Cryptomarket Analysis: Optimal liquidation and Market Efficiency Study

Author
Julián Fernando Sánchez López

A thesis submitted as a requirement to obtain the degree of
Doctor of Economics

Supervisor
Hugo Eduardo Ramírez Jaime

Universidad del Rosario

Facultad de Economía
Doctorado en Economía

Bogotá - Colombia
2024

Abstract

This doctoral thesis presents an in-depth investigation into the optimal liquidation of cryptocurrency assets and the analysis of market dynamics under the Weak Efficient Market Hypothesis (WEMH) across various cryptocurrencies. The first paper of the thesis focuses on the liquidation of Binance Coin (BNB), analyzing temporary and permanent price impacts using Limit Order Book (LOB) data from Binance exchange. It introduces linear and quadratic models for Permanent Price Impact (PPI) and derives optimal liquidation strategies through closed-form solutions. The study finds that the quadratic PPI model notably outperforms linear models in capturing permanent price impacts in financial trading.

The second paper extends this investigation by applying finite differences and optimal policy iteration to numerically solve the liquidation problem under different scenarios of price impact estimation. It characterizes optimal liquidation policies based on various parametrizations and compares their performance with naive and common strategies, highlighting the importance of the inventory's functional form in determining revenue-maximizing policies.

The third paper diverges to examine market dynamics, introducing a novel methodology for identifying epochs of upward trends, downward trends, and mean reversion using statistical techniques. It seeks to evaluate market efficiency within these periods under the Weak Efficient Market Hypothesis (WEMH) through methods including the Hurst index, Shannon entropy, and autocorrelation tests. The analysis concludes that cryptocurrency markets do not uniformly adhere to WEMH principles. While certain periods and frequencies display efficiency, others exhibit predictability and inefficiency, highlighting the complex and fluctuating nature of market efficiency across different cryptocurrencies and market conditions.

Overall, this thesis contributes to the field by providing nuanced insights into asset liquidation strategies in the context of cryptocurrency trading and advancing the understanding of market efficiency dynamics in these novel financial markets.

Acknowledgements

First and foremost, I extend my deepest gratitude to my advisor, Hugo Eduardo Ramirez Jaime, for his invaluable guidance, patience, and support throughout this research journey. His expertise and insights have been instrumental in shaping this thesis.

I am profoundly grateful to a number of other professors, Carlos Castro and Rafael Serrano, whose insightful comments, suggestions, and contributions have been decisive in shaping the direction and outcome of this research. Their expertise and thoughtful guidance have been immensely valuable and deeply appreciated.

I am especially grateful to my close colleagues and friends Camilo Castillo, Juan Carlos Urueña, not only for their constant support and motivation but also for their valuable contributions that have enriched my research and academic growth. Their willingness to share knowledge and provide insightful feedback has been a fundamental part of this journey.

I am immensely thankful to my family for their unwavering love, understanding, and support throughout my academic pursuits. Special thanks to my parents, José Sánchez and Isabel López, and my siblings, Yesid, Martha, Leonardo and Carlos Sánchez, for their constant encouragement. I am equally grateful to my brothers-in-law, Danilo Díaz and Aleida Blandón, and my nephews, Yesica and Diego Sánchez, for their unconditional support. A very special acknowledgment to my dear girlfriend, Catalina García, whose unwavering support has been indispensable. Her belief in me has been a continuous source of strength and inspiration, making this project possible.

Lastly, I want to acknowledge Constanza García, the Director of the Mathematics Department at Universidad Externado de Colombia, whose unconditional support was vital to the completion of this thesis.

Additionally, I extend my appreciation to ChatGPT, a language model developed by OpenAI, for its invaluable assistance in proofreading and enhancing the clarity and coherence of this thesis. The support received from ChatGPT was instrumental in refining the language and ensuring the overall quality of this work.

Contents

List of Figures

List of Tables

1 Analytical Insights to Optimal Liquidation of Cryptocurrencies and Empirical Calibration of Temporary and Permanent Price Impacts

Abstract

This paper investigates the optimal liquidation of cryptocurrency assets, focusing on Binance Coin (BNB),[1] within the context of temporary and permanent price impacts. Utilizing Binance exchange Limit Order Book (LOB) data, we explore three distinct market scenarios: underestimation, overestimation, and average estimation which are defined according to the relationship between the volume of shares to be liquidated and the available volumes in the LOB. For each scenario, we calibrate linear and quadratic models for Permanent Price Impact (PPI), and linear models for Temporary Price Impact (TPI), to accurately reflect the scenario-specific market dynamics. Our research methodology emphasizes analytical over numerical solutions, leading to the derivation of closed-form solutions for the $\mathcal{HJB}$ partial differential equation, which underpins the formulation of optimal liquidation strategies. Through simulation of the liquidation process for an initial inventory of BNB, we evaluate the efficacy of these strategies against a baseline naive approach. The findings highlight the superior performance of quadratic PPI models in capturing the complexity of permanent price effects across various market scenarios. This paper contributes to the broader understanding of asset liquidation within the cryptocurrency trading landscape, underscoring the importance of precise modeling of price impacts.

Keywords: Optimal Liquidation, Cryptocurrency Trading, Binance Coin (BNB), Limit Order Book (LOB) Data, Temporary Price Impact (TPI), Permanent Price Impact (PPI), Market Scenarios (Underestimation, Overestimation, Average Estimation), Analytical Solutions, Financial Trading Models, Price Impact Calibration.

1.1 Introduction

The rapid evolution of the digital economy witnesses the significant rise of cryptocurrencies, a new class of digital assets that offer both security and transparency. These attributes have positioned cryptocurrencies as a notable addition to the portfolio of investment alternatives Blanco Encinosa (2021). As digital representations of value that utilize cryptographic functions for secure transactions, cryptocurrencies stand at the forefront of financial technology innovation.

The rapid development of financial technology (Fintech) has continuously promoted innovation in financial models, reshaping the supply chain and value chain within the financial industry. Cryptocurrencies, as a significant product of Fintech, reduce the reliance on financial intermediaries, thus contributing to the growth of

[1]BNB, short for Binance Coin, is the native cryptocurrency of the Binance exchange platform.

the digital economy. Key data processing technologies further enhance resource allocation efficiency and facilitate financial market reforms Zhao (2021).

Cryptocurrencies operate at the intersection of finance and technology, posing unique regulatory and operational challenges. Their rapid adoption by institutional investors, such as life-insurance companies, investment banks, and hedge funds, underscores their potential to transform financial systems. These digital assets also offer financial independence, as evidenced by El Salvador's adoption of Bitcoin as legal tender. However, issues such as volatility, environmental impact, and regulatory concerns remain significant hurdles Urquhart and Lucey (2022).

The integration of cryptocurrencies into financial portfolios represents a pivotal shift towards decentralized financial instruments. This trend necessitates a deep understanding of their implications on traditional financial systems and the broader economy. The academic community has responded by conducting extensive research on cryptocurrency trading, addressing areas such as trading platforms, strategies, and risk management. This body of work highlights both the opportunities and challenges associated with cryptocurrency trading, emphasizing the need for further study to fully grasp their impact Fang et al. (2022).

The burgeoning interest and inclusion of crypto assets in investment portfolios by institutional investors underline the relevance and emerging focus of research in cryptocurrency trading Fang et al. (2022). Furthermore, the swift pace of developments and a notable increase in interest and activities in cryptocurrency trading since 2018 signal the emergence of this field as a new area of study within financial trading. This paradigm shift invites a comprehensive analysis to unravel the intricate dynamics of cryptocurrency markets and their potential to reshape the financial landscape Fang et al. (2022).

Another fundamental aspect of this study centers on the impact of price changes. A seminal contribution to this area is Kyle's model of insider trading, which highlights the dynamics of how private information is gradually incorporated into market prices through sequential auctions. The model shows that the informed trader, acting as an intertemporal monopolist, strategically exploits his private information while noise traders provide camouflage, enabling the insider to earn profits without being detected by market makers. This process ultimately ensures that all private information is incorporated into prices by the end of trading, reflecting a continuous auction equilibrium where prices follow a Brownian motion and market liquidity remains constant Kyle (1985).

This investigation builds upon the foundational work by Almgren and Chriss (2000), which delineates the concepts of temporary and permanent price impacts on asset prices. According to their analysis, temporary impacts are immediate, short-term price movements away from equilibrium, resulting from trading-induced imbalances in supply and demand. Permanent impacts, on the other hand, are lasting changes in the equilibrium price attributable to trading activity, persisting for the duration of the liquidation period. Within the volatile and decentralized realm of cryptocurrencies, these impacts manifest uniquely, prompting our study to explore both linear and non-linear (specifically, quadratic) models of permanent price impact, informed by cryptocurrency trading data, to delineate their implications

for optimal trading strategies. The inclusion of quadratic models is particularly novel, addressing scenarios not fully captured by linear models alone and reflecting the complex dynamics of cryptocurrency markets. While various approaches to non-linear models of impact functions exist in markets beyond cryptocurrencies, exemplified by studies conducted by Gueant (2014) or Barger and Lorig (2018), they often employ distinct overarching models.

By calibrating the Temporary Price Impact (TPI) and Permanent Price Impact (PPI) functions using simple polynomial functions, we find analytical solutions to the liquidation problem, which is based on the optimal liquidation model proposed by Cartea et al. (2015) and we obtain optimal strategies in three distinct market scenarios: underestimation, overestimation, and average-estimation. Each of these scenarios is defined based on the relationship between the volume of shares to be liquidated and the available volumes in the LOB. This analytical approach allows for a precise understanding of how different price impact models influence trading decisions in the cryptocurrency domain, although the model can be perfectly extrapolated to other markets. Key findings indicate that PPI quadratic models more accurately capture the complexities of market behavior, offering important insights for developing advanced trading strategies. Future directions of this research will involve expanding the analytical framework to include broader variations of the model and applying numerical methods to address the optimal settlement problem in a more generalized context, thereby expanding the understanding of cryptocurrency trading strategies and their implications on the market.

This paper is structured as follows: Section 1.2 introduces the model, based on Cartea et al. (2015). Section 1.3 presents the estimation of the parameters for linear TPI, linear, and quadratic PPI using BNB cryptocurrency's order book data, inspired by Cartea et al. (2015); Cartea and Jaimungal (2016a). In Section 1.4, we provide analytical solutions for two extensions of the liquidation problem from the perspective of optimal stochastic control. Additionally, we conduct a performance comparison between analytical and naive trading strategies through simulations of optimal liquidation using order book (LOB) data. Finally, in Section 1.5, we present our concluding remarks.

1.2 The Model

We work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and build upon the model proposed by Cartea et al. (2015). The agent holds an amount Υ of shares of an asset that she aims to liquidate in the time interval $[0, T)$ to obtain the most profit. The inventory Q_t^ν changes over time with dynamics $dQ_t^\nu = -\nu_t dt$, where ν_t represents the trade rate. The mid price of the asset follows the Stochastic Differential Equation (SDE)

$$dS_t^\nu = -g\left(\nu_t\right) dt + \sigma dW_t, \tag{1.1}$$

where W_t is a standard Brownian motion, $\sigma \in \mathbb{R}^+$ indicates the volatility, and $g : \mathbb{R}^+ \to \mathbb{R}^+$ is the permanent price impact function of the trade rate. When an agent sells shares of the stock, the actual transaction price, known as the execution

price, follows

$$\hat{S}_t^\nu = \left(S_t^\nu - \frac{1}{2}\Delta\right) - f\left(\nu_t\right), \tag{1.2}$$

where Δ is the bid-ask spread in the LOB, and $f : \mathbb{R}^+ \to \mathbb{R}^+$ represents the temporary price impact. The agent seeks to maximize her final expected utility wealth, which is the earnings from transactions,

$$\mathbb{E}\left\{\int_0^T \hat{S}_t^\nu \nu_t dt\right\}. \tag{1.3}$$

The agent controls the rate ν_t at each point in time t, influencing the asset's liquidation price. Thus, the optimization problem is represented by the agent's value function

$$H\left(t,S,q\right) = \sup_{\nu_t \in \mathcal{A}(t,T)} \mathbb{E}\left\{\int_t^T \left(\left(S_r^\nu - \frac{1}{2}\Delta\right) - f\left(\nu_r\right)\right)\nu_r dr\right\}, \tag{1.4}$$

where $\mathcal{A}(t,T)$ denotes the admissible set of non-negative bounded strategies, and equation (1.4) satisfies the $\mathcal{HJB}$ partial differential equation (PDE) [2].

$$0 = \partial_t H + \frac{1}{2}\sigma^2 \partial_{ss} H + \sup_{\nu \in \mathcal{A}}\left\{-g(\nu)\partial_s H - \nu \partial_q H + \left(S - \frac{1}{2}\Delta - f(\nu)\right)\nu\right\}. \tag{1.5}$$

The model imposes two conditions:

$$H(T,S,q) \;\to\; -\infty, \text{ when } t \to T \text{ and } q > 0, \tag{1.6}$$

$$H(t,S,0) \;\to\; 0, \tag{1.7}$$

The first condition imposes a severe penalty for having positive inventory when reaching T, ensuring its avoidance due to the substantial revenue loss it entails, while the second condition ensures zero revenue at time t for nil inventory.

Although Cartea et al. (2015) and Cartea and Jaimungal (2016b) treat linear Permanent Price Impact (PPI), our approach differs due to the type of penalty, which is a boundary condition going to $-\infty$ in the range $[0,T)$. This condition does not allow the manager to have any inventory at expiry. In contrast, Cartea et al. (2015) and Cartea and Jaimungal (2016b) use a quadratic penalty, allowing the manager to end up with some inventory. We propose linear TPI and linear and quadratic PPI, under which we find analytical solutions to the problem.

1.3 Model Calibration

In this section, we describe the implementation of a methodology, inspired by the work of Cartea et al. (2015), for estimating the functional forms of Temporary Price Impact (TPI) and Permanent Price Impact (PPI), along with the parameters

[2] For a detailed explanation, refer to chapter 6 of Cartea et al. (2015)

associated with these impacts. Cartea's empirical analysis focuses on linear estimations for both TPI and PPI by observing order flow and price changes within five-minute intervals and employing robust linear regression techniques. For TPI, the temporary effect is assumed to be linear with respect to the volume traded, with regression analysis based on snapshots of the order book. In our study, we extend their framework, employing linear models for TPI, and exploring both linear and quadratic models for PPI to capture the possible non-linear behaviors observed in market data. Additionally, for estimating the volatility, we use the realized volatility, i.e., $\hat{\sigma} = \sqrt{\sum_{t=1}^{T} \left(\ln(S_t) - \ln(S_{t-1})\right)^2}$, consistent with the model.

1.3.1 *Temporary Price Impact*

To determine the functional form of the TPI, following ideas from Cartea et al. (2015), we use LOB data with n ticks of depth taken at N_t intervals of τ seconds. In each of these N_t time intervals, we simulate the execution of different sizes of Sell Market Orders, i.e., liquidating the volumes $Q = \{Q_1, Q_2, ...Q_i, ...Q_m\}$ against the existing LOB, walking the book when necessary. The execution price is defined as the weighted average of the execution prices for each volume on the bid side, using LOB prices, represented for each available tick j by S_j and its corresponding number of Limit Orders by Q_{S_j}, chosen increasingly up to a significant maximum relative to the LOB's total capacity. Since this procedure is performed every τ seconds, we consider not only volumes but also liquidation rates (ν) in our simulations. The TPI at time t due to the liquidation rate $\nu_i = \dfrac{Q_i}{\tau}$ is $TPI^{bid}_{i,t} = \hat{S}^{bid}_{i,t} - S^{bid}_t$. We use the average of $TPI^{bid}_{i,t}$ over time, denoted by TPI^{bid}_i, against ν_i to fit the linear model $TPI^{bid}(\nu) = f(\nu) = a_1\nu + a_2$, via least squares.

1.3.2 *Permanent Price Impact*

To estimate a functional form of the permanent price impact, we analyze the data every τ seconds, executing simulated Market Sell Orders of different volume sizes and, consequently, different trading liquidation rates at each of those moments. We then observe the impact of each trading rate on the midprice $S^{mid}_t = \dfrac{S^{bid}_t + S^{ask}_t}{2}$. Using least squares, we estimate the parameters of the models $PPI^{bid}(\nu) = g(\nu) = b_1\nu$ and $PPI^{bid}(\nu) = g(\nu) = c_1\nu^2 + c_2\nu$.

1.4 Results

1.4.1 Analytical Solutions

Finding analytical solutions is always satisfactory. This section presents some cases in which we can derive a closed-form formula for solutions. The following lemmas present the analytical solutions for the $\mathcal{HJB}$ PDE (1.5) subject to conditions (1.6)

and (1.7), when TPI is linear and PPI is both linear and quadratic. These lemmas serve as extensions of the solution provided in section 6.3 of Cartea et al. (2015).

Lemma 1.4.1 (**Linear Permanent Price Impact**). *Assuming that TPI and PPI are linear functions, specifically* $f(\nu_t) = a_1\nu_t + a_2$ *, and* $g(\nu_t) = b_1\nu_t$ [3] *respectively, then the solution for the* $\mathcal{HJB}$ *differential equation* (1.5)*, subject to* (1.6) *and* (1.7)*, is*

$$H(t,S,q) = q\left(S - \frac{1}{2}\Delta - a_2\right) - \left(\frac{b_1}{2} + \frac{a_1}{T-t}\right)q^2, \tag{1.8}$$

and the optimal strategy for liquidating q_t *shares in the time interval* $[t,T)$ *is*

$$\nu_t^* = \frac{q_t}{T-t}. \tag{1.9}$$

Proof. The aim is to solve the $\mathcal{HJB}$ differential equation (1.5) to find the policy that optimizes the liquidation of q_t shares in the timeframe $[t,T)$. This involves solving

$$\partial_t H + \frac{1}{2}\sigma^2\partial_{ss}H + \sup_{\nu\in\mathcal{A}}\left\{-b_1\nu\partial_s H - \nu\partial_q H + \left(S - \frac{1}{2}\Delta - a_1\nu - a_2\right)\nu\right\} = 0. \tag{1.10}$$

Applying the first-order conditions, we calculate

$$\nu = \frac{\left(S - \frac{1}{2}\Delta - a_2\right) - b_1\partial_s H - \partial_q H}{2a_1},$$

and substituting in (1.10), we obtain

$$\partial_t H + \frac{1}{2}\sigma^2\partial_{ss}H + \frac{\left(S - \frac{1}{2}\Delta - a_2 - \partial_q H - b_1\partial_s H\right)^2}{4a_1} = 0. \tag{1.11}$$

We employ the educated guess

$$H(t,S,q) = q\left(S - \frac{1}{2}\Delta - a_2\right) - \frac{b_1}{2}q^2 + h(t,q), \tag{1.12}$$

and substituting in the differential equation (1.11) simplifies to

$$\partial_t h + \frac{1}{4a_1}(\partial_q h)^2 = 0, \tag{1.13}$$

which is solvable by using separation of variables $h(t,q) = q^2 h_2(t)$, leading to the solution

$$H(t,s,q) = q\left(S - \frac{1}{2}\Delta - a_2\right) - \left(\frac{b_1}{2} + \frac{a_1}{(T-t)}\right)q^2, \tag{1.14}$$

$$\nu_t^* = \frac{q_t}{T-t}. \tag{1.15}$$

□

[3] We utilize a permanent impact function without the constant coefficient to derive an analytical solution to the problem.

Lemma 1.4.2 (**Quadratic permanent price impact**). *Let us have a linear TPI and a quadratic PPI, that is* $f(\nu_t) = a_1\nu_t + a_2$, *and* $g(\nu_t) = c_1\nu_t^2 + c_2\nu_t$ [4] *respectively, then the analytical solution for the* $\mathcal{HJB}$ *PDE* (1.5) *subject to* (1.6) *and* (1.7) *is*

$$H(t,S,q) = \begin{cases} q\left(S - \frac{1}{2}\Delta - a_2\right) - \frac{c_2}{2}q^2 - \frac{4(c_1q+a_1)^3}{9c_1^2(T-t)} & if \quad q \neq 0 \\ 0 & if \quad q = 0, \end{cases} \tag{1.16}$$

and the optimal strategy for the liquidation of q_t *shares in the time interval* $[t,T)$ *is*

$$\nu_t^* = \frac{2(c_1q_t + a_1)}{3c_1(T-t)} \tag{1.17}$$

Proof. Again the objective is to solve the $\mathcal{HJB}$ differential equation (1.5), and to find the optimal policy. Thus we have to solve

$$\partial_t H + \frac{1}{2}\sigma^2\partial_{ss}H + \sup_{\nu\in\mathcal{A}}\left\{\left(-c_1\nu^2 - c_2\nu\right)\partial_s H - \nu\partial_q H + \left(S - \frac{1}{2}\Delta - a_1\nu - a_2\right)\nu\right\} = 0 \tag{1.18}$$

Using the first order conditions we calculate

$$\nu = \frac{\left(S - \frac{1}{2}\Delta - a_2\right) - c_2\partial_s H - \partial_q H}{2(c_1\partial_s H + a_1)}$$

and replacing in (1.18) , we obtain

$$\partial_t H + \frac{1}{2}\sigma^2\partial_{ss}H + \frac{\left(S - \frac{1}{2}\Delta - a_2 - \partial_q H - c_2\partial_s H\right)^2}{4(c_1\partial_s H + a_1)} = 0 \tag{1.19}$$

Using the educated guess

$$H(t,S,q) = q\left(S - \frac{1}{2}\Delta - a_2\right) - \frac{c_2}{2}q^2 + h(t,q) \tag{1.20}$$

and replacing in the differential equation (1.19) we simplify to

$$\partial_t h + \frac{(\partial_q h)^2}{4(c_1q + a_1)} = 0 \tag{1.21}$$

thus finding the following solutions

$$H(t,S,q) = \begin{cases} q\left(S - \frac{1}{2}\Delta - a_2\right) - \frac{c_2}{2}q^2 - \frac{4(c_1q+a_1)^3}{9c_1^2(T-t)} & if \quad q \neq 0 \\ 0 & if \quad q = 0 \end{cases} \tag{1.22}$$

$$\nu_t^* = \frac{2(c_1q_t + a_1)}{3c_1(T-t)} \tag{1.23}$$

Where q and s are non-negative, bounded and $a_1/c_1 > 0$. □

[4] We utilize a permanent impact function without the constant coefficient to derive an analytical solution to the problem.

Remark. *By directly integrating the relation $dQ_t = -\nu_t dt$ and using the initial condition $Q_0 = \Upsilon$, we solve the optimal inventory path (optimal changes of inventory through time) Q_t for the two scenarios described in the lemmas. For linear PPI in lemma 1.4.1, we have*

$$Q(t) = \left(1 - \frac{t}{T}\right) \Upsilon, \tag{1.24}$$

and for quadratic PPI,

$$Q(t) = \left(1 - \frac{t}{T}\right)^{2/3} \Upsilon + \left(\left(1 - \frac{t}{T}\right)^{2/3} - 1\right) \frac{a_1}{c_1}. \tag{1.25}$$

We observe that both optimal inventory paths include a common term, $\left(1 - \frac{t}{T}\right)^{d_2} \Upsilon$. This term is significant as it implies optimal revenues in some sense.

Remark. *We define the solution function $H(t, S, q)$ explicitly in a piecewise manner, reflecting the intrinsic timing of the completion of liquidation, which invariably occurs before T. This is captured through the solution for t, expressed as $t = T\left(1 - \left(\frac{a_1}{\Upsilon c_1 + a_1}\right)^{3/2}\right)$, where the expression within the parentheses underscores the fact that liquidation concludes prior to T. Such a formulation ensures that $H = 0$ when $q = 0$, adhering to the boundary condition that necessitates H to be zero if the inventory is depleted. This characteristic is pivotal for a liquidation problem, where negative inventories are prohibited, mandating that H adjust according to the dynamic nature of the inventory level.*

1.4.2 Data

It is important to compare the quality of traditional markets and cryptocurrency markets because cryptocurrencies represent a rapidly growing sector with unique characteristics and challenges. As we utilize data from cryptocurrency markets in our analysis, understanding their market quality is crucial. This helps to contextualize our findings and assess the reliability and stability of the data from these emerging markets.

For the analysis of market quality between traditional stocks and cryptocurrencies, we used daily data from January 2021 to December 2023. This comparison is based on the studies by Griffith and Clancey-Shang (2023) and Barbon and Ranaldo (2022). Mann-Whitney tests revealed statistically significant differences in volatility (p = 0.0) and liquidity (p = 1.513e-101) between the two markets.

The comparison of Amihud illiquidity and volatility between stocks and cryptocurrencies is illustrated in Figures 1.1 and 1.2.

The table 1.1 and figures 1.1, 1.2 show significant differences in the market quality between traditional stocks and cryptocurrencies. Cryptocurrencies exhibit higher volatility and Amihud illiquidity compared to traditional stocks, indicating less stable and less liquid markets.

An analysis of market quality within the cryptocurrency market highlights BNB as a significant asset. Figures 1.3 and 1.4 illustrate the comparison of volatility

Table 1.1: Metrics for stocks and cryptocurrencies

	volatility		amihud_illiquidity	
Ticker	mean	std	mean	std
Stocks				
AAPL	0.265573	0.086375	1.046198e-12	7.861875e-13
CSCO	0.225529	0.077011	1.021845e-12	7.641294e-13
GOOGL	0.306937	0.090546	3.693187e-12	2.874899e-12
MSFT	0.270610	0.080046	1.591732e-12	1.204809e-12
TSLA	0.567867	0.158046	1.139629e-12	9.341485e-13
Cryptocurrencies				
ADA-USD	0.737618	0.334655	1.173601e-10	1.648036e-10
BNB-USD	0.660820	0.477135	1.092709e-12	3.776636e-12
BTC-USD	0.490966	0.180052	2.351309e-17	2.249761e-17
ETH-USD	0.628880	0.271060	1.030768e-15	1.020742e-15
XRP-USD	0.806997	0.502449	3.033944e-11	3.256654e-11

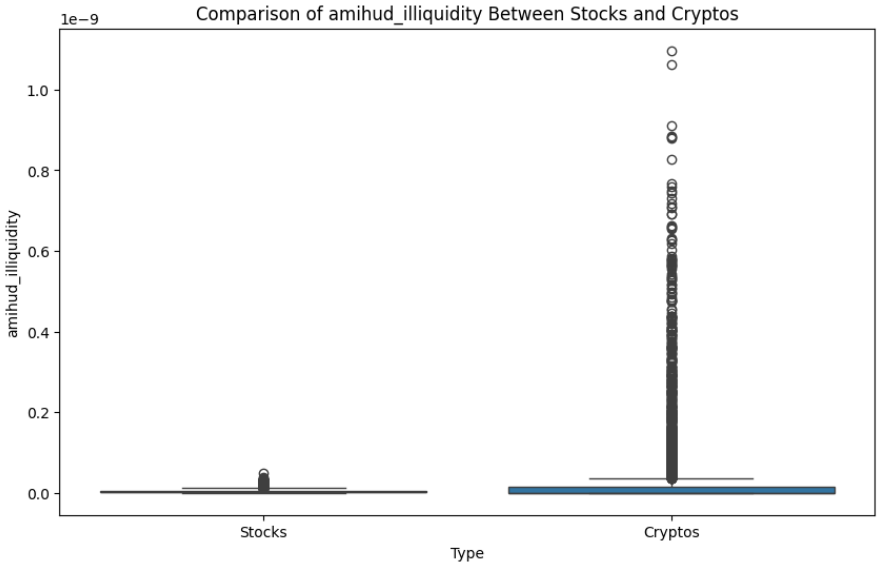


Figure 1.1: Comparison of Amihud illiquidity between stocks and cryptocurrencies

and Amihud illiquidity among various cryptocurrencies, emphasizing the market quality of BNB.

Table 1.2 provides Mann-Whitney test results for volatility and Amihud illiquidity, comparing BNB against other major cryptocurrencies.

The figures 1.3, 1.4 and table 1.2 show that BNB exhibits relatively lower volatility and Amihud illiquidity compared to other major cryptocurrencies, suggesting a more stable and liquid market. We selected Binance Coin (BNB) for this thesis due to its outstanding stability and liquidity, its central role on the Binance platform,

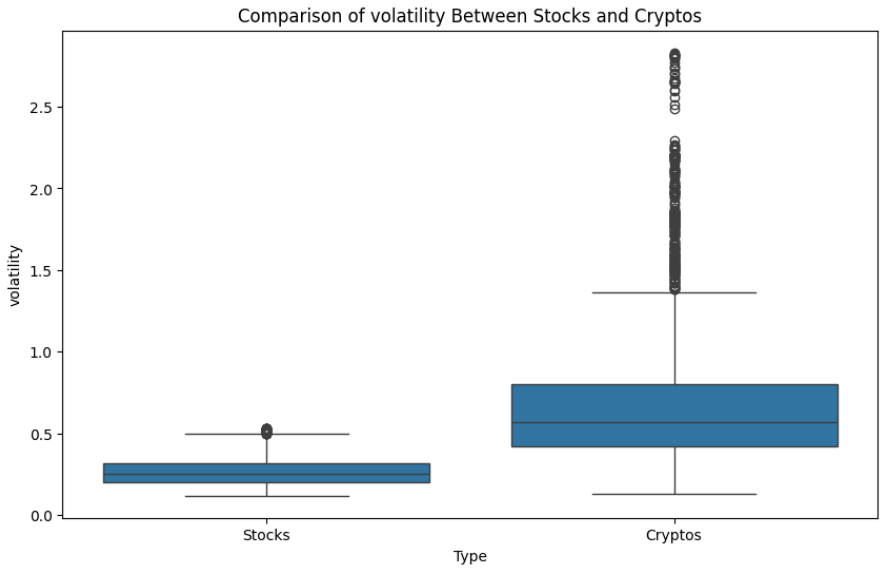


Figure 1.2: Comparison of volatility between stocks and cryptocurrencies

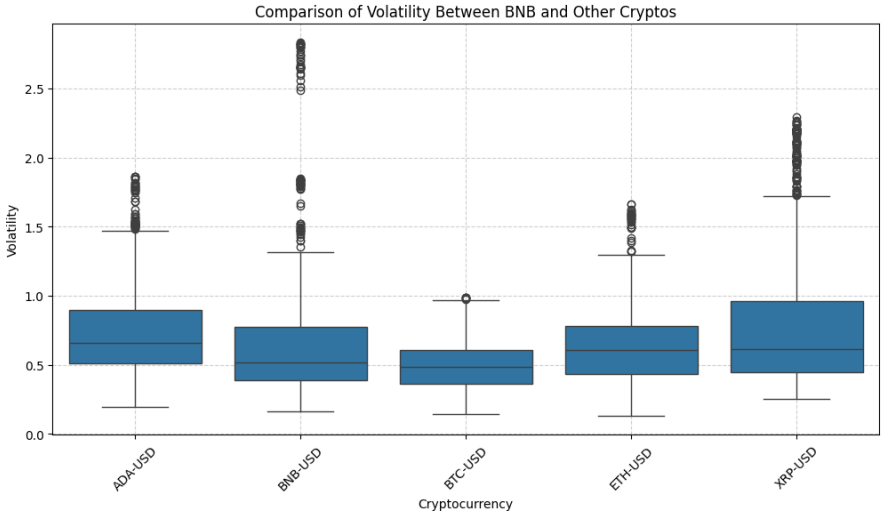


Figure 1.3: Comparison of Volatility Between BNB and Other Cryptos

and its strategic positioning close to market leaders like BTC and ETH, highlighting its prominence in the global cryptocurrency market.

We utilize LOB intraday data with $n = 100$ ticks of depth and a frequency of $\tau = 5$ seconds for $N_t = 1440$ (two hours), from the Binance API for BNB on February 6, 2022.

