INSTITUTO
UNIVERSITÁRIO
DE LISBOA

An Assessment of Historical Simulation Techniques for VaR

Diogo Filipe Maia Gomes

Master in Finance

Supervisor:
PhD Joaquim Paulo Viegas Ferreira de Carvalho, Assistant Professor, ISCTE-IUL

September 2024

BUSINESS SCHOOL

Department of Finance

An Assessment of Historical Simulation Techniques for VaR

Diogo Filipe Maia Gomes

Master in Finance

Supervisor:
PhD Joaquim Paulo Viegas Ferreira de Carvalho, Assistant Professor, ISCTE-IUL

September 2024

Acknowledgments

I would like to express my sincere gratitude to my supervisor, professor Paulo Viegas de Carvalho, for his dedicated guidance, constant availability, and insights throughout this dissertation.

To my family, especially parents and grandparents, I extend my eternal gratitude for their unconditional support, strength, and constant encouragement. Thank you for believing in me and for providing the conditions necessary for me to achieve my goals.

I would also like to thank my friends. Your friendship, support, and motivation were essential in overcoming the challenges and difficult moments.

To all of you, my sincere thanks.

Resumo

Esta dissertação pretende avaliar a adequabilidade de abordagens não paramétricas para estimar o *Value at Risk*. O estudo incide sobre três técnicas de simulação histórica, com o objetivo de perceber qual obtém a estimativa mais precisa do *VaR*, em condições normais e extremas de mercado (crise da COVID-19). As abordagens estudadas são Simulação Histórica simples, a de BRW (Boudoukh, Richardson e Whitelaw, 1998) e a de Hull e White (1998). As três abordagens baseiam-se todas em dados históricos, mas exploram diferentes formas de estimar a distribuição e, consequentemente, o *VaR*. Para testar a adequabilidade das abordagens mencionadas, o *VaR* é estimado para uma carteira composta por quatro importantes índices de mercado (S&P500, FTSE100, etc.). Depois, são realizados *backtestings* para cada modelo, para identificar a técnica mais confiável para estimar o *VaR*, no período estudado. Este estudo visa fornecer uma visão mais recente sobre a confiança de abordagens não paramétricas na estimação do *VaR.* Os resultados indicam que a abordagem de Hull e White (1998) é a mais precisa comparando com os outros dois métodos, e que a Simulação Histórica simples exibe resultados consistentes. Por outro lado, o método de BRW (1998) aparenta ser o método menos fiável.

Palavras-chave: Simulação Histórica*, BRW, HW, Value at Risk, Backtesting*

Abstract

This dissertation assesses the adequacy of non-parametric approaches to estimate Value at Risk. The study reports on three historical simulation techniques to seek which gives a more accurate estimation for VaR, under normal and extreme market conditions (COVID-19 crisis). The approaches addressed are simple Historical Simulation, BRW (Boudoukh, Richardson and Whitelaw, 1998), and Hull and White's (1998). The three approaches are all based on historical data but explore different ways to estimate the distribution and as a consequence, VaR. To test the adequacy of the mentioned approaches, we estimate VaR for a portfolio composed of four market indices (S&P500, FTSE100, etc.) and then perform backtesting for each, to find out the most reliable one to estimate VaR, in the period studied. This study aims to give a more recent insight into the reliability of non-parametric approaches in VaR estimation. The results indicate that Hull and White's (1998) approach offers higher accuracy when compared to the other two methods and that simple Historical Simulation method shows consistent performance. On contrary, the BRW (1998) method proves to be the least reliable.

Keywords: Historical Simulation, BRW, HW, Value at Risk, Backtesting

Index

1. INTRODUCTION 1

2. LITERATURE REVIEW 3

2.1. VALUE AT RISK AND EXPECTED SHORTFALL 3
2.2. HISTORICAL SIMULATION APPROACH 5
2.3. BRW APPROACH 5
2.4. HULL AND WHITE'S APPROACH 6
2.4.1. GARCH model 7
2.4.2. EWMA model 8
2.5. BACKTESTING 8
2.5.1. Backtesting Value at Risk 9
2.5.2. Backtesting Expected Shortfall 14

3. METHODOLOGY AND DATA 19

4. RESULTS 23

4.1. VAR BEFORE AND AFTER COVID-19 CRISIS 23
4.2. VAR IN MARKET STRESS – COVID-19 CRISIS 29
4.3. BACKTESTING 32

5. CONCLUSION 35

REFERENCES 39

ANNEX A: GRAPHIC REPRESENTATION OF EWMA ESTIMATED VOLATILITIES FOR EACH MARKET INDEX 43

ANNEX B: BACKTESTING RESULTS 47

ANNEX C: PYTHON CODE – HW LAMBDAS ESTIMATE 57

ANNEX D: OTHER RELEVANT TABLES 59

Table Index

Table 1: Indices x weights x amount invested and exchange rates 20
Table 2: Maximum and Minimum Variations, and Volatility per Index and Exchange Rate 21
Table 3: Historical Simulation VaR, before the COVID-19 crisis 23
Table 4: BRW VaR, before the COVID-19 crisis 24
Table 5: HW VaR, before the COVID-19 crisis 25
Table 6: Lambda estimates for each asset in the HW approach 25
Table 7: Historical Simulation VaR, after the COVID-19 crisis 26
Table 8: BRW VaR, after the COVID-19 crisis 27
Table 9: HW VaR, after the COVID-19 crisis 27
Table 10: Total and FX VaR for the portfolio, as a percentage of the notional invested (before and after Covid-19) 27
Table 11: Historical Simulation VaR, Covid-19 crisis 29
Table 12: BRW VaR, the COVID-19 crisis 29
Table 13: HW VaR, the COVID-19 crisis 30
Table 14: Total and FX VaR for the portfolio, as a percentage of the notional invested (before and after the COVID-19) 30
Table 15: Number of violations observed in the 3 periods for the Historical Simulation model 47
Table 16: Number of violations observed in the 3 periods for the BRW model 47
Table 17: Number of violations observed in the 3 periods for the HW model 47
Table 18: Backtesting results (per index and portfolio) for Historical Simulation method, before COVID-19 Crisis 48
Table 19: Backtesting results (per index and portfolio) for Historical Simulation method, COVID-19 Crisis 49
Table 20: Backtesting results (per index and portfolio) for Historical Simulation method, after COVID-19 Crisis 50
Table 21: Backtesting results (per index and portfolio) for BRW method, before COVID-19 Crisis 51
Table 22: Backtesting results (per index and portfolio) for BRW method, COVID-19 Crisis 52
Table 23: Backtesting results (per index and portfolio) for BRW method, after COVID-19 Crisis 53
Table 24: Backtesting results (per index and portfolio) for HW method, before COVID-19 Crisis 54
Table 25: Backtesting results (per index and portfolio) for HW method, COVID-19 Crisis 55
Table 26: Backtesting results (per index and portfolio) for HW method, after COVID-19 Crisis 56
Table 27: Backtesting, number of non-violations per index and portfolio for HS method 59
Table 28: Backtesting, number of non-violations per index and portfolio for BRW method 59
Table 29: Backtesting, number of non-violations per index and portfolio for HW method 59

Table 30: Backtesting, percentage of violations per index and portfolio for HS method 59
Table 31: Backtesting, percentage of violations per index and portfolio for BRW method 59
Table 32: Backtesting, percentage of violations per index and portfolio for HW method.................... 59
Table 33: Backtesting, percentage of non-violations per index and portfolio for HS method............. 60
Table 34: Backtesting, percentage of non-violations per index and portfolio for BRW method.......... 60
Table 35: Backtesting, percentage of non-violations per index and portfolio for HW method 60
Table 36: Parameters used for the independence test – HS, before COVID-19 Crisis 60
Table 37: Parameters used for the independence test – BRW, before COVID-19 Crisis...................... 60
Table 38: Parameters used for the independence test – HW, before COVID-19 Crisis 61
Table 39: Parameters used for the independence test – HS, COVID-19 Crisis 61
Table 40: Parameters used for the independence test – BRW, COVID-19 Crisis 61
Table 41: Parameters used for the independence test – HW, COVID-19 Crisis 61
Table 42: Parameters used for the independence test – HS, after COVID-19 Crisis 62
Table 43: Parameters used for the independence test – BRW, after COVID-19 Crisis......................... 62
Table 44: Parameters used for the independence test – HW, after COVID-19 Crisis 62
Table 45: VaR as percentage of value invested – HS before COVID-19 Crisis 63
Table 46: VaR as percentage of value invested – BRW before COVID-19 Crisis............................... 63
Table 47: VaR as percentage of value invested – HW before COVID-19 Crisis 63
Table 48: VaR as percentage of value invested – HS COVID-19 Crisis.. 63
Table 49: VaR as percentage of value invested – BRW COVID-19 Crisis 63
Table 50: VaR as percentage of value invested – HW COVID-19 Crisis 64
Table 51: VaR as percentage of value invested – HS after COVID-19 Crisis 64
Table 52: VaR as percentage of value invested – BRW after COVID-19 Crisis.................................. 64
Table 53: VaR as percentage of value invested – HW after COVID-19 Crisis 64

Figures Index

Figure 1: Graphic representation of conditional volatilities estimated for S&P500 43
Figure 2: Graphic representation of conditional volatilities estimated for FTSE100 43
Figure 3: Graphic representation of conditional volatilities estimated for EUROSTOXX 50 44
Figure 4: Graphic representation of conditional volatilities estimated for CAC 40 44
Figure 5: Graphic representation of conditional volatilities estimated for USDEUR 45
Figure 6: Graphic representation of conditional volatilities estimated for GBPEUR........................... 45

Glossary

BRW – Boudoukh, Richardson and Whitelaw

ES – Expected Shortfall

EUR – Euro

EWMA - Exponentially Weighted Moving Average

FX VaR – Currency Value at Risk

GBP – British Pound Sterling

GARCH - Generalized Autoregressive Conditional Heteroskedasticity

HS – Historical Simulation

HW – Hull and White

LR – Likelihood Ratio

USD – United States Dollar

VaR – Value at Risk

1. Introduction

Financial institutions are exposed to numerous types of risks, including market risk. Market risk is related to the possibility of losses in the trading book of the financial institution, more accurately, due to variations in interest rates, credit spreads, stock prices, and exchange rates, among other factors for which its value is subject to the conditions of financial markets.

Market risk can be divided into systematic risk and idiosyncratic risk. The first is associated with the risk inherent to the entire market (economic, sociopolitical, etc.). Idiosyncratic risk refers to the intrinsic factors that can impact specific individual securities. To reduce and manage market risk exposure, it is fundamental to measure it. Two common measures used for this purpose are Value at Risk and Expected Shortfall. Both are part of the Basel Committee's regulations for managing market risk, with Value at Risk being the primary measure since its introduction in 1994.

