Quadratic Hedging in a Non-causal AR(1) Model

Caleb Kwame Danquah

A Thesis

in

The Department

of

Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Science (Mathematics) at

Concordia University

Montreal, Quebec, Canada

December 2023

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Caleb Kwame Danquah

Entitled: Quadratic Hedging in a Non-causal AR(1) Model

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality. Signed by the final Examining Committee:

Signed by the final Examining Committee:

_____ Thesis Supervisor

Dr. P. Gaillardetz

_____ Thesis Supervisor

Dr. Y. Lu

_____ Examiner

Dr. X. Zhou

Approved by _____

Dr. L. Popovic (Graduate Program Director)

_____

Dr. P. Sicotte (Dean of Faculty)

_____

Date

Abstract

Quadratic Hedging in a Non-causal AR(1) Model

Caleb Kwame Danquah

This thesis explores hedging strategies for European-type derivatives under the non-causal AR(1) Cauchy model. Recently, such non-causal models have raised much attention in the finance literature due to their ability to replicate bubbles often observed in the cryptocurrency market, as well as their tractability for pricing standard European options. However, these discrete-time models are incomplete, meaning that it is impossible to perfectly replicate a derivative's payoff in such a market. This thesis explores the use of quadratic hedging approaches to manage the risk of a derivative trader.

Acknowledgments

I thank God Almighty for seeing me through this research. I would also like to express my deepest gratitude to my supervisors, Dr. Patrice Gaillardetz and Dr. Yang Lu, of the Department of Mathematics and Statistics at Concordia University. Their guidance, encouragement and directions made it possible for me to complete this work. My last thanks go to my family for their continuous support throughout my studies. This work was supported by the Faculty of Arts and Science Graduate Fellowship, and the Department of Mathematics and Statistics at Concordia University.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Cryptocurrencies like Bitcoin (BTC) have inspired a lot of research in Economics and Finance. Several papers showed that their price feature bubbles [see e.g. Garcia, Tessone, Mavrodiev, and Perony (2014), and Kristoufek (2015)], that is, their price process feature epochs of steady increase, followed by a sharp decrease. Investors who buy BTC have been concerned about how their investment can be protected against abrupt fluctuations. It is also known that traditional ARMA and GARCH-type models cannot well replicate this kind of bubble phenomenon.

Noncausal models have received much attention in the time series literature in the past 15 years [see Lanne et al. (2011), Hencic and Gouriéroux (2015), Gourieroux and Jasiak (2016), Davis and Song (2020), Gourieroux, Hencic, and Jasiak (2021)], thanks to their ability to capture bubbles. Some of their theoretical properties have been investigated by Gouriéroux and Zakoïan (2017), which showed that the noncausal Cauchy AR(1) model is a stationary martingale. This makes this model not just empirically relevant, but also compatible with the no-arbitrage condition.

Recently, Gouriéroux and Lu (2023) showed that such noncausal models have also the advantage of leading to semi-analytical formulas for the price of many derivatives. However, these discrete-time models are incomplete, meaning that it is impossible to perfectly replicate the payoff of a derivative in such a market. In replicating the payoff of a derivative, a popular approach is to conduct hedging.

Hedging is a risk management strategy that seeks to find a portfolio whose value at maturity is as close as possible to the payoff of underlying assets. Here, the closeness is measured by the mean-squared error. The literature has mainly considered two hedging strategies: local and global. Local quadratic hedging is a strategy that minimizes the mean squared error at each time interval. In contrast, global quadratic hedging minimizes the error over the entire lifespan without removing or adding funds to the portfolio (also referred to as a self-financing portfolio). To the best of our knowledge, an application of these techniques in a noncausal AR(1) model is yet to be identified. Our motivation for this thesis is to fill this gap.

1.2 Literature Review

1.2.1 Modelling and forecasting of cryptocurrencies

Diverse methodologies have been employed to better understand the complexities of cryptocurrency price dynamics. Yiying and Yeze (2019) used Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) Recurrent Neural Networks to analyze the price behaviour of major cryptocurrencies such as BTC, Ethereum, and Ripple. Hayes (2017) focused on the cost of production as a primary determinant of cryptocurrency value, analyzing factors like production rate and mining difficulty across 66 cryptocurrencies. Matic et al. (2023) also proposed jump-diffusion models.

To effectively replicate the bubble phenomena observed in cryptocurrency markets, this thesis will adopt the noncausal Cauchy model used by Gouriéroux and Zakoïan (2017). As an illustration, Figure 1.1 shows a simulated path of a non-causal Cauchy process with $\rho = 0.9$, and this is compared with Figure 1.2 displaying the historical BTC price evolution. Figure 1.1 below clearly demonstrates recurring bubble patterns, illustrating how the non-causal Cauchy model can effectively capture the kinds of trends seen in real cryptocurrency markets.

Figure 1.2 depicts a distinctly time-varying BTC market behaviour across a specific period. Placing Figure 1.2 side by side with Figure 1.1 unveils visual parallels: The BTC bullish (blue) segment replicates the bubble-like rises and falls captured in the $\rho = 0.9$ scenario of Figure 1.2. The sharp BTC price dip in March 2020 mirrors the sudden bubble collapse in the non-causal model. These visual overlaps suggest that the non-causal AR(1) model may provide a robust framework for capturing the multifaceted behaviours of the BTC market. While this visual similarity isn't definitive evidence, it presents a compelling starting point

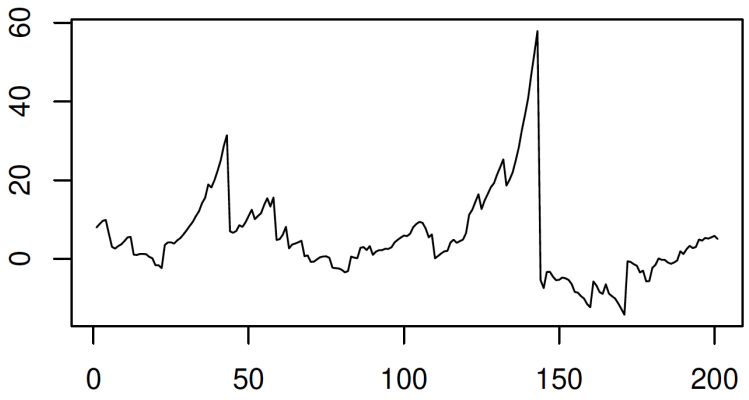

Figure 1.1: Simulated path of noncausal Cauchy process with $\rho = 0.9$. Source: Figure 3 of Gouriéroux and Zakoïan (2017)

for deeper, more rigorous investigations.

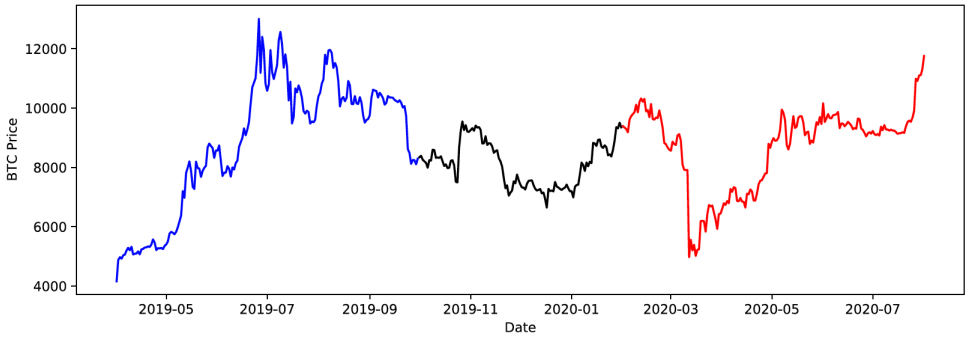

Figure 1.2: BTC Closing Price Trajectory from figure 1 of Matic et al. (2023)

1.2.2 Pricing of cryptocurrencies

Recently, notable progress has been made in the pricing of cryptocurrency literature. Example, Hou et al. (2020) highlighted the importance of jumps and cojumps in cryptocurrency markets, particularly BTC, for option pricing. They suggested using a stochastic volatility model with a correlated jump (SVCJ) to account for the market's unique behaviours, including speculation, sentiment, and the absence of typical market fundamentals. This model, however, does not allow to capture bubbles. Noncausal AR(1) models have also been used to price financial derivatives by Gouriéroux and Lu (2023). However, these papers are based on the existence of a unique pricing measure, which assumes liquid derivative and underlying asset markets. Currently, the market for cryptocurrency derivatives is highly non-liquid. Therefore, such derivatives need to be "hedged" in an incomplete market framework, using techniques such as quadratic hedging.

1.2.3 Quadratic hedging

The two main techniques of quadratic hedging considered in the literature are local and global quadratic hedging and were intensively explored by Schweizer (Schweizer (1988), Schweizer (1991), and Schweizer (1995)). Augustyniak, Godin, and Simard (2017) investigated the effectiveness of these two quadratic hedging strategies under GARCH models. They found that for long-term vanilla options (a year or longer), global quadratic hedging offers a significant risk reduction advantage, especially evident at three-year maturities. Rémillard and Rubenthaler (2013) rectified misconceptions about local optimization techniques that had been previously presented as globally optimal solutions. Their research established explicit formulas that extended the work of Černỳ and Kallsen (2007) (on the structure of general mean-variance hedging strategies) and Schweizer (1995), focusing particularly on situations where the asset value process is either Markovian or forms part of a Markov process. Using dynamic programming approximation techniques, they derived an optimal solution for the discrete-time local quadratic hedging problem, emphasizing the superiority of this optimal hedging over conventional delta hedging, especially when considering models like geometric random walks and NGARCH.

Matic et al. (2023) adopted affine jump-diffusion models and infinite activity Lévy processes to investigate hedging strategies for the BTC market. They used both Monte Carlo simulations to replicate the dynamics of the BTC market. They leveraged various hedging strategies, including the Greeks (Delta, Delta–Gamma, Delta–Vega) and the global quadratic approach. The difference between these strategies is that quadratic hedging aims to reduce the total expected hedging errors ensuring the portfolio remains self-financing. Delta hedging protects a hedged position against first-order changes in the underlying asset, whereas Delta–Gamma and Delta–Vega hedging (or multiple instrument hedges) protects the position against higher-order sensitivities in the underlying asset and volatility respectively. A small gamma value implies gradual fluctuations in delta, thereby indicating a reduced frequency of rebalancing. Whereas a high gamma value suggests that the delta is very sensitive to changes in the price of the underlying asset. In their analysis, they found that multiple-instrument hedges substantially reduced tail risk for longer-dated options.

This thesis contributes to the literature by addressing quadratic hedging of cryptocurrency derivatives, such as European options, under the noncausal AR(1) Cauchy model of Gouriéroux and Zakoïan (2017). This modelling has several advantages. First, under this model, the asset price is a martingale, hence compatible

with non-arbitrage conditions. Second, it can replicate the bubble phenomenon. We start with a simple framework of a one-period derivative. Then we extend to multiple periods.

1.3 Plan of the Thesis

The rest of the thesis is structured as follows. The non-causal AR(1) model's properties are reviewed and the groundwork is laid for simulating the paths of AR(1) processes in Chapter 2. Chapter 3 discusses quadratic hedging starting with an overview of hedging as a way to reduce risk, then gradually introduces global, and local hedging techniques, breaking down how they work and what they can be used for. We also show how these hedging strategies work in real life by looking at how they work with one-period European vanilla options. In Chapter 4, we evaluate the chosen hedging strategies in the volatile cryptocurrency markets. This chapter is mostly made up of three analytical sections: the relationship between at-the-money (ATM) local hedging strategy values, node counts, and steps; the interaction between asset maturity periods and hedging strategies; and a comparison of hedging techniques across a range of strike prices.

Chapter 2

Non-Causal Process

This chapter reviews the non-causal autoregressive model of order 1 (AR(1)) and discusses its effectiveness in capturing complex patterns such as financial bubbles. A standard (causal) AR(1) process is defined by:

$$X_t = \phi X_{t-1} + \varepsilon_t, \tag{2.1}$$

where (ε_t) represents i.i.d. sequence, and $|\phi| < 1$ such that the process (X_t) is stationary. While the AR(1) model has its merits and is widely applied in various fields, it may not always sufficiently capture the intricacies of financial time series. For instance, phenomena like bubbles or sudden market crashes, often observed in cryptocurrency markets like Bitcoin, can elude traditional AR structures. This prompts the exploration of more advanced models, specifically the non-causal AR(1) Cauchy model. Before we talk about the Non-Causal Cauchy model, we first review the Cauchy distribution.

2.1 The Cauchy Distribution

The Cauchy distribution is a continuous probability distribution on $\mathbb{R}$. Its probability density function (pdf) as given by Feller (1991) is:

$$f_X(x) = \frac{1}{\pi\gamma\left[1 + \left(\frac{x-x_0}{\gamma}\right)^2\right]}, \quad \forall x \in \mathbb{R}, \tag{2.2}$$

where the scale parameter $\gamma > 0$ and x_0 is the location parameter (indicating the peak of the distribution). The distribution is also symmetric about the axis $x = x_0$.

The cumulative distribution function (cdf) is given by:

$$F_X(x) = \frac{1}{\pi}\tan^{-1}\left(\frac{x - x_0}{\gamma}\right), \tag{2.3}$$

and the characteristic function is given by

$$\varphi(u) = \mathbb{E}[e^{iXu}] = \int_{-\infty}^{\infty} e^{-ixu} f_X(x; x_0, \gamma)dx = e^{ix_0u - \gamma|u|}. \tag{2.4}$$

The Cauchy distribution has an undefined mean and variance (Johnson, Kotz, and Balakrishnan (1995)). It also belongs to the larger family of α-stable distributions. That is, for n independent Cauchy distributions, $X_1, X_2, \ldots, X_n$ with scale parameters $\gamma_1, \gamma_2, \ldots, \gamma_n$ and location parameters $x_1, x_2, \ldots, x_n$, the sum $\sum_{i=1}^{n} c_i X_i$, where $c_1, c_2, \ldots, c_n$ are real numbers is also Cauchy with scale $\sum_{i=1}^{n} |c_i|\gamma_i$ and location $\sum_{i=1}^{n} c_i x_i$. See appendix A for more information on α-stable distributions.

2.1.1 Graphical representation

Figure 2.1 plots the pdfs of the Cauchy distributions with different scale parameters: $\gamma = 1$, 2, and 5 and the same location parameter $x_0 = 0$.

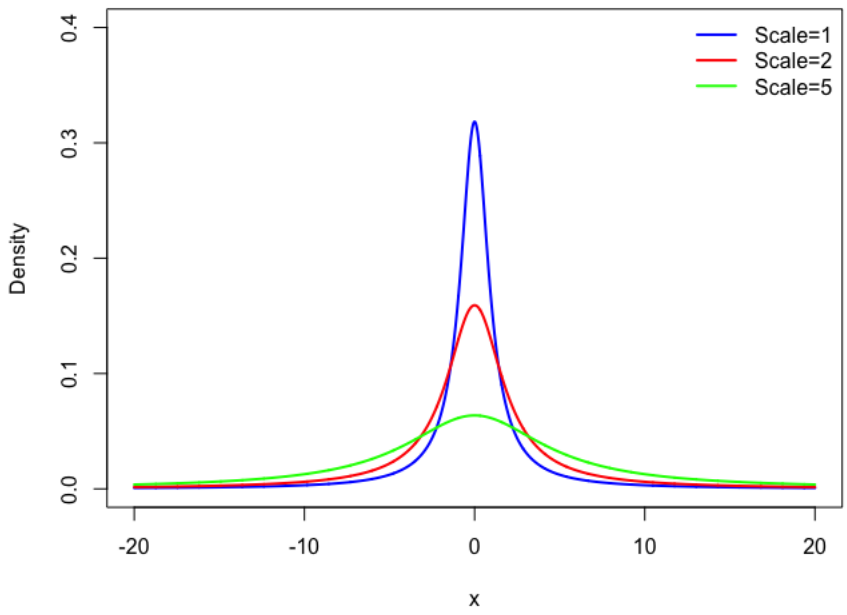

Figure 2.1: Pdfs of the Cauchy distribution for fixed shape and different scale parameters.

