University of Brasília
Institute of Exact Sciences
Department of Statistics

Master's Dissertation

A Multi-Armed Bandit Framework for Portfolio Allocation

by

Gustavo Maia Rodrigues Gomes

Brasília, December 21, 2022

A Multi-Armed Bandit Framework for Portfolio Allocation

by

Gustavo Maia Rodrigues Gomes

Dissertation submitted to the Departament of Statistics at the University of Brasília, as part of the requirements required to obtain the Master's Degree in Statistics.

Advisor: Prof. Raul Yukihiro Matsushita, PhD

Brasília, December 21, 2022

Dissertation submitted to the Graduate Program in Statistics of the Departament of Statistics at the University of Brasília, as part of the requirements required to obtain the Master's Degree in Statistics.

Approved by:

Prof. Raul Yukihiro Matsushita, PhD
Advisor/Dissertation Examining Committee Chair, EST/UnB

Prof. Regina Célia Bueno da Fonseca, PhD
Outside Member, MAT/IFG

Prof. Eraldo Sérgio Barbosa Da Silva, PhD
Outside Member, ECO/UFSC

José Augusto Fiorucci, PhD
Additional Inside Member, EST/UnB

"My life seemed to be a series of events and accidents.

Yet when I look back, I see a pattern."

Benoît B. Mandelbrot

Dedico aos meus familiares que me apoiaram ao longo de toda a minha jornada educacional. Em especial, dedico aos meus pais que sempre fizeram de tudo para que não faltasse o essencial, e aos meus irmãos e animais que sempre formaram minha rede de proteção. Por último, dedico em memória daqueles que nos deixaram, Iramy Rodrigues Gomes, Walmir de Almeida Gomes e Benoît B. Mandelbrot.

Acknowledgments

Thanks to the faculty of PPGEST/UnB for all the knowledge taught to me throughout this master's degree. Thanks to the Department of Statistics for understanding the need to postpone the presentation of this dissertation. Thanks to Professor Raul Yukihiro Matsushita for accepting to participate in the development of this dissertation, for collaborating throughout the project and for always showing enthusiasm for the challenges. Finally, thanks to Guilherme Maia Rodrigues Gomes for being responsible for the python implementation of the MMAR model and providing several insights about the proposed model.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Resumo Expandido

Um estrutura Multi-Armed Bandit para alocação em portfólio.

Palavras-Chave. Fronteira Multi-Armada, Modelo Multifractal de Retornos de Ativo, Limite de Confiança Superior Tunado, Teoria Moderna de Portfólio, Teoria de Portfólio de Dominância Estocástica Axiomática de Segunda Ordem, Método de Dupla-Barreira

Keywords. Multi-Armed Bandit, Multifractal Model of Asset in Return, Upper Confidence Bound - Tuned, Modern Portfolio Theory, Axiomatic Second-order Stochastic Dominance Portfolio Theory, Double Barrier Method

Introdução

Há mais de um século, a comunidade acadêmica estuda o mercado financeiro na tentativa de entender seu comportamento para maximizar os lucros. Este trabalho procura maneiras de maximizar os resultados no mercado financeiro criando um procedimento de duas fases que chamamos de MAB-MMAR.

Primeiro, estabelece-se modelos generativos individuais para cada ativo, para simular, via Monte Carlo, retornos futuros, usando *Multifractal Model of Asset Returns* (Mandelbrot, Fisher, and Calvet, 1997) que é capaz de multiescalar os momentos da distribuição de retorno sob escalas temporais, sendo uma alternativa às representações do tipo ARCH que tem sido o foco de pesquisas empíricas sobre a distribuição de preços nos últimos anos.

Em segundo lugar, constrói-se uma estrutura de *Multi-Armed Bandit* (MAB) aplicando o algorítimo *Upper Confidence Bound (UCB)-Tuned* (Kuleshov and Precup, 2014) sobre os caminhos simulados, a fim de realizar escolhas entre ativos que otimizem a alocação de recursos.

Além disso, como camada de proteção para as operações, propomos o Método da Dupla Barreira, onde a operação é encerrada se uma barreira inferior for tocada.

Objetivo

Consideramos o problema de alocação de portfólio para um dado orçamento η. Seja X_{it} o retorno associado ao ativo i de um dado conjunto de ativos $A = [1, ..., a]$ no tempo t onde $t \in \{1, ..., T\}$. E definimos α_{it} como a proporção de η que será alocada ao ativo i no tempo t. Assim, nosso objetivo, ao propor o MAB-MMAR é encontrar $\tilde{\alpha}_t = [\tilde{\alpha}_{it} : i = 1, ..., a]$ tal que maximizemos o retorno geral no tempo t, ou seja,

$$\hat{\alpha}_t = \arg\max_{\alpha_t} \sum_{i=1}^{a} \alpha_{it} X_{it} \tag{1}$$

O primeiro desafio surge porque observamos retornos apenas até $t - h$ no tempo t. Assim, X_{it} na Eq (1) precisa ser estimado usando $X^*_{it-h} := \{X_{it-h}, ..., X_{i1}\}$, ou seja,

$$\widehat{\boldsymbol{\alpha}}_t = \arg\max_{\alpha_t} \sum_{i=1}^{a} \tag{2}$$

$$\text{onde } \widehat{X}_{it} = f\left(X^*_{it-h}\right) \tag{3}$$

Neste trabalho, definimos uma função $f(\cdot)$ na Eq (3) usando uma abordagem Multi-Armed Bandit (MAB) onde modelos generativos individuais serão considerados para compor nosso conjunto de braços via *Upper Confidence Bound - Tuned* (UCB-Tuned) (Kuleshov and Precup, 2014). A maior parte do trabalho nesta área concentra-se em medir a incerteza associada aos retornos das ações no tempo $t + h$, enquanto em nossa abordagem, tendemos a criar uma configuração mais poderosa, incluindo a incerteza associada a cada modelo generativo individual sobre o caminho h na estrutura MAB.

Um dos principais desafios dessa perspectiva é como estimar com confiança a incerteza associada a cada modelo generativo individual. Isso é especialmente importante nos dias de hoje com a enorme quantidade de modelos de caixa-preta que surgiram em todos os campos. No geral, este trabalho visa derivar o que chamamos de mapeamento quantico de incerteza através de simulações geradas por um modelo generativo não tão popular: *Multifractal Model of Asset Returns* - MMAR (Mandelbrot, Fisher, and Calvet, 1997).

Além disso, como segundo desafio deste trabalho, propomos um Método de Dupla Barreira baseado no Método de Tripla Barreira (De Prado, 2018) onde consideramos níveis de um limite inferior (stop loss) para cada ativo i durante o caminho abrangendo $[t_{i0}, t_{ih}]$. Se um nível inferior for tocado para o ativo i antes do final do caminho, o ativo i sai do portfólio. Múltiplas barreiras inferiores podem ser testadas.

Metodologia

MAB-MMMAR pode ser definido como um método em duas etapas:

1. Estima-se modelos generativos individuais $I_{i,k}$ para cada ativo i em cada tempo k. Simula-se caminhos de tamanho h usando Monte-Carlo para fornecer uma previsão de retorno $\hat{\mu}_{i,l}$ e sua variância $\hat{\sigma}^2_{i,l}$ ao longo do caminho $[k+1, k+h]$ para qualquer $l \in [k+1, k+h]$;

2. Propõe-se um algoritmo MAB M que usa os resultados fornecidos na primeira etapa para fazer escolhas sob incerteza ao longo do caminho $[k+1, k+h]$ selecionando os ativos a serem investido no tempo k até $r+h$. Os ativos que forem selecionados ao final dos p últimos tempos em $[k+1, k+h]$ são aqueles utilizados para investimento, enquanto que a intensidade de seleção é utilizada para proporcionalmente dividir o orçamento.

Como comparação de desempenho, foram testados os modelos *One-Asset* que descreve-se como sendo um investimento pautado na escolha de um único ativo; $1/n$ que descreve-se como sendo um investimento pautado em alocações de igual tamanho entre ativos; Já as metodologias *Modern Portfolio Theory* (MPT) (Markowitz, 1952) e *Axiomatic Second-order Stochastic Dominance Portfolio Theory (ASSDPT)* (Ruszczynski and Vanderbei (2003)), por possuírem uma matemática mais rebuscada, podem ser consultadas em suas respectivas referências.

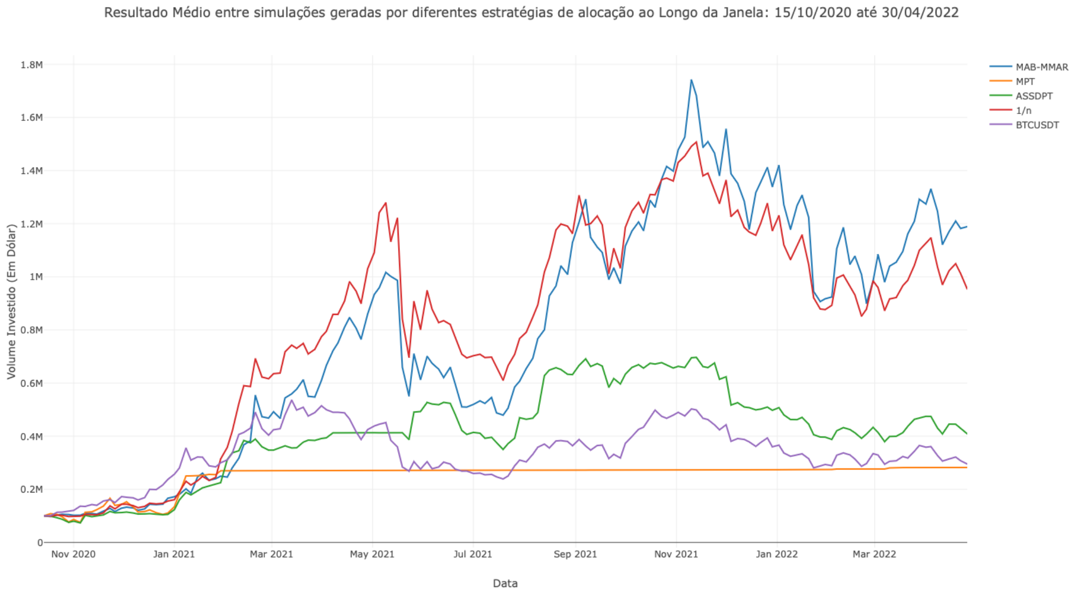
Experimentos

O conjunto de dados é composto pelos preços de fechamento de 15 criptomoedas, em janelas de 5 minutos, variando de 21/08/2020 10:04:59.999 a 30/04/2022 23:59:59.999. Foi coletado do github da binance (`https://github.com/binance/binance-public-data`). Todas as criptomoedas selecionadas já estavam sendo negociadas durante toda a janela de tempo do conjunto de dados.

Usamos o Tether (USDT) como nossa moeda base, pois é um *stablecoin* com paridade em dólares americanos. Além disso, o Tether (USDT) tem a vantagem de ser uma moeda com acesso direto a muitas outras criptomoedas, o que permite uma liquidez mais fácil e direta.

Para cada Estratégia de Alocação de Portfólio, consideramos $h = 1024$, ou seja, cada abrangência de caminho de alocação comprime 5120 minutos, que é 1024×5 já que temos dados de 5 minutos. Então, utilizamos todo o conjunto de dados para detalhar as janelas de de investimentos por trimestre. Dessa forma, avaliamos se os efeitos de alocação têm resultados satisfatórios independentemente do momento de início do procedimento, trazendo robustez em relação a resultados que são dependente do caminho; Parametrizamos todos os modelos que podem ser parametrizados. Desta forma, separamos os modelos por quão arriscados eles são; Aplicamos o Método da Dupla Barreira a todas as Fronteiras Propostas. Desta forma, podemos avaliar se um procedimento com barreira pode superar o não uso da mesma; Repetimos o procedimento por cada estratégia por 4 vezes. Desta forma, podemos produzir resultados mais robustos.

Resultados e Discussão



Todos os modelos, exceto o MPT, não tiveram dificuldade em se afastar do eixo de investimento inicial. Observando-se o comportamento gerado pelos métodos 1/n e One Asset, nota-se que a metodologia 1/n obteve resultados melhores. Já incluíndo ASSDPT e MAB-MMAR na análise, observa-se que o MAB-MMAR foi quem, em geral, obteve o melhor desempenho.

Portanto, temos que o modelo MAB-MMAR se mostrou viável para alocação em portfólio, permitindo realizar operações com um nível de confiança tunável acerca de se obter resultados não-negativos e apresentando vantagens nítidas em relação a outros procedimentos de alocação de recursos presentes na literatura.

Conclusões

Nossos resultados são promissores, onde se revela que entre os modelos utilizados como base de comparação, o MAB-MMAR foi o que melhor desempenhou nos mais variados cenários experimentados.

Referências

KULESHOV, V.; PRECUP, D. Algorithms for multi-armed bandit problems. arXiv preprint arXiv:1402.6028, 2014.

MANDELBROT, B. B.; FISHER, A. J.; CALVET, L. E. A multifractal model of asset returns. Cowles Foundation discussion paper, 1997.

MARKOWITZ, H. Portfolio selection, journal of inance. Markowitz HM—1952.—No, p. 77–91, 1952.

PRADO, M. L. D. Advances in inancial machine learning. [S.l.]: John Wiley Sons, 2018.

