Examining the Impact of Estimation Window Size on the Applicability of Value-at-Risk and Expected Shortfall in a Stressed Cryptocurrency Market

Lappeenranta-Lahti University of Technology LUT

Master's Programme in Strategic Finance, Master's Thesis

2024

Eemeli Leponiemi

Examiners: Post-Doctoral Researcher Jan Stoklasa

Associate Professor Sheraz Ahmed

Abstract

Lappeenranta-Lahti Univeristy of Technology LUT

LUT School of Business and Management

Business Administration

Eemeli Leponiemi

Examining the Impact of Estimation Window Size on the Applicability of Value-at-Risk and Expected Shortfall in a Stressed Cryptocurrency Market

Master's thesis

2024

68 pages, 8 figures, 13 tables and 8 appendices

Examiners: Post-Doctoral Researcher Jan Stoklasa
Associate Professor Sheraz Ahmed

Keywords: Cryptocurrency; Bitcoin; Ethereum; Value-at-Risk; Expected Shortfall; Backtesting; Estimation window

During the past few years, both the stock and the cryptocurrency markets have faced significant market stress due to unforeseen events such as the energy crisis, uncertainties in monetary policies, the US banking crisis, and the burst of the speculative cryptocurrency bubble. These developments have increased market risk, especially associated with cryptocurrencies, which can cause fluctuations that challenge the accuracy of risk models. This thesis introduces different size estimation windows to forecast the market risk of prominent cryptocurrencies with VaR and ES. The accuracy of the model is evaluated using different backtest methods, including POF, CC, CCI, and UC tests. The results suggest that ES remains a robust risk measure, even when faced with increased market stress, particularly when the HS method and the Student's t GARCH(1,1) model are used. The study did not uncover any obstacles to using VaR models in the context of cryptocurrencies. Additionally, while different-sized estimation windows did not lead to significant variations in risk estimates, a 250-day estimation window generally appeared to be the most stable among various estimation methods when applied to different cryptocurrencies.

Tiivistelmä

Lappeenrannan–Lahden teknillinen yliopisto LUT

LUT-kauppakorkeakoulu

Kauppatieteet

Eemeli Leponiemi

Arviointi estimointi-ikkuna koon vaikutuksesta Value-at-Riskin ja odotetun vajeen soveltuvuuteen stressaantuneilla kryptovaluuttamarkkinoilla

Pro Gradu –tutkielma

2024

68 sivua, 8 kuvaa, 13 taulukkoa ja 8 liitettä

Tarkastajat: Tutkijatohtori Jan Stoklasa
Tutkijaopettaja Sheraz Ahmed

Avainsanat: Kryptovaluutta; Bitcoin; Ethereum; Value-at-Risk; Odotettu vaje; Toteumatestaus; Estimointi-ikkuna

Viime vuosina sekä osake- että kryptovaluuttamarkkinat ovat joutuneet kohtaamaan merkittävää markkinapainetta, joka on johtunut odottamattomista tapahtumista, kuten, rahapolitiikan epävarmuudesta, Yhdysvaltain pankkikriisistä ja spekulatiivisen kryptovaluuttakuplan puhkeamisesta. Nämä tapahtumat ovat lisänneet erityisesti kryptovaluuttoihin liittyvää markkinariskiä, joka voi aiheuttaa riskimallien tarkkuutta haastavia hinnan heilahteluja. Tässä tutkielmassa esitellään erikokoisia estimointi-ikkunoita, joiden avulla voidaan ennustaa tunnettujen kryptovaluuttojen markkinariskiä VaR:n ja ES:n avulla. Mallien tarkkuutta arvioidaan käyttämällä erilaisia toteumatestausmenetelmiä, kuten POF-, CC-, CCI- ja UC-testejä. Tulokset viittaavat siihen, että ES on edelleen vankka riskimittari, vaikka markkinastressi lisääntyisi, erityisesti kun käytetään HS-menetelmää ja Studentin t GARCH(1,1)-mallia. Tutkimuksessa ei myöskään havaittu esteitä VaR-mallien käytölle kryptovaluuttojen yhteydessä. Lisäksi vaikka erikokoiset estimointi-ikkunat eivät johtaneet merkittäviin eroihin riskiestimaateissa, 250 päivän estimointi-ikkuna vaikutti yleisesti ottaen vakaimmalta eri estimointimenetelmillä, kun sitä sovellettiin eri kryptovaluutoihin.

Symbols and abbreviations

Roman characters

D	Predictive distribution
F	Real distribution
n	Sample size
p	Probability
P	Price
R	Logarithmic return
t	Time
T	Student's t-distribution
U	Unknown distribution
v	Count of violations
S	Scoring function
X	Asset outcome
W_E	Estimation window
W_T	Testing window

Greek characters

α	Significance level
β	Coefficient of lagged conditional variance
γ	Coefficient of the lagged squared residuals
Γ	Gamma function
ϵ	Error term
η	Violation indicator
κ	Constant
μ	Mean
ν	Degrees of freedom
ρ	Risk measure
σ^2	Variance
σ	Volatility
σ_t	Conditional volatility
ϕ	Density function
Φ	Normal cumulative distribution

Abbreviations

BCBS The Basel Committee on Banking Supervision

BLCA Bitcoin-like crypto asset

CC Conditional Coverage

CCI Conditional Coverage Independence

DeFi Decentralized Finance

DLT Distributed Ledger Technology

ES Expected Shortfall

GARCH Generalized Autoregressive Conditional Heteroskedasticity

HS Historical Simulation

MiCA Markets in Crypto-Asset Regulation

NS Normalized Shortfall

POF Proportion of Failure

VaR Value-at-Risk

VCV Variance-Covariance

VR Violation Rate

Table of Contents

List of Figures

List of Tables

1 Introduction

“You’re worse off relying on misleading information than not having any information at all. If you give a pilot an altimeter that is sometimes defective, he will crash the plane. Give him nothing, and he will look out the window.” Said Taleb (1996), a statistician and risk specialist, while being interviewed about risk management and Value-at-Risk (VaR). As deficiencies within the VaR measure came to light, Artzner et al. (1999) proposed a more sophisticated alternative, Expected Shortfall (ES), aimed at addressing the unideal properties of VaR.

The challenge in risk management involves constructing models that consider infrequent but impactful events across market, credit, operational, and insurance risks (McNeil, 1999). The 2008 global financial crisis diminished trust in the conventional monetary system, leading to the emergence of digital currencies or cryptocurrencies (Dyhrberg, 2016). Cryptocurrencies, characterized by rapid growth and high volatility, attract significant attention from various stakeholders (Obeng 2021; Bruzgė et al. 2023; Trucíos and Taylor 2023). Beyond their intrinsic value, their prices are influenced by external factors such as market news and speculation (Osterrieder et al., 2016).

Recent years have witnessed significant market stress in both the stock and cryptocurrency markets due to unforeseen events, including uncertainty in monetary policies, the US banking crisis, and the burst of the speculative cryptocurrency bubble (Likitratcharoen et al., 2023). Notable episodes of market turmoil include the collapse of TerraUSD (UST) in May 2022 and the collapse of the FTX crypto platform in November 2022 (Cornelli et al., 2023). Financial turmoil and extreme observations can adversely affect the forecast performance of volatility models (Boudt et al. 2013; Hotta and Trucíos 2018), and also cause market risk measures to lose accuracy and potentially underestimate risk (Kourouma et al.; 2011; Mavani 2020).

Although traditional markets are regulated and risk measures are standard in financial institutions, as mandated by the Basel II and Basel III accords, cryptocurrency markets currently lack regulation, and the formal use of risk measures is not obligatory (Trucíos & Taylor, 2023). However, significant regulatory developments are on the horizon in the management of cryptocurrency market risk. At the end of 2022, the The Basel Committee on Banking Supervision (BCBS) released minimum capital requirements for cryptocurrencies, set to take effect at the beginning of 2025 (BCBS, 2022). Furthermore, the European Union for crypto assets has established the new regulation Markets in Crypto-Asset Regulation (MiCA) to assess the risk of the cryptocurrency market, with the implementation date contingent on the approval of the European Commission, the European Parliament, and the Council of the European Union (ESMA, 2023).

Despite the widespread use of VaR and ES for traditional financial assets, limited research on the applicability of VaR to cryptocurrencies can only be found in a few studies (see, e.g., Obeng 2021; Likitratchoaroen et al. 2021). Given the market turmoil experienced by cryptocurrencies in recent years, studies providing information on the performance of VaR and ES during increased market stress are almost non-existent (Likitratcharoen et al. 2023; Trucíos and Taylor 2023). Furthermore, the chosen estimation windows significantly impact risk estimates (Buczyński & Chlebus, 2022). The BCBS has mandated a 250-day estimation window to build risk models (BCBS, 2019b), yet no explicit justification has been given for this specific time frame. Additionally, there is a lack of knowledge regarding the impact of the size of the estimation window, even in the case of traditional assets (see, e.g., Righi and Ceretta 2015; Berens et al. 2018; Buczyński and Chlebus 2019), and this knowledge gap is even more present in the realm of cryptocurrencies.

1.1 Background and motivation

The increase in the market value of cryptocurrencies has also led to a tremendous surge in the number of new currencies. Between 2017 and 2023, the number of cryptocurrencies has increased by more than 3000% (Statista, 2023). This growing demand for cryptocurrencies has also spurred a growing interest in academic research focused on cryptocurrencies and discussions regarding legislative aspects. This is evident in publications such as those of Masciandaro (2018) and Castrén et al. (2022). The primary focus of the business revolves around topics such as its impact on banking, profitability, blockchain technology, the emergence of financial bubbles, and potential associations with illicit activities (Azarenkova et al., 2018). However, limited research has evaluated the efficacy of widely employed risk measures such as VaR and ES in managing market risk for cryptocurrencies. Given their extensive use in traditional financial assets and their acceptance by regulatory authorities, these risk measures have the potential to offer valuable information in the field of cryptocurrency risk management.

Likitratcharoen et al. (2018) studied the performance of VaR in the cryptocurrency market. VaR was estimated using non-parametric Historical Simulation (HS) and parametric method Variance-Covariance (VCV) under the assumption of normality and concluded that HS resulted more accurate risk estimates. (Liu et al., 2020) advanced approaches by implementing the conditional volatility model Generalized Autoregressive Conditional Heteroskedasticity (GARCH) as part of the measurement of the market risk of Bitcoin. Jiménez et al. (2020b) further researched the applicability of ES to Bitcoin by implementing conditional volatility models, and the study showed promising results in the performance of ES using different volatility models for Bitcoin. Because most studies have focused only on Bitcoin because it is the largest by market capitalization, Obeng (2021) advanced the usage of GARCH VaR models with different distributional assumptions from other cryptocurrencies and concluded that the EGARCH models performed best. Although these studies have offered valuable information on the applications of VaR and ES, they do not consider the impact of estimation windows on risk estimates.

During the last few years, unexpected events such as the COVID-19 pandemic, the energy crisis, and conflicts have adversely affected the stock markets (Açikgöz & Günay, 2020). It is particularly noteworthy that the cryptocurrency market has seen increased volatility, presenting superior returns compared to traditional assets (Lahmiri & Bekiros, 2020). Corbet et al. (2020) researched does Bitcoin provide diversification advantages during the COVID-19 pandemic. The findings suggested that instead of providing hedges or safe haven features in times of intense market pressure, crypto assets such as Bitcoin appeared to amplify contagion. A similar finding was also made by Conlon and McGee (2020). Al Mamun et al. (2020) researched how geopolitical risk and the influence of uncertainty in global economic policy impact both the volatility and the risk premiums of Bitcoin. The findings indicated that these factors significantly affect Bitcoin volatility and risk premiums, particularly during tense market situations.

Mavani (2020)'s research findings indicate that while VaR models are adept at capturing the market risk of traditional assets under normal market conditions, they are inadequate to address market risk during financial crises. However, Likitratchoaroen et al. (2021) further examined VaR models during the COVID-19 pandemic on the Bitcoin market. The findings revealed that these models accurately captured the potential adverse losses of BTC, especially at a 99% VaR confidence level, even in the face of the influence of the pandemic on both the stock markets and the cryptocurrency market. Trucíos and Taylor (2023) studied the precision of the VaR and ES models during turbulent market conditions by implementing different volatility modeling techniques. They found that surprisingly combined models like CaViaR did not perform as well as individual models, either globally or over time. Likitratcharoen et al. (2023) examined the efficiency of VaR models for cryptocurrency markets during uncertainty in monetary policies and the Russia-Ukraine war and found that the non-parametric VaR model is the most suitable for predicting extreme losses in cryptocurrency markets under stress. Bruzgė et al. (2023) conducted an extensive study of the fluctuations and risk assessments associated with the volatility of cryptocurrency. The findings indicate a pattern of clustering of volatility observed in Bitcoin, Ethereum, and Ripple. Furthermore, the VaR and ES estimates revealed that cryptocurrencies pose a higher risk as investments compared to technology stocks during stressed market conditions.

According to the BCBS (2019b), the construction of a risk model like VaR or ES requires a minimum of 250 days of historical data to build a market risk model for traditional assets. However, the current regulatory framework for cryptocurrencies is not yet in place, which is why no requirements for estimation windows have been set. Righi and Ceretta (2015) conducted research on determining an ideal estimation period for the VaR and ES risk models in the traditional stock market. They used Monte Carlo simulation as the method of estimation, examining various window sizes ranging from 250 to 2000 days. Their investigation revealed that models using estimation windows of 250 or 500 days exhibited the best performance. This outcome contradicted the conventional belief that longer time windows would produce more accurate risk estimates compared to shorter ones.

Berens et al. (2018) investigated the impact of varying the sizes of the estimation window on industry-standard market risk models. They tested various estimation window tactics across a range of basic parametric, semi-parametric, and non-parametric VaR and ES models. Using DAX daily return data for risk estimation, they illustrated substantial performance disparities arising from the choice of the estimation window strategy. Their findings suggest that forecast combinations emerge as the preferred strategy for estimation windows. Buczyński and Chlebus (2019) examined various VaR methodologies in different sample sizes, employing VaR models with window durations ranging from 50 to 2000, to assess whether there exists a threshold beyond which increasing the sample size does not significantly improve quality. The analysis incorporated parametric, non-parametric, and semi-parametric estimation techniques. Their suggestion put forth that a minimum of 900 or 1000 observations should be employed in conjunction with the aforementioned estimation methodologies. Considering that previous studies have primarily focused on conventional assets and their outcomes have been oriented towards longer time frames, this research seeks to broaden the examination of estimation windows to the context of cryptocurrencies. The goal is to provide additional information on the applicability of using shorter estimation windows in this domain.

1.2 Research objectives and questions

The cryptocurrency market has undergone profound transformations in recent years, marked by the emergence of numerous new cryptocurrencies, the use of diverse technologies, and multiple crises that have exerted immense pressure on the market (Likitratcharoen et al., 2023). The exponential growth of cryptocurrencies and the rapidly changing market landscape have triggered discussions among institutions about the necessity for various forms of legislation (see, e.g., BCBS 2019a; EBA 2019; ESMA 2019; BaFin 2023).

VaR and ES have undergone a detailed examination by academics and regulators in the context of traditional financial instruments (see, e.g., Dowd and Blake 2006; Jorion 2007; Yamai, Yoshiba, et al. 2002; Acerbi and Szekely 2014). On the contrary, the existing literature that examines the applicability of VaR and ES in the cryptocurrency market, especially during periods of increased volatility, is notably limited. This scarcity underscores the rationale for the first research question.

I. *"Can Value-at-Risk and Expected Shortfall accurately capture the potential adverse losses of prominent cryptocurrencies during increased market stress?"*

In this context, the accuracy of a risk measure is related to its ability to estimate the risk in close proximity to the realized value. Simply put, the risk estimate should not be an overestimation or an underestimation of the actual risk level. This study evaluates the accuracy of the risk measure using various methods, the most straightforward being Violation Rate (VR) and Normalized Shortfall (NS), which, according to Daníelsson (2011), are directly analogous. These backtest methods are assigned a value of 1 when the risk estimate precisely captures the realized risk. A value below 1 indicates an overestimation of risk, while a higher value signifies an underestimation of risk. Additionally, the study uses statistical backtest methods to assess the precision of the measures, and detailed methodologies and hypotheses for these measures are expounded in the methodology section.

The different sizes of the estimation windows affect the quality of the VaR and ES models (Buczyński & Chlebus, 2020). However, only a few studies investigate the influence of the duration of the estimation window on the VaR and ES risk estimates, and the few existing studies focus predominantly on traditional financial assets. Furthermore, existing research has mainly focused on the 250-day estimation windows, which yield somewhat conflicting results. Given that shorter estimation windows facilitate quicker adaptation to the swift changes in volatility witnessed in the cryptocurrency market, this study aims to specifically investigate the impacts of shorter estimation windows on the accuracy of risk estimates.

II. *"What is the impact of the size of the estimation window on the applicability of VaR and ES risk estimates in times of heightened market stress?"*

1.3 Contribution to existing literature

According to the best knowledge of the author, this study is one of the first to examine the accuracy of VaR and ES risk measures in measuring the market risk of cryptocurrencies under increased market stress, and at least the very first study that also examines the impact of estimation windows on these previously mentioned risk measures. Likitratcharoen et al. (2023) studied VaR as a market risk measure for cryptocurrencies during periods of increased market risk, but the study did not use the ES measure. On the other hand, Trucíos and Taylor (2023) extensively studied the market risk of cryptocurrencies, employing various estimation methods and both risk measures. However, the time frame used in their study did not specifically focus on the period of increased volatility.

The second contribution of this research is to assess how the size of the estimation window affects the calculated risk estimates between different estimation methods. As mentioned earlier, there has not been much research on this issue, which makes the choice of estimation window size ambiguous. Since shorter estimation windows have not been studied and those can adapt more quickly to large changes in volatility, this study aims to investigate how the estimation window for traditional assets, set by BCBS, performs against shorter estimation windows.

1.4 Structure of the thesis

The remainder of this thesis is organized as follows. Section 2 provides a concise overview of the historical background, classification, and current market landscape of cryptocurrencies. Furthermore, the chapter explores the regulatory framework that governs market risk measures, namely the VaR and ES models, delves into the current state of the regulatory landscape of the cryptocurrency market, outlines the risk measures applied in this thesis, their specific characteristics, and offers a summary of the process involved in the backtesting of these risk measures. In Section 3 the utilized data and its characteristics are introduced. Additionally, the methodologies used to measure market risk and the methodologies used to backtest risk measures are presented. In Section 4, the results are reported. Section 5 provides a summary, conclusions, limitations, and future directions of the study.

2 Theoretical framework

This chapter briefly outlines the historical background of cryptocurrencies and examines their characteristics compared to traditional financial assets. This is followed by the definition of risk used in this thesis and the rationale behind the need to quantify market risk, taking into account the attributes of different asset classes and relevant regulations. Subsequently, it introduces the risk measures used in the thesis, elucidating their characteristics, limitations, and the techniques used for backtesting.

2.1 Cryptocurrencies

On 31 October 2008, just under two months after the Lehman Brothers bankruptcy filing, Satoshi Nakamoto (2009) introduced an entirely new payment system, based on digital signatures and a peer-to-peer network, which allows online payments to be sent from one party to another without the involvement of financial institutions or any other third party. The rationale behind the proposal was to achieve cost efficiency and expedite transactions, which was made feasible by relying on cryptographic proof rather than placing trust in financial institutions. This payment system is now widely recognized as the world's first cryptocurrency, Bitcoin.

Cryptocurrency is a type of digital currency that operates independently without the supervision of any third party or central bank. Typically, cryptocurrencies use encryption techniques for security and are decentralized, relying on blockchain technology, a distributed ledger system, to record transactions and verify the integrity of the network. Instead of relying on a single company, cryptocurrency transactions are confirmed and performed by a community of participants, providing increased security and eliminating the need for intermediaries. (Balaji et al., 2023)

The emergence and popularity of Bitcoin has also spurred the growth of alternative cryptocurrencies, often referred to as "Altcoins", which aim to ameliorate the inherent constraints of Bitcoin, encompassing issues related to transaction speed, practicality, and utility. Nevertheless, it is noteworthy that Bitcoin retains its preeminent status as the world's largest cryptocurrency by market capitalization, with Ethereum (ETH), Tether (USDT), Binance (BNB), Ripple (XRP), and numerous others following suit in the rankings. These digital currencies are widely considered potential contenders for the role of future currencies. (Berentsen & Schär, 2018).

However, some studies, for example, Bianchi (2020) and Doumenis et al. (2021), suggest that cryptocurrencies are considered more appropriately as speculative investments or assets for hedging, rather than as a medium of exchange. Although there is significant potential for cryptocurrency to integrate more seamlessly into global financial and payment networks, its market remains highly volatile, with numerous cryptocurrency transactions being regarded as speculative ventures. Daily price fluctuations of cryptocurrencies can be up to ten times greater than those of traditional money markets ("Volatility in the Cryptocurrency Market", 2019).

In addition to their notable volatility, cryptocurrencies face various additional obstacles that impede their incorporation into the financial ecosystem. These include cybersecurity risks, vulnerabilities to possible hacking, and susceptibility to market manipulation. Due to these challenges related to cryptocurrencies, various legislators have refrained from acknowledging them as financial assets (IMF 2019; IFRS 2019b), and in some cases, countries such as China have even prohibited their use (Griffith & Clancey-Shang, 2023).

2.1.1 A simplified classification of cryptocurrencies

The growing pace of digital assets and the extensive debate on classification methods in the crypto community, where even current classifications are constantly changing, make it difficult to identify different concepts (Glas, 2022). This is why a simple classification of the different digital assets is provided, to make the thesis easier to follow and understand. Loosely defined, a digital asset is an intangible asset generated, exchanged, and held in digital form. In the context of blockchain technology, digital assets include cryptocurrencies and cryptographic tokens. For a more comprehensive discussion of the categorization of digital assets (see, e.g., Nakavachara et al. 2019; Pele et al. 2020; Castrén et al. 2022).

According to IMF (2019, p. 4), cryptocurrencies can be described as "digital representations of value, made feasible through advances in cryptography and Distributed Ledger Technology (DLT)". IMF (2019) classifies cryptocurrencies into two main categories: Bitcoin-like crypto assets (BLCAs) and digital tokens. BLCAs represent digital assets built on distributed ledger technology, with their main purpose being to function as a medium of exchange (e.g., Bitcoin and Ethereum, as used in this thesis). On the other hand, tokens exist within a platform constructed on an established blockchain, as exemplified by the numerous ERC-20 tokens within the Ethereum ecosystem.[1]

Bitcoin-like crypto assets (BLCAs)

Bitcoin (BTC) constitutes a decentralized digital currency originally expounded in a 2008 white paper attributed to an individual or collective entity operating under the pseudonym Satoshi Nakamoto. Bitcoin operates as a peer-to-peer exchange medium, where transactions occur directly among autonomous participants within its network, obviating the need for any intermediaries to authorize or facilitate these exchanges. Although this decentralization confers immunity to governmental manipulation or intervention, it concomitantly foregoes a central governing body to ensure the seamless functioning of the Bitcoin ecosystem or to underwrite the value of a Bitcoin unit. Bitcoins are generated digitally through a computational process known as "mining". Mining is a process that involves validating transactions on the Bitcoin network and adding them to the public ledger known as the blockchain. To be eligible to add a piece of information (block) to the blockchain, a miner must solve complex mathematical problems. These distinctive attributes distinguish Bitcoin from fiat currencies, the latter being underpinned by the complete faith and credit of their respective governments. The fundamental concept of Bitcoin revolves around the deployment of cryptographic techniques for the governance of money creation and transfer, eschewing the reliance on central authorities. (CoinMarketCap, 2023b)

Ethereum (ETH) constitutes a decentralized open-source blockchain framework characterized by its native cryptocurrency, Ether (ETH). The genesis of Ethereum can be traced back to 2013, when Buterin (2013) articulated its principles in a white paper. ETH serves as both a digital currency and a foundational infrastructure for a multitude of other digital assets, as well as the facilitation of decentralized smart contracts. Ethereum has played a pioneering role in the introduction of the concept of a blockchain-based smart contract platform, by which smart contracts are autonomous computer programs designed to automatically execute

[1]ERC-20 serves as the prescribed protocol for generating interchangeable tokens on the Ethereum blockchain.

the necessary actions to fulfill agreements between multiple online parties. [2] (CoinMarketCap, 2023d).

Tether (USDT), which began its operations in 2014, is a blockchain-based platform specifically designed to allow the use of fiat currencies in the digital form. The Tether is poised to challenge the traditional financial system by offering a contemporary paradigm for the handling of monetary transactions. A key innovation brought forth by Tether is its provision of a means for customers to engage in transactions involving conventional currencies on the blockchain, thus circumventing the typical challenges of volatility and complexity that accompany digital currencies. USDT distinguishes itself by being guaranteed by Tether to maintain a fixed value in alignment with the US dollar. According to Tether's assertion, when creating new USDT tokens, they allocate an equivalent sum of USD to their reserves, thereby substantiating the complete backing of USDT with cash and cash equivalents. (CoinMarketCap, 2023g)

Binance (BNB) was established in July 2017 and is widely recognized as the world's leading cryptocurrency exchange, distinguished by its daily trading volume. Binance is resolutely committed to elevating cryptocurrency exchanges to a prominent position within the realm of international financial activities. The nomenclature "Binance" is emblematic of the platform's desire to symbolize a new paradigm in the global financial landscape, denoting "Binary Finance" or simply "Binance". (CoinMarketCap, 2023a)

Ripple (XRP) was introduced in 2021 by a company called OpenCoin, led by technology entrepreneur Chris Larsen. Much like Bitcoin, Ripple operates as a hybrid, serving as both a digital currency and a payment system. Ripple constitutes the underlying currency unit, having a mathematical foundation similar to that of Bitcoin. Ripple stands as a prominent contender for the role of the successor to Bitcoin and employs a consensus algorithm that relies on the use of subnets within the broader network, which are collectively trusted for its operation. (CoinMarketCap, 2023f)

Digital tokens

A token is a digital entity that grants access to and enables participation in the broader cryptocurrency ecosystem. Tokens belong to the category of cryptocurrencies as a subset (Coin-

[2] In many contexts, Ethereum can also be classified as a utility token, which again shows the complexity of classification methods.

MarketCap, 2023h). According to IMF (2019), digital tokens can be categorized into four distinct types according to their fundamental economic functions, which are the following.

a) Payment tokens: those designed to function as universal BLCAs and serve as units of account, stores of value, and mediums of exchange across platforms, without being limited to a particular platform (e.g., Litecoin).

b) Utility tokens: those created with the aim of granting holders future access to services through a DLT-based application. Instances of such applications include those for file storage, social messaging, and trading (e.g., ERC-20 and Filecoin).

c) Asset tokens: those that indicate obligations or ownership stakes in the issuing entity. They produce interest for the holder or pledge a portion of the company's future earnings, respectively.

d) Hybrid tokens: those that are part utility and part asset or payment token.

2.1.2 Cryptocurrencies as an asset class

Cryptocurrencies or crypto assets are novel and innovative assets that can be loosely described as highly volatile and uncorrelated with traditional financial assets such as stocks, bonds and currencies (Baur & Dimpfl, 2018). Due to their unconventional nature in financial markets, they are usually not classified as financial assets by academics and investors. However, there has been an ongoing discussion among lawmakers and economists regarding the classification of cryptocurrencies, with some debating whether they should be considered currencies or speculative investment instruments (see, e.g., Glaser et al. 2014; Dyhrberg 2016). One point of view argues that cryptocurrencies function as a medium of exchange within a decentralized network, while another point of view points out that their notably volatile returns seem to align with the functions of conventional assets ("Volatility in the Cryptocurrency Market", 2019). Classifying cryptocurrencies poses a challenge due to the continuous emergence of new ones and evolving technologies, potentially impacting existing categorizations. However, delving into this subject remains pertinent, as defining cryptocurrencies as financial assets could likely lead to varying regulatory frameworks, such as mandatory market risk management, making the discussion on this matter relevant.

Before we can discuss whether cryptocurrencies should be considered as currencies, financial instruments, or financial assets, we must first define these concepts. For a currency to be recognized as such, according to how Central Banks treat traditional currency, it should ideally serve three fundamental roles: acting as a unit of account, a store of value, and a medium of exchange. Generally, cryptocurrencies with significant market capitalization have the potential to fulfill all these criteria, while most other cryptocurrencies find it challenging to meet even one of them.

Unit of account the primary purpose of the currency is to serve as a unit of account, enabling the measurement of value in various units and simplifying comparisons. Digital currencies are composed of consistent, distinguishable, and measurable units of account. As long as these units can be easily traded, this function is fulfilled, as their value can be determined and compared. Consequently, high-cap cryptocurrencies effectively operate as efficient units of account. (Kim et al., 2018)

Store of value involves maintaining the capacity to preserve purchasing power in the future, making it more, less or equally valuable for subsequent transactions. This requires a certain degree of predictability in the future value of the asset, a challenge faced by cryptocurrencies due to their high volatility. Both precious metals like gold and digital coins have the ability to store value, operate independently of fiat currencies, and serve as safe havens during economic crises. However, it is important to note that only gold consistently retains these attributes over the long term. Daily trading fluctuations of specific digital assets, including Bitcoin (BTC), Ethereum (ETH), and Litecoin (LTC), have frequently exceeded the annual inflation rates of countries experiencing economic downturns, such as Mexico and South Africa. This implies that holding the Mexican peso might be less risky than holding top cryptocurrencies due to their considerable volatility and the potential for security breaches. Therefore, the certainty that cryptocurrencies serve as a secure store of value remains uncertain until the market stabilizes. (Kim et al., 2018)

Medium of exchange demands that a currency be widely recognized and easily exchangeable for all available goods and services, serving as an intermediary to overcome the limitations of barter transactions. Currently, most cryptocurrencies do not fulfill this requirement, as they are not easily accessible for everyday transactions. Nevertheless, BTC, LTC, ETH, and the United States Dollar Tether (USDT) facilitate access to other crypto assets and act as intermediaries between fiat money and cryptocurrencies. Generally, cryptocurrencies can be recognized as a medium of exchange for crypto assets, but this role is still evolving and is predominantly observed among the most prominent crypto-based coins, rather than being prevalent throughout the entire class. (Kim et al., 2018)

According to IFRS (2007) IAS 32 ”Financial Instruments: Presentation” a financial asset is one that is:

· Cash

· A financial instrument representing ownership in another entity.