2.2 Non-Causal Cauchy Model

The non-causal AR(1) model is obtained by time-reversing the standard, causal AR(1) model.

$$X_t = \rho X_{t+1} + \varepsilon_t, \tag{2.5}$$

where $|\rho| < 1$ and (ε_t) is an i.i.d. sequence.

It is well known that if (ε_t) follows the Gaussian distribution, this process has also a standard causal AR(1) representation [see Breidt et al. (1992) and Hamilton (2020)]. However, when (ε_t) is non-Gaussian, the resulting noncausal process can no longer be written as a standard causal Autoregressive Moving Average (ARMA) process.

In particular, Gouriéroux and Zakoïan (2017) studied the case where (ε_t) is α-stable, such as Cauchy or Lévy. The Cauchy case is particularly interesting for financial applications since we will see later on that this leads to a martingale model, which is compatible with the no-arbitrage condition. Moreover, in this case, the conditional distribution has closed-form density, making the model particularly tractable for prediction and hedging purposes. We will consider the case where (ε_t) is Cauchy, with location parameter 0 and scale parameter σ, denoted $\mathcal{C}(0, \sigma)$.

Proposition 1. *The stationary distribution of the noncausal Cauchy linear AR(1) process is Cauchy* $\mathcal{C}\left(0, \frac{\sigma}{1-|\rho|}\right)$.

Proof. The strictly stationary solution of (2.5) defined above is:

$$X_t = \sum_{h=0}^{\infty} \rho^h \varepsilon_{t+h}.$$

When $\alpha = 1$, the characteristic function of X_t is given by

$$\begin{aligned}
\mathbb{E}[e^{iuX_t}] &= \prod_{h=0}^{\infty} \mathbb{E}[e^{iu\rho^h \varepsilon_{t+h}}] \\
&= \exp\left\{ -\sum_{h=0}^{\infty} \sigma |u| |\rho|^h \right\} \\
&= \exp\left\{ \frac{-\sigma|u|}{1-|\rho|} \right\},
\end{aligned} \tag{2.6}$$

which corresponds to a Cauchy distribution with location parameter 0 and scale parameter $\frac{\sigma}{1-|\rho|}$. Note

that if (ε_t) follows other non-Cauchy α-stable distribution, the marginal distribution is still α-stable, see Gouriéroux and Zakoïan (2017). □

From now on, we will consider the case where ρ is real and between 0 and 1, which is the most relevant case for financial modelling. For instance, in an application to NASDAQ price data, Gouriéroux and Zakoïan (2017) found that $\rho = 0.978$. The next proposition gives the conditional distribution of the noncausal process in the causal direction.

Proposition 2. *The h-step causal transition pdf of the noncausal cauchy linear AR(1) process as proposed by Gouriéroux and Zakoïan (2017) is given by:*

$$f_{X_t}(x_t|X_{t-h} = x_{t-h}) = \frac{1}{\pi\sigma_h}\frac{1}{1+(x_{t-h}-\rho^h x_t)^2/\sigma_h^2}\frac{\sigma^2+(1-\rho)^2x_{t-h}^2}{\sigma^2+(1-\rho)^2x_t^2}, \tag{2.7}$$

where $\sigma_h = \sigma\frac{1-\rho^h}{1-\rho}$.

In particular, if $h = 1$, we get

$$f_{X_t}(x_t|X_{t-1} = x_{t-1}) = \frac{1}{\pi\sigma}\frac{1}{1+(x_{t-1}-\rho x_t)^2/\sigma^2}\frac{\sigma^2+(1-\rho)^2x_{t-1}^2}{\sigma^2+(1-\rho)^2x_t^2}. \tag{2.8}$$

Recall that $\frac{1}{\pi\sigma}\frac{1}{1+(x_{t-1}-\rho x_t)^2/\sigma^2}$ is Cauchy pdf, but the additional multiplier term $\frac{\sigma^2+(1-\rho)^2x_{t-1}^2}{\sigma^2+(1-\rho)^2x_t^2}$ makes $f_{X_t}(x_t|X_{t-1} = x_{t-1})$ non Cauchy.

Proof. At horizon $h+1$, the forward recursive equation is given by,

$$\begin{aligned} X_{t-1} &= \rho^{h+1}X_{t+h} + (\varepsilon_{t-1} + \rho\varepsilon_t + \cdots + \rho^h\varepsilon_{t+h-1}) \\ &= \rho^{h+1}X_{t+h} + \varepsilon_{t-1,h}. \end{aligned} \tag{2.9}$$

The distribution of $(\varepsilon_{t-1,h})$ is Cauchy with location parameter 0 and scale parameter σ_h. We denote its pdf by $f_{\varepsilon,h}$ and the pdf of X_{t-1} given X_{t+h} by $f_{\varepsilon,h}(x_{t-1} - \rho^{h+1}x_{t+h})$. Recall from Bayes' theorem that, if X and Y are two continuous random variables, the conditional density is defined by,

$$f_X(x|Y = y) = \frac{f_Y(y|X = x)f_X(x)}{f_Y(y)}. \tag{2.10}$$

Applying Bayes' formula above, we write the conditional pdf of X_{t+h} given X_{t-1} as:

$$f_{X_{t+h}}(x_{t+h}|X_{t-1}=x_{t-1}) = \frac{f_{\varepsilon,h}(x_{t-1}-\rho^{h+1}x_{t+h})f_{X_{t+h}}(x_{t+h})}{f_{X_{t-1}}(x_{t-1})}. \tag{2.11}$$

Therefore, similarly, we can write the pdf of X_t given X_{t-h} as:

$$\begin{aligned} f_{X_t}(x_t|X_{t-h}=x_{t-h}) &= \frac{f_{\varepsilon,h-1}(x_{t-h}-\rho^h x_t)f_{X_t}(x_t)}{f_{X_{t-h}}(x_{t-h})} \\ &= \frac{1}{\pi\sigma_h}\frac{1}{1+(x_{t-h}-\rho^h x_t)^2/\sigma_h^2}\frac{\sigma^2+(1-\rho)^2x_{t-h}^2}{\sigma^2+(1-\rho)^2x_t^2}, \end{aligned} \tag{2.12}$$

after substituting the various distributions, where $\sigma_h = \sigma\frac{1-\rho^h}{1-\rho}$. Note that, the conditional distribution $f_{\varepsilon,h-1}$ corresponds to the distribution of $\sum_{i=0}^{h-1}\rho^i\varepsilon_{t-i}$, which is $\mathcal{C}(0,\sigma_h)$, and $X_t \sim \mathcal{C}\left(0,\frac{\sigma}{1-\rho}\right)$ as established in Proposition 1. □

The conditional CDF of $X_t|X_{t-1}$ can be obtained by calculating the primitive of the conditional pdf (2.8). Gouriéroux and Zakoïan (2017) show that when $\sigma = 1$,

$$\begin{aligned} F_{X_t}(x_t|X_{t-1}=x_{t-1}) &= \frac{\alpha(x_{t-1},\rho)}{\pi}\log\left\{\frac{1+(1-\rho)^2x_t^2}{1+(x_{t-1}-\rho x_t)^2}\frac{\rho^2}{(1-\rho)^2}\right\} \\ &+ \frac{\beta(x_{t-1},\rho)}{\pi}\left\{\frac{\pi}{2}-\tan^{-1}(x_{t-1}-\rho x_t)\right\} \\ &+ \frac{1-\beta(x_{t-1},\rho)}{\pi}\left\{\tan^{-1}[(1-\rho)x_t]+\frac{\pi}{2}\right\}, \end{aligned} \tag{2.13}$$

where,

$$\begin{aligned} \alpha(x_{t-1},\rho) &= \frac{\rho(1-\rho)^2x_{t-1}}{(1-2\rho)^2+(1-\rho)^2x_{t-1}^2}, \\ \beta(x_{t-1},\rho) &= \frac{\rho\{(1-\rho)^2x_{t-1}^2-(1-2\rho)\}}{(1-2\rho)^2+(1-\rho)^2x_{t-1}^2}. \end{aligned}$$

Remark 1. *For a noncausal Cauchy process satisfying $X_t = \rho X_{t+1}+\varepsilon_t$ where the Cauchy sequence (ε_t) has scale parameter σ, then $\frac{X_t}{\sigma} = \rho\frac{X_{t+1}}{\sigma}+\frac{\varepsilon_t}{\sigma}$. That is, $Y_t = \frac{X_t}{\sigma}$ is noncausal Cauchy such that the error term $\varepsilon_t := \frac{\varepsilon_t}{\sigma}$ is Cauchy with unit scale parameter.*

Thus, by change of variable, we also get the conditional CDF in the case where σ is an arbitrary positive

number:

$$F_{Y_t}(y_t|Y_{t-1}=y_{t-1}) = \frac{\alpha(y_{t-1},\rho)}{\pi}\log\left\{\frac{1+(1-\rho)^2(\sigma y_t)^2}{1+\sigma^2(y_{t-1}-\rho y_t)^2}\frac{\rho^2}{(1-\rho)^2}\right\} + \frac{\beta(y_{t-1},\rho)}{\pi}\left\{\frac{\pi}{2}-\tan^{-1}[\sigma(y_{t-1}-\rho y_t)]\right\} + \frac{1-\beta(y_{t-1},\rho)}{\pi}\left\{\tan^{-1}[\sigma(1-\rho)y_t]+\frac{\pi}{2}\right\}, \tag{2.14}$$

where,

$$\alpha(y_{t-1},\rho) = \frac{\sigma\rho(1-\rho)^2 y_{t-1}}{(1-2\rho)^2+(1-\rho)^2(\sigma y_{t-1})^2},$$
$$\beta(y_{t-1},\rho) = \frac{\rho\{(1-\rho)^2(\sigma y_{t-1})^2-(1-2\rho)\}}{(1-2\rho)^2+(1-\rho)^2(\sigma y_{t-1})^2}.$$

Gouriéroux and Zakoïan (2017) also derived (see their proposition 3.5) the first and second-order moment for the non-causal Cauchy AR(1) process when $\rho \in (0,1)$:

$$\mathbb{E}(X_t|X_{t-1}) = X_{t-1}, \tag{2.15}$$
$$\mathbb{E}(X_t^2|X_{t-1}) = \frac{1}{\rho}X_{t-1}^2 + \frac{\sigma^2}{\rho(1-\rho)}. \tag{2.16}$$

In other words, (X_t) is a martingale. Moreover, the existence of the conditional first and second-order moments of X_t indicates the existence of the conditional variance of X_t, in contrast to the first-order marginal moment, which does not exist. This suggests that quadratic hedging based on (conditional) variance criterion can likely be considered.

2.3 The Limiting Behavior while in the Bubble

Recently, Fries and Zakoian (2019) showed that in a non-causal AR(1) model, given $|X_{t-1}| > a$, the distribution of $\frac{X_t}{X_{t-1}}$ converges to a binomial distribution with discrete values 0 and $\frac{1}{\rho}$, when a goes to infinity. However, the conditioning set $|X_{t-1}| > a$ they used is not the conditioning set used for predictive distribution.

In this section, we prove a similar result for the noncausal AR(1) process, concerning the conditional distribution of $X_t|X_{t-1}$. Indeed, in the forecasting and hedging context, the conditional distribution of $X_t|X_{t-1}$ is more relevant. Also, because of the scaling property discussed in the previous section, it suffices to consider the case where $\sigma = 1$.

We first introduce $R_t = \frac{X_t}{X_{t-1}}$, and consider the rescaled conditional distribution of $f_{R_t}(r_t|X_{t-1} = x_{t-1})$. By change of variable formula, its CDF is:

$$\begin{aligned} F_{R_t}(r_t|X_{t-1} = x_{t-1}) = &\frac{\alpha(x_{t-1},\rho)}{\pi} \log\left\{\frac{1+(1-\rho)^2(r_t x_{t-1})^2}{1+(x_{t-1}-\rho r_t x_{t-1})^2}\frac{\rho^2}{(1-\rho)^2}\right\} \\ &+ \frac{\beta(x_{t-1},\rho)}{\pi}\left\{\frac{\pi}{2} - \tan^{-1}(x_{t-1}-\rho r_t x_{t-1})\right\} \\ &+ \frac{1-\beta(x_{t-1},\rho)}{\pi}\left\{\tan^{-1}[(1-\rho)r_t x_{t-1}] + \frac{\pi}{2}\right\}, \end{aligned} \tag{2.17}$$

where

$$\alpha(x_{t-1},\rho) = \frac{\rho(1-\rho)^2 x_{t-1}}{(1-2\rho)^2 + (1-\rho)^2 x_{t-1}^2}, \tag{2.18}$$

$$\beta(x_{t-1},\rho) = \frac{\rho\{(1-\rho)^2 x_{t-1}^2 - (1-2\rho)\}}{(1-2\rho)^2 + (1-\rho)^2 x_{t-1}^2}. \tag{2.19}$$

Theorem 2.1. *When $X_{t-1} \to +\infty$, $R_t|X_{t-1}$ converges in distribution to the binomial distribution with masses at 0 and $1/\rho$ and probabilities $1-\rho$ and ρ, respectively.*

Proof. It suffices to show that $F_{R_t}(r_t|X_{t-1} = x_{t-1})$ converges to the CDF of a binomial distribution. To find $\lim_{X_{t-1}\to\infty} F_{R_t}(r_t|X_{t-1} = x_{t-1})$, each component is considered separately below.

The expression (2.18) can be simplified as:

$$\alpha(x_{t-1},\rho) = \frac{\rho(1-\rho)^2/x_{t-1}}{(1-2\rho)^2/x_{t-1}^2 + (1-\rho)^2},$$

and $\lim_{x_{t-1}\to\infty} \alpha(x_{t-1},\rho) \to 0$. From (2.19), we also have

$$\beta(x_{t-1},\rho) = \frac{\rho\{(1-\rho)^2 - (1-2\rho)/x_{t-1}^2\}}{(1-2\rho)^2/x_{t-1}^2 + (1-\rho)^2},$$

and $\lim_{x_{t-1}\to\infty} \beta(x_{t-1},\rho) \to \rho$.