Since its implementation, VaR has been developed in the financial industry and several approaches have been devised to accurately measure risk (parametric, non-parametric, and semiparametric). This dissertation will examine non-parametric methods, specifically assessing their effectiveness under normal and extreme market conditions, such as the financial crisis due to the COVID-19 pandemic. Selecting the appropriate method to evaluate market risk is critical because it directly impacts the amount of regulatory capital that financial institutions must hold to cope with market events. Financial institutions place considerable emphasis on the assessment of capital requirements due to their substantial costs and lack of profitability.

Non-parametric methods operate without the need for assumptions about the statistical distribution inherent in the sample data. Pérignon and Smith (2010), Mehta et al. (2012), and Sharma (2012) show evidence of the models used by large financial institutions and investigate the models they employ for VaR estimation. The authors conclude that non-parametric models are the most used by financial institutions with over half using the Historical Simulation method. Pérignon and Smith (2010) refer to the causes for the popularity of the Historical Simulation method. Firstly, financial institutions have complex positions over several financial instruments, which makes parametric models hard to use in practice. As the complexity of parameter estimation for distributions increases, so does the uncertainty, leading to more significant repercussions on Value-at-risk estimates. In contrast, the Historical Simulation method uses the same return distribution over the years, which allows financial institutions to maintain their capital requirement estimation.

This study aims to assess the efficacy and performance of non-parametric models, in specific Historical Simulation approaches, and their advantages and disadvantages. To implement Historical Simulation, several techniques can be used with the major difference among them being how each models the returns distribution. Historical Simulation (HS) uses a distribution based on past returns, with no further assumptions. Other alternative approaches were introduced by Boudoukh, Richardson and Whitelaw (1998), and Hull and White (1998). The first one, the BRW approach, assigns greater importance to recent historical returns by utilizing an exponentially weighted moving average. Hull and White's (1998) approach adjusts past observations to current market volatility.

To assess the performance of the Historical Simulation approaches, the market risk metric used is Value at Risk. VaR is estimated for a portfolio composed of four market indices (S&P500, FTSE100, EUROSTOXX50, and CAC40). Then, the performance and efficacy of the referred HS approaches are determined through backtesting each.

In this study, the Hybrid approach, or BRW method, often underestimates risk as it depends more on recent observations. On the other hand, the volatility-adjusted model presented by Hull and White (1998) shows more consistent results when it comes to estimating VaR. Moreover, simple Historical Simulation led to reasonable VaR estimates, making it also a quite consistent model. The results suggest that both Hull and White's (1998) approach and simple Historical Simulation are reasonable models to be used by financial institutions when estimating market risk.

This dissertation is organized as follows. The second chapter, the literature review, defines the metrics, the models used, and recent conclusions on the theme. The third chapter describes the methodology and data used. The following chapter presents a detailed interpretation of the VaR estimates and their backtesting results. Lastly, the conclusions are in chapter five.

2. Literature Review

Market Risk is defined as the risk of losses arising from movements in market prices (BIS, 2019). The Basel Committee on Banking Supervision (BIS) outlines two key measures within the framework of market risk regulation: Value at Risk and Expected Shortfall. Subsequent section (2.1) elaborates on both measures based on existing literature. Following this, a literature review is conducted on Historical Simulation, BRW, and HW approaches. Finally, an overview of backtesting, along with relevant literature, is provided.

2.1. Value at Risk and Expected Shortfall

The most widely used risk measure in the financial system is Value at Risk, introduced by J.P. Morgan in the early ’90s (Danielsson and Zhou, 2016). Value at Risk is the standard measure used to quantify market risk and it is defined as the worst expected loss in value of a portfolio of financial instruments, with a predetermined probability (confidence level), for a given time horizon. One of the reasons for the popularity of this method is that it is conceptually simple since it reduces the market risk of a portfolio to the potential loss associated with a given probability (Manganelli and Engle, 2001).

The VaR at a confidence level $\alpha \in (0,1)$, denoted as VaR_α, of a random variable X is defined as (see Nolde and Ziegel, 2017):

$$VaR_\alpha(X) = \inf\{x \mid F_X(x) \geq \alpha\} \tag{1}$$

where F_X is the cumulative distribution function of X. In the expression above, positive values of X are interpreted as losses.

The Bank for International Settlements (2019) specifically calls for VaR_α values for α = 0.99, which we commonly label as the standard Basel VaR level.

An alternative risk measure to Value at Risk within the BIS regulation framework is Expected Shortfall. VaR has been criticized since it does not account for “tail risk” (Du and Escanciano, 2015). ES is strictly related to VaR and complements it. ES is also referred to in its current literature as Conditional Value at Risk (CVaR), Expected Tail Loss (ETL), and Average

Value at Risk (AVaR) and it is interpreted as the average loss value when the α quantile is exceeded. ES of an integrable random variable X at level $\upsilon \in (0,1)$ is given by (see Nolde and Ziegel, 2017):

$$ES_{\upsilon}(X) = \frac{1}{1-\upsilon}\int_{\upsilon}^{1} VaR_{\alpha}(X)\,d\alpha \tag{2}$$

where $VaR_{\alpha}(X)$ denotes Value at Risk at confidence level α and X represents the distribution function. The Bank for International Settlements (2019) suggests $\upsilon = 0.975$ as the standard Basel Expected Shortfall (ES) level and it is expected to produce a comparable level of risk to VaR_{α} with α = 0.99 under the standard normal distribution.

Chen (2018) emphasizes an important mathematical distinction between Value at Risk and Expected Shortfall. VaR, while widely used, lacks subadditivity, an important property for risk measures. In contrast, ES demonstrates coherence, enhancing its reliability as a risk metric. However, ES falls short in being elicitable, a quality necessary for effective backtesting, unlike VaR. Four crucial conditions must be satisfied for a risk measure to be considered coherent. Following Chen (2018), a risk measure ρ applied to γ, γ_1 and γ_2 is coherent if and only if it meets all these conditions:

i. Subadditivity: $\rho(\gamma_1 + \gamma_2) \leq \rho(\gamma_1) + \rho(\gamma_2)$
ii. Monotonicity: $R(\gamma_1) \geq R(\gamma_2) \Rightarrow \rho(\gamma_1) \leq \rho(\gamma_2)$, where R is the return associated with a given portfolio
iii. Linear homogeneity: $\rho(\lambda\gamma) = \lambda\rho(\gamma)$, with $\lambda > 0$
iv. Translation invariance: $\rho(\gamma + c) = \rho(\gamma) - c$

VaR meets all conditions except subadditivity. Failing this property results in an unexpected outcome where a portfolio VaR sum may exceed the sum of the VaRs for each component.

2.2. Historical Simulation approach

The main advantage of the simple Historical Simulation technique is that no assumption is made on the model generating returns. This approach estimates the probability density function (pdf) by the histogram of observed returns, given a sample of T daily returns $(r_1, \ldots, r_T)$. (Ballotta and Fusai, 2017). It purely relies on historical return data to estimate future risk. Implicitly, it assumes that past data offers insights into future possibilities.

Return volatility is somewhat constrained in this approach, which may slow down adjustments to market changes. HS assumes that all returns within the observation window have the same distribution, hence all returns from the entire time series have the same distribution (Žiković and Aktan, 2010). This can be the cause of overestimating risk when the market is calm and underestimating it when there is high volatility. Because it operates under the assumption that the returns' distribution is i.i.d., i.e., that the historical returns are equally likely to occur again, it can lead to clusters of similar events during calm and turbulent market periods. Basu (2009) states that we may even observe a so-called "ghost effect", where significant past losses disproportionally influence our risk estimation. Moreover, the choice of historical data influences the accuracy of our risk assessments. If we look at longer timeframes, the economic landscape might be different from what it is today. On the other hand, shorter timeframes might not capture enough extreme events to paint an accurate picture of current economic conditions and the tail end of the distribution (Alexander, 2009).

2.3. BRW approach

The idea that recent returns are a better representation of the current portfolio's risk is a commonly accepted assumption. Boudoukh, Richardson, and Whitelaw (1998) developed an alternative approach to the simple Historical Simulation. To implement it, weights are attributed to the past returns, to give more influence on recent observations when estimating risk. The weights are computed using a factor – *lambda* - that decays exponentially and assumes a value between 0 and 1. The decay factor expresses the exponential decay in the weight of observation with time, where the closer to 1, the slower the decay rate. Thus, the weights of a return that occurred *i* days ago are given by the following expression:

$$w_i = \frac{\lambda^i(1-\lambda)}{1-\lambda^n} \tag{3}$$

where the sum of all weights is equal to 1 and the λ is the decay factor.

The original authors test this approach and estimate risk using, among others, a decay factor of 0.99, but do not specify a method to estimate $lambda$.

2.4. Hull and White's approach

Hull and White (1998) presented another alternative approach to the simple Historical Simulation. It considers volatility changes over time when estimating the portfolio distribution. To carry out this approach, the historical returns are adjusted using a volatility factor, to make them sensible to current market volatility levels. This factor is given by the relation between the current market volatility level and the volatility on the day of the return. Thus, returns are adjusted to volatility using the following equation:

$$r_t^* = \frac{\sigma^T}{\sigma^t} r^t \tag{4}$$

where r^t represents the return in day *t*, σ^t represents the volatility observed in day *t* and σ^T denote the current market volatility level.

The relationship between r_t^* and r^t becomes more amplified when the current market volatility surpasses the estimated volatility for day *t*. This comparison helps us clarify whether the current market conditions are more or less turbulent than those in the past.

There are essentially two primary methods to adjust volatility, GARCH and EWMA. In 1998, Hull and White's model was adjusted using an EWMA with a decay factor of 0.94. Therefore, this study uses the EWMA approach to adjust volatility. Regardless, both methods are explored in detail over the next two subchapters.

In the method's implementation, volatility is explicitly factored in, making risk measures more responsive to changes in return distributions. However, this also introduces instability in the estimated risk measures, potentially exceeding the maximum loss recorded in the previous time series before adjustment.

According to empirical findings (Mehta et al., 2012), financial institutions employing this method apply a constraint: adjustments are made to returns only if the volatility factor exceeds 1. The fluctuation in the volatility of financial asset returns over time reveals clusters of volatility - periods of either high or low volatility tend to persist from the recent past into the near future. Therefore, accurate volatility forecasts are essential for this method to reflect these patterns in risk measures effectively (Basu, 2011).

2.4.1. GARCH model

In 1986, Bollerslev introduced the GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model. This model was developed specifically to identify clusters of volatility. It operates under the assumption that variance is conditional, meaning it relies on historical squared errors and instantaneous variance. The GARCH model can be written as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-1}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-1}^2 \tag{5}$$

where $\omega > 0$, $\alpha_i \geq 0$, $i = 1, \ldots, p$, $\beta_j \geq 0, j = 1, \ldots, q$ e $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$.