· A contractual entitlement to obtain cash or an alternative financial asset from another entity

· A contractual entitlement to swap financial assets or financial liabilities with another entity based on specific conditions

· A specific agreement that is either certain or potentially settled using the company's own shares

Despite the fact that we have previously stated that some cryptocurrencies could be considered a medium of exchange, IFRS (2019a) highlighted that the ownership of cryptocurrencies is not classified as a financial asset. This determination arises from the fact that cryptocurrency does not represent cash or an equity instrument in another entity. IMF (2019) revisited the definition of cash outlined in IAS 32. They noted that this definition suggests that cash is anticipated to serve as a medium of exchange (i.e., used for transactions involving goods or services) and as the monetary unit for pricing goods or services to the extent that it forms the fundamental basis for measuring and recognizing all transactions in financial statements. On the basis of this interpretation, the Committee reached the conclusion that cryptocurrencies do not qualify as cash.

There is also an ongoing debate among academics on how cryptocurrencies should be classified, and many academics also have differing opinions. However, it is widely recognized that cryptocurrencies show a distinct behavior compared to traditional asset classes, including currencies (Baur and Dimpfl 2018; Auwera et al. 2020). Cryptocurrencies, in fact, exhibit not only greater volatility and risk compared to traditional asset classes, but also demonstrate more pronounced extreme outcomes. The question of whether cryptocurrencies can be categorized as financial assets continues to be a topic of ongoing debate.

According to Glaser et al. (2014), most cryptocurrency users perceive these digital assets as speculative assets rather than as a medium of exchange. Investing in cryptocurrencies is not considered safe; instead, it constitutes a bet on the underlying blockchain technology or project. In some respects, this aligns cryptocurrency investments with investing in high-tech companies (Bianchi, 2020). Furthermore, the authors, for example, Brown (2019) and Haykir and Yagli (2022) assert that cryptocurrencies are clearly a speculative bubble; the predominant perspective in contemporary research leans toward the notion that cryptocurrencies are progressively developing into a distinct and unique asset category.

For example, Dyhrberg (2016) studied the similarities between Bitcoin, gold and the US dollar. Instead of categorizing Bitcoin as a currency or an asset class, it was categorized as a conceptually placed hybrid between a mere medium of exchange and a mere store of value. Gold and cryptocurrencies share common characteristics, as they have limited supply, exhibit significant price fluctuations, and serve as alternative investments for those who have doubts about fiat currencies and monetary policies (Bianchi, 2020). Glas (2022) conducted an extensive study comparing traditional assets with cryptocurrencies. He highlighted that while significant digital assets such as Bitcoin may be viewed as a legitimate asset category, they should be regarded more as an independent asset class. This distinction arises from the absence of specific regulations that would typically characterize a "serious" asset class, which crypto assets currently lack. Furthermore, persistent concerns about hacking, fraud, and illicit activities remain significant challenges within this asset category, which require effective resolution before it earns the designation of a "serious" asset class.

Thus, cryptocurrencies can be classified in nearly as many manners as there are classifiers and classification systems. However, this does not diminish the importance of assessing the current situation. Establishing a standardized classification approach for cryptocurrencies would not only delineate their categorization but would also determine methods for managing and quantifying the market risk associated with them. However, recent regulatory changes such as MiCA hint at the potential of cryptocurrencies to be recognized and treated as some form of asset. With the introduction of the BCBS (2022) framework, expected to be enforced in 2025 for cryptocurrencies, an adapted version of ES, subject to stress testing, is expected to be employed.

2.1.3 Market condition of cryptocurrencies

The market value of cryptocurrencies and digital assets has increased exponentially in recent years. In April 2023, the total market capitalization of cryptocurrencies reached 1.230 trillion dollars, while in April 2014, the total market value of virtual currencies stood at 29 billion dollars (CoinMarketCap, 2023e). The increase in the market value of cryptocurrencies has also led to a tremendous surge in the number of new currencies. In January 2017, there were 636 "alive" cryptocurrencies, while in January 2023 there were more than 20,000 different cryptocurrencies, of which 8,856 were considered "alive" (Statista, 2023).

In particular, the continued interest of investors in finding new alternative investment avenues beyond traditional options can be seen as a significant factor in this growth. Furthermore, the rise of technological and financial innovations has led to the emergence of new investment opportunities in addition to cryptocurrencies within the crypto sphere, including stablecoins such as Tether (USDT), utility tokens such as Filecoin, and Decentralized Finance (DeFi) tokens such as Terra (LUNA) (Scharfman, 2022).

In 2017, the cryptocurrency market faced a significant increase, marking a notable departure from the previous four years. The total market capitalization of all cryptocurrencies reached approximately $600 billion in 2017, subsequently plummeting to $130.66 billion in 2018, and then experiencing a modest upturn to $192.49 billion at the end of 2019. The years 2020, 2021 and 2022 were characterized by a period of extreme volatility. Despite the global economic devastation caused by the COVID-19 pandemic, 2020 marked the beginning of a substantial bull run in the cryptocurrency market. In May 2021, the market capitalization of cryptocurrencies reached its zenith at approximately $2.4 trillion, only to decline to $1.26 trillion in mid-July. However, the market capitalization of cryptocurrencies resuspended, culminating in an all-time high of $2.8 trillion at the end of 2021. At the beginning of 2022, the crypto market entered yet another bear run, and by June 2022 the total market capitalization of cryptocurrencies fell below $900 billion. (CoinMarketCap, 2023e)

Cryptocurrency price changes in the past few years have been affected by similar factors that impact stock market prices. However, the crypto market has faced unique events that have intensified its volatility and price changes. For instance, China's ban of crypto-related activities in 2021 resulted in a sharp drop in cryptocurrency market capitalization, lingering at their lowest point post the ban announcement. This restriction significantly increased the momentary volatility of cryptocurrencies, as shown in the findings of Griffith and Clancey-Shang (2023).

The precipitous decrease observed in 2022 can be attributed to the demise of Terra and its ripple effect on the cryptocurrency market. Terra had held the rank of the third most substantial cryptocurrency ecosystem in 2022, trailing only Bitcoin and Ethereum, before it experienced a dramatic downfall within a mere three days in May 2022, Terra suffered a substantial downfall, completely erasing its market capitalization due to a 50 billion devaluation. The focal point of this decline revolved around the operational run of Anchor, a blockchain-based lending and borrowing protocol that had extended attractive promises of elevated returns to its stablecoin investors who use TerraUSD (UST). During the past year, the combined market size of cryptocurrencies has predominantly hovered around the trillion dollar threshold. (Liu et al., 2023)

At the time of writing, in October 2023, Bitcoin and Ethereum collectively control a substantial majority of the overall cryptocurrency market, accounting for 68.7% of the total market share. In particular, Bitcoin continues to maintain its leading position as the largest cryptocurrency, with a market value representing 51.42% of the total. Following Bitcoin and Ethereum, the next three largest cryptocurrencies are USDT (7.41%), BNB (2.86%) and XRP (2.45%), while the remaining cryptocurrencies collectively constitute the remaining share, which approximately amounts to 18.6% (CoinMarketCap, 2023e). In light of the prevailing market dynamics and to maintain the focus of this study, our analysis is exclusively focused on Bitcoin and Ethereum.

2.2 Risk and risk management

In order to examine risk management and risk measure, it is necessary first to define what risk is and what it encompasses and excludes. The definition of risk is not trivial. If we define risk as the potential occurrence of anticipated or unforeseen events, we must consider the nature of these events. Even if these events could be identified in such a way that the information obtained from them could be used, it should also be possible to quantify their probability. An individual who chooses to invest in cryptocurrency could face substantial economic risk without the accompanying physical hazards. On the contrary, someone engaged in bull-running incurs physical risk, but without the prospect of financial gain. This thesis explores the economic risks that cryptocurrency investors may encounter and aims to clarify the concept of financial risk in this context.

2.2.1 Defining risk

The financial literature often refers to risk but lacks a widely accepted concrete definition for it. This absence is not coincidental, as risk is a concept that is difficult to precisely define due to its unintuitive nature. It relies on the ideas of exposure and uncertainty, both of which cannot be precisely operationalized. As a result, creating an operational definition of risk is impossible; at most, we can attempt to operationally define our perception of risk. However, defining perceived risk in operational terms is intricate because it encompasses various aspects. To simplify this effort, some elements of perceived risk can be operationally defined. (Holton, 2004) Risk measures such as VaR and ES serve this purpose in practice.

As discussed previously, defining and distinguishing risk in precise terms can be complex. However, a fundamental concept in finance is that seeking rewards often involves accepting risks. It is crucial to acknowledge that not all risks yield the same potential rewards. Some risks are chosen because their benefits outweigh the costs involved. Since both risk and reward pertain to future outcomes, the goal is to take profitable risks where the expectation of gain offsets the anticipation of loss (Engle, 2004). This thesis concentrates on examining the market risk disparity between cryptocurrencies and conventional investments. Therefore, our attention is directed towards events such as market price fluctuations that can cause financial losses. For this purpose, within this study, we utilize asymmetric risk measures, VaR and ES, which specifically target potential losses within the distribution of returns.

Financial risk can be divided into four main types: market risk, credit risk, liquidity risk, and operational risk. These risks are not unambiguous due to their diversity; for example, hedge funds can assess two distinct categories of market risk: the initial category refers to the risk inherent within the underlying positions that comprise the hedge fund portfolio, while the second category encompasses the risk attributable both to the investment strategy used and the risk mitigation measures employed by the hedge fund manager (Duc & Schorderet, 2008, p. 99).

Furthermore, it can be difficult to differentiate between risks because they can arise from identical factors. For example, both market risk and credit risk are influenced by the same economic variables, and the distinction between these two types of risk becomes even less clear when dealing with cryptocurrencies compared to traditional financial instruments. This complexity arises from the fact that cryptocurrencies are affected not only by conventional factors, but also by elements that do not necessarily impact traditional assets. What differentiates cryptocurrencies in terms of market risk is the absence of financial oversight, making coin prices susceptible to manipulations, pump-and-dump schemes, and other fraudulent activities. However, credit risk, defined as the situation in which profits and losses on the value of an abandoned and considered "dead" cryptocurrency position, differs from traditional credit risk because these seemingly defunct coins can undergo multiple revivals. (Fantazzini & Zimin, 2020)

2.2.2 Market risk management

In traditional finance, market risk refers to financial losses incurred due to sudden market movements or volatility in market prices. In a cryptocurrency context, market risk refers to the profits and losses associated with the value of the position or portfolio of 'alive' cryptocurrency, resulting from fluctuations in market prices on central and decentralized exchanges. Fundamentally, the distinctions between credit risk and market risk in the context of cryptocurrencies are primarily related to quantity and timing, rather than being qualitative in nature. In essence, if financial losses and technical issues can be addressed using existing financial and technical capabilities, it constitutes a market event. However, if financial losses become excessive and technical challenges prove insurmountable, it leads to a credit event, ultimately resulting in the demise of the cryptocurrency. Market risk emerges as a result of changes in market factors, which can include asset prices, interest rates, foreign exchange rates, or, in our context, the prices of cryptocurrencies. (Fantazzini & Zimin, 2020)

Commonly, market risk can be categorized into two distinct types: idiosyncratic or non-systematic risk, which refers to risks that can be spread out through diversification, and systematic risk, which pertains to risks that impact the entire economy and cannot be mitigated through diversification. Market risk is inherent in all securities within a particular category. For example, all stocks are consistently exposed to the same market risk. This risk is not reducible through diversification. On the contrary, the risk of the bond market arises from interest rate fluctuations, while stock prices are influenced by a wide range of factors that include company performance, economic conditions, and significant political events of national importance (Szylar, 2013).

The significance of market risk is increasing within the realm of banking management. As indicated by Penza and Bansal (2001), this increase can be attributed to four key factors:

1. The process of securitization, which has led to the replacement of conventional asset forms with financial assets traded on a secondary market, which consequently have an ascertainable market price. Securitization has facilitated the adoption of marking-to-market methods, enabling the prompt monitoring of profits and losses associated with a collection of traded instruments.

2. The increasing complexity of financial instruments, particularly derivatives, frequently transacted by banks, has emphasized the need to replace conventional risk assessment methods with a comprehensive risk measure. This measure is envisioned as the stronghold of a unified and consistent framework for financial risk management.

3. The increased volatility observed in interest rates, foreign exchange rates, and stock prices over the past two to two and a half decades, particularly following the dismantling of the gold exchange standard in 1973, has been amplified by the globalization of financial markets. This increased volatility is, to some extent, attributable to the financial crises that plagued certain financial institutions in the late 1980s.

4. The increased trading activity of the banks and the resulting volatility of income.

Traditional risk management approaches encompass at least two perspectives: one from within the organization (internal viewpoint) and the other dictated by regulatory requirements (regulatory viewpoint). The internal perspective encompasses the view of risk management as a fundamental aspect of the overall business operations of banking and financial institutions. In this perspective, risk management is considered crucial, regardless of regulatory requirements, similar to how manufacturing companies supervise their production and distribution processes. On the other hand, regulations have a specific objective; preventing the failure of one bank from triggering a widespread banking crisis. When an individual bank encounters difficulties, it affects its depositors; however, if multiple banks face challenges simultaneously, it poses a significant threat to the entire economy. (Penza & Bansal, 2001)

Managing risk is a formidable task due to the continuous operation of financial markets and the growing emphasis on intricate financial instruments, making it challenging to evaluate their risks. While it is possible to pinpoint portfolio exposures with a degree of certainty, the actual losses that may result from these exposures remain uncertain. The information used to gauge potential losses relies on historical prices and rates, rather than those anticipated in the future. Mitigating risk is made more complicated by the opportunity cost in financial markets between risk and expected return, since generally higher returns require greater risk taking. An essential concern within contemporary risk management revolves around the measurement of risk. For example, regulatory authorities assess the risk exposure of financial institutions to determine the requisite capital reserves that must be maintained as a buffer against unforeseen financial setbacks. Similarly, exchange-affiliated clearinghouses are tasked with establishing margin prerequisites for investors participating in trading activities on their respective platforms. (McNeil et al., 2005)

Traditional financial markets are subject to regulatory oversight and mandate the use of risk measures as stipulated by the Basel II and Basel III accords. On the contrary, the current regulatory environment for cryptocurrencies is ambiguous, but there is no formal obligation to adopt risk measures. However, the examination of risk measures within cryptocurrency markets is of significant relevance to a range of stakeholders, including investors, hedge funds, market makers, and traders. This relevance stems from their utility in optimizing order limits, formulating effective option pricing strategies, and designing robust trading systems. Furthermore, considering risk assessment methodologies for cryptocurrencies is of paramount importance in the design of prospective regulatory frameworks (Trucíos & Taylor, 2023).

2.2.3 History of VaR and ES

The concept of volatility was found around 1860 and has been used since in various industries. Volatility is still the main measure of financial risk (Daníelsson, 2011). Many well-known academics such as Markowitz (1952) and Tobin (1958) also linked risk to volatility in the value of the portfolio. Before VaR was established, volatility was only a risk measure (Daníelsson, 2011).

As the transition from the 1970s to the 1980s occurred, the volatility of financial markets increased. Companies were becoming more leveraged and there was a growing demand for new risk measures. The resources necessary to implement VaR were becoming accessible, but mainly viewed as a theoretical concept within portfolio theory. Companies were looking for a means to evaluate market risk in various asset types, although they had not fully grasped how VaR could address this requirement. In the United States, regulators were laying the groundwork to facilitate its adoption. (Holton, 2002)

In the late 1980s, JP Morgan developed a comprehensive VaR, system, also known as Risk-Metrics, for their entire firm (Guldimann, 1995). This system considered several hundred risk factors and relied on a covariance matrix that was updated quarterly based on historical data. On a daily basis, different trading units would communicate their position changes in relation to each risk factor by email. These changes were then aggregated to represent the overall value of the portfolio as a linear combination of these risk factors. Using this information, they calculated the standard deviation of the portfolio value. They used various VaR models, including a one-day 95% VaR in USD, which was calculated under the assumption that the portfolio value followed a normal distribution. (Guldimann, 2000)

By 1993, many financial institutions were using their unique VaR methodologies to assess market risk, distribute capital, or maintain close eye on market risk thresholds. These methodologies came in various forms, but were most frequently influenced by the framework of Markowitz (1952). Over time, the value of these internal VaR measures gained recognition and BCBS officially approved their adoption by banks for regulatory capital calculations.

At the turn of the 2000s Artzner et al. (1999) introduced ideal properties of risk measure and criticized VaR about now having these properties. In the same context, he proposed an alternative risk measure for VaR which is commonly known as ES. However, after this proposal, several years passed before the widespread adoption of ES. This delay was due to the ongoing debate surrounding the mathematical properties of the measure, suggesting that it could not be backtested. A significant example of this is the publication of Gneiting (2011) that highlights the shortcomings of ES in terms of elicitability.

However, in 2012, BCBS published a discussion paper that outlines the plans for ES to replace VaR because, due to its mathematical properties, it would be more capable of measuring risk, especially during periods of high market stress (BCBS, 2012). Due to the ongoing discussion on ES, BCBS chose to delay its implementation, preserving VaR as the predominant risk measure (BCBS, 2013). In 2014, (Acerbi & Szekely, 2014) proved in their publication that ES can be subjected to backtesting, despite the lack of elicitability. Since then, both risk measures have become central to risk management, to the extent that regulatory bodies such as BCBS have mandated their application in the financial sector (BCBS 1996; BCBS; 2019b). The regulatory perspective of these risk measures is discussed in more detail in the following sections.

2.3 Regulation of market risk management

In essence, each market participant has the power to determine how to allocate its assets, and, as such, each player is responsible for its own actions. Therefore, the question is whether banking regulation is necessary at all. However, banking regulation is widely seen as necessary, even indispensable, for the stability of financial markets. The purpose of regulation is to ensure that banks are sufficiently solvent in times of crisis to avoid a large-scale financial crisis. For the ordinary commercial bank, the main regulatory issues are mainly externalities and deposit insurance. This section outlines the laws governing the VaR and ES risk measures used to assess market risk, along with an examination of the existing regulatory framework on cryptocurrencies.

2.3.1 Basel I

In 1988, the G10, which consists of the ten most industrialized nations, entered into an agreement aimed at regulating banks. Today, numerous additional countries have become signatories to this agreement. Although member nations have the authority to enact stricter regulations for their banks, they are required to adhere to the fundamental principles outlined in the agreement. This pact requires that banks in member countries maintain a minimum level of capital reserves to hedge a range of risks. The original agreement focused only on credit risk. The agreement required each bank to allocate a capital reserve (Cooke ratio) of 8% of the value of the securities that represent the credit risk in the bank's portfolio. The purpose of this ratio was to provide a financial safety net for the bank. The specific weight assigned to the various financial securities was rather arbitrary when the regulation was first applied. (BCBS, 1988)

The original Basel I Accord faced significant criticism for its omission of market risk and its overly cautious approach to credit risk, as it failed to account for the potential benefits of risk diversification and offset of positions through "netting", which involves matching the maturities of long and short positions. In 1995, the practice of netting for credit risk, including derivative-related positions, was authorized. In 1996, the first revision of Basel I was introduced to incorporate market risk and allow the use of internal market risk models. This framework is based on the proper allocation of positions between the trading book and the non-trading (or banking) book. Within this regulatory context, financial institutions were mandated to assess market risk, in addition to credit risk, using the VaR measure. (BCBS, 1996)

2.3.2 Basel II

A significant assessment of operational risk and credit risk occurred in 2004 with the implementation of Basel II measures, enforced in 2006 by BCBS. However, the 2007 financial crisis hampered its widespread implementation in many countries. During this period, the contentious nature of VaR and its lack of coherency fueled extensive discussions, leading to the suggestion of ES as an alternative risk measure to replace VaR. (BCBS, 2006). However, BCBS opted not to substitute VaR as a risk measure due to uncertainties surrounding the backtesting of the ES measure, due to its lack of elicitability. However, Basel II implemented more stringent criteria concerning VaR usage, evident in the mandate for banks to provide justifications for any omitted pricing elements in VaR risk estimations (BCBS, 2009).

Basel II is based on three mutually reinforcing pillars.

Pillar I includes minimum capital requirements for credit, market, and operational risk, which are based on financial models rather than accounting rules.

Pillar II comprises a comprehensive evaluation conducted by both the supervisor and the regulator to verify capital sufficiency, thus transferring accountability to the regulator, who is tasked with overseeing the bank's performance and promptly identifying any deficiencies. However, the primary focus is on the proactive measures taken by the supervisor. The second pillar also encompasses risks that are not addressed in the first pillar, including but not limited to interest rate risk.

Pillar III emphasizes on fostering market discipline among banks and strives to promote greater transparency in the information they provide. Pilar III also includes specific disclosure mandates.

In response to significant financial losses suffered by banks during the global financial crisis, the BCBS introduced the Basel 2.5 framework in 2009. 2009 The primary contributor to excessive leverage and risk-taking was identified as the trading book. Recognizing that existing regulations did not adequately address certain key risks, the Committee introduced an incremental capital risk charge to complement the VaR-based trading book framework. Additionally, stressed VaR calculations became mandatory, as losses in the trading books of numerous banks far exceeded the minimum capital requirements set by the previous regulation. The implementation of Basel 2.5 in banks coincided with regulators developing a more comprehensive strategy in response to the lessons learned from the 2007 financial crisis (BCBS, 2009).

2.3.3 Basel III

Basel III brings in additional capital demands aimed at safeguarding banks and enhancing their oversight of liquidity risks. This agreement requires banks to improve risk management, strengthening regulatory oversight. Risk managers within banks will need to maintain a greater degree of autonomy compared to CEOs. The agreement also emphasizes increased transparency and the reinforcement of long-term capital reserves. Meanwhile, speculations surfaced about VaR being substituted by a subadditive ES. However, this transition did not materialize when Basel III was released in 2010 (revised in 2011). (BCBS, 2011)

In 2016 (revised in 2019), the BCBS published a minimum capital requirement for market risk. The proposed changes to market risk requirements included several key adjustments. These changes involved transitioning from the use of VaR to ES as the primary risk measure, extending the liquidity horizon beyond the existing 10-day limit for illiquid trades, and implementing a revised standardized approach that would become mandatory, even for banks currently employing the internal model approach. The rationale for the use of the ES measure lies in its ability to account for tail risk prudently, which will help ensure capital adequacy under stressed market conditions. Since ES represents the expected tail loss beyond VaR, these modifications have the potential to increase market risk capital requirements, although regulators tend to reduce the confidence level from 99% to 97.5% for the internal model-based approach. However, the introduction of longer liquidity horizons, exceeding the 10-day threshold, is poised to definitively raise the market risk capital demands for portfolios containing illiquid assets. (BCBS, 2019b)

2.3.4 Current state of cryptocurrency regulation

Governments around the world strive to understand and regulate emerging currencies, but establishing effective regulations poses significant challenges. Cryptocurrencies operate across a global network and are mined intermittently to avoid governmental oversight, making it difficult for any nation to establish unilateral rules. Collaboration among multiple nations is essential to construct a robust legal framework that prevents arbitrage opportunities. Furthermore, unlike traditional currencies controlled by central banks or governments, cryptocurrencies are managed by decentralized networks, adding to the complexity of regulation. Adapting existing regulations to encompass cryptocurrencies presents difficulties due to their multifaceted nature. Cryptocurrencies exhibit characteristics that span three main categories: they can be viewed as securities, commodities, or currencies. This complexity complicates efforts to fit them neatly within pre-existing regulatory frameworks. (Auwera et al., 2020, pp. 44–46) This stands as the primary rationale behind our previous conversation on how cryptocurrencies are or should be considered and treated as assets (recall 2.1.2).

Developing a unilateral regulatory structure for cryptocurrencies presents challenges due to varying perspectives on legislation in different countries. For example, China enforces a complete prohibition on cryptocurrencies, banning all related activities, including mining, trading, and initial coin offerings (Griffith & Clancey-Shang, 2023). On the contrary, Germany adopts a different regulatory position, where the country's regulatory body, BaFin, categorizes Bitcoins as units of account according to German law, thus identifying them as financial instruments (Klöhn & Parhofer, 2018).

BCBS has expressed the view that the rapid development of cryptocurrency assets and associated products and services could introduce greater risks to banks, raising concerns about the stability of the financial system. Consequently, BCBS (2019a) released its initial discussion paper, seeking feedback from relevant stakeholders, in order to formulate a framework for the "prudent treatment of banks' exposure to crypto assets". Subsequently, two consultation papers were issued outlining preliminary and revised proposals on this matter.

The BCBS introduced a minimum capital requirement for banks' direct involvement with cryptoassets toward the close of 2022. Although the standard is not currently legally enforceable, it requires transposition into EU law by January 1, 2025. Additionally, the report touched on market risk management with respect to cryptocurrencies. When assessing market risk capital requirements for Group 1 cryptoassets using the internal model approach (IMA), banks will be required to calculate an overall stressed expected shortfall (SES) as a capital measure. (BCBS, 2022)

The MiCA establishes consistent market rules throughout the European Union for crypto assets. It addresses crypto assets that are not currently under the purview of existing financial services regulation. Important aspects concerning the issuance and trading of crypto assets, including asset reference tokens and e-money tokens, encompass transparency, disclosure, authorization, and transaction supervision. This new legal framework is designed to enhance the integrity of the market and financial stability. It achieves this by regulating public offerings of crypto assets and ensuring that consumers are better informed about the risks associated with these assets. The regulation includes a significant number of Level 2 and Level 3 measures that need to be developed within a 12- to 18-month timeframe, depending on the mandate, before the new rules come into effect. The date of implementation of these measures is contingent on approval by the European Commission, the European Parliament, and the Council of the European Union.

2.4 Risk measures

Risk measures play an essential role in risk management and are particularly important when it comes to managing potential losses in the financial or insurance sector. The risk measure captures the riskiness of an investment in a single figure that allows the risks of different investments to be compared, even if they depend on different market variables. Risk measures are used, for example, to estimate the amount of capital required to cover risks and to compare and limit the riskiness of different investment portfolios (Rockafellar & Uryasev, 2002).

The well-known mean-variance framework introduced by Markowitz (1952) integrates both the mean and the standard deviation (or variance) into a single value, effectively encompassing both the central tendency and the spread of the probability distribution. The mean-variance framework is considered a symmetric risk measure, which means that it does not distinguish between positive deviations that represent portfolio profits and negative deviations that indicate portfolio losses (Alexander et al., 2009). The contemporary financial landscape is considerably more complex than in the late 1950s. Over time, several major markets have undergone deregulation, leading to the emergence of new investment possibilities, especially in the growing economies of Asia and South America. Currently, investment banks manage large trading portfolios that include a multitude of positions in various products. The use of derivative instruments for both speculative and risk mitigation purposes has seen a substantial increase. All of these advances pose novel challenges for investors, traders, and especially risk managers (Hubbert, 2012).

Contemporary risk analysis frequently places emphasis on tail risk, operating under the assumption that extreme events occur infrequently. In this context, quantiles represent a valuable tool for assessing these infrequent events. This orientation has led to several asymmetric risk measures, with prominent examples such as VaR and ES that focus specifically on quantiles within the return distribution (Liu & Wang, 2021). This thesis focuses on asymmetric risk measures, specifically VaR and ES, which differentiate between profits and losses, accentuating the downside risk and the potential magnitude of losses.

2.4.1 Value-at-Risk

VaR constitutes a singular and succinct statistical measure used to assess potential losses within a portfolio. It is primarily designed to gauge the losses of standard market fluctuations. Instances where losses exceed the VaR threshold occur infrequently and are dependent on a predefined low probability (Linsmeier & Pearson, 2000).

According to Hull (2018a) VaR, is a function characterized by two key parameters, namely the holding period and the confidence level. The holding period denotes the time over which losses may occur, and thus the time risk is forecast over this period of interest. The confidence level denotes the probability that the losses exceed the chosen VaR level. In instances where n days serve as the designated time horizon and $(1-\alpha)$ represents the confidence level, VaR can be described as the magnitude of the loss corresponding to the α^{th} quantile within profit-loss distribution observed in the portfolio value during the period n days. Figure 1 illustrates VaR 95% ($\alpha = 0.05$) under the assumption that profits and losses are normally distributed.

Artzner et al. (1999) define VaR at the confidence level of $(1-\alpha)$ as follows.

$$VaR_{\alpha}(X) = \text{inf}\{x \mid P[X \leq x] > \alpha\}, \tag{2.1}$$

where X is a random variable that denotes asset outcome, inf $\{x \mid A\}$ is the lower limit of x given event A, and inf$\{x \mid P[X \leq x] > \alpha\}$, indicates the α^{th} quantile of profit-loss distribution (Artzner et al., 1999).