The first term in (2.17) can be simplified as:

$$\frac{\alpha(x_{t-1},\rho)}{\pi} \log\left\{\frac{1+(1-\rho)^2(r_t x_{t-1})^2}{1+(x_{t-1}-\rho r_t x_{t-1})^2}\frac{\rho^2}{(1-\rho)^2}\right\} = \frac{\alpha(x_{t-1},\rho)}{\pi} \log\left\{\frac{1/x_{t-1}^2 + (1-\rho)^2 r_t^2}{1/x_{t-1}^2 + (1-\rho r_t)^2}\frac{\rho^2}{(1-\rho)^2}\right\}. \tag{2.20}$$

Let us now analyze this limit for different values of r_t. For an arbitrary value z, we have that:

- If $r_t = 0$,

$$\begin{aligned}\frac{\alpha(x_{t-1},\rho)}{\pi}\log\left\{\frac{1/x_{t-1}^2+(1-\rho)^2 r_t^2}{1/x_{t-1}^2+(1-\rho r_t)^2}\frac{\rho^2}{(1-\rho)^2}\right\} &= \frac{\alpha(x_{t-1},\rho)}{\pi}\log\left\{\frac{1/x_{t-1}^2}{1/x_{t-1}^2+1}\frac{\rho^2}{(1-\rho)^2}\right\}\\ &\sim \lim_{z\to 0} z\log\{z\}\\ &\to 0. \end{aligned} \tag{2.21}$$

- If $r_t = 1/\rho$,

$$\begin{aligned}\frac{\alpha(x_{t-1},\rho)}{\pi}\log\left\{\frac{1/x_{t-1}^2+(1-\rho)^2 r_t^2}{1/x_{t-1}^2+(1-\rho r_t)^2}\frac{\rho^2}{(1-\rho)^2}\right\} &= \frac{\alpha(x_{t-1},\rho)}{\pi}\log\left\{\frac{1/x_{t-1}^2+(1/\rho-1)^2}{1/x_{t-1}^2}\frac{\rho^2}{(1-\rho)^2}\right\}\\ &\sim \lim_{z\to 0} z\log\{z\}\\ &\to 0. \end{aligned} \tag{2.22}$$

- If $r_t < 0$, or $0 < r_t < 1/\rho$, or $r_t > 1/\rho$, we have

$$\frac{\alpha(x_{t-1},\rho)}{\pi}\log\left\{\frac{1/x_{t-1}^2+(1-\rho)^2 r_t^2}{1/x_{t-1}^2+(1-\rho r_t)^2}\frac{\rho^2}{(1-\rho)^2}\right\} \to 0. \tag{2.23}$$

This implies that, regardless of the value of r_t, the term (2.20) goes to zero when x_{t-1} goes to infinity.

Let us now analyze the second term in (2.17). We have:

$$\begin{aligned}\lim_{x_{t-1}\to\infty}\frac{\beta(x_{t-1},\rho)}{\pi}\left(\frac{\pi}{2}-\tan^{-1}(x_{t-1}-\rho r_t x_{t-1})\right) &= \frac{\rho}{\pi}\lim_{x_{t-1}\to\infty}\left(\frac{\pi}{2}-\tan^{-1}(x_{t-1}-\rho r_t x_{t-1})\right)\\ &= \frac{\rho}{2}-\frac{\rho}{\pi}\lim_{x_{t-1}\to\infty}\tan^{-1}(x_{t-1}-\rho r_t x_{t-1}). \end{aligned} \tag{2.24}$$

Similarly, the last term in (2.17):

$$\begin{aligned}&\lim_{x_{t-1}\to\infty}\frac{1-\beta(x_{t-1},\rho)}{\pi}\left(\tan^{-1}[(1-\rho)r_t x_{t-1}]+\frac{\pi}{2}\right)\\ &\qquad= \frac{1-\lim_{x_{t-1}\to\infty}\beta(x_{t-1},\rho)}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho)r_t x_{t-1}]+\frac{\pi}{2}\right)\\ &\qquad= \frac{1-\rho}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho)r_t x_{t-1}]+\frac{\pi}{2}\right)\\ &\qquad= \frac{1-\rho}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho)r_t x_{t-1}]\right)+\frac{1-\rho}{2}. \end{aligned} \tag{2.25}$$

The following table provides $\tan^{-1}[(1-\rho r_t)x_{t-1}]$ and $\tan^{-1}[(1-\rho)r_t x_{t-1}]$ for different values of r_t, when x_{t-1} goes to infinity.

Table 2.1: Evaluating the $\tan^{-1}(.)$ term in Equation (2.17) given different r_t conditions.

	$r_t < 0$	$r_t = 0$	$0 < r_t < 1/\rho$	$r_t = 1/\rho$	$r_t > 1/\rho$
$\lim_{x_{t-1}\to\infty} \tan^{-1}[(1-\rho r_t)x_{t-1}]$	$\pi/2$	$\pi/2$	$\pi/2$	0	$-\pi/2$
$\lim_{x_{t-1}\to\infty} \tan^{-1}[(1-\rho)r_t x_{t-1}]$	$-\pi/2$	0	$\pi/2$	$\pi/2$	$\pi/2$

Combining all three terms, the behaviour of the limiting CDF under the conditions presented in the above table is examined.

When $r_t < 0$,

$$\begin{aligned}
\lim_{x_{t-1}\to\infty} F_{R_t}(r_t|X_{t-1}=x_{t-1}) &= \frac{1}{2} - \frac{\rho}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho r_t)x_{t-1}]\right) + \frac{1-\rho}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho)r_t x_{t-1}]\right) \\
&= \frac{1}{2} - \frac{\rho}{\pi}\left(\frac{\pi}{2}\right) + \frac{1-\rho}{\pi}\left(-\frac{\pi}{2}\right) \\
&= \frac{1}{2} - \frac{\rho}{2} - \frac{1-\rho}{2} \\
&= 0. \qquad (2.26)
\end{aligned}$$

When $r_t = 0$,

$$\begin{aligned}
\lim_{x_{t-1}\to\infty} F_{R_t}(r_t|X_{t-1}=x_{t-1}) &= \frac{1}{2} - \frac{\rho}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho r_t)x_{t-1}]\right) + \frac{1-\rho}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho)r_t x_{t-1}]\right) \\
&= \frac{1}{2} - \frac{\rho}{\pi}\left(\frac{\pi}{2}\right) + \frac{1-\rho}{\pi}(0) \\
&= \frac{1}{2} - \frac{\rho}{2} \\
&= \frac{1-\rho}{2}. \qquad (2.27)
\end{aligned}$$

When $0 < r_t < 1/\rho$,

$$\begin{aligned}
\lim_{x_{t-1}\to\infty} F_{R_t}(r_t|X_{t-1}=x_{t-1}) &= \frac{1}{2} - \frac{\rho}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho r_t)x_{t-1}]\right) + \frac{1-\rho}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho)r_t x_{t-1}]\right) \\
&= \frac{1}{2} - \frac{\rho}{\pi}\left(\frac{\pi}{2}\right) + \frac{1-\rho}{\pi}\left(\frac{\pi}{2}\right) \\
&= \frac{1}{2} - \frac{\rho}{2} + \frac{1-\rho}{2} \\
&= \frac{2-2\rho}{2} \\
&= 1-\rho. \qquad (2.28)
\end{aligned}$$

When $r_t = 1/\rho$,

$$\begin{aligned}
\lim_{x_{t-1}\to\infty} F_{R_t}(r_t|X_{t-1}=x_{t-1}) &= \frac{1}{2} - \frac{\rho}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho r_t)x_{t-1}]\right) + \frac{1-\rho}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho)r_t x_{t-1}]\right) \\
&= \frac{1}{2} - \frac{\rho}{\pi}(0) + \frac{1-\rho}{\pi}\left(\frac{\pi}{2}\right) \\
&= \frac{1}{2} + \frac{1-\rho}{2} \\
&= \frac{2-\rho}{2}. \qquad (2.29)
\end{aligned}$$

When $r_t > 1/\rho$,

$$\begin{aligned}
\lim_{x_{t-1}\to\infty} F_{R_t}(r_t|X_{t-1}=x_{t-1}) &= \frac{1}{2} - \frac{\rho}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho r_t)x_{t-1}]\right) + \frac{1-\rho}{\pi}\left(\lim_{x_{t-1}\to\infty}\tan^{-1}[(1-\rho)r_t x_{t-1}]\right) \\
&= \frac{1}{2} - \frac{\rho}{\pi}\left(-\frac{\pi}{2}\right) + \frac{1-\rho}{\pi}\left(\frac{\pi}{2}\right) \\
&= \frac{1}{2} + \frac{\rho}{2} + \frac{1-\rho}{2} \\
&= 1. \qquad (2.30)
\end{aligned}$$

□

Figure 2.2 below is a plot of the limiting CDF computed above for $\rho \in \{0.978, 0.8, 0.5\}$. It can be seen that, for each ρ value, the distribution of the limiting function is equal to the CDF of the binomial distribution with masses at 0 and $1/\rho$ and probabilities $1-\rho$ and ρ, respectively.

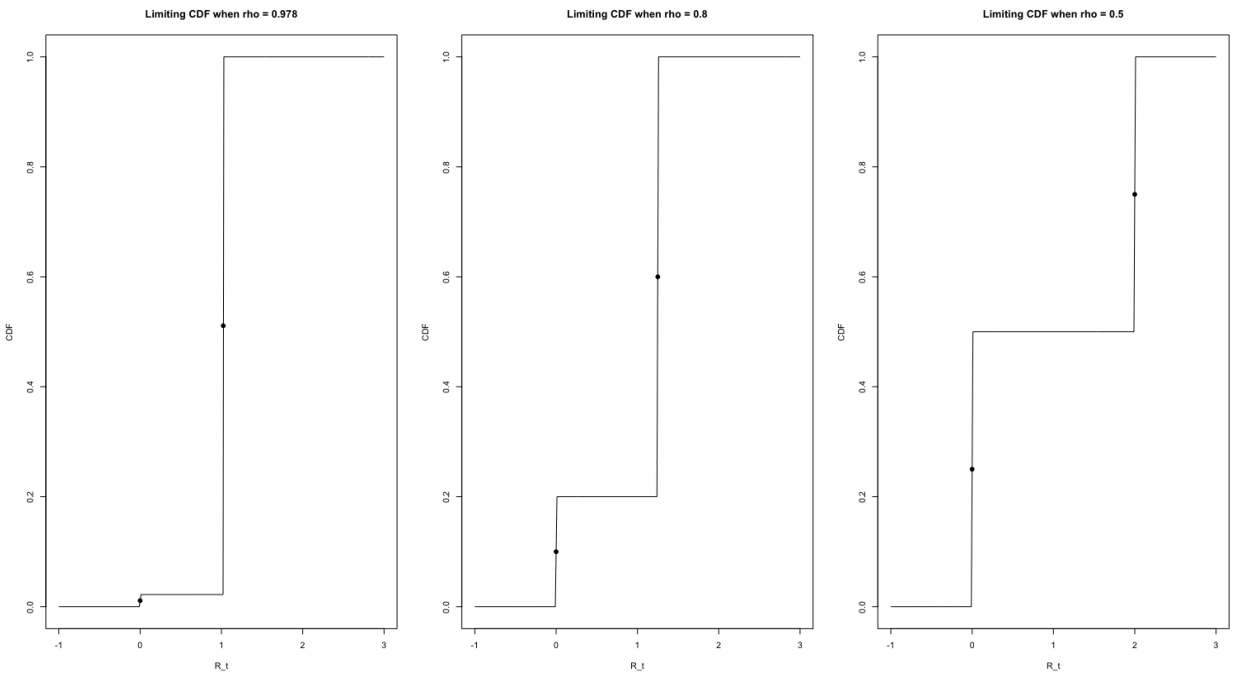

Figure 2.2: Convergence of $F_{R_t}(r_t|X_{t-1} = x_{t-1})$ to the binomial CDF with masses at 0 and $1/\rho$ and probabilities $1-\rho$ and ρ, respectively when $x_{t-1} \to +\infty$, for different values of ρ.

2.4 Illustration

As an illustration, $f_{R_t}(r_t|X_{t-1} = x_{t-1})$ is plotted for different values of $x_{t-1} \in \{100, 250, 500\}$, from top to down and $\rho \in \{0.978, 0.8, 0.5\}$, from left to right respectively. It can be seen that the peak of the distribution occurs at 0 and $\frac{1}{\rho}$ as expected.

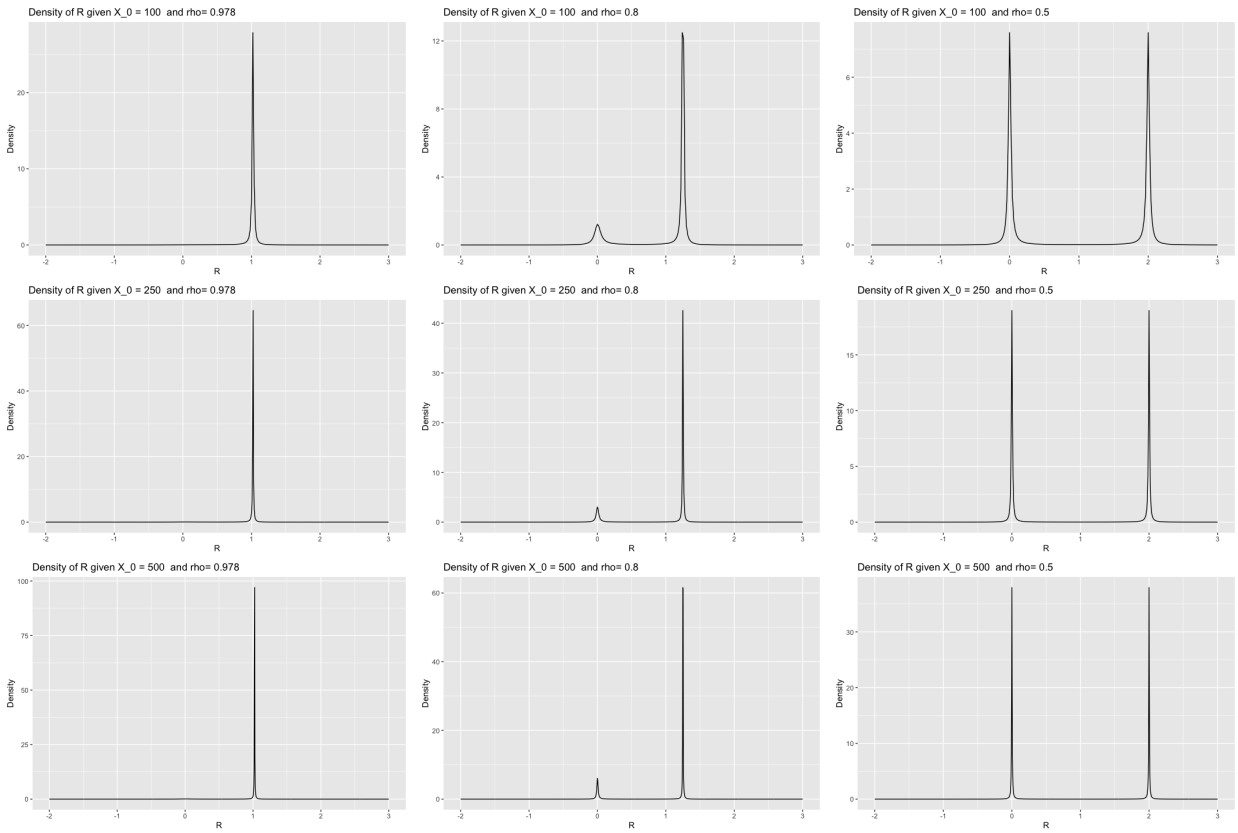

Figure 2.3: Probability density curves for $f_{R_t}(r_t|X_{t-1} = x_{t-1})$.