In the realm of market risk assessment, the GARCH (1,1) model stands out as one of the most used versions. Given the frequent data processing involved in evaluating market risk for banks' trading portfolios daily, it is commonly assumed that asset returns are negligible, with $r^t = \varepsilon_t$. For the GARCH (1,1) model to forecast conditional variance at time t (σ_t^2), historical returns (r_{t-1}) and conditional variance σ_{t-1}^2 at the preceding time step are necessary (Hull and White, 1998):

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{6}$$

where ω, α, and β denote, respectively, the mean reversion of the long-term variance, the intensity of the variance reaction to market events, and the variance resistance, that is, the impact that current volatility has on future volatility.

2.4.2. EWMA model

EWMA (Exponentially Weighted Moving Average) model derives from a specific case of the GARCH (1,1). In the model construction, there is no inclusion of mean reversion for long-term variance. This is reflected in $\alpha = 1 - \lambda$ and $\beta = \lambda$. Consequently, the model is non-stationary since $\alpha + \beta = 1$. To predict conditional variance at time t (σ_t^2) using this model, it is necessary to compute the weighted average between historical returns (r_{t-1}) and the forecasted conditional variance σ_{t-1}^2 for the preceding time step.

Therefore, EWMA volatility for a series of returns can be written as:

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)r_{t-1}^2 \tag{7}$$

where σ_t is the volatility at time t, λ is the decay factor and r_t is the return at time t.

This method ensures that more recent returns have a higher impact on the estimated volatility, providing a more responsive measure to changes in market conditions.

2.5. Backtesting

Literature in finance denotes backtesting as either an assessment of the hypothetical historical performance of a trading strategy or an evaluation of risk models using historical data on risk forecast and profit and loss (P&L) realizations (Christoffersen, 2008). More precisely, it tests how well the calculation for a risk measure would have worked in the past. The idea is to test the daily risk measure estimated against the daily realized portfolio's loss from the day analyzed (Christoffersen, 2008). The most commonly used criteria/tests in the literature to backtest VaR are Conditional Coverage, Unconditional Coverage, and Independence Testing. On the other hand, backtesting ES is not so straightforward since it does not have elicitability. Recent research on how to backtest Expected Shortfall is addressed by Chu and Bhattacharyya's (2020).

2.5.1. Backtesting *Value at Risk*

As VaR is the most popular market risk measure for practitioners, the need for methods to backtest it was soon noted by regulators. Shortly after, in 1995 and 1996, Kupiec and Hendricks, respectively, published the first research on backtesting VaR. Backtesting VaR involves comparing actual losses to the predicted VaR to assess the accuracy of the forecasts (Pajhede, 2015). To do so, firstly, we define a "hit sequence" of VaR violations (see Christoffersen, 2008):

$$I_{t+1} = \begin{cases} 1, if\ PL_{t+1} > VaR_{t+1}^{\alpha} \\ 0, if\ PL_{t+1} < VaR_{t+1}^{\alpha} \end{cases} \tag{8}$$

where VaR_{t+1}^{α} is a number constructed on day *t* such that the portfolio losses on day *t+1* will only be larger than the VaR estimated for that day, and PL_{t+1} is profit and loss realization at *t+1*. If the loss incurred on a particular day exceeds the Value at Risk (VaR) threshold predicted for that day, the hit sequence will return a 1 on the subsequent day (*t + 1*). Else, it returns a 0. To perform backtesting to a risk model, one must define a sequence $\{I_{t+1}\}_{t+1}^{T}$ over T days indicating when the past violations or hits occurred (Christoffersen, 2008). Therefore, the hit sequence is essentially a binary time series that shows whether a loss exceeding the *VaR* at time *t*, also known as a violation or hit, occurred (Pajhede, 2015).

Pajhede (2015) mentions that a VaR forecast is valid, if and only if its hit sequence complies with the three criteria pointed out by Christoffersen (1998). The criteria are:

i. *Unconditional coverage criteria:* the probability of a violation occurring, without considering any conditions, must be exactly equal to the coverage rate κ (1-α):

$$H_{UC} : P(I_t = 1) = \kappa \tag{9}$$

ii. *Independence criteria:* the conditional probability of a hit or violation must be constant:

$$H_{Ind} : P(I_t = 1|\mathcal{F}_{t-1}) = P(I_t = 1) \tag{10}$$

iii. *Conditional coverage criteria:* corresponds to the combination of the two criteria above. The probability of a violation or hit must be equal to the coverage rate κ and constant:

$$H_{CC} : P(I_t = 1|\mathcal{F}_{t-1}) = P(I_t = 1) = \kappa \tag{11}$$

Christoffersen (1998) concludes that the hit sequence of a valid *VaR* forecast is a sequence of i.i.d. Bernoulli distributed variables:

$$I_t \sim_{i.i.d.} Bernoulli(\kappa), \qquad t=1,\ldots,T$$

Unconditional coverage test

The unconditional coverage test statistically evaluates whether the observed number of violations over a given period aligns with the predefined confidence level, κ. Specifically, this test assesses whether the observed probability of a violation in the hit sequence, $P(I_t = 1)$, deviates significantly from the predetermined probability of occurrence.

Christoffersen (2008) demonstrates that unconditional coverage can be tested using a likelihood ratio test. The likelihood of a *i.i.d.* Bernoulli(π) hit sequence is expressed as:

$$L(\pi) = \prod_{t+1}^{T} (1-\pi)^{1-I_{t+1}} \pi^{I_{t+1}} = (1-\pi)^{T_0} \pi^{T_1} \tag{12}$$

where T_0 and T_1 represent the observations 0 and 1 in the sample. The observed fraction of violations in the sample gives us an estimation for π:

$$\hat{\pi} = \frac{T_1}{T} \tag{13}$$

And the likelihood of the *i.i.d.* Bernoulli(π) becomes:

$$L(\hat{\pi}) = \left(1 - \frac{T_1}{T}\right)^{T_0} \left(\frac{T_1}{T}\right)^{T_1} \tag{14}$$

The null hypothesis affirms that the VaR coverage rate $\kappa = \pi$. Then the likelihood is:

$$L(\kappa) = \prod_{t+1}^{T} (1-\kappa)^{1-I_{t+1}} \kappa^{I_{t+1}} = (1-\kappa)^{T_0} \kappa^{T_1} \tag{15}$$

To test the unconditional coverage criteria, we can use a likelihood ratio test:

$$LR_{UC} = -2ln\left[\frac{(1-\kappa)^{T_0}\kappa^{T_1}}{\left(1-\frac{T_1}{T}\right)^{T_0}\left(\frac{T_1}{T}\right)^{T_1}}\right] \tag{16}$$

As the number of observations T increases, the test will be distributed as a χ^2 with one degree of freedom.

Conditional coverage and Independence test

The goal of the Conditional Coverage test is to measure the frequency and temporal distribution of occurrences, while also analyzing the independence of these events. Substantial empirical evidence suggests that daily asset returns exhibit time-varying volatility Pritsker (2005). Models that fail to incorporate these dynamics risk slow responsiveness to changing market conditions, leading to temporal clustering of violations in the sequence of hits, as noted by Christoffersen (2008). Consequently, following a violation within a hit sequence, the probability of another violation occurring the subsequent day is elevated.

Christoffersen (2008) elaborates on independence testing and implements it using a likelihood approach. The hit sequence can be expressed as a first-order Markov sequence:

$$\pi_1 = \begin{bmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{bmatrix} \tag{17}$$

The transition probabilities matrix denotes that the probability of tomorrow being a violation is π_{01}, conditional on today not being a violation. The probability of tomorrow being a violation, given that today is a violation is π_{11}. $1-\pi_{01}$ and $1-\pi_{11}$ complement the probabilities explained, respectively. Mathematically, we have that $\pi_{01} = P(I_{t+1} = 1|I_t = 0)$ and $\pi_{11} = P(I_{t+1} = 1|I_t = 1)$.

First-order Markov likelihood function is given by:

$$L(\Pi_1) = (1-\pi_{01})^{T_{00}}\pi_{01}^{T_{01}}(1-\pi_{11})^{T_{10}}\pi_{11}^{T_{11}} \tag{18}$$

where T is the sample of observations, with T_{ij}, *i, j= 0, 1* corresponds to the number of observations with a *j* following a *i.* By calculating the first derivatives concerning π_{01} and π_{11}

and equating these derivatives to zero, it is possible to determine the Maximum Likelihood estimates:

$$\hat{\pi}_{01} = \frac{T_{01}}{T_{00} + T_{01}} \tag{19}$$

and

$$\hat{\pi}_{11} = \frac{T_{11}}{T_{10} + T_{11}} \tag{20}$$

The independence between the sequence elements shows that $\pi_{01} = \pi_{11} = \pi$. To test this independence hypothesis that $\pi_{01} = \pi_{11}$, Christoffersen (2008) elaborates on a likelihood ratio test:

$$LR_{IND} = -2ln\left[\frac{L(\hat{\pi})}{L(\hat{\Pi}_1)}\right] \tag{21}$$

where LR_{IND} is asymptotically χ_1^2, under the null hypothesis that $\hat{\pi}_{01} = \hat{\pi}_{11}$.

The conditional coverage criterium is then obtained by:

$$LR_{CC} = LR_{IND} + LR_C \tag{22}$$

where LR_{CC} converges asymptotically to χ_2^2 and the null hypothesis assumes that $\pi_{01} = \pi_{11} = \pi = \alpha$.

2.5.2. Backtesting Expected Shortfall

Despite not being the objective of this dissertation, it is important to give some context to the recent literature on Expected Shortfall backtesting.

Regulatory authorities have recently given strong importance to Expected Shortfall as a risk measure, and the unavailability of simple tools to backtest this measure is a problem for financial institutions. Since Expected Shortfall does not have elicitability, there was a belief that it could not be backtested. Acerbi and Szekely (2014) presented the possibility of doing unconditional backtesting to ES and developed three distribution-independent and non-parametric tests. The authors also note that elicitability is not key to model testing, instead, it is to model selection. Du and Escanciano (2015) complete the literature on this subject and elaborate on conditional backtests on ES, based on cumulative violations.

Elicitability

In 2011, it was discovered that Expected Shortfall lacks the property of elicitability, which is a feature found in Value at Risk (VaR) (Acerbi and Szekely, 2014). Acerbi and Szekely (2014) show a simple definition of the property. A statistic $\psi(Y)$ of a random variable Y is said to be elicitable if it minimizes the expected value of a scoring function *S:*

$$\psi = \underset{x}{argmin}\mathbb{E}[S(x, Y)] \tag{23}$$

Consider x_t a history of point prediction for the statistics and y_t the realizations of the random variable. To evaluate the forecast model, we can naturally require the mean score to be as low as possible. The mean of the scoring function *S* can be expressed as:

$$\bar{S} = \frac{1}{T}\sum_{t=1}^{T} S(x_t y_t) \tag{24}$$

Acerbi and Szekely (2014) approach

The standard hypothesis testing framework for Expected Shortfall is similar to the Value at Risk framework. The independence of tail events is evaluated separately. The null hypothesis considers that ES is not estimated correctly. The rejection of the null hypothesis shows an underestimation of risk.