The choice of $N = 100$ ticks of depth for the calibration is important to capturing the true dynamics of the order book. This specific depth was selected to ensure a comprehensive representation of both bid and ask sides, allowing for a detailed analysis of the liquidity and market depth over time. The two figures 1.5 and 1.6, taken within a short interval, demonstrate how the order book evolves even

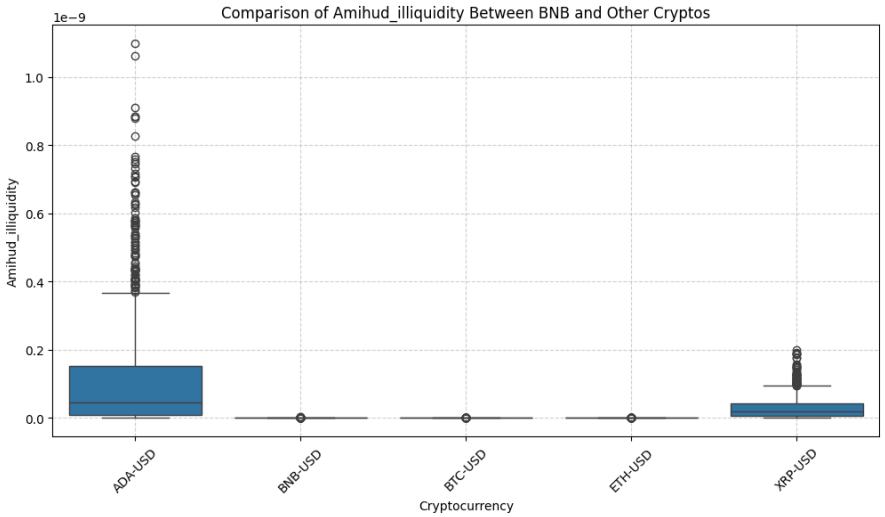


Figure 1.4: Comparison of Amihud Illiquidity Between BNB and Other Cryptos

Table 1.2: Mann-Whitney Test Results for BNB Compared to Other Cryptos

Comparison	Metric	P-Value
BNB-USD vs ADA-USD	volatility	9.66318e-24
BNB-USD vs BTC-USD	volatility	1.57327e-10
BNB-USD vs ETH-USD	volatility	9.79975e-04
BNB-USD vs XRP-USD	volatility	4.24401e-18
BNB-USD vs ADA-USD	amihud_illiquidity	0.0
BNB-USD vs BTC-USD	amihud_illiquidity	0.0
BNB-USD vs ETH-USD	amihud_illiquidity	0.0
BNB-USD vs XRP-USD	amihud_illiquidity	0.0

over a few seconds. This rapid change underscores the importance of selecting an appropriate depth to capture the microstructure of the market.

In practical terms, achieving a depth of 100 ticks requires substantial trading volumes. The figures 1.5 and 1.6 illustrate that even with high-frequency trading, substantial volumes are present at each price level, maintaining the order book's depth. In real market conditions, such volumes are typically observed in highly liquid assets like BNB, especially within an active trading platform like Binance. This depth ensures that the order book data used for calibration is robust and reflects real trading activities, providing a reliable basis for subsequent analyses.

The figures 1.5 and 1.6 demonstrate that significant bid and ask volumes exist at various price levels, ensuring that the chosen depth of 100 ticks captures the necessary market depth and provides a comprehensive view of market liquidity. This justification supports the robustness of our calibration method, leveraging high-frequency data to analyze market microstructure effectively.

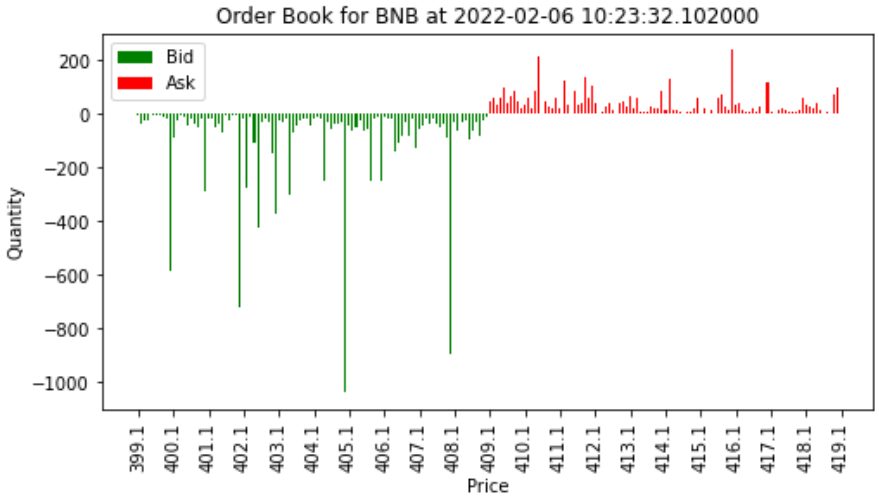


Figure 1.5: Order Book for BNB on 2022-02-06 at 10:23:32.102000

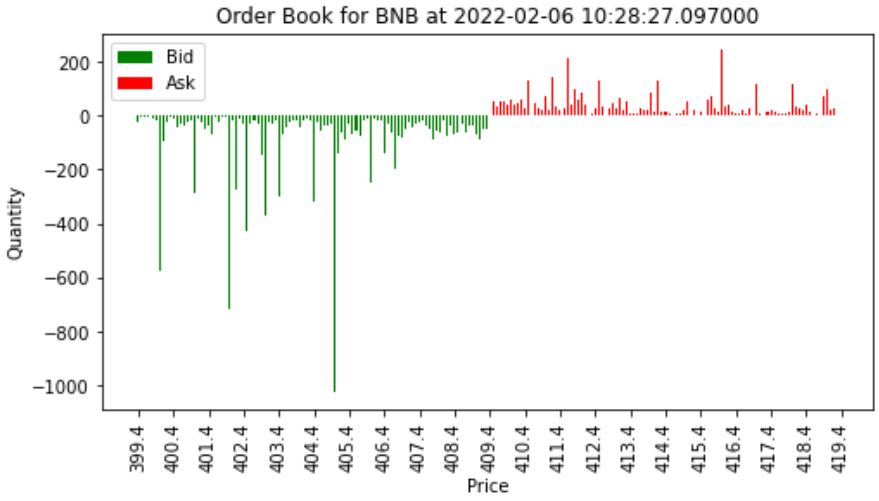


Figure 1.6: Order Book for BNB on 2022-02-06 at 10:28:27.097000

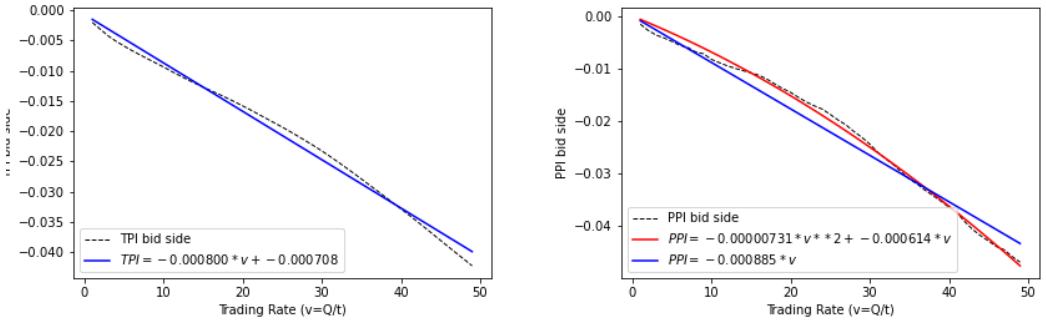
1.4.3 Model parameters calibration

A primary consideration in the calibration involves capturing, from the LOB, the total available volume at each moment that can be liquidated. Specifically, for each moment, we sum up the available volumes (number of shares per tick) of the 100 ticks. This information plays a crucial role in calibrating the functional forms of both the Temporary Price Impact (TPI) and the Permanent Price Impact (PPI). If the volume of shares to be liquidated over a time lapse τ is small relative to the available volumes in the LOB, the transaction barely walks the book, indicating an underestimation of the impact. In this scenario, the calibration generates a quadratic functional form ($ax^2 + bx$ with $a < 0$), termed underestimation. Conversely, if the volume of shares to be liquidated exceeds the available volumes, the impact is

overestimated, as transactions have to walk the book. This leads to a quadratic functional form ($ax^2 + bx$ with $a > 0$), known as overestimation. Moreover, in the overestimated scenario, after reaching a certain volume, the full depth of the LOB is utilized, and therefore the impact does not increase further.

For the estimation, we focus on liquidation rates, observing the effect on $\nu = Q/\tau$ since liquidation volumes are considered within a time lapse τ. Given the data, for the underestimation scenario, we consider $\nu_{max} = 50$; for the overestimation scenario, $\nu_{max} = 7000$; and for an intermediate value between these two cases, we use $\nu_{max} = 1200$. This intermediate scenario leads to a calibration of the functional forms of TPI and PPI that is very close to a linear function ($ax^2 + bx$ with $a = 0$), named average-estimation. In all scenarios, we also calibrate the linear functions of TPI and PPI. A key limitation of this scenario-based approach is the specific quantity that defines each scenario. This definition relies on assumptions that may not always hold true across different market conditions, necessitating further studies to refine these definitions and ensure their applicability. However, this study offers a first approximation to understand how considering specific scenarios, based on inventories to be liquidated versus inventories on hand, is an adequate way to face the information limitations imposed by the market itself.

The underestimated scenario is illustrated in figure 1.7, where subfigure (1.7(a)) displays the plot of trading rates ν_i versus TPI_i^{bid}, for volumes $1 \leq i \leq m$, as well as the plot of the linear calibrated model $TPI^{bid}(\nu) = f(\nu) = a_1\nu + a_2$. Subfigure (1.7(b)) presents plots of trading rates ν_i versus PPI_i^{bid} and the calibrated PPI models $PPI^{bid}(\nu) = g(\nu) = b_1\nu$ and $PPI^{bid}(\nu) = g(\nu) = c_1\nu^2 + c_2\nu$. A preliminary visual assessment indicates that the quadratic function might provide a closer fit to the data than the linear function; however, a definitive comparison will be made through the analysis of residuals and the coefficient of determination. We employ both functional forms to compare the efficacy of the respective optimal policies.

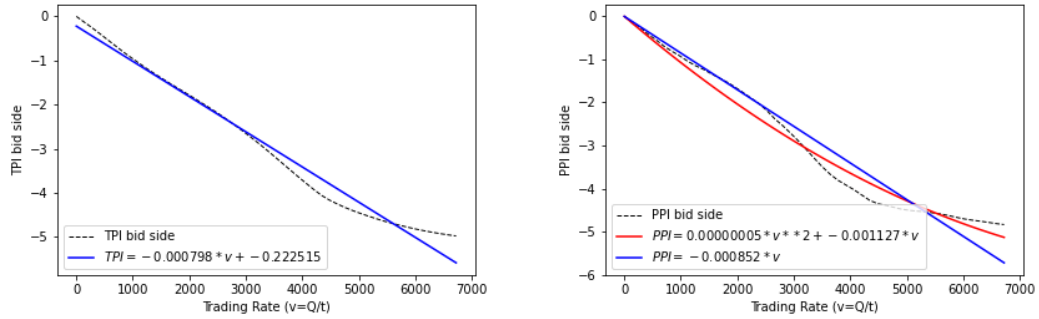


(a) TPI

(b) PPI

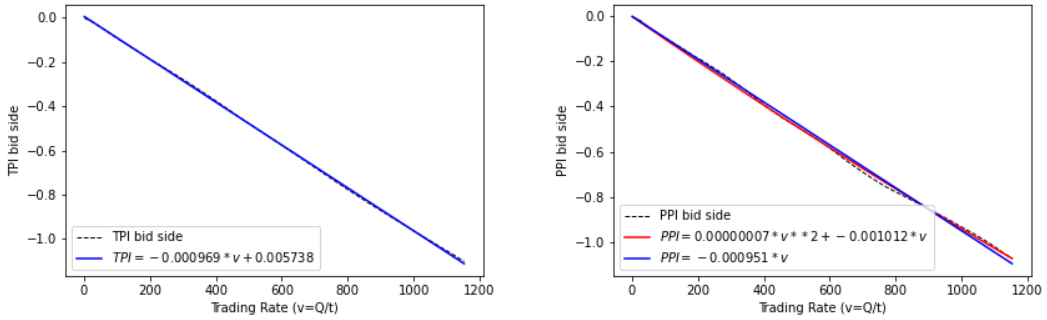
Figure 1.7: Underestimation of linear TPI and linear and quadratic functional forms of PPI with $\nu_{max} = 50$, $m = 50$

Similarly, in figure 1.8, we present the overestimated scenario and specifically note the changes in the concavity of the calibrations, transitioning from concave in



(a) TPI (b) PPI

Figure 1.8: Overestimation of linear TPI and linear and quadratic functional forms of PPI with $\nu_{max} = 7000$,$m = 50$



(a) TPI (b) PPI

Figure 1.9: Average-estimation of linear TPI and linear and quadratic functional forms of PPI with $\nu_{max} = 1200$,$m = 50$

underestimation to convex in overestimation. Initial visual observations suggest that the quadratic form may align more closely with the empirical data than the linear one. Additionally, the average-estimation scenario is depicted in figure 1.9.

The tables (1.3, 1.4, 1.5) show the calibrated parameters for the different functional forms in different scenarios. Table 1.3 presents the results for underestimation, where all coefficients have negative and significant values with a narrow confidence interval, indicating precise parameter estimation. In this scenario, the ratio a_1/c_1 is positive, which is crucial because only if this ratio is positive, the optimal policy derived from a quadratic PPI is applicable, ensuring a positive trading rate. This implies that simulations are feasible only in this scenario. Table 1.4 and Table 1.5 show results where the ratio a_1/c_1 is negative, indicating that the optimal policy is not applicable in these scenarios due to the resulting negative trading rate. The results across all scenarios demonstrate that all calibrated parameters are significant with a 95% confidence level, underscoring the robustness and reliability of the parameter estimations. The consistent significance and narrow confidence

Table 1.3: Underestimation

Parameter	Coefficient	t-statistic	p-value	95% Confidence Interval
a1	-0.0007997	-71.6855	0	[-0.000822, -0.000777]
a2	-0.0007083	-2.1105	0.032	[-0.001353, -0.000064]
b1	-0.0008852	-80.2525	0	[-0.000907, -0.000863]
c1	-7.308e-06	-15.6916	0	[-0.000008, -0.000006]
c2	-0.0006139	-34.3821	0	[-0.000650, -0.000578]

Table 1.4: Overestimation

Parameter	Coefficient	t-statistic	p-value	95% Confidence Interval
a1	-0.0007981	-48.8813	0	[-0.000831, -0.000765]
a2	-0.2225147	-3.4936	0.00105	[-0.350646, -0.094384]
b1	-0.0008516	-63.3839	0	[-0.000879, -0.000825]
c1	5.401e-08	8.0855	1.9e-10	[0.000000, 0.000000]
c2	-0.0011266	-32.0681	0	[-0.001197, -0.001056]

intervals of the parameters suggest that the functional specifications of the TPI and PPI are correct.

To determine which of the two functional forms, linear or quadratic, best fits the average Permanent Price Impact (PPI) across different scenarios, we employ squared residuals, defined as $\hat{\epsilon}^2_{i,t} = (y_i - \hat{y}_i)^2_t$, the corresponding R-squared and Adjusted R-squared where y_i represents PPI^{bid}_i and $\hat{y}_i$ is $PPI^{bid}(\nu_i)$.

The comparison of squared residuals and R^2 values for Permanent Price Impact (PPI) is shown in Table 1.6. This assessment includes both linear and quadratic models across scenarios of overestimation, average estimation, and underestimation. In scenarios with overestimated impacts, the quadratic model shows a superior fit, evidenced by significantly lower total and mean squared residuals, and a higher R^2 value of 0.978687, compared to the linear model's 0.949041. This trend continues in the average estimation scenario, where the quadratic model almost perfects the fit with an R^2 value of 0.999017, surpassing the linear model's 0.997896. The underestimation scenario further highlights the quadratic model's dominance, showing the lowest squared residuals and an R^2 value of 0.995832, in contrast to the linear model's 0.973995. The consistent outperformance of the quadratic model across all metrics and scenarios clearly illustrates its enhanced capacity to capture the nuances of PPI, affirming its general advantage over the linear model for modeling the permanent price impact. Additionally, the adjusted R^2 values confirm the robustness of the quadratic model, maintaining superior fit across all scenarios.

The Epps-Singleton test results, summarized in Table 1.7, further validate these findings. This test evaluates the null hypothesis that the two samples come from the same distribution. Lower p-values indicate rejection of the null hypothesis,

Table 1.5: Average Estimation

Parameter	Coefficient	t-statistic	p-value	95% Confidence Interval
a1	-0.0009689	-489.0427	0	[-0.000973, -0.000965]
a2	0.0057379	4.3262	7.86e-05	[0.003070, 0.008406]
b1	-0.0009505	-300.9221	0	[-0.000957, -0.000944]
c1	7.082e-08	7.3184	2.69e-09	[0.000000, 0.000000]
c2	-0.0010124	-115.9461	0	[-0.001030, -0.000995]

Table 1.6: Comparison of Squared Residuals and R^2 for PPI ($\hat{\epsilon}^2$)

Scenario	Model	Total Sum	Mean	Std Dev	R^2	Adj R^2
Overestimation	Linear	6.4601408	0.1318396	0.1798706	0.949041	0.949757
Overestimation	Quadratic	2.7019078	0.0551410	0.0488995	0.978687	0.977760
Average Estimation	Linear	0.0105163	0.0002646	0.0002028	0.997896	0.997851
Average Estimation	Quadratic	0.0049152	0.0001003	0.0000446	0.999017	0.998974
Underestimation	Linear	0.0002361	0.0000063	0.0000044	0.973995	0.973442
Underestimation	Quadratic	0.0000378	0.0000010	0.0000011	0.995832	0.995561

suggesting that the samples do not come from the same distribution. For the overestimation scenario, the quadratic model exhibits a higher p-value (0.424950) compared to the linear model (0.049153), indicating that the residuals from the quadratic model are more likely to come from the same distribution as the actual data, thus a better fit. Similarly, in the average estimation scenario, the quadratic model shows a p-value of 0.997836, significantly higher than the linear model's 0.966549, reinforcing the superior fit of the quadratic model. In the underestimation scenario, the quadratic model again shows higher p-values, confirming its robustness across different scenarios.

These results collectively demonstrate that the quadratic model consistently outperforms the linear model in fitting the Permanent Price Impact, providing a more accurate and reliable representation of the market dynamics.

The empirical evidence supporting the functional forms for price impacts is

Table 1.7: Consolidated Epps-Singleton Test Results

Scenario	Model	Test Statistic	p-value
Overestimation	Linear	9.529088	0.049153
Overestimation	Quadratic	3.862384	0.424950
Average Estimation	Linear	0.568035	0.966549
Average Estimation	Quadratic	0.134536	0.997836
Underestimation	Linear	3.903857	0.419174
Underestimation	Quadratic	1.678384	0.794640

presented through various captures of order books (LOBs) at different points in time. Figures 1.10 and 1.11 illustrate the order book for BNB at two distinct timestamps on 2022-02-06, demonstrating the evolution of bid and ask volumes. These captures validate the chosen depth of 100 ticks for initial calibrations, showing that the order book maintains significant depth and volume over time.

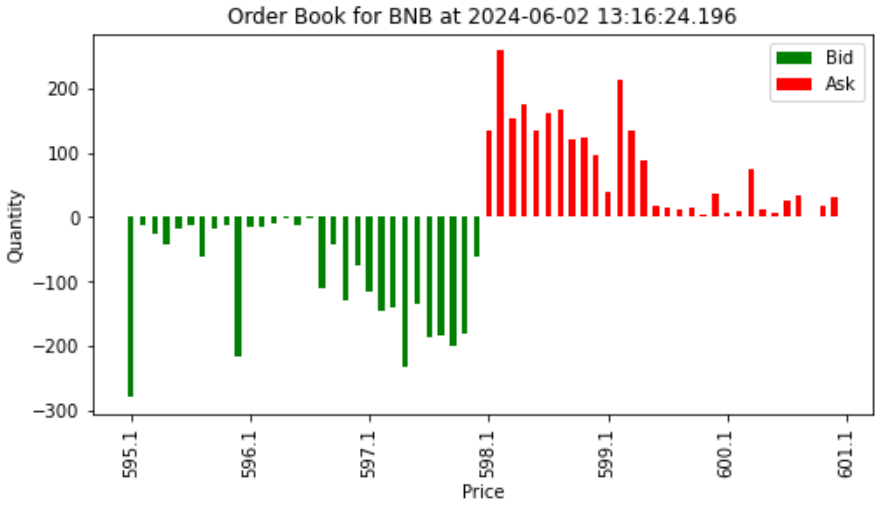


Figure 1.10: Order Book for BNB at 2022-02-06 10:23:32.102000

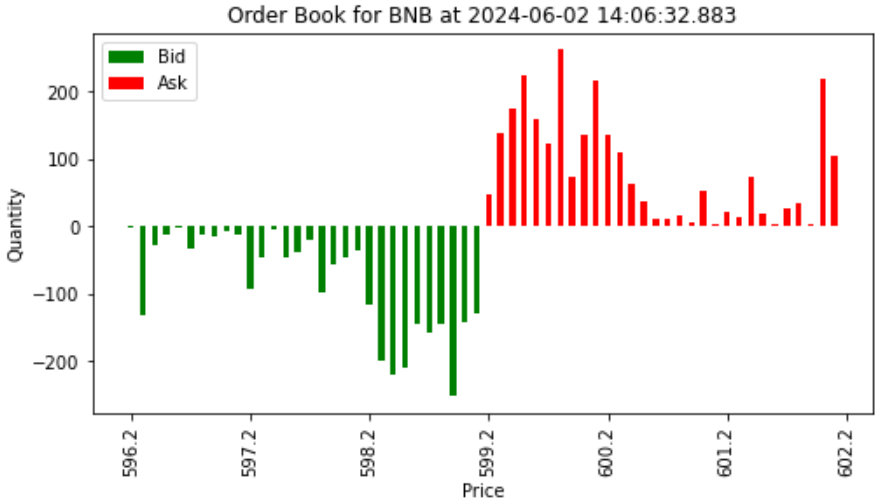


Figure 1.11: Order Book for BNB at 2022-02-06 10:28:27.097000

The stability of the functional forms over time is analyzed through the comparison of these LOB captures. The consistent presence of significant bid and ask volumes at various price levels indicates that the chosen depth and frequency of data collection (5 seconds) are appropriate for capturing the dynamics of the order book.

Further, the analysis explores the presence of convex and concave shapes in the data, particularly at higher levels of order book depth. Figures 1.12, 1.13, and

1.14 illustrate scenarios of underestimation, overestimation, and average estimation using a depth of 30 ticks, taken at a different time and frequency (5 minutes) compared to the initial 100 ticks. These figures demonstrate that the functional forms remain consistent across different depths and frequencies.

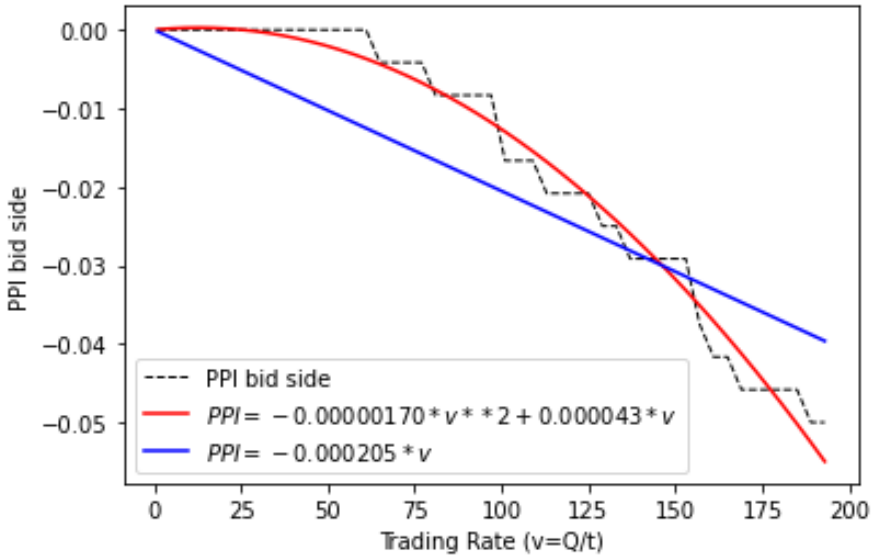


Figure 1.12: Underestimation Scenario with Quadratic PPI at 30 Ticks Depth

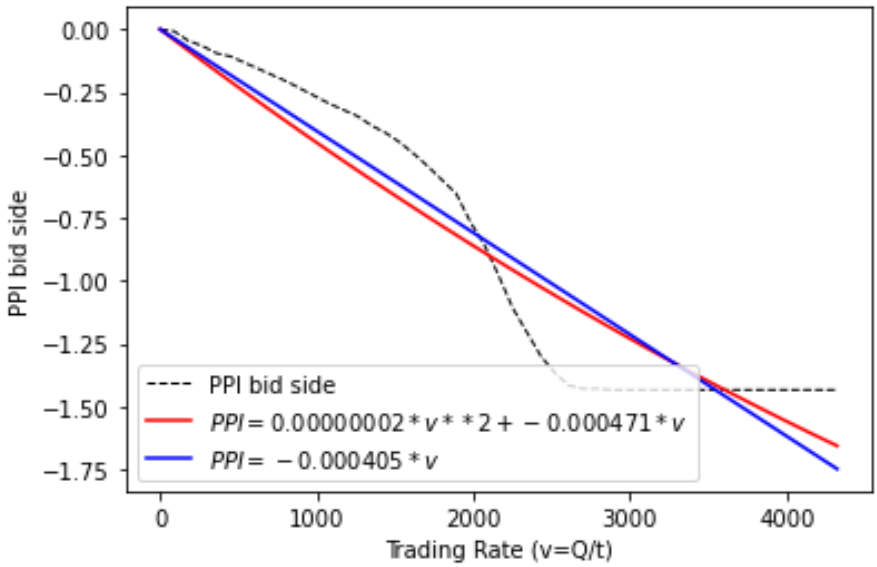


Figure 1.13: Overestimation Scenario with Quadratic PPI at 30 Ticks Depth

The figures 1.10 and 1.11 show that the functional forms for price impacts, whether convex or concave, remain stable over different time frames and depths. This stability confirms the robustness of the forms across varying market conditions

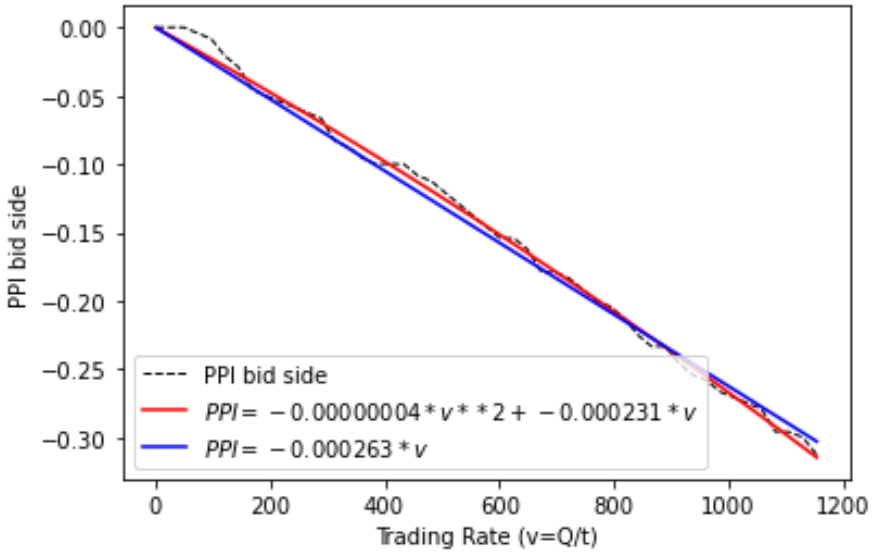


Figure 1.14: Average Estimation Scenario with Quadratic PPI at 30 Ticks Depth

and data collection parameters, providing strong empirical support for their use in this study.

1.4.4 Simulation

This section presents simulations, in the form of backtesting using actual data from Binance, for the liquidation of Υ shares of BNB using the analytically calculated optimal policies. The simulation involves taking an initial inventory Υ and liquidating it over the period $[0, T]$, walking the book when necessary, and following each of the policy paths. We then compare the performance of the analytical optimal strategies against the naive strategy (NS), which consists of liquidating all cryptocurrencies at a single prespecified time.

Following Cartea and Jaimungal (2016b), we assume that the order book is resilient, meaning, after executing a Market Order (MO), the book clears that order and returns to its original state, without any modifications.

Recall, if $f(\nu_t) = a_1\nu_t + a_2$, and $g(\nu_t) = b_1\nu_t$, then the optimal strategy for liquidating q_t shares within the interval $[t, T)$ is $\nu_t^* = \frac{q_t}{T-t}$. Similarly, if $f(\nu_t) = a_1\nu_t + a_2$, and $g(\nu_t) = c_1\nu_t^2 + c_2\nu_t$, the optimal strategy for liquidating q_t shares within the interval $[t, T)$ becomes $\nu_t^* = \frac{2(c_1 q_t + a_1)}{3c_1(T-t)}$, where q values are non-negative, bounded, and $a_1/c_1 > 0$.

Previously, we calibrated the functional forms for the TPI and PPI impact models using BNB LOB data, specifically $f(\nu_t) = a_1\nu_t$ and $g(\nu_t) = c_1\nu_t^2 + c_2\nu_t$, determining the values of all coefficients. We paid special attention to the parameters a_1 and c_1 because the ratio a_1/c_1 needs to be positive for the optimal policy derived from a quadratic permanent price impact to be applicable. As previously found, this condition is met only in the UNDER scenario. Therefore, simulations were developed exclusively for this scenario.

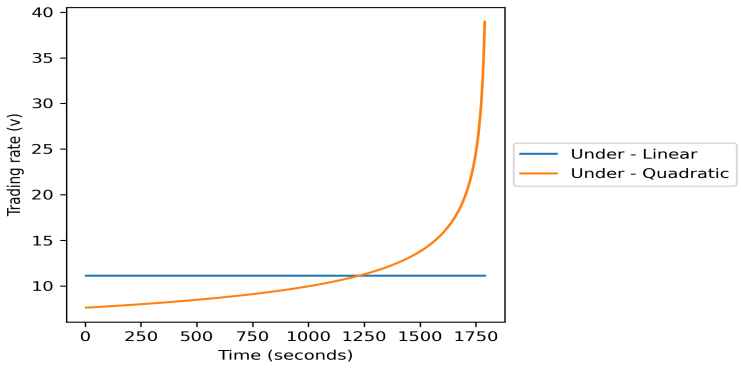


Figure 1.15: Analytical optimal policies in underestimation scenario of linear and quadratic functional forms of PPI

The behavior of these policies becomes observable through the behavior of the inventory on hand as each strategy is executed. Figure 1.16 shows that the inventory associated with the policy derived from linear PPI decreases steadily, forming a straight line; in the case of the inventory associated with the policy derived from quadratic PPI, we observe a certain concavity in the functional form of the available inventory. This aspect will be analyzed in greater detail in the next study.

The graphs of the analytical optimal policies appear in Figure 1.15 and are constructed for a time horizon of 1800 seconds, using the calibrated parameters a_1 and c_1 with $\Upsilon = 4000$. The policy derived from linear PPI remains constant for a fixed inventory Q_t, similar to when there is no permanent impact as per Cartea et al. (2015). The policy derived from quadratic PPI starts by liquidating more slowly than the linear one and accelerates over time to liquidate all inventory before reaching time T, avoiding penalties.