Also, the weak convergence through the convergence of some integrals is illustrated. It is well known that the weak convergence of a sequence of probability measures implies the convergence of the sequence of Laplace transforms. One consequence of the convergence of distribution of $R_t|X_{t-1}$ is that, we have

$$\int_{-\infty}^{\infty} e^{-iur} f_{R_t}(r_t|X_{t-1} = x_{t-1})dr \xrightarrow[X_{t-1}\to+\infty]{} \int_{-\infty}^{\infty} e^{-iur}[\rho\delta_0(r) + (1-\rho)\delta_{\frac{1}{\rho}}(r)]dr \tag{2.31}$$

$$= (1-\rho) + \rho e^{-i\frac{u}{\rho}}, \tag{2.32}$$

for all real u, where i is the pure imaginary number such that $i^2 = -1$. That is, its real part converges to $(1-\rho) + \rho cos(\frac{u}{\rho})$ and its imaginary part converges to $\rho sin(\frac{u}{\rho})$. We take $u \in \{0.1, 0.2\}$, $\rho \in \{0.5, 0.8, 0.978\}$ and plot the left-hand side of (2.31) as a function of X_{t-1}, and report also the limiting value (2.32) in the same figure, where the left-hand side of (2.31) is computed by simulation. Figure 2.4 displays $\int_{-\infty}^{\infty} cos(ur) f_{R_t}(r_t|X_{t-1} = x_{t-1})dr$ and $(1-\rho) + \rho cos(\frac{u}{\rho})$ (in red), which represents the real part of (2.32). For the imaginary case, Figure 2.5 also displays $\int_{-\infty}^{\infty} sin(ur) f_{R_t}(r_t|X_{t-1} = x_{t-1})dr$ and $\rho sin(\frac{u}{\rho})$ (in red). The difference of the limiting values is explained by the fact that $sin(0) = 0$ and $cos(0) = 1$.

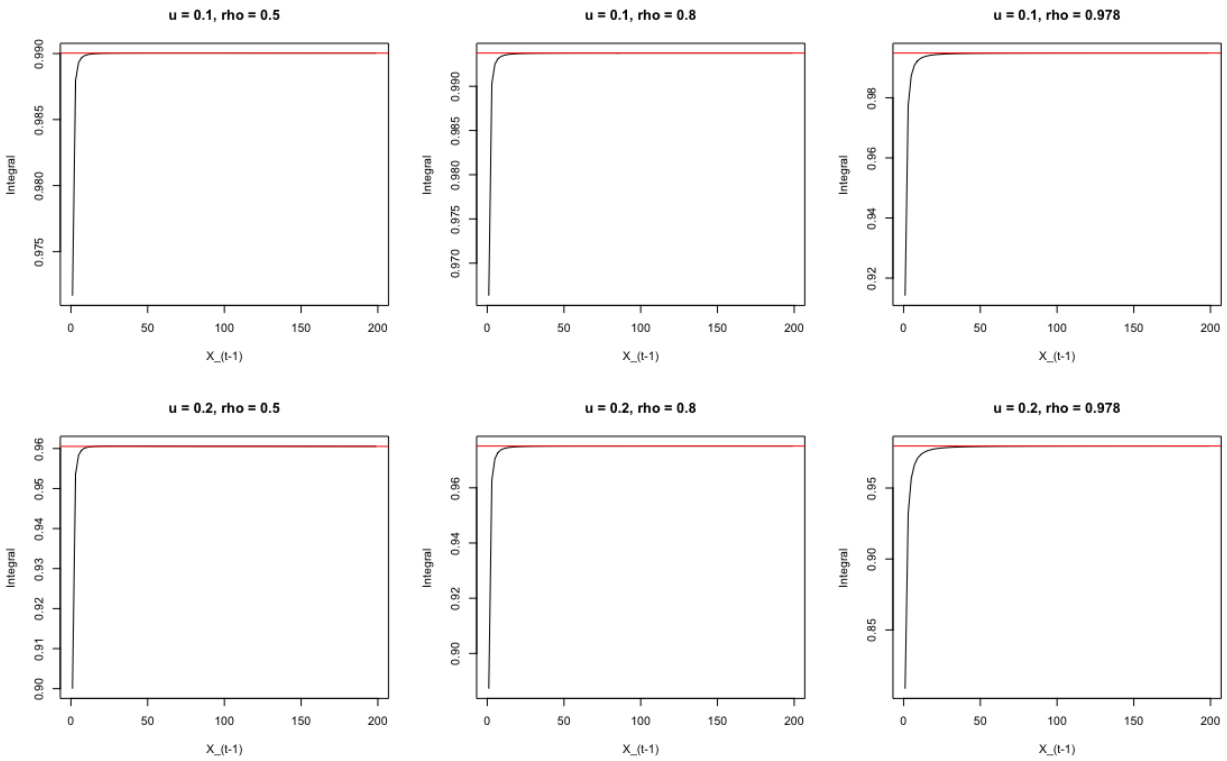

Figure 2.4: Convergence of $\int_{-\infty}^{\infty} cos(ur) f_{R_t}(r_t|X_{t-1} = x_{t-1})dr$ to $(1-\rho)+\rho cos(\frac{u}{\rho})$ (in red) when $x_{t-1} \to +\infty$, for different values of u and ρ.

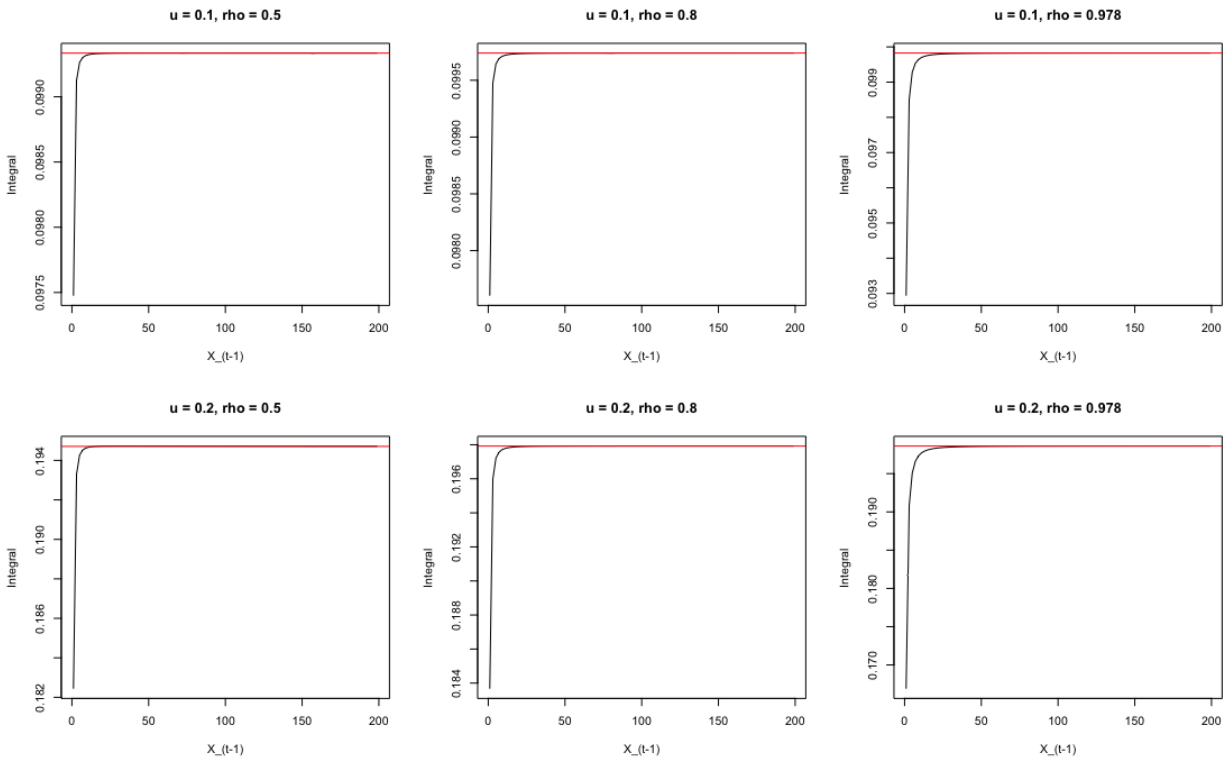

Figure 2.5: Convergence of $\int_{-\infty}^{\infty} sin(ur) f_{R_t}(r_t|X_{t-1} = x_{t-1})dr$ to $\rho sin(\frac{u}{\rho})$ (in red) when $x_{t-1} \to +\infty$, for different values of u and ρ.

2.5 Simulation of Trajectories of AR(1) Processes

There are two primary approaches to simulating trajectories of AR(1) processes. The first involves generating a path from a terminal condition, working backward without fixing an initial value, as used by Gouriéroux and Zakoïan (2017). This approach requires known terminal conditions which is not the case in practice. The second, which is our chosen approach, simulates forwardly given an initial value and requires the use of conditional density. This research uses parameters detailed in Section 7 of Gouriéroux and Zakoïan (2017), particularly $\sigma = 1.048$ and $\rho = 0.978$. We set the initial condition, $X_0 = 100$, and observe the changes in the distribution as the parameter ρ approaches 0.5. The following subsections briefly discuss some simulation methods that can be used for simulating trajectories of the AR(1) process.

2.5.1 Inverse Transform method

Given a continuous random variable X with a known CDF $F(x)$, the Inverse Transform Method can be used to generate random variables following the distribution of X. This method is particularly efficient when $F^{-1}(u)$ has a closed form. The algorithm for this method involves first generating a random variable $U \sim \mathcal{U}(0,1)$. After, $X = F^{-1}(U)$ is computed to obtain a random variable following the distribution $F(x)$. The CDF of the non-causal AR(1) Cauchy model exists, therefore, this method could be used. Specifically, the CDF (2.13) and (2.14) could be used when $\sigma = 1$ and $\sigma \neq 1$ respectively by computing their inverse CDF and following the algorithm.

2.5.2 Sampling Importance-Resampling method

Sampling Importance Resampling (SIR) is a simulation technique used to generate a random sample with a given density by generating random variables from another density that is different from the desired one, similar to the concept of importance sampling. Subsequently, the algorithm performs resampling from this pool of generated values to obtain the desired random vector. Specifically, suppose X is a random variable with density $f(x)$, the algorithm begins by selecting an alternative (or proposal) distribution $g(x)$ that is relatively simple to sample from and close to the target distribution $f(x)$. Next, a set of N samples $\{x_i\}_{i=1}^N$ is drawn from this proposal distribution. For each sample x_i, an importance weight w_i is computed as $w_i = \frac{f(x_i)}{g(x_i)}$, which assesses how representative each sample is to the target distribution. These weights are then normalized to ensure their sum equals 1, giving $\hat{w}_i = \frac{w_i}{\sum_{j=1}^N w_j}$. The algorithm proceeds by resampling N times with replacement from the set $\{x_i\}$ based on the normalized weights $\{\hat{w}_i\}$. This step adjusts the distribution of the samples to more closely resemble $f(x)$. The resulting set of resampled points now follows the initial distribution $f(x)$. For the non-causal model (2.8), a proposal distribution capturing the peaks at 0 and $\frac{X_{t-1}}{\rho}$ could be used. That is, a mixed Cauchy density weighted with $(1-\rho)$ and ρ and location parameters 0 and $\frac{X_{t-1}}{\rho}$ as presented below could be used:

$$g(x) = (1-\rho) \times \frac{1}{\pi} \frac{\sigma}{x^2 + \sigma^2} + \rho \times \frac{1}{\pi} \frac{\sigma}{(x - \frac{x_{t-1}}{\rho})^2 + \sigma^2}. \tag{2.33}$$

2.5.3 Acceptance-Rejection method

The Acceptance-Rejection Method is another method for simulating random variables, especially when an explicit formula for the inverse of the CDF is not readily available or computationally feasible. Given a continuous random variable X with target pdf $f(x)$, this method uses a proposal distribution G with pdf

$g(x)$ which can be easily sampled, and is bound by the relation:

$$0 \leq \frac{f(x)}{g(x)} \leq c \quad \forall x \in \mathbb{R}$$

where c is a constant. To implement this method, begin by generating a random variable Y distributed as G, the proposal distribution. Then, generate an independent random variable $U \sim \mathcal{U}(0,1)$ so that if $U \leq \frac{f(Y)}{cg(Y)}$, set $X = Y$ (i.e., "accept" Y); else, "reject" Y.

All the three simulation methods presented above can be used to simulate the non-causal AR(1) process, however, the acceptance-rejection method was used in this thesis. Figure 2.3 indicates that the non-causal AR(1) Cauchy model has two components, as seen in the last column, which is the $\rho = 0.5$ case. Therefore, we ran our simulations using the acceptance-rejection algorithm with the constant c chosen such that, $c = \max\left\{\frac{f(x)}{g(x)}\right\}$. The mixed Cauchy density (2.33) was used as the proposal density. Using a single Cauchy distribution would lead to a very low acceptance rate (approximately 2% in our simulation) compared to approximately 78% with the mixed distribution. Considering a time horizon of $h = 1$, we simulate X_1 given $X_0 = 100$ and the constant $c = 3.733, 6.039, 3.999$ when $\rho = 0.978, 0.8, 0.5$ respectively. Figure 2.6 below shows the probability density curves for both the simulations and the non-causal AR(1) Cauchy model.

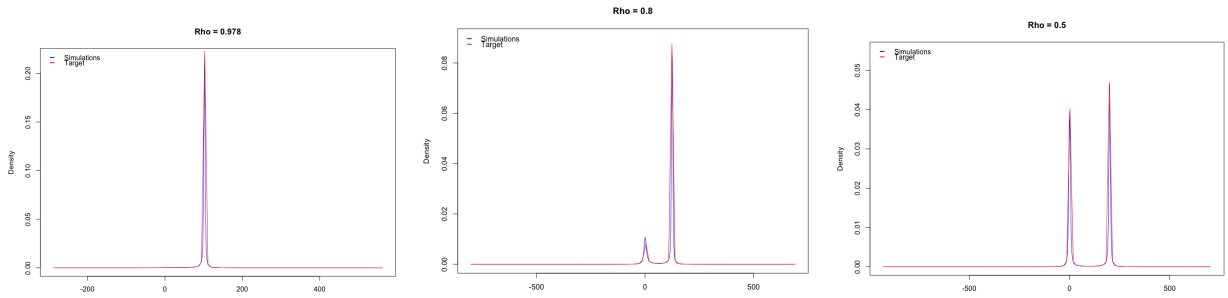

Figure 2.6: Probability density curves for simulated (blue) and target(red) non-causal model with $X_0 = 100$.