Let X_t represent a financial institution's profit-loss at day $t = 1, \dots, T$. X_t is then distributed along a real unknown distribution F_t, predicted by the distribution P_t, conditional to the information used to estimate VaR and ES. Random variables $\vec{X} = \{X_t\}$ are independent but not identically distributed. Additionally, the variability of variables F_t and P_t is not restricted over time. $VaR^F_{\alpha,t}$ and $ES^F_{\alpha,t}$ denote Value at Risk and Expected Shortfall, respectively, when $X \sim F$. ES can be expressed as:

$$ES_{\alpha,t} = -\mathbb{E}\left[X_t \middle| X_t + VaR_{\alpha,t} < 0\right] \tag{25}$$

Testing ES after VaR (TEST 1)

This test is incited by the conditional expectation, from which we derive:

$$\mathbb{E}\left[\frac{X_t}{ES_{\alpha,t}} + 1 \middle| X_t + VaR_{\alpha,t} < 0\right] = 0 \tag{26}$$

Assuming that VaR has been already tested we can separately test the magnitude of the realized exceptions against model predictions. The test statistic is defined as:

$$Z_1(\vec{X}) = \frac{\sum_{t=1}^{T} \frac{X_t I_t}{ES_{\alpha,t}}}{N_T} + 1 \tag{27}$$

where $I_t = (X_t + VaR_{\alpha,t} < 0)$ denotes the indicator function of an α-exception, and $N_T = \sum_{t=1}^{T} I_t > 0$.

The null hypothesis for this test is:

$$H_0: P_t^{[\alpha]} = F_t^{[\alpha]}, \qquad \forall t \tag{28}$$

where $P_t^{[\alpha]}(x) = \min(1, \frac{P_t(x)}{\alpha})$ denotes the distribution tail for $x < -VaR_{\alpha,t}$.
On the other hand, the alternative hypothesis is defined as:

$$H_1: ES_{\alpha,t}^F \geq ES_{\alpha,t}, \qquad \forall t \tag{29}$$

$$H_1: VaR_{\alpha,t}^F = VaR_{\alpha,t}, \qquad \forall t \tag{30}$$

Since, this test is implemented after a VaR test, the predicted VaR_α is still correct under H_1. Under the conditions that $\mathbb{E}_{H_0}[Z_1|N_T > 0] = 0$ and $\mathbb{E}_{H_1}[Z_1|N_T > 0] < 0$, the realized value $Z_1(\vec{x})$ is expected to be zero under the null hypothesis. If contrary, there is evidence of underestimation of risk.

Testing ES directly (TEST 2)

The authors elaborate a second test on the unconditional expectation:

$$ES_{\alpha,t} = -\mathbb{E}\left[\frac{X_t I_t}{\alpha}\right] \tag{31}$$

and the statistic of the test becomes:

$$Z_2(\vec{X}) = \sum_{t=1}^{T} \frac{X_t I_t}{T\alpha ES_{\alpha,t}} + 1 \tag{32}$$

The null hypothesis considers that $P_t^{[\alpha]} = F_t^{[\alpha]},\ \forall t$. The alternative considers that:

$$H_1: ES_{\alpha,t}^F \geq ES_{\alpha,t}, \qquad \forall t \tag{33}$$

$$H_1: VaR_{\alpha,t}^F \geq VaR_{\alpha,t}, \qquad \forall t \tag{34}$$

Under these conditions, and without requirement for the independence of $X_t's$, we have that $\mathbb{E}_{H_0}[Z_2] = 0$ and $\mathbb{E}_{H_1}[Z_2] < 0$.

This test allows us to evaluate the magnitude and frequency of tail events by the relation:

$$Z_2 = 1 - (1 - Z_1)\frac{N_T}{T\alpha} \tag{35}$$

where $\mathbb{E}_{H_0}[N_T] = T\alpha$.

3. Methodology and Data

In this dissertation, simple HS, BRW, and HW models are assessed to find out the most adequate one to estimate VaR.

The main objective of this dissertation is to assess the adequacy and reliability of the approaches mentioned before, under normal and extreme market conditions, such as the COVID-19 Crisis. The literature on this subject shows that the three models have advantages and disadvantages, however, it is important to investigate the techniques available to the present day, to perform a more recent study.

The main difference in these models is the method they use to model the distribution. Additionally, these methods are the most widely used by banks, and since all Historical Simulation techniques are based on past information, the data is easily available. By using Historical Simulation approaches, we assume that the returns are distributed based on the portfolio's past returns rather than distributed normally or another parametric distribution.

To have a broad and diversified study that captures normal market conditions and extreme market events, the time interval for historical data begins on 01/06/2017 and ends on 01/06/2023. This period of data was chosen to have a recent sample that could be used to perform a more updated study on this issue.

To that purpose, VaR is estimated with a 99% confidence level for 3 different periods: before, during, and after the March 2020 – COVID-19 crisis. The 3 periods mentioned are:

- 1st June 2017 to 31st May 2019 (before COVID-19 Crisis)
- 3rd June 2019 to 31st May 2021 (COVID-19 Crisis)
- 1st June 2021 to 1st June 2023 (after COVID-19 Crisis)

The estimation is done using the three approaches mentioned above, regarding the last date of each period. Moreover, this specific time interval allows us to assess the efficacy of the approaches studied under market stress conditions. The implementation of these methodologies requires the use of returns, which can be determined by two models - arithmetic or geometric. The most commonly used in finance is the geometric, i.e., logarithmic returns:

$$r_t = log\left(\frac{P_t}{P_{t-1}}\right) \tag{36}$$

where P_t denotes the asset price at time *t*. Returns in this dissertation are computed as above.

To test the adequacy and reliability of the mentioned models, backtesting VaR is fundamental. Therefore, backtesting is performed for each Historical Simulation approach.

Regarding the data selection, to capture large and different types of risks, the portfolio tested is composed of four major market indices from different parts of the world:

- S&P500
- FTSE100
- Euro Stoxx 50
- CAC 40

Data is captured on a daily frequency. Because the indices are from all over the world, exchange rates are also considered to capture the Total VaR. All indices are converted to EUR. To estimate VaR based on Historical Simulation approaches, the P&L of the portfolio is computed daily. The total amount invested is 1,000,000 EUR, with an investment of 250,000 EUR in each index, which implies an equally weighted portfolio, rebalanced all days. The portfolio is also exposed to currency risk since it has foreign currency indices.

Table 1 describes the composition of the portfolio by name, correspondent weight and amount invested, and the exchange rates used to convert foreign currencies to EUR:

Index	Weight	Amount
S&P 500	25%	250,000 €
FTSE 100	25%	250,000 €
EUROSTOXX 50	25%	250,000 €
CAC 40	25%	250,000 €
EUR/USD	-	-
EUR/GBP	-	-

Table 1: Indices x weights x amount invested and exchange rates

Table 2 shows the volatility, maximum, and minimum returns for each asset, including the exchange rates, in EUR. It is observed that exchange rates are the less volatile assets since they also have smaller maximum and minimum variations. Regarding the indices, EURO STOXX

50 and CAC 40 have registered the minimum variations. On the other hand, S&P 500 and EURO STOXX 50 registered the maximum variations. In the period studied, from 2017 to 2023, all indices noted more volatility than the exchange rates.

As expected, the assets used in this study play an important part in the assessment of the Historical Simulation models. Therefore, having some volatility might help to verify if the VaR estimates are sufficiently accurate.

value in EUR

	S&P500	FTSE100	EUROSTOXX 50	CAC 40	USDEUR	GBPEUR
Maximum Variation	9.6%	7.8%	8.8%	8.1%	2.8%	2.0%
Minimum Variation	-12.5%	-11.7%	-13.2%	-13.1%	-1.8%	-3.7%
Volatility	1.3%	1.1%	1.2%	1.2%	0.5%	0.5%

Table 2: Maximum and Minimum Variations, and Volatility per Index and Exchange Rate

4. Results

The evaluation of the Historical Simulation approaches is executed over three different periods, before, during, and after stress market conditions. Firstly, we show evidence of the results obtained before and after the COVID-19 crisis. Subsequently, the results for the market stress period are presented. VaR is estimated at a 99% confidence level for all methodologies, and for all periods. Finally, we exhibit the results obtained concerning the backtesting.

4.1. VaR before and after COVID-19 Crisis

In this subchapter, an analysis is conducted on the estimates of Value at Risk (VaR) utilizing three distinct methodologies, both before and following the market stress period induced by the COVID-19 crisis.

The table below (Table 3) presents VaR estimates, computed using the Historical Simulation method, before COVID-19. Additionally, Table 3 also shows VaR for the currency factor (FX VaR), if exposed to a foreign currency. The results are shown by index and for the total portfolio.

value in EUR

Historical Simulation	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	7,253	5,304	5,231	4,977	**19,451**
FX VaR	2,445	3,013	-	-	**4,687**

Table 3: Historical Simulation VaR, before the COVID-19 crisis

To implement Boudoukh, Richardson, and Whitelaw's (1998) approach it is required to have a decay factor – *lambda*. The value used by Boudoukh, Richardson, and Whitelaw (1998) is 0.99. Therefore, that is the value used for the decay factor in this study.

Table 4 shows the VaR estimates, determined by the BRW approach, referring to before the market stress period. Similarly to the previous table, it is shown the Total VaR and FX VaR for each index, and the VaR in both components for the total aggregated portfolio.

value in EUR

BRW	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	6,831	6,457	5,315	5,337	**20,416**
FX VaR	2,160	3,085	-	-	**4,687**

Table 4: BRW VaR, before the COVID-19 crisis

Comparing the two methodologies, we can denote that the BRW approach estimates are quite similar to the HS estimates before the market stress period. BRW estimates are higher for most indices and for the total portfolio, which may indicate that recent losses were higher.

Boudoukh, Richardson, and Whitelaw (1998) noted significant evidence of an improvement in the statistical performance of the Hybrid Approach (BRW), in comparison to the Historical Simulation. The results obtained in Table 4 might prove it. Further analysis of backtesting results shall test the estimates obtained.

The Hull and White approach requires the computation of the conditional volatility. In the original paper, Hull and White estimated conditional volatility using an EWMA model with a decay factor of 0.94.

In an attempt to replicate the original model and to simplify the computational process, in this dissertation, conditional volatility is estimated by an EWMA model. To estimate the parameters of the Exponentially Weighted Moving Average (EWMA) model for volatility, we implement an iterative search procedure using the maximum likelihood method. EWMA return volatilities are computed as shown in equation 7.

Then, the loglikelihood function for the EWMA model was maximized:

$$max \sum_{t=1}^{n} \left[-\ln(\sigma_t^2) - \frac{r_t^2}{2\sigma_t^2} \right] \tag{37}$$

The estimated parameter *lambda* varies between 0 and 1.