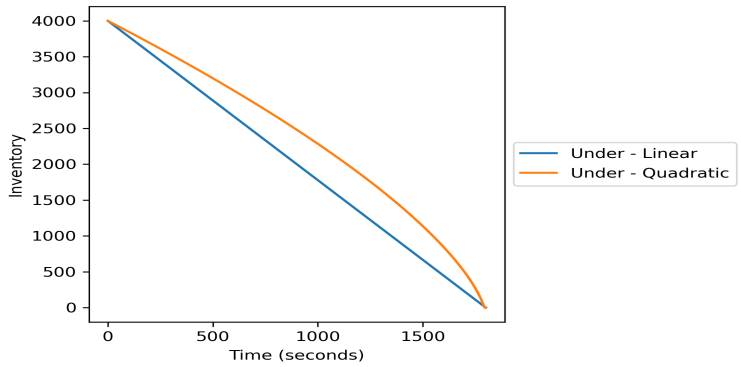


Figure 1.16: Analytical optimal policies in underestimation scenario of linear and quadratic functional forms of PPI

When we simulate the liquidation of $\Upsilon = 4000$ BNB [5] over 1800 seconds, executing every 5 seconds using the analytically derived optimal policies from

[5]we choose a value of 4000 to ensure that this inventory is liquidatable at a single point in time

linear and quadratic PPI functions in the UNDER scenario, and compare their performance against NS, we find that both policies outperform the Naive Strategy. For these two cases, the results show 0.9672% higher performance for the linear case and 1.0077% for the quadratic case, confirming that a better fit of the permanent impact function to the empirically found functional form leads to improved results in terms of maximizing the agent's revenue, the observed concavity of the optimal inventory path also appears to play a significant role in outlining optimal policies that yield higher revenues. After simulating 2000 LOBs with 100 ticks of depth at 5-second intervals over two hours and simulating the execution of linear PPI strategies against NS and quadratic PPI against NS, we obtained the following 95% confidence intervals: 95% Confidence Interval around 0.9672%: $(0.9605, 0.9739)$ and 95% Confidence Interval around 1.0077%: $(1.0009, 1.0145)$, indicating that, with 95% confidence, the improvement of the optimal linear PPI and quadratic PPI strategies over the naive (NS) strategy is not zero, i.e., they are better than the naive strategy at this level of confidence. The nominal improvements are as follows: Nominal improvement for linear PPI strategy: 15722.60 and Nominal improvement for quadratic PPI strategy: 16379.87.

1.5 Discussion and Conclusions

In this work, we study the problem of optimal liquidation by analyzing the consequences of having different functional forms, specifically linear and quadratic, for permanent price impact. We apply this theory to the cryptocurrency market, recognizing cryptocurrencies as rapidly evolving digital assets that not only enhance the diversity of investment portfolios but also stand at the forefront of financial technology innovation, prompting a need for in-depth analysis to understand their impact on the traditional financial systems and the broader economy. Therefore, we use LOB data for BNB cryptocurrency, obtained from the Binance API.

In the *Analytical Solutions* section, we extend closed-form solutions for the $\mathcal{HJB}$ Partial Differential Equation, considering scenarios with linear Temporary Price Impact (TPI) and both linear and quadratic Permanent Price Impact (PPI). The lemmas provide solutions for the PDE $\mathcal{HJB}$ under specific conditions (1.6) and (1.7), also offering solutions for the respective optimal liquidation strategy. We note that the linear PPI derivative strategy remains constant while the quadratic PPI derivative initially settles slowly to increase in speed before reaching the set time horizon.

We use Order Book (LOB) intraday data for Binance Coin (BNB) from the Binance API. This data, comprising 100 ticks at a frequency of 5 seconds over a two-hour period, is used to calibrate the TPI and PPI models. The calibration process is nuanced, incorporating three distinct scenarios based on the relationship between the volume of shares to be liquidated and the available volumes in the LOB: underestimation, overestimation, and average-estimation. These scenarios influence the functional forms of TPI and PPI, which in this study we approximate through linear and quadratic expressions. We can also conclude from the empirical

evidence that the scenarios hold at different frequencies, depths and days.

The key findings from this analysis reveal that the quadratic model for PPI consistently outperforms the linear model across all three scenarios, evidenced by a comparative analysis of squared residuals and R-squared values. Particularly in overestimated and underestimated scenarios, the quadratic model shows a significantly better fit, with lower squared residuals and higher R-squared values. In the average-estimation scenario, the quadratic model achieves a near-perfect fit, indicated by an R-squared value of 0.999017, compared to the linear model's 0.997896. These results highlight the superior capacity of the quadratic model to accurately capture the nuances of Permanent Price Impact in various market conditions, asserting its overall effectiveness over the linear model in assessing permanent price impact in financial trading.

In the simulation section, backtesting with Binance data assesses BNB liquidation strategies, comparing analytically derived optimal policies to a naive strategy. Utilizing calibrated TPI and PPI models, the study finds that both linear and quadratic policies outperform the naive strategy, with the quadratic policy demonstrating superior performance. The calibrated parameters a_1 and c_1 from the BNB LOB data are crucial in this analysis, revealing that only the underestimation scenario allows for feasible simulations. These simulations confirm the effectiveness of these policies, notably the quadratic approach, highlighting the importance of accurately modeling permanent price impact in financial trading.

The scope of this study is limited, primarily by the available LOB data and the liquidity requirements of the assets being liquidated, essential for processing inventories within constrained time horizons. The study recognizes the need for more general functional forms in the models, as finding analytical solutions for price impacts often requires numerical methods. The current model does not account for several key variables that could significantly influence optimal liquidation policy, such as volatility variability and dynamic market factors. Future research aims to integrate more general model forms, including numerical methods, to broaden the study's scope and practical applicability.

2 Optimal liquidation with temporary and permanent price impact, evidence from the cryptocurrency market

Abstract

This paper studies a real-life application of the optimal liquidation of stocks in the presence of temporary and permanent price impacts. Using data from the order book of the BNB cryptocurrency, we estimate the functional form of the temporary and permanent price impact in three different scenarios: underestimation, overestimation, and average estimation, finding different functional forms for each scenario. Using finite differences and optimal policy iteration, we solve the problem numerically and observe interesting changes in the optimal liquidation policy when applying calibrated linear and power forms for the temporary and permanent price impacts. Then, with these optimal policies, we identify optimal liquidation trajectories and simulate the liquidation of initial inventories to compare the performance among the optimal strategies under different parametrizations and against a naive strategy (NS) and a common strategy (CS). Finally, we characterize the optimal policies based on a proposed functional form for the inventory and find that policies generating the highest revenue are those starting with a low trading rate and increasing with time.

Keywords: Cryptocurrency Liquidation, Price Impact Analysis, Optimal Trading Strategies, Finite Difference Methods, Stochastic Modeling, Order Book Microstructure

2.1 Introduction

The market of cryptocurrencies is increasing in importance as the net value of the market grows from \$3.8 billion USD in 2015 to \$943.41 billion USD in the second quarter of 2022, reaching a maximum of over \$3000 Billion USD[1]. Likewise, the use of cryptocurrencies offers advantages such as its articulation with Decentralized Finance (DeFi)[2], security, and transparency, among others, which, according to Blanco Encinosa (2021), make them very convenient investment alternatives. However, this type of investment is highly volatile, and many investors may need or wish to acquire or liquidate their positions optimally at specific times. In this paper, we focus on the optimal strategies for liquidation of assets considering price impacts, specifically on different forms of temporary and permanent price impact calibrated from cryptocurrencies.

[1]Data from https://www.statista.com/statistics/730876/cryptocurrency-maket-value/ and https://coinmarketcap.com/

[2]In accordance with Piñeiro-Chousa et al. (2022) DeFi can be understood as the process that makes use of the blockchain for the development and implementation of novel financial products and services.

Optimal liquidation consists of finding the optimal strategy an agent must follow when selling large amounts of shares over a finite period of time to minimize adverse effects, the consequences of her actions. The time dependence is crucial because if the agent executes in a short time, by selling fast, the rate of execution increases, and so does the impact on the price of shares. But if she sells at a slow pace over a long period, holdings are exposed to greater uncertainty due to the volatility, crucial in the case of cryptocurrencies. Thus, in our setting, the agent must sell a number Υ of shares within a given timeframe $[0, T)$, where she must find the optimal selling rate ν_t such that the revenue from liquidation is maximized. To this end, we adopt an optimization problem based on a model that includes linear and non-linear impact functions, calibrated with data from the Limit Order Book (LOB) of the BNB cryptocurrency. To tackle the problem, we employ stochastic control techniques which have been incorporated into the literature since the seminal paper Bertsimas and Lo (1998), where they offer a precise definition of optimal execution, based on the concept of minimization of execution costs using stochastic dynamic programming. To generalize the methodology, we also utilize numerical methods, that is, finite differences over a non-linear Partial Differential Equation (PDE). As a sanity check for the numerical solution, we work with two extensions of analytical solutions of Cartea et al. (2015) which are presented in Sánchez and Ramírez (2024).

The well-known model for optimal execution of transactions in Almgren and Chriss (2000) establishes two types of impact for the asset price: temporary and permanent. In their work, temporary impact refers to "temporary imbalances in supply and demand caused by our trading leading to temporary price movements away from equilibrium" and permanent impact "means changes in the equilibrium price due to our trading which remain at least for the life of our liquidation". Theoretical approaches agree that the permanent price impact must be linear, Gatheral (2010) shows this condition is necessary to avoid dynamic arbitrage. He presents *the principle of no-dynamic-arbitrage* which states that the trading cost is non-negative for any round-trip trade strategy, i.e. $\int_0^T \nu_t dt = 0$, and shows that non-linear permanent market impact is inconsistent with the principle of no-dynamic-arbitrage. However, our study, focusing only on liquidating shares, indicates that the sum of this sequence of trades cannot be zero; therefore, the principle of no-dynamic arbitrage is not applicable. Consequently, our model considers non-linear permanent price impacts, which are estimated from data, and studies the effect of such impact functions on the optimized trading rate.

Additionally, empirical evidence indicates that market impact is not linear, but rather similar to a square root. That is, it is proportional to the square root of the volume of shares executed, as seen in Toth et al. (2016) and Almgren et al. (2005), or to the square root of the trade duration, as in Bershova and Rakhlin (2013). Moreover, different approaches using non-linear temporary or permanent price impact have been studied in the literature, for example, Gueant (2014) relates to the result in Gatheral (2010) and extends some theoretical results to the non-linear setting. Barger and Lorig (2018) offers a theoretical approach, considering

stochastic temporary and permanent price impacts. Alfonsi and Schied (2010) proposes a model where the impact depends on the theoretical shape of LOB given by a density (exponential). Whereas our model does not rely on any density and is thus capable of capturing price impacts from the actual book using simulations. More recently, Brunovský et al. (2018) solves for optimal liquidation with linear price impact, although importantly shows some theoretical advances by adding the possibility of acquiring shares as well as selling.

Although related literature presents different perspectives on the solution of the optimal liquidation problem for large amounts of shares, in this paper, we study the problem based on the setting proposed in Cartea et al. (2015), but using more general functional forms for temporary and permanent price impact (TPI and PPI respectively). This study, unlike the work of Cartea et al. (2015), generalizes the solution using finite differences and the functional forms of PPI and TPI. Additionally, we introduce the concept of a scenario associated with the quantity of assets to be liquidated versus the availability in the order book and characterize the optimal policies obtained through the optimal available inventory. We find that concave inventories represent optimal policies that generate higher revenues across all scenarios. Since analytical solutions are scarce for these general models, we apply numerical methods, specifically finite differences and optimal policy iteration, and solve the problem for different forms of the TPI and PPI functions, i.e., functions resulting from calibration.

Alternating Direction Implicit (ADI) finite differences are a numerical method used for solving partial differential equations (PDEs) that are parabolic, elliptic, or hyperbolic. This method is based on the idea of decomposing the solution of the problem into a series of simpler steps by alternating the direction of the finite difference at each step. In each step, a system of linear equations is solved implicitly in one direction, while in the other direction, an explicit scheme is used. This approach reduces computational complexity and improves stability compared to fully explicit methods Peaceman and Ratchford Jr (1955); Thomas (1995); Duffy (2006).

This type of problem, i.e., a PDE with two dimensions in space and only one diffusion, is frequently numerically solved using the ADI scheme, but has some stability issues, as shown in the classical work of Peaceman and Ratchford Jr (1955), hence also, for example, in Thomas (1995), Tavella and Randall (2000), and more recently by Duffy (2006). To improve convergence in the multi-dimensional finite differences, we propose the use of an alternative scheme, namely the implicit directional scheme.

On the other hand, using the technique in Cartea et al. (2015), we calibrate, in a cryptocurrency market, the functional forms of the TPI and PPI functions. A first aspect to highlight is that calibrated parameters of the functional forms for the TPI and PPI change with the depth in ticks and therefore the volumes of the LOB. That is, given a fixed LOB setting (depth in ticks and volumes), if the inventory to liquidate is sufficiently small, the liquidation can be done barely walking the book, so that the impact would be small and increasing as the liquidation target is reached, this configures a power form x^a, where $a > 1$. The opposite case occurs

when the inventory to be liquidated is large for the LOB setting, since at first the impact would be large, but decreasing as the liquidation continues, this configures a power form x^a but having $0 < a < 1$. As a third scenario, we find an in-between point, where the TPI and PPI approach a linear form, i.e., x^a with $a = 1$. Using these three configurations for calibration gives different functions for TPI and PPI, and we are able to identify the different optimal trading policies for each of these scenarios. Thus, in accordance with Lillo et al. (2003) and Almgren et al. (2005), modeling the problem must include power forms for TPI and PPI, but for these more complex and general scenarios, the use of numerical techniques to find solutions is compulsory.

We find that the optimal liquidation strategies change for different TPI and PPI functions, and that they outperform a naive strategy of liquidating all inventory at a single point in time (NS) and a common strategy of liquidating all inventory proportionally and constantly over the entire time period (CS), as seen in Cartea et al. (2015), which we confirmed with simulations. Also, we observe that, we may describe these strategies in a common functional form of the optimal inventory, as a non-increasing monotonic function, $q(t) = \Upsilon \left(1 - \frac{t}{T}\right)^{d_2}$, subject to $q(0) = \Upsilon$ and $q(T) = 0$. Then, under certain conditions and using real-life data simulations, similar to paper trading, we recognize that the optimal policies starting with a slow rate and increasing it as time passes have the highest cumulative revenue. Thus, our results, in some sense, confirm the conclusions about the concavity of market impact in Curato et al. (2017), stating that concave functions of the inventory generate more profit. Therefore, encouraged by two analytical extensions of the results in Cartea et al. (2015), we reaffirm that trading policies generating concave inventory paths are the ones having better performance, implying that the dependence between optimal policies and functional form of TPI and PPI is not binding. This has important potential in practice, as it eliminates the hustle of calibrating models and thus saves time by implementing directly the optimal strategy.

This paper is structured as follows: Section 2.2 introduces the model, based on Cartea et al. (2015). Section 2.3 presents a general numerical solution using finite differences, for general forms of TPI and PPI. Section 2.4 shows the estimation of possible functional forms of TPI and PPI and the calibration of the parameters for the BNB cryptocurrency, using the LOB data. In Section 2.5, we find optimal policies and analyze these by varying some of the calibrated parameters. We also compare the performance of the numerical strategies with the naive and common strategies that simulate the optimal liquidation of initial inventories on LOB data. Finally, we establish a characterization of the optimal policies from the optimal inventory available, and independent of the calibration. Section 2.6 gives some final remarks.

2.2 The model

We work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and build on the model proposed by Cartea et al. (2015), the agent has an amount Υ of shares of an asset that she wants to liquidate in the time interval $[0, T)$ and obtain the most profit. The inventory changes through time and has dynamics $dQ_t^\nu = -\nu_t dt$, thus ν_t is the trade rate. The mid price of the asset is governed by the SDE

$$dS_t^\nu = -g\left(\nu_t\right) dt + \sigma dW_t, \tag{2.1}$$

where W_t is a standard Brownian motion, $\sigma \in \mathbb{R}^+$ represents the volatility, and $g : \mathbb{R}^+ \to \mathbb{R}^+$ is the permanent price impact function of the trade rate, which is desirable to belong to C^2 to facilitate the optimization process. When an agent sells shares of the stock, the actual price of the transaction is called the execution price and follows

$$\hat{S}_t^\nu = \left(S_t^\nu - \frac{1}{2}\Delta\right) - f\left(\nu_t\right), \tag{2.2}$$

where Δ is the bid-ask spread in the LOB and the function $f : \mathbb{R}^+ \to \mathbb{R}^+$ is the temporary price impact, which should ideally be in C^2 for optimizing the process. Finally, the agent wants to maximize her final expected utility wealth, earnings from transactions, that is

$$\mathbb{E}\left\{\int_0^T \hat{S}_t^\nu \nu_t dt\right\} \tag{2.3}$$

the rate ν_t is controlled by the agent at each time point t and her actions affect the asset's liquidation price, thus, the optimization problem can be represented by the agent's value function

$$H\left(t, S, q\right) = \sup_{\nu_t \in \mathcal{A}(t,T)} \mathbb{E}\left\{\int_t^T \left(\left(S_r^\nu - \frac{1}{2}\Delta\right) - f\left(\nu_r\right)\right) \nu_r dr\right\} \tag{2.4}$$

where $\mathcal{A}(t, T)$ represents the admissible set of non-negative bounded strategies and the equation (2.4) satisfies the $\mathcal{HJB}$ partial differential equation (PDE) [3]

$$0 = \partial_t H + \frac{1}{2}\sigma^2 \partial_{ss} H + \sup_{\nu \in \mathcal{A}} \left\{-g(\nu)\partial_s H - \nu \partial_q H + \left(S - \frac{1}{2}\Delta - f(\nu)\right)\nu\right\} \tag{2.5}$$

subject to

$$H(T, S, q) \;\to\; -\infty, \text{ when } t \to T \text{ and } q > 0, \tag{2.6}$$

$$H(t, S, 0) \;\to\; 0, \tag{2.7}$$

the first condition corresponds to the penalty, negative revenue, for reaching T with positive inventory, and the second is complementary and guarantees that at time t the revenue is zero for nil inventory.

[3] For a detailed explanation, refer to chapter 6 of Cartea et al. (2015)

Although linear PPI is treated in Cartea et al. (2015) and Cartea and Jaimungal (2016b), our approach is essentially different because of the type of penalty, in the range $[0, T)$. Since our penalty is a boundary condition going to $-\infty$, we are not allowing the manager to have any inventory at expiry. On the other hand, Cartea et al. (2015) and Cartea and Jaimungal (2016b) have a quadratic penalty, which, although costly, allows the manager to end up with some inventory. Next we propose the numerical approach for the solution having more general TPI and PPI, namely power functions.

2.3 Numerical approximation

2.3.1 *Finite differences*

We use *implicit finite differences* to find the solution of the PDE (2.5). This technique approximates the partial derivatives through discrete expressions constructed from Taylor expansions and thereby giving an approximate solution of the differential equation.

We denote the discretized version of the function H by $H^k_{i,j} = H\left(k\Delta t, i\Delta s, j\Delta q\right)$, where $t = k\Delta t$, $s = i\Delta s$ and $q = j\Delta q$, additionally we set $S_{max} = N_s\Delta s$, $Q_{max} = N_q\Delta q$ and $T = N_t\Delta t$. As boundary conditions are of great importance in this technique, we state these next:

- When the price is $S = 0$, we think of it as an absorbent state and the possible revenue for potential trades at this price is 0, thus

$$H(t, 0, q) = 0. \tag{2.8}$$

- If the inventory reaches the $q = 0$ level, there is no more shares to sell, having

$$H(t, S, 0) = 0. \tag{2.9}$$

- If the price reaches a maximum (we have to fix a maximum price S_{max} for the numerical approach) the revenue is the amount of shares sold multiplied by the fixed price. That is the inventory multiplied by the fixed maximum price.

$$H(t, S_{max}, q) = S_{max}q. \tag{2.10}$$

- We avoid the case of having any inventory $q > 0$ at final time T, thus we introduce a heavy penalty (minus infinity to be sure it never happens)

$$H(T, S, q) = -\infty. \tag{2.11}$$

By applying central differences on S, forward differences on q, and backward differences on t, we obtain the discretized version of Equation 2.5 that must be

optimized

$$\begin{aligned}\frac{H_{i,j}^{k+1}-H_{i,j}^{k}}{\Delta t}+\frac{\sigma^2}{2}\left(\frac{H_{i+1,j}^{k}-2H_{i,j}^{k}+H_{i-1,j}^{k}}{(\Delta s)^2}\right)-\nu\left(\frac{H_{i,j}^{k}-H_{i,j-1}^{k}}{\Delta q}\right)\\ -g(\nu)\left(\frac{H_{i+1,j}^{k}-H_{i-1,j}^{k}}{2\Delta s}\right)+\left(i\Delta s-\frac{1}{2}\Delta-f(\nu)\right)\nu=0,\end{aligned} \tag{2.12}$$

which has approximation of orders $O(\Delta t), O(\Delta s^2), O(\Delta q)$.

By grouping similar terms and leaving the unknown terms (time $k\Delta t$ and space $\geq j\Delta q$) on the same side of the equality we obtain

$$AH_{i-1,j}^{k}+BH_{i,j}^{k}+CH_{i+1,j}^{k}=Z_{i,j}^{k*} \tag{2.13}$$

where

$$Z_{i,j}^{k*}=DH_{i,j}^{k+1}+EH_{i,j-1}^{k}+F_i(\nu)$$

$$F_i(\nu)=-\left(i\Delta s-\frac{1}{2}\Delta-f(\nu)\right)\nu$$

and A, B, C, D, and E are defined as follows:

$$\begin{aligned}A&=\frac{\sigma^2}{2(\Delta s)^2}+\frac{g(\nu)}{2\Delta s}\\ B&=-\frac{1}{\Delta t}-\frac{\sigma^2}{(\Delta s)^2}-\frac{\nu}{\Delta q}\\ C&=\frac{\sigma^2}{2(\Delta s)^2}-\frac{g(\nu)}{2\Delta s}\\ D&=-\frac{1}{\Delta t}\\ E&=-\frac{\nu}{\Delta q}\end{aligned}$$

Since there is no diffusion in the q variable the traditional ADI scheme for 2 dimensions leads to oscillations and instabilities. That is why we prefer this approach, namely *Directional 2-dimensional finite differences*, where we use implicit solutions for vectors in the variable S but having the variable q constant, this approach additionally simplifies to a tridiagonal system. More precisely, for each time $k\Delta t$ before expiry, we begin by calculating a boundary condition at $q=0$, and once solved for $q=(j-1)\Delta q$, we step to $q=j\Delta q$ and use the tridiagonal system in (2.13) to solve for all the values of S, and iterate in ascending order over q up to Q_{max}. That is, on a uniform grid, we find the values of the function H for all values of i while iterating in ascending order over j and then stepping backwards through time in k, as shown in figure (2.1).

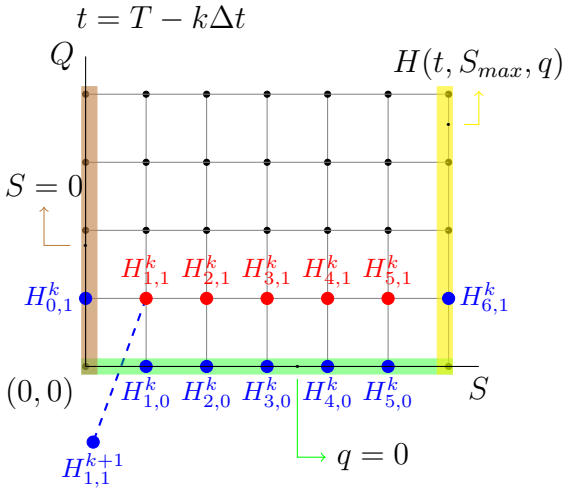


Figure 2.1: This figure shows an example of the uniform grid at time $t = T - k\Delta t$, where the brown highlighted line corresponds to the boundary at $S = 0$, the yellow to the boundary at $S = S_{max}$ and the green to the boundary at $q = 0$. The blue dots correspond to known values and the red dots correspond to unknown values that will be found simultaneously at each step through q using the tridiagonal system (2.13)

2.3.2 *Optimization*

For the optimization process, recalling that $f(\nu)$ is the TPI and $g(\nu)$ is the PPI, we apply optimal policy iteration using first order condition (FOC) on the control ν, that is

$$-g'(\nu)\partial_s H - \partial_q H + S - \frac{1}{2}\Delta - \Big(f'(\nu)\nu + f(\nu)\Big) = 0 \tag{2.14}$$

Specifically, we treat the following cases:

- For linear TPI and quadratic PPI, that is $f(\nu) = a_1\nu + a_2$, and $g(\nu) = c_1\nu^2 + c_2\nu + c_3$ the FOC is

$$-(2c_1\nu + c_2)\,\partial_s H - \partial_q H + S - \frac{1}{2}\Delta - (2a_1\nu + a_2) = 0. \tag{2.15}$$

- For power TPI and power PPI, that is $f(\nu) = r_1\nu^{r_2} + r_3$, and $g(\nu) = p_1\nu^{p_2} + p_3$ the FOC is

$$-p_1p_2\partial_s H\nu^{p_2-1} - \partial_q H + S - \frac{1}{2}\Delta - (r_1(r_2+1)\nu^{r_2} + r_3) = 0. \tag{2.16}$$

Note that linear cases are included in the power form just by taking $r_2 = 1$ in $f(\nu)$ and/or $p_2 = 1$ in $g(\nu)$.

2.4 Model calibration

This section presents the implementation of a method, based on Cartea and Jaimungal (2016a) and Cartea et al. (2015), to estimate functional forms of TPI and PPI. Additionally, to estimate the volatility, we use the realized volatility, i.e, $\hat{\sigma} = \sqrt{\sum_{t=1}^{T} \left(\ln(S_t) - \ln(S_{t-1})\right)^2}$, which is consistent with the model.

2.4.1 *Temporary Price Impact*

To find the functional form of the TPI, following the ideas from Cartea et al. (2015), we use LOB data with n ticks of depth taken at N_t intervals of τ seconds. In each of those N_t time intervals, we simulate the execution of different sizes of Sell Market Orders, i.e. we liquidate the m volumes (As orders of different sizes) $Q = \{Q_1, Q_2, ...Q_i, ...Q_m\}$ against the existing LOB, walking the book when necessary. LOB prices are represented, for each available tick j, by S_j and its corresponding number of Limit Orders by Q_{S_j}. These volumes are chosen increasingly up to a significant maximum relative to the LOB's total capacity (all the posted LOB orders to buy). To find the execution price we define it as the weighted average of the execution prices for each volume, in bid side, that is

$$\hat{S}^{bid}_{i,t} = \frac{\sum_{j=1}^{N_i} S_j Q^*_{s_j}}{Q_i}, \qquad \text{where } Q^*_{s_j} = \begin{cases} Q_i - \sum_{k=1}^{j-1} Q_{s_k} & \text{if } Q_i \leq \sum_{k=1}^{j} Q_{s_k} \\ Q_{s_j} & \text{else.} \end{cases},$$

and $N_i = \min\left\{ l \in \mathbb{N}, \sum_{k=1}^{l} Q_{s_k} \geq Q_i \right\}$, and $0 \leq t \leq N_t \tau$. Since we perform this procedure every τ seconds, we consider not only volumes, but also liquidation rates (ν) in our simulations, i.e. volumes per τ seconds. Thus, by letting S^{bid}_t be the best price on the bid side, the TPI at time t due to the liquidation rate $\nu_i = \frac{Q_i}{\tau}$ is $TPI^{bid}_{i,t} = \hat{S}^{bid}_{i,t} - S^{bid}_t$. Finally, we use the average of $TPI^{bid}_{i,t}$ over time, denoted by TPI^{bid}_i, versus ν_i to fit the models $TPI^{bid}(\nu) = f(\nu) = a_1\nu + a_2$ and $TPI^{bid}(\nu) = f(\nu) = r_1\nu^{r_2} + r_3$, via least squares.

2.4.2 *Permanent Price Impact*

To estimate a functional form of the permanent price impact, we use the same idea as for the temporary impact, i.e. going through the data every τ seconds and at each of those moments executing simulated Market Sell Orders of different volume sizes and therefore different trading liquidation rates. Then we capture the impact of each trading rate, in this case, on the midprice $S^{mid}_t = \frac{S^{bid}_t + S^{ask}_t}{2}$, that is, the difference between the mean price before and after the liquidation of Q_i shares in τ seconds. Using least squares, we estimate the parameters of the models $PPI^{bid}(\nu) = g(\nu) = b_1\nu + b_2$ and $PPI^{bid}(\nu) = g(\nu) = p_1\nu^{p_2} + p_3$.

2.5 Results

2.5.1 Model parameters calibration

We use the same data from the previous paper (Sánchez and Ramírez (2024)), i.e. LOB intraday data with $n = 100$ ticks of depth and a frequency of $\tau = 5$ seconds for $N_t = 1440$ (two hours), from Binance API for BNB [4] on February 6, 2022.

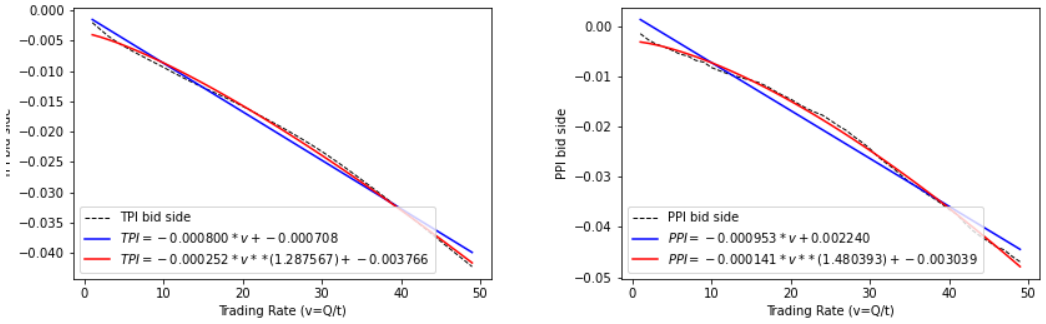
A first aspect to consider in the calibration is to capture, from the LOB, the total available volume at each moment that could be liquidated, i.e., for each moment add up the available volumes (number of shares per tick) of the 100 ticks. This information is very important in the calibration of the functional form of both the TPI and the PPI, because if the amount of shares to be liquidated over a time lapse τ is small compared to the available volumes in the LOB, the transaction would barely walk the book, namely the impact will be underestimated. In this case, the calibration generates a power functional form (x^α with $\alpha > 1$), we call this scenario underestimation. In the opposite case, if the amount of shares to be liquidated is higher than available volumes, the impact will be overestimated, since transactions have to walk the book. This leads to a power functional form (x^α with $0 < \alpha < 1$), we call this scenario overestimation. Moreover, in the overestimated scenario, after a certain volume the full depth of the LOB will be reached and therefore the impact will not grow more. Finally, for the estimation we use liquidation rates, that is, we see the effect on $\nu = Q/\tau$ since liquidation volumes are considered in a time lapse τ. Thus, for the data in question, we define three scenarios: an underestimation scenario with $\nu_{\max} = 50$, an overestimation scenario with $\nu_{\max} = 7000$, and an intermediate scenario using $\nu_{\max} = 1200$, positioned between the former two. This intermediate scenario leads to a calibration of the functional forms of TPI and PPI very close to linear function (x^α with $\alpha = 1$), we call this scenario average-estimation.