Chapter 3

Hedging Strategies

Financial markets are often classified into two main categories: complete and incomplete. In a complete market, every contingent claim can be perfectly hedged. An example of this concept is the Black-Scholes-Merton model (Black and Scholes (1973) and Merton (1973)) in continuous time. In the real-world context, many markets are incomplete. An example of an incomplete market is one where certain types of risks, like currency risk or political risk, cannot be fully hedged due to a lack of corresponding financial instruments or inefficiencies in the market. In such markets, perfect hedging is unattainable. To address this limitation, super-replicating strategies can be employed. A super-replicating portfolio is any self-financing strategy where the portfolio's value at maturity is always greater than or equal to the contingent claim value. While effective, these strategies are expensive. An alternative is quadratic hedging, offering a middle ground between cost and risk. This approach, which Föllmer and Sondermann (1986) first used in the framework of finance, looks for a strategy that reduces the total expected squared hedging errors. Within quadratic hedging, two approaches will be considered: global and local strategies. Global quadratic hedging aims to determine a sequence of hedging decisions to minimize the expected squared terminal hedging error, ensuring the hedging portfolio remains self-financing. The local quadratic hedging, on the other hand, allows for some intermediate errors and dynamically minimizes the local hedging errors based on new information. This approach identifies optimal positions using a blinded approach by minimizing quadratic costs for each time interval. The upcoming sections will introduce the mathematical terminologies and tools that are essential for a detailed understanding of these methods.

3.1 Terminology

This subsection introduces some tools and notations that will be used in explaining the hedging methods. Consider a series of trading events that happen at regular intervals within a predetermined finite time frame, denoted by $T \in \mathbb{N}$. The sequence $\mathcal{T}$ is used to represent these discrete times, where each element corresponds to an integer ranging from 0 up to and including T. $\{X_t\}_{t\in\mathcal{T}}$ is a stochastic process, defined on a probability space $(\Omega, \mathcal{F}_T, \mathbb{P})$ where the filtration $\mathcal{F}_t$ represents stock price information till time t, and $\mathbb{P}$ represents the real-world probability measure. $\{\beta_t\}_{t\in\mathcal{T}}$ is also a risk-free asset price process assumed to grow at a risk-free rate r such that $\beta_t = e^{rt}$. A trading strategy is defined as a pair of stochastic processes $\{\phi_t\}_{t\in\mathcal{T}} = \{\phi_{t,1}, \phi_{t,2}\}_{t\in\mathcal{T}}$, where the $\mathcal{F}_t$-measurable random variables $\phi_{t,1}$ and $\phi_{t,2}$ respectively denotes the quantity of risky asset and the amount of risk-free asset held at time t, where $t = 0, 1, ..., T-1$. The hedging portfolio value sums up the values of all hedging investments up to time t, and it is given by,

$$V_t = \phi_{t,1}X_t + \phi_{t,2}. \tag{3.1}$$

The payoff P for a European option with maturity T is the maximum of either the difference between the asset price X_T at maturity and a predetermined price, known as the strike price K, or zero. Specifically, for a European call option, the payoff is $\max(X_T - K, 0)$, and for a European put option, it is $\max(K - X_T, 0)$. Let W_t denote the accumulated value of a position where $\phi_{t-1,1}$ shares of the risky asset are bought at time $t-1$, and a cash position $\phi_{t-1,2}$ is held from time $t-1$ to t. That is,

$$W_t = \phi_{t-1,1}X_t + e^r\phi_{t-1,2}, \tag{3.2}$$

for $t = 1, \ldots, T.$, where $\phi_{t-1,2} = V_{t-1} - \phi_{t-1,1}X_{t-1}$ since we consider two assets.

A special hedging strategy is a self-financing strategy. This strategy is one where the hedger neither adds nor withdraws money from the hedging portfolio, regardless of the changes in the prices of assets. In our setting, this means that there's no difference between the accumulated portfolio value (W_t) and the hedging strategy value (V_t) at time $t = 0, \ldots, T$. It only allows errors at maturity if any. Let Φ_T denote the set of all self-financing hedging strategies between 0 and T.

Proposition 3. *For a self-financing strategy with two assets, the value of the strategy at time t is given by,*

$$V_t = \beta_t \left(V_0 + \sum_{i=1}^{t} \phi_{i-1,1} \Delta X_i \right), \tag{3.3}$$

for $t = 1, \ldots, T-1$, where $\Delta X_i = \beta_i^{-1} X_i - \beta_{i-1}^{-1} X_{i-1}$.

Proof. By induction, and from the definition of self-financing, at time 1 the value of the strategy is equal to the accumulated portfolio value (i.e. $V_1 = W_1$). Using the definition of W_t in (3.2), we have

$$\begin{aligned} V_1 &= \phi_{0,1} X_1 + \beta_1 \phi_{0,2} \\ &= \phi_{0,1} X_1 + \beta_1 (V_0 - \phi_{0,1} X_0) \\ &= \beta_1 (V_0 + \phi_{0,1} \beta_1^{-1} X_1 - \phi_{0,1} X_0) \\ &= \beta_1 (V_0 + \phi_{0,1} \Delta X_1), \end{aligned} \tag{3.4}$$

which verifies that (3.3) is true for $t = 1$. Next, assume that (3.3) is true for t.

Now at time $t+1$, $V_{t+1} = W_{t+1}$ by the definition of self-financing. So we have

$$\begin{aligned} V_{t+1} &= \phi_{t,1} X_{t+1} + \frac{\beta_{t+1}}{\beta_t} \phi_{t,2} \\ &= \phi_{t,1} X_{t+1} + \frac{\beta_{t+1}}{\beta_t} (V_t - \phi_{t,1} X_t) \\ &= \frac{\beta_{t+1}}{\beta_t} (V_t + \phi_{t,1} \frac{\beta_t}{\beta_{t+1}} X_{t+1} - \phi_{t,1} X_t) \\ &= \frac{\beta_{t+1}}{\beta_t} \left(\beta_t \left(V_0 + \sum_{i=1}^{t} \phi_{i-1,1} \Delta X_i \right) + \phi_{t,1} \frac{\beta_t}{\beta_{t+1}} X_{t+1} - \phi_{t,1} X_t \right) \\ &= \beta_{t+1} \left(V_0 + \sum_{i=1}^{t} \phi_{i-1,1} \Delta X_i + \phi_{t,1} \beta_{t+1}^{-1} X_{t+1} - \phi_{t,1} \beta_t^{-1} X_t \right) \\ &= \beta_{t+1} \left(V_0 + \sum_{i=1}^{t} \phi_{i-1,1} \Delta X_i + \phi_{t,1} \Delta X_{t+1} \right) \\ &= \beta_{t+1} \left(V_0 + \sum_{i=1}^{t+1} \phi_{i-1,1} \Delta X_i \right), \end{aligned} \tag{3.5}$$

which completes the proof. □

3.2 Binomial Framework

In this section, we explore the binomial asset-pricing model, also known as the CRR model (Cox, Ross, and Rubinstein (1979)), a fundamental framework for understanding concepts in arbitrage pricing theory. We begin with the most basic form of this model, which is the one-period model. Then we look at the more general multi-period case. The binomial model creates a replicating portfolio by holding a certain number of shares of the underlying stock and a risk-free asset, ensuring perfect payoffs. This model, under specific conditions, aligns with the results discussed in Section 2.3, where the conditional distribution of asset returns converges to a binomial distribution. The discussion in this section is mostly based on content from Shreve (2005).

3.2.1 One-period binomial model

Consider a discrete-time one-period binomial model. Suppose at time 0 we have the price of one share of stock $X_0 > 0$, at time 1 the price either goes up to $X_1^1 = uX_0$ or down to $X_1^0 = dX_0$. For the binomial model to be arbitrage-free and complete; the up factor (u) and down factor (d) must satisfy $u > e^r > d$. Unless stated otherwise, superscripts in this section would represent the number of up movements. If we are considering the European call option, for instance, its payoff at maturity under a one-period model is given as $P = \max(X_1 - K, 0)$. We want to create a hedging portfolio that will mimic the payoff of the option. Hence, we want our accumulated hedging portfolio at time one to be equal to the option payoff. That is,

$$W_1 = P. \tag{3.6}$$

It means that the hedging portfolio replicates the payoff of this option at time one by choosing ϕ_0 such that, the hedging strategy at time one in the up state (W_1^1) is equal to the value of the payoff in the up state (P^1) and similarly, for the down states. That is, we want $W_1^1 = P^1$ and $W_1^0 = P^0$. Using (3.2), we get

$$\phi_{0,1}X_1^1 + e^r(V_0 - \phi_{0,1}X_0) = P^1, \tag{3.7}$$

$$\phi_{0,1}X_1^0 + e^r(V_0 - \phi_{0,1}X_0) = P^0. \tag{3.8}$$

Solving (3.7) and (3.8) for $\phi_{0,1}$, we get:

$$\phi_{0,1} = \frac{P^1 - P^0}{X_1^1 - X_1^0}, \tag{3.9}$$

and substituting (3.9) back into (3.7), we have:

$$V_0 = e^{-r}P^1 - \left[\frac{P^1 - P^0}{X_1^1 - X_1^0}\right]\left(e^{-r}X_1^1 - X_0\right). \tag{3.10}$$

The no-arbitrage condition imposes that the value of the option is V_0. Equation (3.10) can further be simplified using the definitions $X_1^1 = uX_0$ and $X_1^0 = dX_0$ to get the value at time zero as,

$$\begin{aligned} V_0 &= \left[\frac{e^r - d}{u - d}\right]e^{-r}P^1 + \left[1 - \frac{e^r - d}{u - d}\right]e^{-r}P^0 \\ &= \widetilde{p}e^{-r}P^1 + (1 - \widetilde{p})e^{-r}P^0 \\ &= \widetilde{\mathbb{E}}[e^{-r}P], \end{aligned} \tag{3.11}$$

where $\widetilde{p} = \frac{e^r - d}{u - d}$ and $\widetilde{\mathbb{E}}[\cdot]$ is the expected value under the risk-neutral probability measure $\mathbb{Q}$. In the special case discussed in the previous chapter where $u = \frac{1}{\rho}$ and $d = 0$, $\widetilde{p} = e^r\rho$.

3.2.2 Multi-period binomial model

The one-period binomial model introduced in the previous subsection can be extended to a multi-period case where instead of moving from time 0 to 1, we go from time 0 to T, where the hedging portfolio can be rebalanced at time $t = 0, \ldots, T-1$. The initial price of the stock is still X_0. Using the same idea as the one-period case, the stock price can either go up or down at time 1. The stock price will continue this process for each time step. That is, given the stock price X_t^n at time t where $n = 0, \ldots, t$ is the total number of up moves since the beginning, the stock price can either go up to $X_{t+1}^{n+1} = uX_t^n$ or down to $X_{t+1}^n = dX_t^n$ at time $t+1$. For example, when maturity $T = 3$, Figure 3.1 represents a three-period recombining binomial model with the bold nodes being the points where the tree recombines.

It can be seen from this figure that, $X_t^n = u^n d^{t-n} X_0$ for $0 \le n \le t$. Similar to the one-period case, the prices of derivative securities can be constructed by a replication portfolio. The expression for ϕ_t and V_t can be obtained recursively. That is, at time $T-1$, we choose a ϕ_{T-1} such that $W_T^{n+1} = P^{n+1}$ and $W_T^n = P^n$ for node $n = 0, \ldots, T-1$. Hence, using (3.2), we get

$$\phi_{T-1,1}X_T^{n+1} + e^r(V_{T-1}^n - \phi_{T-1,1}X_{T-1}^n) = P^{n+1}, \tag{3.12}$$

$$\phi_{T-1,1}X_T^n + e^r(V_{T-1}^n - \phi_{T-1,1}X_{T-1}^n) = P^n. \tag{3.13}$$

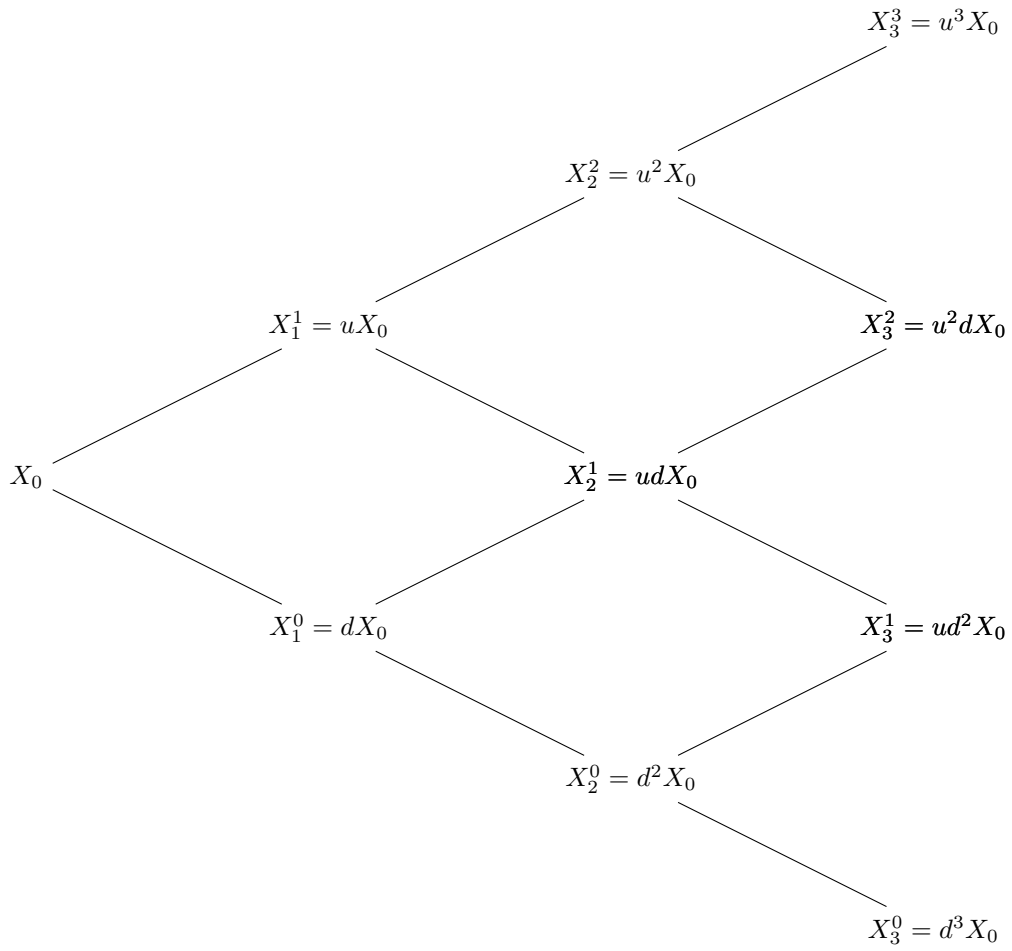

Figure 3.1: A special case of the binomial tree with maturity $T = 3$.