RUSZCZYNSKI, A.; VANDERBEI, R. J. Frontiers of stochastically nondominated portfolios. Econometrica, Wiley Online Library, v. 71, n. 4, p. 1287–1297, 2003.

Abstract

For over a century, the academic community has studied the financial market in an attempt to understand its behavior to maximize profits. This work looks for ways to maximize results in the financial market by creating a two-phase procedure that we call MAB-MMAR. First, individual generative models are established for each asset, to simulate, via Monte Carlo, future returns, using the Multifractal Model of Asset Returns, which is able to multiscale the moments of the return distribution under time scales, being an alternative to representations of the ARCH type, which are the representations that have been the focus of empirical research on the distribution of prices in recent years. Second, a Multi-Armed Bandit (MAB) structure is built by applying the Upper Confidence Bound (UCB)-Tuned algorithm on the simulated paths, in order to make choices between assets that optimize the allocation of resources. Furthermore, as a layer of protection for operations, we propose the Double Barrier Method, where the operation is terminated if a lower barrier is touched. As a performance comparison, the One-Asset, 1/n, Modern Portfolio Theory (MPT) and Axiomatic Second-order Stochastic Dominance Portfolio Theory (ASSDPT) models were tested. Our results are promising, revealing that, in general, the MAB-MMAR performed best in the most varied scenarios.

Keywords. Multi-Armed Bandit, Multifractal Model of Asset in Return, Upper Confidence Bound - Tuned, Modern Portfolio Theory, Axiomatic Second-order Stochastic Dominance Portfolio Theory, Double Barrier Method

Contents

List of Tables

List of Figures

Abbreviations and Acronyms

ARCH	Autoregressive Conditional Heteroskedasticity
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
FIGARCH	Fractional Integrated Generalized Autoregressive Conditional Heteroskedasticity
MMAR	Multifractal Model of Asset Returns
MPT	Modern Portfolio Theory
ASSDPT	Axiomatic Second-order Stochastic Dominance Portfolio Theory
MAB	Multi-Armed Bandit
UCB	Upper Confidence Bound
MAB-MMAR	Multi-Armed Bandit - Multifractal Model of Asset Returns

Chapter 1

Introduction

Financial market is the name of the entire universe that involves the purchase and sale of financial assets, where these assets can be securities, commodities or currencies. It first appeared in 1611 with the Dutch East India Company in Amsterdam, which was the first public company for commercial exchanges, well known for its involvement in the slavery industry (Vink (2003)). With years of market evolution, the format has changed a lot, where more assets were included, more companies and currencies started to be traded on a global level, leading to a much higher volume of operations.

This Market, which influences humanity in several themes like social aspects, such as an increase in inequality (Claessens and Perotti (2007)), or economics, with economic growth (Wong and Zhou (2011)), or technology, with its advances (Hsu, Tian, and Xu (2014)), is governed merely by future expectations, where the pricing of asset ends up being the result of a collection of future expectations regarding it and its relation with the environment. Several ways of trying to control this future expectation in the market have been tried, such as applying forecast models that seek to estimate the future return in finance (for clarification, we will call them as **Generative Models**), or applying optimization models that seeks to improve the choice when resource allocating (also, for clarification, we will call them as **Portfolio Allocation Strategies**).

Nowadays, much work have already been done when it comes to forecasting financial times series, such as development of the class of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models commonly used for forecast changes in the volatility, or such as the class of multifractal models, in special, the Multifractal Model of Asset Returns (MMAR) which combines a fractional brownian motion with an stochastic trading time to provide estimates of the return.

When it comes to the portfolio allocation problem, it was first studied by Harold Markowitz, in 1952, creating the Modern Portfolio Theory (MPT) which uses risk, return and correlation measure based on the expected values to allocate resources, however the risk measure is given by the variance of the actual return where its use may suggest a portfolio which is always outperformed by another portfolio (Ruszczynski and Vanderbei, 2003). In contrast, the Axiomatic Second-Order Stochastic Dominance Portfolio Theory (ASSDPT) improves the mean-risk approach proposed by Markowitz by including the notion of Stochastic Dominance: a random variable Z is dominated by Y if $[U(Y)] \geq [U(Z)]$ for all non decreasing concave function $U(.)$ which these expected values are finite (Ruszczynski and Vanderbei, 2003). It allows a decision maker prefer a portfolio Y over a portfolio Z, but it is computationally very difficult (Ruszczynski and Vanderbei, 2003).

When it comes to machine learning field, many portfolio allocation strategies have been proposed using a Multi-Armed Bandit (MAB) Framework (Huo and Fu, 2017, Zhu et al., 2019, Hoffman, Brochu, De Freitas, et al., 2011). In a classical Multi-Armed Bandit Problem, there is an agent with access to several slots machines and he needs to selects an arm to receive a random reward drawn i.i.d. from the unknown distribution of reward to the selected arm. Then, an MAB algorithm design a way to select the arm to be played in order to maximize rewards over time. The use of this structure seems to be promising for a problem that we intend to solve: Would it be feasible to implement a MAB strategy to select the assets to be included in a portfolio by defining the arms as predictions simulated by individual generative models for each asset? In this sense, this work aims to build a MAB framework capable of making decisions under uncertainty to select the basket of assets expected to bring positive returns with less variability, i.e., with less risk.

The documentation is divided into 6 chapters. Chapter 2 presents a literary review of forecast models, MAB algorithms and portfolio management models. Chapter 3 states the definitions around the problems we want to solve. Chapter 4 presents the methodolgy behind our setup. Chapter 5 presents our experiments showing our experimental setup and discusses the results. Chapter 6 concludes the dissertation providing final considerations.

Chapter 2

Literature Review

This section presents a literature review of the development of the fields of the forecast models, the portfolio allocation models and the multi-armed bandit models. It is presented their advances and some points that must be taken into account for the construction of the MAB framework.

2.1 Forecast Models

For more than a century, forecasting financial time series has been studied from the academic point of view to the industry one (Mandelbrot and Hudson (2010)). Since its first probabilistic formalization by Bachelier (Bachelier (1900)) based on the premise of the independent and gaussian distribution of price changes, it has been developed by several extensions such as the family of ARCH/GARCH models (Mandelbrot, Fisher, and Calvet (1997)).

Over the decades, empirical results began to show something else. First, the assumption of independence was disagreed since price changes commonly display temporal dependence with long-memory. Second, the assumption of gaussian distribution in price changes was also dropped since the empirical price changes usually show a much fatter tail on their histogram than the predicted by a normal distribution (Mandelbrot, Fisher, and Calvet (1997)).

In order to introduce temporal long memory on the analysis, the FIGARCH model, a set of infinite-dimensional restrictions upon its ARCH parameters (Günay (2016)), was proposed and achieve its objective parsimoniously (Mandelbrot, Fisher, and Calvet (1997)), despite to be scale-inconsistent (Günay (2016)). To overcome the scale-inconsistency, one could use the weak-GARCH class of models, however, this class is too restrict assuming that only the parameters (α, β) are the best linear predictors for the model.

In the search to deal with the challenges found by the classic literature in the study of financial returns, Benoit Mandelbrot suggests a Multifractal Model of Asset Returns (MMAR) based on the concept of multifractality where, as in FIGARCH, martingale properties with long memory are combined. Although there are some other models to analyse return from a multifractal prospective such as the class of Markov Switching Multifractal (MSM) models, or the Multifractal Random Walk (MRW) model (Jiang et al. (2019)), this work handles multifractality only on the prospective of MMAR. The MMAR stands out in comparison to the classic models for dealing with the problem of long memory by considering the return signal to be fractal brownian motion like; for allowing that long tails in the distribution of returns can happen; and for incorporating the trading time property, responsible for relating observable time to unobservable natural time (Mandelbrot, Fisher, and Calvet (1997)). Also, the MMAR, which is build on the top of *self-affine* processes, is scale-consistent since it defines scaling rules to relate returns over different sampling intervals (Mandelbrot, Fisher, and Calvet (1997)).

As an illustration about the empirical behaviour of a price changes serie, a daily closing price serie of Bovespa Index (IBOV) ranging from 2000-01-01 to 2005-12-31 was studied and summarized on Figure 2.1 where 2.1(a) is the actual price time serie of the Bovespa Index; 2.1(b) is the histogram of the daily Returns where Return is defined as $X(t) = log(P(t)) - log(P(t-1))$; 2.1(c) is the time serie of the normalized price changes; and, 2.1(d) shows the ACF of Bovespa Index Price serie. The range was selected in order to be the same chosen on Gaio and Sáfadi (2008), not because of any interpretations they stated about their results, but merely for didactic purposes, since they had already estimated the values of the GARCH and FIGARCH parameters handling the Bovespa Index serie on this extent which are useful on generating simulated measure paths for comparison (Figure 2.2) with the actual path (Figure 2.1(c)).

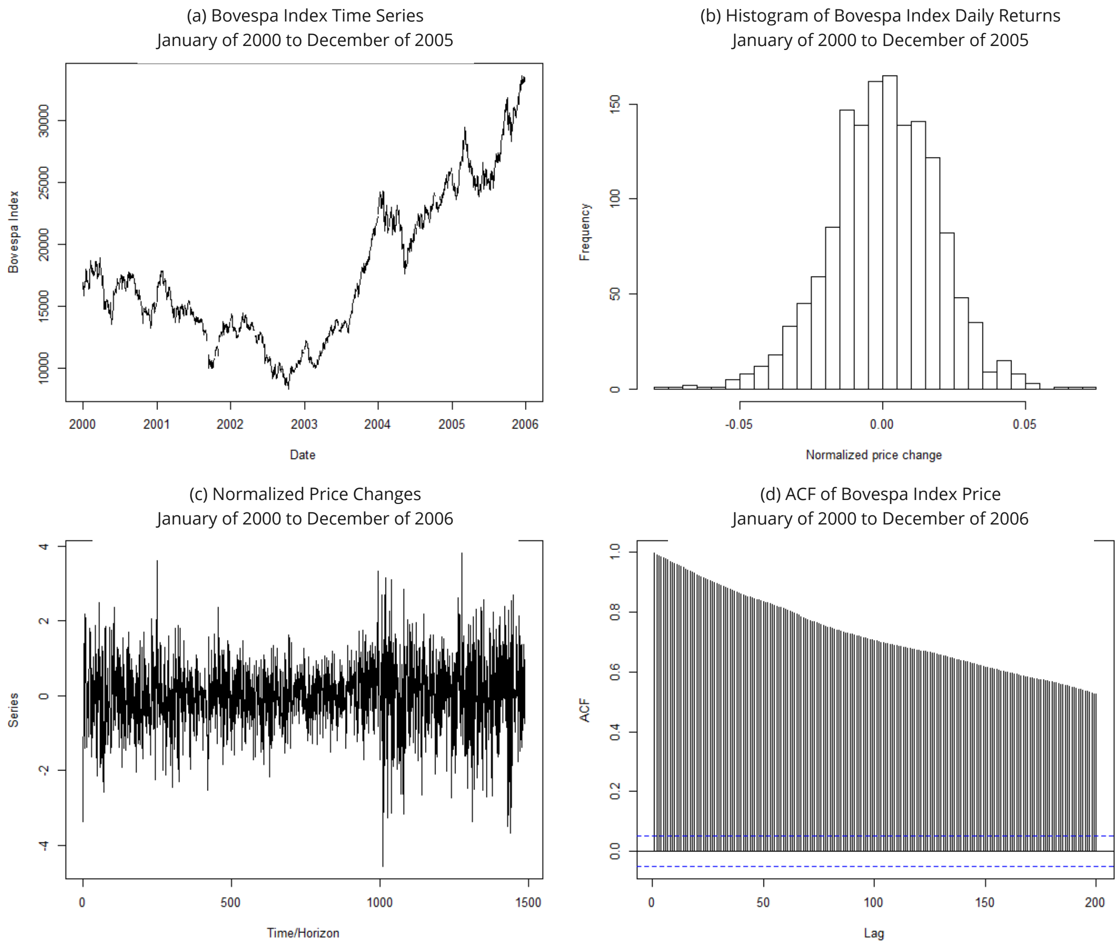


Figure 2.1: Bovespa Analysis. (a) The Actual Time Series of Bovespa Index from January of 2000 to December of 2005. There is evidence of small and large oscillations occurring in cycles. (b) Histogram of Bovespa Index Daily Returns. There is evidence of fat-tail on the distribution of Returns. (c) Time series of the Normalized Price Changes. Also, shows evidence of periods of large and small fluctuations which is different from the white noise behaviour. (d) ACF of Bovespa Index Price. Reveals an autocorrelation that decays slowly, regarding the long memory structure of a financial time serie.