The underestimated scenario is shown in figure 2.2, where subfigure (2.2(a)) shows the plot of trading rates ν_i vs TPI_i^{bid}, for volumes $1 \leq i \leq m$, as well as the plots of the linear and power calibrated models $TPI^{bid}(\nu) = f(\nu) = a_1\nu + a_2$ and $TPI^{bid}(\nu) = f(\nu) = r_1\nu^{r_2} + r_3$. Subfigure (2.2(b)) shows plots of trading rates ν_i vs. PPI_i^{bid} and the calibrated PPI models $PPI^{bid}(\nu) = g(\nu) = b_1\nu + b_2$ and $PPI^{bid}(\nu) = g(\nu) = p_1\nu^{p_2} + p_3$. We can see that the shape of the TPI and PPI functions can be calibrated more closely through a power function than through a linear function. However the above consideration, we resolve to find optimal policies in both scenarios, power (x^α with $\alpha > 0$ and $\alpha \neq 1$) and linear (x^α with $\alpha = 1$) impact forms.

Similarly, in figure 2.3, we have the overestimated scenario and specially note the changes in the concavity of the calibrations, i.e. from concave in underestimation to convex in overestimation.

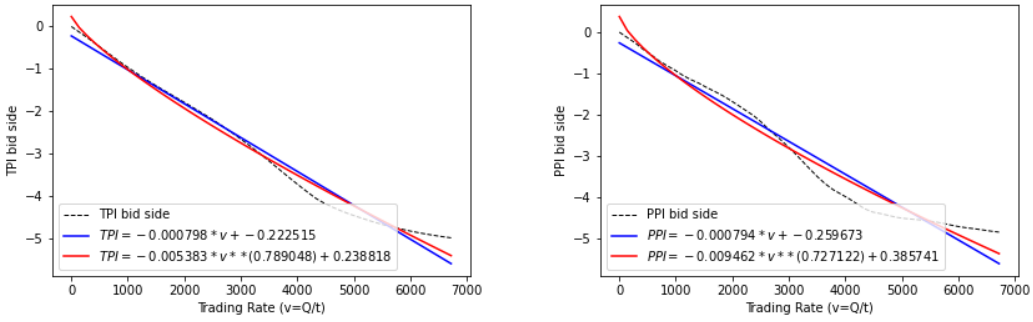
Additionally, the average-estimation scenario is in figure 2.4.

[4] BNB is the official cryptocurrency of the Binance exchange and powers the BNB Chain ecosystem



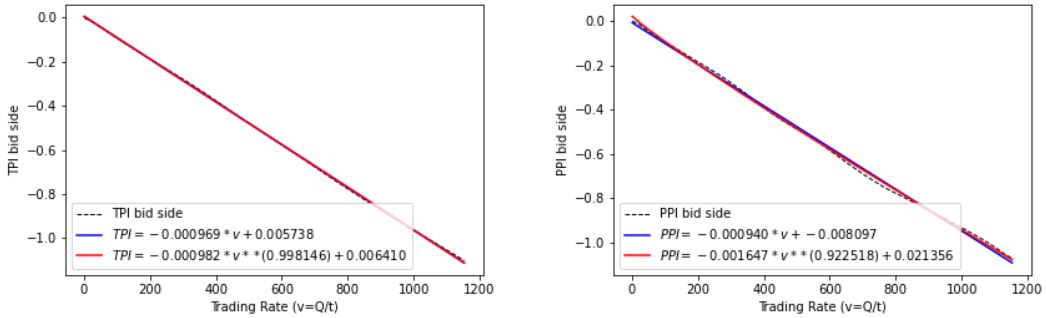
(a) TPI (b) PPI

Figure 2.2: Underestimation of power and linear functional forms of TPI and PPI with $\nu_{max} = 50$, $m = 50$



(a) TPI (b) PPI

Figure 2.3: Overestimation of power and linear functional forms of TPI and PPI with $\nu_{max} = 7000$,$m = 50$



(a) TPI (b) PPI

Figure 2.4: Average-estimation of power and linear functional forms of TPI and PPI with $\nu_{max} = 1200$,$m = 50$

As in Sánchez and Ramírez (2024), data from the BNB order book with a depth of 30 ticks and a frequency of 5 minutes from June 2, 2024, were used to calibrate the functional forms of TPI and PPI and observe if the scenarios remain consistent across different frequencies, depths, and days. Figures 2.5, 2.6, 2.7, 2.8, 2.9, and 2.10 show that despite changes in the LOB depth, frequency, and data from another day, the scenarios remain consistent. Specifically, in the underestimation scenario, we obtain concave functional forms, in the overestimation scenario, convex functional forms, and in the average estimation scenario, nearly linear forms. These findings demonstrate the robustness of approaching the study through scenarios since they remain invariant to the considered variables.

The most significant limitation of this perspective is identifying the size of the inventories that allow for a scenario change in relation to the order book availability, information that is also typically difficult to obtain.

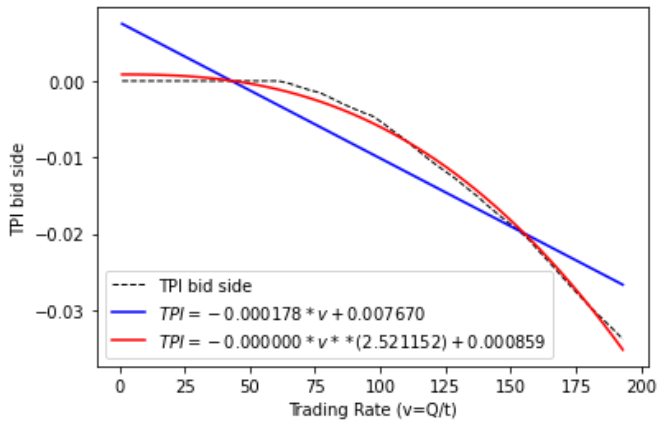


Figure 2.5: TPI bid side underestimation scenario.

In order to determine whether linear models can replace power specifications of TPI and PPI, we conducted the Harvey-Collier Harvey and Collier (1977) and

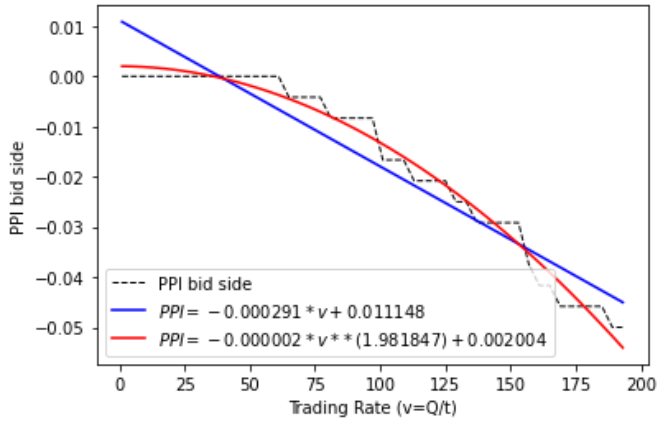


Figure 2.6: PPI bid side underestimation scenario.

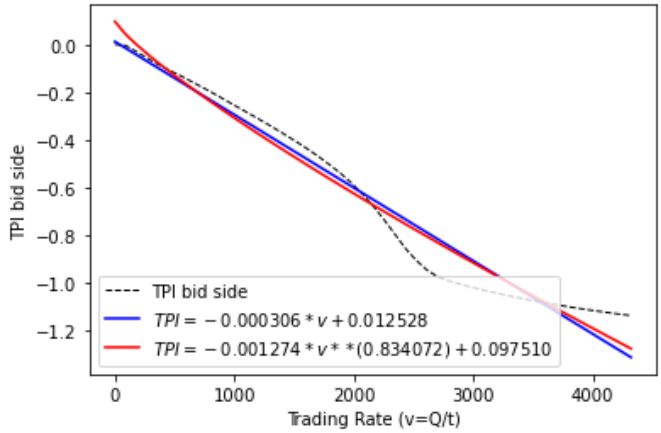


Figure 2.7: TPI bid side overestimation scenario.

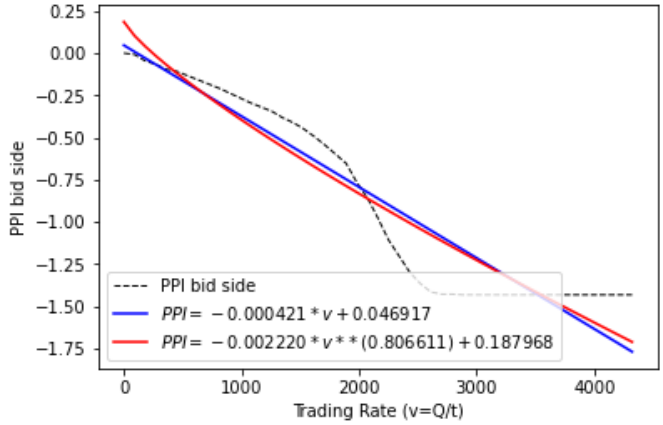


Figure 2.8: PPI bid side overestimation scenario.

Ramsey RESET Ramsey (1969) linearity tests. The results of the Harvey-Collier test, in Table 2.1, consistently show statistically significant p-values well below

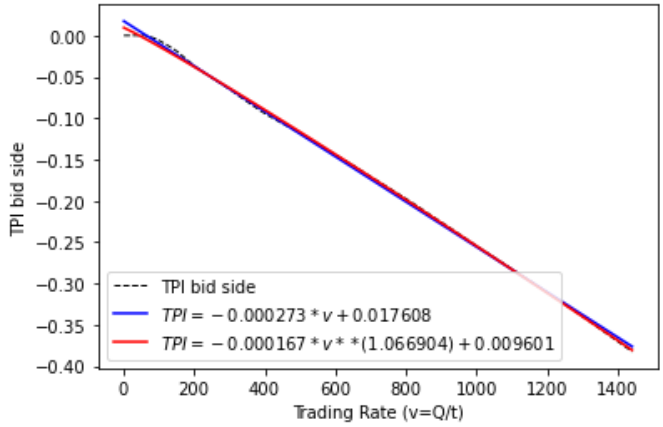


Figure 2.9: TPI bid side average estimation scenario.

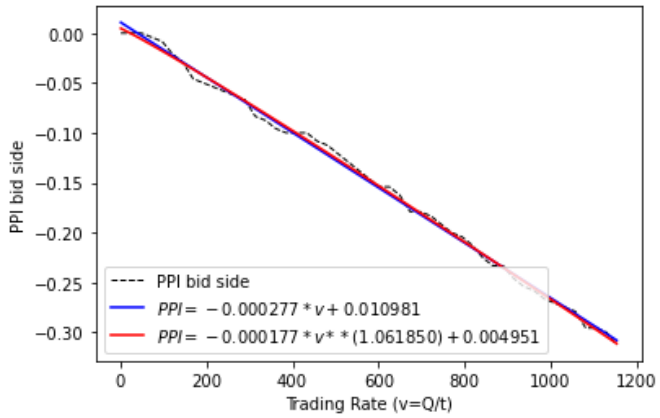


Figure 2.10: PPI bid side average estimation scenario.

the conventional significance level of 0.05. These findings indicate a violation of linearity in the relationship between the trading rate ν and TPI or PPI in the regression models. In other words, the data suggests that a linear function may not accurately represent the relationship between ν and TPI or PPI. Additionally, the results of the Ramsey RESET test, also in Table 2.1, demonstrate extremely small p-values (close to zero). This provides strong evidence against the null hypothesis of linearity.

Consequently, we can conclude that there is substantial evidence supporting the presence of non-linearity in the relationship between ν and TPI or PPI. These results imply that a linear model may not adequately capture the true nature of the relationship between ν and TPI or PPI in underestimation and overestimation scenarios.

To measure which of the two functional forms, linear or power, fits best the average TPI or PPI on the different scenarios, we also use the squared residuals, defined as $\hat{\epsilon}^2_{i,t} = (y_i - \hat{y}_i)^2_t$ and the corresponding R-squared, where y_i is TPI^{bid}_i or PPI^{bid}_i and $\hat{y}_i$ is $TPI^{bid}(\nu_i)$ or $PPI^{bid}(\nu_i)$ (respectively). In appendix .1, we

Table 2.1: p-values of linearity tests

Scenario	Harvey-Collier test	Ramsey RESET test
TPI Underestimation	7.085738088374458e-14	0.0
TPI Overestimation	6.006211572347987e-22	3.041347079865239e-290
PPI Underestimation	1.1678166840620367e-12	0.0
PPI Overestimation	1.2348654834514251e-23	3.718888799866515e-225

show these results and we observe that, in all cases, both the total sum of the residuals and the average of the residuals are lower for the power functional form than for the linear form. Likewise, the R^2 and the adjusted- R^2 is higher in all cases for the power functional form than for the linear form. This implies a better fit in all cases for the power functional form, even in the average scenario where the differences are minimal. Finally, we also obtain, for a time lapse of 5 seconds, an average spread of $\Delta = 0.100069$ and an estimated constant volatility for a 5 seconds timestep of $\sigma = 0.009388$.

2.5.2 Optimal policies

This section refers to the algorithm optimizing the trading rate, that is the numerics for solving the $\mathcal{HJB}$ PDE and using the optimal policies in (2.14) and (2.16).

Error and Convergence analysis

To verify our numerical algorithm, we perform some tests, first we calculate the absolute error (when feasible) comparing the analytical [5] and numerical value function H at $t = 0$, that is for linear and quadratic PPI, obtaining an accuracy of 10^{-5} and 10^{-4} respectively.

Additionally, we were able to confirm the convergence rate of the numerical method for each of the independent variables of the value function H, i.e. inventory q, price s and time t, finding that the convergence rate is 1 for q as expected. As for the price s, convergence rate is very close to the theoretical value 2 for all points of the grid, likewise, the convergence rate for time t is around its theoretical value 1.

Finally, to verify that there are no significant differences between the theoretical vs numerical optimal policies ν, we applied the Chi-square goodness of fit test between numerical and analytical policies for the linear and quadratic PPI. We verify there are no significant differences between these policies, i.e. in essence the policy found through the numerical method is the same as the one found analytically. After performing these tests, we are confident that our algorithm is correct and is working properly.

[5] In the first chapter of this dissertation, i.e. Sánchez and Ramírez (2024) we present analytical optimal solutions for H and ν.

Analysis of optimal policies

Next, we show the optimal policies obtained using the numerical approach in section 2.3, and we explore the results of using the calibrated linear and power TPI ($f(\nu)$) and PPI ($g(\nu)$) for the scenarios under, over and average estimation.

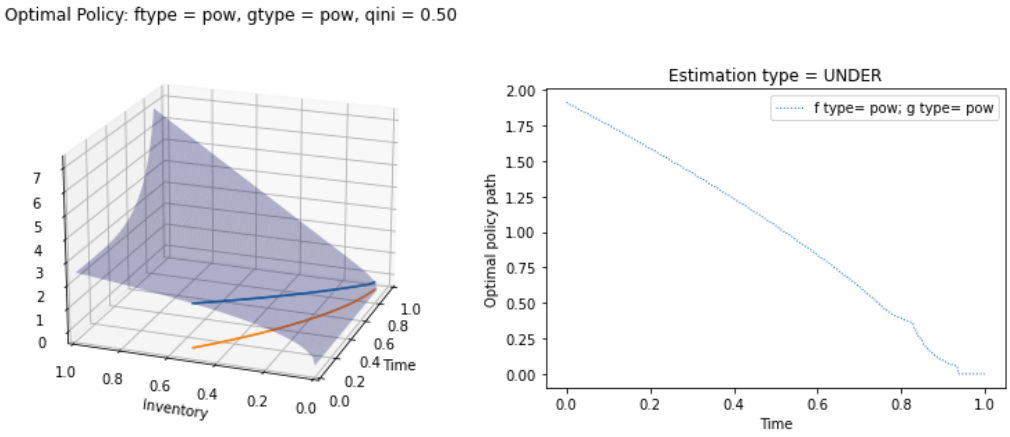
Figure (2.11(a)) presents the optimal policy's plot, featuring a policy path (in blue) and an optimal inventory path (in orange) starting from an inventory level of 0.5, in the context of power TPI and PPI within the underestimation scenario. To construct the path following the optimal policy and the optimal inventory, starting at q_0 inventory, with the corresponding optimal trading rate $\nu(0, q_0)$, we find $q_1 = q_0 - \nu(0, q_0)\Delta t$. If q_1 lies on the grid, the procedure continues, otherwise the respective value of $\nu(1, q_1)$ is interpolated. We repeat the procedure using $q_k = q_{k-1} - \nu(k-1, q_{k-1})\Delta t$ from $t = 0$ to $t = T = 1$. Figure (2.11(b)) shows the corresponding 2D graph of the policy path. Note that the figure (2.11(a)) is a plot of $\nu(t, q)$, i.e. not depending on S, and that is because optimal policies are independent of the price.

By changing the starting point of the inventory to 0.2, figures (2.12(a)) and (2.12(b)) show the 3D optimal policy, the optimal policy path, the optimal inventory path, and the corresponding 2D optimal policy path, respectively.

The graphs shown in Figures 2.11(b) and 2.12(a) display the paths of the optimal policy, constructed from the corresponding policies. In reality, what might appear as an abrupt drop is the natural construction of the path. This path emerges naturally by following the optimal policy (which is a surface in $\mathbb{R}^3$ for a constant price). The paths indicate that at a specific point, the agent has liquidated all inventory, and thus from this moment, the path of the optimal policy is zero. This is not a drop but rather the optimal trading speed, which is zero when the inventory is also zero.

Comparing the two graphs, it can be observed that when constructing these paths, if the inventory starts at 0.5, the agent finishes liquidating very close to the predetermined time horizon. However, when the construction of the optimal policy path starts with an inventory of 0.2, the agent finishes liquidating well before the predetermined time horizon. This behavior will be described later in terms of the concavity of the available optimal inventory, through which we can anticipate that an optimal policy that liquidates more quickly at the beginning is less profitable than those that start liquidating more slowly. In essence, between these two policies, we would achieve higher revenue by liquidating the same amount of stock with the policy that starts more slowly.

Since $Q_{max} = 1$, to identify a path from the numerical optimal policy found, it is necessary to define a q_0 such that $0 < q_0 < 1$, which not only allows us to identify that optimal policy path but also the optimal inventory path. Thus, we determine the optimal inventory path associated with a specific Υ as $q_0^* = \frac{\Upsilon}{q_0} q_0, q_1^* = \frac{\Upsilon}{q_0} q_1, ..., q_n^* = \frac{\Upsilon}{q_0} q_n$. We observe some changes in the liquidation policy, since starting with an inventory of 0.5 implies liquidation at rates close to 0 after approximately time $t = 0.9$, while starting with an inventory of 0.2, these low rates are reached around $t = 0.6$. One possible interpretation of this situation is that,



(a) 3D optimal policy, optimal policy path and optimal inventory path

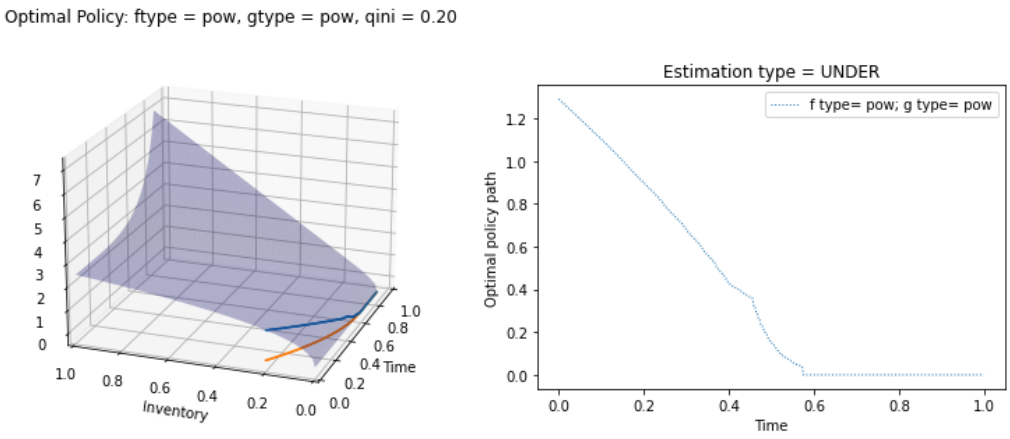
(b) 2D optimal policy path

Figure 2.11: Optimal policy for constant S along with a path for the policy and inventory when it begin at 0.5 with power TPI and PPI in underestimation scenario

in an underestimation scenario, a small impact causes the agent to liquidate early without much impact and avoid the risk of exposure. We opt to show only these graphs because we think they are interesting since power TPI and PPI have the best fit. However, as the functional forms of TPI and PPI change, so does the policy.

Using parameters $T = 1$, $S_{max} = 300$, $Q_{max} = 1$, $N_t = 360$, $N_s = 10$, $N_q = 100$, $q_0 = 0.5$, the calibrated coefficients showed in appendix .2, and using the mid point price $S = 150$ as a constant, we plot the optimal policy paths. The confidence intervals in Appendix .2 demonstrate that no coefficient is significantly equal to zero, thereby confirming the correctness of the functional form specifications. In figure 2.13, we observe that if PPI is linear, the policy path suggests to liquidate more slowly near initial time, even zero, to increase the liquidation rate to a maximum near time T, in which the inventory is zero, then the trading rate drop to zero. However, when PPI is power, the policy starts liquidating quickly and decreasing the trading rate over time, except towards the end of the liquidation period, because having positive inventory causes the trading rate to increase slightly until all inventory is liquidated. This may be caused by the fear of the heavy penalty imposed in the boundary to avoid positive inventory at final time T.

We notice a huge difference between the policies, for different TPI and PPI functions, and notably when g changes from linear to power, because the policy changes from increasing to decreasing. This may be caused by the concavity induced in the underestimation scenario, since not walking the book implies a faster execution and as there are less shares to liquidate, the trading rate decreases. However, this only happens when PPI is power, the opposite happens when PPI is linear, which suggests that in this scenario linear PPI has a greater impact on the optimal policy, than power PPI.



(a) 3D optimal policy, optimal policy path and optimal inventory path

(b) 2D optimal policy path

Figure 2.12: Optimal policy for constant S along with a path for the policy and inventory when it begin at 0.2 with power TPI and PPI in underestimation scenario

Next we find the optimal policy paths for over and average-estimation of linear and power TPI and PPI, these values represent the amount of shares to be liquidated per period of time. Figures 2.14 and 2.15 show a comparison of the optimal policy paths for different TPI and PPI. Interestingly, in the average-estimation scenario we observe, that the optimal policy paths are similar if PPI is power, regardless of the functional form of TPI, which begins by not liquidating and then increases until a certain point at which it drops to zero. For policies where PPI is linear, we observe that these start with a high liquidation rate, decreasing over time and quickly reach a zero trading rate. This is the opposite of the situation observed in the underestimation scenario and suggests that in the average-estimation scenario, power PPI has a greater impact than linear PPI.

Similar situation is presented in the overestimation scenario, in figure 2.15, only that in such scenario we observe that policy paths that liquidate early (Linear PPI) do so more quickly than those liquidating late (power PPI). Moreover, by having even a small inventory, the penalty imposed towards the end of the time period causes a very large liquidation rate, which happens when PPI is power.

According to the above, the optimal policy obtained is highly sensitive to the model calibration process and therefore the optimal policy depends on the calibration scenario, the functional form of both TPI and PPI and the LOB data, since model coefficients depend entirely on this information.

2.5.3 Simulation

This section presents simulations, in the form of backtesting using actual data from Binance, of the liquidation of Υ shares of BNB using the optimal policies calculated from the previous section. The simulation consists of taking an initial inventory

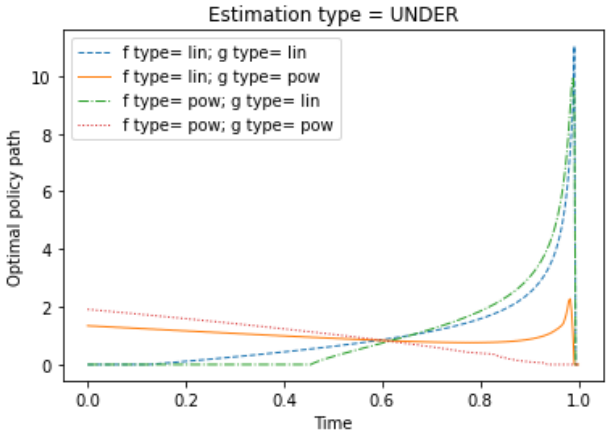


Figure 2.13: Graphical comparison of optimal policies from underestimated TPI and PPI, with $q_0 = 0.5$

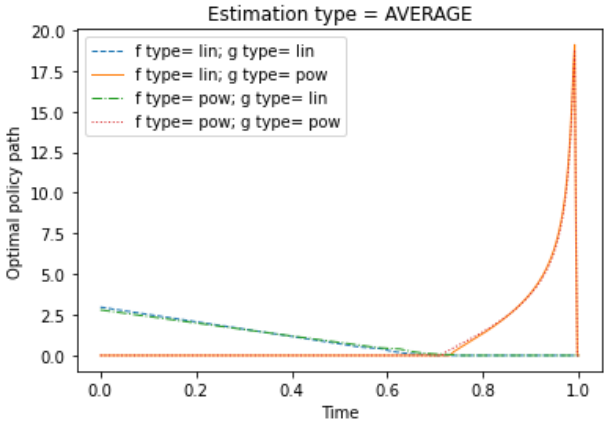


Figure 2.14: Graphical comparison of optimal policies from average-estimated TPI and PPI, with $q_0 = 0.5$

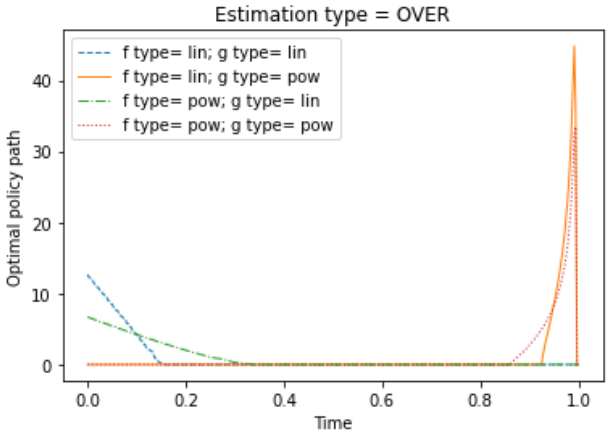


Figure 2.15: Graphical comparison of optimal policies from overestimated TPI and PPI, with $q_0 = 0.5$

Υ and liquidating that inventory over time $[0, T]$ walking the book when needed, following each of the policy paths and then compare the behavior of the numerical optimal strategies versus the naive strategy (NS), which consists of liquidating all cryptocurrencies in a single prespecified time and a commun strategy (CS), which consists of liquidating all the inventory in a constant and proportional manner within the time established for this purpose, strategy that was found in Cartea et al. (2015); Sánchez and Ramírez (2024).

In these simulations we assume, in the sense of Cartea and Jaimungal (2016b), that the order book is resilient, i.e. when executing a MO, at the next moment in time the book has again cleared that order and therefore the order book is taken as it has been downloaded, without any modification.

To refer to the numerical optimal strategies we use the following notation: the first letter refers to the estimation scenario, i.e. under (U), over(O) and average(A), the second letter refers to TPI and therefore will always be T, the third letter refers to the functional form of TPI, i.e. linear (L) or power (P), the fourth letter will always be P in reference to PPI and the fifth and last letter refers to the functional type of PPI, i.e. L or P. For example, the optimal policy obtained in the underestimation scenario with linear TPI and power PPI is represented by the letters UTLPP. With these references, we use the policy paths starting with inventory 0.5 for UTLPL, UTLPP, UTPPL, UTPPP, OTLPL, ATLPL, ATLPP, ATPPL, ATPPP, OTLPP, OTPPL and OTPPP. The importance in reviewing all these scenarios is to test the different results that occur under different assumptions, such as book size and functional form of the impact.

With respect to the policies obtained numerically we show in tables (2.2) and (2.3) the cumulative revenue obtained by simulating the execution of each strategy as a percentage of NS and CS respectively, when we liquidate $\Upsilon = 4000$ BNB cryptocurrencies in 1800 seconds at intervals of 5 seconds.

The 95% confidence intervals provided around the percentage values for each strategy compared to NS and CS allow us to evaluate the statistical significance of the observed differences. In the comparison with NS, all strategies show higher percentages, with confidence intervals that do not include the value of 1. This indicates that each strategy is significantly superior to the naive strategy (NS), demonstrating that these strategies generate greater accumulated revenues. In particular, the strategies ATLPP and ATPPP stand out with the highest percentages, confirming their superior performance over NS. The confidence intervals were constructed from the simulation of 1000 order books with a depth of 100 ticks, liquidating 4000 BNB in 1800 seconds.

When comparing the strategies to CS, the situation is more nuanced. The strategies ATLPP and ATPPP again show the highest percentages, with confidence intervals excluding the value of 1, indicating they are significantly superior to the common strategy (CS) as well. However, the strategies OTLPL and OTPPL present the lowest percentages, with confidence intervals that also do not include the value of 1, suggesting they are significantly inferior to CS. This distinction highlights the relative effectiveness of the strategies: while all strategies outperform NS, only some strategies outperform CS, and others, like OTLPL and OTPPL,

Table 2.2: Cumulative revenue as a percentage of NS

Strategy	Percentage	Confidence Interval (95%)
Under Linear TPI Power PPI (UTLPL)	1.010922	(1.010813, 1.011031)
Under Linear TPI Linear PPI (UTLPP)	1.009504	(1.009400, 1.009608)
Under Power TPI Linear PPI (UTPPL)	1.011207	(1.011094, 1.011320)
Under Power TPI Power PPI (UTPPP)	1.008875	(1.008767, 1.008983)
Average Linear TPI Power PPI (ATLPL)	1.008409	(1.008296, 1.008522)
Average Linear TPI Linear PPI (ATLPP)	1.011459	(1.011340, 1.011578)
Average Power TPI Linear PPI (ATPPL)	1.008456	(1.008344, 1.008568)
Average Power TPI Power PPI (ATPPP)	1.011478	(1.011359, 1.011597)
Over Linear TPI Power PPI (OTLPL)	1.006268	(1.006147, 1.006389)
Over Linear TPI Linear PPI (OTLPP)	1.010701	(1.010580, 1.010822)
Over Power TPI Linear PPI (OTPPL)	1.007157	(1.007040, 1.007274)
Over Power TPI Power PPI (OTPPP)	1.011091	(1.010969, 1.011213)
Naive Strategy (NS)	1.000000	(1.000000, 1.000000)
Common Strategy (CS)	1.009673	(1.009575, 1.009771)

Table 2.3: Percentage of each strategy compared to CS

Strategy	Percentage	Confidence Interval (95%)
Under Linear TPI Power PPI (UTLPL)	1.001238	(1.001192, 1.001284)
Under Linear TPI Linear PPI (UTLPP)	0.999833	(0.999796, 0.999870)
Under Power TPI Linear PPI (UTPPL)	1.001519	(1.001464, 1.001574)
Under Power TPI Power PPI (UTPPP)	0.999210	(0.999158, 0.999262)
Average Linear TPI Power PPI (ATLPL)	0.998748	(0.998684, 0.998812)
Average Linear TPI Linear PPI (ATLPP)	1.001769	(1.001698, 1.001840)
Average Power TPI Linear PPI (ATPPL)	0.998795	(0.998733, 0.998857)
Average Power TPI Power PPI (ATPPP)	1.001787	(1.001717, 1.001857)
Over Linear TPI Power PPI (OTLPL)	0.996628	(0.996538, 0.996718)
Over Linear TPI Linear PPI (OTLPP)	1.001019	(1.000933, 1.001105)
Over Power TPI Linear PPI (OTPPL)	0.997508	(0.997428, 0.997588)
Over Power TPI Power PPI (OTPPP)	1.001405	(1.001323, 1.001487)
Naive Strategy (NS)	0.990420	(0.990322, 0.990518)
Common Strategy (CS)	1.000000	(1.000000, 1.000000)

underperform relative to CS.