Consequently, the EWMA model used to estimate the return volatility for each asset can be written as:

$$S\&P500\text{: } \sigma_t^2 = 0.887779\sigma_{t-1}^2 + (1 - 0.887779)r_{t-1}^2$$
$$FTSE100\text{: } \sigma_t^2 = 0.860219\sigma_{t-1}^2 + (1 - 0.860219)r_{t-1}^2$$
$$EUROSTOXX\ 50\text{: } \sigma_t^2 = 0.91372\sigma_{t-1}^2 + (1 - 0.91372)r_{t-1}^2$$
$$CAC\ 40\text{: } \sigma_t^2 = 0.901283\sigma_{t-1}^2 + (1 - 0.901283)r_{t-1}^2$$
$$USDEUR\text{: } \sigma_t^2 = 0.958592\sigma_{t-1}^2 + (1 - 0.958592)r_{t-1}^2$$
$$GBPEUR\text{: } \sigma_t^2 = 0.851849\sigma_{t-1}^2 + (1 - 0.851849)r_{t-1}^2$$

Note that the parameters estimated considered logarithmic returns from 01/06/2017 to 01/06/2023, to have a more robust estimate of the parameters. Hence, the *lambdas* estimated above will be used to estimate HW VaR in the other two time periods.

Table 5 shows evidence of the VaR estimates, obtained using Hull and White's approach. The parameter *lambda* used for each asset in the above equations is also exhibited in Table 6. The VaR estimates for this methodology are presented by index, and as an aggregated portfolio, and the estimates correspond to Total VaR and FX VaR.

value in EUR

HW	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	13,126	6,171	4,893	4,792	**22,325**
FX VaR	4,405	5,667	-	-	**8,203**

Table 5: HW VaR, before the COVID-19 crisis

	Lambda
S&P 500	0.8878
FTSE 100	0.8602
EUROSTOXX	0.9137
CAC 40	0.9013
USDEUR	0.9586
GBPEUR	0.8518

Table 6: Lambda estimates for each asset in the HW approach

We can observe that VaR estimates obtained from the Hull and White approach are higher, both for Total VaR and FX VaR, for the aggregated portfolio and for the index S&P500 and FTSE100, when compared to the Historical Simulation. When compared to the BRW methodology, the estimates are also higher for Total VaR and FX VaR, for the aggregated portfolio, and S&P500. The index S&P500 appears to have the most influence in raising the VaR estimates by the HW approach.

This variant of the Historical Simulation adjusts historical returns by a volatility factor, making it sensitive to volatility changes in the current market. In the Historical Simulation method, returns' volatility is fixed, which may cause an overestimate or underestimate in risk. In this specific case, the VaR estimates adjusted by current volatility changes are higher than the VaR estimates obtained with simple Historical Simulation. It may be explained by an increase in the estimated volatility of recent observations.

Annex A represents graphically the conditional volatility estimated regarding each asset of the portfolio, including exchange rates. All graphics reveal that during the COVID-19 crisis, conditional volatility went to all-time highs. Figure 1 shows that the S&P500 estimated volatility is kept between 0% and 3%, being higher in the period after the Crisis. Figures 2, 3, and 4, referring to the other three indices, reflect that the estimated volatility increases after the market stress period. We can also note that the FTSE 100 index has a recent spike in conditional volatility. This may increase VaR estimates since volatility-adjusted returns would be higher than historical returns.

The following table (Table 7) shows evidence of the VaR estimates obtained with the simple Historical Simulation after the COVID-19 crisis, in a similar format as the above tables.

value in EUR

Historical Simulation	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	8,193	7,780	8,905	8,841	**28,301**
FX VaR	3,294	3,065	-	-	**4,698**

Table 7: Historical Simulation VaR, after the COVID-19 crisis

Table 8 exhibits the estimated values for VaR, for the period after the market stress, utilizing the Hybrid approach (BRW).

value in EUR

BRW	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	7,961	7,838	8,935	9,121	**28,586**
FX VaR	3,842	2,770	-	-	**5,115**

Table 8: BRW VaR, after the COVID-19 crisis

Lastly, Table 9 denotes the VaR estimates regarding the volatility-adjusted VaR (HW), also for the period after the COVID-19 crisis.

value in EUR

HW	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	16,094	11,261	11,691	10,521	**45,198**
FX VaR	5,205	5,714	-	-	**7,393**

Table 9: HW VaR, after the COVID-19 crisis

As observed in Tables 7, 8, and 9 the same effect happened in the period after the COVID-19 crisis. BRW and HS methodologies estimate very similar risk values. However, the HW approach estimates are higher individually for all assets and the aggregated portfolio, in comparison with the other two approaches. A reason for this may be the fact that conditional volatility estimated is higher in the period after COVID-19 and consequently, estimates using HW methodology are higher.

To give another perspective of the results, Table 10 exhibits the estimated Total and FX VaR regarding the portfolio, as a percentage of the notional invested (1 million EUR), for the three methodologies in the periods before and after COVID-19.

	Before Covid-19 Crisis			After Covid-19 Crisis		
	HS	**BRW**	**HW**	**HS**	**BRW**	**HW**
Total VaR	1.95%	2.04%	2.23%	2.83%	2.86%	4.52%
FX VaR	0.47%	0.47%	0.82%	0.47%	0.51%	0.74%

Table 10: Total and FX VaR for the portfolio, as a percentage of the notional invested (before and after COVID-19)

The currency VaR has almost no weight in relative terms to the investment. However, Total VaR stays around 2% to 4% of the total investment. Also note that the estimates obtained by HS and BRW approaches give very similar results, while HW methodology gives higher estimates both on Total VaR and Currency VaR. Tables 45 to 53 from Annex D, exhibit VaR as a percentage of the value invested for all models and all periods studied.

4.2. VaR in market stress – COVID-19 Crisis

VaR estimates are even more relevant during market stress periods, such as the COVID-19 crisis. Thus, models must accurately predict possible losses in investors' portfolios. In this subchapter, an analysis of the VaR estimates is carried out to compare the results obtained by the three methodologies, in market stress conditions.

Firstly, an exhibit of the results estimated with a simple Historical Simulation approach, over market stress conditions is shown. Table 11 shows the Total VaR and FX VaR for each index and the portfolio of indices.

value in EUR

Historical Simulation	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	11,345	10,704	10,620	10,950	**42,987**
FX VaR	2,533	3,145	-	-	**4,320**

Table 11: Historical Simulation VaR, COVID-19 crisis

Implementation of the BRW approach in this period uses the same *lambda* - 0.99 - as before and after the COVID-19 crisis periods. Table 12 shows the VaR estimates, determined by the BRW approach, referring to the market stress period. Similarly to the previous table, it shows the Total VaR and FX VaR for each index and the VaR in both components for the total aggregated portfolio.

value in EUR

BRW	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	7,085	8,589	9,317	8,594	**30,882**
FX VaR	2,157	2,823	-	-	**4,024**

Table 12: BRW VaR, the COVID-19 crisis

Regarding the HW methodology, the results were estimated using the same process as the other periods analyzed. The *lambda* parameters used to perform the EWMA for each asset correspond to the ones shown in Table 6, as mentioned in the subchapter before. The results achieved using the HW approach, which adjusts return volatilities to current market volatility are presented in Table 13, below. It is possible to observe Total and FX VaR by index and as an aggregated portfolio.

value in EUR

HW	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	57,187	53,134	43,732	58,614	**212,677**
FX VaR	3,198	6,186	-	-	**6,898**

Table 13: HW VaR, the COVID-19 crisis

Comparing the three methods, it is evident that the HW approach estimates VaR higher than the other two methodologies. A market stress scenario tends to impact market volatility, and hence, VaR estimates by this model. On the other hand, the BRW model seems to underestimate risk when compared to Historical Simulation, which might be evidence that there were no consistent major losses in recent observations, to the date of 01/06/2021. Because HS takes equal weights for all observations it gives more emphasis to the actual period of the Crisis than BRW. The hybrid approach (BRW) gives more weight to recent observations, and therefore less weight to the observations in 2020 which causes this method to underestimate risk.

Table 14 presents VaR as a percentage of the total value invested. The HW approach Total VaR estimates are much higher than HS and BRW ones. HS and BRW indicate similar percentages both for Total VaR. As for FX VaR, the values are close between the three methods. Compared to Table 10, there is an increase in percentage values as expected during a market stress period.

	Covid-19 Crisis		
	HS	**BRW**	**HW**
Total VaR	4.30%	3.09%	21.27%
FX VaR	0.43%	0.40%	0.69%

Table 14: Total and FX VaR for the portfolio, as a percentage of the notional invested (before and after the COVID-19)

To give a more consistent response to the accuracy of the models, the next subchapter shows evidence of the results from the backtesting.

4.3. Backtesting

To assess the effectiveness of non-parametric methods in responding to market events, we must evaluate the consistency and suitability of the risk estimates produced by risk evaluation methods. This will be done using backtesting. The test period is divided into three: before, after, and during the market stress period. The period before the crisis contains 522 daily comparisons. In the period after the crisis, there are 523 daily comparisons, and in the market stress period, 521 daily comparisons are done. To consider diversification effects and, therefore, evaluate the three VaR methodologies more effectively, backtestings are performed for each index and also for the total aggregated portfolio, described in Table 1.

Backtesting of the Value at Risk (VaR) is conducted at a 99% confidence level, and it gives us the number of days where losses exceed the estimated VaR. Annex B provides tables detailing the number of exceptions observed and the test statistics, when estimating VaR using the three distinct methodologies, both during the market stress period – COVID-19 crisis – and the period before and after it.

Overall, as observed in Tables 15, 16, and 17 of the mentioned Annex, there is evidence of a higher number of violations associated with the HS and BRW models. On the other hand, the HW approach has the lowest number of violations. If we compare the models in the period after the COVID-19 crisis, the HW model continues to be the one with the lowest number of violations, whereas the HS and BRW models have almost the same number of violations. However, before the COVID-19 crisis, the results observed allow us to say that the three models have almost the same number of violations, where the HW approach has the lowest number and HS has the highest one. During the market stress period, i.e., the COVID-19 crisis, there is a clear difference in the number of violations. HW model registered zero violations in this period, whereas HS and BRW revealed the highest number of violations. However, the number of violations noticed with HS is less than half of the number of violations registered by BRW, which indicates that BRW has more difficulties estimating risk in market stress periods. The number of violations per methodology allows us to have an idea of the effectiveness of each model.

To evaluate statistically whether the number of violations observed is consistent with the VaR confidence level, it is applied Kupiec's test. The results of this test are presented by index and for the aggregated portfolio in Annex B. Tables 18, 19, and 20 exhibit the results for the Historical Simulation approach. Tables 21, 22, and 23 show the results for the BRW model. Lastly, Tables 24, 25, and 26 show evidence of the results obtained for the HW model.

Comparing the results of the test before the COVID-19 crisis, there is evidence that all methods are fit to evaluate VaR for the indices and the total portfolio. The null hypothesis for all models was not rejected once and therefore, one can confirm that the number of violations observed is consistent with the VaR confidence level of 99%. As previously mentioned, the number of violations observed per model, in this period, is the lowest when compared to the other two periods (Tables 15, 16, and 17) which helps to confirm the results obtained.