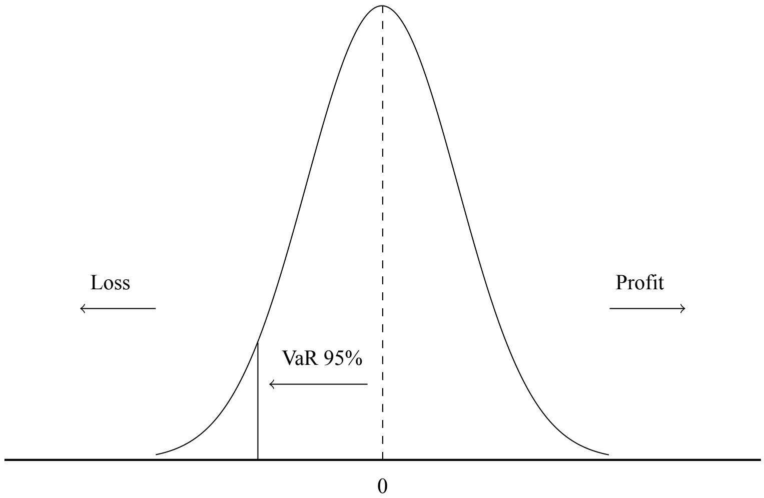

Figure 1: Profit-Loss Distribution and VaR

Although widely used, VaR as a risk measure faces several limitations that impact its accuracy. First, it assumes a constant portfolio composition within the specified time frame. This holds well for short periods but becomes less reliable over extended periods due to potential changes in portfolio positions, rendering the originally calculated VaR outdated (Röman, 2017). Second, VaR relies on the assumption of a normal distribution, which does not always hold in the financial and cryptocurrency markets. Anomalies such as thick tails, high peaks, and skewness, observed in various studies (e.g., Wong and Vlaar 2003; Koutmos 2018; Liu and Tsyvinski 2021), deviate from the theoretical normal distribution, commonly witnessed empirically (Wolke, 2017, p. 64). Third, VaR does not cover losses beyond a preset threshold. Lastly, Artzner et al. (1999) argue that VaR lacks coherence as a risk measure because it overlooks the advantages stemming from diversification. The concept of coherence and its constituents as a risk measure are extensively debated in the following section.

2.4.2 Coherent risk measure

According to Artzner et al. (1999), a risk measure must be coherent to be validly used to measure risk. A risk measure that adheres to the four axioms of translation invariance, subadditivity, positive homogeneity, and monotonicity is called a coherent risk measure. VaR as a risk measure is not considered a coherent measure because it does not meet the axiom of subadditivity (Roccioletti, 2016, p. 275). This shortcoming is considered significant, as cases where subadditivity is not taken into account properly can produce severe consequences from a risk management perspective. The lack of subadditivity refers to a situation where diversification of a portfolio does not reduce its riskiness. In the interim, it can lead to a situation where the financial institution makes suboptimal investment decisions.

We can illustrate a coherent risk measure in a following way. Suppose that X and Y denotes two different assets and ρ is a risk measure which maps observations of an asset, like X, onto a risk measurement. Further define constant κ. Then ρ is coherent risk measure if it satisfies the following four axioms (Steland, 2012):

AXIOM I. Monotonicity. A risk measure ρ is monotone when:

$$\rho(Y) \geq \rho(X) \quad \text{if} \quad X \leq Y. \tag{2.2}$$

The rationale behind this is that if a portfolio consistently underperforms compared to another portfolio, it should be considered riskier and require a greater amount of capital (Hull, 2018b, p. 275).

AXIOM II. Subadditivity. A risk measure ρ is subbadditive, if:

$$\rho(X+Y) \leq \rho(X) + \rho(Y). \tag{2.3}$$

If a bank is composed of two units, it should have to reserve less buffer capital compared to the sum of buffer capital required for each unit treated as individual entities, that is, diversification should be encouraged (Hult, 2012, p. 162).

AXIOM III. Positive homogeneity. A risk measure ρ is positively homogeneous if:

$$\kappa < 0, \quad \text{and} \quad \rho(\kappa X) = \kappa \rho(X). \tag{2.4}$$

Assuming that the portfolio is not too large, if the size of the portfolio doubles we can assume that it also requires twice the amount of capital. This is because as the size of a portfolio grows, its liquidity decreases, and a larger amount of capital may be necessary to maintain the appropriate proportions (Hull, 2018b, p. 275).

AXIOM IV. Translation invariance. A risk measure ρ is translation invariant, if:

$$\rho(X + \kappa) = \rho(X) - \kappa. \tag{2.5}$$

Adding the amount of capital κ to a portfolio should result in a decrease in the risk measure of the portfolio ρ by κ (Hull, 2018b, p. 275).

To demonstrate the non-subadditivity of VaR, consider this straightforward illustration. Consider a portfolio comprising two uncorrelated zero-coupon bonds, namely Bond X and Bond Y, each having a 4% probability of default. These bonds have a face value of $1000 and have only a chance to zero payout and in the event of default, their value reduces to $0. When examining VaR at the 95% confidence level, the VaR for both positions is zero (Table 1).[3]

Table 1: Subadditivity: Individual positions

Bond X			**Bond Y**		
Loss	Probability	Cumulative	Loss	Probability	Cumulative
$1000	4%	4%	$1000	4%	4%
$0	96%	100%	$0	96%	100%
VaR 95% = $0			VaR 95% = $0		

When the positions are combined, the chance of zero return decreases below 95%, so VaR of the combined portfolio is higher than the sum of the individual positions. This suggests that VaR is not subadditive, as diversification should reduce risk (Table 2).

[3] Please note that losses are presented as a positive number.

Table 2: Subadditivity: Combined positions

Combined		
Loss	Probability	Cumulative
$2000	0.16%	0.16%
$1000	**7.68%**	**7.84%**
$0	92.16%	100%
VaR 95% = $1000		

According to Daníelsson (2011) VaR is subadditive in the special case of normally distributed returns, but as we know, that is an unusual property of financial returns which is shown in various studies (see, e.g.,). Furthermore, if the assets in a portfolio have continuous payout functions, the VaR behaves subadditively. In contrast, when assets exhibit discontinuous payout functions near the critical VaR level, it tends to create problems with subadditivity. In practice, reality often falls somewhere between these two extremes, making it a more nuanced concept. However, in situations involving extensive and diversified portfolios, it could be argued that reality tends to align more closely with the continuous scenario (Miller, 2019). Given the lack of coherency of VaR as a risk measure, another risk measure, ES, has emerged as a more coherent alternative. ES satisfies all four properties of the axioms proposed by Artzner et al. (1999) for coherency.

2.4.3 Expected Shortfall

Artzner et al. (1999) initially suggested employing ES (also known as "conditional VaR", "mean excess loss", "beyond VaR" or "tail VaR") as a coherent risk measure to address the shortcomings of VaR. ES represents the average expected loss since the loss exceeds the threshold of VaR (Figure 2). Therefore, by its very definition, ES takes into account losses beyond the VaR level. Furthermore, it has been shown that ES exhibits more favorable mathematical properties compared to VaR. ES is subadditive, which ensures its coherency as a risk measure. (Sarykalin et al., 2008). Yamai, Yoshiba, et al. (2002) define $ES_\alpha(X)$ as follows: [4]

[4] $E[-X|B]$ represents the conditional expectation of the random variable $-X$ given the occurrence of event B. Typically, the profit-loss X tends to be negative when the loss exceeds VaR. Using $-X$ in the definition ensures that ES is expressed as a positive number. (Yamai, Yoshiba, et al., 2002)

$$ES_\alpha(X) = \mathbb{E}[-X \mid -X \geq \text{VaR}_\alpha(X)], \tag{2.6}$$

where X is a random variable that denotes asset outcome. VaR_α is the VaR at the $(1-\alpha)$% confidence level or interchangeably α^{th} quantile (Yamai, Yoshiba, et al., 2002).

The visual representation in Figure 2 illustrates the difference between VaR and ES when considering a confidence level of 95% and assuming a normal distribution. In this context, VaR represents the α quantile within the distribution of profits and losses. On the contrary, ES functions as a conditional expectation, illustrating the average magnitude of losses that exceed a specified VaR threshold at a given confidence level. The vertical centerline in the figure represents the mean of the distribution, which is 0 in the case of a normal distribution. The VaR and ES threshold are derived from the value of $\alpha = 0.05$, from which we obtain a critical value, i.e., the z-value. In the case of the VaR depicted on the chart, this z-value is approximately 1.65.

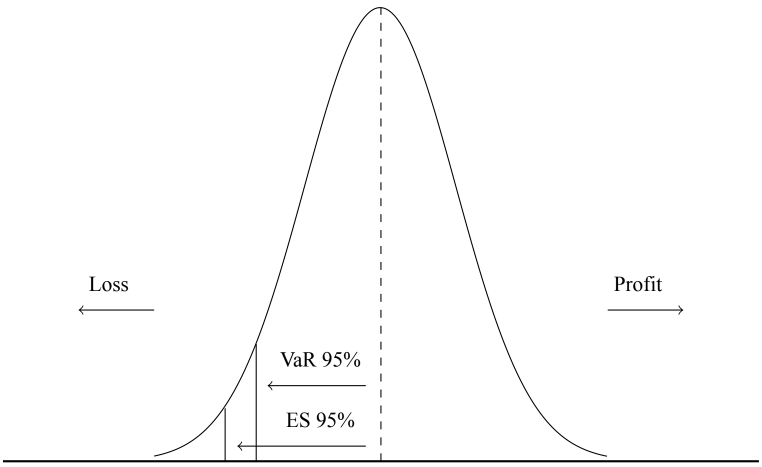

Figure 2: Profit-Loss Distribution, VaR and ES

Given that ES is derived from VaR, it acquires the inherent characteristics associated with VaR, including properties such as translation invariance, monotonicity, and positive homogeneity. Furthermore, it can be shown that ES also conforms to the principle of subadditivity, establishing itself as a coherent risk measure (Artzner et al., 1999). This particular advantage means that ES, in contrast to VaR, can encompass the diversification effect when evaluating the level of risk associated with a portfolio composed of various assets, each characterized by its unique risk attributes. However, ES has its share of drawbacks. First, ES is prone to errors associated with tail losses, as the magnitude of the expected value depends on all tail losses, and there is considerable uncertainty in the estimation of the tail distributions of these losses (see, e.g., Kondor 2014; Lönnbark 2013). Second, ES lacks a consistent scoring function (elicitability), making it impossible to reliably compare the estimated values with the actual ones that have a scoring function (Gneiting, 2011). This topic is covered in more detail in the following section.

2.4.4 Elicitability

Different estimation methods give different estimates of the magnitude of the risk, even if the same risk measure is used. When comparing estimates calculated with the same risk measure, it is useful if the goodness of the estimate can be judged against a pre-selected rule. A common way to examine the goodness of the calculated estimate is to compare the estimated value with the observed value. Gneiting (2011) discusses the making and evaluation of point forecast estimates in his publication. A natural way to assess the goodness of a computed estimate is to compare the estimated value to the observed value. Elicitability is a fundamental attribute of a risk measure in theoretical discourse due to its ability to facilitate the validation and comparative analysis of various risk measures derived from historical data. Risk measures that permit such validation and comparative evaluation are appropriately characterized as elicitable (Patton et al., 2019).

Next, a scoring function is defined according to Gneiting (2011). Consider the interval I as the possible spectrum of outcomes, where I equals the set of real numbers $\mathbb{R}$ for a real-valued quantity or I equals the open interval $[0, \infty]$ for a strictly positive quantity. Then, let $x \in I$ be point forecasts and $y \in I$, $I \subset \mathbb{R}$ observed values. The function S: $I \times I \Rightarrow [0,\infty]$ is then a scoring function that describes the magnitude of the prediction error x when the value y is observed. For example, we can think of the function $S(x, y)$ that provides the magnitude of the average error by examining the average goodness of the estimate based on n observations as a mean:

$$\bar{S} = \frac{1}{n} \sum_{i=1}^{n} S(x_i, y_i), \tag{2.7}$$

Based on the scoring function, we can choose the optimal point Various scoring functions suitable for different situations have been discussed, for instance, in Gneiting (2011) publication. Table 3 provides examples of commonly used scoring functions.

Table 3: Commonly used scoring functions (Gneiting, 2011)

$S(x, y) = (x - y)^2$	squared error (SE)
$S(x, y) = \|x - y\|$	absolute error (AE)
$S(x, y) = \|(x - y)/y\|$	absolute percentage error (APE)
$S(x, y) = \|(x - y)/x\|$	relative error (RE)

Next, a consistent scoring function is defined acording to Gneiting (2011) publication. Estimators and predictions are often statistical functionals, meaning they are mappings from the set of probability distributions $\mathcal{P}$ to Euclidean space. Risk measures are also statistical functionals. When the future value Y of the estimable variable is associated with an unknown distribution U, for a certain set of distributions $\mathcal{P}$, a scoring function $S(x, y)$ defined for a risk measure ρ: $\mathcal{P} \Rightarrow \mathbb{R}$ is consistent if, for all $U \in \mathcal{P}$, $\hat{p} \in \rho(U)$, and $x \in I$, the following holds:

$$\mathbb{E}_U[S(\hat{p}, Y] \leq \mathbb{E}_U[S(x, Y)] \tag{2.8}$$

The scoring function is strictly consistent if the equality holds only when $x \in \hat{p}(U)$.

If there exists a consistent scoring function for the functional, then for the optimal point forecast $\hat{x}$ of the functional, the following holds:

$$\hat{x} = \arg\min_x \mathbb{E}_U[S(x, Y)]. \tag{2.9}$$

When the scoring function is strictly consistent, it ensures that the optimal estimate is unambiguous. Therefore, in cases where there exists a consistent scoring function for a given functional, it becomes a valuable tool for comparing estimates derived from different methodologies. This attribute proves particularly beneficial when faced with multiple statistical models for estimation, allowing a rigorous comparison to select the most appropriate one. (Gneiting, 2011)

The presence of a truly consistent scoring function for a functional in a range of probability distributions in the set $\mathcal{P}$ is termed "elicitable". Gneiting (2011) demonstrates that elicitability is a mathematical property satisfied by VaR but not ES, which implies that backtesting of ES is more complicated than backtesting of VaR. Acerbi and Szekely (2014) later proved that ES is backtestable and that elictability is not a mandatory property for absolute model selection. The feature makes the comparison easier, but does not prevent the implementation testing of models. The VaR risk measure is consistently scoring, and its strictly consistent scoring function can be defined as follows:

$$S(x, y) = (1\{x \geq y\} - \alpha)(g(x) - g(y)), \tag{2.10}$$

where g is an increasing function and 1 denotes the indicator function (Gneiting, 2011).

2.4.5 Summary of risk measure characteristics

ES is a more comprehensive measure of risk than VaR, as it clarifies the potential magnitude of losses in adverse scenarios. On the contrary, the measure VaR has been criticized for its limited ability to identify losses above its defined threshold, leading to an incomplete description of the true extent of potential losses, especially when it does not take into account the possibility of significantly larger losses. Furthermore, the non-subadditive nature of VaR highlights its inconsistency, since the combined risk of the positions may in some cases exceed the sum of the risks of the individual positions.

ES, in contrast, is a subadditive and coherent risk measure that offers a more robust framework for risk assessment. However, ES is not a nonproblematic risk measure; it lacks consistent scorable attributes, which implies the absence of a scoring function for consistent comparisons between the estimated and actual values of ES. Furthermore, its sensitivity to errors in tail losses, stemming from its dependence on cumulative tail losses, poses a challenge considering the inherent uncertainty associated with tail loss distribution estimates (Jiménez et al., 2020b).

Table 4: Summary of the properties of VaR and ES risk measures

Property	VaR	ES
Translation invariance	Yes	Yes
Subadditivity	No	Yes
Positive homogeneity	Yes	Yes
Monotonicity	Yes	Yes
Coherency	No	Yes
Elicitability	Yes	No
Backtestability	Yes	Yes

2.5 Risk measure estimation windows

The estimation window denotes a defined time frame used to calculate statistical factors or inputs within risk models. Its importance lies in shaping the accuracy and dependability of forecasts or estimates within risk modeling. A carefully selected estimation window has the potential to increase the foresight capabilities of risk models by encompassing pertinent market conditions, thus improving the accuracy of future risk assessments (Berens et al., 2018). The estimation window is illustrated in Figure 3 according to Daníelsson (2011). The total sample n is divided into an estimation window of size W_E and a testing window of size $W_T = n - W_E$.

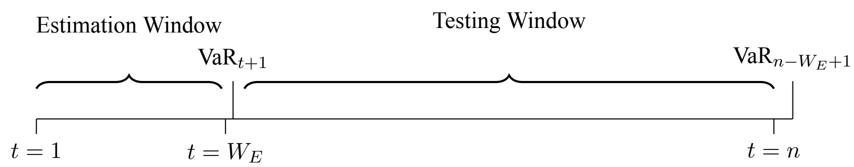

Figure 3: Estimation and Testing Windows

The characteristics of the asset under evaluation, including its market liquidity, volatility, and dynamics, significantly influence the optimal estimation window duration. Selecting the right estimation window involves a trade-off: too few observations create high volatility in the training sample, while too many introduce biases not present during the modeling period. Therefore, determining the ideal training window size is not straightforward. Assets exhibiting higher volatility may require shorter windows for accurate estimates, while illiquid or less volatile assets might benefit from longer windows to capture unusual patterns. (Righi & Ceretta, 2015)

Understanding how the estimation window affects the accuracy of the risk estimation is essential. According to the BCBS (2019b), the construction of a risk model requires a minimum of 250 days of historical data to construct a market risk model for traditional assets. However, the current regulatory framework for cryptocurrencies is not yet in place, which is why there is no minimum requirement for the length of estimation window. In traditional financial studies, the norm is to assume an extended estimation window, often spanning 1000 to 2000 observations or more (Danielsson 2002; Buczyński and Chlebus 2019; Patton et al. 2019).

However, assessing market risk in cryptocurrencies does not allow such extended windows due to their relatively recent emergence compared to traditional assets. In particular, in periods of high volatility, an even shorter estimation window becomes necessary to evaluate quick changes in risk levels. Despite numerous studies that introduce and compare estimation methods, there is a relative oversight in understanding the significance of the information from the estimation window. This is evident in the limited number of comprehensive studies on this topic, which contributes to the lack of empirical consensus (Righi and Ceretta 2015; Buczyński and Chlebus 2020). Since the study of cryptocurrencies does not allow the use of long estimation windows and shorter estimation windows that adapt more quickly to large changes in volatility, this study aims to investigate how the estimation window set by BCBS performs against shorter estimation windows.

2.6 Overview of VaR and ES backtesting

Risk mitigation comprises two fundamental components: the estimation of the loss distribution and the calculation of the risk magnitude based on this distribution using a selected risk measure. The specific risk measure chosen to gauge the magnitude of the risk, along with the method used to estimate the loss distribution, collectively constitutes the risk model (Hult, 2012). The preceding chapter dived into the theoretical characteristics of risk measures that should be considered when determining the most suitable risk measures for risk estimation. In addition to selecting the risk measure, it is imperative to choose an appropriate method to conduct the risk estimation. Different methods can produce notably disparate risk assessments, even when employing the same risk measure. As a result, the estimated risk values are subject to empirical validation and validation against real-world data.

Backtesting involves statistically testing the estimates of risk measures by comparing them with actual losses from a previous period. If an internal model computes the recurrent capital charge, regulatory requirements require regular backtesting of the risk measure estimation model. Failure of performance tests that indicate inconsistencies between risk estimates and observed losses can lead to the prohibition of using the internal model. Furthermore, the internal model risk estimates are adjusted by coefficients that affect the size of the capital requirement based on the results of the performance test (Daníelsson, 2011).

When estimating VaR, we establish a specific confidence level that defines the position (quantile) within the profit-loss distribution that we are examining. Backtesting VaR is rather straightforward, the simplest way to evaluate the accuracy of VaR, by using α we calculate the expected amount of observations which should fall to the left quantile of this particular quantile in profit-loss distribution. Thus, if we want to estimate VaR with a 99% level of confidence, we will analyze the first quantile, represented by $\alpha = 0.01$. For instance, given a dataset comprising 1000 observations, we would expect approximately 0.01 * 1000 = 10 observations (also called violations), falling beyond the VaR threshold. To calculate the difference between expected and observed violations, we can calculate VR as follows:

Let X_t denote asset outcome, η_t denote whether a violation occurs at time t, let v_1 denote the count of violations and v_0 count of non-violation.

VaR is said to be violated at time t when $\eta_t = 1$, where:

$$\eta_t = \begin{cases} 1, & \text{if } X_t \leq -\text{VaR}_t \\ 0, & \text{if } X_t > -\text{VaR}_t \end{cases}$$

The amounts of violations and non-violations are counted as follows:

$$v_1 = \sum_{W_{E+1}}^{W_T} \eta_t, \tag{2.11}$$

$$v_0 = W_T - v_1. \tag{2.12}$$

From where we can calculate VR:

$$VR = \frac{v_1}{\alpha W_T}. \tag{2.13}$$

VR greater than one means VaR underestimates the risk, while VR smaller than mean VaR overestimates the risk.

The testing of ES accuracy is considerably more complicated compared to VaR due to the infinite potential outcomes for each day. Additionally, a significant distinction in the models' performance testing lies in the fact that most evaluations of ES performance necessitate data regarding the daily return distribution, or at the very least information about the distribution's tail. As a result, backtests for ES inherently rely on approximations and are highly susceptible to inaccuracies in the predicted VaR. (Daníelsson, 2011)

This is because while VaR examines a single quantile, ES evaluates an expectation. In simpler terms, we can pinpoint when VaR is violated, but detecting such violations for ES is challenging. However, there is a straightforward approach to the test of ES, similar to the violation ratios used in the evaluation of VaR. When VaR is violated on a specific day, according to Daníelsson (2011) NS is calculated as follows:

$$NS_t = \frac{X_t}{ES_t}, \tag{2.14}$$

where ES_t is the empirical ES on day t.

From the definition of ES, the expected X_t given that VaR is violated is as follows:

$$\frac{\mathbb{E}[X_t | X_t < -VaR_t]}{ES_t} = 1 \tag{2.15}$$

Therefore, the average NS, $\overline{NS}$ should be one:

$$H_0 : \overline{NS} = 1 \tag{2.16}$$

Daníelsson (2011) defines a useful rule of thumb for evaluating the accuracy of VaR and ES using VR and NS as follows:

Table 5: Violation Ratios (Daníelsson, 2011)

VR(NS) [0.8, 1.2]	Good
VR(NS) [0.5, 0.8] or VR(NS)[1.2, 1.5]	Acceptable
VR(NS) [0.3, 0.5] or VR(NS)[1.5, 2]	Bad
VR(NS) < 0.3 or VR(NS) > 2	Useless

Note: In presenting the results of the study, a similar scale and color coding is used to evaluate the accuracy of VaR and ES.

Chapter 4.5 introduces additional backtesting methods in addition to VR and NS. These methods possess unique characteristics compared to those previously discussed in this chapter. Consequently, their calculation and focus vary slightly from these basic ones, offering further insight into the performance of VaR and ES risk measures.

3 Data and methodology

In this chapter, the used data are presented and the properties of time series are explored using various statistics and tests. The parametric and non-parametric methods utilized to estimate the VaR and ES models are then presented. Furthermore, the time windows used and their relation to the calculation of risk estimates are outlined. Finally, different backtest methods are presented to assess the accuracy of the VaR and ES risk measures.

3.1 Data and descriptive statistics

This research incorporates the price data for Bitcoin and Ethereum along with S&P 500 data, serving as a reference. Cryptocurrency data are obtained from https://coinmarketcap.com/, while S&P 500 index data is obtained from https://finance.yahoo.com/. All analyzes are performed using MATLAB software (The MathWorks Inc., 2023), and VaR and ES models, including their backtesting, were implemented using the Risk Management Toolbox (MathWorks, 2023).

Bitcoin and Ethereum collectively control a substantial majority of the overall cryptocurrency market, accounting for 68. 7% of the total market share (CoinMarketCap, 2023c), which is why they are chosen for this study. Respectively S&P500 is the world largest index by market capitalization. The study analyzes information spanning from the beginning of 2021 until the end of November 2023. Consequently, cryptocurrency data include 1,060 observations, whereas the S&P 500 comprises 730 observations within the same time frame, reflecting daily trading of cryptocurrencies.

To maintain the most accurate representation of its volatility, the missing values for the S&P500 are not filled in using moving averages. The chosen time frame is specifically selected to explore how market risk measures operate during periods of increased volatility, thus omitting the use of moving averages for the S&P500 to ensure a realistic comparison. Figure 4 shows the performance of both the S&P 500 and cryptocurrencies throughout the evaluation period. In particular, market volatility, particularly for cryptocurrencies, is evident from early 2021 until July 2022. An observable trend across all assets is a substantial price decline that occurs within a few months, followed by a subsequent price increase that began around the end of June or July 2022.

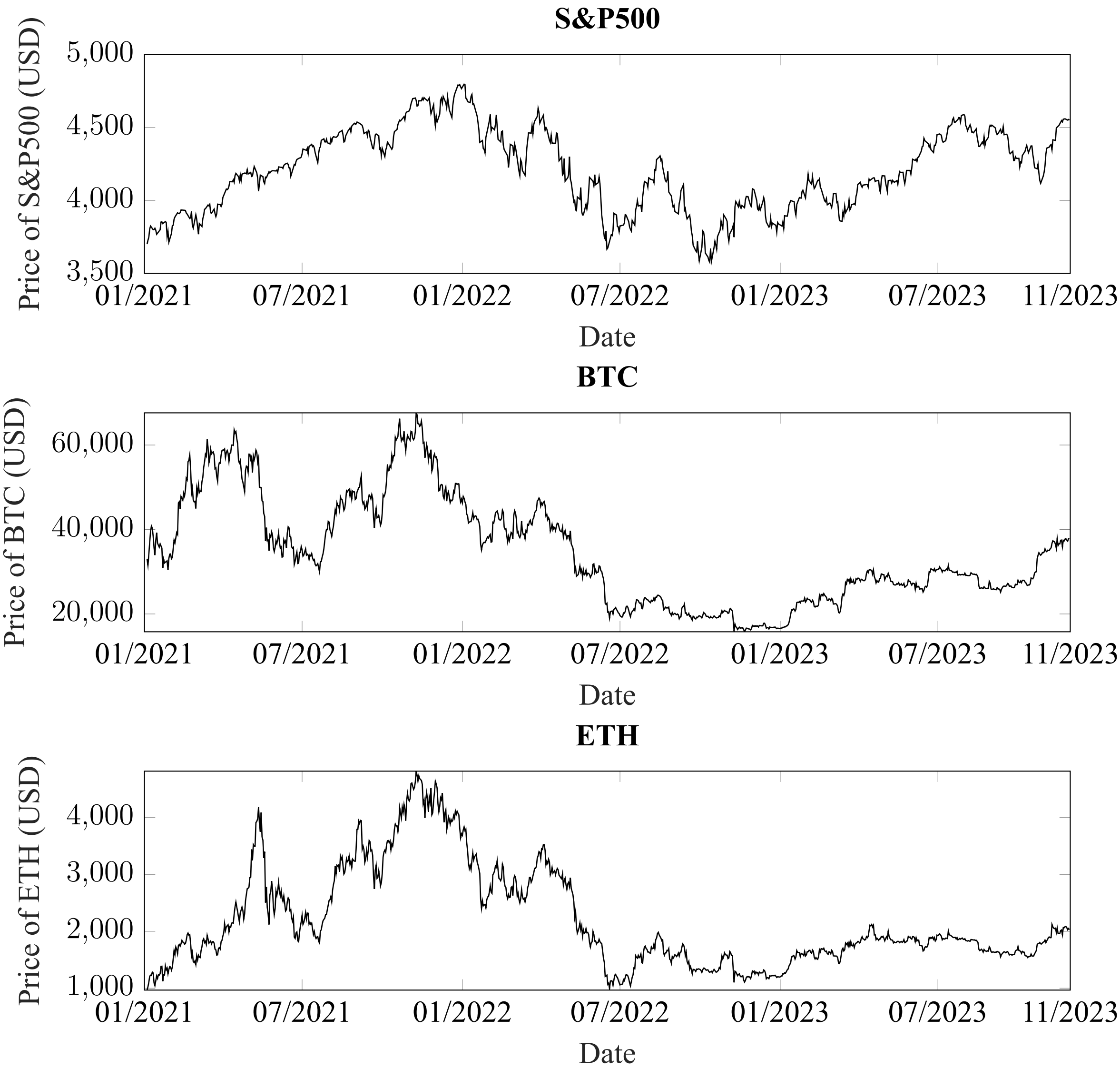

Figure 4: Returns of S&P500, BTC and ETH

Daily logarithmic returns are calculated using Equation (3.1). As illustrated in Figure 5, the volatility of cryptocurrencies is significantly elevated when compared to the S&P500. Consequently, we can readily observe indications of volatility clustering with both BTC and ETH. Instances of pronounced volatility are succeeded by abrupt fluctuations in returns, whereas periods lacking such spikes generally exhibit more subdued movements. Since the volatility of the S&P 500 being significantly lower, it is noticeably more challenging to draw equally clear conclusions solely from the figure. However, it can be observed that volatility reaches its peak between 2022 and 2023.

$$R_t = log\left(\frac{P_t}{P_{t-1}}\right), \tag{3.1}$$

where R_t denotes the daily logarithmic return of an asset on day t, P_t denotes the closing price of an asset on day t and P_{t-1} denotes the closing price of an asset on day $t-1$.