These results indicate that the observed differences are not only statistically significant but also economically relevant. Strategies with higher percentages generate greater accumulated revenues, and those with lower percentages generate lower accumulated revenues compared to the reference strategies. The robustness of these results is confirmed by the consistency of the confidence intervals, providing greater confidence in the relative effectiveness of each strategy. This detailed analysis allows for a better understanding of the economic implications of choosing one strategy over another, confirming the superior performance of certain strategies and highlighting the shortcomings of others.

For practical reasons, next we focus our attention on the top two policies ATPPP and ATLPP, and the lowest two policies OTLPL, OTPPL.

From the results in table 2.2, we observe that interestingly the increasing policy paths are the ones that earn the highest revenue. The policy path that generated the highest revenue was ATPPP with a cumulative revenue of 1.1478% above NS and 0.1787% above CS and in second place was ATLPP with a revenue of 1.1459% above NS and 0.1769% above CS. This, although surprising, may be justified by the fact that the number of shares to liquidate gives an average estimation when compared to the LOB setting in the data. In average-estimation scenario, these policy paths are the ones that start liquidating slowly, even zero, increasing the trading rate until a maximum point and when the inventory is zero, dropping to zero. Similarly, in the overestimation and underestimation scenarios, the policy paths that obtained the highest cumulative revenue were those that started with a low trading rate.

Likewise, the policy paths that obtained the lowest cumulative revenue were those that liquidated faster at the beginning. Such is the case of OTLPL, which was the one that obtained the lowest revenue with a 0.6268% above NS and -0.003372% below CS, and OTPPL, whose cumulative revenue was 0.7157% above NS and -0.002492% below CS. Something interesting that we can observe is that the CS strategy was better in some cases than the numerical strategies, a characterization of which we will present later on. To confirm that the behavior is not subject to the initial inventory, we test the solution with different initial inventories Υ and reaffirm that increasing policies generate more revenue.

This situation can be described in terms of optimal inventory. If an optimal policy path starts by liquidating slowly and then increase the liquidation rate, the optimal inventory path will decrease slowly at first and then decrease more rapidly, i.e. it is a concave function. In the opposite case, when the policy path liquidates quickly at the beginning and then decreases the liquidation rate, the inventory path will decrease quickly at the beginning and then decrease more slowly, i.e. a convex function. Then, the optimal inventory path is a non-increasing function, which when concave describes an optimal policy path that, as shown before, generates higher cumulative revenue than convex optimal inventory functions, see figure (2.16).

According to this characterization, since the optimal inventory path associated with the CS strategy is a straight line, its performance is better than strategies

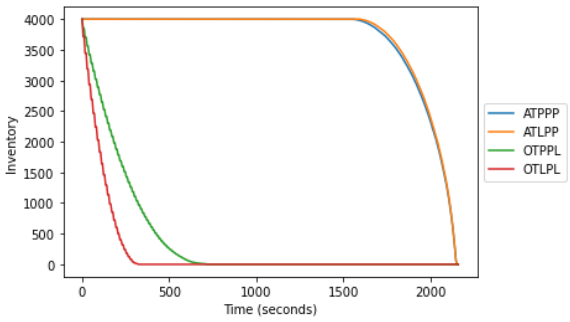


Figure 2.16: Optimal inventory paths for which, when concave, such as ATPPP and ATLPP, generate higher cumulative revenues than convex paths, such as OTPPL and OTLPL.

whose optimal inventory paths are convex and worse than strategies whose optimal inventory paths are concave.

Next we present a further exploration of the empirical hypothesis that concave optimal inventory paths are the ones generating higher cumulative revenue and that, on the contrary, policy paths that start with a high trading rate and decrease as time passes (corresponding to convex inventory paths) are the ones generating less cumulative revenue. This empirical hypothesis, together with the theoretical results found in Sánchez and Ramírez (2024), lead us to consider that the optimal inventory trajectories can be modeled by a non-increasing monotonic function

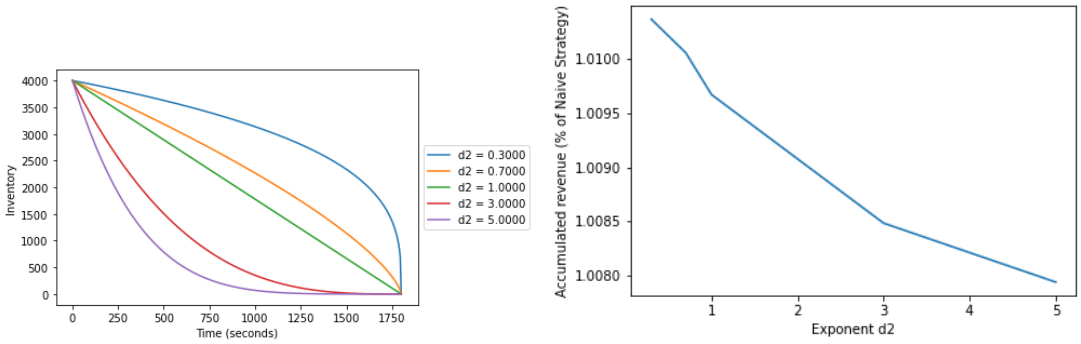
$$q: \quad \{t \in \mathbb{Z} : 0 \leq t \leq T\} \quad \rightarrow \quad [0, \Upsilon] \tag{2.17}$$

subject to $q(0) = \Upsilon$ and $q(T) = 0$, then, a possible function is

$$q(t) = \Upsilon \left(1 - \frac{t}{T}\right)^{d_2} \tag{2.18}$$

For the parameters $\Upsilon = 4000$, $T = 360$, which represents a time horizon of 1800 seconds, and taking the exponent d_2 in $[0.3, 0.7, 1, 3, 5]$ to obtain changes in the concavity of the function $q(t)$, we build simulations of optimal inventory paths, see figure (2.17(a)). With each of these inventory paths, we simulate the execution of the corresponding optimal policy path, thus obtaining its cumulative revenue as a percentage of NS, characterized by d_2, see figure (2.17(b)).

We observe that when the exponent d_2 is less than 1, i.e. $q(t)$ is concave, the cumulative revenues as a percentage of NS obtained are higher than the other strategies. It seems that by increasing the value of the exponent d_2, the cumulative revenue decreases. However, if the value of the exponent d_2 is very close to zero or very large, these inventory paths begin to approximate those of the NS strategy and therefore the corresponding cumulative revenue obtained decrease. Figure (2.18) shows the cumulative revenue as a percentage of NS obtained by simulating optimal inventories paths when the exponent d_2 varies from 0.001 to 5 with increments of 0.01. We can observe that for values of d_2 close to zero, the cumulative revenue is close to that obtained by NS and that for large values of the exponent d_2 the



(a) Simulated optimal inventory paths from equation $q(t) = 4000\left(1 - \frac{t}{360}\right)^{d_2}$

(b) Accumulated revenue from simulated optimal inventory paths

Figure 2.17: Simulation of optimal inventory paths and corresponding cumulative revenue as a percentage of NS.

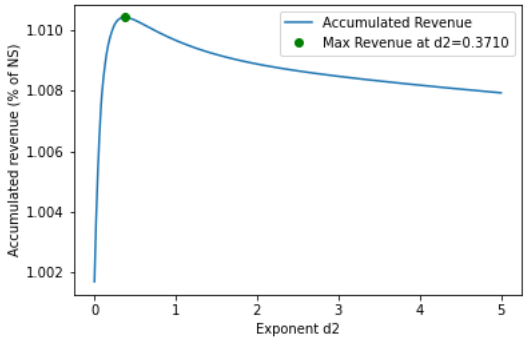
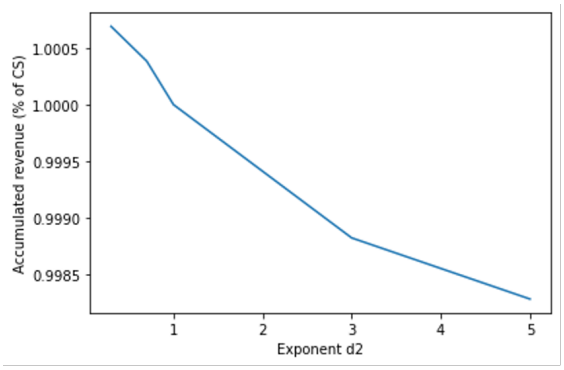


Figure 2.18: Accumulated revenue as a percentage of NS from simulated optimal inventories paths for d_2 from 0.001 to 5 with steps of 0.01, liquidating $\Upsilon = 4000$.

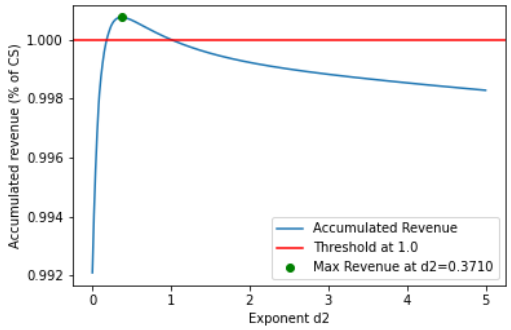
cumulative revenue decreases. This situation leads us to conclude that in an interval between 0 and 1 for the exponent d_2, which in this case is 0.371 the concave optimal inventories paths generate a higher cumulative revenue than convex optimal inventories paths.

A similar situation is observed when we look at the behavior of the simulated policies with respect to the CS strategy, as shown in Figure 2.19. The big difference lies in the fact that the CS strategy is better than many of the simulated strategies, particularly those simulated from convex optimal inventories and some strategies with concave optimal inventories, especially those that are close to the NS strategy, but it is highlighted that many strategies with concave optimal inventories are better than the CS strategy and that the behavior described above remains the same. We also highlight that the value of the exponent d_2 that maximizes the accumulated income as a percentage of CS is also 0.371.

In terms of optimal policies, those that start with a slow trading rate and increase



(a) Accumulated revenue from simulated optimal inventory paths



(b) Accumulated revenue as a percentage of CS from simulated optimal inventories paths for d_2 from 0.001 to 5 with steps of 0.01, liquidating $\Upsilon = 4000$.

Figure 2.19: Simulation of optimal inventory paths and corresponding accumulated revenue as a percentage of CS.

over time, as opposed to those that start liquidating more quickly and decrease over time, generate higher cumulative revenue. Likewise, we emphasize that this behavior is only observed if such policy paths do not reach the extreme of resembling the NS, since in that case their cumulative revenue decreases.

2.6 Discussion and conclusions

In this work we study the problem of optimal liquidation by analyzing the consequences of having different functional forms, specifically linear and power, for temporary and permanent price impact. We apply this theory in the cryptocurrency market given its importance due to its evident growth, high volatility and its articulation in the world of Decentralized Finance (DeFi) and secondly due to the ease of accessing historical and LOB information. Therefore, we use LOB data for BNB cryptocurrency, information downloaded from the Binance API.

By using different sizes of inventories to liquidate in the estimation approach over the LOB data, lead us to three scenarios, Average estimation, Underestimation and Overestimation, each one configuring different TPI and PPI functions, i.e. convex/concave. This functions are represented in the form of power functions x^α with $\alpha > 0$. And we confirm by the Harvey-Collier and Ramsey RESET tests that this functions are actually significantly different and lead to different optimal policies, more specifically different from the classical linear approach.

We observe that the shape of the optimal policy paths can change depending on the starting inventory, the functional forms of TPI and PPI, and the values of the calibrated parameters. However, the functional description of the policy paths does not seem obvious we carry out this characterization through the concavity of the optimal inventory paths. Leading us to acknowledge the importance of the ratio between the initial amount of shares to be liquidated and the availability of shares in the LOB when characterizing the performance of optimal policies, and that these do not depend strongly on the functional form of TPI and PPI.

In this problem, where an amount Υ of cryptocurrencies is to be liquidated in a pre-established period of time T, the optimal inventory path can be described through a non-increasing monotonic function $q(t) = \Upsilon\left(1 - \frac{t}{T}\right)^{d_2}$, which describes a family of optimal policy paths. Under these conditions, we observe that concave optimal inventory paths, i.e., $0 < d_2 < 1$, are the ones that generate the greatest cumulative revenue, under the condition that the concavity is not so dramatic, this shows that even slight deviations from a linear optimal policy can result in significant changes, leading to increased earnings. Also, this leads us to characterize the optimal winning policy paths are those whose trading rate is low at the beginning and increases over time.

Further research tends to examine this phenomenon in greater detail, identifying, for example, the exact conditions and points at which the described behavior changes, or, based on families of optimal policies, identifying the "most" optimal among them, the functional forms of these families of optimal policies, among other

options. In this context, an understanding of singular control theory could provide valuable insights. Singular control, a specialized area within optimal control theory, offers unique perspectives on how systems behave under discontinuous or singular control actions.

Singular control is typically represented in a system described by a differential equation:

$$dx(t) = f(x(t), u(t), t)dt, \tag{2.19}$$

where $x(t)$ represents the state of the system at time t, and $u(t)$ is the control variable, with f characterizing the system's dynamics.

In singular control, the function $u(t)$ often includes impulses or switches at specific instances, leading to a state trajectory $x(t)$ that may exhibit sudden changes or discontinuities.

The objective in singular control problems is to optimize a control $u(t)$ to achieve goals such as minimizing a cost function:

$$\min_{u(t)} \int_0^T g(x(t), u(t), t)dt, \tag{2.20}$$

subject to the dynamics of the system and constraints on $x(t)$ and $u(t)$.

Singular control is particularly relevant in financial applications, including the optimal liquidation of assets, where decisions can significantly and instantaneously impact the system's state, such as market prices.

The field of singular control theory spans various applications, as demonstrated in studies by Reppen (2018); Agram et al. (2019); Ata et al. (2023); Archankul et al. (2023); Budhiraja and Ross (2006); Huang et al. (2012).

Reppen (2018) explores the theoretical aspects of singular control in financial economics. Agram et al. (2019) investigates singular control in managing SPDEs with space-mean dynamics, highlighting its role in complex system modeling. Ata et al. (2023) focuses on computational methods in singular control of reflected Brownian motion, relevant for systems engineering. Archankul et al. (2023) examines singular control in cash management under ambiguity, demonstrating its utility in decision-making. Budhiraja and Ross (2006) presents foundational work on the existence of optimal controls in constrained singular control problems. Additionally, Huang et al. (2012) discusses iterative methods for singular control in GMWB pricing, showing the theory's adaptability in financial product analysis.

These studies highlight the breadth of singular control theory across various domains, offering valuable insights for financial applications such as optimizing asset liquidation amidst market complexities. The extension of this thesis to include such applications is reserved for future research endeavors.

The scope of this study is limited, initially by the LOB data available, by the liquidity of the asset to be liquidated since high liquidity is required so that the inventories can be liquidated, in this case in time horizons of seconds. Other limitations of this study are the model used does not take into consideration other variables that could affect the optimal liquidation policy such as volatility variability, liquidating using limit orders, new market and limit order flows by other

agents, the liquidation dynamics that may imply jumps due to ticks, among others. Therefore, considering several additional variables to the problem, different models, different characteristics of financial assets and markets in which they are traded are part of the future research agenda.

3 A Multifaceted Analysis of Efficiency in Cryptocurrencies Market

Abstract

This descriptive study provides a multifaceted examination of market dynamics under the Weak Efficient Market Hypothesis (WEMH) for various cryptocurrencies. An innovative statistical approach is used to identify epochs of distinct market trends, with the Hurst index calculation indicating a general trend towards inefficiency across various cryptocurrencies. Shannon entropy is employed to quantify predictability and efficiency, with results pointing to greater inefficiency during mean reversion periods. The Ljung-Box and Dickey-Fuller tests reveal no autocorrelations, upholding the WEMH, yet they identify stationarity in trend epochs, hinting at potential inefficiencies. The study also incorporates a joint binomial distribution as a complementary statistical approach to analyze the interplay between persistence and predictability of market behaviors. Emphasizing the importance of data granularity, a frequency dependence analysis is conducted, providing a more intricate understanding of efficiency dynamics in cryptocurrency markets.

Keywords: Cryptocurrency, Market Efficiency, Weak Efficient Market Hypothesis, Trend Analysis, Hurst Index, Shannon Entropy, Autocorrelation, Stationarity, High-Frequency Data

3.1 Introduction

In the landscape of financial economics, the Efficient Market Hypothesis (EMH) postulated by Fama (1970) remains a cornerstone, categorizing markets into three levels of efficiency: weak, semi-strong, and strong. The weak form asserts that all past trading information is already reflected in stock prices, rendering technical analysis futile; the semi-strong form suggests that all publicly available information is accounted for in stock prices, thus discounting the effectiveness of fundamental analysis; and the strong form posits that all information, both public and private, is fully integrated into stock prices, implying that no investor has monopolistic access to relevant information for gaining abnormal returns.

Many studies illuminate the nuances of market efficiency. For instance, Shternshis et al. (2022) examines the Moscow Stock Exchange, revealing a significantly low degree of market efficiency from 2012 to 2021, achieved through the use of innovative entropy-based measures and volatility estimation methods. Duarte and Mascareñas Perez-Iñigo (2013) synthesize over fifteen years of research, uncovering variability in market efficiency dependent on development levels, global regions, study periods, and economic openness. Moreover, Lim and Brooks (2011) provide an extensive review of the literature, highlighting the evolution of the weak-form

market efficiency hypothesis and the increasing recognition of its time-varying characteristics. These studies collectively represent the continuous scholarly pursuit to unravel the complexities of efficient markets.

In the context of emerging markets, particularly those of cryptocurrencies, numerous studies have addressed the topic of market efficiency. Kakinaka and Umeno (2022) explore the asymmetric multifractality and efficiency of major cryptocurrencies during the COVID-19 pandemic and find that the outbreak influenced the efficiency of price behaviors differently across short- and long-term horizons, with a noted increase in asymmetry in market behaviors. Wei (2018) examines the liquidity of 456 cryptocurrencies, showing that return predictability decreases in highly liquid cryptocurrencies and that while Bitcoin is showing signs of efficiency, many cryptocurrencies still display autocorrelation and non-independence.

Furthering this inquiry, Chu et al. (2019) investigate the high-frequency markets of Bitcoin and Ethereum against the Euro and US Dollar, confirming the adaptive market hypothesis (AMH) by demonstrating that market efficiency varies over time. Similarly, Noda (2020) measures the degree of market efficiency over time using a generalized least squares-based time-varying model, finding that Bitcoin generally exhibits a higher level of market efficiency than Ethereum. On the contrary, Nadarajah and Chu (2017) address the efficiency of Bitcoin by applying power transformations to return series, challenging the notion that Bitcoin returns do not satisfy the efficient market hypothesis. Sensoy (2019) compares the weak-form efficiency of Bitcoin against the US Dollar and Euro, finding that higher trading frequencies correspond to lower pricing efficiency, and that liquidity and volatility significantly influence informational efficiency. Lastly, Zhang et al. (2020a) utilize algorithms to identify bull and bear market phases in high-frequency markets for Bitcoin, Ethereum, and Litecoin, observing that market efficiency and liquidity vary significantly between these states.

This descriptive study falls within the scope of the literature examining the efficiency of the cryptocurrency market. Our primary objective is to provide a comprehensive, multifaceted description of market dynamics under the Weak Efficient Market Hypothesis (WEMH) for various cryptocurrencies. We utilize a range of methodologies to describe, rather than definitively verify or reject, the WEMH for different cryptocurrencies across epochs of uptrend, downtrend, or mean reversion. These epochs are determined using a novel methodology based on the observation that, excluding significant price jumps, expected returns in uptrends are positive, in downtrends are negative, and in periods of mean reversion are approximately zero. This approach is supported by the historical performance of financial assets and corroborated by subsequent empirical evidence Evans and Lewis (1992). While our findings offer a detailed descriptive analysis of market efficiency, they represent a preliminary step in understanding the multifaceted nature of cryptocurrency market dynamics.

For example, the empirical evidence from Jegadeesh and Titman (1993) and Mitra and Bawa (2017) suggest that financial assets exhibit momentum and trend persistence, which can be capitalized upon to generate positive returns. Specifically,

Jegadeesh and Titman (1993) demonstrate that a strategy of buying past winners and selling past losers generates significant positive returns over 3- to 12-month holding periods, indicating that, omiting large jumps, upward trending periods are associated with positive returns. Similarly, Mitra and Bawa (2017) find that significant positive returns can be generated in periods of upward trend persistence, as identified by the time-varying Hurst exponent. Conversely, during periods of downward trend persistence, the expected returns are negative, as future index price movement is anticipated to decrease Mitra and Bawa (2017). Furthermore, both papers indirectly mention the concept of mean reversion, where Jegadeesh and Titman (1993) observe that the abnormal returns tend to diminish over time, and Mitra and Bawa (2017) suggest initiating a "Sell" based on the expectation of mean reversion, indicative of returns converging to zero.

In this descriptive study, we first approximate the observed prices to the efficient prices using the methodology of uncertainty zones Robert and Rosenbaum (2012), finding no significant differences between the two price series. Therefore, we utilize the observed prices throughout the paper. Given that this is a descriptive analysis and the market efficiency metrics employed are non-parametric, it is unnecessary to assume a specific data generation process. We then implement the methodology for identifying periods of downward, upward, and mean-reversion trends, observing that the length of the windows is crucial in this determination process. The size of the window significantly influences the frequency of trend changes detected, thereby affecting the epoch assignment. A smaller window size tends to identify frequent shifts between different trends, often resulting in a classification of mean reversion epochs. Conversely, a larger window may fail to adequately model trend or mean reversion epochs because over extended periods, price movements can mask underlying trend behaviors by averaging out significant fluctuations. This can lead to a misrepresentation of market conditions in such long windows. Thus, a balanced window size is essential, one that is sufficiently robust to identify the respective epochs without overly frequent shifts or oversimplification of market dynamics. For this analysis, windows of size 24 data points have been determined as the optimal balance, providing a robust foundation for accurately identifying the trend epochs.

Once the epochs are identified, we find the Hurst index of each of these epochs for each cryptocurrency following the methodology we have dubbed the p-th variation Cont and Das (2024). After conducting a sanity check of the methodology implementation using simulations of Fractional Brownian motions, we find that the prices have Hurst indices H less than 0.5 for all epochs and all cryptocurrencies, indicating a certain tendency towards market inefficiencies.

Subsequently, we employ the concept of Shannon entropy Shannon (1948) to study the predictability of the returns of each epoch for each cryptocurrency. When returns are predictable, these periods signify inefficiency Shternshis et al. (2022). We use the ratio proposed in Shternshis et al. (2022) that it allows us to evaluate the degree of efficiency of the different epochs, finding that, unlike BTCBUSD and ETHBUSD, the rest of the cryptocurrencies studied tend to be more inefficient in all the epochs studied.

Then, we use the Ljung-Box and Dickey-Fuller tests to identify autocorrelations and unit roots, respectively (Zhang et al. (2020a) and Cox et al. (2004)). In the first case, the presence of autocorrelations indicates inefficiency because past information can be used to predict future returns. In the second case, stationarity indicates inefficiency in that constant statistics over time could allow the prediction of future returns. We find no evidence of autocorrelations in returns for any epoch or cryptocurrency, supporting the Weak-Form Efficient Market Hypothesis (WEMH), but we do find evidence of stationarity in returns, especially for the trending epochs.

Moreover, following Lo and MacKinlay (1999)'s assertion on the necessity of empirically robust statistical models for analyzing the stochastic evolution of prices, this study employs a joint binomial distribution to investigate the simultaneous behavior of the Hurst index and Shannon entropy. This approach adheres to the call for rigorous empirical modeling and adds a comprehensive statistical dimension to our analysis of market efficiency. By examining the interplay between these two measures, we aim to provide a more holistic view of market dynamics. Specifically, we explore how the combination of a high Hurst index and low Shannon entropy might suggest a market phase with strong trend persistence but low return predictability, highlighting the nuanced ways markets can exhibit efficiency or inefficiency.

The application of the joint binomial distribution analysis in our study methodically integrates the Hurst and Shannon efficiency indices, elucidating the multifaceted nature of market efficiency across various cryptocurrencies. This descriptive analysis offers a comprehensive perspective that balances the persistence measured by the Hurst index with the unpredictability captured by Shannon entropy. While the Ljung-Box and Augmented Dickey-Fuller tests independently address autocorrelations and unit roots, our joint distribution approach synthesizes these dimensions of market behavior into a cohesive framework. The resulting heatmaps visually articulate the complexity and variability inherent in the cryptocurrency markets.

Finally, we carry out an analysis of the frequency dependence of the previous results, changing the initial frequency of the data from 5 minutes to 10 and 30 minutes, respectively. We find that some metrics, such as the efficiency ratio derived from Shannon entropy, tend to indicate a more efficient market.

This paper is structured as follows: Section 3.2 outlines the models used to assess cryptocurrency market efficiency, incorporating statistical methods for analyzing market dynamics. These include the "Uncertainty Zones" model (Robert and Rosenbaum (2012)) for price data refinement, the trend period identification method for trend detection, the p-th variation method (Cont and Das (2024)) for Hurst exponent estimation, and the Shannon Entropy model (Shternshis et al. (2022)) for quantifying market unpredictability. Additionally, we apply the Ljung-Box test and the Augmented Dickey-Fuller test to probe autocorrelations and price series stationarity, respectively. A novel joint binomial distribution analysis combines Hurst index and Shannon entropy results for a comprehensive market behavior overview. Section 3.3 delineates the empirical findings, highlighting the

differentiated impact of these models on understanding market efficiency, including the discovery of frequency-dependent efficiency and the variability in cryptocurrency behavior under different market conditions. Concluding remarks in Section 3.4 synthesize these insights, reinforcing the complex nature of cryptocurrency market efficiency and proposing avenues for future inquiry. Through a detailed interrogation of market dynamics using a diverse analytical framework, this descriptive paper significantly enhances the understanding of efficiency across cryptocurrency markets.

3.2 Models

In exploring market efficiency within cryptocurrency markets, this paper introduces a series of models tailored to address the unique challenges presented by the high-frequency and discretized nature of financial data. The novelty of our approach lies in adapting and integrating established and emerging statistical techniques to uncover underlying market dynamics and inefficiencies. We carefully select and refine each model to capture specific aspects of market behavior, from price approximation to trend detection and the quantification of efficiency. The "uncertainty zones" methodology, for instance, allows us to refine raw price data into a form more amenable to theoretical analysis, setting the stage for a more nuanced interpretation of price movements. By combining these models, we present a holistic approach to analyzing the WEMH, offering fresh insights into the predictability and efficiency of cryptocurrencies.

The ensuing subsections delineate each model's framework and its specific contribution to our investigation, unveiling the multifaceted nature of market efficiency and the innovative analytical lens through which we examine it.

3.2.1 Uncertainty Zones

Financial data presents challenges due to the discretized nature of observed prices, particularly given the tick sizes in which they are quantized. To address this, we employ the methodology of "uncertainty zones", allowing us to approximate observed prices to theoretical (efficient) prices. This method is based on the work of Robert and Rosenbaum (2012).

We consider $(X_t)_{t\geq 0}$ to represent the efficient asset price on a filtered probability space $\left(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$. It is assumed that the logarithm of the efficient price, $(Y_t)_{t\geq 0}$, is an $\mathcal{F}_t$-adapted continuous semi-martingale. This can be represented as:

$$Y_t = \log X_t = \log X_0 + \int_0^t a_u du + \int_0^t \sigma_{u^-} dW_u \tag{3.1}$$

Here, $(W_t)_{t\geq 0}$ denotes a standard $\mathcal{F}$-Brownian motion, $(a_t)_{t\geq 0}$ is a measurable process with locally bounded sample paths, and $(\sigma_t)_{t\geq 0}$ is a positive $\mathcal{F}$-adapted process with càdlàg sample paths.

A tick grid is defined as $\{k\alpha; k \in \mathbb{N}\}$ where α is the tick price size. The zone U_k encompasses $[0, \infty) \times (d_k, u_k)$ with $d_k = \left(k + \frac{1}{2} - \eta\right)\alpha$ and $u_k = \left(k + \frac{1}{2} + \eta\right)\alpha$,

where $0 < \eta < 1$, see Figure 3.1. The process $(P_t)_{t\geq 0}$ depicts the last traded price, a càdlàg piecewise constant process representing the observed price.

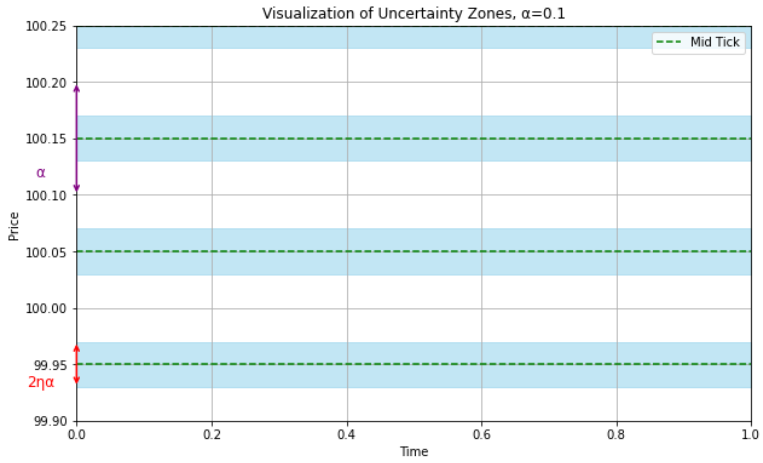


Figure 3.1: Visualization of Uncertainty Zones

To estimate $X^t_{\tau_i}$, $i \geq 0$, the formula used is:

$$\hat{X}^t_{\tau_i} = P_{t_i} - \alpha\left(\frac{1}{2} - \hat{\eta}_{\alpha,t}\right)\text{sign}\left(P_{t_i} - P_{t_{i-1}}\right) \tag{3.2}$$

where the sequence $(\tau_i)_{i\geq 0}$ represents the exit times from the uncertainty zones which will lead to a change in the transaction price. In this model, the parameter η is essential in defining the boundaries of the uncertainty zones, which are crucial for aligning observed prices with efficient prices. The parameter η is estimated using $\hat{\eta}_{\alpha,t}$, which is based on an analysis of the frequency of alternations (reversals in price trend) and continuations (consistent price trends) in price changes. Assuming that the jump sizes are bounded, with m denoting their maximal value, for a given $k = 1, \ldots, m$ and $t > 0$, $\hat{\eta}_{\alpha,t}$ is defined as:

$$\hat{\eta}_{\alpha,t} = \left(0 \vee \sum_{k=1}^{m} \lambda_{\alpha,t,k} u_{\alpha,t,k}\right) \wedge 1 \tag{3.3}$$

The concept of this estimator lies in the assumption that $u_{\alpha,t,k}$ serve as consistent estimator for η for each k. Consequently, $\lambda_{\alpha,t,k}$ act as weighting coefficients. It is important to highlight that the ratio $N^{(c)}_{\alpha,t,1}/N^{(a)}_{\alpha,t,1}$ functions as an estimator for 2η. Here, $\lambda_{\alpha,t,k}$ quantifies the proportion of specific price change magnitudes, and $u_{\alpha,t,k}$ assesses the directional tendency of these changes. The value of $\hat{\eta}_{\alpha,t}$ thus reflects the behavior of the market at a microstructural level, adjusting observed prices to estimate the efficient price.