On the other hand, an analysis of the results obtained for the market stress period – COVID-19 crisis – shows, that the most consistent models are the HS and the HW, with no rejection of the null hypothesis. The BRW approach verifies the rejection of the null hypothesis for all indices and the total portfolio. Thus, in this period, the number of violations for HS and HW models is coherent with the 99% confidence level assumed for the VaR.

After the COVID-19 crisis, there is evidence of rejection of the null hypothesis regarding FTSE100 using the HW model. Furthermore, HS and BRW models show evidence of rejection of the null hypothesis for S&P500, FTSE100, and the total portfolio. Therefore, the HW model is the most consistent between the number of violations observed and the 99% confidence level used to estimate VaR.

Boudoukh, Richardson, and Whitelaw (1998) originally observed that their approach was more consistent than the simple Historical Simulation. However, the results noted in this dissertation indicate that the approach often underestimates risk in market stress periods. In this case, giving weight to observations might not be ideal since the peak of the Crisis was observed almost one year before the estimation date. Because Historical Simulation gives equal weight to all observations, it gives more weight to an event that occurred one year ago.

Before the COVID-19 crisis, all models performed well with close results among them. After the COVID-19 crisis, there is evidence that the HW approach evaluates risk more appropriately. Hull and White's (1998) model also appears to be the fittest to evaluate risk, during the market stress period. Overall, adjusting current volatility to past observations comes across as a good practice to estimate risk both in normal and extreme market conditions (COVID-19 crisis).

Christoffersen (2008) demonstrated the importance of independence testing. The goal of the independence test is to find out if the violations observed are independent amongst themselves. As mentioned in the literature review, once to test the conditional coverage, one has to perform the independence test. Therefore, both tests are performed for all models in all periods studied, by index, and for the total portfolio. The results are exhibited in Annex B, where Tables 18, 19, and 20 show the results associated with the simple Historical Simulation, Tables 21, 22, and 23 indicate the results for the Hybrid approach (BRW), and Tables 24, 25 and 26 denote the results obtained for the volatility adjusted approach (HW). The null hypothesis is never rejected for the three models, which indicates that the number of violations that occurred did not depend on the previous day.

To completely backtest VaR, one has to perform the conditional coverage test. The test allows us to evaluate the unconditional coverage test and the independence test simultaneously. In other words, it evaluates simultaneously if the number of violations observed is coherent with the confidence level and if the violations observed are independent among themselves. The results corroborate the ones of the unconditional coverage test. That is, before the COVID-19 Crisis we do not reject the null hypothesis for any model, in the market stress period the null hypothesis is rejected in the BRW model for all indices and total portfolio, and after the COVID-19 crisis, the null hypothesis is rejected for S&P500, FTSE100 and total portfolio regarding HS and BRW models, and it is also rejected for FTSE100 using HW approach.

To sum up, generally, adjusting volatility using Hull and White (1998) model demonstrates to be a good fit for the market indices and the total portfolio studied, in all periods. Thus, one can conclude that the VaR estimates produced by this model are more reliable when compared to HS and BRW models in the period studied.

5. Conclusion

This dissertation explores the effectiveness of non-parametric models for estimating market risk, both under normal market conditions and under market stress conditions (COVID-19 Crisis). This study focuses on models frequently employed by financial institutions to assess their market risk exposure. To this end, we defined a portfolio composed of four indices, two of which are in foreign currency. The assets used exhibit some volatility, which is an important factor in this study, as it allows us to assess if Value at Risk estimates reflect market dynamics.

To better understand the effectiveness of the non-parametric models studied, our analysis uses data from 1st June 2017 to 1st June 2023. Value at Risk is estimated for three distinct periods, of which two periods with normal market conditions and one with extreme market conditions, corresponding to the COVID-19 crisis.

The results observed in normal market conditions indicate that the BRW method tends to underestimate risk when recent observations register profit. This happens because recent profits or losses are given more weight, which may imply that this method is highly dependent on recent performance. During the market stress period, this method often underestimated VaR. In 1998, Boudoukh, Richardson, and Whitelaw indicated that it performed better than simple Historical Simulation, however, in this study it performs similarly during normal market conditions periods and registers a worse performance when it comes to the market stress period.

Simple Historical Simulation makes reasonable estimates of market risk, before and during the COVID-19 crisis. After that, this model shows difficulties in estimating risk, for some assets.

The Hull and White (1998) model presents the most consistent results when compared to the other two methodologies. This model shows that adjusting past observations to current market volatility might be a good fit to estimate market risk. Despite positive backtesting results, the simple HS method underestimates risk more often than the HW method. Thus, one could say that the HW approach assesses market risk better than HS.

In summary, with this set of data and portfolio composition, the volatility-adjusted model introduced by Hull and White (1998) is the most appropriate one to estimate market risk through VaR. Simple Historical Simulation demonstrates solid results, however, tends to underestimate risk more often than the HW method. The worst fit appears to be the Hybrid approach (BRW), which tends to underestimate risk more often than the other two methods, especially during market stress periods such as the COVID-19 crisis.

VaR estimates by non-parametric models are inherently dependent on the specific dataset selected, which can constrain the results and potentially lead to an underestimation or overestimation of market risk. Consequently, the choice of the observation window is a critical and somewhat arbitrary parameter that can significantly influence the results. In this study, each observation period is composed by two years of historical data.

Future research on the effectiveness of non-parametric models should consider a way to estimate the decay factor for the BRW approach. Additionally, exploring other ways to estimate the parameters used in the HW approach, such as using a GARCH model, could be of interest. This might lead to different outcomes in what comes to the evaluation of the models.

References

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity, 308 – 313. Accessed in: *https://public.econ.duke.edu/~boller/Published_Papers/joe_86.pdf*

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *The J. of Derivatives.*

Boudoukh, J., Richardson, M. & Whitelaw, R. F. (1998). *The Best of Both Worlds: A Hybrid Approach to Calculating Value at Risk,* 1 - 6. Accessed in: *https://papers.ssrn.com/sol3/papers.cfm?abstract_id=51420*

Hull, J., White, A. (1998). Incorporating Volatility Updating Into The Historical Simulation Method For Value At Risk. *Journal of Risk. 5 – 9.* https://www-2.rotman.utoronto.ca/~hull/downloadablepublications/VaR_VolUpdating.pdf

Christoffersen, P. F. (1998). Evaluating Interval Forecasts. International economic review, 841-862. https://www.jstor.org/stable/2527341?origin=crossref

Manganelli, S. & Engle, R. F. (2001). Value at Risk models in Finance. *European Central Bank*, 5. Accessed in: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=356220

Pritsker, M. (2005). The hidden dangers of historical simulation. *Journal of Banking & Finance (30)*, 562-565. Acessed in: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=278438

Christoffersen, P. (2008). Backtesting. *University of Toronto - Rotman School of Management; Copenhagen Business School; Aarhus University*, 1 - 8. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2044825

Alexander, C. (2009). Market Risk Analysis - Value at Risk Models (Vol. 4). *John Wiley & Sons*

Basu, S. (2009). Stress testing for Market Risk: A comparison of VaR method, 11. Accessed in: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1370965

Žiković, S. & Aktan, B. (2010). Decay Factor Optimisation in Time Weighted Simulation - Evaluating VAR Performance, 2-3. *University of Rijeka - Faculty of Economics; University of Bahrain*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1622704

Pérignon, C., & Smith, D. R. (2010). The level and quality of Value-at-Risk disclosure by commercial banks, 363 – 367. Accessed in: https://www.sciencedirect.com/science/article/pii/S0378426609001940?via%3Dihub

Basu, S. (2011). Comparing simulation models for market risk stress testing. *European journal of operational research*, 336. Accessed in: https://doi.org/10.1016/j.ejor.2011.02.023

Sharma, M. (2012). *The Historical Simulation Method for Value-at-Risk: A Research Based Evaluation of the Industry Favourite.* 1 - 3. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2042594

Mehta, A., Neukirchen, M., Pfetsch, S., & Poppensieker, T. (2012). Managing market risk : Today and tomorrow. *McKinsey & Company McKinsey Working Papers on Risk*. Accessed in: https://www.binhphuongnguyen.com/wp-content/uploads/2019/09/The-McKinsey-on-Risk-Evaluation-System.pdf

Acerbi, C., & Szekely, B. (2014). Backtesting expected shortfall. *Risk, 27(11), 1 - 5*. Accessed in: https://www.msci.com/documents/10199/22aa9922-f874-4060-b77a-0f0e267a489b

Pajhede, T. (2015). Backtesting Value-at-Risk: A Generalized Markov Framework, 2 – 4. Accessed in: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2693504

Du, Z., Escanciano J. (2015). Backtesting Expected Shortfall: Accounting for Tail Risk, 3 – 5. Acessed in: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2548544

Danielsson, J., Zhou, C. (2016). Why Risk Is So Hard to Measure. *De Nederlandsche Bank Working Paper No. 494*, 2. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2597563

Ballotta, L., Fusai, G. (2017). A Gentle Introduction to Value at Risk, 36. Accessed in: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2942138

Nolde, N., Ziegel J. (2017). Elicitability And Backtesting: Perspectives For Banking Regulation, 1835. Accessed in:
https://web.archive.org/web/20200221181826id_/https://boris.unibe.ch/108729/1/AOAS1041.pdf

Chen, J. (2018). On Exactitude in Financial Regulation: Value-at-Risk, Expected Shortfall, and Expectiles. *Risks*. Accessed in: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3136278

Bank for International Settlements *BIS* (2019). Minimum capital requirements for market risk. *Bank for International Settlements*, 10, 89, 94 https://www.bis.org/bcbs/publ/d457.pdf

Chu, S-W. & Bhattacharyya, R. (2020). Expected Shortfall – An Alternative Risk Measure to Value-at-Risk.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3514508

Annex A: Graphic representation of EWMA estimated volatilities for each market index