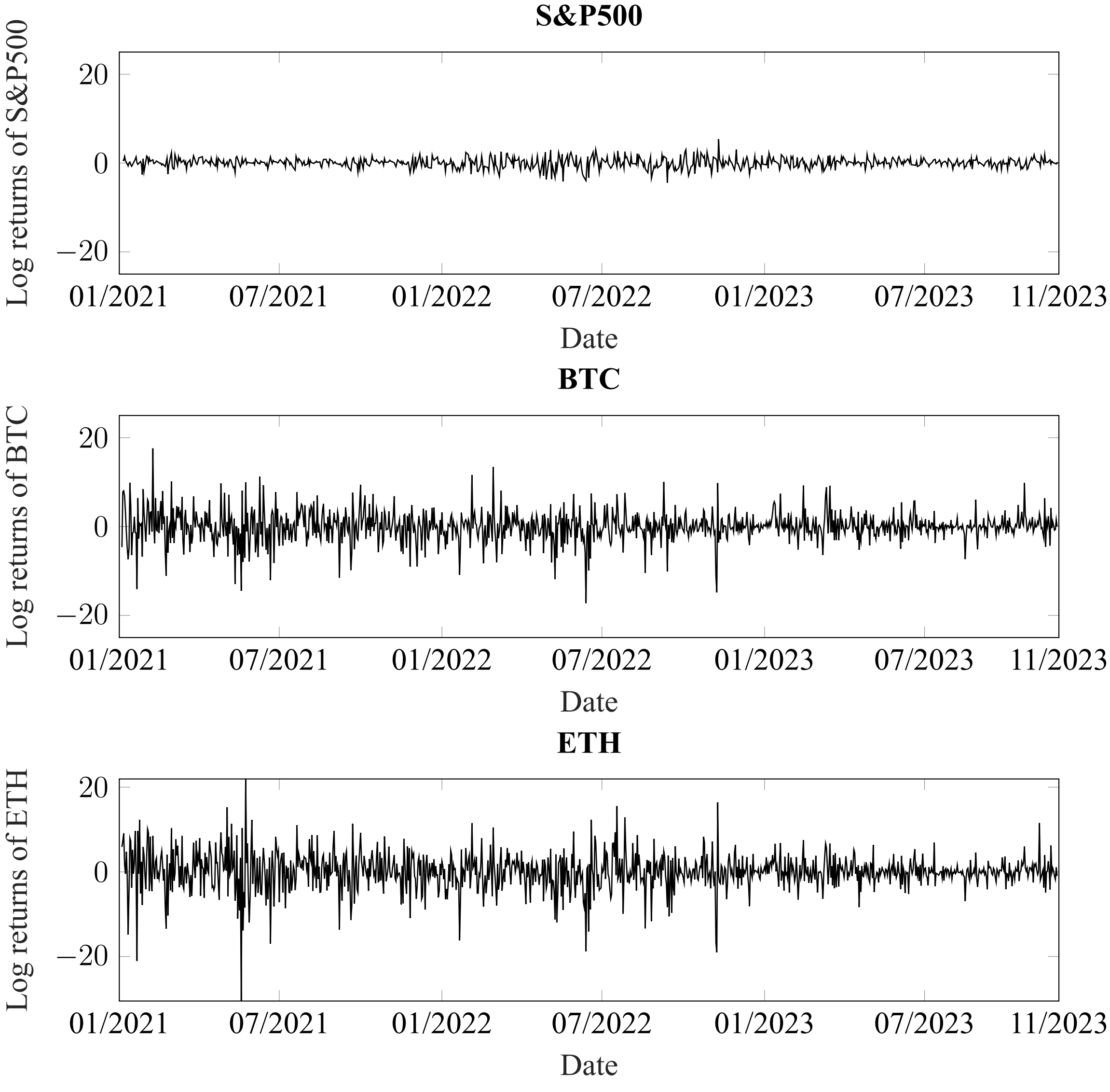

Figure 5: Logarithmic returns of S&P500, BTC and ETH

Table 6 provides the descriptive statistics of the calculated logarithmic returns. Skewness and kurtosis are calculated using Equations (3.2) and (3.3). Bitcoin has the lowest mean and median, while Ethereum has the highest mean and median leaving S&P500's mean and median of S&P500 between both cryptocurrencies. The means and medians are quite similar for all, i.e. close to zero, but there are significant differences in the minimum and maximum values, indicating the existence of outliers. In the case of cryptocurrencies display more extreme outcomes compared to S&P500. Therefore, it is unsurprising that cryptocurrencies exhibit a higher variance and standard deviation compared to S&P500.

$$Skewness(X) = \frac{1}{(n-1)s^3} \sum_{t=1}^{n} (x_t - \hat{\mu})^3. \tag{3.2}$$

$$Kurtosis(X) = \frac{1}{(n-1)s^4} \sum_{t=1}^{n} (x_t - \hat{\mu})^4. \tag{3.3}$$

Table 6: Descriptive statistics

Statistic	BTC	ETH	S&P500
Mean	0.0129	0.0700	0.0283
Std. Dev.	3.3919	4.3858	1.1201
Variance	11.5047	19.2353	1.2547
Min	-17.2520	-30.5201	-4.4199
Q1	-1.3805	-1.8383	-0.6151
Median	-0.0269	0.0675	0.0277
Q3	1.5727	2.2809	0.7043
Max	17.6026	21.9406	5.3953
Skewness	-0.1962	-0.5449	-0.1922
Kurtosis	6.4725	7.8347	4.5802

The skewness and kurtosis of the normal distribution are constants that are equal to 0 and 3, respectively.

All assets, especially BTC and ETC, exhibit high kurtosis, meaning that return distributions are much more leptokurtic than in a normal distribution. Negative skewness indicates that the left tail of the distribution is longer than is assumed in the normal distribution. These results are not in line with the study of Auwera et al. (2020) in which the logarithmic returns of both cryptocurrencies and traditional assets were examined. The study found that the logarithmic returns of cryptocurrencies show positive skewness, a characteristic not commonly observed in traditional capital markets. Because the statistics indicated that time series are not normally distributed, Student's t-distribution is fitted to the data by using the parameters (mean, standard deviation, and degrees of freedom) obtained from the data. The fitted Student's t-distributions are illustrated in Figure 6, ν denotes the degree of freedom of Student's t-distribution.

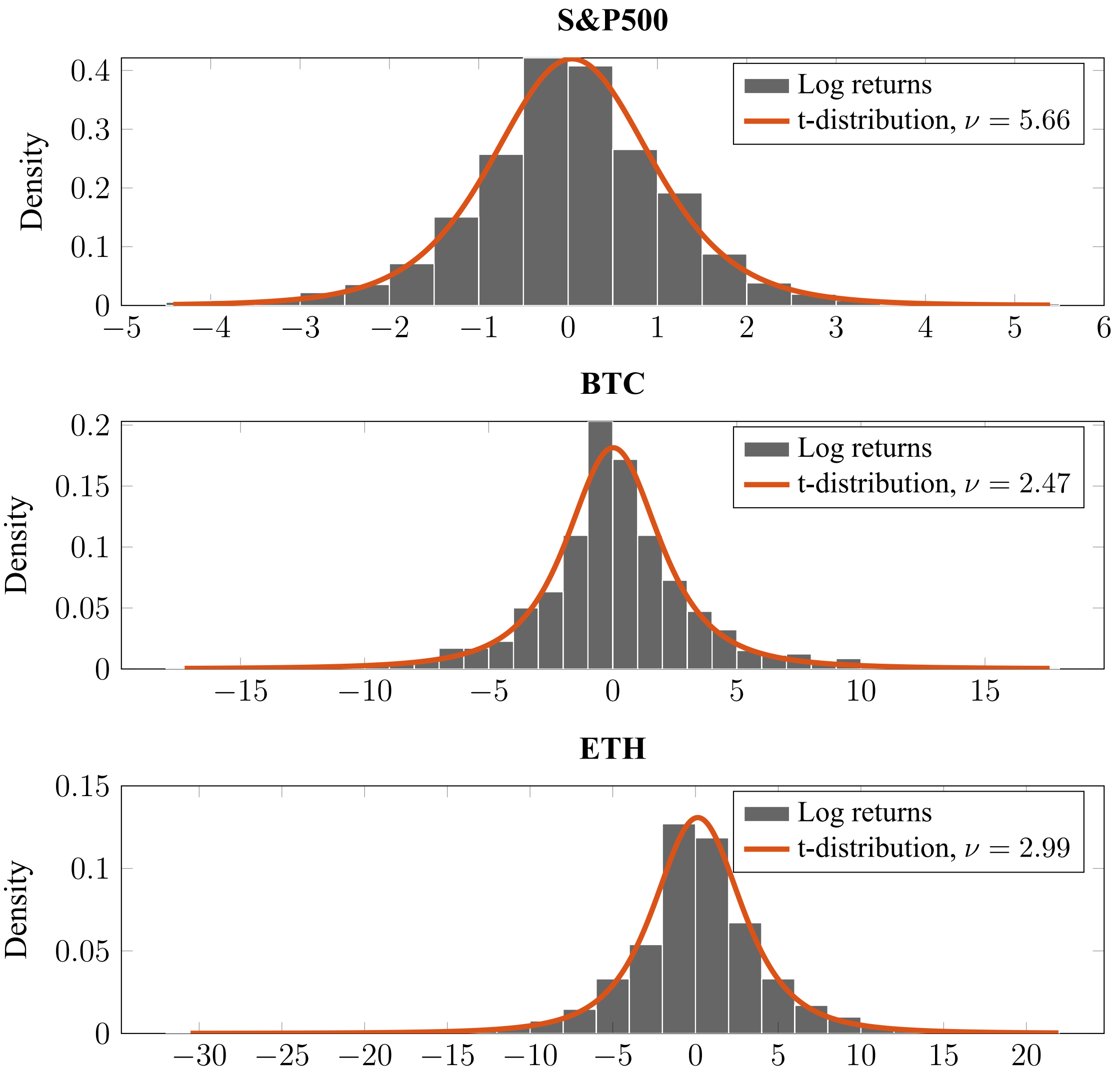

Figure 6: Distribution fitting

As depicted in Figure 6, each time series has different shaped distributions. Additionally, if we consider the Student's t-distribution, which is expected to have a similar fit to the normal distribution, it would have approximately $\nu = 30$. The degrees of freedom parameter plays a crucial role in shaping the Student's t-distribution, offering control over characteristics such as tail thickness. Specifically, an increase in degrees of freedom results in a more peaked distribution with thinner tails, whereas a decrease leads to a less peaked distribution with thicker tails. Therefore, it is not surprising that cryptocurrencies, characterized by more extreme outcomes, require lower degrees of freedom in comparison to the S&P 500 when fitting their distributions.

The results of various statistical tests applied to the time series used and their statistical significances are summarized in Table 7. Considering the high skewness of all assets, revealed in the descriptive statistics, it is expected that all the series would deviate from normality, as indicated by the rejection of the null hypothesis of normality in the Jacque-Bera (J.B) test. Both the J.B test and descriptive statistics lead to the conclusion that all distributions exhibit thick tails. This implies that estimation methods that assume normality may not be the most suitable for VaR and ES estimation.

The Ljung-Box Q-test (LB) for residual autocorrelation suggests that there is a lack of substantial evidence to dismiss the null hypothesis, which asserts the absence of residual autocorrelation for up to 20 lags. This suggests that the autocorrelations observed at these lags do not reach statistical significance, and consequently, there is insufficient support to affirm the existence of notable autocorrelation within the residuals. However, a repeated LB test (LB-2) over 20 lags for the squared returns reveals significant ARCH effects in the residuals of the returns.

Table 7: Statistical hypothesis testing

Statistic	BTC	ETH	S&P500
J.B	538.9***	1083.8***	80.6***
LB	17.7	30.8	20.6
LB-2	81.5***	202.2***	242.6***
Engle	8.7**	17.7***	7.3**
KPSS	.12	.13	.09

(*) p < .05 (**) p < .01 (***) p < .001.

Respectively, the Engle test for residual heteroskedasticity rejects the null hypothesis of no conditional heteroskedasticity for all the assets in concern. Thus, we can conclude that there are significant ARCH effects in the return series, which makes it appropriate for GARCH modeling. Additionally, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test fails to reject the trend-stationary null hypothesis, meaning that all the return series showed the ability to achieve stationarity. This is to be expected, as the use of logarithmic returns implies that a logarithmic transformation took place, which stabilizes the variance of the time series.

3.2 Procedure

After the statistical tests are performed, the next phase involves constructing a model and providing a comprehensive overview of the estimation procedure. This research evaluates the losses of individual assets using VaR and ES risk measures, employing four distinct estimation methods (Gaussian VCV, Gaussian GARCH (1,1), Student's t GARCH (1,1) and HS). Furthermore, by using these methods, five different estimation windows are applied to each model: $W_E = [50, 100, 150, 200, 250]$. The fixed-size estimation window is composed of observations from 2021. The actual length of the estimation window is kept the same for easier numerical comparability. However, it is crucial to highlight that when dealing with S&P 500 returns, a 250-day estimation window encompasses nearly all yearly observations, while this is not the case with cryptocurrencies due to their continuous daily trading.

The testing window W_T begins at observation 252 for the S&P500 and at observation 365 for the cryptocurrencies, corresponding to the initial available observations in 2022. Consequently, the first estimates of VaR and ES are calculated at $W_T + 1$ and the last estimates at $n - W_E + 1$. It is standard practice to compute the VaR and ES estimates for a single day ahead (Daníelsson, 2011). The testing phase incorporates the rolling-window method to assess model stability and performance over time. This method involves shifting the window forward one observation at a time, with the process repeated until the end of the dataset is reached (Table 8). The confidence levels in the models adhere to the BCBS mandates: 99% for VaR and 97.5% for ES, which are approximately equivalent to each other (BCBS, 2019b).

Table 8: Estimation process

S&P500				**Cryptocurrencies**			
t	$t + W_E - 1$	VaR$(t + W_E)$	ES$(t + W_E)$	t	$t + W_E - 1$	VaR$(t + W_E)$	ES$(t + W_E)$
1	252	VaR(253)	ES(253)	1	365	VaR(366)	ES(366)
2	253	VaR(254)	ES(254)	2	364	VaR(367)	ES(368)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
479	731	VaR(732)	ES(732)	696	1059	VaR(1060)	ES(1060)

After the estimation process, the accuracy of the models is tested using various backtesting methods, after which their accuracy is evaluated. In the next sections, the estimation methods and backtesting methods used are presented, and their characteristics are briefly discussed.

3.3 Parametric methods

After analyzing the stylized facts of the returns, different estimation methods are introduced, starting from parametric methods. Numerous parametric methods exist to estimate VaR and ES. Nadarajah et al. (2017) identify over 100 methods that encompass various distributional assumptions (e.g., Gaussian, Johnson family, Student's t distributions, etc.). These approaches can utilize various econometric models (e.g., EWMA, ARMA, or GARCH methods). The downside of parametric methods are that they rely on assuming a specific model, thus carrying substantial model risk.

3.3.1 Gaussian Variance-Covariance

The VCV has various names, such as the delta-normal method and parametric VaR, however, the term delta is closely related to option pricing and may imply a linear relationship between VaR and risk factors, which is not the case when dealing with portfolios containing derivatives. Similarly, the term "normal" can be misleading, since the risk factor can also be log-normal or follow a Student's t-distribution. Hence, the term "Variance-Covariance Method" is used.

In this study, for the estimation of VaR and ES using the VCV method under the assumption of Gaussian distribution, we make the assumption that the loss X_{t+1} follows a multivariate normal distribution, unconditionally or conditionally. Furthermore, we assume that the linearized loss, expressed in terms of risk factors, serves as a reasonably precise approximation of the actual loss. Subsequently, we utilize the unbiased estimators for the mean and variance of the population, denoted as μ and σ^2, respectively. These estimators are calculated as follows:

$$\hat{\mu} = \frac{1}{n}\sum_{t=1}^{n} x_t, \qquad (3.4) \qquad \hat{\sigma}^2 = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}(x_t - \mu)^2} \qquad (3.5)$$

$\text{VaR}_{\alpha,\text{t}}$ under the assumption of Gaussian distribution is estimated as follows:

$$\widehat{VaR}_{\alpha,t} = \hat{\mu} + \hat{\sigma}\Phi^{-1}(\alpha), \tag{3.6}$$

and ES_{α}, respectively:

$$\widehat{ES}_{\alpha,t} = \hat{\mu} + \hat{\sigma}\frac{\phi(\Phi^{-1}(\alpha))}{\alpha}, \tag{3.7}$$

where $\Phi^{-1}(\alpha)$ is the normal inverse cumulative distribution function and ϕ is the density function (McNeil et al., 2005).

The VCV provides a straightforward analytical approach to address the problem of risk measurement. However, this simplicity comes with the trade-off of relying on two oversimplified assumptions. First, the linearization approach may not provide an accurate approximation of the connection between the actual loss distribution and changes in risk factors. Second, it is improbable that normality is able to accurately represent the distribution of risk-factor changes, particularly when working with daily data and possibly even with weekly or monthly data. (McNeil et al., 2005, p. 49) One significant issue arises from the presence of heavy tails in the return distribution of most financial and digital assets (see, e.g., Wong and Vlaar 2003; Koutmos; 2018; Auwera et al. 2020). These heavy tails are of particular concern because VaR aims to depict the behavior of the portfolio's returns in the left tail. In such circumstances, a model that relies on a normal distribution would underestimate the number of extreme events, leading to an underestimated VaR (Jorion, 2007, p. 262). This is also a concern in this study since, as we previously concluded, our data exhibit thicker tails than normal distribution.

3.3.2 GARCH(1,1) applications

The GARCH model, an extension developed by Bollerslev (1986) following Engle's ARCH model, is designed to address volatility clustering and the leptokurtic distribution of price returns, both recognized as stylized statistical properties of returns (Barjašić & Antulov-Fantulin, 2021). In this study, statistical tests have revealed the presence of significant ARCH effects and leptokurtic distributions. To accommodate these characteristics, a GARCH(1,1) model is introduced that considers time-dependent volatility. Empirical studies suggest that a GARCH (1,1) model with a lag of past variance and one lag of past innovation is generally sufficient to capture the dynamics of volatility (Asgharian et al., 2021). Various studies (for example, Trucíos and Taylor 2023; Jiménez et al. 2020a; Acereda et al. 2020), utilize GARCH(1,1) applications with different distributional assumptions.

The conditional variance of the GARCH(1,1) model is calculated as follows:

$$\sigma_t^2 = \kappa + \gamma\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{3.8}$$

where σ_t^2 is conditional variance at time t, κ is the constant term representing the long-term average of the conditional variance, ϵ^2 is lagged squared residuals, γ is the coefficient of the lagged squared residuals and β is the coefficient of lagged conditional variance.

Thus, $\mathrm{VaR}_{\alpha,\mathrm{t}}$ under the assumption of Gaussian distribution is estimated as follows:

$$\widehat{VaR}_{\alpha,t} = \hat{\mu} + \hat{\sigma}_t\Phi^{-1}(\alpha), \tag{3.9}$$

and $\mathrm{ES}_{\alpha,\mathrm{t}}$, respectively:

$$\widehat{ES}_{\alpha,t} = \hat{\mu} + \hat{\sigma}_t\frac{\phi(\Phi^{-1}(\alpha))}{\alpha}. \tag{3.10}$$

where $\hat{\sigma}_t$ is estimated conditional volatility.

The assumption of a normal distribution of returns simplifies the estimation of risk estimates. However, findings from studies such as Auwera et al. (2020) and Fung et al. (2022) propose that the distribution of returns in cryptocurrencies is better approximated by Student's t-distribution. Taking into account the aforementioned results and descriptive statistics, Student's t applications are included in this study. Unlike the Gaussian distribution, which incorporates only location parameter μ (mean) and scale parameter σ (standard deviation), the Student's t-distribution also includes a shape parameter ν (degrees of freedom).

According to Gaunt (2021) the Student's t-distribution can be expressed as follows. Let $X \sim T_\nu$ follow Student's t-distribution with ν greater than 0 and PDF:

$$f_X(x) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\pi\nu}\Gamma(\nu/2)}\left[1+\frac{x^2}{\nu}\right]^{-\frac{\nu+1}{2}} \tag{3.11}$$

where Γ denotes gamma function and ν denotes degrees of freedom.

From the above equation, we can derive the Student's t-distribution into the GARCH (1,1) process. Thus, $\text{VaR}_{\alpha,\text{t}}$ under the assumption of Student's t-distribution is estimated as follows:

$$\widehat{VaR}_{\alpha,t} = \hat{\mu} + \hat{\sigma}_t T_{\nu}^{-1}(\alpha), \tag{3.12}$$

and $\text{ES}_{\alpha,\text{t}}$, respectively:

$$\widehat{ES}_{\alpha,t} = \hat{\mu} + \hat{\sigma}_t \frac{\phi(T_{\nu}^{-1}(\alpha))}{\alpha}, \tag{3.13}$$

where T_{ν} is an inverse cumulative distribution function of Student's t-distribution with ν degrees of freedom.

In the process of estimating with different parametric methods, we opt for simplicity by not estimating the mean, denoted as μ. The assumption that $\mu = 0$ is considered relatively inconsequential, given that the error remains negligible at the daily level, as noted by Daníelsson (2011). This observation is also consistent with the findings of this study (recall Table 6).

3.4 Non-parametric methods

Non-parametric models are approaches used to estimate VaR and ES without making explicit assumptions about the underlying probability distribution of asset returns. Non-parametric models offer flexibility, proving beneficial in situations where data deviates from standard distributions or during financial turbulence when distributions are variable. However, they can be more prone to the influence of outliers compared to parametric models.

3.4.1 Historical Simulation

HS is widely used non-parametric method for estimating VaR and ES (see, e.g., Likitratcharoen et al. 2023; Pele and Mazurencu-Marinescu-Pele 2019). The risk estimates are calculated from the empirical distribution of the outcomes of the assets. The method is a substantive non-parametric approach to calculate VaR and ES due to its straightforwardness and simplicity. For the empirical VaR estimator, historical losses $x_1, \ldots, x_t$ are first ranked in ascending order, and the i-th smallest observation is denoted as $x(i)$ (McNeil et al., 2005). In this case, VaR at the confidence level α can be estimated as follows.

$$\widehat{VaR}_{\alpha,n}(x) = x_{([n\alpha])}, \tag{3.14}$$

where $[n\alpha]$ is the largest integer smaller than $n\alpha$.

The ES indicates the expected value of the loss when it exceeds the loss level given by the VaR of the same confidence level. The simplest way to estimate this expected value is to calculate the average of the losses above the VaR level:

$$\widehat{ES}_{\alpha,n}(x) = \frac{1}{n - [na] + 1} \sum_{i=[n[n\alpha]]}^{n} x_{(i)}, \tag{3.15}$$

Using HS avoids typical problems associated with the parametric approach, the three most significant being the assumption of a normal distribution of returns, the assumption of constant correlation, and the assumption of a constant delta (Szylar, 2013). Instead of making assumptions about the return distribution, HS relies on actual data, using the empirical distribution, which is a significant advantage over parametric methods because it avoids estimation errors (Dowd & Blake, 2006). Although HS is widely used and easy to execute, the method also has its drawbacks. Since most extreme observations fluctate a lot more than observations that are less extreme, HS benefits from a larger sample size. However, the downside is that the old data may not be representative. Furthermore, if the data have structural breaks, HS tends to perform inaccurately (Daníelsson, 2011). The precision of HS is based on having the right data points that truly reflect the changing dynamics of the market. Therefore, the estimation method needs sufficient data without including all historical dynamics because markets never perfectly mirror the past.

3.5 Backtesting procedures

After the creation of the models and before practical applications, it is crucial to thoroughly evaluate the reliability of the models. Furthermore, during its implementation, it is vital to consistently assess its efficacy. A pivotal aspect of confirming the accuracy of the models involves performing various backtests. In this study, VaR is backtested with three different backtests and ES with one backtest. Furthermore, VRs and NSs (recall Section 2.6) are calculated, respectively.

3.5.1 Kupiec's proportion of failure test

The widely used VaR backtest method in the context of traditional assets and cryptocurrency (see, e.g., Likitratcharoen et al. 2018; Jiménez et al. 2020b; Odel et al. 2019), was presented by Kupiec (1995). The test mimics the framework of Bernoulli trials. Next, the Bernoulli trials framework is defined according to Daníelsson (2011).

The Bernoulli density on day t is given by:

$$(1-p)^{1-\eta_t}(p)^{\eta_t}, \quad \eta_t = 0, 1 \tag{3.16}$$

The likelihood function is given by:

$$LU(\hat{p}) = \prod_{t=W_E+1}^{n} (1-\hat{p})^{1-\eta_t}(\hat{p})^{\eta_t} \tag{3.17}$$

The restricted likelihood function is defined as follows:

$$LR(p) = \prod_{t=W_E+1}^{n} (1-p)^{1-\eta_t}(p)^{\eta_t} \tag{3.18}$$

Using Proportion of Failure (POF) we test whether $LR = LU$ or equivalently, whether $p = \hat{p}$:

$$LR_{POF} = -2log\frac{(1-\hat{p})^{v_0}(\hat{p})^{v_1}}{(1-p)^{v_0}(p)^{v_1}} \sim \chi_1^2. \tag{3.19}$$

The test follows a chi-square (χ_1^2) distribution with one degree of freedom and evaluates the null hypothesis, which posits that the actual violation rate is equal to the observed violation rate. The POF test is two-sided, which means that it is capable of testing if there are too many or too few violations. If the statistic value of LR_{POF} exceeds the 99th percentile of a chi-square distribution with one degree of freedom, Kupiec (1995) suggests rejecting the null hypothesis.

3.5.2 Christoffersen's tests

Backtesting a VaR model involves assessing the probability of encountering violations within a given sample period. However, if these violations occur in clusters for consecutive days, it can compromise the efficacy of the VaR model (Daníelsson, 2011). In response, Christoffersen (1998) introduced a Conditional Coverage Independence (CCI) test to determine whether the likelihood of observing a violation on a specific day depends on the occurrence of a previous violation. The test can be presented as a Markov chain with two states. Christoffersen's Conditional Coverage (CC) jointly tests previously mentioned POF and CCI tests using the chi-square with two degrees of freedom instead of one. Both of these tests are widely used to test VaR in a cryptocurrency context (see, e.g., Jiménez et al. 2020a; Liu et al. 2020).

Following the framework outlined by Daníelsson (2011), the two-state Markov chain is subsequently characterized. The p_{00} denotes the probability for a period without violations is succeeded by another period without violations. Similarly, p_{10} signifies the probability that a day with a violation is followed by a day without any violations. Furthermore, p_{01} represents the probability that a day without violations is succeeded by a day with a violation, while p_{11} captures the probabilities of two consecutive violations. This can be represented equivalently $p_{ij} = P(\eta_t = j|\eta_{t-1} = i)$, $p_{ij} \in [0, 1]$ for all $i, j \in [0, 1]$. The first-order transformation probability matrix of a Markov chain with states can be defined as follows:

$$\Pi_1 = \begin{pmatrix} 1 - p_{01} & p_{01} \\ 1 - p_{11} & p_{11} \end{pmatrix}$$

Thus, the likelihood function is:

$$L_1(\Pi_1) = (1 - p_{01})^{v_{00}} p_{01}^{v_{01}} (1 - p_{11})^{v_{10}} p_{11}^{v_{11}}, \tag{3.20}$$

where v_{ij} is the number of observations where j follows i.

$$\therefore \ \widehat{\Pi}_1 = \begin{pmatrix} \frac{v_{00}}{v_{00}+v_{01}} & \frac{v_{01}}{v_{10}+v_{11}} \\ \frac{v_{10}}{v_{10}+v_{11}} & \frac{v_{11}}{v_{10}+v_{11}} \end{pmatrix}$$

Then $p_{01} = p_{11} = p$ and the transition matrix is:

$$\Pi_2 = \begin{pmatrix} 1-p & p \\ 1-p & p \end{pmatrix}$$

And the maximum likelihood estimate is:

$$\hat{p} = \frac{v_{01} + v_{11}}{v_{00} + v_{10} + v_{01} + v_{11}} \tag{3.21}$$

The likelihood function then is:

$$L_2(\Pi) = (1-p)^{v_{00}+v_{10}} p^{v_{01}+v_{11}} \tag{3.22}$$

$$\therefore\ \widehat{\Pi}_2 = \begin{pmatrix} 1-\hat{p} & \hat{p} \\ 1-\hat{p} & \hat{p} \end{pmatrix}$$

From where we can derive the test statistic for Christoffersen (1998) CCI test:

$$LR_{CCI} = -2log\left(\frac{(1-\widehat{\Pi})^{v_o}\widehat{\Pi}^{v_1}}{(1-\widehat{\Pi}_{01})^{v_{00}}\widehat{\Pi}_{01}^{v_{01}}(1-\widehat{\Pi}_{11})^{v_{10}}\widehat{\Pi}_{11}^{v_{11}}}\right) \sim \chi_1^2, \tag{3.23}$$

The test follows a chi-square (χ_1^2) distribution with one degree of freedom and evaluates the null hypothesis, which posits that there is no clustering in violations ($\Pi_1 = \Pi_2$). Consequently, if significant differences exist between the probabilities, the test indicates the unreliability of the VaR model. (Christoffersen, 1998)

CC is a joint test derived from the likelihood ratio of Kupiec's (1995) POF test and Christoffersen's (1998) CCI test. The test statistic of Christoffersen's (1998) CC test can be defined as follows:

$$LR_{CC} = (LR_{POF} + LR_{CCI}) \sim \chi_2^2. \tag{3.24}$$

The conditional coverage likelihood ratio test is asymptotically distributed as a chi-square with two degrees of freedom. In the CC test, the null hypothesis of unconditional coverage is tested against the alternative hypothesis of the independence test. The joint test has less power to reject a VaR model which only satisfies one of the two properties. Testing clustering effects on violations is of significant economic importance, as clusters of violations can pinpoint VaR models that provide inaccurate risk predictions. Especially, when the appearance of clusters is least favorable — during economic downturns and financial crises — when there is a clustering of extreme losses in investments due to a persistent increase in volatility levels (Ziggel et al., 2014).