Solving (3.12) and (3.13) for $\phi_{T-1,1}$, we get

$$\phi_{T-1,1} = \frac{P^{n+1} - P^n}{X_T^{n+1} - X_T^n}, \tag{3.14}$$

and substituting (3.14) back into (3.12), we have

$$V_{T-1}^n = e^{-r}P^{n+1} - \left[\frac{P^{n+1} - P^n}{X_T^{n+1} - X_T^n}\right]\left(e^{-r}X_T^{n+1} - X_{T-1}^n\right). \tag{3.15}$$

Given the information available at time $T-1$ ($\mathcal{F}_{T-1}$) and the no-arbitrage condition, the option value at time $T-1$ is V_{T-1}. Using $X_T^{n+1} = uX_{T-1}^n$ and $X_T^n = dX_{T-1}^n$, Equation (3.15) can be simplified to get the

value at time $T-1$ as,

$$\begin{aligned} V_{T-1}^n &= \widetilde{p}e^{-r}P^{n+1} + (1-\widetilde{p})e^{-r}P^n \\ &= \widetilde{\mathbb{E}}[e^{-r}P|\mathcal{F}_{T-1}]. \end{aligned} \tag{3.16}$$

Similarly, going from time $t+1$ to t, $\phi_{t,1}$ and V_t can be found by replacing P^n in (3.14) and (3.16) with V_{t+1}^n, given X_t^n as,

$$\phi_{t,1} = \frac{V_{t+1}^{n+1} - V_{t+1}^n}{X_{t+1}^{n+1} - X_{t+1}^n}, \tag{3.17}$$

$$V_t^n = \widetilde{\mathbb{E}}[e^{-r}V_{t+1}|\mathcal{F}_t]. \tag{3.18}$$

This recursive process will continue until time 0 to find $\phi_{0,1}$ and V_0.

In this process where the portfolio's composition is adjusted at each step (or node in the binomial tree) without injecting or withdrawing funds, the strategy is self-financing. Also to point out, the relation between the binomial model and the non-causal process in our case is that, as established in Section 2.3, the conditional distribution of $\frac{X_t}{X_{t-1}}$ given X_{t-1} converges to a binomial distribution with $d=0$ and $u=\frac{1}{\rho}$.

3.3 Quadratic Hedging: One-period

The one-period quadratic hedging involves setting a hedging strategy at the beginning of the period and not adjusting it until the period ends. This constitutes a portfolio where $\phi_{0,1}$ shares of the risky asset are held without any further adjustment until maturity. The optimal value of $\phi_{0,1}$ (shares of the risky asset) and initial value V_0 are determined through the optimization problem

$$\min_{\phi_0} \ \mathbb{E}\left[(\beta_1^{-1}(P - W_1))^2 | X_0\right].$$

The expression in the expectation can be expanded by using the definition of W_1 as follows:

$$\begin{aligned}\beta_1^{-1}(W_1 - P) &= \beta_1^{-1}(P - (\phi_{0,1}X_1 + \beta_1\phi_{0,2})) \\ &= \beta_1^{-1}[P - (\phi_{0,1}X_1 + \beta_1(V_0 - \phi_{0,1}X_0))] \\ &= \beta_1^{-1}P - (\beta_1^{-1}\phi_{0,1}X_1 + V_0 - \phi_{0,1}X_0) \\ &= \beta_1^{-1}P - (V_0 + \phi_{0,1}(\beta_1^{-1}X_1 - X_0)). \end{aligned} \tag{3.19}$$

Therefore, the values for V_0 and $\phi_{0,1}$ can be obtained by solving the optimization problem

$$\min_{\phi_{0,1},V_0} \ \mathbb{E}\left[(\beta_1^{-1}P - V_0 - \phi_{0,1}(\beta_1^{-1}X_1 - X_0))^2|X_0\right].$$

Suppose the objective function of the optimization problem is represented by $\mathcal{O}(\phi_{0,1}, V_0)$, or shortly $\mathcal{O}$. That is,

$$\mathcal{O}(\phi_{0,1}, V_0) = \mathbb{E}\left[(\beta_1^{-1}P - V_0 - \phi_{0,1}(\beta_1^{-1}X_1 - X_0))^2|X_0\right]. \tag{3.20}$$

To find the values V_0 and $\phi_{0,1}$ that minimizes $\mathcal{O}$, we find the derivative of $\mathcal{O}$ with respect to V_0 and $\phi_{0,1}$, equate to zero. Differentiating the expectation in the optimization problem is possible because of the existence of the first and second moments of the noncausal process. That is,

$$\frac{\partial \mathcal{O}}{\partial V_0} = -2\,\mathbb{E}[\beta_1^{-1}P - V_0 - \phi_{0,1}(\beta_1^{-1}X_1 - X_0)]. \tag{3.21}$$

Solving $\frac{\partial \mathcal{O}}{\partial V_0} = 0$, we have

$$0 = \mathbb{E}[\beta_1^{-1}P - V_0 - \phi_{0,1}(\beta_1^{-1}X_1 - X_0)], \tag{3.22}$$

and we get

$$V_0 = \mathbb{E}[\beta_1^{-1}P - \phi_{0,1}(\beta_1^{-1}X_1 - X_0)], \tag{3.23}$$

$$= \mathbb{E}[\beta_1^{-1}P] - \phi_{0,1}\,\mathbb{E}[(\beta_1^{-1}X_1 - X_0)]. \tag{3.24}$$

Now for $\phi_{0,1}$,

$$\frac{\partial \mathcal{O}}{\partial \phi_{0,1}} = -2\,\mathbb{E}[(\beta_1^{-1}X_1 - X_0)(\beta_1^{-1}P - V_0 - \phi_{0,1}(\beta_1^{-1}X_1 - X_0))]. \tag{3.25}$$

Again, when $\frac{\partial \mathcal{O}}{\partial \phi_{0,1}} = 0$ we have

$$\begin{aligned}
0 &= \mathbb{E}[(\beta_1^{-1}X_1 - X_0)(\beta_1^{-1}P - V_0 - \phi_{0,1}(\beta_1^{-1}X_1 - X_0))] \\
&= \mathbb{E}[\beta_1^{-1}P(\beta_1^{-1}X_1 - X_0)] - V_0\,\mathbb{E}[\beta_1^{-1}X_1 - X_0] - \phi_{0,1}\,\mathbb{E}[(\beta_1^{-1}X_1 - X_0)^2] \\
&= \mathbb{E}[\beta_1^{-1}P(\beta_1^{-1}X_1 - X_0)] - \mathbb{E}[\beta_1^{-1}P]\,\mathbb{E}[\beta_1^{-1}X_1 - X_0] + \phi_{0,1}(\mathbb{E}[\beta_1^{-1}X_1 - X_0])^2 - \phi_{0,1}\,\mathbb{E}[(\beta_1^{-1}X_1 - X_0)^2] \\
&= Cov[\beta_1^{-1}P, (\beta_1^{-1}X_1 - X_0)] - \phi_{0,1}Var[\beta_1^{-1}X_1 - X_0], \qquad (3.26)
\end{aligned}$$

and we get

$$\phi_{0,1} = \frac{Cov[\beta_1^{-1}P, (\beta_1^{-1}X_1 - X_0)]}{Var[\beta_1^{-1}X_1 - X_0]}. \qquad (3.27)$$

3.4 Quadratic Hedging: Multiple-period

While one-period quadratic hedging involves setting a hedging strategy at the beginning of the period and holding it until the period ends, multiple-period quadratic hedging involves hedging over multiple periods by adjusting the strategy at each period or considering the entire hedging horizon as a single optimization problem. The one-period local and global quadratic hedging reduces to one-period quadratic hedging due to the absence of intermediate rebalancing opportunities.

3.4.1 Local approach

The local quadratic hedging employs a backward recursive scheme, refining the hedging strategy based on the most recent observations. The fundamental objective of this strategy is to minimize the squared expected local hedging errors. That is, keeping the difference between the projected hedge and actual market movement as small as possible. In essence, we seek to find the trading strategy ϕ that minimizes:

$$\min_{\phi_{t-1}} \mathbb{E}[e^{-r}(V_t - W_t)^2 | \mathcal{F}_{t-1}],$$

for $t = 1, \ldots, T$, where $V_T = P$ to get an expression for the optimal quantity ϕ of the risky asset to be held at time t. To do this, it can be seen that the above optimization problem is similar to the one-period quadratic approach where P is replaced with V_t. Thus, using (3.27) we have

$$\phi_{t-1,1} = \frac{Cov[e^{-r}V_t,\, \Delta_t | \mathcal{F}_{t-1}]}{Var[\Delta_t | \mathcal{F}_{t-1}]}, \qquad (3.28)$$

where $\Delta_t = X_t e^{-r} - X_{t-1}$. Similarly, using (3.23), V_{t-1} is computed as:

$$V_{t-1} = \mathbb{E}[e^{-r} V_t - \phi_{t-1,1} \Delta_t | \mathcal{F}_{t-1}]. \tag{3.29}$$

The relationship in (3.29) ties our current portfolio's value to its anticipated value for the ensuing period, leveraging all information up to $t-1$. The quadratic hedging strategy has a mean self-financing property. A mean self-financing trading strategy has an expected hedging error of zero.

The local quadratic hedging can be implemented using dynamic programming. Begin by initializing V_T as the payoff of the option. Next, compute V_{t-1} and $\phi_{t-1,1}$ recursively for $t = T, \ldots, 1$ as

$$\phi_{t-1,1} = \frac{A_t}{D_t}, \tag{3.30}$$

where

$$A_t = e^{-r}\, \mathbb{E}[V_t \Delta_t | \mathcal{F}_{t-1}] - e^{-r}\, \mathbb{E}[V_t | \mathcal{F}_{t-1}]\, \mathbb{E}[\Delta_t | \mathcal{F}_{t-1}], \tag{3.31}$$

$$D_t = \mathbb{E}[\Delta_t^2 | \mathcal{F}_{t-1}] - (\mathbb{E}[\Delta_t | \mathcal{F}_{t-1}])^2, \tag{3.32}$$

$$V_{t-1} = e^{-r}\, \mathbb{E}[V_t | \mathcal{F}_{t-1}] - \frac{A_t}{D_t}\, \mathbb{E}[\Delta_t | \mathcal{F}_{t-1}]. \tag{3.33}$$

Begin by computing A_T using $V_T = P$, and calculate also D_T. These values will be used to calculate V_{T-1} and $\phi_{T-1,1}$. V_{T-1} will be used to calculate A_{T-1} which is used for V_{T-2} and $\phi_{T-2,1}$. This algorithm continues recursively for $t = T, \ldots, 1$.

3.4.2 Global approach

Global hedging involves finding a self-financing strategy ϕ, and initial value V_0 that minimizes the squared expected difference between the payoff of the derivative, and the value of the portfolio. Schweizer (1995) presented a theoretical solution to the global quadratic hedging problem. This solution was efficiently implemented by Bertsimas, Kogan, and Lo (2001) and Černỳ (2004) using dynamic programming. If we have a derivative paying P at time T, with the present time being 0, the optimal global hedging challenge entails finding the initial investment V_0, and the sequence of investment quantities $\phi_{t-1,1}, t = 1, ..., T$ that minimizes:

$$\min_{\{\phi_0, \ldots, \phi_{T-1}\} \in \Phi} \mathbb{E}[\beta_T^{-2} (P - W_T)^2 | \mathcal{F}_0].$$

The recursive formula for $\phi_{t-1,1}$ is given by Equation 4.8 of Föllmer and Schweizer (1988):

$$\phi_{t-1,1} = \frac{Cov\left(\beta_T^{-1}P - \sum_{i=t}^{T}\phi_{i,1}\Delta X_{i+1},\ \Delta X_t|\mathcal{F}_{t-1}\right)}{Var[\Delta X_t|\mathcal{F}_{t-1}]}. \tag{3.34}$$

For the rest of the section, we assume $r = 0$. In this case, we can simplify the above expression using the martingale property of X_t, we have

$$E[\Delta X_t|\mathcal{F}_{t-1}] = 0.$$

The linearity of covariance makes the numerator of (3.34)

$$Cov\left(P - \sum_{i=t}^{T}\phi_{i,1}\Delta X_{i+1},\ \Delta X_t|\mathcal{F}_{t-1}\right) = Cov(P, \Delta X_t|\mathcal{F}_{t-1}) - \sum_{i=t}^{T}\phi_{i,1}Cov(\Delta X_{i+1}, \Delta X_t|\mathcal{F}_{t-1}). \tag{3.35}$$

Considering the orthogonal increments of a martingale, for any $i \neq t$, the increments ΔX_i and ΔX_t are uncorrelated. Thus, $Cov(\Delta X_{i+1}, \Delta X_t|\mathcal{F}_{t-1}) = 0$ for $i \neq t$. Hence, (3.35) becomes:

$$Cov\left(P - \sum_{i=t}^{T}\phi_{i,1}\Delta X_{i+1},\ \Delta X_t|\mathcal{F}_{t-1}\right) = Cov(P, \Delta X_t|\mathcal{F}_{t-1}). \tag{3.36}$$

Combining the derived results, we get a similar result as the one-period case:

$$\phi_{t-1,1} = \frac{Cov\left(P,\ \Delta X_t|\mathcal{F}_{t-1}\right)}{Var[\Delta X_t|\mathcal{F}_{t-1}]}. \tag{3.37}$$

From (3.23), the value is given as

$$V_{t-1} = \mathbb{E}[P - \phi_{t-1,1}\Delta X_t|\mathcal{F}_{t-1}]. \tag{3.38}$$

However, due to the martingale property of X_t, the value process can be expressed as $V_t = \mathbb{E}[P|\mathcal{F}_t]$ for $t = 0\ldots, T$. The expression (3.37) involves the conditional expectation and co-variances with respect to X_t instead of X_0. But for the non-causal AR(1) Cauchy model, the conditional distribution of X_{t+h} given X_t belongs to the same parametric family for any value of h [see Proposition 4 of Gouriéroux and Zakoïan (2017)].

The recursive formula for ϕ_t aligns with the goal of minimizing $\mathbb{E}[(P - W_T)^2|\mathcal{F}_0]$ by adjusting our hedge at each time point. This ensures the hedging strategy remains adaptive to market changes, thus maintaining minimized total hedging error.

Schweizer (1995) found that the local and global quadratic hedging strategies are equivalent when discounted asset prices are assumed to be martingales. We derive that the two strategies are also equivalent for the noncausal process. That is, from (3.37) we have for $t = 1, \ldots, T$ that

$$\begin{aligned}
\phi_{t-1,1} &= \frac{\mathbb{E}[P(X_t - X_{t-1})|\mathcal{F}_{t-1}] - \mathbb{E}[P|\mathcal{F}_{t-1}]\,\mathbb{E}[X_t - X_{t-1}|F_{t-1}]}{Var[X_t - X_{t-1}|\mathcal{F}_{t-1}]} \\
&= \frac{\mathbb{E}\left[\mathbb{E}[P(X_t - X_{t-1})|\mathcal{F}_t]|\mathcal{F}_{t-1}\right] - \mathbb{E}[\mathbb{E}[P|\mathcal{F}_t]|\mathcal{F}_{t-1}]\,\mathbb{E}[X_t - X_{t-1}|F_{t-1}]}{Var[X_t|\mathcal{F}_{t-1}]} \\
&= \frac{\mathbb{E}\left[V_t(X_t - X_{t-1})|\mathcal{F}_{t-1}\right] - \mathbb{E}[V_t|\mathcal{F}_{t-1}]\,\mathbb{E}[X_t - X_{t-1}|F_{t-1}]}{Var[X_t|\mathcal{F}_{t-1}]} \\
&= \frac{Cov[V_t,\, (X_t - X_{t-1})|F_{t-1}]}{Var[X_t|\mathcal{F}_{t-1}]},
\end{aligned} \tag{3.39}$$

which is equal to (3.28) when $r = 0$.

Therefore the two approaches are indeed equivalent, and the dynamic programming algorithm presented for the local approach could be used for the global. Even though these two approaches are equivalent, depending on how they are implemented numerically, they might yield different results.