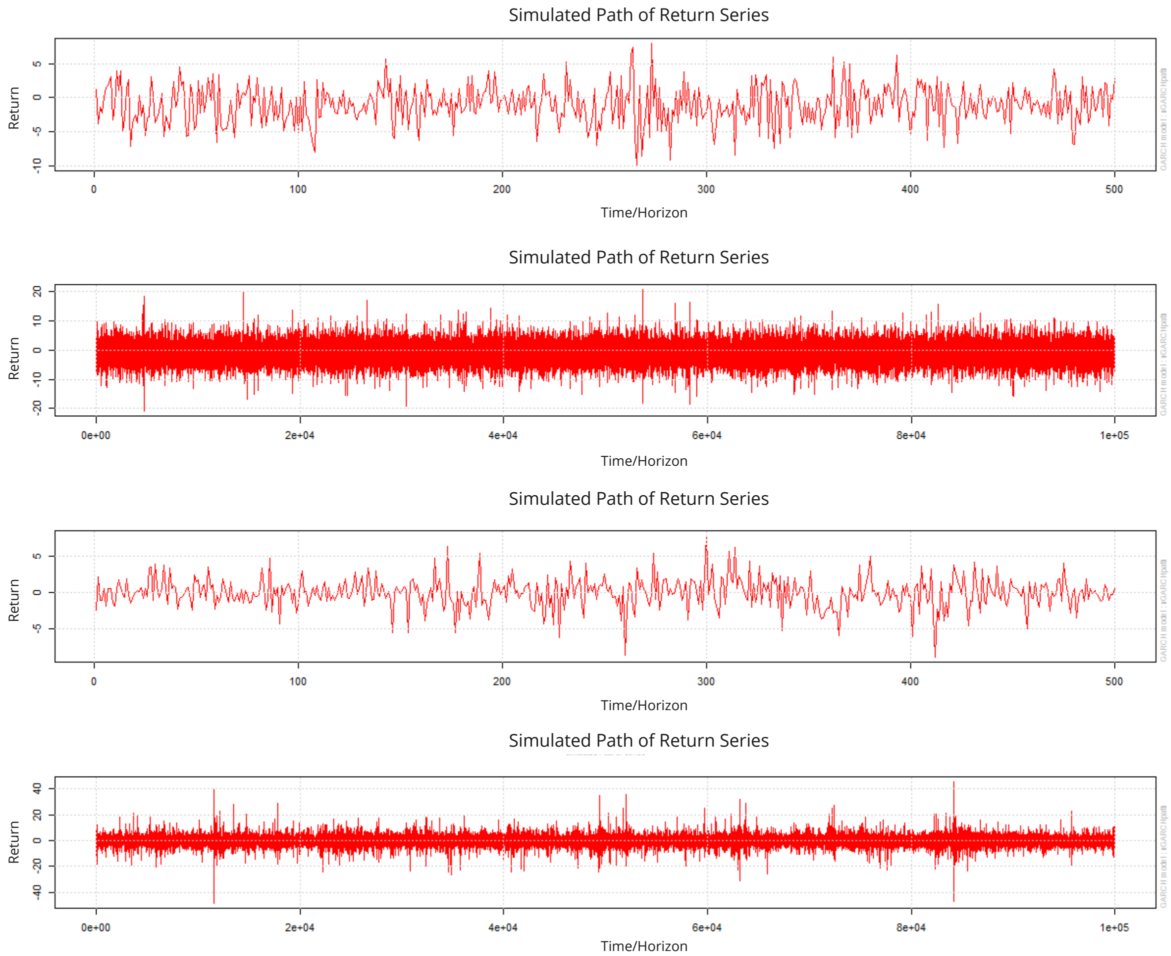


Figure 2.2: Garch Comparison. (From top to bottom) The first one is a simulated path of size = 500 from GARCH(1,1) using $(\alpha = 0.174, \beta = 0.663)$(Gaio and Sáfadi (2008)), while the second one is a simulated path using the same model but for size = 100000, showing that there is some proportionally wild fluctuation in the smaller sample, however when increasing the sample size, the wildness shows to be not too wild being better described as white noise, touching a variation of 20 standard deviations in only two occasions. (From top to bottom) The last two charts are shown simulated paths from the FIGARCH(1,d,1) using $(\alpha = -0.075, \beta = 0.169, d = 0.241)$(Gaio and Sáfadi (2008)) with sizes of 500 and 100000 steps, respectively. On the short-run, the FIGARCH model simulated path has shown to be as wild as the GARCH one with the benefit of being less random, while on the long-run, the FIGARCH gets much better than the vanilla GARCH, showing values with deviations higher than 40 standard deviations and clusters of volatility with periods with small and large oscillations which is better described as fractional gaussian noise.

2.2 Multi-Armed Bandit (MAB) Algorithms

The Multi-Armed Bandit algorithms are models that seek to choose among a limited number of resources, the ones that would bring the highest gain by exploring the trade-off between exploration and exploitation. In an exploration phase, the a multi-armed bandit model tries to learn knowledge the environment, while in an exploitation phase, the model tries to optimize its decisions based on its accumulated knowledge.

It was first introduced by Robbins when he proposes a sequential design of experiments (Robbins (1952)). In his paper, he discuss that the design of sequential experiments should not have a fixed value for sample size or a fixed composition. Instead, it should be a function of the knowledge acquired while the sequential experiment happens.

In a multi-armed bandit model, we just need to have a set of G probabilities distributions $\{D_1, ..., D_G\}$ and their corresponding expected values $\{\mu_1, ..., \mu_G\}$ and variances $\{\sigma_1^2, ..., \sigma_G^2\}$, so, in a financial prospective, we could could think in at least two cases where it looks to handle uncertainty:

- **Ensemble of Forecast Models** Choose among a basket of forecast models, the ones that would bring the highest gain;
- **Portolio Management** Choose among a basket of stock assets, the ones to invest (Portfolio Management).

A generic multi-armed bandit algorithm starts with the definition that D_i is initially unknown to the system. At every next step in $t = 1, 2, ...,$, the system selects an element (could be a model, or an asset) with index $j(t)$ and receives a reward $r(t)\ D_{j(t)}$. Then the system needs to gain as much rewards as possible while the time goes by, and also needs to find out how the distributions of rewards are.

There are several algorithms to handle a multi-armed bandit problem. In this work, we will discuss the $\epsilon -$ **greedy** algorithm, the **Boltzmann Exploration** algorithm and the **Upper Con-**

fidence Bounds (UCB) algorithm. Those algorithms can have theirs performances evaluated through the *total expected regret* which is defined for any fixed turn T as:

$$R_T = T\mu^* - \sum_{t=1}^{T} \mu_{j(t)} \tag{2.1}$$

where $\mu^* = \max_{i=1,\dots,K} \mu_i$ is the expected reward from the best element. Can also be written as:

$$R_T = T\mu^* - \mu_{j(t)} \sum_{k=1}^{K} \mathbb{E}(T_k(T)) \tag{2.2}$$

where $T_k(T)$ defines a random variable that expresses the number of times the element k was chosen during the first T turns. Using the result given by (**empty citation**), it is possible to states that for any suboptimal element k:

$$\mathbb{E}(T_k(T)) \geq \frac{\ln T}{D(p_k||p^*)} \tag{2.3}$$

where $D(p_k||p^*)$ is the Kullback-Leibler Divergence between the reward density p_k of the sub-optimal element and the reward density p^* of the optimal element, defined as:

$$D(p_k||p^*) = p_j \int \ln \frac{p_j}{p^*}. \tag{2.4}$$

2.3 Portfolio Management Models

The history of portfolio theory starts in 1952 with Harry Markowitz publication *Portfolio Selection* (Robbins (1952)). In his work, he proposes that the volatility of a stock portfolio can be compared to a broader market and can be measure by a β variable. By this way, one could choose a portfolio based on how volatile the market can be. However, it has some sources on fragility since the model assumes prior knowledge of the parameters (Taleb (2012)). Also, he assumes that bias ω_A and fragility ω_b are both equal to zero (Taleb (2012)). Fragility is a mea-

sure proposed heuristically through Nassim Taleb's work *A New Heuristic Measure of Fragility and Tail Risks: Application to Stress Testing* (Taleb et al. (2012))
Over time, more models were being developed with different approaches. It was created a multicriteria model (Costa and Soares* (2004)) which uses concepts from operations Research, also, it was published the paper *Incorporating Markov Decision Process on Genetic Algorithms to Formulate Trading Strategies for Stock Markets* (Chang and Lee (2017)) that uses Markov Decision Processes, and published *Event-Based Optimization for POMDPs and Its Application in Portfolio Management* (Wang and Cao (2011)) and *An Approximate Solution Method for Large Risk-Averse Markov Decision Processes* (Petrik and Subramanian (2012)) that both use Partially Observed Markov Decision Processes to solve the portfolio allocation Problem. From a deep learning prospective, *A Deep Reinforcement Learning FrameWork for the Financial Porfolio Management Problem* provides a framework using Ensemble of Identical Independent Evaluators (EIIE) topology. Under a MAB framework, the publication *Risk-aware multi-armed bandit problem with application to portfolio selection* (Huo and Fu (2017)) incorporates risk but using value-at-risk as a risk measure.

Chapter 3

Problem Definition

We consider the problem of portfolio allocation for a given budget η. Let X_{it} be the return associated with asset i of a given pool of assets $A = [1, ..., a]$ at time t where $t \in \{1, ..., T\}$. Now, define α_{it} as the proportion of η that will be allocated to asset i at time t. Thus, we aim to find $\underset{\sim}{\alpha}_t = [\underset{\sim}{\alpha}_{it} : i = 1, ..., a]$ such that maximizes the overall return at time t, i.e.

$$\widehat{\boldsymbol{\alpha}}_t = \arg\max_{\alpha_t} \sum_{i=1}^{a} \alpha_{it} X_{it} \tag{3.1}$$

The first challenge arise since we observe returns only until $t - h$ at time t. Thus, X_{it} in Eq (3.1) needs to be estimated using $X^*_{it-h} := \{X_{it-h}, ..., X_{i1}\}$, i.e.

$$\widehat{\boldsymbol{\alpha}}_t = \arg\max_{\alpha_t} \sum_{i=1}^{a} \alpha_{it} \widehat{X}_{it}(h) \tag{3.2}$$

$$\text{where } \widehat{X}_{it} = f\left(X^*_{it-h}\right) \tag{3.3}$$

In this work, we define a function $f(\cdot)$ in Eq (3.3) using a Multi-Armed Bandit approach where individual generative models will be consider in order to compose our set of of arms. Most of the work in this area, focus on measuring the uncertainty associated with the stock returns at time $t + h$, while in our approach, we tend to create a more powerful setting by including the uncertainty associated with each individual generative model over the path h.

One of the main challenges of this perspective is how to estimate with confidence the uncertainty associated with each individual generative model. This is specially important in these days with the enormous amount of black-box models that have been arising across all fields. Overall, this work aims to derive what we called an uncertainty map quantification across simulations generated by a not so popular generative model: Multifractal Model of Asset Returns.

Also, as a second challenge of this work, we propose a Double-Barrier Method based on the Triple-Barrier Method (De Prado, 2018) where we consider levels of a lower bound (stop loss) for each asset i during the path spanning $[t_{i0}, t_{ih}]$. If a lower level is touched for asset i before the end of the path, asset i leaves the portfolio. Multiple lower barriers can be tested.

3.1 Questions we want to answer

1. Is it feasible to implement a multi-armed strategy to allocate assets in a portfolio using predictions made by a generative model?

2. Using a multi-armed bandit framework, at the end of each trial, can we select assets with non-negative returns with a high level of statistical significance?

3. Does the multi-armed framework bring better advantages against other portfolio allocation models presented in literature?

4. Does the Double-Barrier Method bring better advantages against using a non-hedge model?

3.2 Proposed Solution

This work aims to build a two-step procedure:

- **First Step.** Estimate individual generative models $I_{i,k}$ for each asset i at each time k. Simulate paths of size h using Monte-Carlo to provide a prediction of return $\hat{\mu}_{i,l}$ and its variance $\hat{\sigma}^2_{i,l}$ at each time in a path spanning $[k+1, k+h]$ where l is inside the path spanning;

- **Second Step.** Proposes a multi-armed algorithm M that uses the results provided in first step to make choices under uncertainty over the path spanning $[k+1, k+h]$ to select the assets to be invested at time k until $r+h$.

3.2.1 Generative Model

- **MMAR.** The MMAR defines a compound process $X(t)$ such as (Mandelbrot, Fisher, and Calvet (1997)):

$$X(t) \equiv B_H[\theta(t)] \tag{3.4}$$

  with the following assumptions:

  - **Assumption I.** $\theta(t)$ is the stochastic trading time while $B_H(t)$ is a fractional Brownian motion.
  - **Assumption II.** $\theta(t)$ and $B_H(t)$ are independent with each other.
  - **Assumption III.** Trading time $\theta(t)$ is the cumulative density function of a multifractal measure μ on an interval $[0, T]$.

  Theorem. *Under Assumptions [I] - [III], the process $X(t)$ is multifractal, with scaling function $\tau_X(q) \equiv \tau_\theta(Hq)$ and stationary increments* (Proof in Appendix A).

Motivation. MMAR stands out in comparison to the classic models for allowing that long tails in the distribution of returns can happen; for incorporating the trading time property, responsible for relating observable time to unobservable natural time (Mandelbrot, Fisher, and Calvet (1997)); and, for dealing with the problem of long memory by considering the return signal to be fractal brownian motion like, which unlike standard Brownian motions allows for dependence on increments. A standard brownian motion is a random process $X = \{X_t : t \in [0, \infty)\}$ with state space $\mathbb{R}$ satisfying the following properties: $X_0 = 0$ with probability 1; X has stationary and independent increments; X_t is normally distributed; and, $t \to X_t$ is continuous on $[0, \infty)$. Also, the MMAR, which is build on the top of *self-affine* processes, is scale-consistent since it defines scaling rules to relate returns over different sampling intervals (Mandelbrot, Fisher, and Calvet (1997)). Properties of the MMAR and their proofs can be seen on appendix B.