3.5.3 Acerbi's and Szekelys' unconditional test

Acerbi and Szekely (2014) introduced several inventive methodologies for the backtest of ES. One of these approaches involves the unconditional testing framework, which tests ES directly. The unconditional test statistic of Acerbi and Szekely (2014) is based on unconditional expectation and jointly evaluates the frequency and magnitude of α-tail events. Let $t = 1, \dots, n, X_t$ represent the outcome of the asset distributed along a real distribution F_t that is forecasted by a predictive distribution D_t and let η_t be an indicator of whether a violation occurs during the period t.

The hypotheses for the test are as follows:

$$H_0 : D_t^{[\alpha]} = F_t^{[\alpha],} \quad \text{for all } t$$

$$H_1 : ES^D_{\alpha,t} \geq ES_{\alpha,t}, \quad \text{for all } t \text{ and} > \text{for some } t$$

$$VaR^F_{\alpha,t} \geq VaR_{\alpha,t}, \quad \text{for all } t$$

Then the unconditional expectation can be defined as follows:

$$ES_{\alpha,t} = -\mathbb{E}\left[\frac{X_t \eta_t}{\alpha}\right], \tag{3.25}$$

That suggests defining:

$$Z_{UC}(X) = \sum_{t=1}^{n} \frac{X_t \eta_t}{n \alpha ES_{a,t}} + 1 \tag{3.26}$$

The test statistic for the UC test is affected by both the magnitude of VaR violations in relation to the ES estimate and the frequency of VaR violations. Consequently, a single exceptionally large VaR violation relative to the ES (or a few instances of significant losses) could result in the rejection of a model within a specific time period. The impact of a substantial loss on a day with a large ES estimate might not be as significant in the test results compared to a day with a smaller ES estimate. Furthermore, a model can be rejected during periods with numerous VaR violations, even if all violations are relatively minor and only slightly exceed the VaR (MathWorks, 2023). However, the UC test continues to demonstrate consistency in critical levels in various tail shapes (Acerbi & Szekely, 2014).

4 Results

Using the models described in Sections 3.3 and 3.4, the study performed calculations for VaR estimates at a 99% confidence level and ES estimates at a 97.5% confidence level for the S&P 500, Bitcoin, and Ethereum. Risk estimates were calculated for a single day using five different estimation windows (W_E) – 50, 100, 150, 200, and 250 – in the estimation process. In addition to using data splitting and a rolling time window method, the reliability of the models is evaluated using various backtest methods detailed in Section 3.5. The results for a 50-day estimation window are presented in Table 9, for a 100-day estimation window in Table 10, for a 150-day estimation window in Table 11, for a 200-day estimation window in Table 12, and for a 250-day estimation window in Table 13, respectively. For models in which Student's t-distribution was applied, the degrees of freedom (ν) used were as follows: SP500: 5.66, BTC 2.47, and ETH 2.99.

When interpreting the results, the ”Asset” field indicates the product under examination, the ”Method” specifies the method used for estimating VaR and ES, and the ”Observations” denotes the number of observations in the test window (W_T). In the VaR 99% column, the results of the statistical backtest are presented along with the calculated VR values. The VR value should be equal to one if the model accurately measures the risk. A higher VR than 1 represents an underestimation and a lower VR than 1 represents an overestimation of risk, respectively. The expected number of VaR exceedances at the 99% confidence level is 4.8 for S&P500 and 6.97 for both cryptocurrencies.

In the ES 97.5% column, UC refers to the unconditional test backtest method introduced earlier in the chapter, the ES column corresponds to the expected severity value, and OS represents the observed severity value. The ES column uses the average ratio of ES to VaR for the VaR violation periods. The OS column shows the average ratio of loss to VaR during periods when VaR was violated. Additionally, the last column presents NS, a value comparable to the corresponding VaR VR. In the case of ES, the expected amount of violations are 12 for S&P 500 and 17.425 for cryptocurrencies, respectively. The significance level for all implemented backtests is 0.05. More detailed results, including p-values, test statistics, and critical values, are provided in the appendices.

Cryptocurrencies show rejection of the Gaussian VCV and GARCH(1,1) models in both the VaR and ES backtests. Violation rates (VRs) are elevated, leading to the rejection of Probability of Failure (POF) tests, which assess the number of violations. Despite this, no clustering of violations has been observed in any asset class. The UC test also rejects the ES estimates, even though its NS results appear to be comparatively better than the VR results. This is because the observed severity is significantly higher than the expected severity, resulting in rejection. Student's t GARCH(1,1) remains unrejected in both the VaR and ES models, suggesting that fitting the distribution to the data could yield benefits. In particular, the estimation of ES using Student's t GARCH appears promising, considering the short estimation window. However, HS models face rejection in the case of ES. Although the BTC and S&P500 VaR estimates using HS are not rejected, they exhibit a noticeable underestimation of risk.

Table 9: Comparison of risk models with 50-day estimation window

Asset	Method	Observations	VaR 99%				ES 97.5%			
			POF	CCI	CC	VR	UC	ES	OS	NS
S&P500	Gaussian VCV	480	Fail to reject	Fail to reject	Fail to reject	1.458	Fail to reject	1.193	1.209	1.333
	Gaussian GARCH (1,1)	480	Fail to reject	Fail to reject	Fail to reject	1.458	Fail to reject	1.193	1.230	1.417
	Student's t GARCH (1,1)	480	Fail to reject	Fail to reject	Fail to reject	0.833	Fail to reject	1.193	1.319	0.667
	HS	480	Fail to reject	Fail to reject	Fail to reject	1.667	**Reject**	1.222	1.294	1.583
BTC	Gaussian VCV	697	**Reject**	Fail to reject	**Reject**	2.009	**Reject**	1.193	1.713	1.148
	Gaussian GARCH(1,1)	697	**Reject**	Fail to reject	**Reject**	2.439	**Reject**	1.193	1.599	1.492
	Student's t GARCH (1,1)	697	Fail to reject	Fail to reject	Fail to reject	1.291	Fail to reject	1.193	1.342	0.689
	HS	697	Fail to reject	Fail to reject	Fail to reject	1.722	**Reject**	1.409	1.767	1.377
ETH	Gaussian VCV	697	**Reject**	Fail to reject	**Reject**	2.152	**Reject**	1.193	1.511	1.377
	Gaussian GARCH (1,1)	697	**Reject**	Fail to reject	**Reject**	2.869	**Reject**	1.193	1.464	1.836
	Student's t GARCH (1,1)	697	Fail to reject	Fail to reject	Fail to reject	1.435	Fail to reject	1.193	1.388	0.976
	HS	697	**Reject**	Fail to reject	Fail to reject	1.865	**Reject**	1.433	1.552	1.377

Extending the estimation window from 50 to 100 days results in a higher number of models that remain unrejected by statistical tests. Specifically, none of the VaR model estimates in the ETH time series is rejected in addition to S&P500 models. Compared to Gaussian VCV, Gaussian GARCH(1,1) and HS were rejected in the previous scenario. Regarding ES, only the Gaussian GARCH model is rejected in the case of ETH, while it also shows a relatively high value of NS (1.607). When evaluating the risk associated with cryptocurrencies using HS models, both VaR and ES models demonstrate favorable results for cryptocurrencies but not for S&P500. In particular, VR and NS closely approach a value of 1 in this context. Student's t GARCH estimates are also notably accurate. Furthermore, in the case of BTC, the VaR model is rejected due to violation clustering. However, Student's t GARCH model provides conservative estimates for both VaR and ES in the context of BTC. The outcomes for the S&P 500 appear somewhat uncertain, since there is an increase in VR results, but statistical tests have not yet led to the rejection of VaR.

Table 10: Comparison of risk models with 100-day estimation window

Asset	Method	Observations	VaR 99%				ES 97.5%			
			POF	CCI	CC	VR	UC	ES	OS	NS
S&P500	Gaussian VCV	480	Fail to reject	Fail to reject	Fail to reject	1.875	**Reject**	1.193	1.201	1.667
	Gaussian GARCH (1,1)	480	Fail to reject	Fail to reject	Fail to reject	1.875	Fail to reject	1.193	1.203	1.417
	Student's t GARCH (1,1)	480	Fail to reject	Fail to reject	Fail to reject	0.625	Fail to reject	1.193	1.145	1.083
	HS	480	Fail to reject	Fail to reject	Fail to reject	1.875	Fail to reject	1.118	1.192	1.333
BTC	Gaussian VCV	697	**Reject**	Fail to reject	Fail to reject	1.865	Fail to reject	1.193	1.572	1.033
	Gaussian GARCH (1,1)	697	**Reject**	Fail to reject	**Reject**	2.296	**Reject**	1.193	1.62	1.205
	Student's t GARCH (1,1)	697	Fail to reject	**Reject**	**Reject**	0.574	Fail to reject	1.193	1.29	0.689
	HS	697	Fail to reject	Fail to reject	Fail to reject	1.148	Fail to reject	1.457	1.443	1.033
ETH	Gaussian VCV	697	Fail to reject	Fail to reject	Fail to reject	1.722	Fail to reject	1.193	1.501	1.09
	Gaussian GARCH(1,1)	697	Fail to reject	Fail to reject	Fail to reject	1.578	**Reject**	1.193	1.415	1.607
	Student's t GARCH (1,1)	697	Fail to Reject	Fail to Reject	Fail to Reject	1.004	Fail to reject	1.193	1.425	0.803
	HS	697	Fail to reject	Fail to reject	Fail to reject	1.148	Fail to reject	1.275	1.41	0.978

The results of HS continue to stay close to the value of 1 regarding cryptocurrencies as the estimation window increases to 150 days. For BTC, the estimates are slightly closer to 1 compared to using a 100-day estimation window, while for ETH, the estimates are slightly farther from the value of 1. Furthermore, none of the statistical tests succeeds in rejecting the VaR or ES model when HS is used. Student's t GARCH models are also not rejected for Ethereum; however, for BTC, the VaR model is rejected due to clustering. In the case of BTC, the ES estimates become even more conservative, whereas for ETH, they remain the same as with a shorter 100-day estimation window. VaR models experience rejection for the first time in statistical tests for S&P 500. On the other hand, estimates from Student's t GARCH and HS models have been converging towards the value of 1 more consistently as the estimation window increases. The Gaussian GARCH model has been rejected by the UC test at each estimation window so far.

Table 11: Comparison of risk models with 150-day estimation window

Asset	Method	Observations	VaR 99%				ES 97.5%			
			POF	CCI	CC	VR	UC	ES	OS	NS
S&P500	Gaussian VCV	480	**Reject**	Fail to reject	**Reject**	2.917	**Reject**	1.193	1.288	1.75
	Gaussian GARCH (1,1)	480	Fail to reject	Fail to reject	Fail to reject	1.875	**Reject**	1.193	1.199	1.5
	Student's t GARCH (1,1)	480	Fail to reject	Fail to reject	Fail to reject	0.625	Fail to reject	1.193	1.187	0.917
	HS	480	Fail to reject	Fail to reject	Fail to reject	1.875	Fail to reject	1.1297	1.255	1.083
BTC	Gaussian VCV	697	**Reject**	Fail to reject	Fail to reject	2.152	Fail to reject	1.193	1.496	1.033
	Gaussian GARCH(1,1)	697	**Reject**	Fail to reject	**Reject**	2.726	**Reject**	1.193	1.468	1.263
	Student's t GARCH (1,1)	697	Fail to reject	**Reject**	Fail to reject	0.861	Fail to reject	1.193	1.452	0.517
	HS	697	Fail to reject	Fail to reject	Fail to reject	1.004	Fail to reject	1.442	1.332	1.033
ETH	Gaussian VCV	697	Fail to reject	Fail to reject	Fail to reject	1.722	Fail to reject	1.193	1.412	1.09
	Gaussian GARCH (1,1)	697	**Reject**	Fail to reject	Fail to reject	1.865	**Reject**	1.193	1.368	1.377
	Student's t GARCH (1,1)	697	Fail to reject	Fail to reject	Fail to reject	1.435	Fail to reject	1.193	1.337	0.803
	HS	697	Fail to reject	Fail to reject	Fail to reject	1.291	Fail to reject	1.293	1.334	0.918

When using a 200-day estimation window, HS continues to demonstrate stability in staying close to the value of 1 with respect to cryptocurrencies when using VaR and ES. Statistical tests also do not reject ES models, but for BTC, the HS VaR model is rejected due to clustering of violations. Studentized GARCH models once again show promising results, although the outcomes remain relatively conservative. However, this may not necessarily be a negative aspect from a risk management perspective. The Gaussian GARCH model, on the other hand, is again rejected based on statistical tests for cryptocurrencies in all cases except for the ETH VaR model. Results for S&P 500 still appear somewhat arbitrary and do not follow a specific trend. For example, VCV estimates have increased as the estimation window grows, leading to more frequent rejections of the models. On the contrary, with regard to ETH, the results of the VCV model have improved as the estimation window increases.

Table 12: Comparison of risk models with 200-day estimation window

Asset	Method	Observations	VaR 99%				ES 97.5%			
			POF	CCI	CC	VR	UC	ES	OS	NS
S&P500	Gaussian VCV	480	**Reject**	Fail to reject	**Reject**	3.542	**Reject**	1.193	1.396	1.667
	Gaussian GARCH (1,1)	480	Fail to reject	Fail to reject	Fail to reject	1.667	**Reject**	1.193	1.213	1.5
	Student's t GARCH (1,1)	480	Fail to reject	Fail to reject	Fail to reject	0.833	Fail to reject	1.193	1.18	1.083
	HS	480	Fail to reject	Fail to reject	Fail to reject	1.875	**Reject**	1.2	1.27	1.583
BTC	Gaussian VCV	697	Fail to reject	Fail to reject	Fail to reject	1.578	Fail to reject	1.193	1.434	1.09
	Gaussian GARCH(1,1)	697	**Reject**	Fail to reject	**Reject**	2.001	**Reject**	1.193	1.505	1.205
	Student's t GARCH (1,1)	697	Fail to reject	Reject	Fail to reject	0.717	Fail to reject	1.193	1.468	0.574
	HS	697	Fail to reject	**Reject**	Fail to reject	0.861	Fail to reject	1.429	1.39	0.976
ETH	Gaussian VCV	697	Fail to reject	Fail to reject	Fail to reject	1.578	Fail to reject	1.193	1.407	1.09
	Gaussian GARCH(1,1)	697	Fail to reject	Fail to reject	Fail to reject	1.578	**Reject**	1.193	1.421	1.32
	Student's t GARCH (1,1)	697	Fail to reject	Fail to reject	Fail to reject	1.004	Fail to reject	1.193	1.378	0.803
	HS	697	Fail to reject	Fail to reject	Fail to reject	1.291	Fail to reject	1.429	1.368	1.033

When increasing the size of the estimation window from 200 to 250, the ES estimates for cryptocurrencies are not rejected by the UC test. Additionally, their NS values are quite close to 1. An exception is the Student's t GARCH model, which, however, does not directly indicate the model's poor performance, but rather suggests that it is conservative. In the case of ETH, the VaR models are also not rejected. On the other hand, for BTC, surprisingly, all models except Gaussian VCV are rejected in at least one test. Regarding S&P500, the results still appear more random, making it more difficult to draw direct conclusions. This may be influenced by the estimation window that covers a proportionally larger part of the year's observations for S&P500 compared to cryptocurrencies. However, the performance of the models does not necessarily improve or worsen directly as the size of the estimate window changes.

Table 13: Comparison of risk models with 250-day estimation window

Asset	Method	Observations	VaR 99%				ES 97.5%			
			POF	CCI	CC	VR	UC	ES	OS	NS
S&P500	Gaussian VCV	480	**Reject**	Fail to reject	**Reject**	3.125	**Reject**	1.193	1.405	1.833
	Gaussian GARCH (1,1)	480	Fail to reject	Fail to reject	Fail to reject	1.667	Fail to reject	1.193	1.225	1.417
	Student's t GARCH (1,1)	480	Fail to reject	Fail to reject	Fail to reject	0.833	Fail to reject	1.193	1.231	0.917
	HS	480	**Reject**	Fail to reject	Fail to reject	2.083	Fail to reject	1.202	1.335	1.333
BTC	Gaussian VCV	697	Fail to reject	Fail to reject	Fail to reject	1.578	Fail to reject	1.193	1.436	1.033
	Gaussian GARCH(1,1)	697	**Reject**	Fail to reject	**Reject**	2.001	Fail to reject	1.193	1.462	1.09
	Student's t GARCH (1,1)	697	Fail to reject	**Reject**	**Reject**	0.574	Fail to reject	1.193	1.468	0.459
	HS	697	Fail to reject	**Reject**	Fail to reject	0.717	Fail to reject	1.414	1.331	1.033
ETH	Gaussian VCV	697	Fail to reject	Fail to reject	Fail to reject	1.435	Fail to reject	1.193	1.413	1.033
	Gaussian GARCH (1,1)	697	Fail to reject	Fail to reject	Fail to reject	1.578	Fail to reject	1.193	1.419	1.09
	Student's t GARCH (1,1)	697	Fail to reject	Fail to reject	Fail to reject	0.861	Fail to reject	1.193	1.387	0.631
	HS	697	Fail to reject	Fail to reject	Fail to reject	1.004	Fail to reject	1.414	1.392	0.861

At this point, it can be stated that the results of the HS and Student's t GARCH(1,1) models were the most promising concerning cryptocurrencies. The HS model managed to produce relatively stable risk estimates with different-sized estimation windows, while the GARCH model using the Student's t distribution provided a conservative approach to risk estimation. Additionally, when considering VaR and ES models as a whole, through which the risk estimates for cryptocurrencies were calculated, HS and Student's t GARCH were relatively less frequently rejected with backtesting methods. Therefore, we could examine the performance of these two models through a few figures.

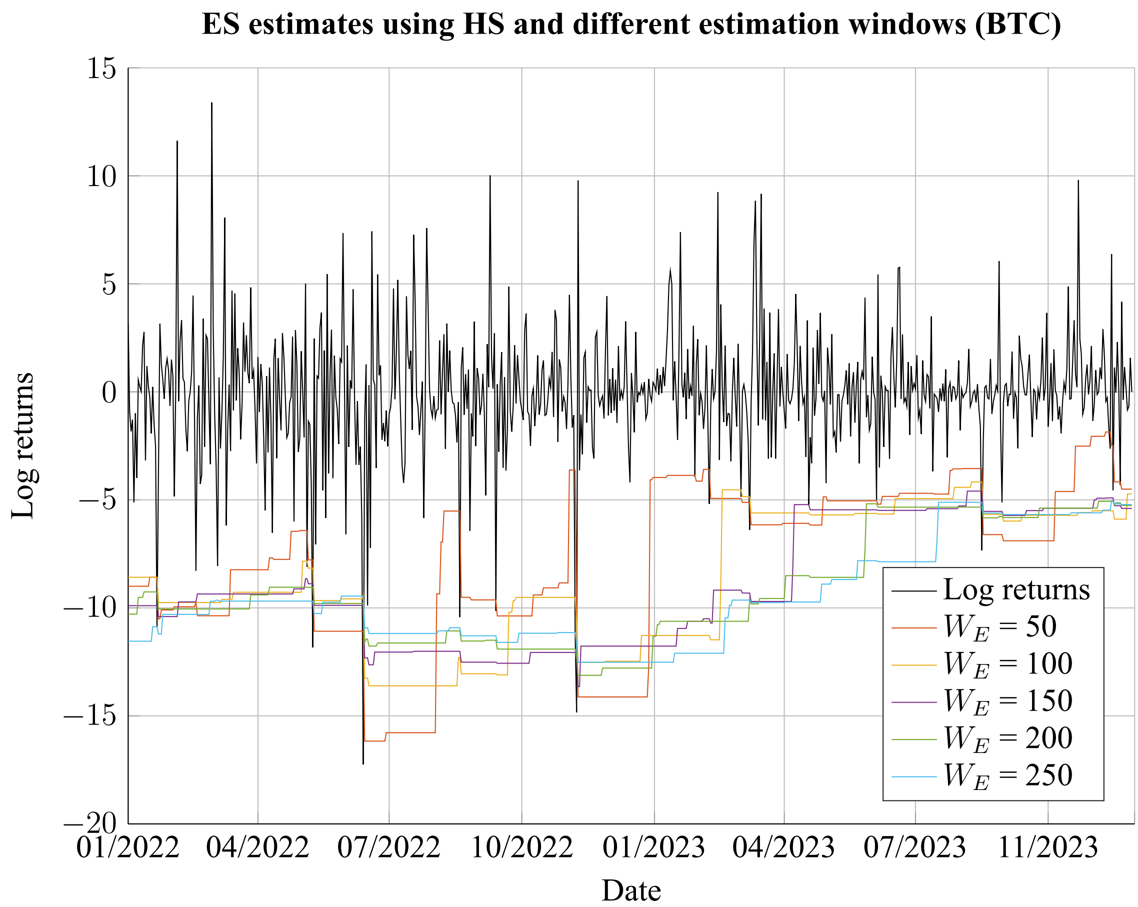

Figure 7: ES estimates using HS and different estimation windows (BTC)

The Figure 7 shows the performance of ES on BTC, using HS and various sizes of estimation windows. From the figure, it is clearly noticeable that when using a 50-day estimation window, ES reacts rapidly to changes in volatility. While this can be beneficial in some cases, generally, even a 50-day estimation window fails to react quickly enough to significant volatility spikes, such as the one before 07/2022 or around 11/2022. However, when using this estimation window, the model tracks the actual logarithmic returns more closely. In contrast, a 250-day estimate window naturally responds more slowly to changes. This can be advantageous if there are no significant changes in volatility and the risk level remains relatively constant. However, like models predicted with other estimation windows, this model also struggles to react to major changes in volatility. It is essential to note that the volatility around 09/2022 could be modeled using 100-250-day estimation windows, while the 50-day estimation window reacted sharply to the earlier decline in volatility, failing to capture the mentioned spike effectively. ES model with a 150-day estimation window was violated fewer times than other models (see Appendix H).

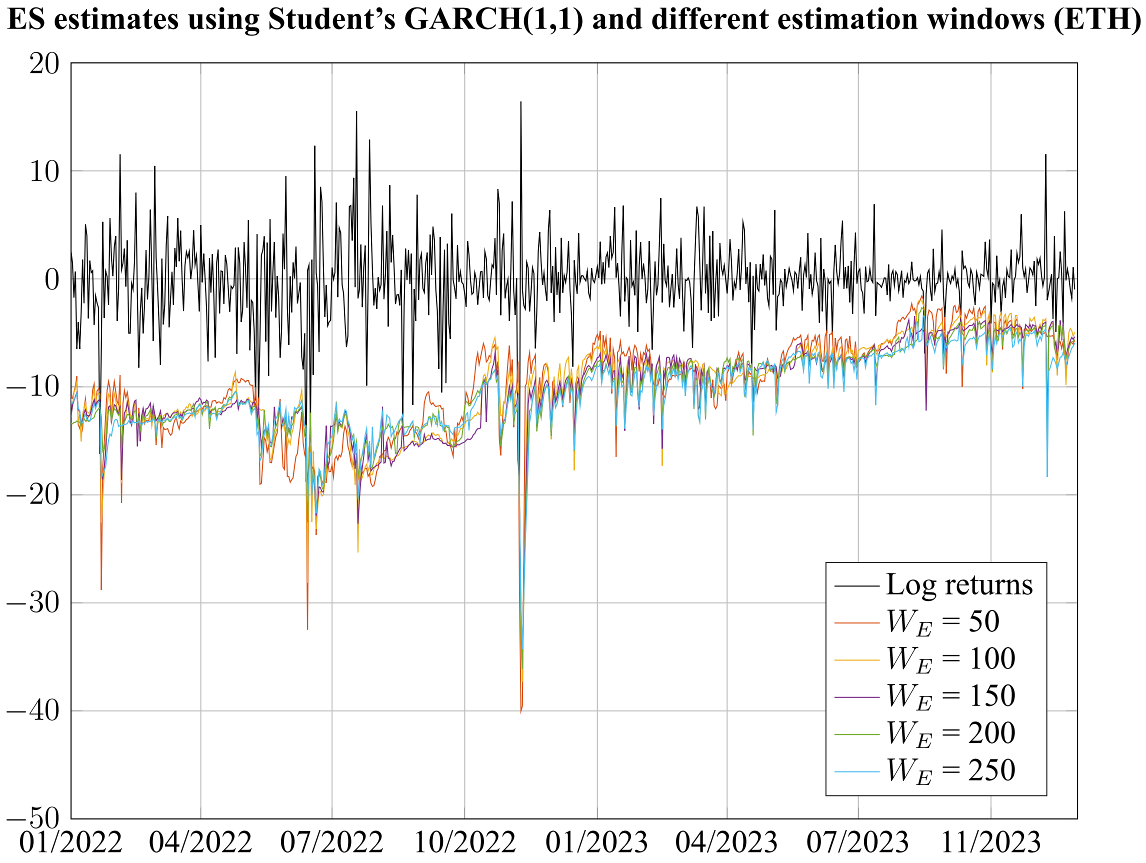

Figure 8: ES estimates using Student's t GARCH(1,1) and different estimation windows (ETH)

From the Figure 8, calculated risk estimates for ETH using ES with different-sized estimation windows can be observed. It is evident from the figure that the impact of the estimation window size is not as pronounced as when using HS. Additionally, modeling volatility using the GARCH(1,1) model with the assumption of a Student's t-distribution produces estimates that more smoothly follow logarithmic returns compared to HS. Although, in general, longer time frames may be recommended for volatility modeling, this figure provides some evidence that the estimation window size may have a smaller impact when using the GARCH model (at least under the assumption of Student's t-distribution) compared to the HS model. Although models with different-sized estimation windows performed quite well overall, violations were minimized when using a 250-day estimation window (see Appendix G). However, all models passed the UC test.

.

5 Summary and conclusion

Cryptocurrencies serve as an investment asset characterized by significant volatility. Given the rapid growth of the cryptocurrency market and its stressed market conditions, evaluating investment risks becomes increasingly important. In this thesis, parametric and non-parametric models were used to calculate VaR and ES for cryptocurrencies and the S&P500, which served as a benchmark. The process included employing five distinct estimation windows for each VaR and ES models, with the HS and Student's t GARCH estimation techniques generally demonstrating better accuracy compared to other models for cryptocurrencies.

To answer the first research question *"What is the impact of the size of the estimation window on the applicability of VaR and ES risk estimates in times of heightened market stress?"* it can be inferred that although VaR and ES are not directly comparable risk measures, the mathematical properties of ES are more favorable for measuring the market risk of the cryptocurrencies used in the study. The result aligns with traditional asset classes. Although some studies suggest that VaR is not a favorable risk measure under stressful market conditions, this study cannot refute the possibility that the measure may have its place when used with caution and with awareness of its limitations. However, VaR is an easy-to-implement risk measure, and the results it produces are easily expressible and understandable.

Additionally, when comparing the accuracy of measures based on the results provided by backtesting methods, it is important to note that ES backtesting methods consider not only the quantity of errors but also their magnitude. As a result, a few severe violations may lead to the rejection of the model, whereas several less significant violations may pass post-testing. The information obtained about the average size of losses is a crucial factor, making ES a more informative risk measure than VaR and, consequently, better suited for cryptocurrency markets, where rapid and substantial changes in volatility occur. In this study, ES demonstrated the ability to produce more accurate risk estimates for Bitcoin and Ethereum, especially during periods of elevated market risk, particularly when using the Student's t-distribution GARCH(1,1) method and the HS method. However, it is crucial to recognize that ES is sensitive to parameter changes; for example, in the case of BTC, the Student's t-distribution GARCH(1,1) model became more conservative as the estimation window increased. This may not necessarily be a drawback and can, in some cases, be a desirable feature depending on the risk manager's perspective. However, it is essential to be aware of this when using the model and its assumptions.

Considering the ease of implementation and promising results, it can be stated that HS in the estimation method for VaR and ES may, in many cases, be a more reliable solution, as its use can help avoid model errors related to assumptions about distributions. On the other hand, VaR provides a simpler approach to risk estimation, but it should be noted that statistical backtesting methods may not necessarily reject VaR models, even if the model's VaR ratio is nearly twice as large as expected. This does not necessarily preclude the use of models, but it is crucial to be aware of this consideration. In conclusion, this research identifies ES as a comprehensive risk measure to assess the market risk of cryptocurrencies. Moreover, it sees no obstacle to the VaR model serving as a helpful tool for risk management in cryptocurrency markets when used with the appropriate parameters.

To answer the second research question *"How estimation window size impact on applicability of VaR and ES risk estimates during periods of increased market stress?"* it can be stated that in this study, the alteration of the estimation window did not produce significant effects on the risk estimates. However, it should be noted that the most promising models, HS and Student's t GARCH(1,1), provided the most accurate risk estimates (i.e., the fewest amount of violations in the backtesting) when a 250-day estimation window was used. This estimation window aligns with the BCBS recommendation. On the other hand, the study does not rule out the possibility that, in certain situations, shorter estimation windows could be used or might yield better results. For example, VaR estimates using the Gaussian VCV method for S&P500 deteriorated as the estimation window increased. This makes the assessment of the impact of the size of the estimation window multifaceted. It is also crucial to note that, in addition to the estimation window size, the time frame to which it is applied has a significant effect.

According to Daníelsson (2011), HS requires a minimum 300-day estimation window, and GARCH models need at least a 500-day estimation window in the context of traditional assets. Although this study did not use these specific estimation windows, it generally does not provide evidence supporting the superior performance of larger estimation windows for HS VaR and ES models. Additionally, none of the Student's GARCH(1,1) models faced rejection in statistical tests for VaR and ES in ETH. Although the results with a 250-day estimation window generally appeared more stable, there is the possibility that even shorter windows could prove effective for both cryptocurrency and traditional asset risk models. In a manner analogous to how VaR and ES can be applicable in various scenarios, the same holds true for different estimation window sizes, as suggested by the findings of this study.