The specific definitions of $\lambda_{\alpha,t,k}$ and $u_{\alpha,t,k}$ are:

$$\lambda_{\alpha,t,k} = \frac{N^{(a)}_{\alpha,t,k} + N^{(c)}_{\alpha,t,k}}{\sum_{j=1}^{m}\left[N^{(a)}_{\alpha,t,j} + N^{(c)}_{\alpha,t,j}\right]} \tag{3.4}$$

$$u_{\alpha,t,k} = \frac{1}{2}\left(k\left(\frac{N^{(c)}_{\alpha,t,k}}{N^{(a)}_{\alpha,t,k}} - 1\right) + 1\right) \tag{3.5}$$

where

$$N^{(a)}_{\alpha,t,k} = \sum_{t_i<t} \mathbb{1}_{\left\{\left(P_{t_i}-P_{t_{i-1}}\right)\left(P_{t_{i-1}}-P_{t_{i-2}}\right)<0 \text{ and } |P_{t_i}-P_{t_{i-1}}|=k\alpha\right\}} \tag{3.6}$$

$$N^{(c)}_{\alpha,t,k} = \sum_{t_i<t} \mathbb{1}_{\left\{\left(P_{t_i}-P_{t_{i-1}}\right)\left(P_{t_{i-1}}-P_{t_{i-2}}\right)>0 \text{ and } |P_{t_i}-P_{t_{i-1}}|=k\alpha\right\}} \tag{3.7}$$

To establish a relationship between $u_{\alpha,t,k}$ and u_k, it is important to note that $u_{\alpha,t,k}$ provides a microscopic view of price dynamics within and around uncertainty zones by analyzing the ratio $N^{(c)}_{\alpha,t,k}/N^{(a)}_{\alpha,t,k}$. This analysis sheds light on the market's propensity for continuations versus alternations in price movements. On the other hand, u_k defines the theoretical boundaries of these zones based on market structure. Together, these concepts play a crucial role in aligning observed prices with efficient prices within the uncertainty zone framework, taking into account both the specific dynamics of the market and the theoretical constraints of the model.

In this model, $\lambda_{\alpha,t,k}$ determines the proportional impact of price changes of size $k\alpha$ at time t. It is computed as the ratio of the sum of alternations and continuations for tick size $k\alpha$ to the total number of these events for all tick sizes up to m. This ratio effectively captures the relative influence or contribution of price changes of size $k\alpha$ against the backdrop of all price changes, thereby underscoring the relative importance of this specific change in price magnitude over the market's overall price dynamics.

Conversely, $u_{\alpha,t,k}$ represents the balance between continuations and alternations for each tick size, formulated as a function of the ratio of continuations to alternations, adjusted by tick size k.

The formula for $N^{(a)}_{\alpha,t,k}$ is designed to count the number of alternations up to time t for a given tick size $k\alpha$. An alternation is identified when consecutive price changes result in a negative product, indicating a trend reversal. In parallel, the formula for $N^{(c)}_{\alpha,t,k}$ tallies the number of continuations up to time t for the same tick size, with a continuation occurring when consecutive price changes yield a positive product.

3.2.2 Trend period identification method

In our analysis, we introduce a statistical methodology for epoch identification within price series, underscoring the importance of empirically based statistical models in capturing the stochastic evolution of prices, a principle highlighted by Lo and MacKinlay (1999). This method, adaptable across various financial markets, not just cryptocurrencies, employs statistical analyses to discern different epochs in the price series, encompassing uptrends, downtrends, and mean reversion accurately. By refining these techniques for broader application, we contribute to the field of financial analysis, offering a more precise understanding of market

dynamics through statistical modeling. We describe the novel statistical trend period identification method as follows:

We partition the price series into disjoint windows of size $n+1$, with the number of windows depending on the length of the price series and n. Within each window, we compute the log-returns, their average μ, and standard deviation σ:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} r_i \tag{3.8}$$

$$\sigma(r) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(r_i - \mu)^2} \tag{3.9}$$

Where r_i denotes the individual log-returns within the window.

To moderate the influence of outliers and ensure accurate trend identification, we constrain the analysis to returns falling within ± 3 standard deviations from the average. This restriction curbs the impact of extreme events, which can distort the identification of trends, especially with our non-parametric methodology. Although studies like Scaillet et al. (2017) emphasize the importance of considering jumps in the analysis of cryptocurrencies, particularly Bitcoin, we do not focus on jumps in this study for several reasons. First, our primary objective is to provide a descriptive analysis of market efficiency using non-parametric methods, which do not require assumptions about the underlying data generation process. Second, controlling for jumps is essential to maintain the robustness of our trend identification approach, as jumps can significantly affect the accuracy of trend detection. Lastly, our methodology aims to identify persistent trends in the market, and the inclusion of jumps could introduce noise that obscures the true underlying trends. From these refined returns, we derive two subsets: one containing the positive returns (r^+) and the other the absolute values of the negative returns ($|r^-|$).

We then perform a difference test of the cumulative returns (CR), assuming that returns follow a t distribution, see Cont (2001):

$$CR(r) = \sum_{i=1}^{n} r_i \tag{3.10}$$

$$t = \frac{CR(r^+) - CR(|r^-|)}{\sqrt{\frac{\sigma(r^+)^2}{n_1} + \frac{\sigma(|r^-|)^2}{n_2}}} \tag{3.11}$$

At a significance level of 0.05, we designate a window as being in an uptrend if the difference in cumulative returns between the positive returns and the absolute values of the negative returns is statistically significant according to a t-test. Conversely, we identify a downtrend when the inverse relationship is statistically significant.

We assign a mean reversion status to a window when no statistically significant differences are observed between the cumulative returns of positive returns and the absolute values of negative returns, as determined by a t-test. Following the determination of the trend status for each window, we ascertain epochs. If two or more successive windows exhibit the same statistically identified trend,

they collectively constitute an epoch of that trend. However, if a subsequent window exhibits a different trend from its predecessor, then the epoch defaults to a mean reversion state. This methodology ensures epochs accurately reflect periods of consistent trend behavior, with transitions in trend direction systematically categorized as mean reversion.

In the context of this study, an epoch is defined as a period during which the price series exhibits a consistent trend behavior, either an uptrend, a downtrend, or mean reversion. Formally, an epoch is constructed by grouping consecutive windows that share the same trend classification. The statistical tests used to determine the trend within each window ensure that each epoch accurately reflects a continuous market condition.

Mathematically, let W_i be the i-th window in the price series, and let $T(W_i)$ denote the trend classification of W_i (i.e., uptrend, downtrend, or mean reversion). An epoch E_k is then defined as a maximal sequence of windows $W_i, W_{i+1}, \ldots, W_j$ such that $T(W_i) = T(W_{i+1}) = \ldots = T(W_j)$.

Formally, an epoch E_k is given by:

$$E_k = \{W_i, W_{i+1}, \ldots, W_j\} \tag{3.12}$$

such that:

$$T(W_i) = T(W_{i+1}) = \ldots = T(W_j) \tag{3.13}$$

and for any W_{i-1} or W_{j+1} not included in E_k, $T(W_{i-1}) \neq T(W_i)$ and $T(W_j) \neq T(W_{j+1})$.

This formal definition ensures that each epoch represents a continuous and homogeneous period of market behavior, allowing for a precise and systematic analysis of market dynamics across different time frames.

3.2.3 p-th variation method

The p-th variation method introduced by Cont and Das (2024) offers a sophisticated approach for estimating the Hurst exponent of a path, a metric pivotal for understanding the inherent roughness or memory in financial time series. This method's robustness is particularly beneficial for analyzing financial markets and provides insights into their efficiency. Following the identification of epochs characterized by upward trends, downward trends, and mean reversions within these series, we seek to evaluate market efficiency in light of the Weak Form Efficiency Hypothesis. Determining the Hurst exponent for each epoch serves as a key technique in this evaluation, enabling us to gauge the degree of market predictability and the presence of long-range dependence in price movements. This approach allows us to assess the efficiency of financial markets during distinct periods, offering a nuanced understanding of their behavior in accordance with the principles of the Weak Form Efficiency Hypothesis.

To estimate the Hurst exponent of a path, consider a sequence of partitions, $\pi = (\pi^n)_{n\geq 1}$, of $[0, T]$. Each partition, π^n, is defined as $\pi^n = \left(0 = t_0^n < t_1^n < \cdots < t_{N(\pi^n)}^n = T\right)$.

A path x that belongs to $C^0\left([0,T],\mathbb{R}\right)$ is said to have a finite p-th variation along the sequence of partitions, π, if there exists a continuously increasing function, $[x]_\pi^{(p)}:[0,T]\to\mathbb{R}_+$, satisfying:

$$\forall t\in[0,T],\ \sum_{\left[t_j^n,t_{j+1}^n\right]\in\pi^n:t_j^n\leq t}|x\left(t_{j+1}^n\right)-x\left(t_j^n\right)|^p\overset{n\to\infty}{\longrightarrow}[x]_\pi^{(p)}(t) \tag{3.14}$$

The requirement for a continuously increasing function, $[x]_\pi^{(p)}$, reflects the need to capture the cumulative and orderly accumulation of variations in the financial path, essential for accurately estimating the Hurst exponent. This continuity and increasing nature ensure the function faithfully represents the evolving volatility and memory of the financial series, fundamental for analyzing market dynamics over time.

This leads to the space, $V_\pi^p\left([0,T],\mathbb{R}\right)$, defined as all paths with finite p-th variation along π.

The variation index of a path, x, along a partition sequence, π, is defined as the smallest $p\geq 1$ for which x has finite p-th variation along π. This is formally represented as:

$$p^\pi(x)=\inf\left\{p\geq 1:x\in V_\pi^p\left([0,T],\mathbb{R}\right)\right\} \tag{3.15}$$

From this, the roughness index of a path, x, along π is defined as:

$$H^\pi(x)=\frac{1}{p^\pi(x)} \tag{3.16}$$

A key component of their methodology involves leveraging the concept of normalized p-th variation. For a sequence of partitions, π, of $[0,T]$ with mesh going to zero and defined by $\pi^n=\left(0=t_0^n<t_1^n<\cdots<t_{N(\pi^n)}^n=T\right)$, a path in $V_\pi^p\left([0,T],\mathbb{R}\right)$ is said to have normalized variation if there exists a continuous function $\omega(x,p,\pi):[0,T]\to\mathbb{R}$ such that:

$$\forall t\in[0,T]\sum_{\pi^n\cap[0,T]}\frac{|x\left(t_{i+1}^n\right)-x\left(t_i^n\right)|^p}{[x]_\pi^{(p)}\left(t_{i+1}^n\right)-[x]_\pi^{(p)}\left(t_i^n\right)}\left(t_{i+1}^n-t_i^n\right)\overset{n\to\infty}{\longrightarrow}\omega(x,p,\pi)(t) \tag{3.17}$$

Then, the authors demonstrate that if p-th variation is absolutely continuous, for all t in $[0,T]$, we have $\omega(x,p,\pi)(t)=t$.

To transition from theory to practice, especially with discrete financial data, Cont and Das (2024) introduce the following estimator, which is the discrete equivalent of the p-th normalized variation:

$$W(L,K,\pi,p,t,X)=\sum_{\pi^K\cap[0,T]}\frac{|x\left(t_{i+1}^K\right)-x\left(t_i^K\right)|^p}{\sum_{\pi^L\cap\left[t_i^K,t_{i+1}^K\right]}|X\left(t_{j+1}^L\right)-X\left(t_j^L\right)|^p}\times\left(t_{i+1}^K-t_i^K\right) \tag{3.18}$$

This estimator, a statistical function, assesses the variability of a financial time series X over $[0, t]$, by calculating the ratio of the sum of the p-th power differences of X within each π^K partition to a corresponding sum over a finer π^L partition. This ratio, scaled by the subinterval length of π^K, incorporates parameters L and K for different sampling frequencies, alongside p for the power parameter, and t as the interval's endpoint. With K as a block frequency and $L \gg K$ as a sampling frequency, where π^K refines into π^L, this setup groups size L samples into K groups of $\frac{L}{K}$ consecutive points. As L and K increase, the estimator converges to the normalized p-th variation: $\lim_{K\to\infty} \lim_{L\to\infty} W(L, K, \pi, p, t, x) = w(x, p, \pi)(t)$.

The variation index estimator $\hat{p}^{\pi}_{L,K}(X)$, corresponding to the signal sampled on π^L, is derived by varying p and computing $W(L, K, \pi, p, t, X)$ to solve for $\hat{p}^{\pi}_{L,K}(X)$ where the equation $W(L, K, \pi, \hat{p}^{\pi}_{L,K}(X), T, X) = T$ is satisfied. Finally, the roughness index estimator is computed as:

$$\hat{H}^{\pi}_{L,K}(X) = \frac{1}{\hat{p}^{\pi}_{L,K}(X)} \tag{3.19}$$

The roughness metric, also known as the Hurst exponent, plays a pivotal role in understanding market dynamics. A Hurst exponent value of 0.5 is indicative of a random walk, which aligns with the principles of the Weak Efficient Market Hypothesis (WEMH) that assumes market prices fully reflect all available information, resulting in an unpredictable price movement pattern. A Hurst exponent greater than 0.5 suggests a persistent behavior, indicating that positive returns are likely to be followed by positive returns and negative returns by negative returns, thus implying a long-term memory effect. Conversely, a Hurst exponent less than 0.5 suggests an anti-persistent behavior, indicating a future price movement is likely to reverse the past trend, also pointing to the existence of long-term memory but in a manner that suggests a mean-reverting pattern. These deviations from 0.5 imply potential market inefficiencies, as they suggest that not all available information is being fully incorporated into market prices. This ability to predict future price movements, to some degree, based on past data contradicts the WEMH premise that prices should be completely unpredictable if they fully integrate all known information. By quantifying these deviations from randomness, the p-th variation method not only enhances the financial analytical toolkit but also provides profound insights into market behavior and efficiency, elucidating the complex dynamics of how markets might deviate from perfectly efficient behavior.

3.2.4 Shannon Entropy

In information theory, Shannon entropy quantifies the unpredictability or randomness inherent in a system. For a discrete random variable X, with a set of possible outcomes $\{x_1, x_2, \ldots, x_n\}$, and a corresponding probability mass function $P(X)$, the entropy $H(X)$ is given by:

$$H(X) = -\sum_{i=1}^{n} P(x_i) \log_2 P(x_i) \tag{3.20}$$

This concept, originated by Shannon (1948), illuminates the degree of uncertainty in a variable's outcomes and is a cornerstone in assessing information efficiency.

When applied to financial markets, entropy becomes a powerful indicator of market dynamics. High entropy in the return series signifies a level of unpredictability consistent with efficient markets, where price movements reflect all available information and thus resemble a random walk. Conversely, low entropy points to potential patterns or predictability, suggesting market inefficiency.

To determine the theoretical maximum of Shannon entropy for a discrete random variable, we initially identify the total number of distinct states or outcomes (n) that the variable can assume. Assuming that each of these n states is equally probable, the probability of each state is $\frac{1}{n}$. We then apply Shannon's entropy formula, $H(X) = -\sum_{i=1}^{n} P(x_i) \log_2 P(x_i)$, where $P(x_i)$ is the probability of the i-th state. With equal probabilities, the formula simplifies to $H(X) = -n\left(\frac{1}{n} \log_2 \frac{1}{n}\right)$, which further simplifies to $H_{\max}(X) = \log_2 n$. This process reveals that the maximum entropy, representing the highest level of uncertainty or disorder in the system, is achieved when there is no information available to predict which state will occur, and it is solely determined by the logarithm in base 2 of the number of possible states.

The theoretical minimum of Shannon entropy for a discrete random variable occurs under conditions of complete certainty, when one outcome has a probability of 1 and all others have a probability of 0. Mathematically, if a random variable X has n possible outcomes, and one specific outcome x_k has a probability $P(x_k) = 1$ while the probabilities of all other outcomes are 0, then we apply Shannon's entropy formula: $H(X) = -\sum_{i=1}^{n} P(x_i) \log_2 P(x_i)$. In this scenario, the entropy simplifies to $H(X) = -1 \log_2 1 = 0$. This indicates the minimum entropy value, representing a state of no uncertainty or complete predictability in the system, where the outcome is known with certainty beforehand.

Fama's Weak Efficient Market Hypothesis (WEMH) posits that financial markets are informationally efficient, making price and return series inherently unpredictable Fama (1970). Shannon's entropy aligns with this hypothesis, offering a metric for randomness and, therefore, efficiency.

This theoretical framework is used in the work of Shternshis et al. (2022), who applied Shannon entropy to assess the efficiency of the Moscow Stock Exchange prior to 2022. By filtering out regularities in the data and evaluating the entropy of the resulting return series, they provided an empirical measure of market efficiency. Their findings indicate that market efficiency fluctuates over time. Specifically, periods characterized by lower entropy, indicating increased predictability, may present opportunities for deploying profitable trading strategies by exploiting market inefficiencies, once transaction costs are considered.

Shannon's entropy thereby serves not just as an abstract concept but as a practical tool in the empirical analysis of financial markets, embodying the principles of WEMH and offering insights into the temporal dynamics of market efficiency.

We employ the methodology used to detect market inefficiency in Shternshis et al. (2022) based on the utilization of Shannon's entropy as a measure of the

unpredictability of returns. We simulate Brownian motion trajectories to establish reference distributions of entropies. These simulated entropies are then used to determine the lower bounds of the 99% confidence intervals (CI). We define the market efficiency rate as the ratio of the observed return series entropy to the lower CI bound. If the efficiency rate is less than 1, we deem the market inefficient for the observed time interval. In this study, we utilize this methodology to find a measure of market efficiency using Shannon entropy.

For a theoretical random walk, we expect Shannon entropy to be high, indicating a high level of unpredictability and randomness, which aligns with the concept of an efficient market. Practically, in financial markets, Shannon entropy provides a benchmark for efficiency; the closer a market's entropy is to that of a theoretical random walk, the more efficient the market is considered. The threshold of 1 in the efficiency rate serves as a critical marker. An efficiency rate less than 1 suggests that the market's behavior is less random and more predictable than an idealized efficient market, indicating potential inefficiencies. Conversely, a rate equal to or greater than 1 implies market behavior akin to a random walk, suggesting efficiency in line with the Weak Efficient Market Hypothesis (WEMH).

3.2.5 Ljung-Box test

The Ljung-Box test, initially proposed by Ljung and Box (1978), is a statistical tool we use to evaluate the presence of significant autocorrelation in a time series. The null hypothesis for the Ljung-Box test, which assesses the absence of significant autocorrelation up to a certain number of lags, is formulated as follows:

$$H_0 : \rho_1 = \rho_2 = \ldots = \rho_m = 0 \tag{3.21}$$

Here, ρ_k denotes the autocorrelation coefficient at the k-th lag, and m represents the total number of lags considered in the test. The null hypothesis asserts that all these autocorrelation coefficients are equal to zero, indicating no significant linear dependence between the values of the time series at different lags up to the m-th lag. This suggests that, under the null hypothesis, the time series is considered to exhibit linear independence at the specified lags, implying a lack of autocorrelation.

Within the context of financial markets, this test is especially relevant when we analyze the returns of assets to determine the presence of any independence in their returns. The null hypothesis for the test states that the data points are independent, suggesting that previous values have no predictive power over future values. This concept is fundamental to the WEMH, which posits that asset prices fully incorporate all available information, thus implying that returns should follow a random walk, characterized by their independence and identical distribution. A rejection of the null hypothesis by the Ljung-Box test indicates detectable patterns or predictability in the returns, which could be interpreted as evidence against market efficiency Zhang et al. (2020b). In this study, we use the Ljung-Box test as a tool to assess the degree of efficiency within financial markets.

3.2.6 Augmented Dickey - Fuller (ADF) test

In the field of financial time series analysis, we utilize the Augmented Dickey-Fuller (ADF) test as a tool to examine the stationarity of returns. As detailed in Cox et al. (2004), the ADF test, under the null hypothesis, examines the presence of a unit root, which signifies a non-stationary process similar to a random walk. By applying the ADF test to the returns of cryptocurrencies within each epoch, we make an inference about whether the market behavior aligns with the random walk hypothesis. Should the ADF test indicate the presence of a unit root, it implies that the cryptocurrency market is efficient, incorporating all available information into prices, thereby rendering them unpredictable in a manner similar to a random walk.

Furthermore, the importance of the ADF test extends to validating the efficiency of the market over time. In the context of cryptocurrencies, if the ADF test across different epochs consistently indicates non-stationarity, it supports the idea that the cryptocurrency market remains efficient during these periods. However, if the market were to exhibit stationarity at certain epochs, this would suggest periods of predictability, contravening the weak efficient market hypothesis and potentially offering arbitrage opportunities. Thus, the ADF test not only helps us understand the dynamic nature of market efficiency but also to identify temporal changes in the informational efficiency of the cryptocurrency market.

3.2.7 Joint Binomial Distribution for Market Efficiency Indices

The construction of a joint binomial distribution for the Hurst exponent and the Shannon efficiency index represents a novel approach we employ to integrate multiple market efficiency metrics into a unified framework. This unification aims to provide a multifactorial perspective on market efficiency, acknowledging that no single measure can fully encapsulate the complex dynamics of financial markets. We predicate the choice of a binomial distribution on the dichotomous nature of each index; both are designed to categorically determine market efficiency from their respective vantage points. This initial attempt at unification allows us to explore the combined effects of temporal correlation and informational unpredictability on market behavior, despite it being a preliminary step towards a more comprehensive metric.

In constructing a joint binomial distribution for the Hurst exponent and the Shannon efficiency index, we assume independence between these two measures. This assumption stems from their distinct conceptual foundations: the Hurst exponent assesses long-term memory and temporal correlation in time series, while the Shannon efficiency index measures unpredictability and randomness, focusing on the generation of new information by price movements. Despite their differences, the mathematical modeling of their joint distribution as a product of their individual distributions is pragmatic, as it allows us to explore how these fundamentally different aspects of the data—temporal correlation versus informational unpredictability—can independently influence market efficiency.

We define the binomial parameters, n_H and n_S, as the number of trials for the Hurst and Shannon indices, respectively. A "success" in this context refers to the occurrence of efficient market behavior as defined by each index: for the Hurst index, efficiency is indicated by values within 0.45 to 0.55, and for the Shannon index, by values greater than or equal to 1. We establish the probability of success, p_H for Hurst and p_S for Shannon, through a calibration process based on these efficiency criteria.

The joint probability mass function, $f_{X,Y}(x, y)$, is given by:

$$f_{X,Y}(x, y) = P(X = x \text{ and } Y = y) = P(X = x) \cdot P(Y = y), \tag{3.22}$$

where X and Y are the random variables representing the outcomes that meet the Hurst and Shannon criteria, respectively. By combining these measures, we provide insights into the multifaceted nature of market efficiency, assuming these factors operate independently across different market conditions.

The integration of the Hurst exponent and Shannon efficiency index through a joint binomial distribution allows for a nuanced examination of market efficiency. By calibrating the binomial parameters n_H and n_S based on specific efficiency criteria—values within 0.45 to 0.55 for the Hurst index and values greater than or equal to 1 for the Shannon index—we ensure that the model accurately reflects both temporal correlation and informational unpredictability. The joint probability mass function, $f_{X,Y}(x, y)$, is constructed under the assumption of independence between the indices, which simplifies the interpretation of combined effects on market efficiency. This approach enables us to separately assess the contributions of each metric while also understanding their collective impact, providing a comprehensive framework for analyzing complex market behaviors.

3.3 Results

We utilize 5-minute frequency OHLCV (Open, High, Low, Close, and Volume) data for a range of cryptocurrencies, including BNBBUSD (Binance Coin against USD), BTCBUSD (Bitcoin against USD), ETHBUSD (Ethereum against USD), ADABUSD (Cardano against USD), XRPBUSD (Ripple's XRP against USD), SOLBUSD (Solana against USD), and DOGEBUSD (Dogecoin against USD). This data spans from $2022-10-01$ to $2023-04-01$, encompassing a total of 52257 data points for each cryptocurrency, obtained via the Binance API. The selected period is intended to capture a representative sample of market conditions. The choice of a 5-minute interval strikes a balance between capturing intraday price movements and managing the volume of data, ensuring that significant price fluctuations and trading activities are adequately represented without overwhelming computational resources. We selected a range of cryptocurrencies to provide a comprehensive overview of the market: BTC and ETH as market leaders, BNB as a representative of an intermediate market characterized by lower volatility and higher liquidity, and ADA, XRP, SOL, and DOGE as smaller yet rapidly growing assets with significant

market potential. This diverse selection ensures that our analysis captures the dynamics across different segments of the cryptocurrency market.

3.3.1 Results - Uncertainty zones

Using the uncertainty zones methodology to estimate an efficient prices process X from observed prices P, we observe in figure (3.2) that there are no apparent differences between the two prices processes for BTCBUSD as a visual example. To confirm or reject this observation, we use the Kolmogorov-Smirnov test. The null hypothesis of the Kolmogorov-Smirnov test is that the two samples being compared are drawn from the same distribution. This allows us to compare whether the two price series, X and P, are equal in distribution for all the cryptocurrencies studied. The results of this test are presented in table (3.1).

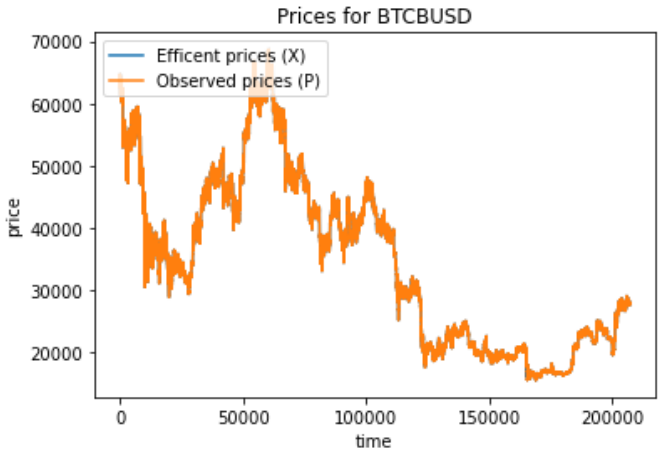


Figure 3.2: Comparison of observed P and efficient X price series for BTCBUSD

Table 3.1: Kolmogorov-Smirnov test for series X-prices and P-prices

Symbol	p-value
BNBBUSD	0.999
BTCBUSD	1.0
ETHBUSD	1.0
ADABUSD	1.0
XRPBUSD	1.0
SOLBUSD	0.999
DOGEBUSD	0.999

The results of the Kolmogorov-Smirnov test, which fail to reject the null hypothesis, indicate that the distributions of the efficient price series X and the observed prices P are not significantly different for all cryptocurrencies. The remarkable similarity between the observed and efficient price series can be attributed, in part,

to the high frequency of the data used. In high-frequency time series, the time interval between consecutive observations is short, which reduces the probability of significant events affecting the price between two time points. As a result, price movements are captured with high accuracy, which may explain the close match between observed and efficient prices. Given the above reasons, and for the sake of simplicity and robustness of our analysis, we decide to use only observed prices. Furthermore, based on our analysis, we affirm that the diffusion model, as initially assumed, is an excellent representation of price behavior.

Although our analysis did not find significant differences between the observed prices P and the efficient price series X, this does not contradict the findings of Aït-Sahalia and Yu (2009). Their research highlights that market microstructure noise is an inherent characteristic of high-frequency price observations, affecting the observed prices. This noise can cause the observed prices to behave similarly to the efficient prices, particularly in high-frequency settings. In this specific section of our study, we used a diffusion price model for efficient prices as described by Robert and Rosenbaum (2012). However, it is important to emphasize that in the rest of our study, we do not assume a specific model for the prices. The methodology used to identify trend epochs and the metrics used to measure efficiency are non-parametric, meaning there is no need to presuppose a specific model for the prices. Therefore, the close match observed in our study is consistent with the understanding that high-frequency prices, despite being influenced by microstructure noise, can still approximate the efficient price process very closely Aït-Sahalia and Yu (2009).

3.3.2 Results - Trend period identification method

For illustrative purposes, we apply the trend period identification methodology to a subset of the data, spanning 4 days with a 5-minute frequency, yielding 1152 data points for BTCBUSD. This subset serves as a representative snapshot, enabling clear visualization of the epoch identification process, which might be more challenging to discern when considering the entire dataset.

The choice of window size is pivotal in epoch identification. Using small windows, e.g., of just 6 data points or a half-hour duration, we observe frequent state transitions, leading most epochs to be categorized as mean reversion epochs, as demonstrated in figure 3.3. Such frequent transitions arise primarily from the constrained temporal scope of these small windows, which renders capturing substantive and extended trends a challenge.

On the other end of the spectrum, employing excessively large windows, such as those encompassing 120 data points or spanning 10 hours, presents its own set of challenges. As illustrated in figure 3.4, these extended windows might oversimplify price movement patterns, erroneously categorizing genuine uptrends and downtrends as mean reversion epochs.

Striking a balance in window size is essential. Windows comprising 24 data points or spanning two hours show promising results in our subset analysis, effectively capturing short-term price movements and facilitating accurate epoch identification, as depicted in figure 3.5.

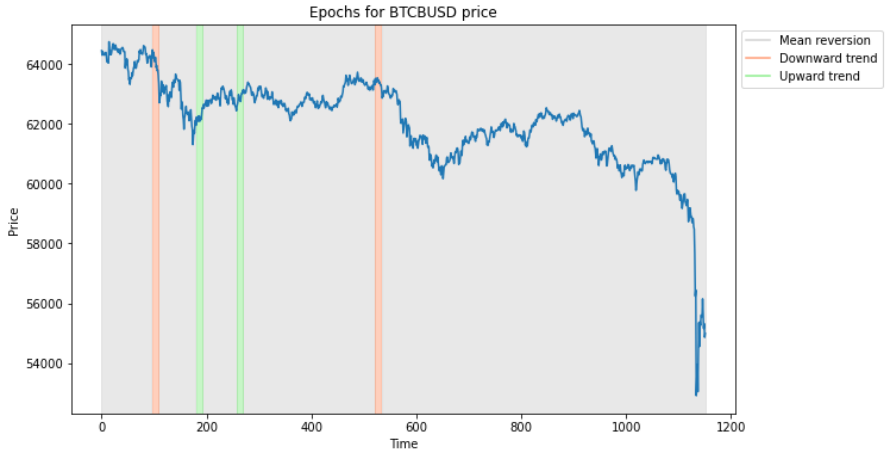


Figure 3.3: BTCBUSD epochs Chart with Window Size 6

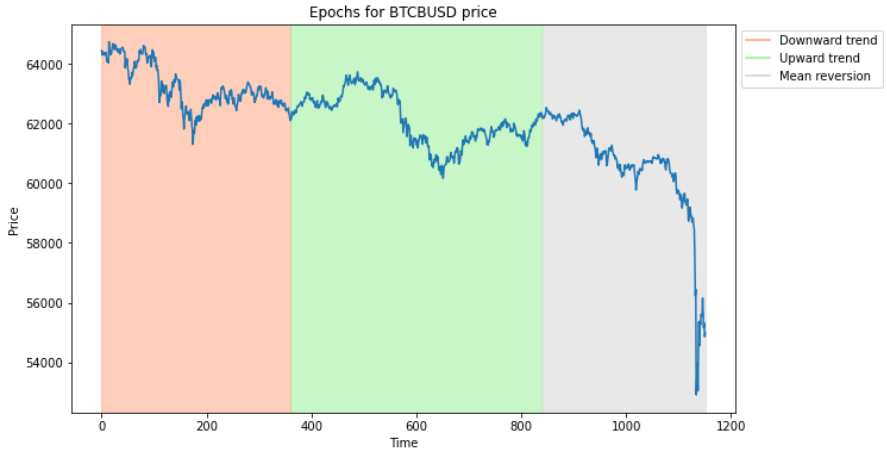


Figure 3.4: BTCBUSD epochs Chart with Window Size 120

A plausible explanation for the effectiveness of the 24-point window size is its balanced approach to capturing short-term market trends while minimizing noise. In the context of 5-minute frequency data, this two-hour window allows for an accurate reflection of rapid market movements without the oversensitivity to minor price changes seen in smaller windows.

This observation suggests that for the complete dataset, spanning from $2022-10-01$ to $2023-04-01$ with 52257 data points for each cryptocurrency, window sizes around 24 data points are appropriate for trend identification.