3.2.2 Multi-Armed Algorithm

- **Upper Confidence Bound Tuned (UCB1-Tuned)** is a multi-armed without any theoretical guarantees, however the authors claim it performs better in practice than UCB1 (another multi-armed algorithm, but with theoretical guarantees based on a sophisticated math). They were both proposed by the same authors, and the main feature of UCB1-Tuned against UCB1 is because it takes into account the variance of each arm (Kuleshov and Precup, 2014). Specifically, at time $l = 1, 2, ..., h$, the algorithm selects the an arm $j(l)$ as:

$$j(l) = \underset{i=1...a}{\arg\max} \left(\hat{\mu}_i + \sqrt{\frac{\ln l}{n_i} \min\left(\frac{1}{4}, V_i(n_i)\right)} \right) \tag{3.5}$$

where $i \in A = [1, ..., a]$ and,

$$V_i(l) = \hat{\sigma}_i^2(l) + \sqrt{\frac{2\ln l}{n_i(l)}}. \tag{3.6}$$

Motivation. This MAB strategy brings variance inside formulation. In this way, we can use the variance for each step in the simulated path spanning to make decisions under the uncertainty of the model in relation to asset i at time l.

Chapter 4

Methodology

First, we present the proposed model for portfolio allocation, MAB-MMAR. Second, we present the Double-Barrier Method and its tested lower frontiers.

4.1 Portfolio Allocation Strategies

4.1.1 MAB-MMAR

MAB-MMAR defines that a return of a portfolio $R_k \in \mathbb{R}$ at time window k for a basket of assets a, will be given by:

$$R_k = \sum_{i=1}^{a} \alpha_{i,k} R_{i,k} \tag{4.1}$$

where $R_{i,k}$ is defined as:

$$R_{i,k} = \log\left(\frac{P_{i,k}}{P_{i,k-h}}\right) \text{ where } P_{i,s} \text{ means price of asset } i \text{ at time } s. \tag{4.2}$$

Since time is a continuous variable, we have that the set $[k-h, k]$ allows subdivisions where we can, for example, divide it into h smaller windows and consider them as pseudo-investment which is a windows that represent investment simulations performed in subdivisions of the time window to be invested. Therefore, if the investment will be made between the time t and $t+h$, where h = 10, we can generate 10 smaller time windows and simulate investments in each of them. What allows us to rewrite:

$$R_k = \sum_{j=1}^{h} \sum_{i=1}^{a} \alpha_{i,k-h+j} R_{i,k-h+j}. \tag{4.3}$$

Once we have observed the returns until $k - h$, we need to estimate the return values $R_{i,k-h+j}$ and their associated uncertainties from $k - h + 1$ to k referring to each asset i. This is done by Monte Carlo simulations generated from the generative process estimated by the MMAR (Eq (3.4)) for each asset i. The above suggests rewriting the previous equation in:

$$\hat{R}_k = \sum_{j=1}^{h} \sum_{i=1}^{a} \alpha_{i,k-h+j} \hat{R}_{i,k-h+j}. \tag{4.4}$$

To find the set $\hat{\alpha}_k$, we divide the window $[k - h, k]$ into 3 blocks. In order to have enough data to carry out the procedure applied to each of the blocks, the following proportions were used, which, as part of future work, may be fine-tuned: The first block includes the initial 25% of the window $[k - h, k - \frac{3}{4}h)$, being dedicated to the exploration of the estimated returns taking into account their uncertainties. The second block is dedicated to the exploitation phase, which are optimized choices, contemplating the window $[k - \frac{3}{4}h, k - \frac{1}{4}h)$. The last block $[k - \frac{1}{4}h, k]$ is also an exploitation period, but it is the only window directly used to estimate the set $\hat{\alpha}_k$. We restrict the estimation of α_k to the third block because we want to choose the assets that are converging the most in the final step of the path spanning $[k - h, k]$.

The uncertainty map quantification procedure starts assuming that during the whole first block $[k - h, k - \frac{3}{4}h)$, all assets were selected and, therefore, we start the choices of Eq (3.5) from block 2 $[k - \frac{3}{4}h, k - \frac{1}{4}h)$, assuming the allocation state left by block 1. The Eq (3.5) is also used in block 3 $[k - \frac{1}{4}h, k]$ from the results generated at the end of block 2.

Given that UCB-Tuned selects one arm at a time, and in our case we want to select a basket of assets, we evaluate the selections generated by UCB-Tuned in the last block $[k - \frac{1}{4}h, k]$ and we divided the allocations of each asset relative to the number of times it was selected on it. So, if a hypothetical asset i is responsible for 70% of the allocations in block 3, while another hypothetical asset i' is responsible for 30% of the allocations in block 3, this relationship will be used to estimate $\hat{\alpha}_k$ and, therefore, allocate the budget η between assets.

Our risk measure is given by the uncertainty of predictions generated using the simulations of each asset at time k. If the estimated probability $\hat{P}_k(i)$ of an asset i to have positive return at time k is lower than a threshold (risk measure) in a path spanning G, it will be not used in UCB-Tuned during G. In this sense, this strategy can generated non-invested path spannings if all the probabilities to all assets are lower than the risk threshold.

4.1.2 1/n

This strategy is the second simplest one. We allocate to each asset i, in every time t, an amount equal to $1/n$ of the corresponding budget η_t.

4.1.3 One Asset

This strategy is the simplest one. We only allocate to a specific asset i during the full time window. So, in every time t, the budget η_t will be whole invested in asset i.

4.1.4 Modern Portfolio Theory

MPT defines that a return of a portfolio R_t at time t can be expressed as:

$$R_t = \sum_i w_{i,t} R_{i,t}, \text{ where } w_{i,t} \in [0,1] \forall i, t \text{ and } \sum_i w_{i,t} = 1. \tag{4.5}$$

Since the original MPT is ambigous about the definition of Return, here we define $R_{i,t}$ such as:

$$R_{i,t} = \log\left(\frac{P_{i,t+h}}{P_{i,t}}\right) \text{ where } P_{i,t} \text{ means price of asset } i \text{ at time } t. \tag{4.6}$$

By linearity of expectation,

$$E[R_t] = \sum_i w_{i,t} E[R_{i,t}]. \tag{4.7}$$

So, if R is a vector of asset returns R_i, w a vector of weights w_i, and σ be the covariance matrix of vector R, then,

$$E[R_t] = R^T w \tag{4.8}$$

and,

$$\sigma_t^2 = w^T \sigma w. \tag{4.9}$$

The MPT defines an optimal portfolio by parameterizing an efficient frontier where, in practice it minimizes the following equation with respect to w:

$$\sigma_t^2 - qE[R_t] = w^T \sigma w - qR^T w \tag{4.10}$$

where $q \geq 0$ is the risk parameter. Greater values of q, means riskier portfolios.

4.1.5 Axiomatic Second-Order Stochastic Dominance Portfolio Theory

The ASSDPT (Ruszczynski and Vanderbei, 2003) defines that each portfolio allocation w has a mean return:

$$\mu(w) = E[R(w)] = \sum_{a=1}^{A} \sum_{t=1}^{T} r_{at} w_a p_t \tag{4.11}$$

where r_{at} is the return and p_t its attained probability, assuming that the returns have a discrete joint distribution.

Also, it defines that each portfolio allocation w has an associated $\rho(w)$ that represents the variability of the return $R(w)$, given by the absolute semideviation. The optimization can be formulated as:

$$\begin{gathered} \max \mu(w) - \lambda\rho(w), \\ \text{subject to } w \in W \end{gathered} \tag{4.12}$$

where λ is a nonnegative parameter representing the desirable exchange rate of mean for risk. So, it defines the model (μ, ρ) as consistent with Second-Order Stochastic Dominance (SSD) with coefficient $\alpha > 0$, if the relation is true:

$$R(w) \succeq_{SSD} R(y) \Rightarrow \mu(w) - \lambda\rho(w) \geq \mu(y) - \lambda\rho(y) \tag{4.13}$$

$\forall 0 \leq \lambda \leq \alpha$.

Then, it defines that the risk as the absolute semideviation:

$$\bar{\delta}(w) = E\{\max(\mu(w) - R(w), 0)\} \tag{4.14}$$

is consistent with the second-order stochastic dominance relation with coefficient 1 (Proof in Appendix C).

Next, since we have used the Python Monte Carlo Efficient Frontier (PyMCEF) package to estimate the allocations, it rewrites the procedure in a way that each portfolio on the efficient frontier is obtained by solving:

$$\begin{gathered} \arg\min_{w} \delta(w) - \lambda\mu(w) \\ \text{subject to } w \in W \end{gathered} \tag{4.15}$$

If the Lagrangier multiplier λ goes to zero, the solution is the least risky allocation, while if it goes to infinity, the solution is the most risky allocation. Since λ can vary to positive infinity, there is an optimization over a continuum structure.

The joint distribution of the returns can be parametric, alternatively, it is possible to work with finite samples of returns to simulate returns by Monte Carlo, and no knowledge about the underlying distribution would be needed. Then, its replace the expectation in δ or μ with statistical mean and work with a stochastic programming problem. With slack variables, it is possible to linearize the constraint and the objective function in a way that the solution is only valid for an interval of λ.

Then, considering K samples of $r : \{r^k\}_{k=1}^K$, it can be formulated as a linear programming problem, where the model, using absolute semideviation as a risk measure, takes the form:

$$\begin{aligned} \arg\min_{v_k, w_a, v_k} \frac{1}{K} \sum_{k=1}^{K} u_k - \lambda \sum_{a=1}^{A} w_a \bar{r}_a \\ \text{subject to } u_k + \sum_{a=1}^{A} w_a (r_a^k - \bar{r}_n) - v_k = 0, \forall k = 1, ..., K \\ \sum_{a=1}^{A} w_a = 1 \\ w_a, u_k, v_k \geq 0 \end{aligned} \tag{4.16}$$

where w_a are decision variables, u_k are auxiliary variables, while v_k are slack variables.

4.2 Double-Barrier Method

The Double-Barrier Method consists of ending the operation for each specific asset i during allocation in the path spanning [t,t+h] if a lower barrier is touched before the end of the operation window, i. e, before $t + h$. The following picture illustrates the procedure:

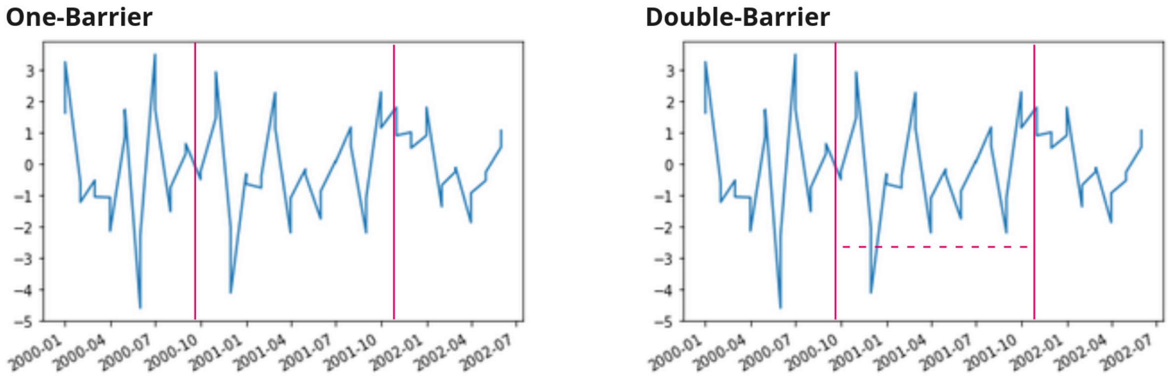


Figure 4.1: One-Barrier and Double-Barrier Methods. One-Barrier: The vertical lines in pink are limiting the investment horizon for an hypothetical asset. The investment starts in the first vertical line and ends in the second vertical line. Double-Barrier: Analogously to the One-Barrier method, it has vertical lines limiting the investment time window, however, it has, also, a horizontal band that defines a lower limit along the investment window, where, if the horizontal line is touched before the second vertical line, the investment will be ended earlier.

4.2.1 Tested Lower Frontiers

5% Level: If an asset i loses 5% in value between the path spanning [t, t+h], it will be sell assuming a loss of 5% in respect to asset i and the respectively path spanning (assuming enough liquity to be sold);

10% Level: If an asset i loses 10% in value between the path spanning [t, t+h], it will be sell assuming a loss of 10% in respect to asset i and the respectively path spanning (assuming enough liquity to be sold);

15% Level: If an asset i loses 15% in value between the path spanning [t, t+h], it will be sell assuming a loss of 15% in respect to asset i and the respectively path spanning (assuming enough liquity to be sold);

20% Level: If an asset i loses 20% in value between the path spanning [t, t+h], it will be sell assuming a loss of 20% in respect to asset i and the respectively path spanning (assuming enough liquity to be sold);

25% Level: If an asset i loses 25% in value between the path spanning [t, t+h], it will be sell assuming a loss of 25% in respect to asset i and the respectively path spanning (assuming enough liquity to be sold);

None: None of the assets will leave operation before the end of the path spanning [t, t+h].