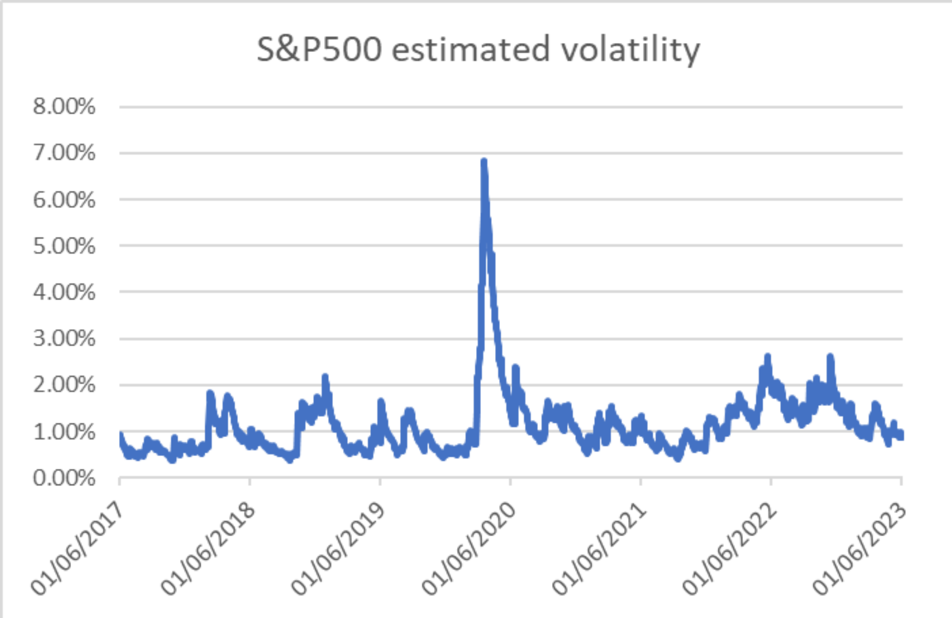

Figure 1: Graphic representation of conditional volatilities estimated for S&P500

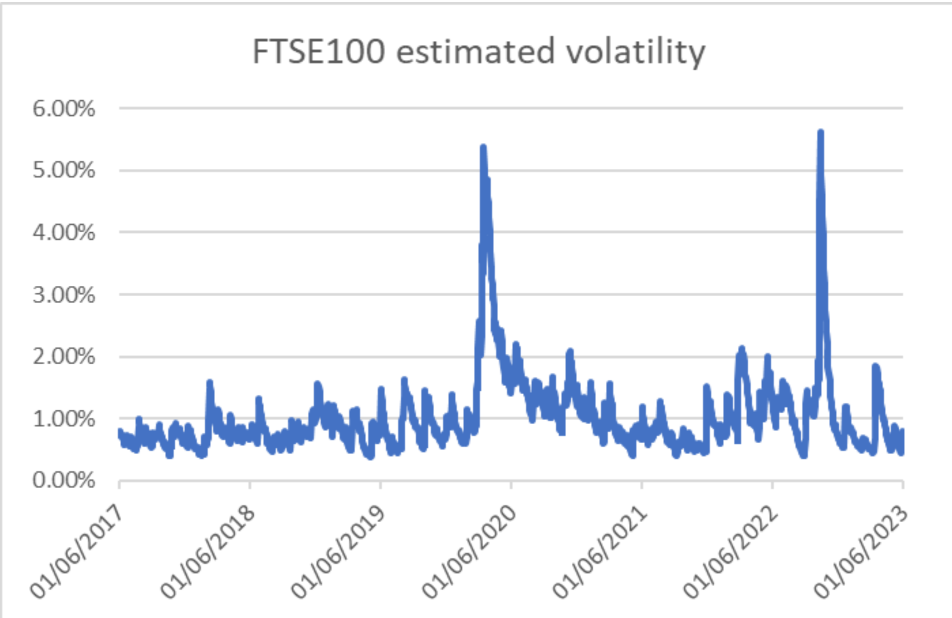

Figure 2: Graphic representation of conditional volatilities estimated for FTSE100

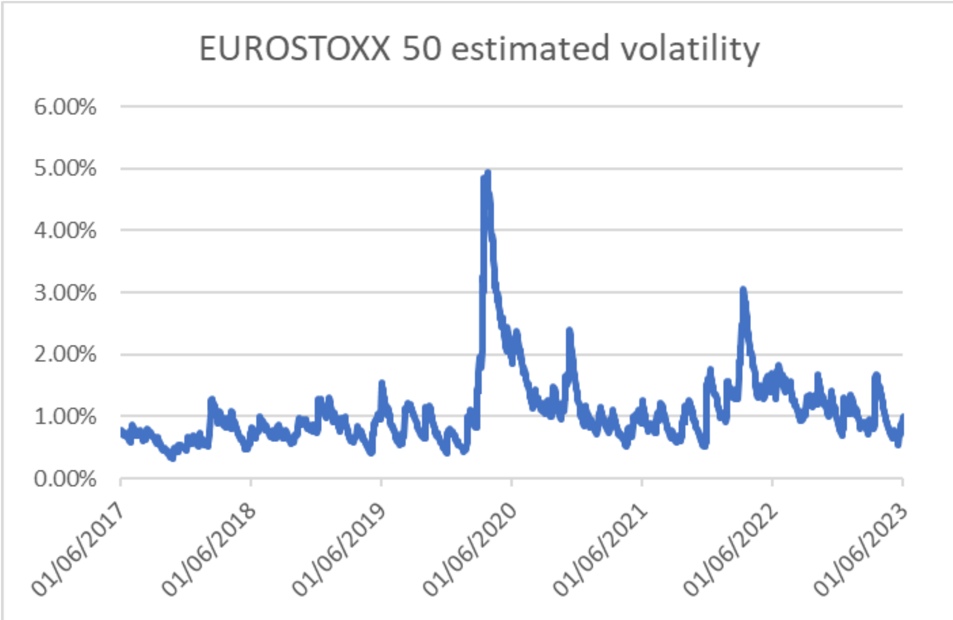

Figure 3: Graphic representation of conditional volatilities estimated for EUROSTOXX 50

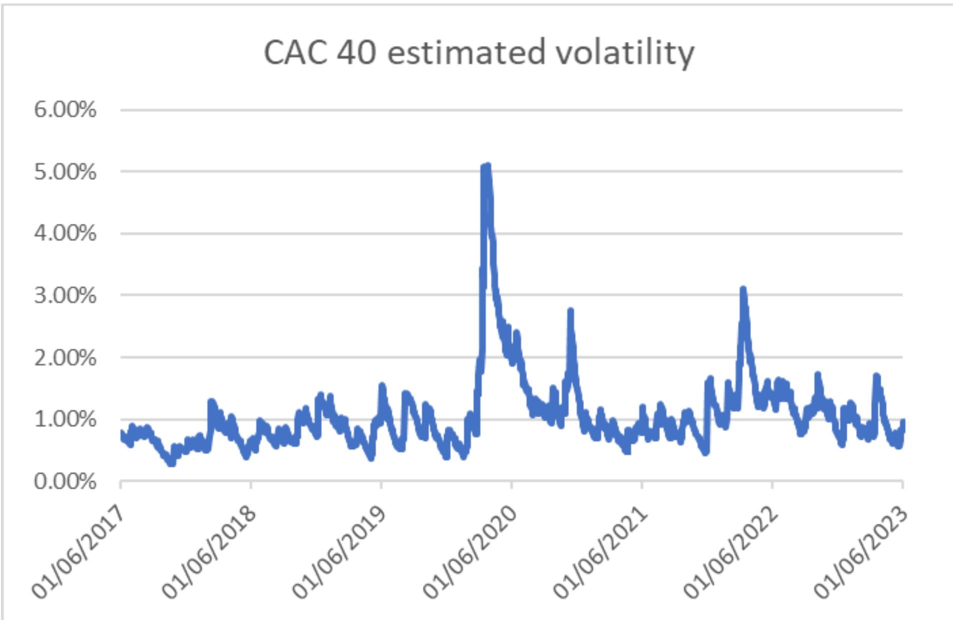

Figure 4: Graphic representation of conditional volatilities estimated for CAC 40

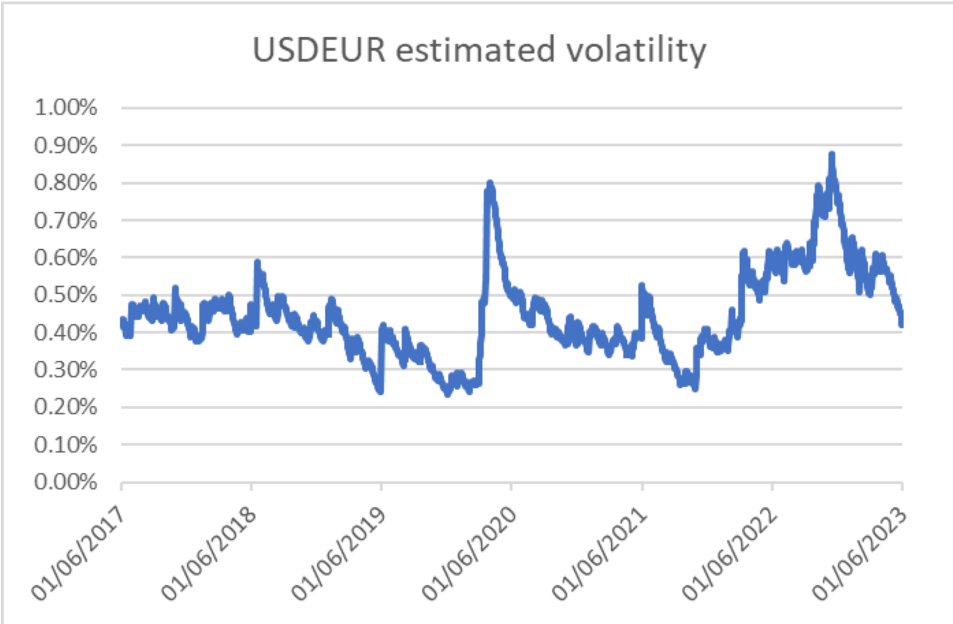

Figure 5: Graphic representation of conditional volatilities estimated for USDEUR

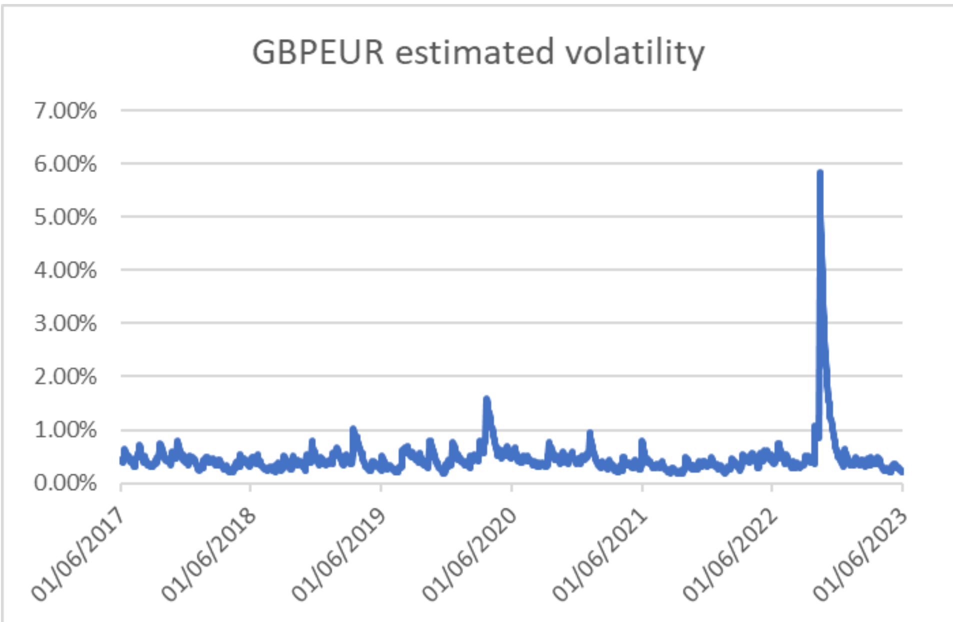

Figure 6: Graphic representation of conditional volatilities estimated for GBPEUR