5.1 Limitations and future directions

This research contributed by adding information to the limited studies on the management of the market risk of cryptocurrencies, especially during periods of increased market stress. It should be noted that this thesis only provides a picture of the "life" of the selected cryptocurrencies, and as a result, the estimates of VaR and ES relate only to the period considered in this study. It is important to note that the time frame used in the study captures a unique moment in the cryptocurrency markets, as their volatility has been at its highest during these years. However, there is currently a lack of additional data on the high-volatility period, and exploring this subject further would be valuable with additional data. On the contrary, the duration of the volatility of the market remains uncertain. The future may be very different, with volatility potentially decreasing or increasing.

Furthermore, the research emphasized shorter estimation windows in contrast to previous studies, resulting in a limited comparison between relatively lengthy and shorter estimation windows. It is crucial to note that for cryptocurrencies, the longest estimation window falls short of covering nearly the entire year's observations, unlike the S&P500 case where it does, potentially impacting the results. The confidence levels for the risk measures used in this thesis align with the BCBS standards, but the use of different confidence levels could produce results divergent from those of this study. Taking into account the above, caution should be exercised when making generalizations. As this study did not explore multiple estimation methods, it could be interesting in the future to compare different volatility models, especially by using a larger set of different estimation windows and distribution assumptions.

References

Acerbi, C., & Szekely, B. (2014). Backtesting Expected Shortfall. *Risk*, *27*(11), 76–81. https://dlu-umich.github.io/docs/Research_Insight_Backtesting_Expected_Shortfall_December_2014.pdf

Acereda, B., Leon, A., & Mora, J. (2020). Estimating the Expected Shortfall of Cryptocurrencies: An Evaluation Based on Backtesting. *Finance Research Letters*, *33*. https://doi.org/10.1016/j.frl.2019.04.037

Açikgöz, Ö., & Günay, A. (2020). The Early Impact of the Covid-19 Pandemic on the Global and Turkish economy. *Turkish journal of Medical Sciences*, *50*(9), 520–526. https://doi.org/10.3906/sag-2004-6

Al Mamun, M., Uddin, G. S., Suleman, M. T., & Kang, S. H. (2020). Geopolitical risk, Uncertainty and Bitcoin Investment. *Physica A: Statistical Mechanics and its Applications*, *540*, 123107. https://doi.org/10.1016/j.physa.2019.123107

Alexander, C., et al. (2009). *Market Risk Analysis Volume IV: Value-at-Risk Models*. John Wiley & Sons Chichester.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent Measures of Risk. *Mathematical finance*, *9*(3), 203–228. https://doi.org/10.1111/1467-9965.00068

Asgharian, H., Christiansen, C., Hou, A. J., & Wang, W. (2021). Long- and short-run components of factor betas: Implications for stock pricing. *Journal of International Financial Markets, Institutions & Money*, *74*. https://doi.org/10.1016/j.intfin.2021.101412

Auwera, E., Schoutens, W., Petracco Giudici, M., & Alessi, L. (2020). *Financial Risk Management for Cryptocurrencies* (1st ed.). Springer.

Azarenkova, G., Shkodina, I., Samorodov, B., Babenko, M., & Onishchenko, I. (2018). The influence of financial technologies on the global financial system stability. *Investment Management and Financial Innovations*, *15*, 229–238. https://doi.org/10.21511/imfi.15(4).2018.19

Balaji, K., Karim, S., Naidu, N. G., Venkatesh, T., Ranjitha, P., & Chandrasekhar, S. (2023). Examining the Potential of Cryptocurrencies as An Asset Class-An Empirical Study, 1–6. https://doi.org/10.1109/ICAISC58445.2023.10200309

Barjašić, I., & Antulov-Fantulin, N. (2021). Time-varying Volatility in Bitcoin Market and Information Flow at Minute-level Frequency. *9*. https://doi.org/10.3389/fphy.2021.644102

Basel Committee on Banking Supervision. (2012). *Fundamental review of the trading book* [Consultative Document]. https://www.bis.org/publ/bcbs219.pdf

Basel Committee on Banking Supervision. (2013). *Fundamental review of the trading book: Second consultative document*. https://www.bis.org/press/p131031.htm

Basel Committee on Banking Supervision (BCBS). (1988). International Convergence of Capital Measurement and Capital Standards. https://www.bis.org/publ/bcbs04a.htm

Basel Committee on Banking Supervision (BCBS). (1996). Amendment to the Capital Accord to Incorporate Market Risks. https://www.bis.org/publ/bcbs24.htm

Basel Committee on Banking Supervision (BCBS). (2006). Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version. https://www.bis.org/publ/bcbs128.htm

Basel Committee on Banking Supervision (BCBS). (2009). Enhancements to the Basel II Framework. https://www.bis.org/publ/bcbs157.pdf

Basel Committee on Banking Supervision (BCBS). (2011). Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems. https://www.bis.org/publ/bcbs189.htm

Basel Committee on Banking Supervision (BCBS). (2019a). Designing a Prudential Treatment for Crypto-assets. https://www.bis.org/bcbs/publ/d490.htm

Basel Committee on Banking Supervision (BCBS). (2019b). Minimum Capital Requirements for Market Risk. https://www.bis.org/bcbs/publ/d457.htm

Basel Committee on Banking Supervision (BCBS). (2022). Prudential treatment of cryptoasset exposures. https://www.bis.org/bcbs/publ/d545.pdf

Baur, D. G., & Dimpfl, T. (2018). Asymmetric Volatility in Cryptocurrencies. *Economics Letters*, *173*, 148–151. https://doi.org/10.1016/j.econlet.2018.10.008

Berens, T., Weiß, G., & Ziggel, D. (2018). Estimation window strategies for value-at-risk and expected shortfall forecasting. *Journal of Risk*, *20*, 33–82. https://doi.org/10.21314/JOR.2018.385

Berentsen, A., & Schär, F. (2018). A Short Introduction to the World of Cryptocurrencies. *Federal Reserve Bank of St. Louis Review*, *First Quarter*, 1–16. https://doi.org/10.20955/r.2018.1-16

Bianchi, D. (2020). Cryptocurrencies As an Asset Class? An Empirical Assessment. *The journal of Alternative Investments*, *23*(2), 162–179. https://doi.org/10.3905/jai.2020.1.105

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Boudt, K., Danielsson, J., & Laurent, S. (2013). Robust forecasting of dynamic conditional correlation GARCH models. *International Journal of Forecasting*, *29*(2), 244–257. https://doi.org/10.1016/j.ijforecast.2012.06.003

Brown, M. A. (2019). Cryptocurrency and Financial Regulation: The SEC's Rejection of Bitcoin-Based ETPs. *NC Banking Inst.*, *23*, 139. https://scholarship.law.unc.edu/ncbi/vol23/iss1/10/

Bruzgė, R., Černevičienė, J., Šapkauskienė, A., Mačerinskienė, A., Masteika, S., & Driaunys, K. (2023). Stylized Facts, Volatility Dynamics and Risk Measures of Cryptocurrencies. *Journal of Business Economics and Management*, *24*(3), 527–550. https://doi.org/10.3846/jbem.2023.19118

Buczyński, M., & Chlebus, M. (2019). Old-fashioned Parametric Models are Still the Best. A Comparison of Value-at-Risk Approaches in Several Volatility States. *IDEAS Working Paper Series from RePEc*. https://ideas.repec.org/p/war/wpaper/2019-12.html

Buczyński, M., & Chlebus, M. (2020). Size Does Matter. A Study on the Required Window Size for Pptimal Quality Market Risk Models. https://ideas.repec.org/p/war/wpaper/2020-09.html

Buczyński, M., & Chlebus, M. (2022). The importance of window size: A study on the required window size for optimal-quality market risk models. *Journal of Risk Model Validation*, *16*(1). https://ssrn.com/abstract=4095796

Buterin, V. (2013). Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform [Accessed on January 8, 2024]. https://blockchainlab.com/pdf/Ethereum_white_paper-a_next_generation_smart_contract_and_decentralized_application_platform-vitalik-buterin.pdf

Castrén, O., Kavonius, I. K., & Rancan, M. (2022). Digital Currencies in Financial Networks. *Journal of Financial Stability*, *60*. https://doi.org/10.1016/j.jfs.2022.101000

Christoffersen, P. F. (1998). Evaluating Interval Forecasts. *International Economic Review*, *39*(4), 841–862. https://doi.org/10.2307/2527341

CoinMarketCap. (2023a). *Binance* [Accessed on 30.09.2023]. https://coinmarketcap.com/currencies/bnb/#About

CoinMarketCap. (2023b). *Bitcoin* [Accessed on 30.09.2023]. https://coinmarketcap.com/currencies/bitcoin/#About

CoinMarketCap. (2023c). *Coinmarketcap - cryptocurrency market capitalizations* [Accessed on 11.10.2023]. https://coinmarketcap.com/charts/

CoinMarketCap. (2023d). *Ethereum* [Accessed on 30.09.2023]. https://coinmarketcap.com/currencies/ethereum/#About

CoinMarketCap. (2023e). *Global Live Cryptocurrency Charts & Market Data* [Accessed on 13.10.2023]. https://coinmarketcap.com/charts/

CoinMarketCap. (2023f). *Ripple* [Accessed on 30.09.2023]. https://coinmarketcap.com/currencies/xrp/#About

CoinMarketCap. (2023g). *Tether* [Accessed on 30.09.2023]. https://coinmarketcap.com/currencies/tether/#About

CoinMarketCap. (2023h). *Token* [Accessed on 30.09.2023]. https://coinmarketcap.com/academy/glossary/token

Conlon, T., & McGee, R. (2020). Safe Haven or Risky Hazard? Bitcoin During the Covid-19 Bear Market. *Finance Research Letters*, *35*, 101607–101607. https://doi.org/10.1016/j.frl.2020.101607

Corbet, S., Larkin, C., & Lucey, B. (2020). The Contagion Effects of the COVID-19 Pandemic: Evidence From Gold and Cryptocurrencies. *Finance Research Letters*, *35*, 101554–101554. https://doi.org/10.1016/j.frl.2020.101554

Cornelli, G., Doerr, S., Frost, J., & Gambacorta, L. (2023). Crypto shocks and retail losses.

Danielsson, J. (2002). The emperor has no clothes: Limits to risk modelling. *26*(7), 1273–1296. https://doi.org/10.1016/S0378-4266(02)00263-7

Daníelsson, J. (2011). *Financial Risk Forecasting: The Theory and Practice of Forecasting Market Risk with Implementation in R and Matlab*. Wiley.

Doumenis, Y., Izadi, J., Dhamdhere, P., Katsikas, E., & Koufopoulos, D. (2021). A Critical Analysis of Volatility Surprise in Bitcoin Cryptocurrency and Other Financial Assets. *Risks*, *9*(11), 207. https://doi.org/10.3390/risks9110207

Dowd, K., & Blake, D. (2006). After VaR: The Theory, Estimation, and Insurance Applications of Quantile-Based Risk Measures. *The Journal of Risk and Insurance*, *73*(2), 193–229. https://doi.org/10.1111/j.1539-6975.2006.00171.x

Duc, F., & Schorderet, Y. (2008). *Market Risk Management for Hedge Funds: Foundations of the Style and Implicit Value-At-Risk*. John Wiley & Sons, Incorporated.

Dyhrberg, A. H. (2016). Bitcoin, Gold and the Dollar – A GARCH Volatility Analysis. *Finance Research Letters*, *16*, 85–92. https://doi.org/10.1016/j.frl.2015.10.008

Engle, R. (2004). Risk and Volatility: Econometric Models and Financial Practice. *The American Economic Review*, *94*(3), 405–420. https://www.jstor.org/stable/3592935

European Banking Authority (EBA). (2019). Report with advice for the European Commission on crypto-assets. https://eba.europa.eu/sites/default/documents/files/documents/10180/2545547/67493daa-85a8-4429-aa91-e9a5ed880684/EBA%20Report%20on%20crypto%20assets.pdf?retry=1

European Securities and Markets Authority (ESMA). (2019). Advice on Initial Coin Offerings and Crypto-Assets. https://www.esma.europa.eu/sites/default/files/library/esma50-157-1391_crypto_advice.pdf

European Securities and Markets Authority (ESMA). (2023). Technical Standards specifying certain requirements of Markets in Crypto Assets Regulation (MiCA) - second consultation paper. https://www.esma.europa.eu/sites/default/files/2023-10/ESMA75-453128700-438_MiCA_Consultation_Paper_2nd_package.pdf

Fantazzini, D., & Zimin, S. (2020). A multivariate approach for the simultaneous modelling of market risk and credit risk for cryptocurrencies. *Economia e politica industriale*, *47*(1), 19–69. https://doi.org/10.1007/s40812-019-00136-8

Federal Financial Supervisory Authority (BaFin). (2023). Services relating to DLT, blockchain and crypto assets [Accessed on 24.11.2023]. https://www.bafin.de/EN/Aufsicht/FinTech/Geschaeftsmodelle/DLT_Blockchain_Krypto/DLT_Blockchain_Krypto_node_en.html

Fung, K., Jeong, J., & Pereira, J. (2022). More to cryptos than bitcoin: A GARCH modelling of heterogeneous cryptocurrencies. *Finance Research Letters*, *47*. https://doi.org/10.1016/j.frl.2021.102544

Gaunt, R. E. (2021). A simple proof of the characteristic function of Student's t-distribution. *Communications in statistics. Theory and methods*, *50*(14), 3380–3383. https://doi.org/10.1080/03610926.2019.1702695

Glas, T. (2022). *Asset Pricing and Investment Styles in Digital Assets : A Comparison with Traditional Asset Classes*. Springer.

Glaser, F., Zimmermann, K., Haferkorn, M., Weber, M., & Siering, M. (2014). Bitcoin - asset or currency? revealing users' hidden intentions. *ECIS 2014 Proceedings - 22nd European Conference on Information Systems*.

Gneiting, T. (2011). Making and Evaluating Point Forecasts. *Journal of the American Statistical Association*, *106*(494), 746–762. https://www.jstor.org/stable/41416407

Griffith, T., & Clancey-Shang, D. (2023). Cryptocurrency regulation and market quality. *Journal of international financial markets, institutions & money*, *84*. https://doi.org/10.1016/j.intfin.2023.101744

Guldimann, T. M. (1995). *Risk measurement framework, riskmetrics—technical document, 3rd ed.* Morgan Guaranty.

Guldimann, T. M. (2000). The story of riskmetrics. *Risk*, *13*(1), 56–58.

Haykir, O., & Yagli, I. (2022). Speculative bubbles and herding in cryptocurrencies. *Financial innovation*, *8*(1), 78. https://doi.org/10.1186/s40854-022-00383-0

Holton, G. A. (2002). History of value-at-risk.

Holton, G. A. (2004). Defining risk. *Financial analysts journal*, *60*(6), 19–25.

Hotta, L. K., & Trucíos, C. (2018). Inference in (m) garch models in the presence of additive outliers: Specification, estimation, and prediction. *Advances in Mathematics and Applications: Celebrating 50 years of the Institute of Mathematics, Statistics and Scientific Computing, University of Campinas*, 179–202. https://doi.org/10.1007/978-3-319-94015-1_8

Hubbert, S. (2012). *Essential Mathematics for Market Risk Management* (2nd ed.). Wiley.

Hull, J. (2018a). *Options, Futures and Other Derivatives* (9th ed.). Pearson Education.

Hull, J. C. (2018b). *Risk Management and Financial Institutions* (5th ed.). Wiley.

Hult, H. (2012). *Risk and Portfolio Analysis Principles and Methods*. Springer New York.

IFRS Foundation's Accounting Standards Advisory Forum (ASAF). (2007). *IAS 32 — financial instruments: Presentation* [Accessed on 23.09.2023]. https://www.iasplus.com/en/standards/ias/ias32

International Financial Reporting Standards. (2019a). Holdings of Cryptocurrencies. https://www.ifrs.org/content/dam/ifrs/supporting-implementation/agenda-decisions/2019/holdings-of-cryptocurrencies-june-2019.pdf

International Monetary Fund. (2019). Treatment of Crypto Assets in Macroeconomic Statistics. https://www.imf.org/external/pubs/ft/bop/2019/pdf/Clarification0422.pdf

Jiménez, I., Mora-Valencia, A., & Perote, J. (2020a). Risk quantification and validation for bitcoin. *48*(4), 534–541.

Jiménez, I., Mora-Valencia, A., & Perote, J. (2020b). Risk quantification and validation for bitcoin. *Operations Research Letters*, *48*. https://doi.org/10.1016/j.orl.2020.06.004

Jorion, P. (2007). *Value at risk : The new benchmark for managing financial risk* (3rd ed.). McGraw-Hill.

Kim, S., Sarin, A., & Virdi, D. (2018). Crypto-Assets Unencrypted. *Journal of Investment Management*. https://ssrn.com/abstract=3117859

Klöhn, L., & Parhofer, N. (2018). Bitcoins Under European and German Banking and Capital Markets Laws. https://doi.org/10.2139/ssrn.3287189

Kondor, I. (2014). Estimation error of expected shortfall. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2400401

Kourouma, L., Dupre, D., Sanfilippo, G., & Taramasco, O. (2011). Extreme Value at Risk and Expected Shortfall during Financial Crisis. *IDEAS Working Paper Series from RePEc*. https://doi.org/10.2139/ssrn.1744091

Koutmos, D. (2018). Return and volatility spillovers among cryptocurrencies. *Economics letters*, *173*, 122–127. https://doi.org/10.1016/j.econlet.2018.10.004

Kupiec, P. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. (95-24). https://EconPapers.repec.org/RePEc:fip:fedgfe:95-24

Lahmiri, S., & Bekiros, S. (2020). The impact of covid-19 pandemic upon stability and sequential irregularity of equity and cryptocurrency markets. *Chaos, Solitons & Fractals*, *138*, 109936. https://doi.org/10.1016/j.chaos.2020.109936

Likitratcharoen, D., Ranong, T. N., Chuengsuksomboon, R., Sritanee, N., & Pansriwong, A. (2018). Value at Risk Performance in Cryptocurrencies. *The Journal of Risk Management and Insurance*, *22*(1), 11–28. https://jrmi.au.edu/index.php/jrmi/article/view/171

Likitratcharoen, D., Chudasring, P., Pinmanee, C., & Wiwattanalamphong, K. (2023). The Efficiency of Value-at-Risk Models during Extreme Market Stress in Cryptocurrencies. *Sustainability*, *15*(5), 4395. https://doi.org/10.3390/su15054395

Likitratchoaroen, D., Kronprasert, N., Wiwattanalamphong, K., & Pinmanee, C. (2021). The Accuracy of Risk Measurement Models on Bitcoin Market during COVID-19 Pandemic. *Risks (Basel)*, *9*(12), 1–16. https://doi.org/10.3390/risks9120222

Linsmeier, T. J., & Pearson, N. D. (2000). Value at risk. *Financial analysts journal*, *56*(2), 47–67. https://www.jstor.org/stable/4480232

Liu, F., & Wang, R. (2021). A theory for measures of tail risk. *Mathematics of Operations Research*, *46*(3), 1109–1128. https://doi.org/10.1287/moor.2020.1072

Liu, J., Makarov, I., & Schoar, A. (2023). Anatomy of a Run: The Terra Luna Crash. *NBER Working Paper Series*. https://doi.org/10.3386/w31160

Liu, W., Semeyutin, A., Lau, C. K. M., & Gozgor, G. (2020). Forecasting Value-at-Risk of Cryptocurrencies with RiskMetrics type models. *Research in International Business and Finance*, *54*. https://doi.org/10.1016/j.ribaf.2020.101259

Liu, Y., & Tsyvinski, A. (2021). Risks and Returns of Cryptocurrency. *The Review of financial studies*, *34*(6), 2689–2727. https://doi.org/10.1093/rfs/hhaa113

Lönnbark, C. (2013). On the role of the estimation error in prediction of expected shortfall. *Journal of Banking & Finance*, *37*(3), 847–853. https://doi.org/10.1016/j.jbankfin.2012.10.013

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

Masciandaro, D. (2018). Central Bank Digital Cash and Cryptocurrencies: Insights from a New Baumol–Friedman Demand for Money. *Australian Economic Review*, *51*(4), 540–550. https://doi.org/10.1111/1467-8462.12304

MathWorks. (2023). *Risk management toolbox™ user's guide*. The MathWorks, Inc. 1 Apple Hill Drive, Natick, MA 01760-2098.

Mavani, M. P. (2020). *The Value at Risk Models in Times of Financial Crisis: Case Study of an Irish Equity Portfolio* (Master's thesis). National College of Ireland. Dublin, Ireland.

McNeil, A. J. (1999). Extreme value theory for risk managers. *Departement Mathematik ETH Zentrum*, *12*(5), 217–37.

McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative Risk Management* (1st ed.). Princeton University Press.

Miller, M. B. (B. (2019). *Quantitative Financial Risk Management*. Wiley.

Nadarajah, S., Chan, S., & Afuecheta, E. (2017). Tabulations for value at risk and expected shortfall. *Communications in statistics. Theory and methods*, *46*(12), 5956–5984. https://doi.org/10.1080/03610926.2015.1116572

Nakamoto, S. (2009). Bitcoin: A Peer-to-Peer Electronic Cash System. http://www.bitcoin.org/bitcoin.pdf

Nakavachara, V., Potipiti, T., & Lertmongkolnam, T. (2019). Should All Blockchain-Based Digital Assets be Classified Under the Same Asset Class? https://doi.org/10.2139/ssrn.3437279

Obeng, C. (2021). Measuring value at risk using garch model - evidence from the cryptocurrency market. *International Journal of Entrepreneurial Knowledge*, *9*(2), 63–84. https://doi.org/10.37335/ijek.v9i2.133

Odel, N., Ahmed, R., & Hedariani, L. (2019). Estimation of value at risk for gold futures contracts. *Journal Of Organizational Behavior Research*, *4*.

Osterrieder, J., Lorenz, J., & Strika, M. (2016). Bitcoin and cryptocurrencies - not for the faint-hearted. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2867671

Patton, A. J., Ziegel, J. F., & Chen, R. (2019). Dynamic semiparametric models for expected shortfall (and Value-at-Risk). *Journal of econometrics*, *211*(2), 388–413. https://doi.org/10.1016/j.jeconom.2018.10.008

Pele, D. T., & Mazurencu-Marinescu-Pele, M. (2019). Using High-Frequency Entropy to Forecast Bitcoin's Daily Value at Risk. *Entropy*, *21*(2). https://doi.org/10.3390/e21020102

Pele, D. T., Wesselhöfft, N., Härdle, W. K., Kolossiatis, M., & Yatracos, Y. G. (2020). A Statistical Classification of Cryptocurrencies. https://doi.org/10.2139/ssrn.3548462

Penza, P., & Bansal, V. K. (2001). *Measuring Market Risk with Value at Risk* (Vol. 17). John Wiley & Sons.

Righi, M., & Ceretta, P. (2015). On the existence of an optimal estimation window for risk measures. *Applied Finance Letters*, *4*, 28. https://doi.org/10.24135/afl.v4i1and2.30

Roccioletti, S. (2016). *Backtesting value at risk and expected shortfall*. https://doi.org/10.1007/978-3-658-11908-9

Rockafellar, R., & Uryasev, S. (2002). Conditional Value-at-Risk for General Loss Distributions. *Journal of banking & finance*, *26*(7), 1443–1471. https://doi.org/10.2139/ssrn.267256

Röman, J. R. M. (2017). *Analytical Finance: Volume II The Mathematics of Interest Rate Derivatives, Markets, Risk and Valuation*. Springer International Publishing.

Sarykalin, S., Serraino, G., & Uryasev, S. (2008). Value- at-Risk vs Conditional Value-at-Risk in Risk Management and Optimization. *Tutorials in Operations Research*. https://doi.org/10.1287/educ.1080.0052

Scharfman, J. A. (2022). *Cryptocurrency Compliance and Operations: Digital Assets, Blockchain and DeFi*. Palgrave Macmillan.

Standards, I. F. R. (2019b). International Accounting Standard 32 Financial Instruments: Presentation. https://www.iasplus.com/en/standards/ias/ias32

Statista. (2023). *Number of cryptocurrencies and tokens worldwide from 2013 to 2022* [Accessed on 29.09.2023]. https://www.statista.com/statistics/863917/number-crypto-coins-tokens/

Steland, A. (2012). *Statistics and Mathematical Finance: Methods, Models and Applications* (1st ed.). Wiley.

Szylar, C. (2013). *Handbook of Market Risk* (1st ed.). John Wiley & Sons, Inc.

Taleb, N. (1996). The World According to Nassim Taleb. https://merage.uci.edu/~jorion/oc/ntaleb.htm

The MathWorks Inc. (2023). *Matlab version: 23.2.0.2380103 (r2023b)*. 1 Apple Hill Drive, Natick, MA 01760-2098. https://www.mathworks.com

Tobin, J. (1958). Liquidity Preference as Behavior Towards Risk. *The Review of economic studies*, *25*(2), 65–86. https://doi.org/10.2307/2296205

Trucíos, C., & Taylor, J. W. (2023). A comparison of methods for forecasting value at risk and expected shortfall of cryptocurrencies. *Journal of forecasting*, *42*(4), 989–1007. https://doi.org/10.1002/for.2929

Volatility in the Cryptocurrency Market. (2019). *Open economies review*, *30*(4), 779–811. https://doi.org/10.1007/s11079-019-09547-5

Wolke, T. (2017). *Risk management*. De Gruyter Oldenbourg.

Wong, A., & Vlaar, P. (2003). Modelling time-varying correlations of financial markets. https://www.researchgate.net/publication/4799583_Modelling_time-varying_correlations_of_financial_markets

Yamai, Y., Yoshiba, T., et al. (2002). On the Validity of Value-at-Risk: Comparative Analyses with Expected Shortfall. *Monetary and economic studies*, *20*(1), 57–85. https://www.imes.boj.or.jp/research/papers/english/me20-1-3.pdf

Ziggel, D., Berens, T., Weiß, G. N., & Wied, D. (2014). A new set of improved value-at-risk backtests. *Journal of banking & finance*, *48*, 29–41. https://doi.org/10.1016/j.jbankfin.2014.07.005

A Appendix: VaR estimates using the Gaussian Variance-Covariance model

This table illustrates the VaR results derived by the Gaussian VCV method for the S&P500. The term ”W_E” represents the size of the estimation window, while ”VaR cl.” indicates the confidence level for which VaR is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ”Violations” column reveals the number of violated, and ”Expected” indicates the anticipated number of model violations. The ”Ratio” represents the proportion between these two quantities. ”FirstViolation” denotes the observation at which the initial violation occurred, and ”Missing” indicates the count of missing observations.

Asset	W_E	VaR cl.	ObservedLevel	Observations	Violations	Expected	Ratio	FirstViolation	Missing
S&P500	50	0.99	0.98542	480	7	4.8	1.4583	44	0
S&P500	100	0.99	0.98125	480	9	4.8	1.875	3	0
S&P500	150	0.99	0.97083	480	14	4.8	2.9167	3	0
S&P500	200	0.99	0.96458	480	17	4.8	3.5417	3	0
S&P500	250	0.99	0.96875	480	15	4.8	3.125	3	0

This table displays the results of the statistical backtests used for model validation. The meanings of the first three columns remain consistent with those in the preceding table. Subsequently, the backtest results are outlined in the following manner: The first column specifies the test name, the subsequent one indicates the Likelihood Ratio of the test, and the third column displays the p-value of the test. To illustrate, CC corresponds to Christoffersen's Conditional Coverage test, LR denotes the Likelihood Ratio of the test, and the p-value represents the p-value of the test. Following this, the ”Observations” column provides the number of observations, and the ”Violations” column reports the count of violations, mirroring the structure of the upper table. The symbol v_{00} signifies the count of days without violations followed by another violation-free period. Similarly, v_{10} denotes the count of days with a violation succeeded by a day without any violations. Furthermore, v_{01} represents the count of days without violations followed by a day with a violation, while v_{11} represents the count of days with two consecutive violations. Lastly, ”TestLevel” indicates the confidence level of the backtests used.

Asset	W_E	VaR cl.	CC	LR	p-value	POF	LR	p-value	CCI	LR	p-value	Observations	Violations	v_{00}	v_{10}	v_{01}	v_{11}	TestLevel.
S&P500	50	0.99	Fail to Reject	1.1	0.57696	Fail to Reject	0.89232	0.34485	Fail to Reject	0.20763	0.64863	480	7	465	7	7	0	0.95
S&P500	100	0.99	Fail to Reject	3.2969	0.19235	Fail to Reject	2.9522	0.085761	Fail to Reject	0.3447	0.55713	480	9	461	9	9	0	0.95
S&P500	150	0.99	Reject	12.595	0.0018411	Reject	11.752	0.0006079	Fail to Reject	0.84314	0.3585	480	14	451	14	14	0	0.95
S&P500	200	0.99	Reject	19.149	6.9472e-05	Reject	18.912	1.3687e-05	Fail to Reject	0.23693	0.62643	480	17	446	16	16	1	0.95
S&P500	250	0.99	Reject	14.494	0.0007123	Reject	14.004	0.00018247	Fail to Reject	0.49047	0.48372	480	15	450	14	14	1	0.95

This table illustrates the VaR results derived by the Gaussian VCV method for the BTC. The term ”W_E” represents the size of the estimation window, while ”VaR cl.” indicates the confidence level for which VaR is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ”Violations” column reveals the number of violated, and ”Expected” indicates the anticipated number of model violations. The ”Ratio” represents the proportion between these two quantities. ”FirstViolation” denotes the observation at which the initial violation occurred, and ”Missing” indicates the count of missing observations.