Chapter 5

Experiments

5.1 Data

The dataset is composed by closing prices of 15 cryptocurrencies, in 5 minute widows, ranging from 2020-08-21 10:04:59.999 to 2022-04-30 23:59:59.999. It was collected from binance gitub (`https://github.com/binance/binance-public-data`). All crypto coins selected were already being negotiated during the whole dataset time window. We selected the following assets:

Asset	Parity
BCHUSDT	BITCOIN CASH / Tether (USDT)
BTCUSDT	BITCOIN / Tether (USDT)
XRPUSDT	XRP / Tether (USDT)
LUNAUSDT	LUNA / Tether (USDT)
XMRUSDT	XRP / Tether (USDT)
XLMUSDT	XLM / Tether (USDT)
USDCUSDT	USDC / Tether (USDT)
ETHUSDT	ETHERIUM / Tether (USDT)
BUSDUSDT	BUSD / Tether (USDT)
ADAUSDT	ADA / Tether (USDT)
DOTUSDT	DOT / Tether (USDT)
SOLUSDT	SOL / Tether (USDT)
DOGEUSDT	DOGE / Tether (USDT
BNBUSDT	BNB / Tether (USDT)
LTCUSDT	LTC / Tether (USDT)

Table 5.1: List of Cryto Coins used for Experimental Purposes.

We used Tether (USDT) as our baseline coin since it is a stabecoin with parity with US Dollars. It has lost parity sometimes, however, it was very stable during the whole path spanning covered by our dataset:

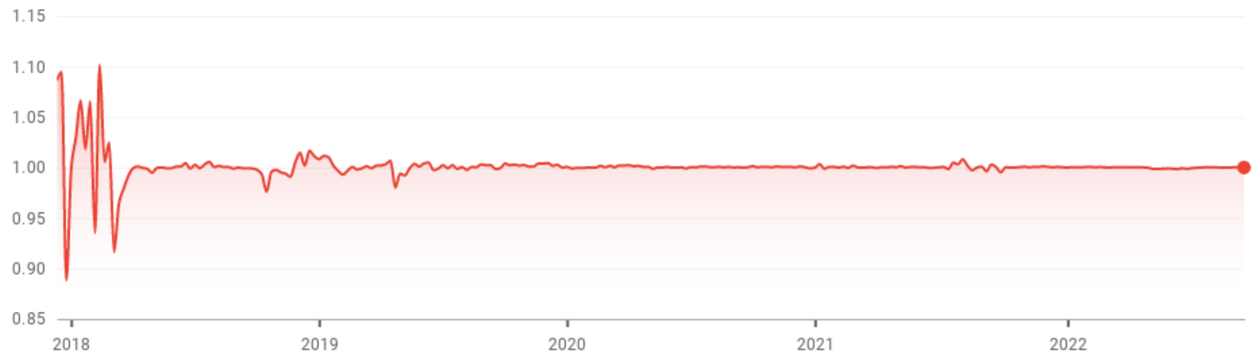


Figure 5.1: Historical of 5 years of USDT parity with US Dollars. Produced by Google Finance.

Also, Tether (USDT) has the advantage of being a coin with direct access to many other cryptocurrencies, which allows easier and direct liquity.

5.2 Experimental Setup

The experiments were carried out by dividing the Portfolio Allocation Strategies proposed in section 3 between the groups of models with low risky tunings γ_1, from the group of medium risky tunings γ_2, and from that group of models with high risky tunings γ_3. The 1/n strategy and the One Asset strategy are not parameterizable on the risk content, therefore, they appear in all groups. First, we present the general procedure applied to all tuned models. Second, we present the tuning procedure applied to each specific class of allocation model.

5.2.1 General Procedure

To each Portfolio Allocation Strategy, we consider an $h = 1024$, i.e., each allocation path spanning compresses 5120 minutes, which is 1024×5 since we have 5-minute data. Then,

1. **We used the whole dataset to break down investment initiation windows by quarter.** In this way, we assess whether the allocation effects have satisfactory results regardless of the time of initiation of the procedure, bringing robustness related to results that are path dependent;

2. **We parameterize all models that can be parameterized.** In this way, we separate the models by how risky they are;

3. **We apply the Double-Barrier Method to all Proposed Frontiers.** In this way, we are able to evaluate if an hedge procedure can overcome a non-hedge one;

4. **Repeat each simulation strategy in respect to each path spanning 4 times.** In this way, we are able to produce more robust results.

Formally speaking, we consider multiple series of returns $R_{a,t}$ where $a = 1, 2, ...$ represents an asset and $t = 0, 1, ..., T$ represents a time. Then,

1. Since the results can be path dependent, for experimental purposes, we create 5 path spanning scenarios: $[0, T], [j, T], [2j, T], [3j, T], [4j, T]$;

2. We define that the budget χ_t given by a model m at time t can be described as $\nu_m(R_{<t}, \chi_t)$;

3. We define that the budget χ_t given by a frontier f applied to a model m at time t can be described as $\tau_f(\nu_m(R_{<t}, \chi_t))$;

4. Given that for each model m, the historical of returns $R_{<t}$ generates allocation estimates, it is expected that when repeating the procedure, allocations change, therefore, for all path spannings, using all models in all boundary specifications, $\tau_f(\nu_m(.))$ was calculated, where, for all cases, the procedure was simulated 4 times.

5.2.2 Tuning Procedure

The models MAB-MMAR, MPT and ASSDPT are the ones that allow tuning of the risk measure. In MAB-MMAR framework, we define the risk measure as being the estimated probability $\hat{P}_k(a)$ of an asset a to have positive return at time k where the tunable parameter is the threshold ι that keeps the asset in the allocation procedure if $\hat{P}_k(a) \geq \iota$. In MPT, the tunable parameter is given by the q in the minimization of $w^T \sigma w - qR^T w$ subject to $w \in W$. While in ASSDPT, the tunable parameter is given by λ in $\arg\min_w \delta(w) - \lambda\mu(w)$ subject to $w \in W$.

Due to computational limitations, the tunings for the MAB-MMAR model were tested using ι equal to $[0.5, 0.53, 0.55]$, while for MPT, q is given from the set $[0, 2500, 5000]$, while for ASSDPT, λ is given from $[0, 10, 20]$.

Then, we defined the following experimental groups:

Group	Class	Tunable Parameter
	MAB-MMAR	0.55
Low Risk Tuning (γ_1)	MPT	0
	ASSDPT	0
	MAB-MMAR	0.53
Medium Risk Tuning (γ_2)	MPT	2500
	ASSDPT	10
	MAB-MMAR	0.50
High Risk Tuning (γ_3)	MPT	5000
	ASSDPT	20

5.3 Results

Given that for each model $\nu_m(R_{<t}, \chi_t)$, at each risk level γ_s for $s = 1, 2, 3$, at each path spanning $[gj, T]$ for $g = 0, 1, 2, 3, 4$, in each boundary specification $\tau_f(\nu_m(.))$ assuming an initial budget $\chi_0 = 100000$ dollars, the allocations were simulated in each path spanning 4 times, therefore, the average value of the budget $\bar{\chi}_{t,m,f,g,s}$ at each time t for each of the specifications was calculated.

For the analysis of the results, the time series of the average results of the budgets $\bar{\chi}_{t,m,f,g,s}$ were split by risk level group γ_s to be plotted. Each plot of each group γ_s has each path spanning on the collumns while each row has a boundary specification of the double-barrier method.

Lastly, it was generated the plot of the time series of the average results of the budgets $\bar{\chi}_{t,m,f,g,s}$ for the class MAB-MMAR, where we added the region given in each time t by the difference between the worst and the best allocations, allowing to assert about the variability of each tunning.

5.3.1 Comparing Models

Since the allocation models generated by the $1/n$ and the *One Asset* methodology do not allow risk level tuning (but allow double-barrier tuning) and their behaviours being better correlated in performance with the high risky tunings, we compared those models only on them.

Also, given that the evaluation of the "pure" MAB-MMAR is the main core of this work, we present the results without the Double-Barrier Method, considering the "none" case, i.e., as a non-hedge procedure. However, to evaluate variability, the Double-Barrier Method is presented.

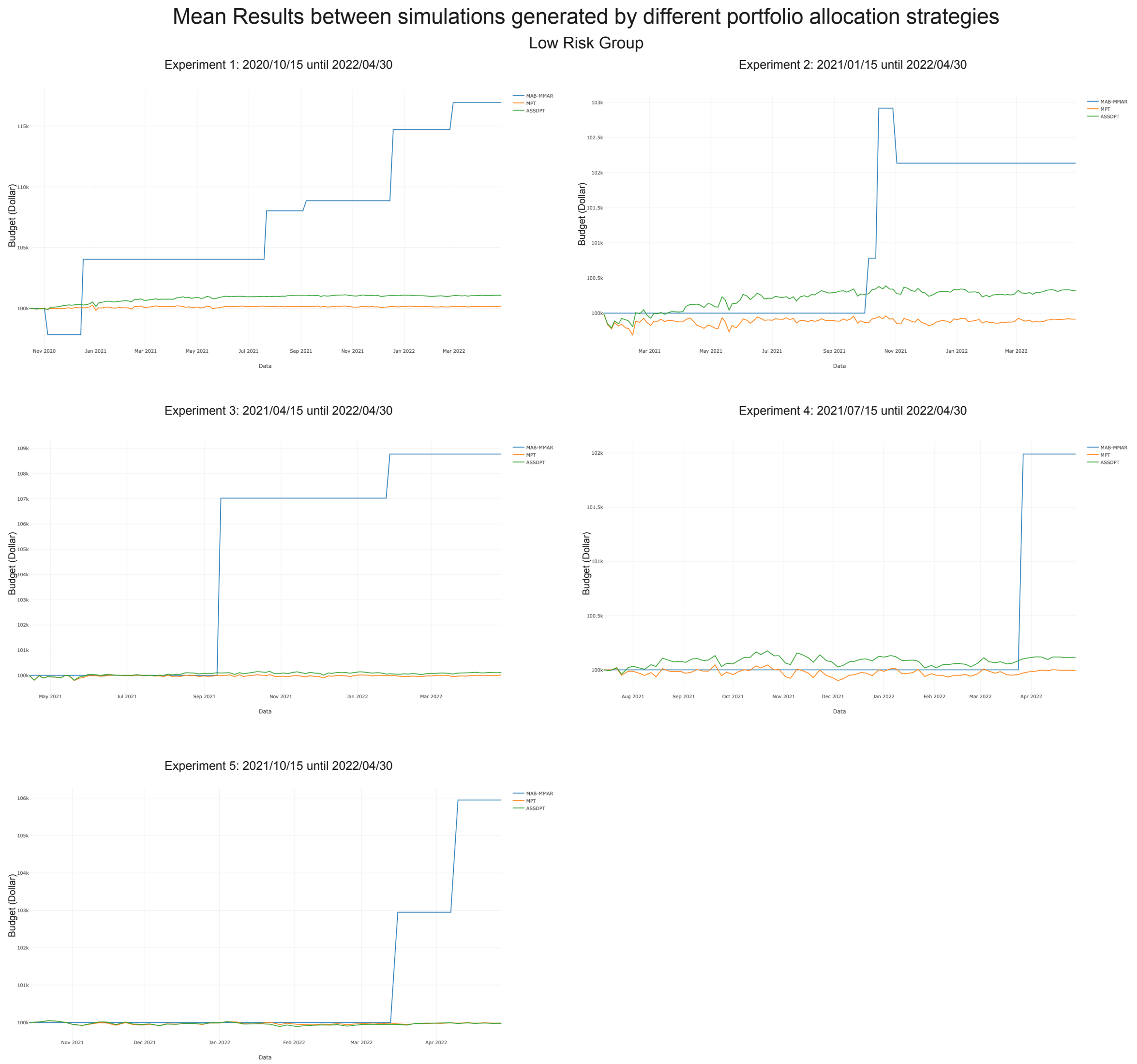


Figure 5.2: Low Risky Tunning of each Allocation Strategy.

From Figure 5.2 which represents the Low Risky Tuning chart, observing the MPT, ASSDPT and MAB-MMAR models, we have that all of them had difficulty in moving away from the initial investment axis, with MAB-MMAR being the one that best positively distanced, having positive results including for path spannings with downturn in the market.