To ensure the appropriateness of selected window sizes in real data analysis, we implement a Monte Carlo simulation, executing 1000 iterations to determine the optimal window size for trend identification within financial time series. Each simulation produces series with a total average length of 1070 data points, each containing 10 to 15 trends with randomly assigned lengths ranging from 30 to 150 data points. The analysis examines window sizes ranging from 5 to 150, in 20-unit increments, to determine their effectiveness in identifying these trends.

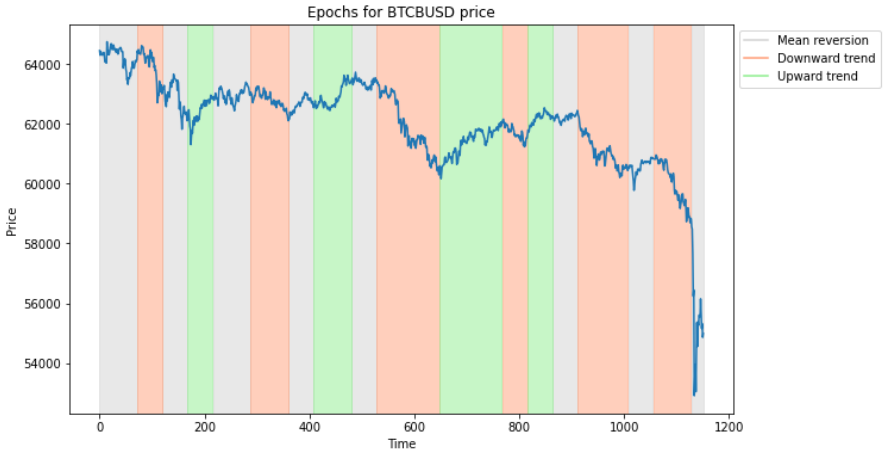


Figure 3.5: BTCBUSD epochs Chart with Window Size 24

The essence of our analytical approach involves comparing the epochs, which were identified through the analysis, against a set of predefined trends. This comparison looks for matches in trend type within overlapping intervals, allowing for an assessment of precision for each window size. Through this process, we are able to quantitatively evaluate the effectiveness of each window size in accurately capturing the underlying trends, thereby validating our window size selection for analyzing real-world financial data.

The empirical results, visualized through histogram analyses, reveal significant insights into the distribution of optimal window sizes. The histogram in Figure 3.6 illustrates the frequency of each window size deemed most precise across all simulations. Notably, window sizes ranging from 24 to 60 units frequently emerge as the most precise, suggest their effectiveness in trend detection across various series lengths and configurations. This range, with a window size around 24 units often being particularly effective, reinforce the notion that these window sizes aptly capture trends. Similarly, the histogram in Figure 3.7 displays the distribution of maximum precision levels attained throughout the simulations, corroborate the proficiency of the methodology. When examining the correlation between the series length and the optimal window size using the Spearman's rank correlation coefficient, the analysis yields a value of -0.029. This result indicates no significant dependence between these variables, further substantiating the robustness of our approach.

3.3.3 Results of the P-th variation methodology

We perform a sanity check of the implementation of the P-th variation methodology using fractional Brownian motions (fBm) simulations for different values of the Hurst parameter H and compare the results of estimation versus the simulation value.

The `fbm` library in Python is used to simulate paths of fBm. The Fractional Brownian motion is characterized by its Hurst coefficient H. The H parameter

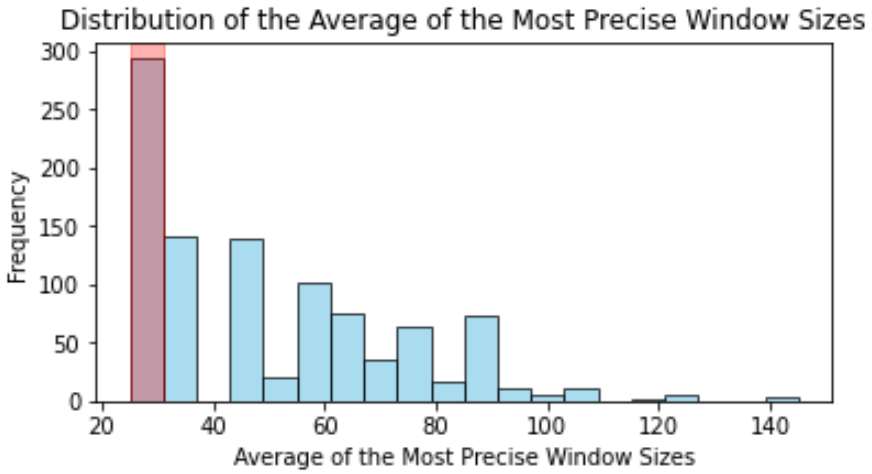


Figure 3.6: Distribution of average optimal window sizes across simulations.

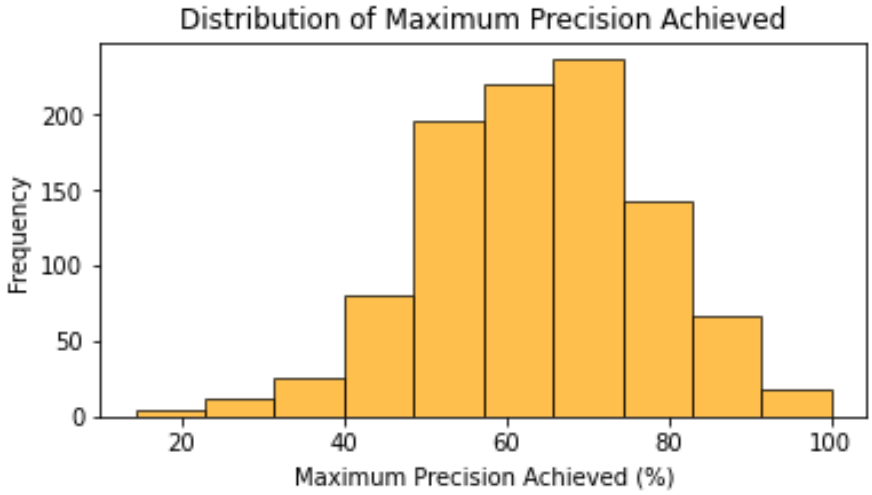


Figure 3.7: Distribution of maximum precision levels achieved across simulations.

determines the roughness of the path and can take values in the range $0 < H < 1$.

The `daviesharte` method, available in the `fbm` library, offers a sophisticated technique to simulate fractional Brownian motion (fBm), proposed by Davies and Harte (1987). It is distinctive for its application of the Fast Fourier Transform (FFT) to synthesize fBm paths. The detailed process of the `daviesharte` method is as follows:

- **Initialization:** it starts with the generation of a sequence of independent and identically distributed (i.i.d.) random variables, specifically Gaussian white noise. These variables serve as the foundational elements for constructing the fBm.
- **Covariance Matrix Construction:** the key to the `daviesharte` method is the formulation of a specific covariance matrix that encapsulates the

statistical properties of fractional Brownian motion (fBm) as a process of random variables. These properties include self-similarity and long-range dependence, fundamental to the fBm's nature. The matrix is designed to represent the correlation structure inherent in the increments of the fBm, effectively capturing the essence of these random variables' statistical behavior.

- **Application of FFT:** the FFT applies to the sequence of random variables. This step is crucial as it transforms the white noise in a way that embeds the fBm's statistical characteristics, as dictated by the covariance matrix.
- **Inverse FFT:** following the transformation, an inverse FFT is performed. The output from this inverse process is a set of values that approximate the increments of fBm.
- **Path Generation:** the final step involves the cumulative summing of these increments to generate a continuous path of fBm.

The `daviesharte` method relies on FFT and makes it particularly efficient for generating large samples of fBm. Besides `daviesharte`, the `fbm` library also includes other methods to simulate fBm, such as the `cholesky` and `hosking` methods. Each of these methods offers unique advantages and is suitable for different scenarios. The advantages of the `daviesharte` method over other fBm simulation methods include its computational efficiency, as it leverages the FFT, and its capability to produce paths that closely resemble the desired statistical properties of fBm.

For the simulations, we choose the following values for the Hurst coefficient $H = [0.1, 0.3, 0.5, 0.8]$. For each H, 2000 simulations are conducted, and then we compute the average of the estimated Hurst coefficients. Given the duration of the epochs previously determined, some of which are short, we decide to simulate series with a length of 50 data points.

The table 3.2 presents the average estimated Hurst coefficients using the method described in section 3.2 for each simulated path:

Table 3.2: Average estimated Hurst coefficients for simulated paths.

Simulated H Value	Average Estimated H
0.1	0.1575
0.3	0.2983
0.5	0.4694
0.8	0.7233

In the table 3.2, we observe that the estimated Hurst coefficients are close to the simulated values, suggesting an initial alignment between our methodology's estimations and the expected characteristics of fractional Brownian motions for various Hurst values H. The simulated paths for different H values seem to reflect the anticipated roughness and temporal behavior corresponding to each H, which could indicate the potential accuracy of the implemented methodology.

Additionally, further simulations were conducted, incrementally increasing both the length of the series and the number of simulations, and the precision of the Hurst coefficient estimations remained consistent. The estimated H values continued to closely match the values set in the simulations, further evidencing the robustness of our methodology and its implementation.

Table 3.3: Average Estimated H Values for Price Series

Asset	Downward Trend	Mean Reversion	Upward Trend
BNBBUSD	0.4085	0.4123	0.4239
BTCBUSD	0.4050	0.3894	0.4032
ETHBUSD	0.4039	0.3995	0.4144
ADABUSD	0.4219	0.4092	0.4374
XRPBUSD	0.4271	0.4258	0.4371
SOLBUSD	0.4217	0.4047	0.4204
DOGEBUSD	0.4238	0.4175	0.4334

Table 3.3 displays the average estimated H values for cryptocurrency price series across various market states. Noteworthy, the table presents H values consistently below 0.5, indicating a level of roughness in all market states for these cryptocurrencies.

For these cryptocurrency price series, the 'Mean Reversion' period typically demonstrates a marginally lower H value compared to the 'Upward Trend' and 'Downward Trend' periods. This suggests a more pronounced anti-persistent behavior during mean-reverting phases, potentially due to faster price corrections.

A uniform observation across all market states is that the H value remains below 0.5. This characteristic implies a persistent inefficiency in the behavior of these markets, contrasting with the random walk theory proposed by the traditional Efficient Market Hypothesis (EMH) Fama (1970).

3.3.4 Results - Shannon Entropy

The Table 3.4 encapsulates the aggregated results of an efficiency analysis we conduct on various cryptocurrencies across identified epochs of mean reversion, downward trend, and upward trend. The average efficiency ratios, derived from the methodology proposed by Shternshis et al. (2022), quantify the predictability of cryptocurrency returns. A higher ratio indicates a market that behaves more randomly, which aligns with efficiency as posited by the Weak Efficient Market Hypothesis. The standard deviation sheds further light on the variability of the efficiency ratios, adding dimension to our understanding of market efficiency stability across different market trends.

In the analysis of the efficiency metrics under different market trends, as illustrated in Figures 3.8, 3.9, and 3.10 and in the Table 3.4, BTCBUSD and ETHBUSD exhibit notable characteristics that stand out. In the context of the weak form of market efficiency hypothesis, which assumes that all past trading information is

Table 3.4: Average Efficiency Ratios for Cryptocurrency Return Series

Asset	Mean Reversion	Downward Trend	Upward Trend
BNBBUSD	0.6868 (± 0.1382)	0.7325 (± 0.1542)	0.7278 (± 0.1490)
BTCBUSD	0.9695 (± 0.1617)	1.0114 (± 0.0700)	1.0083 (± 0.0710)
ETHBUSD	0.9963 (± 0.1576)	1.0148 (± 0.0714)	1.0210 (± 0.0765)
ADABUSD	0.7782 (± 0.1468)	0.8251 (± 0.1327)	0.8165 (± 0.1405)
XRPBUSD	0.8150 (± 0.1509)	0.8497 (± 0.1326)	0.8538 (± 0.1248)
SOLBUSD	0.7764 (± 0.1646)	0.8325 (± 0.1476)	0.8202 (± 0.1556)
DOGEBUSD	0.8805 (± 0.1512)	0.9297 (± 0.1109)	0.9366 (± 0.0990)

already reflected in asset prices, both BTCBUSD and ETHBUSD show a median efficiency ratio that frequently dips below the threshold of 1, suggesting a potential for predictive patterns based on past price movements. However, their median efficiency ratio often is closer to the threshold compared to other cryptocurrencies, which could imply that these markets are, on average, better at incorporating historical price information. This situation is also confirmed through Table 3.4, where we can observe that the average values of the ratio for BTCBUSD and ETHBUSD are close to the threshold 1.

The observed variability through the extent of the interquartile ranges (IQRs) and the presence of outliers provide additional insight. BTCBUSD and ETHBUSD tend to have lower variability, especially in the upward and downward trend scenarios, which could indicate a more stable incorporation of past information into current prices, aligned with the weak form efficiency hypothesis. Yet, during the mean reversion trend, the highest variability, particularly in BTCBUSD, indicates moments where the market efficiency significantly deviates from the median. This deviation could reflect instances where market sentiment or external shocks impact price movements more heavily, thus challenging the notion that past information is fully reflected in current prices.

When contrasted with BTCBUSD and ETHBUSD, other cryptocurrencies reflect lower efficiency during trends and mean reversion conditions, with medians consistently below the threshold of 1. This observation indicates a widespread departure from the weak form of market efficiency, suggesting that these markets may not fully integrate historical price information into current values. Such a trend implies the presence of persistent inefficiencies that could potentially be exploited by strategies based on historical data.

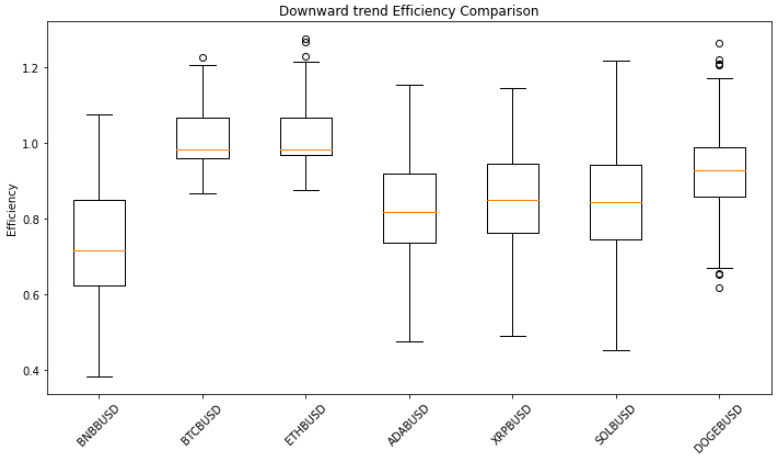


Figure 3.8: Downward trend Efficiency Comparison of cryptocurrencies.

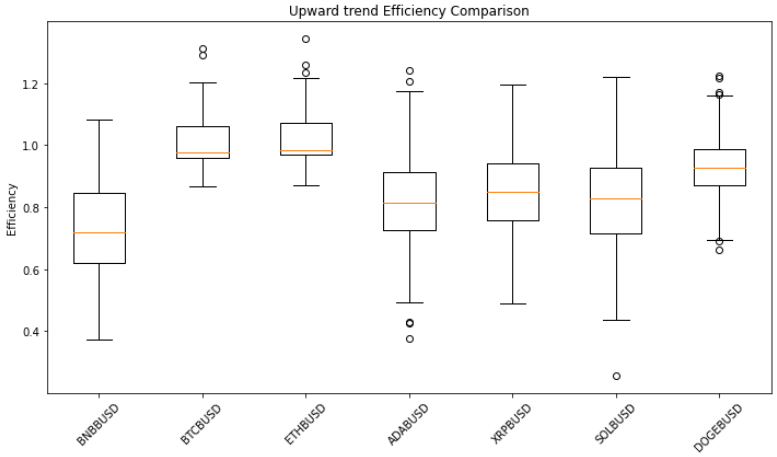


Figure 3.9: Upward trend Efficiency Comparison of cryptocurrencies.

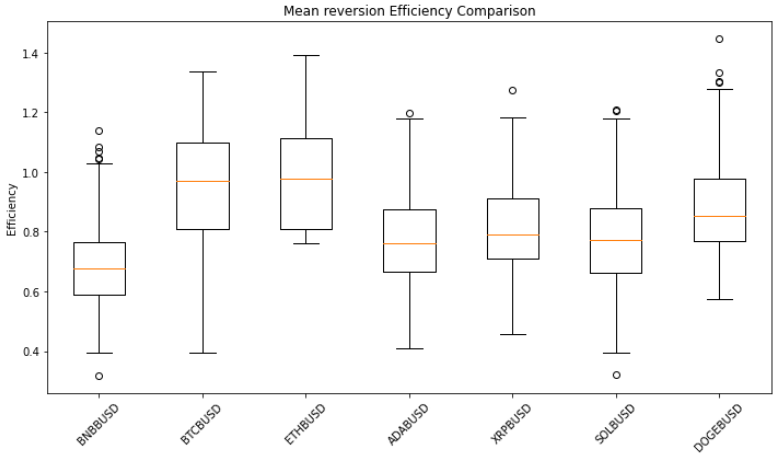


Figure 3.10: Mean reversion Efficiency Comparison of cryptocurrencies.

Collectively, these statistical analyses may suggest that market efficiency in the cryptocurrency space is not uniform and is highly influenced by market trends. This differential behavior suggests that external factors, investor sentiment, and inherent market mechanisms play significant roles in shaping the efficiency of these digital asset markets.

3.3.5 Results - Ljung-Box test

For each identified trend—upward, downward, and mean reversion—across different cryptocurrencies, we administer the Ljung-Box test for lags from 1 to 10. Subsequently, we calculate average p-values for each type of trend for every cryptocurrency across each lag. This approach allows us to explore discernible patterns or predictability in returns. According to Zhang et al. (2020b), the rejection of the null hypothesis by the Ljung-Box test suggests detectable patterns or predictability, challenging the Weak Form of the Efficient Market Hypothesis.

Table 3.5: Average p-values for Ljung-Box Test across Different Trends for Each Cryptocurrency for Selected Lags (1, 5, and 10).

Cryptocurrency	Mean Reversion	Downward Trend	Upward Trend
BNBBUSD	0.415, 0.440, 0.445	0.399, 0.388, 0.405	0.444, 0.401, 0.442
BTCBUSD	0.439, 0.425, 0.467	0.431, 0.394, 0.415	0.452, 0.410, 0.452
ETHBUSD	0.430, 0.433, 0.477	0.428, 0.433, 0.460	0.452, 0.437, 0.466
ADABUSD	0.439, 0.404, 0.457	0.437, 0.418, 0.432	0.471, 0.417, 0.454
XRPBUSD	0.444, 0.448, 0.479	0.448, 0.439, 0.453	0.460, 0.403, 0.430
SOLBUSD	0.466, 0.459, 0.473	0.452, 0.442, 0.442	0.461, 0.403, 0.431
DOGEBUSD	0.454, 0.441, 0.465	0.473, 0.435, 0.441	0.453, 0.396, 0.439

The summary of average p-values for the Ljung-Box test, as depicted in Table 3.5, provides valuable insights into the predictability of cryptocurrency markets under different trends. Specifically, our analysis reveals that average p-values across various cryptocurrencies and lags always exceed the 0.05 significance threshold. This suggests a lack of sufficient evidence to reject the null hypothesis, indicating that past prices do not provide reliable indications for future prices. Such findings are consistent with the Weak Form of the Efficient Market Hypothesis, which posits that asset prices fully reflect all available historical information. As supported by Zhang et al. (2020b), our inability to reject the null hypothesis through the Ljung-Box test in these contexts reaffirms the market's efficiency, challenging the potential for detectable patterns or predictability that could be exploited for trading advantages. The implication of these results not only reinforces the notion of market efficiency but also underscores the complexity of predicting cryptocurrency price movements based on historical data alone.

3.3.6 Results - Augmented Dickey - Fuller (ADF) test

The table 3.6 presents the average p-values from the Dickey-Fuller test for the returns of the cryptocurrencies studied during distinct market conditions: mean reversion, downward trends, and upward trends. It indicates that there is more substantial evidence of stationarity, as shown by lower p-values, when the cryptocurrencies exhibit pronounced trends, either upward or downward, as opposed to periods characterized by mean reversion. Specifically, the p-values falling below the standard significance level of 0.05 in the trend columns imply we reject the null hypothesis of a unit root, suggesting stationary behavior in the return series under these trending conditions. On the other hand, the higher p-values in the mean reversion scenarios point to a weaker rejection of the null hypothesis, hinting that the return series may not be stationary when the returns are expected to revert to the mean.

Table 3.6: Average p-values from the Dickey-Fuller test for different epochs of returns for each cryptocurrency.

Cryptocurrency	Mean reversion	Downward trend	Upward trend
BNBBUSD	0.0583	**0.0047**	**0.0085**
BTCBUSD	0.0843	**0.0232**	**0.0094**
ETHBUSD	0.0632	**0.0136**	**0.0083**
ADABUSD	0.0879	**0.0118**	**0.0076**
XRPBUSD	0.0806	**0.0224**	**0.0119**
SOLBUSD	0.1066	**0.0253**	**0.0189**
DOGEBUSD	0.1053	**0.0226**	**0.0165**

Our findings from the Dickey-Fuller tests on cryptocurrency returns feed directly into our broader market efficiency study. The evidence of stationarity in returns during trending periods suggests that price movements in these phases are not random but potentially influenced by persistent factors, which might be predictable to some extent. This could challenge the Weak Efficient Market Hypothesis (WEMH), which posits that asset prices reflect all available information and are thus unpredictable. The higher p-values in the mean reversion period, however, align more closely with the WEMH, indicating that returns may revert to a mean without discernible patterns. Our market efficiency analysis must, therefore, consider these differing behaviors across market conditions to ascertain the degree to which cryptocurrency markets adhere to or deviate from the principles of the WEMH.

3.3.7 Results - Joint Binomial Distribution for Market Efficiency Indices

This section presents the findings derived from the implementation of the joint binomial probability function for the Shannon and Hurst indices across various cryptocurrencies in distinct epochs characterized as periods of mean reversion, upward trends, and downward trends. By integrating these two metrics into a

singular framework, we aim to unveil a multifaceted perspective on market efficiency, capturing both the temporal correlation and informational unpredictability of price movements. This joint distribution approach not only facilitates a nuanced assessment of market behavior under varying conditions but also serves as a foundational step towards synthesizing diverse efficiency metrics into a coherent, unified metric. The results herein explore the effectiveness of this preliminary model in distinguishing efficient from inefficient market dynamics, illustrating the potential for a more comprehensive and multifactorial metric of market efficiency.

The joint probability distribution derived from these indices provides insights into the fluctuations and efficiency trends within the cryptocurrency market, reflecting the unique dynamics of each cryptocurrency during these distinct market epochs. Based on the results from the Ljung-Box test, which suggest no significant autocorrelations, the markets are seen as consistently efficient across all epochs. On the other hand, the Augmented Dickey-Fuller test indicates consistent inefficiency for epochs other than mean reversion. However, when considering a significance level of 0.1, mean reversion might also be deemed inefficient for almost all cryptocurrencies. This dichotomy in test outcomes encourages us to construct the joint distribution focusing solely on the Hurst and Shannon efficiency variables.

We begin the analysis by verifying the independence between the Hurst and Shannon indices using Spearman's rank correlation coefficient. This method allows us to assess both the strength and the direction of the association between these indices. To confirm their independence, we rely on the p-value obtained from Spearman's test, considering the indices to be independent if the p-value exceeds the significance level of 0.05.

Table 3.7: Correlation coefficient for Shannon and Hurst indices across different cryptocurrencies and market epochs.

Cryptocurrency	Mean reversion	Downward trend	Upward trend
BNBBUSD	0.3093	0.2447	0.1912
BTCBUSD	-0.0358	-0.3614	0.0208
ETHBUSD	0.0022	0.0756	0.0531
ADABUSD	-0.0225	0.2712	0.5937
XRPBUSD	0.1251	0.1985	0.1328
SOLBUSD	-0.0234	0.2352	0.2328
DOGEBUSD	0.1599	0.1250	0.1910

The correlation table 3.7 provides a nuanced view of the relationship between the Shannon and Hurst indices across different cryptocurrencies and market conditions. For instance, BNBBUSD shows a moderate positive correlation across all market conditions, indicating a tendency for these indices to move together. BTCBUSD, however, exhibits mostly weak negative correlations, especially in downward trends, suggesting an inverse relationship between the indices during market declines. In contrast, ETHBUSD demonstrates minimal correlation across conditions, indicating little to no relationship between the indices. Interestingly, ADABUSD shows

a strong positive correlation in upward trends, suggesting a more pronounced concurrent movement of the indices during market rises. Other cryptocurrencies like XRPBUSD and SOLBUSD display varied correlations across different market epochs, reflecting the diverse and complex nature of these relationships in the cryptocurrency market.

Table 3.8: Independence test results for Shannon and Hurst indices across different cryptocurrencies and market epochs.

Cryptocurrency	Epoch	Correlation	P-value	Independent
BNBBUSD	Mean reversion	0.3093	0.0289	No
BNBBUSD	Downward trend	0.2447	0.1387	Yes
BNBBUSD	Upward trend	0.1912	0.2194	Yes
BTCBUSD	Mean reversion	-0.0358	0.7936	Yes
BTCBUSD	Downward trend	-0.3614	0.0258	No
BTCBUSD	Upward trend	0.0208	0.8884	Yes
ETHBUSD	Mean reversion	0.0022	0.9869	Yes
ETHBUSD	Upward trend	0.0531	0.7290	Yes
ETHBUSD	Downward trend	0.0756	0.6428	Yes
ADABUSD	Mean reversion	-0.0225	0.8770	Yes
ADABUSD	Upward trend	0.5937	2.1542e-05	No
ADABUSD	Downward trend	0.2712	0.0683	Yes
XRPBUSD	Mean reversion	0.1251	0.3768	Yes
XRPBUSD	Downward trend	0.1985	0.1715	Yes
XRPBUSD	Upward trend	0.1328	0.4078	Yes
SOLBUSD	Mean reversion	-0.0234	0.8682	Yes
SOLBUSD	Downward trend	0.2352	0.0810	Yes
SOLBUSD	Upward trend	0.2328	0.1655	Yes
DOGEBUSD	Mean reversion	0.1599	0.2777	Yes
DOGEBUSD	Upward trend	0.1909	0.2199	Yes
DOGEBUSD	Downward trend	0.1250	0.4420	Yes

The independence test results shown in table 3.8 provide valuable insights into the relationship between the Shannon and Hurst indices across various cryptocurrencies and market epochs. The majority of epochs for each cryptocurrency propose a 'Yes' for independence, as indicated by the p-values being above the significance level of 0.05. This suggests that, in most cases, there is no statistically significant correlation between the Shannon and Hurst indices, supporting the assumption of their independence. For instance, BTCBUSD and ETHBUSD predominantly exhibit independence across different market conditions. However, some specific cases like BNBBUSD in mean reversion and BTCBUSD in downward trends show dependency, highlighting instances where the indices move together in a statistically significant manner.

In Table 3.9, the calibrated binomial parameters for the Hurst index reflect the likelihood of market efficiency in different cryptocurrencies. Specifically, higher

Table 3.9: Binomial Parameters for Hurst Index across Different Market Trends for Each Cryptocurrency.

Cryptocurrency	Mean reversion (n, p)	Downward trend (n, p)	Upward trend (n, p)
BNBBUSD	(50, 0.24)	(38, 0.2895)	(43, 0.3488)
BTCBUSD	(56, 0.125)	(38, 0.2368)	(48, 0.1667)
ETHBUSD	(56, 0.125)	(40, 0.2)	(45, 0.3111)
ADABUSD	(50, 0.26)	(46, 0.2391)	(44, 0.3864)
XRPBUSD	(52, 0.2115)	(49, 0.2245)	(41, 0.3171)
SOLBUSD	(53, 0.2264)	(56, 0.2679)	(37, 0.4054)
DOGEBUSD	(48, 0.1875)	(40, 0.225)	(43, 0.3953)

probability values in this context suggest a greater likelihood of observing efficient market conditions, as defined by the Hurst criterion (efficiency is indicated by Hurst values within the range of 0.45 to 0.55). Cryptocurrencies like SOLBUSD and ADABUSD demonstrate this trend, showing a higher probability of efficiency especially during upward trends. This implies that in such market conditions, these cryptocurrencies frequently exhibit behavior aligned with efficient market theories, such as random walk characteristics.

Table 3.10: Binomial Parameters for Shannon Index across Different Market Trends for Each Cryptocurrency.

Cryptocurrency	Mean reversion (n, p)	Downward trend (n, p)	Upward trend (n, p)
BNBBUSD	(50, 0.2)	(38, 0.2368)	(43, 0.1395)
BTCBUSD	(56, 0.3214)	(38, 0.4211)	(48, 0.4375)
ETHBUSD	(56, 0.4107)	(40, 0.325)	(45, 0.4444)
ADABUSD	(50, 0.28)	(46, 0.3478)	(44, 0.2727)
XRPBUSD	(52, 0.3654)	(49, 0.2449)	(41, 0.3902)
SOLBUSD	(53, 0.2075)	(56, 0.3214)	(37, 0.3514)
DOGEBUSD	(48, 0.375)	(40, 0.4)	(43, 0.4186)

Table 3.10 shows the calibrated binomial parameters for the Shannon index, indicating the likelihood of market efficiency across various cryptocurrencies and market epochs. In this context, higher probability values (p) signal a greater chance of efficient market behavior, as defined by the Shannon criterion (efficiency is assumed for index values ≥ 1). Cryptocurrencies like BTCBUSD and ETHBUSD, for instance, display notably high probabilities in upward trends, pointing towards a tendency for these currencies to exhibit efficient market behavior in such periods.

The data in the tables 3.9 and 3.10 illustrates the variable nature of market efficiency as measured by the Shannon and Hurst indices, dependent on both the cryptocurrency in question and the market conditions. This variation emphasizes the dynamic nature of cryptocurrency markets and the importance of utilizing multiple metrics for a comprehensive assessment of market efficiency.

As examples of the joint binomial distributions constructed, the analysis presents

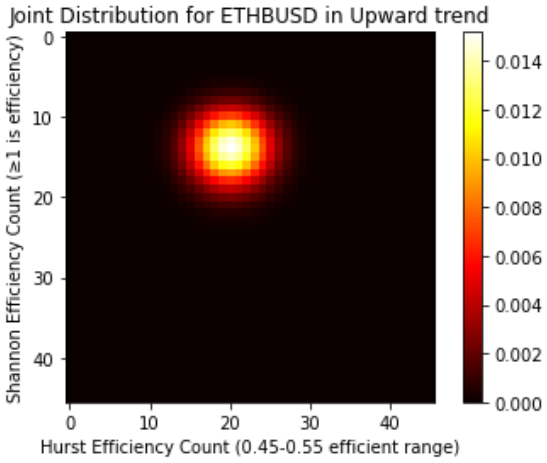


Figure 3.11: Joint Distribution for ETHBUSD in Upward Trend.

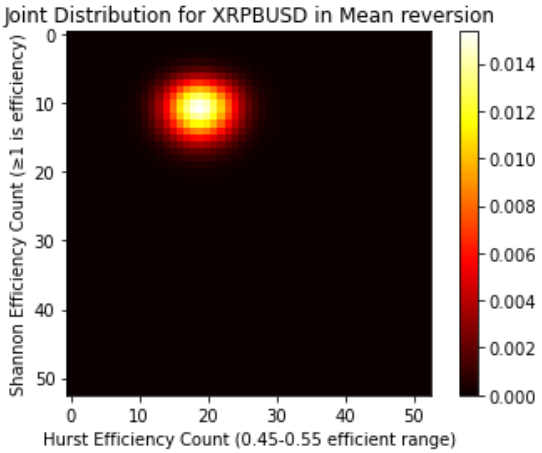


Figure 3.12: Joint Distribution for XRPBUSD in Mean Reversion.

heat maps for three cryptocurrencies during each of the identified epochs. The heat maps reveal the intricate dynamics of market efficiency as defined by the Hurst and Shannon indices. In the heat map for the upward trend of ETHBUSD, shown in Figure 3.11, the dense central concentration of high values indicates a prevalence of efficient market behavior within counts of 12 to 28 for the Hurst and 5 to 21 for Shannon. This observation suggests episodes of efficiency, albeit not uniformly distributed across all count ranges, aligning with the notion that market efficiency is not static but fluctuates.