Asset	W_E	VaR cl.	ObservedLevel	Observations	Violations	Expected	Ratio	FirstViolation	Missing
BTC	50	0.99	0.97991	697	14	6.97	2.0086	21	0
BTC	100	0.99	0.98135	697	13	6.97	1.8651	21	0
BTC	150	0.99	0.97848	697	15	6.97	2.1521	21	0
BTC	200	0.99	0.98422	697	11	6.97	1.5782	21	0
BTC	250	0.99	0.98422	697	11	6.97	1.5782	21	0

This table displays the results of the statistical backtests used for model validation. The meanings of the first three columns remain consistent with those in the preceding table. Subsequently, the backtest results are outlined in the following manner: The first column specifies the test name, the subsequent one indicates the Likelihood Ratio of the test, and the third column displays the p-value of the test. To illustrate, CC corresponds to Christoffersen’s Conditional Coverage test, LR denotes the Likelihood Ratio of the test, and the p-value represents the p-value of the test. Following this, the ”Observations” column provides the number of observations, and the ”Violations” column reports the count of violations, mirroring the structure of the upper table. The symbol v_{00} signifies the count of days without violations followed by another violation-free period. Similarly, v_{10} denotes the count of days with a violation succeeded by a day without any violations. Furthermore, v_{01} represents the count of days without violations followed by a day with a violation, while v_{11} represents the count of days with two consecutive violations. Lastly, ”TestLevel” indicates the confidence level of the backtests used.

Asset	W_E	VaR cl.	CC	LR	p-value	POF	LR	p-value	CCI	LR	p-value	Observations	Violations	v_{00}	v_{10}	v_{01}	v_{11}	TestLevel
BTC	50	0.99	Reject	6.7153	0.034817	Reject	5.5402	0.018584	Fail to Reject	1.175	0.27837	697	14	669	13	13	1	0.95
BTC	100	0.99	Fail to Reject	5.6086	0.060548	Reject	4.1995	0.040435	Fail to Reject	1.4091	0.2352	697	13	671	12	12	1	0.95
BTC	150	0.99	Reject	7.996	0.018352	Reject	7.0269	0.0080296	Fail to Reject	0.96912	0.3249	697	15	667	14	14	1	0.95
BTC	200	0.99	Fail to Reject	3.979	0.13676	Fail to Reject	2.0017	0.15712	Fail to Reject	1.9773	0.15968	697	11	675	10	10	1	0.95
BTC	250	0.99	Fail to Reject	3.979	0.13676	Fail to Reject	2.0017	0.15712	Fail to Reject	1.9773	0.15968	697	11	675	10	10	1	0.95

This table illustrates the VaR results derived by the Gaussian VCV method for the ETH. The term ”W_E” represents the size of the estimation window, while ”VaR cl.” indicates the confidence level for which VaR is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ”Violations” column reveals the number of violated, and ”Expected” indicates the anticipated number of model violations. The ”Ratio” represents the proportion between these two quantities. ”FirstViolation” denotes the observation at which the initial violation occurred, and ”Missing” indicates the count of missing observations.

Asset	W_E	VaR cl.	ObservedLevel	Observations	Violations	Expected	Ratio	FirstViolation	Missing
ETH	50	0.99	0.97848	697	15	6.97	2.1521	21	0
ETH	100	0.99	0.98278	697	12	6.97	1.7217	21	0
ETH	150	0.99	0.98278	697	12	6.97	1.7217	21	0
ETH	200	0.99	0.98422	697	11	6.97	1.5782	21	0
ETH	250	0.99	0.98565	697	10	6.97	1.4347	21	0

This table displays the results of the statistical backtests used for model validation. The meanings of the first three columns remain consistent with those in the preceding table. Subsequently, the backtest results are outlined in the following manner: The first column specifies the test name, the subsequent one indicates the Likelihood Ratio of the test, and the third column displays the p-value of the test. To illustrate, CC corresponds to Christoffersen's Conditional Coverage test, LR denotes the Likelihood Ratio of the test, and the p-value represents the p-value of the test. Following this, the ”Observations” column provides the number of observations, and the ”Violations” column reports the count of violations, mirroring the structure of the upper table. The symbol v_{00} signifies the count of days without violations followed by another violation-free period. Similarly, v_{10} denotes the count of days with a violation succeeded by a day without any violations. Furthermore, v_{01} represents the count of days without violations followed by a day with a violation, while v_{11} represents the count of days with two consecutive violations. Lastly, ”TestLevel” indicates the confidence level of the backtests used.

Asset	W_E	VaR cl.	CC	LR	p-value	POF	LR	p-value	CCI	LR	p-value	Observations	Violations	v_{00}	v_{10}	v_{01}	v_{11}	TestLevel
ETH	50	0.99	Reject	7.996	0.018352	Reject	7.0269	0.0080296	Fail to Reject	0.96912	0.3249	697	15	667	14	14	1	0.95
ETH	100	0.99	Fail to Reject	4.6907	0.095814	Fail to Reject	3.0157	0.082459	Fail to Reject	1.675	0.1956	697	12	673	11	11	1	0.95
ETH	150	0.99	Fail to Reject	4.6907	0.095814	Fail to Reject	3.0157	0.082459	Fail to Reject	1.675	0.1956	697	12	673	11	11	1	0.95
ETH	200	0.99	Fail to Reject	3.979	0.13676	Fail to Reject	2.0017	0.15712	Fail to Reject	1.9773	0.15968	697	11	675	10	10	1	0.95
ETH	250	0.99	Fail to Reject	3.4948	0.17422	Fail to Reject	1.1727	0.27884	Fail to Reject	2.3221	0.12755	697	10	677	9	9	1	0.95

B Appendix: VaR estimates using the Gaussian GARCH(1,1) model

This table illustrates the VaR results derived by the Gaussian GARCH(1,1) model for the S&P500. The term ”W_E” represents the size of the estimation window, while ”VaR cl.” indicates the confidence level for which VaR is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ”Violations” column reveals the number of violated, and ”Expected” indicates the anticipated number of model violations. The ”Ratio” represents the proportion between these two quantities. ”FirstViolation” denotes the observation at which the initial violation occurred, and ”Missing” indicates the count of missing observations.

Asset	W_E	VaR cl.	ObservedLevel	Observations	Violations	Expected	Ratio	FirstViolation	Missing
S&P50	50	0.99	0.98542	480	7	4.8	1.4583	3	0
S&P50	100	0.99	0.98125	480	9	4.8	1.875	3	0
S&P50	150	0.99	0.98125	480	9	4.8	1.875	3	0
S&P50	200	0.99	0.98333	480	8	4.8	1.6667	3	0
S&P50	250	0.99	0.98333	480	8	4.8	1.6667	3	0

This table displays the results of the statistical backtests used for model validation. The meanings of the first three columns remain consistent with those in the preceding table. Subsequently, the backtest results are outlined in the following manner: The first column specifies the test name, the subsequent one indicates the Likelihood Ratio of the test, and the third column displays the p-value of the test. To illustrate, CC corresponds to Christoffersen's Conditional Coverage test, LR denotes the Likelihood Ratio of the test, and the p-value represents the p-value of the test. Following this, the ”Observations” column provides the number of observations, and the ”Violations” column reports the count of violations, mirroring the structure of the upper table. The symbol v_{00} signifies the count of days without violations followed by another violation-free period. Similarly, v_{10} denotes the count of days with a violation succeeded by a day without any violations. Furthermore, v_{01} represents the count of days without violations followed by a day with a violation, while v_{11} represents the count of days with two consecutive violations. Lastly, ”TestLevel” indicates the confidence level of the backtests used.

Asset	W_E	VaR cl.	CC	LR	p-value	POF	LR	p-value	CCI	LR	p-value	Observations	Violations	v_{00}	v_{10}	v_{01}	v_{11}	TestLevel
S&P50	50	0.99	Fail to Reject	1.1	0.57696	Fail to Reject	0.89232	0.34485	Fail to Reject	0.20763	0.64863	480	7	465	7	7	0	0.95
S&P50	100	0.99	Fail to Reject	3.2969	0.19235	Fail to Reject	2.9522	0.085761	Fail to Reject	0.3447	0.55713	480	9	461	9	9	0	0.95
S&P50	150	0.99	Fail to Reject	3.2969	0.19235	Fail to Reject	2.9522	0.085761	Fail to Reject	0.3447	0.55713	480	9	461	9	9	0	0.95
S&P50	200	0.99	Fail to Reject	2.0666	0.35583	Fail to Reject	1.7948	0.18034	Fail to Reject	0.27178	0.60214	480	8	463	8	8	0	0.95
S&P50	250	0.99	Fail to Reject	2.0666	0.35583	Fail to Reject	1.7948	0.18034	Fail to Reject	0.27178	0.60214	480	8	463	8	8	0	0.95

This table illustrates the VaR results derived by the Gaussian GARCH(1,1) model for the BTC. The term ”W_E” represents the size of the estimation window, while ”VaR cl.” indicates the confidence level for which VaR is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ”Violations” column reveals the number of violated, and ”Expected” indicates the anticipated number of model violations. The ”Ratio” represents the proportion between these two quantities. ”FirstViolation” denotes the observation at which the initial violation occurred, and ”Missing” indicates the count of missing observations.

Asset	W_E	VaR cl.	ObservedLevel	Observations	Violations	Expected	Ratio	FirstViolation	Missing
BTC	50	0.99	0.97561	697	17	6.97	2.439	21	0
BTC	100	0.99	0.97704	697	16	6.97	2.2956	21	0
BTC	150	0.99	0.97274	697	19	6.97	2.726	21	0
BTC	200	0.99	0.97991	697	14	6.97	2.0086	21	0
BTC	250	0.99	0.97991	697	14	6.97	2.0086	21	0

This table displays the results of the statistical backtests used for model validation. The meanings of the first three columns remain consistent with those in the preceding table. Subsequently, the backtest results are outlined in the following manner: The first column specifies the test name, the subsequent one indicates the Likelihood Ratio of the test, and the third column displays the p-value of the test. To illustrate, CC corresponds to Christoffersen's Conditional Coverage test, LR denotes the Likelihood Ratio of the test, and the p-value represents the p-value of the test. Following this, the ”Observations” column provides the number of observations, and the ”Violations” column reports the count of violations, mirroring the structure of the upper table. The symbol v_{00} signifies the count of days without violations followed by another violation-free period. Similarly, v_{10} denotes the count of days with a violation succeeded by a day without any violations. Furthermore, v_{01} represents the count of days without violations followed by a day with a violation, while v_{11} represents the count of days with two consecutive violations. Lastly, ”TestLevel” indicates the confidence level of the backtests used.

Asset	W_E	VaR cl.	CC	LR	p-value	POF	LR	p-value	CCI	LR	p-value	Observations	Violations	v_{00}	v_{10}	v_{01}	v_{11}	TestLevel
BTC	50	0.99	reject	11.031	0.0040231	reject	10.401	0.0012596	Fail to Reject	0.63055	0.42715	697	17	663	16	16	1	0.95
BTC	100	0.99	reject	9.4383	0.008923	reject	8.6498	0.0032709	Fail to Reject	0.78841	0.37458	697	16	665	15	15	1	0.95
BTC	150	0.99	reject	14.634	0.00066408	reject	14.258	0.00015936	Fail to Reject	0.37594	0.53978	697	19	659	18	18	1	0.95
BTC	200	0.99	reject	6.7153	0.034817	reject	5.5402	0.018584	Fail to Reject	1.175	0.27837	697	14	669	13	13	1	0.95
BTC	250	0.99	reject	6.7153	0.034817	reject	5.5402	0.018584	Fail to Reject	1.175	0.27837	697	14	669	13	13	1	0.95

This table illustrates the VaR results derived by the Gaussian GARCH(1,1) model for the ETH. The term ”W_E” represents the size of the estimation window, while ”VaR cl.” indicates the confidence level for which VaR is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ”Violations” column reveals the number of violated, and ”Expected” indicates the anticipated number of model violations. The ”Ratio” represents the proportion between these two quantities. ”FirstViolation” denotes the observation at which the initial violation occurred, and ”Missing” indicates the count of missing observations.

Asset	W_E	VaR cl.	ObservedLevel	Observations	Violations	Expected	Ratio	FirstViolation	Missing
ETH	50	0.99	0.97131	697	20	6.97	2.8694	5	0
ETH	100	0.99	0.98422	697	11	6.97	1.5782	21	0
ETH	150	0.99	0.98135	697	13	6.97	1.8651	21	0
ETH	200	0.99	0.98422	697	11	6.97	1.5782	21	0
ETH	250	0.99	0.98422	697	11	6.97	1.5782	21	0

This table displays the results of the statistical backtests used for model validation. The meanings of the first three columns remain consistent with those in the preceding table. Subsequently, the backtest results are outlined in the following manner: The first column specifies the test name, the subsequent one indicates the Likelihood Ratio of the test, and the third column displays the p-value of the test. To illustrate, CC corresponds to Christoffersen's Conditional Coverage test, LR denotes the Likelihood Ratio of the test, and the p-value represents the p-value of the test. Following this, the ”Observations” column provides the number of observations, and the ”Violations” column reports the count of violations, mirroring the structure of the upper table. The symbol v_{00} signifies the count of days without violations followed by another violation-free period. Similarly, v_{10} denotes the count of days with a violation succeeded by a day without any violations. Furthermore, v_{01} represents the count of days without violations followed by a day with a violation, while v_{11} represents the count of days with two consecutive violations. Lastly, ”TestLevel” indicates the confidence level of the backtests used.

Asset	W_E	VaR cl.	CC	LR	p-value	POF	LR	p-value	CCI	LR	p-value	Observations	Violations	v_{00}	v_{10}	v_{01}	v_{11}	TestLevel
ETH	50	0.99	reject	17.536	0.00015564	reject	16.352	5.2592e-05	Fail to Reject	1.1836	0.27662	697	20	656	20	20	0	0.95
ETH	100	0.99	Fail to Reject	3.979	0.13676	Fail to Reject	2.0017	0.15712	Fail to Reject	1.9773	0.15968	697	11	675	10	10	1	0.95
ETH	150	0.99	Fail to Reject	5.6086	0.060548	reject	4.1995	0.040435	Fail to Reject	1.4091	0.2352	697	13	671	12	12	1	0.95
ETH	200	0.99	Fail to Reject	3.979	0.13676	Fail to Reject	2.0017	0.15712	Fail to Reject	1.9773	0.15968	697	11	675	10	10	1	0.95
ETH	250	0.99	Fail to Reject	3.979	0.13676	Fail to Reject	2.0017	0.15712	Fail to Reject	1.9773	0.15968	697	11	675	10	10	1	0.95

C Appendix: VaR estimates using the Student's t GARCH(1,1) model

This table illustrates the VaR results derived by the Student's t GARCH(1,1) model for the S&P500. The term "W_E" represents the size of the estimation window, while "VaR cl." indicates the confidence level for which VaR is being estimated. The "ObservedLevel" column indicates the actual confidence level. The "Violations" column reveals the number of violated, and "Expected" indicates the anticipated number of model violations. The "Ratio" represents the proportion between these two quantities. "FirstViolation" denotes the observation at which the initial violation occurred, and "Missing" indicates the count of missing observations.

Asset	W_E	VaR cl.	ObservedLevel	Observations	Violations	Expected	Ratio	FirstViolation	Missing
S&P500	50	0.99	0.99167	480	4	4.8	0.83333	3	0
S&P500	100	0.99	0.99375	480	3	4.8	0.625	3	0
S&P500	150	0.99	0.99375	480	3	4.8	0.625	3	0
S&P500	200	0.99	0.99167	480	4	4.8	0.83333	3	0
S&P500	250	0.99	0.99167	480	4	4.8	0.83333	3	0

This table displays the results of the statistical backtests used for model validation. The meanings of the first three columns remain consistent with those in the preceding table. Subsequently, the backtest results are outlined in the following manner: The first column specifies the test name, the subsequent one indicates the Likelihood Ratio of the test, and the third column displays the p-value of the test. To illustrate, CC corresponds to Christoffersen's Conditional Coverage test, LR denotes the Likelihood Ratio of the test, and the p-value represents the p-value of the test. Following this, the "Observations" column provides the number of observations, and the "Violations" column reports the count of violations, mirroring the structure of the upper table. The symbol v_{00} signifies the count of days without violations followed by another violation-free period. Similarly, v_{10} denotes the count of days with a violation succeeded by a day without any violations. Furthermore, v_{01} represents the count of days without violations followed by a day with a violation, while v_{11} represents the count of days with two consecutive violations. Lastly, "TestLevel" indicates the confidence level of the backtests used.

PortfolioID	W_E	VaR cl.	CC	LR	p-value	POF	LR	p-value	CCI	LR	p-value	Observations	Violations	v_{00}	v_{10}	v_{01}	v_{11}	TestLevel
S&P500	50	0.99	Fail to Reject	0.21014	0.90026	Fail to Reject	0.14277	0.70554	Fail to Reject	0.067369	0.79521	480	4	471	4	4	0	0.95
S&P500	100	0.99	Fail to Reject	0.8246	0.66212	Fail to Reject	0.78679	0.37507	Fail to Reject	0.037815	0.84581	480	3	473	3	3	0	0.95
S&P500	150	0.99	Fail to Reject	0.8246	0.66212	Fail to Reject	0.78679	0.37507	Fail to Reject	0.037815	0.84581	480	3	473	3	3	0	0.95
S&P500	200	0.99	Fail to Reject	0.21014	0.90026	Fail to Reject	0.14277	0.70554	Fail to Reject	0.067369	0.79521	480	4	471	4	4	0	0.95
S&P500	250	0.99	Fail to Reject	0.21014	0.90026	Fail to Reject	0.14277	0.70554	Fail to Reject	0.067369	0.79521	480	4	471	4	4	0	0.95

This table illustrates the VaR results derived by the Student's t GARCH(1,1) model for the BTC. The term "W_E" represents the size of the estimation window, while "VaR cl." indicates the confidence level for which VaR is being estimated. The "ObservedLevel" column indicates the actual confidence level. The "Violations" column reveals the number of violated, and "Expected" indicates the anticipated number of model violations. The "Ratio" represents the proportion between these two quantities. "FirstViolation" denotes the observation at which the initial violation occurred, and "Missing" indicates the count of missing observations.

Asset	W_E	VaR cl.	ObservedLevel	Observations	Violations	Expected	Ratio	FirstViolation	Missing
BTC	50	0.99	0.98709	697	9	6.97	1.2912	21	0
BTC	100	0.99	0.99426	697	4	6.97	0.57389	164	0
BTC	150	0.99	0.99139	697	6	6.97	0.86083	129	0
BTC	200	0.99	0.99283	697	5	6.97	0.71736	129	0
BTC	250	0.99	0.99426	697	4	6.97	0.57389	164	0

This table displays the results of the statistical backtests used for model validation. The meanings of the first three columns remain consistent with those in the preceding table. Subsequently, the backtest results are outlined in the following manner: The first column specifies the test name, the subsequent one indicates the Likelihood Ratio of the test, and the third column displays the p-value of the test. To illustrate, CC corresponds to Christoffersen's Conditional Coverage test, LR denotes the Likelihood Ratio of the test, and the p-value represents the p-value of the test. Following this, the "Observations" column provides the number of observations, and the "Violations" column reports the count of violations, mirroring the structure of the upper table. The symbol v_{00} signifies the count of days without violations followed by another violation-free period. Similarly, v_{10} denotes the count of days with a violation succeeded by a day without any violations. Furthermore, v_{01} represents the count of days without violations followed by a day with a violation, while v_{11} represents the count of days with two consecutive violations. Lastly, "TestLevel" indicates the confidence level of the backtests used.

Asset	W_E	VaR cl.	CC	LR	p-value	POF	LR	p-value	CCI	LR	p-value	Observations	Violations	v_{00}	v_{10}	v_{01}	v_{11}	TestLevel
BTC	50	0.99	Fail to Reject	0.78276	0.67612	Fail to Reject	0.54695	0.45957	Fail to Reject	0.23581	0.62725	697	9	678	9	9	0	0.95
BTC	100	0.99	reject	7.6281	0.022059	Fail to Reject	1.5102	0.21911	reject	6.1179	0.013382	697	4	689	3	3	1	0.95
BTC	150	0.99	Fail to Reject	4.4914	0.10586	Fail to Reject	0.14309	0.70522	reject	4.3483	0.037047	697	6	685	5	5	1	0.95
BTC	200	0.99	Fail to Reject	5.7514	0.056378	Fail to Reject	0.62385	0.42962	reject	5.1275	0.023549	697	5	687	4	4	1	0.95
BTC	250	0.99	reject	7.6281	0.022059	Fail to Reject	1.5102	0.21911	reject	6.1179	0.013382	697	4	689	3	3	1	0.95

This table illustrates the VaR results derived by the Student's t GARCH(1,1) model for the ETH. The term "W_E" represents the size of the estimation window, while "VaR cl." indicates the confidence level for which VaR is being estimated. The "ObservedLevel" column indicates the actual confidence level. The "Violations" column reveals the number of violated, and "Expected" indicates the anticipated number of model violations. The "Ratio" represents the proportion between these two quantities. "FirstViolation" denotes the observation at which the initial violation occurred, and "Missing" indicates the count of missing observations.

Asset	W_E	VaR cl.	ObservedLevel	Observations	Violations	Expected	Ratio	FirstViolation	Missing
ETH	50	0.99	0.98565	697	10	6.97	1.4347	21	0
ETH	100	0.99	0.98996	697	7	6.97	1.0043	21	0
ETH	150	0.99	0.98565	697	10	6.97	1.4347	21	0
ETH	200	0.99	0.98996	697	7	6.97	1.0043	21	0
ETH	250	0.99	0.99139	697	6	6.97	0.86083	21	0

This table displays the results of the statistical backtests used for model validation. The meanings of the first three columns remain consistent with those in the preceding table. Subsequently, the backtest results are outlined in the following manner: The first column specifies the test name, the subsequent one indicates the Likelihood Ratio of the test, and the third column displays the p-value of the test. To illustrate, CC corresponds to Christoffersen's Conditional Coverage test, LR denotes the Likelihood Ratio of the test, and the p-value represents the p-value of the test. Following this, the "Observations" column provides the number of observations, and the "Violations" column reports the count of violations, mirroring the structure of the upper table. The symbol v_{00} signifies the count of days without violations followed by another violation-free period. Similarly, v_{10} denotes the count of days with a violation succeeded by a day without any violations. Furthermore, v_{01} represents the count of days without violations followed by a day with a violation, while v_{11} represents the count of days with two consecutive violations. Lastly, "TestLevel" indicates the confidence level of the backtests used.

Asset	W_E	VaR cl.	CC	LR	p-value	POF	LR	p-value	CCI	LRI	p-value	Observations	Violations	v_{00}	v_{10}	v_{01}	v_{11}	TestLevel
ETH	50	0.99	Fail to Reject	1.4643	0.48088	Fail to Reject	1.1727	0.27884	Fail to Reject	0.29156	0.58923	697	10	676	10	10	0	0.95
ETH	100	0.99	Fail to Reject	0.14237	0.93129	Fail to Reject	0.00013024	0.99089	Fail to Reject	0.14224	0.70607	697	7	682	7	7	0	0.95
ETH	150	0.99	Fail to Reject	1.4643	0.48088	Fail to Reject	1.1727	0.27884	Fail to Reject	0.29156	0.58923	697	10	676	10	10	0	0.95
ETH	200	0.99	Fail to Reject	0.14237	0.93129	Fail to Reject	0.00013024	0.99089	Fail to Reject	0.14224	0.70607	697	7	682	7	7	0	0.95
ETH	250	0.99	Fail to Reject	0.24744	0.88363	Fail to Reject	0.14309	0.70522	Fail to Reject	0.10435	0.74667	697	6	684	6	6	0	0.95

D Appendix: VaR estimates using the Historical Simulation model

This table illustrates the VaR results derived by the HS method for the S&P500. The term ”W_E” represents the size of the estimation window, while ”VaR cl.” indicates the confidence level for which VaR is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ”Violations” column reveals the number of violated, and ”Expected” indicates the anticipated number of model violations. The ”Ratio” represents the proportion between these two quantities. ”FirstViolation” denotes the observation at which the initial violatin occurred, and ”Missing” indicates the count of missing observations.

Asset	W_E	VaR cl.	ObservedLevel	Observations	Violations	Expected	Ratio	FirstViolation	Missing
S&P500	50	0.99	0.98333	480	8	4.8	1.6667	23	0
S&P500	100	0.99	0.98125	480	9	4.8	1.875	23	0
S&P500	150	0.99	0.98125	480	9	4.8	1.875	23	0
S&P500	200	0.99	0.98125	480	9	4.8	1.875	23	0
S&P500	250	0.99	0.97917	480	10	4.8	2.0833	23	0

This table displays the results of the statistical backtests used for model validation. The meanings of the first three columns remain consistent with those in the preceding table. Subsequently, the backtest results are outlined in the following manner: The first column specifies the test name, the subsequent one indicates the Likelihood Ratio of the test, and the third column displays the p-value of the test. To illustrate, CC corresponds to Christoffersen's Conditional Coverage test, LR denotes the Likelihood Ratio of the test, and the p-value represents the p-value of the test. Following this, the ”Observations” column provides the number of observations, and the ”Violations” column reports the count of violations, mirroring the structure of the upper table. The symbol v_{00} signifies the count of days without violations followed by another violation-free period. Similarly, v_{10} denotes the count of days with a violation succeeded by a day without any violations. Furthermore, v_{01} represents the count of days without violations followed by a day with a violation, while v_{11} represents the count of days with two consecutive violations. Lastly, ”TestLevel” indicates the confidence level of the backtests used.

Asset	W_E	VaR cl.	CC	LR	p-value	POF	LR	p-value	CCI	LR	p-value	Observations	Violations	v_{00}	v_{10}	v_{01}	v_{11}	TestLevel
S&P500	50	0.99	Fail to Reject	2.0666	0.35583	Fail to Reject	1.7948	0.18034	Fail to Reject	0.27178	0.60214	480	8	463	8	8	0	0.95
S&P500	100	0.99	Fail to Reject	3.2969	0.19235	Fail to Reject	2.9522	0.085761	Fail to Reject	0.3447	0.55713	480	9	461	9	9	0	0.95
S&P500	150	0.99	Fail to Reject	3.2969	0.19235	Fail to Reject	2.9522	0.085761	Fail to Reject	0.3447	0.55713	480	9	461	9	9	0	0.95
S&P500	200	0.99	Fail to Reject	3.2969	0.19235	Fail to Reject	2.9522	0.085761	Fail to Reject	0.3447	0.55713	480	9	461	9	9	0	0.95
S&P500	250	0.99	Fail to Reject	4.763	0.092413	Reject	4.3365	0.037304	Fail to Reject	0.42647	0.51373	480	10	459	10	10	0	0.95

This table illustrates the VaR results derived by HS method for the BTC. The term ”W_E” represents the size of the estimation window, while ”VaR cl.” indicates the confidence level for which VaR is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ”Violations” column reveals the number of violated, and ”Expected” indicates the anticipated number of model violations. The ”Ratio” represents the proportion between these two quantities. ”FirstViolation” denotes the observation at which the initial violation occurred, and ”Missing” indicates the count of missing observations.

Asset	W_E	VaR cl.	ObservedLevel	Observations	Violations	Expected	Ratio	FirstViolation	Missing
BTC	50	0.99	0.98278	697	12	6.97	1.7217	21	0
BTC	100	0.99	0.98852	697	8	6.97	1.1478	21	0
BTC	150	0.99	0.98996	697	7	6.97	1.0043	21	0
BTC	200	0.99	0.99139	697	6	6.97	0.86083	21	0
BTC	250	0.99	0.99283	697	5	6.97	0.71736	129	0

This table displays the results of the statistical backtests used for model validation. The meanings of the first three columns remain consistent with those in the preceding table. Subsequently, the backtest results are outlined in the following manner: The first column specifies the test name, the subsequent one indicates the Likelihood Ratio of the test, and the third column displays the p-value of the test. To illustrate, CC corresponds to Christoffersen's Conditional Coverage test, LR denotes the Likelihood Ratio of the test, and the p-value represents the p-value of the test. Following this, the ”Observations” column provides the number of observations, and the ”Violations” column reports the count of violations, mirroring the structure of the upper table. The symbol v_{00} signifies the count of days without violations followed by another violation-free period. Similarly, v_{10} denotes the count of days with a violation succeeded by a day without any violations. Furthermore, v_{01} represents the count of days without violations followed by a day with a violation, while v_{11} represents the count of days with two consecutive violations. Lastly, ”TestLevel” indicates the confidence level of the backtests used.

Asset	W_E	VaR cl.	CC	LR	p-value	POF	LR	p-value	CCI	LR	p-value	Observations	Violations	v_{00}	v_{10}	v_{01}	v_{11}	TestLevel
BTC	50	0.99	Fail to Reject	4.6907	0.095814	Fail to Reject	3.0157	0.082459	Fail to Reject	1.675	0.1956	697	12	673	11	11	1	0.95
BTC	100	0.99	Fail to Reject	3.3218	0.18997	Fail to Reject	0.14676	0.70165	Fail to Reject	3.1751	0.07477	697	8	681	7	7	1	0.95
BTC	150	0.99	Fail to Reject	3.7106	0.1564	Fail to Reject	0.00013024	0.99089	Fail to Reject	3.7105	0.054071	697	7	683	6	6	1	0.95
BTC	200	0.99	Fail to Reject	4.4914	0.10586	Fail to Reject	0.14309	0.70522	Reject	4.3483	0.037047	697	6	685	5	5	1	0.95
BTC	250	0.99	Fail to Reject	5.7514	0.056378	Fail to Reject	0.62385	0.42962	Reject	5.1275	0.023549	697	5	687	4	4	1	0.95

This table illustrates the VaR results derived by HS method for the ETH. The term ”W_E” represents the size of the estimation window, while ”VaR cl.” indicates the confidence level for which VaR is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ”Violations” column reveals the number of violated, and ”Expected” indicates the anticipated number of model violations. The ”Ratio” represents the proportion between these two quantities. ”FirstViolation” denotes the observation at which the initial violatin occurred, and ”Missing” indicates the count of missing observations.