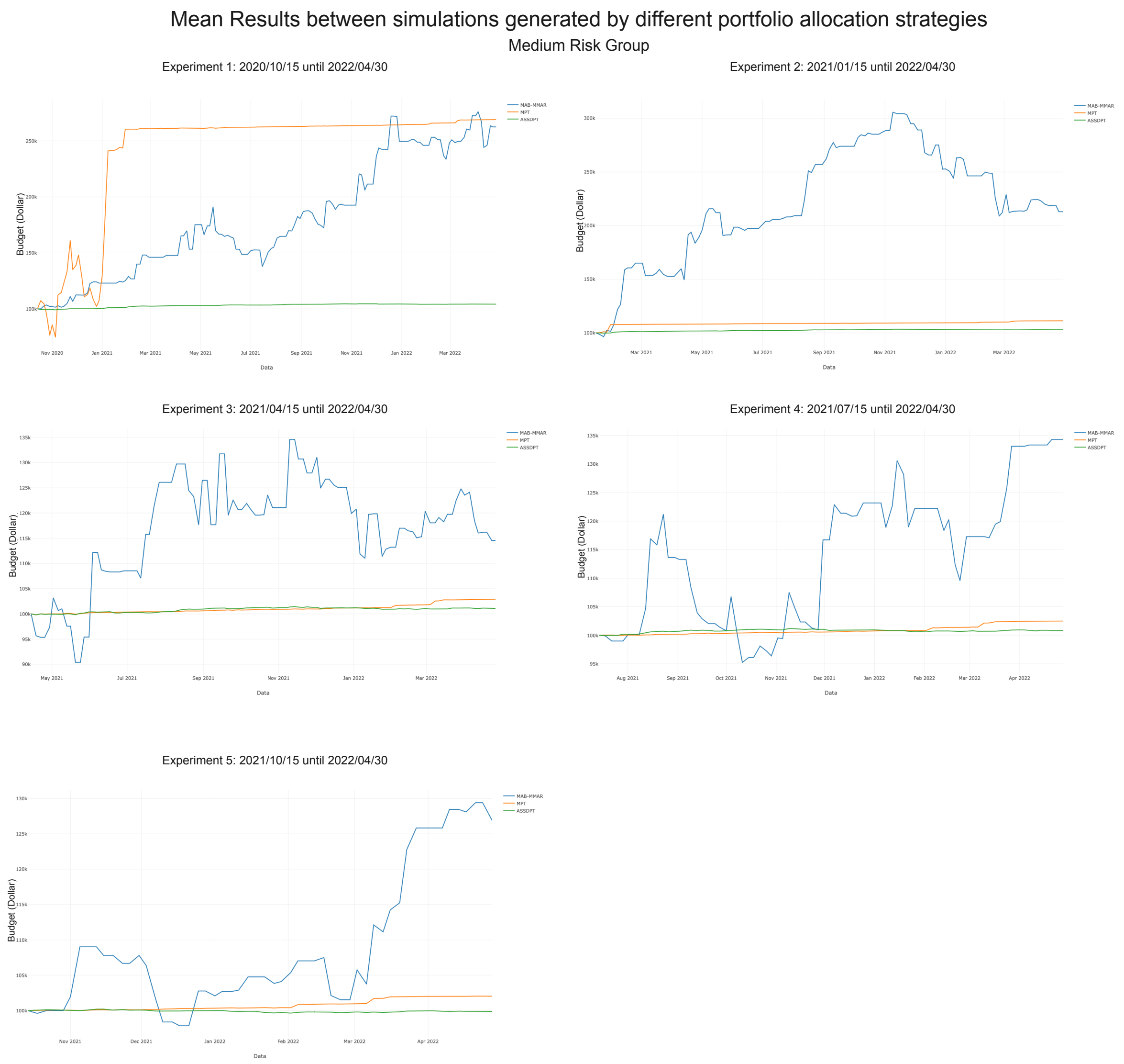


Figure 5.3: Medium Risky Tunning of each Allocation Strategy.

From Figure 5.3 which represents the Medium Risky Tuning chart, MPT and ASSDPT had difficulty in moving away from the initial investment axis, while MAB-MMAR not, having positive results including for path spannings that had downturn in the market. Also, MAB-MMAR had the best overall results comparing all models ($1/n$ and *One Asset* included), for the time window 20210415.

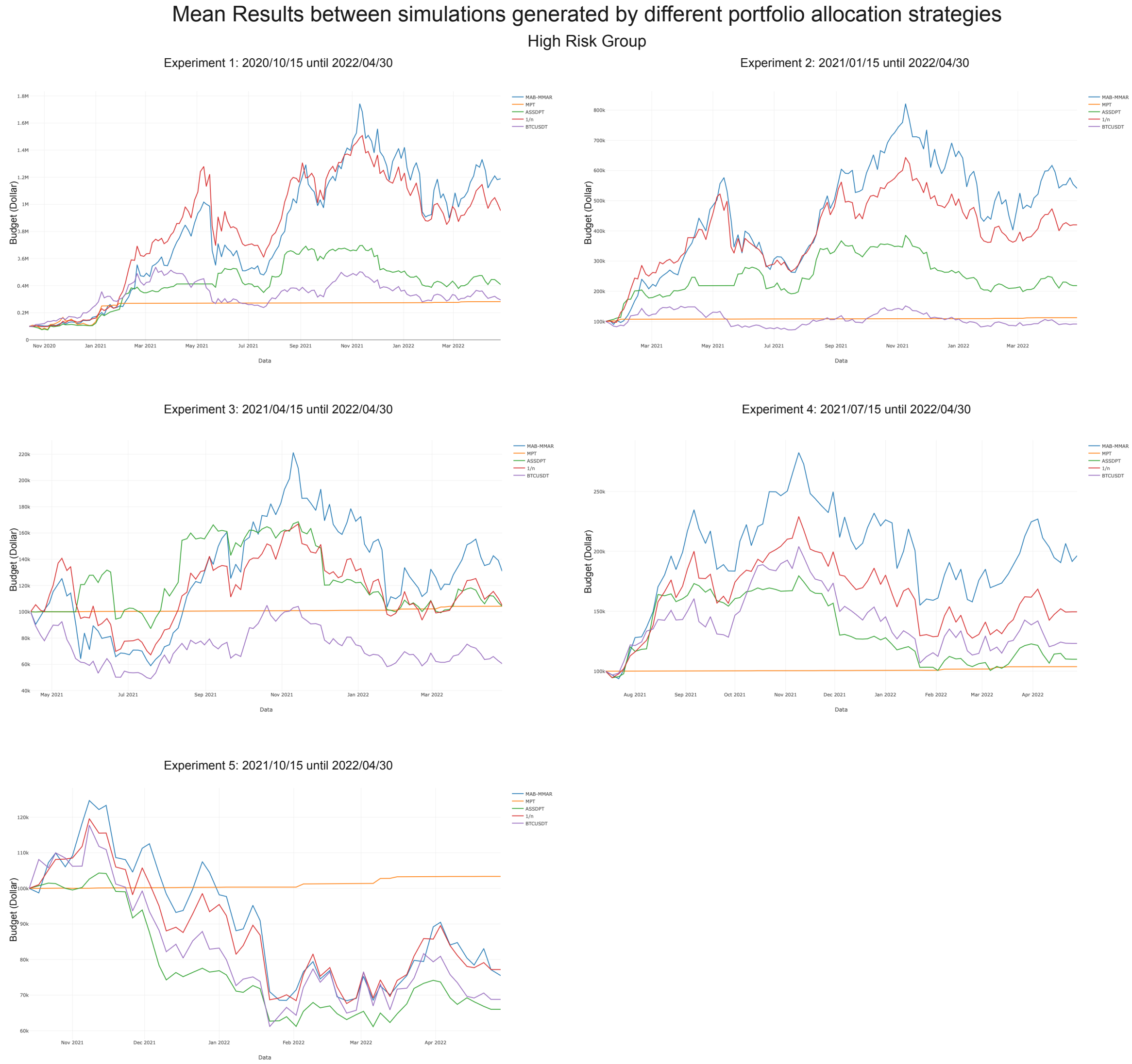


Figure 5.4: High Risky Tunning of each Allocation Strategy.

From Figure 5.4 which represents the High Risky Tuning chart, observing all models, we have that all of them, excepting MPT, had no difficulty in moving away from the initial investment axis, with MAB-MMAR being the one that, in general, had the best performance, mainly if no hedge procedure is applied. The only path spanning that the MAB-MMAR had lost money was the smallest one which had a downturn in the market, however, excepting MPT which has moved almost nothing from the initial investiment axis, all the other models hd also lost money.

Observing the behavior generated by the $1/n$ and *One Asset* methodologies, it is noticed that the $1/n$ methodology has, in general, a superior performance than the *One Asset* methodology, having behaviors in direction of the series quite close, where, when the bitcoin series goes up, the $1/n$ series tends to rise as well, suggesting a strong correlation between cryptocurrencies with bitcoin.

5.3.2 Evaluating Variability

In Figure 5.5, the gray area is given by the maximum and minimum values of each allocation tuning in respect to each time window, where each allocation tuning in each time window were simulated 4 times. The green line presents the mean of each allocation tuning in respect to each time window. Observing the variability chart of the MAB-MMAR model, it is verified that the higher the risk level, the lower the volatility of generated allocation results, where the models with less risky tunings were those that generated the largest gray regions. The Double-Barrier Method does not appear to have significant effects on the distribution of the variability of the results.

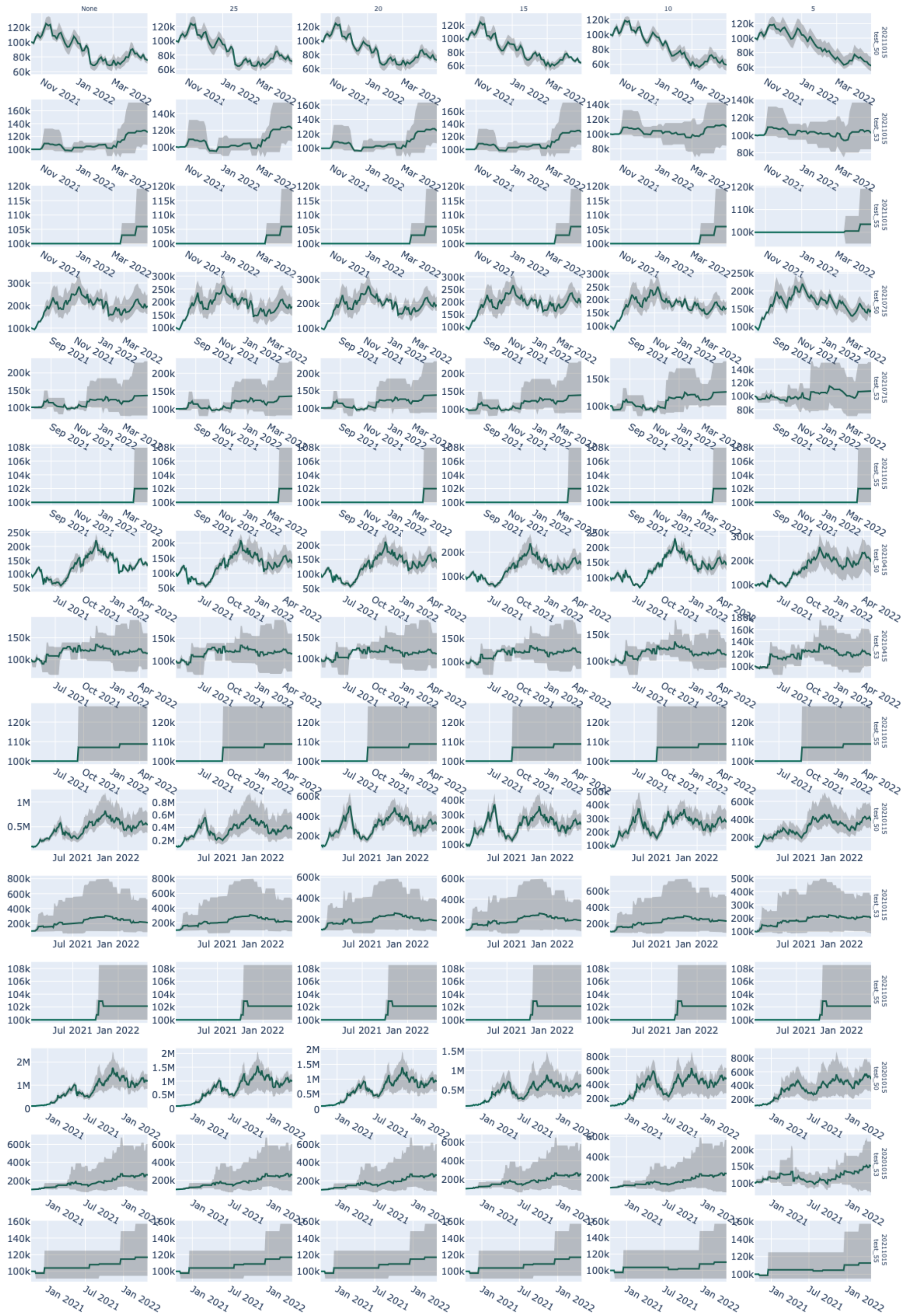


Figure 5.5: Variability of each tunning generated by MAB-MMAR strategy.

Chapter 6

Final Considerations

This dissertation aimed to propose a new methodology for allocating resources in the financial market by making decisions on the perspective of uncertainty. This work was only possible through the availability of intraday historical series made available by Binance.

The proposed MAB-MMAR model sought to answer the following questions:

1. **Feasibility of MAB-MMAR.** Is it feasible to implement a multi-armed strategy to allocate assets in a portfolio using predictions made by a generative model?
2. **Confidence in Non-Negative Returns.** Using a multi-armed bandit framework, at the end of each trial, can we select assets with non-negative returns with a high level of statistical significance?
3. **MAB-MMAR - Comparison.** Does the multi-armed framework bring better advantages against other portfolio allocation models presented in literature?
4. **Double-Barrier Method - Comparison.** Does the Double-Barrier Method bring better advantages against using a non-hedge model?

6.1 Feasibility of MAB-MMAR

Is it feasible to implement a multi-armed strategy to allocate assets in a portfolio using predictions made by a generative model?

Yes. The allocation model for assets in the financial market proposed by this MAB-MMAR dissertation was able to achieve interesting results in all path spannings tested. Remember that it is interesting to test different path spanning scenarios, because the results are dependent on the real path and, if the real path has been positive along the path spanning, even if the model had performed bad allocations, it would also be expected a positive result.

In particular, it is interesting to note that the proposed risk measure $\hat{P}_k(a) \geq \iota$, where ι is the threshold that limits the minimum expected probability of a positive outcome at the end of the round, is able to assert the most varied risk profiles. We have that the more aggressive the risk profile, the lower the value of ι, the more allocations are made, the more intense the results generated and the greater the dependence of the allocation regarded to the path.

