93–96.

Ravipati, A. (2012). *Markowitz's portfolio selection model and related problems*. PhD thesis, Rutgers University-Graduate School-New Brunswick.

Redelmeier, A., Jullum, M., and Aas, K. (2020). Explaining predictive models with mixed features using Shapley values and conditional inference trees. In *Machine Learning and Knowledge Extraction: 4th IFIP TC 5, TC 12, WG 8.4, WG 8.9, WG 12.9 International Cross-Domain Conference, CD-MAKE 2020, Dublin, Ireland, August 25–28, 2020, Proceedings 4*, pages 117–137. Springer.

Rezaei, H., Faaljou, H., and Mansourfar, G. (2021). Intelligent asset allocation using predictions of deep frequency decomposition. *Expert Systems with Applications*, 186:115715.

Ribeiro, M. T., Singh, S., and Guestrin, C. (2016). " why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144.

Rossi, A. G. and Utkus, S. P. (2020). Who benefits from robo-advising? Evidence from machine learning. *Evidence from Machine Learning (March 10, 2020)*.

Shapley, L. S. (1953). A value for n-person games. In Kuhn, H. W. and Tucker, A. W., editors, *Contributions to the Theory of Games II*, pages 307–317. Princeton University Press, Princeton.

Štrumbelj, E. and Kononenko, I. (2014). Explaining prediction models and individual predictions with feature contributions. *Knowledge and information systems*, 41:647–665.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288.

Xu, Q., Zhou, Y., Jiang, C., Yu, K., and Niu, X. (2016). A large CVaR-based portfolio selection model with weight constraints. *Economic Modelling*, 59:436–447.