Chapter 4

Numerical Analysis

As discussed in the previous chapter, quadratic hedging seeks strategies that reduce the total expected squared hedging errors. Whether we take a broad view (global) or a myopic view (local). We begin the analysis by recalling some notations. $\phi_{0,1}^{Bin}$, $\phi_{0,1}^{VH}$, and $\phi_{0,1}^{GH}$ are the initial quantities of the risky assets to invest using the binomial model, local quadratic hedging, and global quadratic hedging approach respectively. Similarly, V_0^{Bin}, V_0^{LH} and V_0^{GH} represent the value of the portfolio at time zero using the three different approaches. h will be used to represent the number of trading steps between time 0 to maturity T. Therefore, for a one-year maturity option, $h = 12$ means that trading occurs monthly and $h = 1$ means trading occurs once a year. K is the strike price. Unless stated, consider option values to be an at-the-money European call option (i.e. $K = X_0$), maturity (T) of one year, $h = 12$, $r = 0$, $X_0 = 250$ and $\rho = 0.978$. Next, the local and global quadratic hedging strategies are implemented for the non-causal process.

Algorithm 1 presents a dynamic programming approach for the local quadratic hedging of European-type options. It computes the optimal hedging strategy over discretized time and value nodes, capturing the dynamics of the underlying asset. It begins by constructing a discretized state space, which represents different prices of the underlying asset. For each node in this space, the algorithm calculates the corresponding probabilities based on the proposed rectangular filtration, by integrating the asset's probability distribution function. Due to the martingale property of the asset process, the proposed rectangular filtration is used in order to capture the mass at zero. Subsequently, the algorithm enters a dynamic programming loop, iterating over time steps and nodes. In each iteration, the algorithm computes the hedging strategy quantity and value for each node then utilizes the "solnp" function, a nonlinear optimization solver, to minimize the squared hedging error, and determines the optimal hedging parameters at each node and time step. Finally,

Algorithm 1 Local Quadratic Hedging

Given: $X_0, K, T, h, n, r, \sigma, \rho$, where n represents the dimension of the rectangular filtration.
Construct discretized state space for the stock price ($X[i]$) and probabilities:
for $i = 1$ **to** n **do**
    $X[i] = \text{Minimize } \left(F(x|X_0) - \frac{i}{n+1}\right)^2$
    for $j = 1$ **to** n **do**
        Calculate probabilities $p_j(X[i])$ for each node:
        if $i = 1$ **then**
            $p_j(X[1]) \leftarrow \int_{-\infty}^{X[1]} f(x|X_0)dx$
        else if $i = n$ **then**
            $p_j(X[n]) \leftarrow \int_{X[n-1]}^{\infty} f(x|X_0)dx$
        else
            $p_j(X[i]) \leftarrow \int_{X[i-1]}^{X[i]} f(x|X_0)dx$
        end if
    end for
end for
Calculate the hedging strategy ϕ_t and portfolio value V_t:
Set $V_T = P$
for $t = T \times h$ **to** 1 **do**
    for $i = 1$ **to** n **do**
        Compute $V[i]$ for the current node and time step
        Update probabilities for the current node and time step
        Define the loss function $L(\phi_t[i]) \leftarrow \sum_{j=1}^{n} p_j(X[i]) \cdot (V[j] - \phi_{t,1}[i] \cdot X[j] - \phi_{t,2}[i] \cdot \beta_1^{-1})^2$
        Minimize $L(\phi_t[i])$ to find optimal values of $\phi_{t,1}[i]$ and $\phi_{t,2}[i]$
    end for
end for
Compute the initial portfolio value as: $V_0 = \phi_{0,1} \cdot X_0 + \phi_{0,2}$
Output: V_0 and $\phi_{0,1}$

the algorithm outputs the portfolio value and hedging strategy at the initial time step.

Algorithm 2 Global Quadratic Hedging

Given: $X_0, K, T, h, n, r, \sigma, \rho, nsimul$, where n represents the number of nodes at each time step.
Define weight parameters w_1 and w_2 to be used for a mixed proposal density, where $w_1 \leftarrow 1-\rho$, $w_2 \leftarrow \rho$
Setup of Distributions:
Define a target density function $f(x)$, which again is $f(X_t|X_{t-1})$ in our case.
Define a proposal density $g(x)$ as a mixture of two cauchy distributions

$$g(x) \leftarrow w_1 \times \frac{1}{\pi}\frac{\sigma}{(x-0)^2+\sigma^2} + w_2 \times \frac{1}{\pi}\frac{\sigma}{(x-\frac{X_t}{\rho})^2+\sigma^2}$$

Compute Multiplier c **such that:**
$c \leftarrow \max\left(\frac{f(x)}{g(x)}\right)$
Simulation of Stock Prices:
for $t = 0, \ldots, T$ **do**
    for $s = 1$ **to** $nsimul$ **do**
        Simulate stock price using acceptance-rejection:
        repeat
            Generate sample X from $g(X)$
            Accept X with probability $\frac{f(X)}{c \cdot g(X)}$
        until X_t samples are accepted
    end for
end for
Compute initial value $V_0 \leftarrow \mathbb{E}[P]$
Compute $\phi_{0,1} \leftarrow \frac{Cov(P, X_1 - X_0)}{\mathbb{E}(X_1 - X_0)^2}$
Output: V_0 and $\phi_{0,1}$

To compute the global hedging strategy, we see from (3.37) that, an expression for the covariance term is needed. However, a closed-form expression for the covariance is not available, so we use simulations. In algorithm 2, the acceptance-rejection simulation method described in Section 2.5.3 is used to simulate $100,000$ asset prices from the noncausal model, and these values are used to compute the covariance and variance terms of (3.37). The target distribution for the simulation is the noncausal density $f(X_t|X_{t-1})$, while the proposal distribution chosen is a mixture of two Cauchy distributions described in Section 2.5.3. The multiplier constant c was calculated as the maximum ratio of the target to the proposal function. After simulating the asset prices, the initial portfolio value V_0 and hedging strategy (ϕ_0) is computed for a given strike price using (3.34).

4.1 One-period Options

In this section, we put the hedging strategy into practice using a one-period European vanilla call and put options with payoffs of $P = (X_1 - K)^+$ and $P = (K - X_1)^+$, respectively. We next observe how the cumu-

lative cost at time 1 is distributed and how hedging errors change for various strikes K at given X_0. Since the local and global quadratic hedging strategies are equivalent in this setting, the hedging strategy for just the global is calculated and compared with the binomial. The local case is implemented in the next section, which involves multi-periods.

Table 4.1: $\phi_{0,1}$ and V_0 values for $\rho = 0.978$ and different K's.

K	$0.8X_0$	$0.85X_0$	$0.9X_0$	$0.95X_0$	X_0	$1.02X_0$	$1.05X_0$	$1.1X_0$	$1.15X_0$	$1.2X_0$
European Call										
$\phi_{0,1}^{GH}$	0.2814	0.2344	0.1865	0.1373	0.0866	0.0666	0.0558	0.0488	0.0443	0.0409
$\phi_{0,1}^{Bin}$	0.2176	0.1687	0.1198	0.0709	0.0220	0.0024	0.0000	0.0000	0.0000	0.0000
V_0^{GH}	54.1313	41.9894	29.8872	17.8574	6.0637	1.8435	0.5945	0.3326	0.2339	0.1775
V_0^{Bin}	54.4000	42.1750	29.9500	17.7250	5.5000	0.6100	0.0000	0.0000	0.0000	0.0000
European Put										
$\phi_{0,1}^{GH}$	-0.7186	-0.7656	-0.8135	-0.8627	-0.9134	-0.9334	-0.9442	-0.9511	-0.9557	-0.9591
$\phi_{0,1}^{Bin}$	-0.7824	-0.8313	-0.8802	-0.9291	-0.9780	-0.9976	-1.0000	-1.0000	-1.0000	-1.0000
V_0^{GH}	3.9518	4.3099	4.7076	5.1778	5.8842	6.6640	12.9149	25.1531	37.5544	49.9979
V_0^{Bin}	4.4000	4.6750	4.9500	5.2250	5.5000	5.6100	12.5000	25.0000	37.5000	50.0000

Table 4.2: $\phi_{0,1}$ and V_0 values for $\rho = 0.5$ and different K's.

K	$0.8X_0$	$0.85X_0$	$0.9X_0$	$0.95X_0$	X_0	$1.02X_0$	$1.05X_0$	$1.1X_0$	$1.15X_0$	$1.2X_0$
European Call										
$\phi_{0,1}^{GH}$	0.6003	0.5755	0.5506	0.5258	0.5010	0.4910	0.4761	0.4513	0.4265	0.4016
$\phi_{0,1}^{Bin}$	0.6000	0.5750	0.5500	0.5250	0.5000	0.4900	0.4750	0.4500	0.4250	0.4000
V_0^{GH}	150.2153	143.9231	137.6345	131.3495	125.0685	122.5569	118.7904	112.5151	106.2425	99.9732
V_0^{Bin}	150.0000	143.7500	137.5000	131.2500	125.0000	122.5000	118.7500	112.5000	106.2500	100.0000
European Put										
$\phi_{0,1}^{GH}$	-0.3997	-0.4245	-0.4493	-0.4742	-0.4990	-0.5089	-0.5239	-0.5487	-0.5735	-0.5984
$\phi_{0,1}^{Bin}$	-0.4000	-0.4250	-0.4500	-0.4750	-0.5000	-0.5100	-0.5250	-0.5500	-0.5750	-0.6000
V_0^{GH}	98.9417	105.1496	111.3610	117.5760	123.7949	126.2834	130.0169	136.2416	142.4689	148.6997
V_0^{Bin}	100.0000	106.2500	112.5000	118.7500	125.0000	127.5000	131.2500	137.5000	143.7500	150.0000

From Tables 4.1 and 4.2, we can see that the quantity needed to be longed, and the initial portfolio value decreases as the call option becomes more and more out of the money. Especially for the Binomial model whose values become zero quickly after the option becomes out of the money. The put case is different, though. In this case, as the strike price rises, so do the amount to be shorted and the portfolio value. Comparing the two methods when $\rho = 0.978$, the portfolio value of the quadratic approach tends to be cheaper when the option is in the money and becomes expensive than the binomial when it is about, or out of the money. When ρ decreases to 0.5, the two approaches a very similar except when the option is very out of, or in the money.

4.2 Local Hedging Convergence

Table 4.3: V_0^{LH} for different steps and nodes.

Nodes	h					
	1	2	3	4	6	12
200	54.4580	53.5727	52.3141	50.6684	46.1158	30.7085
500	54.7711	54.3826	53.9193	53.4569	51.1996	39.6751
1000	54.9151	54.6950	54.4797	54.1651	53.2858	47.4838
1500	54.9740	54.8158	54.6621	54.4389	53.9302	50.4397

Table 4.3 demonstrates how the hedging strategy values for the local quadratic hedging approach behave for different steps and number of nodes. These values are compared with the one year maturity hedging strategy value for the global approach, which is 55.6753, since varying h for does not impact the V_0. It can be seen that, the local hedging strategy value converges to the global when the number of nodes is increased. This observation particularly stands out when h remains constant and the nodes increase. For instance, if we fix $h = 3$, the value refines from 52.3141 with 200 nodes to 54.4797 with 1000 nodes. This behaviour suggests that a more granular node count could lead to a more accurate representation in our model. However, increasing the number of nodes for higher h value increases the computation time significantly. Therefore, we will use 1000 nodes for the rest of our analysis.

4.3 Comparing Strike Prices for Various Hedging Methods

Table 4.4: $\phi_{0,1}$ and V_0 values for different K's.

K	$0.8X_0$	$0.85X_0$	$0.9X_0$	$0.95X_0$	X_0	$1.02X_0$	$1.05X_0$	$1.1X_0$	$1.15X_0$	$1.2X_0$
European Call										
$\phi_{0,1}^{Bin}$	0.3874	0.3491	0.3109	0.2726	0.2343	0.2190	0.1960	0.1577	0.1194	0.0811
$\phi_{0,1}^{LH}$	0.3872	0.3470	0.3067	0.2663	0.2258	0.2096	0.1853	0.1450	0.1051	0.0664
$\phi_{0,1}^{GH}$	0.4300	0.3958	0.3615	0.3270	0.2924	0.2786	0.2578	0.2234	0.1894	0.1562
V_0^{Bin}	96.8573	87.2859	77.7145	68.1431	58.5717	54.7431	49.0002	39.4288	29.8574	20.2860
V_0^{LH}	85.6299	75.9346	66.3347	56.8436	47.4838	43.7826	38.2841	29.2911	20.5878	12.3891
V_0^{GH}	92.6620	83.2533	73.9464	64.7474	55.6753	52.0896	46.7621	38.0499	29.6150	21.6132
European Put										
$\phi_{0,1}^{Bin}$	-0.6126	-0.6509	-0.6891	-0.7274	-0.7657	-0.7810	-0.8040	-0.8423	-0.8806	-0.9189
$\phi_{0,1}^{LH}$	-0.6128	-0.6530	-0.6933	-0.7337	-0.7742	-0.7904	-0.8147	-0.8550	-0.8949	-0.9336
$\phi_{0,1}^{GH}$	-0.5944	-0.6286	-0.6629	-0.6974	-0.7320	-0.7458	-0.7666	-0.8010	-0.8350	-0.8682
V_0^{Bin}	46.8573	49.7859	52.7145	55.6431	58.5716	59.7431	61.5002	64.4288	67.3574	70.2860
V_0^{LH}	35.6299	38.4346	41.3347	44.3436	47.4837	48.7826	50.7841	54.2911	58.0878	62.3891
V_0^{GH}	41.3962	44.4876	47.6807	50.9817	54.4096	55.8239	57.9964	61.7841	65.8492	70.3474

Table 4.4 presents a comparison across different strike prices relative to X_0 for the three approaches when $\rho = 0.978$. The hedging strategies, represented by $\phi_{0,1}$, differ across all methods. For instance, at $K = 0.8X_0$, the global hedging exhibits a notably higher $\phi_{0,1}$ than either binomial or local hedging, suggesting a more

aggressive strategy under the global method for this particular strike price. A closer look reveals that as the strike price moves closer to X_0, the differences between the $\phi_{0,1}$ values of the different approaches tend to decrease, indicating a similarity in hedging strategies as the option becomes at the money. Regarding the hedging strategy values V_0, for both call and put options, the value of the binomial approach is higher than the quadratic approaches with high differences observed when the option because more in or out of the money.

4.4 Impact of ρ

Table 4.5: $\phi_{0,1}$ and V_0 values for different values of ρ.

ρ	0.5	0.8	0.978
European Call			
$\phi_{0,1}^{Bin}$	0.9998	0.9313	0.2343
$\phi_{0,1}^{LH}$	0.5960	0.3980	0.2258
$\phi_{0,1}^{GH}$	0.9529	0.9345	0.2924
V_0^{Bin}	249.9390	232.8201	58.5717
V_0^{LH}	151.5789	100.4808	47.4837
V_0^{GH}	238.6421	232.4055	55.6753

ρ	0.5	0.8	0.978
European Put			
$\phi_{0,1}^{Bin}$	-0.0002	-0.0687	-0.7657
$\phi_{0,1}^{LH}$	-0.4040	-0.6020	-0.7742
$\phi_{0,1}^{GH}$	0.0003	-0.0734	-0.7320
V_0^{Bin}	249.9390	232.8201	58.5716
V_0^{LH}	151.5789	100.4808	47.4837
V_0^{GH}	251.9637	233.2690	54.4096

Table 4.5 explores how the three hedging approaches respond to different values of the ρ parameter, for at-the-money European Call and Put options. The results show differences in both trading strategies and initial portfolio values as ρ changes, with noticeable distinctions across methods and option types. The global quadratic hedging's behaviour varies between European Call and Put options, highlighting the interplay between option type, hedging strategy, and the underlying ρ parameter. It can be seen that the values for the three approaches seem to be reasonably close when $\rho = 0.978$ than other values of ρ. However, recall from the one-period case that, the values were rather similar for $\rho = 0.5$ than they were for 0.978. This suggests that when the frequency of trading is increased, a bigger ρ will produce more accurate results.