It is observed that among all the tested values of $\iota \in [0.5, 0.53, 0.55]$ for the MAB-MMAR model, the riskiest version $[0.5]$ was the one that obtained, in general, the most expressive results where, in some cases, the initial amount was multiplied by more than 10 times. However, it is observed that for the smallest time window tested, the one where there was a fall in the market, the riskier version obtained results that reduced the initial amount, while the less risky tunings $[0.53, 0.55]$ managed to increase the investment value.

From the above, for future work we suggest that, even though the tuning $[0.5]$ has obtained, in general, the best results, some value between $[0.5, 0.53]$ may be the one that can bring expressive results like a riskier profile, with less dependence on the path like a riskless profile. Also, something that could be tested would be the addition of a third intermediate step between steps 1 and 2 of the MAB-MMAR, which could possibly be another MAB structure, responsible for tuning ι.

6.2 Confidence in Non-Negative Returns

Using a multi-armed bandit framework, at the end of each trial, can we select assets with non-negative returns with a high level of statistical significance?

To answer the question in this section, we calculated the mean percentage of the signs generated by the 4 simulations of allocations produced by the MAB-MMAR in all of the Path Spannings without considering the application of boundaries for the Double-Barrier Method, that is, considering the "None" case. If χ_t is a budget at time t, a sign can be defined as:

$$\text{sign}(\chi_t) = \begin{cases} 1, & \chi_t \geq 0 \\ -1, & \chi_t < 0 \end{cases} \tag{6.1}$$

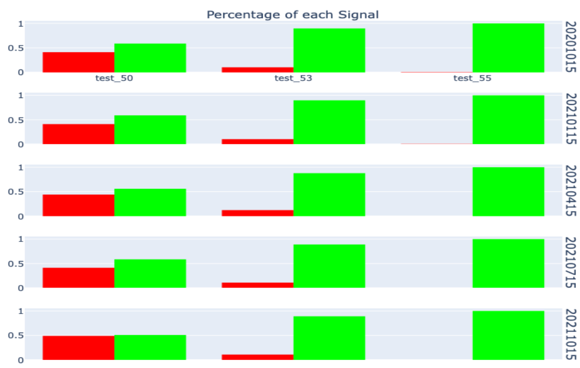


Figure 6.1: Percentage of Mean Signs in MAB-MMAR. Red: Percentange of the average of negative signs, i.e., percentage of cases where $\chi_t < 0$. Green: Percentange of the average of nonnegative signs, i.e., percentage of cases where $\chi_t \geq 0$.

From the Figure 6.1, we have that the level of statistical significance of non-negative signs increases as the level of aggressiveness of the investor profile decreases, where, although in all of the cases we had non-negative results greater than negative signs, we have that **the level of statistical significance is dependent on the risk profile parameter ι.**

6.3 MAB-MMAR - Comparison

Does the multi-armed framework bring better advantages against other portfolio allocation models presented in literature?

In order to answer this question, we calculated the average result of the allocations between the 4 samples generated by each of the models $\nu_m(.)$ in each tuning γ_s where $s = 1, 2, 3$ in each path spanning $[gt, T]$ for $g = 0, 1, 2, 3, 4$ when $t = T$, that is, at the end of the whole process. In order to isolate the analysis from the procedure, we chose to evaluate the case where no boundary is applied for the Double-Barrier Method, that is, considering the "None" case.

Figure 6.2: MAB-MMAR - Comparison. Mean amount at the end of the path spanning per model in each risk group without a Double-Barrier Frontier. Pink: Stochastically Nondominated Portfolio Strategy. Red: Bitcoin Strategy (One Asset Strategy). Green: MAB-MMAR Strategy. Blue: Modern Portfloio Theory Strategy. Yellow: $1/n$ Strategy.

From the left-side chart in Figure 6.2, comparing all the models, it is possible to see that the $1/n$ model performed, in general, equally or better than the groups of os less risky tunings, while $one - asset$ (BTCUSDT) was usually equal or worse than the other models for all risk tunings. When evaluating the riskiest profile group, excepting for the smallest time window 20211015, in all other time spannings, MAB-MMAR were the one that best performed.

from the right-side chart in Figure 6.2, comparing MAB-MMAR against, exclusively, MPT and ASSDPT, the only case were a model had a better performance than the MAB-MMAR was in the riskier profile group for the smallest time window 20211015.

From the above extend, **the MAB-MMAR setup could generate better results asserting every specific risk profile**, then it had better advantages against other methods compared in this dissertation.

6.4 Double-Barrier Method - Comparison

Does the Double-Barrier Method bring better advantages against using a non-hedge model?

To answer this question, as in the previous section, we calculated the average result of the allocations at time $t = T$ between the 4 samples generated by an allocation model $\bar{\chi}_{T,g}$, however, we evaluated the results using $\nu_m(.)$ given by MAB-MMAR, which is the core model of this dissertation. Each risk group γ_s where $s = 1, 2, 3$ and each path spanning $[jg, T]$ where $g = 0, 1, 2, 3, 4$ were tested.

In a second moment, a table was generated with the average results $\bar{\bar{\chi}}_T$ at the end of the allocation $t = T$ between time windows by frontier level for the Double-Barrier Method. Also, it is presented the corresponding standard deviation $\sigma_{\bar{\chi}_T}$ between time windows.

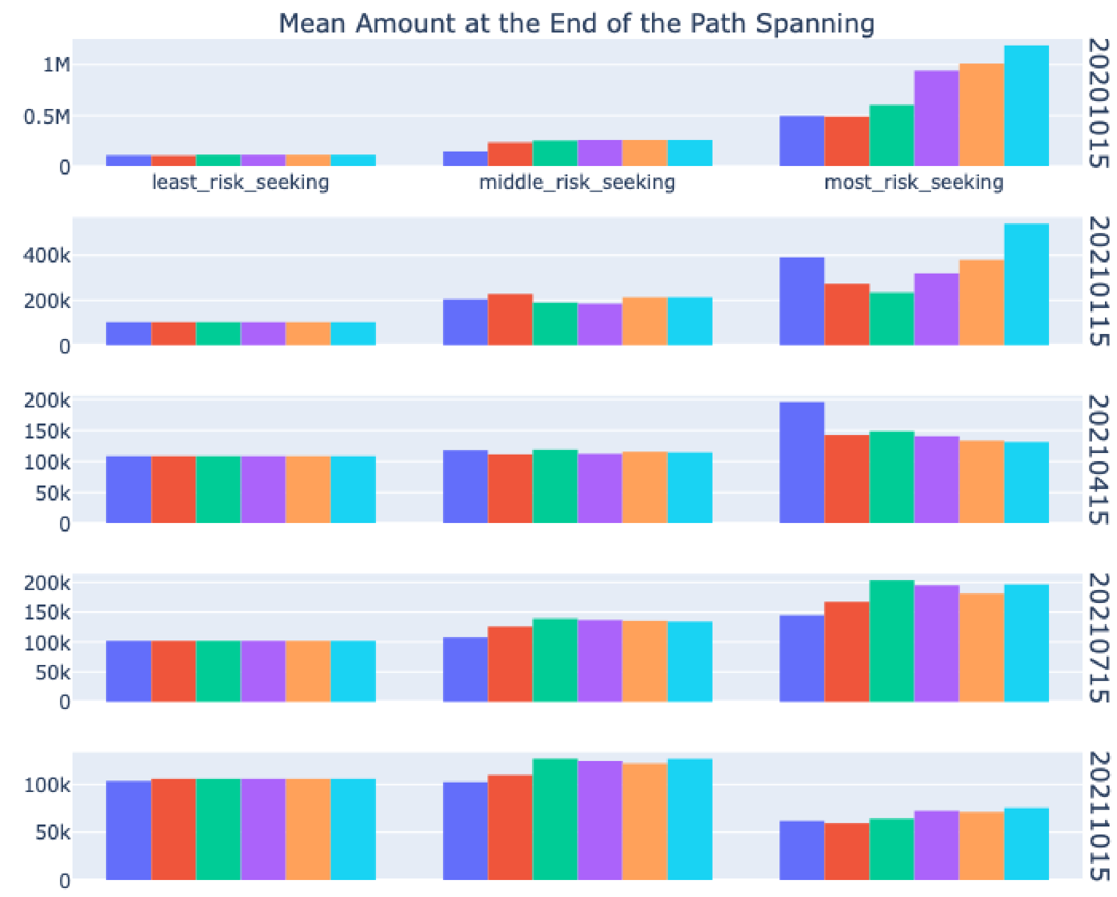


Figure 6.3: Double-Barrier Method - Comparison. Mean amount at the end of the path spanning using MAB-MMAR as the model in each risk group for each Frontier for the Double-Barrier Method. Dark Blue: 5% Frontier Level. Red: 10% Frontier Level. Green: 15% Frontier Level. Violet: 20% Frontier Level. Orange: 25% Frontier Level. Light Blue: Without Frontier.

From the Figure 6.3, it is possible to verify that the riskier is the profile group, more changes happen between frontier levels. Evaluating the most risk group, it is possible to see that without a barrier, in general, the results were similar or better than the other frontiers. However, for the time spanning 20210415 which compresses 1 year of investiment, a lower barrier would bring better results.

Considering the table of average final results between time windows, it can be seen that the less risky the tuning profile, the less levels of the Double-Barrier Method are reached, where for the least risky profile, the highest frontier reached was the frontier of 10%. The results reveal that, on average, using a non-hedge procedure woulb bring higher results. However, a high level of frontier can bring results close to those generated by not using anything, with the benefit of allowing the framework to leave an operation if at some point there is an extreme drop in the invested asset, assuming enough liquity on the market.

From the above, it is suggested that if an aggressive risk profile is used in the tuning of the MAB-MMAR, some level for the Double-Barrier Method will bring more safety to the operation. In opposition, if a less agrressive risk profile is used for tuning MAB-MMAR, then it might not be necessary the use of an hedge procedure, but if so, it can brings a second layer of protection. **So, a Double-Barrier Frontier is not always going to bring advantages against using a non-hedge model, but it brings safety**.

Group (γ_s)	Double-Barrier Frontier	**mean**	**std**
least_risk_seeking	5	105816.184368	4711.187121
least_risk_seeking	10	105788.632451	3718.630105
least_risk_seeking	15	107154.573936	6155.369323
least_risk_seeking	20	107154.573936	6155.369323
least_risk_seeking	25	107154.573936	6155.369323
least_risk_seeking	None	107154.573936	6155.369323
middle_risk_seeking	5	136877.040865	42351.740373
middle_risk_seeking	10	162892.646140	64991.787729
middle_risk_seeking	15	166219.024896	57309.349981
middle_risk_seeking	20	164033.297640	61408.414174
middle_risk_seeking	25	169718.611131	64849.340383
middle_risk_seeking	None	170236.000204	64427.280449
most_risk_seeking	5	258193.368546	180249.008627
most_risk_seeking	10	226427.264228	165706.103970
most_risk_seeking	15	251236.033507	208382.452040
most_risk_seeking	20	333560.817868	351728.424334
most_risk_seeking	25	355310.945550	384585.822921
most_risk_seeking	None	426562.524730	463129.615526

Table 6.1: Mean Allocation Results at $t = T$ among time windows for the Double-Barrier Method by Risk Group (γ_s) using MAB-MMAR.

Appendix A

Multifractality of $X(t)$

From the original publication: *A MULTIFRACTAL MODEL OF ASSET RETURNS* (Mandelbrot, Fisher, and Calvet, 1997).

$X(t)$ is a compound process:

$$X(t) \equiv B_H[\theta(t)] \tag{A.1}$$

with the following assumptions:

- **Assumption I.** $\theta(t)$ is the stochastic trading time while $B_H(t)$ is a fractional Brownian motion.
- **Assumption II.** $\theta(t)$ and $B_H(t)$ are independent with each other.
- **Assumption III.** Trading time $\theta(t)$ is the cumulative density function of a multifractal measure μ on an interval $[0, T]$.

Theorem A.1. *Under Assumptions [I] - [III], the process $X(t)$ is multifractal, with scaling function $\tau_X(q) \equiv \tau_\theta(Hq)$ and stationary increments.*

Proof. from:

$$E(|X(t)|^q) = E[E(|X(t)|^q|\theta(t) = u)],$$

since the trading time $\theta(t)$ and the self-affine process $\{B_H(t)\}$ are independent, conditioning on $\theta(t)$ yields:

$$E(|X(t)|^q|\theta(t) = u) = E[|B_H(u)|^q|\theta(t) = u],$$

$$E(|X(t)|^q|\theta(t) = u) = \theta(t)^{Hq}E[|B_H(1)|^q],$$

and then,

$$E[|X(t)|^q] = E[\theta(t)^{Hq}]E[|B_H(1)|^q],$$

so, the process $X(t)$ satisfies the multiscaling relation (A.1.), with $\tau_X(q) \equiv \tau_\theta(Hq)$ and $c_X(q) \equiv c_\theta(Hq)E[|B_H(1)|^2]$.

A.1 Multiscaling Relation

Definition A.1. *A stochastic process $\{X(t)\}$ is called multifractal if it has stationary increments and satisfies*:

$$E[|X(t)|^q] = c(q)t^{\tau(q)+1}, \textit{ for all } t \in T, q \in Q,$$

where T and Q are intervals on the real line, $\tau(q)$ and $c(q)$ are functions with domain Q. Moreover, it is assumed that T and Q have positive lengths, and that $0 \in T$,$[0, 1] \subseteq Q$.