Asset	W_E	VaR cl.	ObservedLevel	Observations	Violations	Expected	Ratio	FirstViolation	Missing
ETH	50	0.99	0.98135	697	13	6.97	1.8651	21	0
ETH	100	0.99	0.98852	697	8	6.97	1.1478	21	0
ETH	150	0.99	0.98709	697	9	6.97	1.2912	21	0
ETH	200	0.99	0.98709	697	9	6.97	1.2912	21	0
ETH	250	0.99	0.98996	697	7	6.97	1.0043	21	0

This table displays the results of the statistical backtests used for model validation. The meanings of the first three columns remain consistent with those in the preceding table. Subsequently, the backtest results are outlined in the following manner: The first column specifies the test name, the subsequent one indicates the Likelihood Ratio of the test, and the third column displays the p-value of the test. To illustrate, CC corresponds to Christoffersen's Conditional Coverage test, LR denotes the Likelihood Ratio of the test, and the p-value represents the p-value of the test. Following this, the ”Observations” column provides the number of observations, and the ”Violations” column reports the count of violations, mirroring the structure of the upper table. The symbol v_{00} signifies the count of days without violations followed by another violation-free period. Similarly, v_{10} denotes the count of days with a violation succeeded by a day without any violations. Furthermore, v_{01} represents the count of days without violations followed by a day with a violation, while v_{11} represents the count of days with two consecutive violations. Lastly, ”TestLevel” indicates the confidence level of the backtests used.

Asset	W_E	VaR cl.	CC	LR	p-value	POF	LR	p-value	CCI	LR	p-value	Observations	Violations	v_{00}	v_{10}	v_{01}	v_{11}	TestLevel
ETH	50	0.99	Fail to Reject	5.6086	0.060548	Reject	4.1995	0.040435	Fail to Reject	1.4091	0.2352	697	13	671	12	12	1	0.95
ETH	100	0.99	Fail to Reject	3.3218	0.18997	Fail to Reject	0.14676	0.70165	Fail to Reject	3.1751	0.07477	697	8	681	7	7	1	0.95
ETH	150	0.99	Fail to Reject	3.2646	0.19548	Fail to Reject	0.54695	0.45957	Fail to Reject	2.7176	0.099245	697	9	679	8	8	1	0.95
ETH	200	0.99	Fail to Reject	3.2646	0.19548	Fail to Reject	0.54695	0.45957	Fail to Reject	2.7176	0.099245	697	9	679	8	8	1	0.95
ETH	250	0.99	Fail to Reject	3.7106	0.1564	Fail to Reject	0.00013024	0.99089	Fail to Reject	3.7105	0.054071	697	7	683	6	6	1	0.95

E Appendix: ES estimates using the Gaussian Variance-Covariance model

This table illustrates the ES results derived by the Gaussian VCV model for the S&P500. The term ”W_E” represents the size of the estimation window, while ”ES cl.” indicates the confidence level for which ES is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ExpectedSeverity tells how severe VaR excaadences are expected to be, whereas ObservedSeverity tells the empirical value. The ”Observations” column indicates the number of observations. ”Violations” denotes the number of violations the model faced, and ”Expected” indicates the number of expected violations. The ”Ratio” represents the proportion between these two quantities and ”Missing” indicates the count of missing observations.

Asset	W_E	ES cl.	ObservedLevel	ExpectedSeverity	ObservedSeverity	Observations	Violations	Expected	Ratio	Missing
S&P500	50	0.975	0.96667	1.1928	1.2085	480	16	12	1.3333	0
S&P500	100	0.975	0.95833	1.1928	1.2012	480	20	12	1.6667	0
S&P500	150	0.975	0.95625	1.1928	1.2881	480	21	12	1.75	0
S&P500	200	0.975	0.95833	1.1928	1.3964	480	20	12	1.6667	0
S&P500	250	0.975	0.95417	1.1928	1.4052	480	22	12	1.8333	0

This table presents the backtesting results of ES using UC. The initial three columns mirror those in the preceding table. The ”UnconditionalTest” column indicates whether the test is rejected or not. The ”p-value” displays the test’s p-value, while the ”TestStatistic” column reveals the actual value of the UC test. In the ”CriticalValue” column, you can find the critical value associated with the UC test. The ”Observations” column denotes the number of observations utilized, and the ”TestLevel” column specifies the confidence level of the UC test.

Asset	W_E	ES cl.	UnconditionalTest	p-value	TestStatistic	CriticalValue	Observations	TestLevel
S&P500	50	0.975	Fail to Reject	0.12186	-0.35095	-0.49713	480	0.95
S&P500	100	0.975	Reject	0.015981	-0.67846	-0.49713	480	0.95
S&P500	150	0.975	Reject	0.0034097	-0.88986	-0.49713	480	0.95
S&P500	200	0.975	Reject	0.0020797	-0.95119	-0.49713	480	0.95
S&P500	250	0.975	Reject	0.00026473	-1.1598	-0.49713	480	0.95

This table illustrates the ES results derived by the Gaussian VCV model for the BTC. The term "W_E" represents the size of the estimation window, while "ES cl." indicates the confidence level for which ES is being estimated. The "ObservedLevel" column indicates the actual confidence level. The ExpectedSeverity tells how severe VaR excaadences are expected to be, whereas ObservedSeverity tells the empirical value. The "Observations" column indicates the number of observations. "Violations" denotes the number of violations the model faced, and "Expected" indicates the number of expected violations. The "Ratio" represents the proportion between these two quantities and "Missing" indicates the count of missing observations.

Asset	W_E	ES cl.	ObservedLevel	ExpectedSeverity	ObservedSeverity	Observations	Violations	Expected	Ratio	Missing
BTC	50	0.975	0.97131	1.1928	1.7132	697	20	17.425	1.1478	0
BTC	100	0.975	0.97418	1.1928	1.572	697	18	17.425	1.033	0
BTC	150	0.975	0.97418	1.1928	1.4955	697	18	17.425	1.033	0
BTC	200	0.975	0.97274	1.1928	1.4339	697	19	17.425	1.0904	0
BTC	250	0.975	0.97418	1.1928	1.436	697	18	17.425	1.033	0

This table presents the backtesting results of ES using UC. The initial three columns mirror those in the preceding table. The "UnconditionalTest" column indicates whether the test is rejected or not. The "p-value" displays the test's p-value, while the "TestStatistic" column reveals the actual value of the UC test. In the "CriticalValue" column, you can find the critical value associated with the UC test. The "Observations" column denotes the number of observations utilized, and the "TestLevel" column specifies the confidence level of the UC test.

Asset	W_E	ES cl.	UnconditionalTest	p-value	TestStatistic	CriticalValue	Observations	TestLevel
BTC	50	0.975	Reject	0.0068092	-0.64853	-0.41331	697	0.95
BTC	100	0.975	Fail to Reject	0.076596	-0.36145	-0.41331	697	0.95
BTC	150	0.975	Fail to Reject	0.11958	-0.29518	-0.41331	697	0.95
BTC	200	0.975	Fail to Reject	0.10477	-0.31078	-0.41331	697	0.95
BTC	250	0.975	Fail to Reject	0.16853	-0.2436	-0.41331	697	0.95

This table illustrates the ES results derived by the Gaussian VCV model for the ETH. The term "W_E" represents the size of the estimation window, while "ES cl." indicates the confidence level for which ES is being estimated. The "ObservedLevel" column indicates the actual confidence level. The ExpectedSeverity tells how severe VaR excaadences are expected to be, whereas ObservedSeverity tells the empirical value. The "Observations" column indicates the number of observations. "Violations" denotes the number of violations the model faced, and ”Expected” indicates the number of expected violations. The ”Ratio” represents the proportion between these two quantities and ”Missing” indicates the count of missing observations.

Asset	W_E	ES cl.	ObservedLevel	ExpectedSeverity	ObservedSeverity	Observations	Violations	Expected	Ratio	Missing
ETH	50	0.975	0.96557	1.1928	1.5111	697	24	17.425	1.3773	0
ETH	100	0.975	0.97274	1.1928	1.5012	697	19	17.425	1.0904	0
ETH	150	0.975	0.97274	1.1928	1.4122	697	19	17.425	1.0904	0
ETH	200	0.975	0.97274	1.1928	1.4067	697	19	17.425	1.0904	0
ETH	250	0.975	0.97418	1.1928	1.4134	697	18	17.425	1.033	0

This table presents the backtesting results of ES using UC. The initial three columns mirror those in the preceding table. The ”UnconditionalTest” column indicates whether the test is rejected or not. The ”p-value” displays the test’s p-value, while the ”TestStatistic” column reveals the actual value of the UC test. In the ”CriticalValue” column, you can find the critical value associated with the UC test. The ”Observations” column denotes the number of observations utilized, and the ”TestLevel” column specifies the confidence level of the UC test.

Asset	W_E	ES cl.	UnconditionalTest	p-value	TestStatistic	CriticalValue	Observations	TestLevel
ETH	50	0.975	Reject	0.003084	-0.74493	-0.41331	697	0.95
ETH	100	0.975	Fail to Reject	0.071002	-0.37236	-0.41331	697	0.95
ETH	150	0.975	Fail to Reject	0.12358	-0.29096	-0.41331	697	0.95
ETH	200	0.975	Fail to Reject	0.12836	-0.28592	-0.41331	697	0.95
ETH	250	0.975	Fail to Reject	0.18704	-0.2241	-0.41331	697	0.95

F Appendix: ES estimates using the Gaussian GARCH(1,1) model

This table illustrates the ES results derived by the Gaussian GARCH(1,1) model for the S&P500. The term ”W_E” represents the size of the estimation window, while ”ES cl.” indicates the confidence level for which ES is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ExpectedSeverity tells how severe VaR excaadences are expected to be, whereas ObservedSeverity tells the empirical value. The ”Observations” column indicates the number of observations. ”Violations” denotes the number of violations the model faced, and ”Expected” indicates the number of expected violations. The ”Ratio” represents the proportion between these two quantities and ”Missing” indicates the count of missing observations.

Asset	W_E	ES cl.	ObservedLevel	ExpectedSeverity	ObservedSeverity	Observations	Violations	Expected	Ratio	Missing
S&P500	50	0.975	0.96458	1.1928	1.2303	480	17	12	1.4167	0
S&P500	100	0.975	0.96458	1.1928	1.2027	480	17	12	1.4167	0
S&P500	150	0.975	0.9625	1.1928	1.1987	480	18	12	1.5	0
S&P500	200	0.975	0.9625	1.1928	1.2127	480	18	12	1.5	0
S&P500	250	0.975	0.96458	1.1928	1.2252	480	17	12	1.4167	0

This table presents the backtesting results of ES using UC. The initial three columns mirror those in the preceding table. The ”UnconditionalTest” column indicates whether the test is rejected or not. The ”p-value” displays the test’s p-value, while the ”TestStatistic” column reveals the actual value of the UC test. In the ”CriticalValue” column, you can find the critical value associated with the UC test. The ”Observations” column denotes the number of observations utilized, and the ”TestLevel” column specifies the confidence level of the UC test.

Asset	W_E	ES cl.	UnconditionalTest	p-value	TestStatistic	CriticalValue	Observations	TestLevel
S&P500	50	0.975	Fail to Reject	0.065189	-0.46119	-0.49713	480	0.95
S&P500	100	0.975	Fail to Reject	0.079002	-0.4285	-0.49713	480	0.95
S&P500	150	0.975	Reject	0.047562	-0.50743	-0.49713	480	0.95
S&P500	200	0.975	Reject	0.043389	-0.52505	-0.49713	480	0.95
S&P500	250	0.975	Fail to Reject	0.067719	-0.4552	-0.49713	480	0.95

This table illustrates the ES results derived by the Gaussian GARCH(1,1) model for the BTC. The term ”W_E” represents the size of the estimation window, while ”ES cl.” indicates the confidence level for which ES is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ExpectedSeverity tells how severe VaR excaadences are expected to be, whereas ObservedSeverity tells the empirical value. The ”Observations” column indicates the number of observations. ”Violations” denotes the number of violations the model faced, and ”Expected” indicates the number of expected violations. The ”Ratio” represents the proportion between these two quantities and ”Missing” indicates the count of missing observations.

Asset	W_E	ES cl.	ObservedLevel	ExpectedSeverity	ObservedSeverity	Observations	Violations	Expected	Ratio	Missing
BTC	50	0.975	0.9627	1.1928	1.5991	697	26	17.425	1.4921	0
BTC	100	0.975	0.96987	1.1928	1.6196	697	21	17.425	1.2052	0
BTC	150	0.975	0.96844	1.1928	1.4684	697	22	17.425	1.2626	0
BTC	200	0.975	0.96987	1.1928	1.5046	697	21	17.425	1.2052	0
BTC	250	0.975	0.97274	1.1928	1.4618	697	19	17.425	1.0904	0

This table presents the backtesting results of ES using UC. The initial three columns mirror those in the preceding table. The ”UnconditionalTest” column indicates whether the test is rejected or not. The ”p-value” displays the test’s p-value, while the ”TestStatistic” column reveals the actual value of the UC test. In the ”CriticalValue” column, you can find the critical value associated with the UC test. The ”Observations” column denotes the number of observations utilized, and the ”TestLevel” column specifies the confidence level of the UC test.

Asset	W_E	ES cl.	UnconditionalTest	p-value	TestStatistic	CriticalValue	Observations	TestLevel
BTC	50	0.975	Reject	0.00011918	-1.0004	-0.41331	697	0.95
BTC	100	0.975	Reject	0.0076527	-0.63643	-0.41331	697	0.95
BTC	150	0.975	Reject	0.017062	-0.55425	-0.41331	697	0.95
BTC	200	0.975	Reject	0.022011	-0.52026	-0.41331	697	0.95
BTC	250	0.975	Fail to Reject	0.089506	-0.33627	-0.41331	697	0.95

This table illustrates the ES results derived by the Gaussian GARCH(1,1) model for the ETH. The term ”W_E” represents the size of the estimation window, while ”ES cl.” indicates the confidence level for which ES is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ExpectedSeverity tells how severe VaR excaadences are expected to be, whereas ObservedSeverity tells the empirical value. The ”Observations” column indicates the number of observations. ”Violations” denotes the number of violations the model faced, and ”Expected” indicates the number of expected violations. The ”Ratio” represents the proportion between these two quantities and ”Missing” indicates the count of missing observations.

Asset	W_E	ES cl.	ObservedLevel	ExpectedSeverity	ObservedSeverity	Observations	Violations	Expected	Ratio	Missing
ETH	50	0.975	0.95409	1.1928	1.4637	697	32	17.425	1.8364	0
ETH	100	0.975	0.95983	1.1928	1.4149	697	28	17.425	1.6069	0
ETH	150	0.975	0.96557	1.1928	1.3677	697	24	17.425	1.3773	0
ETH	200	0.975	0.967	1.1928	1.4214	697	23	17.425	1.3199	0
ETH	250	0.975	0.97274	1.1928	1.4193	697	19	17.425	1.0904	0

This table presents the backtesting results of ES using UC. The initial three columns mirror those in the preceding table. The ”UnconditionalTest” column indicates whether the test is rejected or not. The ”p-value” displays the test’s p-value, while the ”TestStatistic” column reveals the actual value of the UC test. In the ”CriticalValue” column, you can find the critical value associated with the UC test. The ”Observations” column denotes the number of observations utilized, and the ”TestLevel” column specifies the confidence level of the UC test.

Asset	W_E	ES cl.	UnconditionalTest	p-value	TestStatistic	CriticalValue	Observations	TestLevel
ETH	50	0.975	Reject	0.0001	-1.2535	-0.41331	697	0.95
ETH	100	0.975	Reject	0.00042208	-0.90613	-0.41331	697	0.95
ETH	150	0.975	Reject	0.013417	-0.57929	-0.41331	697	0.95
ETH	200	0.975	Reject	0.014342	-0.57294	-0.41331	697	0.95
ETH	250	0.975	Fail to Reject	0.11741	-0.29746	-0.41331	697	0.95

G Appendix: ES estimates using the Student's t GARCH(1,1) model

This table illustrates the ES results derived by the Student's t GARCH(1,1) model for the S&P500. The term ”W_E” represents the size of the estimation window, while ”ES cl.” indicates the confidence level for which ES is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ExpectedSeverity tells how severe VaR excaadences are expected to be, whereas ObservedSeverity tells the empirical value. The ”Observations” column indicates the number of observations. ”Violations” denotes the number of violations the model faced, and ”Expected” indicates the number of expected violations. The ”Ratio” represents the proportion between these two quantities and ”Missing” indicates the count of missing observations.

Asset	W_E	ES cl.	ObservedLevel	ExpectedSeverity	ObservedSeverity	Observations	Violations	Expected	Ratio	Missing
S&P500	50	0.975	0.98333	1.1928	1.319	480	8	12	0.66667	0
S&P500	100	0.975	0.97292	1.1928	1.1448	480	13	12	1.0833	0
S&P500	150	0.975	0.97708	1.1928	1.1873	480	11	12	0.91667	0
S&P500	200	0.975	0.97292	1.1928	1.1798	480	13	12	1.0833	0
S&P500	250	0.975	0.97708	1.1928	1.2307	480	11	12	0.91667	0

This table presents the backtesting results of ES using UC. The initial three columns mirror those in the preceding table. The ”UnconditionalTest” column indicates whether the test is rejected or not. The ”p-value” displays the test's p-value, while the ”TestStatistic” column reveals the actual value of the UC test. In the ”CriticalValue” column, you can find the critical value associated with the UC test. The ”Observations” column denotes the number of observations utilized, and the ”TestLevel” column specifies the confidence level of the UC test.

Asset	W_E	ES cl.	UnconditionalTest	p-value	TestStatistic	CriticalValue	Observations	TestLevel
S&P500	50	0.975	Fail to Reject	0.5	0.26277	-0.49713	480	0.95
S&P500	100	0.975	Fail to Reject	0.43352	-0.039761	-0.49713	480	0.95
S&P500	150	0.975	Fail to Reject	0.5	0.08757	-0.49713	480	0.95
S&P500	200	0.975	Fail to Reject	0.39407	-0.071562	-0.49713	480	0.95
S&P500	250	0.975	Fail to Reject	0.5	0.054166	-0.49713	480	0.95

This table illustrates the ES results derived by the Student's t GARCH(1,1) model for the BTC. The term "W_E" represents the size of the estimation window, while "ES cl." indicates the confidence level for which ES is being estimated. The "ObservedLevel" column indicates the actual confidence level. The ExpectedSeverity tells how severe VaR excaadences are expected to be, whereas ObservedSeverity tells the empirical value. The "Observations" column indicates the number of observations. "Violations" denotes the number of violations the model faced, and "Expected" indicates the number of expected violations. The "Ratio" represents the proportion between these two quantities and "Missing" indicates the count of missing observations.

Asset	W_E	ES cl.	ObservedLevel	ExpectedSeverity	ObservedSeverity	Observations	Violations	Expected	Ratio	Missing
BTC	50	0.975	0.98278	1.1928	1.3416	697	12	17.425	0.68867	0
BTC	100	0.975	0.98278	1.1928	1.2895	697	12	17.425	0.68867	0
BTC	150	0.975	0.98709	1.1928	1.355	697	9	17.425	0.5165	0
BTC	200	0.975	0.98565	1.1928	1.2728	697	10	17.425	0.57389	0
BTC	250	0.975	0.98852	1.1928	1.3343	697	8	17.425	0.45911	0

This table presents the backtesting results of ES using UC. The initial three columns mirror those in the preceding table. The "UnconditionalTest" column indicates whether the test is rejected or not. The "p-value" displays the test's p-value, while the "TestStatistic" column reveals the actual value of the UC test. In the "CriticalValue" column, you can find the critical value associated with the UC test. The "Observations" column denotes the number of observations utilized, and the "TestLevel" column specifies the confidence level of the UC test.

Asset	W_E	ES cl.	UnconditionalTest	p-value	TestStatistic	CriticalValue	Observations	TestLevel
BTC	50	0.975	Fail to Reject	0.5	0.22544	-0.41331	697	0.95
BTC	100	0.975	Fail to Reject	0.5	0.25552	-0.41331	697	0.95
BTC	150	0.975	Fail to Reject	0.5	0.41325	-0.41331	697	0.95
BTC	200	0.975	Fail to Reject	0.5	0.38763	-0.41331	697	0.95
BTC	250	0.975	Fail to Reject	0.5	0.48643	-0.41331	697	0.95

This table illustrates the ES results derived by the Student's t GARCH(1,1) model for the ETH. The term "W_E" represents the size of the estimation window, while "ES cl." indicates the confidence level for which ES is being estimated. The "ObservedLevel" column indicates the actual confidence level. The ExpectedSeverity tells how severe VaR excaadences are expected to be, whereas ObservedSeverity tells the empirical value. The "Observations" column indicates the number of observations. "Violations" denotes the number of violations the model faced, and ”Expected” indicates the number of expected violations. The ”Ratio” represents the proportion between these two quantities and ”Missing” indicates the count of missing observations.

Asset	W_E	ES cl.	ObservedLevel	ExpectedSeverity	ObservedSeverity	Observations	Violations	Expected	Ratio	Missing
ETH	50	0.975	0.97561	1.1928	1.3876	697	17	17.425	0.97561	0
ETH	100	0.975	0.97991	1.1928	1.4248	697	14	17.425	0.80344	0
ETH	150	0.975	0.97991	1.1928	1.3366	697	14	17.425	0.80344	0
ETH	200	0.975	0.97991	1.1928	1.3776	697	14	17.425	0.80344	0
ETH	250	0.975	0.98422	1.1928	1.3867	697	11	17.425	0.63128	0

This table presents the backtesting results of ES using UC. The initial three columns mirror those in the preceding table. The ”UnconditionalTest” column indicates whether the test is rejected or not. The ”p-value” displays the test’s p-value, while the ”TestStatistic” column reveals the actual value of the UC test. In the ”CriticalValue” column, you can find the critical value associated with the UC test. The ”Observations” column denotes the number of observations utilized, and the ”TestLevel” column specifies the confidence level of the UC test.

Asset	W_E	ES cl.	UnconditionalTest	p-value	TestStatistic	CriticalValue	Observations	TestLevel
ETH	50	0.975	Fail to Reject	0.28406	-0.13495	-0.41331	697	0.95
ETH	100	0.975	Fail to Reject	0.5	0.040267	-0.41331	697	0.95
ETH	150	0.975	Fail to Reject	0.5	0.099677	-0.41331	697	0.95
ETH	200	0.975	Fail to Reject	0.5	0.072052	-0.41331	697	0.95
ETH	250	0.975	Fail to Reject	0.5	0.26611	-0.41331	697	0.95

H Appendix: ES estimates using the Historical Simulation model

This table illustrates the ES results derived by the HS model for the S&P500. The term "W_E" represents the size of the estimation window, while "ES cl." indicates the confidence level for which ES is being estimated. The "ObservedLevel" column indicates the actual confidence level. The ExpectedSeverity tells how severe VaR excaadences are expected to be, whereas ObservedSeverity tells the empirical value. The "Observations" column indicates the number of observations. "Violations" denotes the number of violations the model faced, and "Expected" indicates the number of expected violations. The "Ratio" represents the proportion between these two quantities and "Missing" indicates the count of missing observations.

Asset	W_E	ES cl.	ObservedLevel	ExpectedSeverity	ObservedSeverity	Observations	Violations	Expected	Ratio	Missing
S&P500	50	0.975	0.96042	1.2218	1.2941	480	19	12	1.5833	0
S&P500	100	0.975	0.96667	1.118	1.1915	480	16	12	1.3333	0
S&P500	150	0.975	0.97292	1.1297	1.2548	480	13	12	1.0833	0
S&P500	200	0.975	0.96042	1.2	1.27	480	19	12	1.5833	0
S&P500	250	0.975	0.96667	1.2015	1.3353	480	16	12	1.3333	0

This table presents the backtesting results of ES using UC. The initial three columns mirror those in the preceding table. The "UnconditionalTest" column indicates whether the test is rejected or not. The "p-value" displays the test's p-value, while the "TestStatistic" column reveals the actual value of the UC test. In the "CriticalValue" column, you can find the critical value associated with the UC test. The "Observations" column denotes the number of observations utilized, and the "TestLevel" column specifies the confidence level of the UC test.

Asset	W_E	ES cl.	UnconditionalTest	p-value	TestStatistic	CriticalValue	Observations	TestLevel
S&P500	50	0.975	Reject	0.015872	-0.67937	-0.49713	480	0.95
S&P500	100	0.975	Fail to Reject	0.080665	-0.42456	-0.49713	480	0.95
S&P500	150	0.975	Fail to Reject	0.23632	-0.20511	-0.49713	480	0.95
S&P500	200	0.975	Reject	0.016194	-0.67667	-0.49713	480	0.95
S&P500	250	0.975	Fail to Reject	0.056975	-0.48063	-0.49713	480	0.95

This table illustrates the ES results derived by the HS model for the BTC. The term "W_E" represents the size of the estimation window, while "ES cl." indicates the confidence level for which ES is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ExpectedSeverity tells how severe VaR excaadences are expected to be, whereas ObservedSeverity tells the empirical value. The ”Observations” column indicates the number of observations. ”Violations” denotes the number of violations the model faced, and ”Expected” indicates the number of expected violations. The ”Ratio” represents the proportion between these two quantities and ”Missing” indicates the count of missing observations.

Asset	W_E	ES cl.	ObservedLevel	ExpectedSeverity	ObservedSeverity	Observations	Violations	Expected	Ratio	Missing
BTC	50	0.975	0.96557	1.4094	1.7671	697	24	17.425	1.3773	0
BTC	100	0.975	0.97418	1.4573	1.4425	697	18	17.425	1.033	0
BTC	150	0.975	0.97418	1.4422	1.3315	697	18	17.425	1.033	0
BTC	200	0.975	0.97561	1.4285	1.3898	697	17	17.425	0.97561	0
BTC	250	0.975	0.97418	1.4136	1.3314	697	18	17.425	1.033	0

This table presents the backtesting results of ES using UC. The initial three columns mirror those in the preceding table. The ”UnconditionalTest” column indicates whether the test is rejected or not. The ”p-value” displays the test’s p-value, while the ”TestStatistic” column reveals the actual value of the UC test. In the ”CriticalValue” column, you can find the critical value associated with the UC test. The ”Observations” column denotes the number of observations utilized, and the ”TestLevel” column specifies the confidence level of the UC test.

Asset	W_E	ES cl.	UnconditionalTest	p-value	TestStatistic	CriticalValue	Observations	TestLevel
BTC	50	0.975	Reject	0.0023404	-0.77227	-0.41331	697	0.95
BTC	100	0.975	Fail to Reject	0.33616	-0.10007	-0.41331	697	0.95
BTC	150	0.975	Fail to Reject	0.4689	-0.011194	-0.41331	697	0.95
BTC	200	0.975	Fail to Reject	0.5	0.020865	-0.41331	697	0.95
BTC	250	0.975	Fail to Reject	0.48019	-0.0036359	-0.41331	697	0.95

This table illustrates the ES results derived by the HS model for the BTC. The term ”W_E” represents the size of the estimation window, while ”ES cl.” indicates the confidence level for which ES is being estimated. The ”ObservedLevel” column indicates the actual confidence level. The ExpectedSeverity tells how severe VaR excaadences are expected to be, whereas ObservedSeverity tells the empirical value. The ”Observations” column indicates the number of observations. ”Violations” denotes the number of violations the model faced, and ”Expected” indicates the number of expected violations. The ”Ratio” represents the proportion between these two quantities and ”Missing” indicates the count of missing observations.

Asset	W_E	ES cl.	ObservedLevel	ExpectedSeverity	ObservedSeverity	Observations	Violations	Expected	Ratio	Missing
ETH	50	0.975	0.96557	1.4325	1.5523	697	24	17.425	1.3773	0
ETH	100	0.975	0.97561	1.2753	1.4095	697	17	17.425	0.97561	0
ETH	150	0.975	0.97704	1.2932	1.334	697	16	17.425	0.91822	0
ETH	200	0.975	0.97418	1.4042	1.3675	697	18	17.425	1.033	0
ETH	250	0.975	0.97848	1.4133	1.3923	697	15	17.425	0.86083	0

This table presents the backtesting results of ES using UC. The initial three columns mirror those in the preceding table. The ”UnconditionalTest” column indicates whether the test is rejected or not. The ”p-value” displays the test’s p-value, while the ”TestStatistic” column reveals the actual value of the UC test. In the ”CriticalValue” column, you can find the critical value associated with the UC test. The ”Observations” column denotes the number of observations utilized, and the ”TestLevel” column specifies the confidence level of the UC test.

Asset	W_E	ES cl.	UnconditionalTest	p-value	TestStatistic	CriticalValue	Observations	TestLevel
ETH	50	0.975	Reject	0.016267	-0.55971	-0.41331	697	0.95
ETH	100	0.975	Fail to Reject	0.34127	-0.096649	-0.41331	697	0.95
ETH	150	0.975	Fail to Reject	0.5	0.048011	-0.41331	697	0.95
ETH	200	0.975	Fail to Reject	0.45593	-0.019879	-0.41331	697	0.95
ETH	250	0.975	Fail to Reject	0.5	0.14323	-0.41331	697	0.95