4.5 Impact of X_0

Table 4.6: $\phi_{0,1}$ and V_0 values for different values of X_0.

X_0	250	350	500	1000
	European Call			
$\phi_{0,1}^{Bin}$	0.2343	0.2343	0.2343	0.2343
$\phi_{0,1}^{LH}$	0.2258	0.2044	0.1749	0.1274
$\phi_{0,1}^{GH}$	0.2924	0.2772	0.2709	0.2481
V_0^{Bin}	58.5717	82.0003	117.1433	234.2866
V_0^{LH}	47.4838	63.1933	80.0357	121.1883
V_0^{GH}	55.6753	78.7166	113.7171	229.5554

X_0	250	350	500	1000
	European Put			
$\phi_{0,1}^{Bin}$	-0.7657	-0.7657	-0.7657	-0.7657
$\phi_{0,1}^{LH}$	-0.7742	-0.7956	-0.8251	-0.8726
$\phi_{0,1}^{GH}$	-0.7320	-0.7333	-0.7458	-0.7474
V_0^{Bin}	58.5716	82.0003	117.1433	234.2866
V_0^{LH}	47.4837	63.1932	80.0356	121.1882
V_0^{GH}	54.4096	76.8738	111.5528	230.2473

Table 4.6 provides insights into the relationship between various hedging strategies for European call and put options, alongside portfolio values for different initial stock values X_0. For the call case, it can be seen that $\phi_{0,1}^{Bin}$ remains unchanged across all X_0 values. In contrast, $\phi_{0,1}^{LH}$ decreases as X_0 increases, whereas the decrement in $\phi_{0,1}^{GH}$ is comparatively moderate. With respect to the hedging strategy values, all methods exhibit an increasing trend with an increasing X_0. Similar behaviour can be observed in the put case. The margin of difference between the binomial and global approach seems to be consistent across the X_0 values. The difference in the local on the other hand becomes more pronounced for bigger X_0 values. Therefore, when a larger X_0 is chosen, the number of nodes in the numerical implementation of the local quadratic hedging needs to be increased for accurate results.

4.6 Impact of T

Table 4.7: V_0 values for different maturity months.

	T (in months) $= h$							
	1	2	3	4	6	12	18	24
V_0^{Bin}	5.5000	10.8790	16.1397	21.2846	31.2374	58.5717	82.4905	103.4207
V_0^{LH}	5.9766	11.2977	16.2780	20.9850	29.6601	47.4838	57.0583	60.8795
V_0^{GH}	6.0637	11.5623	16.7114	21.6490	30.9075	55.6753	77.5285	97.1838

Table 4.7 provides insights into the behaviour of three strategies as the maturity varies. In this table, the portfolio values of each approach are observed for an equal number of steps (in months) and time to maturity (also in months). For instance, 4 represents a four-month maturity option with four trading months. It can be seen that as T increases, the values for all approaches also increase, which is expected given the greater uncertainty associated with longer durations. When $T = 3$, the values for all three approaches are closer. After, when T is more than 4 months, the value of the binomial becomes higher than the quadratic approach.

4.7 Local Hedging Strategy Values for Different h's and T's

Table 4.8: V_0 values for different steps and maturities.

T (in years)	V_0^{LH} for different values of h						V_0^{GH}
	1	2	3	4	6	12	
1	54.9151	54.6950	54.4797	54.1651	53.2858	47.4838	55.6753
2	93.8443	92.6463	91.6311	88.0435	75.4782	56.0505	97.1838
3	123.6416	117.5850	108.4273	99.2819	81.9881	59.2179	133.4878

In Table 4.8, the V_0^{LH} values are tabulated for varying maturities (in years) and step counts (in months). The V_0^{GH} and V_0^{Bin} values are also presented for different maturity years. For a given maturity, there is a noticeable decrease in V_0^{LH} as the step count rises, indicating that less frequent adjustments are associated with reduced initial portfolio value. Concurrently, for a fixed step count, an increase in maturity generally corresponds to a rise in V_0^{LH}, highlighting the time value of option values. This behaviour is observed for the V_0^{GH} values too. Comparing the local and global portfolio values, it is observed that when $h = 1$, the V_0^{LH} values are closest to the V_0^{GH} values for all maturity years. As the frequency of adjustments increases, the local portfolio values become more different and smaller compared with the global values. This is because, increasing the frequency of adjustments reduces the risk of the portfolio, thereby decreasing its value.

The discussion in this chapter offers a perspective on the behaviour of different hedging strategies across various parameters. It can be seen that selecting the optimal strategy requires a balance between precision (as seen with the choice of node count) and the inherent behaviour of the strategy across different maturities and strike prices.

Chapter 5

Conclusion

This research explored quadratic hedging using the noncausal AR(1) Cauchy model by Gouriéroux and Zakoïan (2017). The noncausal AR(1) Cauchy model effectively preserves the martingale property of cryptocurrency prices as implied by Schilling and Uhlig (2019). This indicates that within our model, future price movements remain uninfluenced by past price dynamics, aligning well with the non-arbitrage principle in financial mathematics. An advantage of our model is its capacity to emulate the bubble phenomena, a key feature of the cryptocurrency markets. We showed that, when X_{t-1} goes to infinity, the limiting distribution of the re-scaled noncausal AR(1) process (2.17) approaches the binomial distribution with masses at 0 and $1/\rho$ and probabilities $1-\rho$ and ρ, respectively.

This thesis implemented two quadratic hedging approaches, local and global, and compared them with the binomial model. For the quadratic hedging, it was observed that the convergence of the local hedging approach towards the global method, as evidenced in Table 4.3, is influenced by the number of steps and nodes in the model. Tables 4.1, 4.2 and 4.5 reveal that the ρ parameter plays a role in yielding accurate results, especially in frequent trading scenarios where a higher ρ tends to yield more accurate results. The sensitivity of the hedging strategies and portfolio values to the initial stock value X_0, as shown in Table 4.6, and the influence of maturity T on option pricing, demonstrated in Tables 4.7 and 4.8, should be considered when implementing the quadratic hedging strategies.

The impact of various parameters when using the noncausal AR(1) Cauchy discussed in this thesis highlights the importance of a tailored approach in derivative trading, especially in complex and dynamic markets like cryptocurrencies. The results extend a broader understanding of hedging strategies and contribute to the

field of mathematical finance. Future research could include the incorporation of real-world cryptocurrency market data to compare with the findings in this thesis.

Appendix A

Stable Distribution

A random variable has a stable distribution, also known as α-stable distribution if the distribution is such that, up to location and scale parameters, the distribution of a linear combination of two independent random variables with this distribution is the same. The stable distribution has four parameters: the stability parameter $\alpha \in (0, 2]$, the skewness or asymmetry parameter $\beta \in [-1, 1]$, the scale parameter $\sigma \in (0, \infty)$, and the location parameter $\mu \in (-\infty, \infty)$. Therefore, we use $X \sim \mathcal{S}(\alpha, \beta, \sigma, \mu)$ to denote a random variable X that follows a stable distribution.

Whereas the probability density function for a stable distribution cannot be defined analytically, the general characteristic function is given as:

$$\psi(s) = e^{(i\mu s \, - \, \sigma^\alpha |s|^\alpha \{1 \, - \, i\beta(\text{sign } s)\Phi\})}, \tag{A.1}$$

$$\log \psi(s) = i\mu s \; - \; \sigma^\alpha |s|^\alpha \{1 \; - \; i\beta(\text{sign } s)\Phi\},$$

given that

$$\Phi = \begin{cases} \tan(\frac{\pi\alpha}{2}) & \text{if } \alpha \neq 1 \\ -\frac{2}{\pi}\log|s| & \text{if } \alpha = 1 \end{cases}$$

and sign(t) is equal to the sign of the constant t.

From (A.1), we can see that when $\alpha = 2$ and $\beta = 0$, we get the characteristic function of a Normal distribution $N(\mu, \sigma^2)$. On the other hand, when $\alpha = 1$ and β remains 0, we get the Cauchy distribution $\mathcal{C}(\mu, \sigma)$.

Appendix B

Complementary Results

This section contains additional results and figures that supplement the main text's findings. We only consider the initial value $X_0 = 100$ for illustration purposes. Figure B.1 shows the density curves for $100,000$ simulations from the noncausal AR(1) process using the acceptance-rejection simulation method, and a single Cauchy distribution centred at X_0 as the proposal density for each $\rho \in \{0.978, 0.8, 0.5\}$.

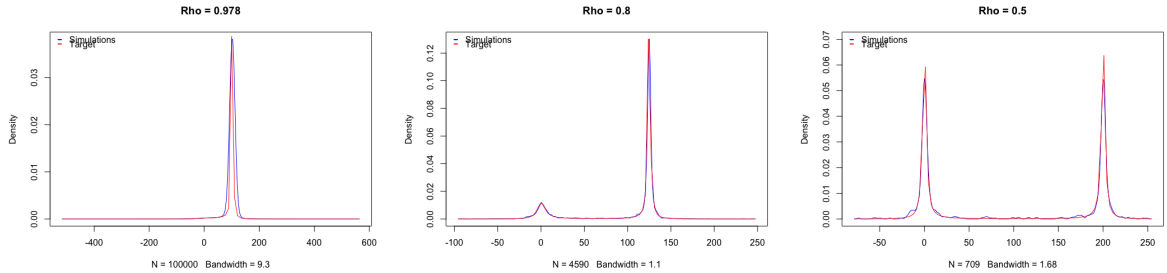

Figure B.1: Density curves for simulated (blue) and target (red) Cauchy AR(1) model for different ρ's.

The table below presents the simulation acceptance rates using the single and mixed Cauchy distribution.

ρ	0.978	0.8	0.5
Mixed Cauchy			
Acceptance rate	0.9606675	0.843947	0.781313
Single Cauchy			
Acceptance rate	0.4734415	0.4734415	0.0267495

References

Augustyniak, M., Godin, F., & Simard, C. (2017). Assessing the effectiveness of local and global quadratic hedging under garch models. *Quantitative Finance*, *17*(9), 1305–1318.

Bertsimas, D., Kogan, L., & Lo, A. W. (2001). Hedging derivative securities and incomplete markets: An ϵ-arbitrage approach. *Operations research*, *49*(3), 372–397.

Black, F., & Scholes, M. (1973). The valuation of options and corporate liabilities. *Journal of political economy*, *81*(3), 637–654.

Breidt, et al. (1992). Time-reversibility, identifiability and independence of innovations for stationary time series. *Journal of Time Series Analysis*, *13*(5), 377–390.

Černỳ, A. (2004). Dynamic programming and mean-variance hedging in discrete time. *Applied Mathematical Finance*, *11*(1), 1–25.

Černỳ, A., & Kallsen, J. (2007). On the structure of general mean-variance hedging strategies. *Annals of probability*, 1479–1531.

Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of financial Economics*, *7*(3), 229–263.

Davis, R. A., & Song, L. (2020). Noncausal vector ar processes with application to economic time series. *Journal of Econometrics*, *216*(1), 246–267.

Feller, W. (1991). *An introduction to probability theory and its applications, volume 2* (Vol. 81). John Wiley & Sons.

Föllmer, H., & Schweizer, M. (1988). Hedging by sequential regression: An introduction to the mathematics of option trading. *ASTIN Bulletin: The Journal of the IAA*, *18*(2), 147–160.

Föllmer, H., & Sondermann, D. (1986). Hedging of non-redundant contingent claims. *Contributions to Mathematical Economics: Essays in Honor of G. Debreu*, 205–223.

Fries, S., & Zakoian, J.-M. (2019). Mixed causal-noncausal ar processes and the modelling of explosive bubbles. *Econometric Theory*, *35*(6), 1234–1270.

Garcia, D., Tessone, C. J., Mavrodiev, P., & Perony, N. (2014). The digital traces of bubbles: feedback

cycles between socio-economic signals in the bitcoin economy. *Journal of the Royal Society Interface*, *11*(99), 20140623.

Gourieroux, C., Hencic, A., & Jasiak, J. (2021). Forecast performance and bubble analysis in noncausal mar (1, 1) processes. *Journal of Forecasting*, *40*(2), 301–326.

Gourieroux, C., & Jasiak, J. (2016). Filtering, prediction and simulation methods for noncausal processes. *Journal of Time Series Analysis*, *37*(3), 405–430.

Gouriéroux, C., & Lu, Y. (2023). Noncausal affine processes with applications to derivative pricing. *Mathematical Finance*, *33*(3), 766–796.

Gouriéroux, C., & Zakoïan, J.-M. (2017). Local explosion modelling by non-causal process. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *79*(3), 737–756.

Hamilton, J. D. (2020). *Time series analysis*. Princeton university press.

Hayes, A. S. (2017). Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin. *Telematics and informatics*, *34*(7), 1308–1321.

Hencic, A., & Gouriéroux, C. (2015). Noncausal autoregressive model in application to bitcoin/usd exchange rates. In *Econometrics of risk* (pp. 17–40). Springer.

Hou, et al. (2020). Pricing cryptocurrency options. *Journal of Financial Econometrics*, *18*(2), 250–279.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous univariate distributions, volume 2* (Vol. 289). John wiley & sons.

Kristoufek, L. (2015). What are the main drivers of the bitcoin price? evidence from wavelet coherence analysis. *PloS one*, *10*(4), e0123923.

Lanne, et al. (2011). Noncausal autoregressions for economic time series. *Journal of Time Series Econometrics*, *3*(3).

Matic, J. L., Packham, N., & Härdle, W. K. (2023). Hedging cryptocurrency options. *Review of Derivatives Research*, *26*(1), 91–133.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of economics and management science*, 141–183.

Rémillard, B., & Rubenthaler, S. (2013). Optimal hedging in discrete time. *Quantitative Finance*, *13*(6), 819–825.

Schilling, L., & Uhlig, H. (2019). Some simple bitcoin economics. *Journal of Monetary Economics*, *106*, 16–26.

Schweizer, M. (1988). *Hedging of options in a general semimartingale model* (Unpublished doctoral dissertation). ETH Zurich.

Schweizer, M. (1991). Option hedging for semimartingales. *Stochastic processes and their Applications*,

37(2), 339–363.

Schweizer, M. (1995). Variance-optimal hedging in discrete time. *Mathematics of Operations Research*, *20*(1), 1–32.

Shreve, S. (2005). *Stochastic calculus for finance i: the binomial asset pricing model*. Springer Science & Business Media.

Yiying, W., & Yeze, Z. (2019). Cryptocurrency price analysis with artificial intelligence. In *2019 5th international conference on information management (icim)* (pp. 97–101).