Appendix B

Properties of MMAR

From the original publication: *A MULTIFRACTAL MODEL OF ASSET RETURNS* (Mandelbrot, Fisher, and Calvet, 1997).

B.1 Theorem 1

If $B_H(t)$ is a Brownian motion without drift, the following properties hold:

1. *If $E(\theta^{1/2})$ is finite, then $\{X(t)\}$ is a martingale with respect to its natual filtration;*
2. *If $E(\theta)$ is finite, the increments of $X(t)$ are uncorrelated, that is $\gamma_X(t) = 0$ for all $t \geq \Delta t$.*

where $\gamma_X(t)$ is the covariance function.

Proof.

1. Letting $\mathcal{F}_t$ and $\mathcal{F}'_t$ to denote the natural filtrations of $\{X(t)\}$ and $\{X(t), \theta(t)\}$, respectively. It is computed $E\{X(t+T)|\mathcal{F}_t\}$ as the iterated expectation:

$$E\{E\{B_H[\theta(t+T)]|\mathcal{F}'_t, \theta(t+T) = u\}|\mathcal{F}_t\} \tag{B.1}$$

for any t, T and $u \geq t$, the independence of B_H and θ implies that:

$$E\{B_H[\theta(t+T)]|\mathcal{F}'_t, \theta(t+T) = u\} = E[B_H(u)|\mathcal{F}'_t] = B_H[\theta(t)]$$

since in this case, $\{B_H(t)\}$ is a martingale. Then it is infered that

$$E[X(t+T)|\mathcal{F}_t] = X(t)$$

2. It is a direct consequence of the that $E[X(t'+\Delta t) - X(t')|\mathcal{F}_t] = 0$ when $t + \Delta t \leq t'$.

This theorem shows that a white spectrum is generated by MMAR when $H = 1/2$.

B.2 Theorem 2

If $E(\theta^{2H})$ is finite, the autocovariance function of the price process $X(t)$ satisfies for all $t \geq \Delta t$:

$$\gamma_X(t) = K\{(t+\Delta t)^m + (t-\Delta t)^m - 2t^m\}$$

where $m = \tau_\theta(2H) + 1$ and $K = c_\theta(2H)Var[B_H(1)]/2$. It is positive when $H > 1/2$, and negative when $H < 1/2$.

Proof.

Considering the conditional expectation:

$$E\{X(0, \Delta t)X(t, \Delta t)|\theta(\Delta t) = u_1, \theta(t) = u_2, \theta(t+\Delta t) = u_3\}. \quad \text{(B.2)}$$

Given that $B_H(t)$ and $\theta(t)$ are independent processes, the expression can be rewrite as:

$$E\{B_H(u_1)[B_H(u_3) - B_H(u_2)]\}$$

or

$$\frac{1}{2}\{|u_3|^{2H} - |u_2|^{2H} + |u_2 - u_1|^{2H} - |u_3 - u_1|^{2H}\}Var[B_H(1)]$$

which can be rewrite in trading time terms:

$$\frac{Var[B_H(1)]}{2}\{|\theta(t+\Delta t)|^{2H} - |\theta(t)|^{2H} + |\theta(t) - \theta(\Delta t)|^{2H} - |\theta(t+\Delta t) - \theta(\Delta T)|^{2H}\}.$$

Taking the expectation yields to theorem B.2.

This theorem shows that the process $B_H(t)$ has long memory, and price increments are positively correlated when $H > 1/2$.

B.3 Theorem 3

If $H \geq 1/2$ and $E(\theta^{Hq})$ is finite, the compound process satisfies:

$$\delta_X(t, q) \geq \delta_\theta(t, Hq)[E[|B_H(1)|^q]]^2$$

for all non-negative q and $t \geq \Delta t$. Moreover, the result holds as equality when $H = 1/2$.

Since X is a stochastic process with stationary increments, it is convenient to define:

$$\delta_X(t, q) = Cov(|X(a, \Delta t)|^q, |X(a + t, \Delta t)|^q)$$

Proof.

Defining the conditional expectation:

$$E\{|X(0,\Delta t)X(t,\Delta t)|^q|\theta(\Delta t)=u_1,\theta(t)=u_2,\theta(t+\Delta t)=u_3\}. \tag{B.3}$$

Since $B_H(t)$ and $\theta(t)$ are independent processes, the conditional expectation can be expressed as,

$$E\{|B_H(u_1)[B_H(u_3)-B_H(u_2)]|^q\}.$$

It is assumed that the Fractal Brownian Motion has either positively correlated ($H > 1/2$) or independent ($H = 1/2$) increments. Then, equation B.3 is equal to, or bounded below by

$$E[|B_H(u_1)|^q]E[|B_H(u_3)-B_H(u_2)|^q] = |u_1|^{Hq}|u_3-u_2|^{Hq}[E|B_H(1)|^q]^2$$

The lower bound can be rewritten in terms of trading time,

$$|\theta(\Delta t)|^{Hq}|\theta(t,\Delta t)|^{Hq}[E|B_H(1)|^q]^2 \tag{B.4}$$

and taking the expectation in equations B.3 and B.4, it is infered that:

$$E[|X(0,\Delta t)X(t,\Delta t)|^q] \geq E[|\theta(0,\Delta t)\theta(t,\Delta t)|^q][E|B_H(1)|^q]^2$$

This theorem shows that the price process has long memory in the absolute value of its increments.

Appendix C

Mean-Risk Consistency with *SSD*

From the original publication: *FRONTIERS OF STOCHASTICALLY NONDOMINATED PORTFOLIOS* (Ruszczynski and Vanderbei, 2003).

For a real random variable V, its first performance function is defined as the right-continuous cumulative distribution function of V:

$$F_V(\eta) = P\{V \leq \eta\} \text{ for } \eta \in \mathbb{R}.$$

A random return V is said to *stochastically dominate* another random return S to the first order, denoted $V \succeq_{FSD} S$ if,

$$F_V(\eta) \leq F_s(\eta) \text{ for all } \eta \in \mathbb{R}.$$

The second performance function $F^{(2)}$ is given by areas below the distribution function F,

$$F_V^{(2)}(\eta) = \int_{-\infty}^{\eta} F_V(\xi)\, d\xi \text{ for } \eta \in \mathbb{R}$$

and defines the weak relation of the *second-order stochastic dominance* (SSD).

That is, random return V stochastically dominates S to the second order, denoted $V \succeq_{SSD} S$ if,

$$F_V^{(2)}(\eta) \leq F_S^{(2)}(\eta) \text{ for all } \eta \in \mathbb{R}.$$

Then, is said that, if x dominates y under the *SSD* rules, then the total return $R(x) \succeq_{SSD} R(y)$ if and only if $E[U(R(x))] \geq E[U(R(y))]$ for every nondecreasing and concave $U(.)$ for which theses expected values are finite. By changing the order of integration, it is possible to write the function $F^{(2)}(.;x)$ as the expected shortfall, such as:

$$F^{(2)}(\eta; x) = E\{\max(\eta - R(w), 0)\}.$$

This function $F^{(2)}(.;x)$ is continuous, convex, nonnegative and nondecreasing.

Definition C.1. *The mean-risk model (μ, ρ) is consistent with SSD with coefficient $\alpha > 0$ if,*

$$R(x) \succeq_{SSD} R(y) \Rightarrow \mu(x) - \lambda\rho(x) \geq \mu(y) - \lambda\rho(y)$$

for all $0 \geq \lambda \geq \alpha$.

Theorem C.1. *The mean-risk model in which the risk is defined as the absolute semideviation:*

$$\bar{\delta}(x) = E\{\max(\mu(x) - R(x), 0)\}$$

is consistent with the second-order stochastic dominance relation with coefficient 1.

Proof. First, the line $\eta - \mu(x)$ is the asymptote of $F^{(2)}(\eta; x)$ for $\eta \to \infty$. Therefore, $R(x) \succeq_{SSD} R(y)$ implies that

$$\mu(x) \geq \mu(y). \tag{C.1}$$

Secondly, setting $\eta = \mu(x)$, then,

$$\bar{\delta}(x) \leq E[\max(0, \mu(x) - R(y))].$$

Since $\mu(x) - \mu(y) \geq 0$, then,

$$\max(0, \mu(x) - R(y)) = \max(0, \mu(x) - \mu(y) + \mu(y) - R(y))$$

$$\max(0, \mu(x) - R(y)) \leq \mu(x) - \mu(y) \max(0, \mu(y) - R(y))$$

Taking the expected value of both sides and combining with the preceding inequality, then,

$$\bar{\delta}(x) \leq \mu(x) - \mu(y)\bar{\delta}(y)$$

which can be rewritten as:

$$\mu(x) - \bar{\delta}(x) \leq \mu(y) - \bar{\delta}(y) \tag{C.2}$$

Combining equations C.1 with C.2. with coefficients $1 - \lambda$ and λ where $\lambda \in [0, 1]$, it is possible to obtain the required result.

Elementary calculations show that for any distribution

$$\bar{\delta}(x) = \frac{1}{2}\delta(x),$$

where $\delta(x)$ is the mean absolute deviation from the mean:

$$\delta(x) = E[R(x) - \mu(x)]. \tag{C.3}$$

Thus, $\delta(x)$ is a consistent risk measure with the coefficient $\alpha = \frac{1}{2}$.

References

Bachelier, Louis (1900). "Théorie de la spéculation". *Annales scientifiques de l'École normale supérieure*. Vol. 17, pp. 21–86.

Chang, Ying-Hua and Lee, Ming-Sheng (2017). "Incorporating Markov decision process on genetic algorithms to formulate trading strategies for stock markets". *Applied Soft Computing* 52, pp. 1143–1153.

Claessens, Stijn and Perotti, Enrico (2007). "Finance and inequality: Channels and evidence". *Journal of comparative Economics* 35.4, pp. 748–773.

Costa, Carlos A Bana E and Soares*, João Oliveira (2004). "A multicriteria model for portfolio management". *The European Journal of Finance* 10.3, pp. 198–211.

De Prado, Marcos Lopez (2018). *Advances in financial machine learning*. John Wiley & Sons.

Gaio, Luiz Eduardo and Sáfadi, Thelma (2008). "Memória longa na volatilidade do Índice BOVESPA: uma análise utilizando modelos da classe ARCH." *Revista de Economia e Administração* 7.2.

Günay, Samet (2016). "Performance of the Multifractal Model of Asset Returns (MMAR): evidence from emerging stock markets". *International Journal of Financial Studies* 4.2, p. 11.

Hoffman, Matthew, Brochu, Eric, De Freitas, Nando, et al. (2011). "Portfolio Allocation for Bayesian Optimization." *UAI*. Citeseer, pp. 327–336.

Hsu, Po-Hsuan, Tian, Xuan, and Xu, Yan (2014). "Financial development and innovation: Cross-country evidence". *Journal of financial economics* 112.1, pp. 116–135.

Huo, Xiaoguang and Fu, Feng (2017). “Risk-aware multi-armed bandit problem with application to portfolio selection”. *Royal Society open science* 4.11, p. 171377.

Jiang, Zhi-Qiang et al. (2019). “Multifractal analysis of financial markets: a review”. *Reports on Progress in Physics* 82.12, p. 125901.

Kuleshov, Volodymyr and Precup, Doina (2014). “Algorithms for multi-armed bandit problems”. *arXiv preprint arXiv:1402.6028*.

Mandelbrot, Benoit B, Fisher, Adlai J, and Calvet, Laurent E (1997). “A multifractal model of asset returns”.

Mandelbrot, Benoit B and Hudson, Richard L (2010). *The (mis) behaviour of markets: a fractal view of risk, ruin and reward*. Profile books.

Markowitz, Harry (1952). “Portfolio Selection, Journal of Finance”. *Markowitz HM—1952.—№*, pp. 77–91.

Petrik, Marek and Subramanian, Dharmashankar (2012). “An approximate solution method for large risk-averse Markov decision processes”. *arXiv preprint arXiv:1210.4901*.

Robbins, Herbert (1952). “Some aspects of the sequential design of experiments”. *Bulletin of the American Mathematical Society* 58.5, pp. 527–535.

Ruszczynski, Andrzej and Vanderbei, Robert J (2003). “Frontiers of stochastically nondominated portfolios”. *Econometrica* 71.4, pp. 1287–1297.

Taleb, Nassim Nicholas (2012). *Antifragile: Things that gain from disorder*. Vol. 3. Random House Incorporated.

Taleb, Nassim Nicholas et al. (2012). “A new heuristic measure of fragility and tail risks: application to stress testing”.

Vink, Markus (2003). “" The World’s Oldest Trade": Dutch Slavery and Slave Trade in the Indian Ocean in the Seventeenth Century”. *Journal of world history*, pp. 131–177.

Wang, De-Xin and Cao, Xi-Ren (2011). “Event-based optimization for POMDPs and its application in portfolio management”. *IFAC Proceedings Volumes* 44.1, pp. 3228–3233.

Wong, Anson and Zhou, Xianbo (2011). “Development of financial market and economic growth: Review of Hong Kong, China, Japan, the United States and the United Kingdom”. *International Journal of Economics and Finance* 3.2, pp. 111–115.

Zhu, Mengying et al. (2019). “Adaptive portfolio by solving multi-armed bandit via thompson sampling”. *arXiv preprint arXiv:1911.05309*.