For XRPBUSD in mean reversion, Figure 3.12 a similar pattern emerges, implying that while efficient market behavior is detected, it is part of a broader context where inefficiencies may still exist and could be exploited by quantitative trading strategies.

The BTCBUSD downward trend heatmap Figure 3.13 displays a smaller yet significant central region of high values, challenging the expected trend of inefficiency

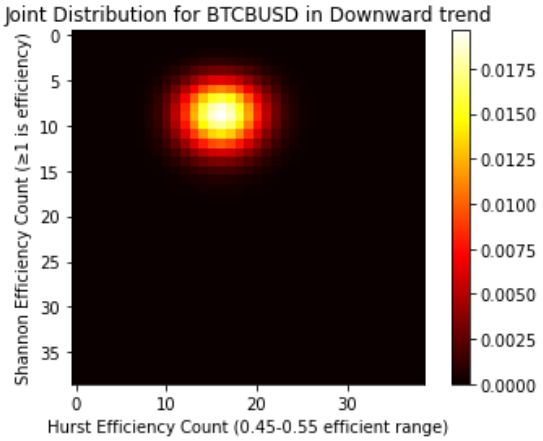


Figure 3.13: Joint Distribution for BTCBUSD in Downward Trend.

in bearish markets and suggesting a moderate probability of efficiency, albeit not pervasive. This resonates with the understanding that market efficiency varies with time and conditions, and while there are intervals of efficient behavior, they coexist with potential inefficiencies.

These visual patterns affirm the financial theory that recognizes varying degrees of efficiency, as posited by Fama, and reflect a market that is efficient at times, but not uniformly or predictably so, aligning with the dynamic and complex nature of cryptocurrency markets.

3.3.8 Frequency-Dependent Efficiency Analysis

In the pursuit of a comprehensive understanding of market efficiency, the study progresses by replicating all stages of the initial methodologies, albeit with a pivotal alteration in the temporal granularity of the data. Where the initial inquiry utilized a 5-minute frequency, we now extend the analysis to encompass data frequencies of both 10 minutes and 30 minutes. This strategic variation aims to elucidate the influence of observational frequency on the partial outcomes of the study. By comparing results across these distinct intervals, we intend to unearth any frequency-dependent nuances in market behavior, thereby offering a more granular perspective on the efficiency patterns that characterize the cryptocurrency markets under different temporal lenses. We find that cryptocurrency markets exhibit a frequency-dependent efficiency, with lower-frequency data showing increased market efficiency as evidenced by the Hurst index, Shannon entropy, Ljung-Box, and Dickey-Fuller tests.

Comparing the average p-values from the Dickey-Fuller tests across various epochs and frequencies for each cryptocurrency (see Appendix **??**), we observe significant variances that may indicate differences in market efficiency. For mean reversion epochs, the p-values are generally higher, suggesting a less strong rejection of the unit root hypothesis, which could indicate a tendency toward a more efficient market as prices correct themselves over time. In contrast, during downward

and upward trends, the lower p-values across the 5-minute, 10-minute, and 30-minute frequencies consistently support trend-stationarity, suggesting that prices are diverging from a mean, potentially indicating inefficiencies or periods where market prices are driven by momentum rather than intrinsic values. Intriguingly, as the frequency decreases from 5 to 30 minutes, there is a general trend of increasing p-values for mean reversion epochs and mixed results for upward and downward trends, which may reflect the complex dynamics of market efficiency over varying time frames.

Upon reviewing the average estimated Hurst (H) values for price series at varying frequencies, see Appendix .4, is evident that all recorded values are consistently below the critical level of 0.5. This finding points to a persistent tendency towards mean reversion within the market, suggesting inherent inefficiencies regardless of whether the data is observed at 5-minute, 10-minute, or 30-minute intervals. The values, while show minor fluctuations across different epochs and frequencies, reinforce the notion that the market does not fully adhere to the weak efficient market hypothesis, allowing for potential predictability in price movements.

Analyzing the average Shannon Efficiency Ratios for cryptocurrency return series across different time intervals reveal higher efficiency in markets with longer sampling periods, as detailed in Appendix .5. This evidence by the generally higher ratios observed in the 30-minute frequency data compared to the 5-minute and 10-minute frequencies. This suggest that shorter-term fluctuations may present more arbitrage opportunities, indicative of less efficient markets, whereas longer observation periods tend to reflect a more efficient market behavior, with ratios approaching or surpassing the threshold of 1. This trend aligns with the theoretical expectation that higher-frequency data is likely to display more market inefficiencies due to the increased noise and transaction costs associated with short-term trading.

The study on market efficiency, across varying epochs and cryptocurrencies, reveals that the Hurst index consistently indicates mean-reverting behavior with values below 0.5, pointing to inherent market inefficiencies at all observed frequencies. The Shannon entropy ratios suggest that efficiency improves at lower frequencies, with longer observation periods reflecting more efficient market behavior. The Ljung-Box test corroborate these findings, with high p-values indicating a lack of autocorrelation and thus efficiency, which is more pronounced at lower data frequencies. Dickey-Fuller test results further support this, with higher p-values in longer frequency data during mean-reverting epochs, hinting at a stronger tendency towards market efficiency. Together, these metrics provide robust evidence that market efficiency varies with frequency and that the intrinsic dynamics of cryptocurrencies exhibit complexity across different time scales.

3.3.9 Microstructure Errors Analysis

The analysis of microstructure errors is crucial, particularly when dealing with high-frequency data. Microstructure noise can significantly affect observed prices, causing deviations from theoretical efficient prices due to factors such as bid-ask spreads, order book dynamics, and transaction costs.

To examine this, we conduct a comparative analysis of observed prices (P) and efficient prices (X) estimated using the uncertainty zones methodology. The Kolmogorov-Smirnov test results, presented in Table 3.1, show no significant differences between the distributions of P and X across all studied cryptocurrencies. This finding suggests that, at high frequency, the observed prices closely approximate the efficient prices, likely due to the rapid adjustment of prices to new information, which minimizes the impact of microstructure noise over short intervals.

Nevertheless, it is essential to acknowledge that microstructure noise may still subtly influence our results. High-frequency data inherently contains noise that can affect the accuracy of trend detection and other metrics. While our current methodology does not explicitly quantify the magnitude of microstructure noise, the close match between P and X in high-frequency settings implicitly indicates that the noise is relatively low during the observed period.

Future research could benefit from more detailed analyses using advanced techniques specifically designed to measure microstructure noise, such as those proposed by Aït-Sahalia and Yu (2009). These methods could offer a more granular understanding of the impact of microstructure errors on observed prices and enhance the robustness of efficiency analyses in high-frequency markets.

In conclusion, while our findings suggest that microstructure noise has a limited impact on our results, recognizing and addressing this factor is important. We recommend further investigation using specialized methodologies to quantify and account for microstructure noise in high-frequency data analysis.

3.4 Conclusions and Discussion

In this descriptive study, we undertake a detailed analysis of market efficiency within cryptocurrency markets, examining periods of upward trends, downward trends, and mean reversion. We synthesize findings from a suite of statistical methods, revealing a complex picture of market behavior.

We begin by establishing a baseline of efficient price processes, where we find observed prices to be a reliable reflection of market dynamics across various cryptocurrencies. This foundational step is crucial for confirming the accuracy of the diffusion model and the sufficiency of using observed prices for robust efficiency analysis.

A critical aspect of our study involves accurately identifying trend periods. We determine that a window size of 24 to 60 data points is optimal for trend categorization. This window size effectively distinguishes between upward, downward, and mean reversion trends, which is confirmed by our Monte Carlo simulation. In this simulation, we execute 1000 iterations, examining window sizes from 5 to 150 in 20-unit increments. The simulation results validate our choice, showing these window sizes capture the essence of market trends with high precision. This process allows us to quantitatively evaluate and confirm the effectiveness of the selected window sizes in identifying trends within financial time series.

The application of the P-th variation methodology and the validation through

fractional Brownian motion simulations consistently produce Hurst values below 0.5. This finding indicates pervasive market roughness and inefficiency, with the mean reversion periods displaying an even greater degree of anti-persistent behavior, suggesting quicker price corrections during such phases.

The application of Shannon Entropy in calculating efficiency ratios offers a detailed perspective on market behavior. For instance, we observe that BTCBUSD and ETHBUSD sometimes reveal efficiency ratios just below 1, suggesting the presence of minor inefficiencies. In contrast, assets like BNBBUSD and ADABUSD exhibit notably lower ratios during periods of trends and mean reversion, reflecting more substantial inefficiencies. This difference underscores the varying degrees of market efficiency captured by Shannon Entropy, where ratios closer to 1 indicate a tendency towards efficiency, and those significantly below 1 point to a greater departure from efficient behavior.

When we assess autocorrelation through the Ljung-Box test, the high p-values suggest an efficient market in line with the Weak Efficient Market Hypothesis. Nevertheless, the variability and outliers observed in the data indicate that the level of market efficiency varies among different cryptocurrencies. However, we detect no significant inefficiencies. The absence of significant inefficiencies, as detected by the Ljung-Box test, does not undermine its suitability; rather, it suggests that the markets examined may truly exhibit efficient dynamics as per the Weak Efficient Market Hypothesis from the point of view of self-correlations.

The Augmented Dickey-Fuller test results indicate stronger evidence of stationarity and non-random price movements during pronounced trends, challenging the hypothesis of market efficiency. Conversely, the mean reversion periods tend to show higher p-values, aligning more closely with efficient market behavior.

Our results from the joint binomial distribution analysis, which incorporates the Hurst and Shannon efficiency indices for various cryptocurrencies, suggest a market that is complex and variable in its efficiency. The Ljung-Box test indicates an absence of significant autocorrelations, pointing towards consistent market efficiency across epochs, while the Augmented Dickey-Fuller test reveals prevalent inefficiencies, except in mean reversion epochs. The latter, however, border on inefficiency for nearly all cryptocurrencies when the significance level is adjusted. These findings, illustrated through heatmaps, underpin our decision to focus on the Hurst and Shannon indices, which collectively provide a more stable indication of market efficiency despite the nuanced and sometimes contradictory signals from other statistical tests.

Our frequency-dependent efficiency analysis reveals that cryptocurrency markets show increased efficiency at lower data frequencies. Through the application of the Hurst index, Shannon entropy, Ljung-Box, and Dickey-Fuller tests, we observe a consistent trend towards more efficient behavior in longer observation periods, except when using the Hurst index. This pattern highlights the complexity of market dynamics and suggests a nuanced understanding of efficiency across various time scales.

We highlight that market efficiency in cryptocurrencies is a multifaceted phenomenon, varying with market trends and data frequencies. We observe that

the cryptocurrency markets do not uniformly conform to the principles of market efficiency. While certain periods and frequencies show signs of an efficient market, others reveal potential for predictability and inefficiency. This dynamic nature of market efficiency, along with the observed heterogeneity across different cryptocurrencies and market conditions, underscores the intricate behavior of the cryptocurrency market.

Investors and market participants would do well to consider these findings, recognizing that the efficiency of cryptocurrency markets is subject to temporal and trend-specific influences. As such, market strategies should be tailored to the nuances of the market state and the frequency of data analysis to harness the potential opportunities that arise from these inefficiencies. While this study does not aim to predict market movements, it describes conditions under which inefficiencies in the cryptocurrency market are identified. This sets a favorable stage for the introduction of predictive methodologies designed to exploit these inefficiencies. By understanding the contexts where the market deviates from efficiency, investors and analysts can develop strategies that capitalize on these moments, thereby potentially enhancing returns or mitigating risks in this volatile market environment.

Bibliography

Agram, N., Hilbert, A., and Øksendal, B. (2019). Singular control of spdes with space-mean dynamics.

Aït-Sahalia, Y. and Yu, J. (2009). High frequency market microstructure noise estimates and liquidity measures. *The Annals of Applied Statistics*, 3(1):422–457.

Alfonsi, A. F. A. and Schied, A. (2010). Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance, Taylor*.

Almgren, R. and Chriss, N. (2000). Optimal execution of portfolio transactions. *J. Risk*, 3:5–39.

Almgren, R., Thum, C., Hauptmann, E., and Li, H. (2005). Direct estimation of equity market impact.

Archankul, A., Ferrari, G., Hellmann, T., and Thijssen, J. J. J. (2023). Singular control in a cash management model with ambiguity.

Ata, B., Harrison, J. M., and Si, N. (2023). Singular control of (reflected) brownian motion: A computational method suitable for queueing applications.

Barbon, A. and Ranaldo, A. (2022). On The Quality Of Cryptocurrency Markets: Centralized Versus Decentralized Exchanges. (22-38).

Barger, W. and Lorig, M. (2018). Optimal liquidation under stochastic price impact. *International Journal of Theoretical and Applied Finance*.

Barles, G. and Souganidis, P. E. (1991). Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Analysis*, 4(3):271–283.

Barunik, J. and Dvorakova, S. (2015). An empirical model of fractionally cointegrated daily high and low stock market prices. *Economic Modelling*, 45:193–206.

Bellman, R. and Dreyfus, S. (1962). Applied dynamic programming. *A report kprepared for United States Air Force project RAND*.

Bershova, N. and Rakhlin, D. (2013). The non-linear market impact of large trades: Evidence from buy-side order flow. *Quantitative Finance*, 13(11):1759–1778.

Bertsimas, D. and Lo, A. W. (1998). Optimal control of execution costs. *Journal of Financial Markets*, 1:1–50.

Blanco Encinosa, L. J. (2021). Criptomonedas. Breve análisis desde la perspectiva económica y financiera. *Cofin Habana*, 15.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, 69(3):542–547.

Brunovský, P., Černý, A., and Komadel, J. (2018). Optimal trade execution under endogenous pressure to liquidate: Theory and numerical solutions. *European Journal of Operational Research*, 264(3):1159–1171.

Budhiraja, A. and Ross, K. (2006). Existence of optimal controls for singular control problems with state constraints. *The Annals of Applied Probability*, 16(4):Pages.

Cartea, A. and Jaimungal, S. (2016a). A closed-form execution strategy to target. *SIAM Journal on Financial Mathematics*, 7(1):760–785.

Cartea, A. and Jaimungal, S. (2016b). Incorporating order-flow into optimal execution. *Mathematics and Financial Economics*, 10(3):339–364.

Cartea, A., Jaimungal, S., and Penalva, J. (2015). *Algorithmic and high-frequency trading*. Cambridge University Press.

Chu, J., Zhang, Y., and Chan, S. (2019). The adaptive market hypothesis in the high frequency cryptocurrency market. *International Review of Financial Analysis*, 64:221–231.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236.

Cont, R. and Das, P. (2024). Rough volatility: Fact or artefact? *Sankhya B: The Indian Journal of Statistics*.

Cox, R., Alimov, A., Chakraborty, D., and Jain, A. (2004). The random walk hypothesis on the bombay stock exchange. *Finance India*, 18:1251–1258.

Curato, G., Gatheral, J., and Lillo, F. (2017). Optimal execution with nonlinear transient market impact. *Quantitative Finance*, 17(1):41–54.

Davies, R. B. and Harte, D. S. (1987). Tests for hurst effect. *Biometrika*, 74(1):95–101.

Duarte, J. and Mascareñas Perez-Iñigo, J. (2013). The efficiency of stock markets: a review. *Análisis financiero*, (122):21–35.

Duffy, D. J. (2006). *Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach*. John Wiley.

Evans, M. D. D. and Lewis, K. K. (1992). Trends in expected returns in currency and bond markets. *NBER- National Bureau of Economics Research*.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383–417.

Fang, F., Ventre, C., Basios, M., Kanthan, L., Martinez-Rego, D., Wu, F., and Li, L. (2022). Cryptocurrency trading: a comprehensive survey. *Financial Innovation*, 8(13).

Gatheral, J. (2010). No-dynamic-arbitrage and market impact. *Quantitative Finance*, 10(7):749–759.

Griffith, T. and Clancey-Shang, D. (2023). Cryptocurrency regulation and market quality. *Journal of International Financial Markets, Institutions and Money*, 84:101744.

Gueant, O. (2014). Permanent market impact can be nonlinear. Preprint.

Harvey, A. C. and Collier, P. (1977). Testing for functional misspecification in regression analysis. *Journal of Econometrics*, 6(1):103–119.

Huang, Y., Forsyth, P. A., and Labahn, G. (2012). Iterative methods for the solution of a singular control formulation of a gmwb pricing problem. *Numerische Mathematik*, Volume(Number):Pages.

Huberman, G. and Stanzl, W. (2004). Price manipulation and quasi-arbitrage. *Econometrica*, 72(4):1247–1275.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1):65–91.

Kakinaka, S. and Umeno, K. (2022). Cryptocurrency market efficiency in short- and long-term horizons during covid-19: An asymmetric multifractal analysis approach. *Finance Research Letters*, 46:102319.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335.

Lillo, F., Farmer, J. D., and Mantegna, R. N. (2003). Master curve for price-impact function. *Nature*, 421(1):129–130.

Lim, K.-P. and Brooks, R. (2011). The evolution of stock market efficiency over time: A survey of the empirical literature. *Journal of Economic Surveys*, 25(1):69–108.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.

Lo, A. W. and MacKinlay, A. C. (1999). *A Non-Random Walk Down Wall Street*. Princeton University Press, Princeton, NJ.

McAleer, M. and Medeiros, M. C. (2008). Realized volatility: A review. *Econometric Reviews*, 27(1-3):10–45.

Mitra, S. and Bawa, J. (2017). Can trade opportunities and returns be generated in a trend persistent series? evidence from global indices. *Physica A: Statistical Mechanics and its Applications*, 469:124–135.

Nadarajah, S. and Chu, J. (2017). On the inefficiency of bitcoin. *Economics Letters*, 150:6–9.

Noda, A. (2020). On the evolution of cryptocurrency market efficiency. *Applied Economics Letters*, pages 1–7.

OpenAI (2024). Chatgpt: A large language model trained by openai. `https://www.openai.com/chatgpt`. Accessed: 2024-06-17.

Peaceman, D. and Ratchford Jr, H. H. (1955). The numerical solution of parabolic and elliptic differential equations. *Journal of the Society for Industrial and Applied Mathematics*, 3(1):28–41.

Pham, H. (2009). *Continuous-time Stochastic Control and Optimization with Financial Applications*. Springer.

Piñeiro-Chousa, J., Ángeles López-Cabarcos, M., Sevic, A., and González-López, I. (2022). A preliminary assessment of the performance of defi cryptocurrencies in relation to other financial assets, volatility, and user-generated content. *Technological Forecasting and Social Change*, 181:121740.

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society: Series B (Methodological)*, 31(2):350–371.

Reppen, A. M. A. (2018). Singular control in financial economics.

Robert, C. and Rosenbaum, M. (2012). Volatility and covariation estimation when microstructure noise and trading times are endogenous. *Mathematical Finance*, 22:133 – 164.

Scaillet, O., Treccani, A., and Trevisan, C. (2017). High-frequency jump analysis of the bitcoin market. *Swiss Finance Institute Research Paper Series*, (17-19).

Sensoy, A. (2019). The inefficiency of bitcoin revisited: A high-frequency analysis with alternative currencies. *Finance Research Letters*, 28:68–73.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423.

Shternshis, A., Mazzarisi, P., and Marmi, S. (2022). Efficiency of the moscow stock exchange before 2022. *Entropy*, 24(1184).

Subramanian, A. (2008). Optimal liquidation by a large investor. *SIAM Journal of Applied Mathematics*, 68:1168–1201.

Sánchez, J. and Ramírez, H. (2024). Cryptomarket analysis: Optimal liquidation and market efficiency study. Preprint.

Tavella, D. and Randall, C. (2000). *Pricing Financial Instruments: The Finite Difference Method*. John Wiley.

Thomas, J. (1995). *Numerical Partial Differential Equations: Finite Difference Methods*. Springer New York, NY.

Toth, B., Eisler, Z., and Bouchaud, J. P. (2016). The square-root impact law also holds for option markets. *Wilmott*, (85):70–73.

Urquhart, A. and Lucey, B. (2022). Crypto and digital currencies — nine research priorities. *Nature*, 604:36–39.

Watorek, M., Drozdz, S., Kwapien, J., Minati, L., Oswiecimka, P., and Stanuszek, M. (2020). Multiscale characteristics of the emerging global cryptocurrency market. *Physics Reports*, 901.

Wei, W. C. (2018). Liquidity and market efficiency in cryptocurrencies. *Economics Letters*, 168:21–24.

Zhang, Y., Chan, S., Chu, J., and Sulieman, H. (2020a). On the market efficiency and liquidity of high-frequency cryptocurrencies in a bull and bear market. *Journal of Risk and Financial Management*, 13:8.

Zhang, Y., Chan, S., Chu, J., and Sulieman, H. (2020b). On the market efficiency and liquidity of high-frequency cryptocurrencies in a bull and bear market. *Journal of Risk and Financial Management*, 13(8).

Zhao, L. (2021). The function and impact of cryptocurrency and data technology in the context of financial technology: introduction to the issue. *Financial Innovation*, 7(84).

.1 Appendix A: Comparison squared residuals and R^2

Impact	Model	Total Sum	Mean	Std Dev	R^2	Adj R^2
Under TPI	Linear	0.0000573	0.0000012	0.0000011	0.990937	0.990744
Under TPI	Power	0.0000164	0.0000003	0.0000006	0.997413	0.997300
Under PPI	Linear	0.0001765	0.0000036	0.0000028	0.980561	0.980147
Under PPI	Power	0.0000174	0.0000004	0.0000004	0.998087	0.998004
Over TPI	Linear	2.4067705	0.0491178	0.0737979	0.980709	0.980299
Over TPI	Power	1.4850594	0.0303073	0.0374185	0.988097	0.987579
Over PPI	Linear	5.6089534	0.1144684	0.1358861	0.955756	0.954814
Over PPI	Power	3.8265454	0.0780928	0.0730643	0.969816	0.968503
Average TPI	Linear	0.0010414	0.0000213	0.0000237	0.999804	0.999799
Average TPI	Power	0.0010392	0.0000212	0.0000230	0.999804	0.999795
Average PPI	Linear	0.0096906	0.0001978	0.0001781	0.998061	0.998020
Average PPI	Power	0.0056340	0.0001150	0.0001242	0.998873	0.998824

Table 11: Comparison of linear and power model fit through squared residuals and R^2 for under, over and average estimated TPI and PPI

.2 Appendix B: Calibrated coefficients

Since the differential equation (2.5) incorporated the direction of the negotiations, i.e., liquidation, through the negative sign for both TPI and PPI, these coefficients must be considered through their additive inverse, except those representing exponents, i.e., r_2 and p_2.

Coefficients obtained for each calibrated model				
Scenario	Model	Coefficient	Value	Confidence Interval (95%)
Under TPI	Linear	a_1	-0.000800	[-0.000822, -0.000777]
		a_2	-0.000708	[-0.001353, -0.000064]
Under TPI	Power	r_1	-0.000252	[-0.000308, -0.000197]
		r_2	1.288	[1.232, 1.343]
		r_3	-0.003766	[-0.004333, -0.003198]
Under PPI	Linear	b_1	-0.000953	[-0.001009, -0.000914]
		b_2	0.002240	[0.001100, 0.003371]
Under PPI	Power	p_1	-0.000141	[-0.000170, -0.000112]
		p_2	1.480	[1.428, 1.533]
		p_3	-0.003039	[-0.003553, -0.002525]
Over TPI	Linear	a_1	-0.000798	[-0.000831, -0.000765]
		a_2	-0.222515	[-0.350646, -0.094384]
Over TPI	Power	r_1	-0.005383	[-0.006044, -0.004722]
		r_2	0.789	[0.712, 0.866]
		r_3	0.239	[0.006044, 0.471593]
Over PPI	Linear	b_1	-0.000794	[-0.000844, -0.000744]
		b_2	-0.259673	[-0.455277, -0.064069]
Over PPI	Power	p_1	-0.009462	[-0.019450, 0.000526]
		p_2	0.727	[0.612, 0.842]
		p_3	0.386	[-0.013370, 0.784653]
Average TPI	Linear	a_1	-0.000969	[-0.000973, -0.000965]
		a_2	0.005738	[0.003070, 0.008406]
Average TPI	Power	r_1	-0.000982	[-0.001007, -0.000957]
		r_2	0.998	[0.986, 1.010]
		r_3	0.006410	[0.001302, 0.011518]
Average PPI	Linear	b_1	-0.000940	[-0.000952, -0.000928]
		b_2	-0.008097	[-0.016236, 0.000042]
Average PPI	Power	p_1	-0.001647	[-0.001969, -0.001325]
		p_2	0.923	[0.896, 0.949]
		p_3	0.021356	[0.008623, 0.034088]

Table 12: Calibrated coefficients

.3 Appendix C: Average p-values from the Dickey-Fuller test

Table 13: Average p-values from the Dickey-Fuller test for different epochs of returns for each cryptocurrency with 5-minute frequency.

Cryptocurrency	Mean reversion	Downward trend	Upward trend
BNBBUSD	0.0583	0.0047	0.0085
BTCBUSD	0.0843	0.0232	0.0094
ETHBUSD	0.0632	0.0136	0.0083
ADABUSD	0.0879	0.0118	0.0076
XRPBUSD	0.0806	0.0224	0.0119
SOLBUSD	0.1066	0.0253	0.0189
DOGEBUSD	0.1053	0.0226	0.0165

Table 14: Average p-values from the Dickey-Fuller test for different epochs of returns for each cryptocurrency with 10-minute frequency.

Cryptocurrency	Mean reversion	Downward trend	Upward trend
BNBBUSD	0.0859	0.0146	0.0097
BTCBUSD	0.1160	0.0133	0.0191
ETHBUSD	0.0963	0.0150	0.0103
ADABUSD	0.1050	0.0142	0.0363
XRPBUSD	0.0773	0.0120	0.0039
SOLBUSD	0.1470	0.0121	0.0063
DOGEBUSD	0.0522	0.0197	0.0168

Table 15: Average p-values from the Dickey-Fuller test for different epochs of returns for each cryptocurrency with 30-minute frequency.

Cryptocurrency	Mean reversion	Downward trend	Upward trend
BNBBUSD	0.0812	0.0126	0.0305
BTCBUSD	0.1243	0.0073	0.0433
ETHBUSD	0.1226	0.0029	0.0199
ADABUSD	0.0744	0.0122	0.0203
XRPBUSD	0.1206	0.0177	0.0189
SOLBUSD	0.0716	0.0185	0.0432
DOGEBUSD	0.0971	0.0367	0.0267

.4 Appendix D: Average Estimated H Values for Price Series

Table 16: Average Estimated H Values for Price Series with 5-minute frequency

Asset	Downward Trend	Mean Reversion	Upward Trend
BNBBUSD	0.4085	0.4123	0.4239
BTCBUSD	0.4050	0.3894	0.4032
ETHBUSD	0.4039	0.3995	0.4144
ADABUSD	0.4219	0.4092	0.4374
XRPBUSD	0.4271	0.4258	0.4371
SOLBUSD	0.4217	0.4047	0.4204
DOGEBUSD	0.4238	0.4175	0.4334

Table 17: Average Estimated H Values for Price Series with 10-minute frequency

Asset	Downward Trend	Mean Reversion	Upward Trend
BNBBUSD	0.4215	0.3959	0.4220
BTCBUSD	0.3911	0.3820	0.3992
ETHBUSD	0.3908	0.3918	0.4128
ADABUSD	0.4250	0.4048	0.4259
XRPBUSD	0.4205	0.4130	0.4203
SOLBUSD	0.4149	0.3932	0.4298
DOGEBUSD	0.4190	0.4010	0.4301

Table 18: Average Estimated H Values for Price Series with 30-minute frequency

Asset	Downward Trend	Mean Reversion	Upward Trend
BNBBUSD	0.4311	0.3914	0.4158
BTCBUSD	0.3917	0.3636	0.3899
ETHBUSD	0.3829	0.3753	0.4067
ADABUSD	0.4049	0.4190	0.4300
XRPBUSD	0.3999	0.3814	0.4186
SOLBUSD	0.3982	0.4026	0.4290
DOGEBUSD	0.4011	0.3839	0.4167

.5 Appendix E: Average Shannon Efficiency Ratios for Return Series

Table 19: Average Shannon Efficiency Ratios for 5-Minute Frequency Cryptocurrency Return Series

Asset	Mean Reversion	Downward Trend	Upward Trend
BNBBUSD	0.6868 (± 0.1382)	0.7325 (± 0.1542)	0.7278 (± 0.1490)
BTCBUSD	0.9695 (± 0.1617)	1.0114 (± 0.0700)	1.0083 (± 0.0710)
ETHBUSD	0.9963 (± 0.1576)	1.0148 (± 0.0714)	1.0210 (± 0.0765)
ADABUSD	0.7782 (± 0.1468)	0.8251 (± 0.1327)	0.8165 (± 0.1405)
XRPBUSD	0.8150 (± 0.1509)	0.8497 (± 0.1326)	0.8538 (± 0.1248)
SOLBUSD	0.7764 (± 0.1646)	0.8325 (± 0.1476)	0.8202 (± 0.1556)
DOGEBUSD	0.8805 (± 0.1512)	0.9297 (± 0.1109)	0.9366 (± 0.0990)

Table 20: Average Shannon Efficiency Ratios for 10-Minute Frequency Cryptocurrency Return Series

Asset	Mean Reversion	Downward Trend	Upward Trend
BNBBUSD	0.7700 (± 0.1625)	0.8123 (± 0.1374)	0.8070 (± 0.1335)
BTCBUSD	0.9880 (± 0.1647)	1.0151 (± 0.0725)	1.0244 (± 0.0750)
ETHBUSD	0.9810 (± 0.1668)	1.0174 (± 0.0673)	1.0272 (± 0.0757)
ADABUSD	0.8559 (± 0.1559)	0.9079 (± 0.1076)	0.8779 (± 0.1321)
XRPBUSD	0.8971 (± 0.1611)	0.9038 (± 0.1134)	0.9151 (± 0.1110)
SOLBUSD	0.8399 (± 0.1714)	0.8956 (± 0.1382)	0.9006 (± 0.1384)
DOGEBUSD	0.9528 (± 0.1602)	0.9732 (± 0.0914)	0.9767 (± 0.0944)

Table 21: Average Shannon Efficiency Ratios for 30-Minute Frequency Cryptocurrency Return Series

Asset	Mean Reversion	Downward Trend	Upward Trend
BNBBUSD	0.8854 (± 0.1515)	0.9252 (± 0.1047)	0.9248 (± 0.0986)
BTCBUSD	0.9614 (± 0.1527)	1.0317 (± 0.0727)	1.0399 (± 0.0827)
ETHBUSD	0.9804 (± 0.1578)	1.0178 (± 0.0558)	1.0570 (± 0.0930)
ADABUSD	0.9084 (± 0.1594)	0.9945 (± 0.0727)	0.9784 (± 0.1014)
XRPBUSD	0.9391 (± 0.1423)	0.9762 (± 0.0766)	1.0026 (± 0.0931)
SOLBUSD	0.9155 (± 0.1510)	0.9789 (± 0.0865)	0.9814 (± 0.1221)
DOGEBUSD	0.9841 (± 0.1705)	1.0163 (± 0.0759)	1.0291 (± 0.0947)