Annex B: Backtesting Results

Historical Simulation No. Violations	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Before Covid-19 Crisis	6	6	6	6	**6**
Covid-19 Crisis	6	6	6	6	**6**
After Covid-19 Crisis	18	20	2	2	**15**

Table 15: Number of *violations* observed in the 3 periods for the Historical Simulation model

BRW No. Violations	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Before Covid-19 Crisis	7	3	5	3	**4**
Covid-19 Crisis	20	14	13	14	**15**
After Covid-19 Crisis	19	20	2	2	**15**

Table 16: Number of *violations* observed in the 3 periods for the BRW model

HW No. Violations	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Before Covid-19 Crisis	0	3	8	9	**1**
Covid-19 Crisis	0	0	0	0	**0**
After Covid-19 Crisis	7	16	0	0	**6**

Table 17: Number of *violations* observed in the 3 periods for the HW model

Before Covid-19 Crisis					
Historical Simulation	**S&P 500**	**FTSE 100**	**EUROSTOXX 50**	**CAC 40**	**TOTAL**
No. violations	*6*	*6*	*6*	*6*	*6*
LR_{uc}	0.11232	0.11232	0.11232	0.11232	0.11232
p-value	*0.73752*	*0.73752*	*0.73752*	*0.73752*	*0.73752*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{ind}	0.13954	0.13954	0.13954	0.13954	3.79685
p-value	*0.70874*	*0.70874*	*0.70874*	*0.70874*	*0.05135*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{cc}	0.25186	0.25186	0.25186	0.25186	3.90917
p-value	*0.61577*	*0.61577*	*0.61577*	*0.61577*	*0.04802*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***

Table 18: Backtesting results (per index and portfolio) for Historical Simulation method, before COVID-19 Crisis

Covid-19 Crisis					
Historical Simulation	**S&P 500**	**FTSE 100**	**EUROSTOXX 50**	**CAC 40**	**TOTAL**
No. violations	*6*	*6*	*6*	*6*	*6*
LR_{uc}	0.11537	0.11537	0.11537	0.11537	0.11537
p-value	*0.73412*	*0.73412*	*0.73412*	*0.73412*	*0.73412*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{ind}	0.13981	3.79320	0.13981	0.13981	0.13981
p-value	*0.70847*	*0.05146*	*0.70847*	*0.70847*	*0.70847*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{cc}	0.25517	3.90857	0.25517	0.25517	0.25517
p-value	*0.61345*	*0.04804*	*0.61345*	*0.61345*	*0.61345*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***

Table 19: Backtesting results (per index and portfolio) for Historical Simulation method, COVID-19 Crisis

After Covid-19 Crisis					
Historical Simulation	**S&P 500**	**FTSE 100**	**EUROSTOXX 50**	**CAC 40**	**TOTAL**
No. violations	*18*	*20*	*2*	*2*	*15*
LR_{uc}	19.27215	24.53824	2.63505	2.63505	12.25469
p-value	*0.00001*	*0.00000*	*0.10453*	*0.10453*	*0.00046*
Result	***Reject***	***Reject***	***Do not reject***	***Do not reject***	***Reject***
LR_{ind}	1.28344	1.75051	0.01536	0.01536	1.00606
p-value	*0.25726*	*0.18581*	*0.90138*	*0.90138*	*0.31585*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{cc}	20.55559	26.28874	2.65041	2.65041	13.26075
p-value	*0.00001*	*0.00000*	*0.10352*	*0.10352*	*0.00027*
Result	***Reject***	***Reject***	***Do not reject***	***Do not reject***	***Reject***

Table 20: Backtesting results (per index and portfolio) for Historical Simulation method, after COVID-19 Crisis

Before Covid-19 Crisis					
BRW	**S&P 500**	**FTSE 100**	**EUROSTOXX 50**	**CAC 40**	**TOTAL**
No. violations	*7*	*3*	*5*	*3*	*4*
LR_{uc}	0.55392	1.12621	0.00950	1.12621	0.31325
p-value	*0.45672*	*0.28858*	*0.92236*	*0.28858*	*0.57569*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{ind}	3.16997	0.03468	0.09671	0.03468	0.06178
p-value	*0.07500*	*0.85226*	*0.75581*	*0.85226*	*0.80371*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{cc}	3.72389	1.16089	0.10621	1.16089	0.37503
p-value	*0.05364*	*0.28128*	*0.74450*	*0.28128*	*0.54028*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***

Table 21: Backtesting results (per index and portfolio) for BRW method, before COVID-19 Crisis

Covid-19 Crisis					
BRW	**S&P 500**	**FTSE 100**	**EUROSTOXX 50**	**CAC 40**	**TOTAL**
No. violations	*20*	*14*	*13*	*14*	*15*
LR_{uc}	24.65430	10.24803	8.31186	10.24803	12.33112
p-value	*0.00000*	*0.00137*	*0.00394*	*0.00137*	*0.00045*
Result	***Reject***	***Reject***	***Reject***	***Reject***	***Reject***
LR_{ind}	1.52978	3.84431	0.97471	0.76648	3.35031
p-value	*0.21615*	*0.04992*	*0.32351*	*0.38131*	*0.06719*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{cc}	26.18408	14.09234	9.28657	11.01450	15.68143
p-value	*0.00000*	*0.00017*	*0.00231*	*0.00090*	*0.00007*
Result	***Reject***	***Reject***	***Reject***	***Reject***	***Reject***

Table 22: Backtesting results (per index and portfolio) for BRW method, COVID-19 Crisis

After Covid-19 Crisis					
BRW	**S&P 500**	**FTSE 100**	**EUROSTOXX 50**	**CAC 40**	**TOTAL**
No. violations	*19*	*20*	*2*	*2*	*15*
LR_{uc}	21.85055	24.53824	2.63505	2.63505	12.25469
p-value	*0.00000*	*0.00000*	*0.10453*	*0.10453*	*0.00046*
Result	***Reject***	***Reject***	***Do not reject***	***Do not reject***	***Reject***
LR_{ind}	1.43288	1.75051	0.01536	0.01536	1.00606
p-value	*0.23130*	*0.18581*	*0.90138*	*0.90138*	*0.31585*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{cc}	23.28343	26.28874	2.65041	2.65041	13.26075
p-value	*0.00000*	*0.00000*	*0.10352*	*0.10352*	*0.00027*
Result	***Reject***	***Reject***	***Do not reject***	***Do not reject***	***Reject***

Table 23: Backtesting results (per index and portfolio) for BRW method, after COVID-19 Crisis

Before Covid-19 Crisis					
HW	**S&P 500**	**FTSE 100**	**EUROSTOXX 50**	**CAC 40**	**TOTAL**
No. violations	*0*	*3*	*8*	*9*	*1*
LR_{uc}	0.00000	1.12621	1.28609	2.27281	5.16937
p-value	*1.00000*	*0.28858*	*0.25677*	*0.13166*	*0.02299*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{ind}	0.00000	0.03468	0.24904	2.20499	0.00384
p-value	*1.00000*	*0.85226*	*0.61775*	*0.13757*	*0.95060*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{cc}	0.00000	1.16089	1.53513	4.47779	5.17321
p-value	*1.00000*	*0.28128*	*0.21534*	*0.03434*	*0.02294*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***

Table 24: Backtesting results (per index and portfolio) for HW method, before COVID-19 Crisis

Covid-19 Crisis					
HW	**S&P 500**	**FTSE 100**	**EUROSTOXX 50**	**CAC 40**	**TOTAL**
No. violations	*0*	*0*	*0*	*0*	*0*
LR_{uc}	0.00000	0.00000	0.00000	0.00000	0.00000
p-value	*1.00000*	*1.00000*	*1.00000*	*1.00000*	*1.00000*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{ind}	1.52978	3.84431	0.97471	0.76648	3.35031
p-value	*0.21615*	*0.04992*	*0.32351*	*0.38131*	*0.06719*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{cc}	1.52978	3.84431	0.97471	0.76648	3.35031
p-value	*0.21615*	*0.04992*	*0.32351*	*0.38131*	*0.06719*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***

Table 25: Backtesting results (per index and portfolio) for HW method, COVID-19 Crisis

After Covid-19 Crisis					
HW	**S&P 500**	**FTSE 100**	**EUROSTOXX 50**	**CAC 40**	**TOTAL**
No. violations	*7*	*16*	*0*	*0*	*6*
LR_{uc}	0.54704	14.46727	0.00000	0.00000	0.10932
p-value	*0.45953*	*0.00014*	*1.00000*	*1.00000*	*0.74092*
Result	***Do not reject***	***Reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{ind}	0.18993	1.13801	0.00000	0.00000	0.13927
p-value	*0.66298*	*0.28607*	*1.00000*	*1.00000*	*0.70901*
Result	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***	***Do not reject***
LR_{cc}	0.73697	15.60528	0.00000	0.00000	0.24859
p-value	*0.39063*	*0.00008*	*1.00000*	*1.00000*	*0.61807*
Result	***Do not reject***	***Reject***	***Do not reject***	***Do not reject***	***Do not reject***

Table 26: Backtesting results (per index and portfolio) for HW method, after COVID-19 Crisis

Annex C: Python Code – *HW lambdas estimate*

```
# Import libraries
import pandas as pd
import numpy as np
from scipy.optimize import minimize_scalar

# Data
file_path = "/Users/diogogomes/Desktop/TESE/returns.xlsx"
returns_df = pd.read_excel(file_path, sheet_name='Sheet1') # read the excel
file
indices = returns_df.columns[1:]  # Indices names
returns_matrix = returns_df.iloc[:, 1:].to_numpy()  # First column is the
date


# Compute EWMA volatility
def calcular_volatilidade_ewma(returns, lambda_):
    n = len(returns)
    volatilities = np.zeros(n)
    volatilities[0] = np.std(returns)  # Start with the standard deviation
of returns

    for t in range(1, n):
        volatilities[t] = np.sqrt(lambda_ * volatilities[t - 1] ** 2 + (1 -
lambda_) * returns[t - 1] ** 2)

    return volatilities


# LogLikelihood function for EWMA
def ewma_log_likelihood(lambda_, returns):
    volatilities = calcular_volatilidade_ewma(returns, lambda_)
    n = len(returns)
    log_likelihood = -np.sum(np.log(volatilities) + (returns ** 2 / (2 *
volatilities ** 2)))
    return -log_likelihood  # Maximize the loglikelihood function
```

```python
# Estimate lambdas between 0 and 1
def estimar_lambda(returns):
    result = minimize_scalar(ewma_log_likelihood, bounds=(0, 0.99),
args=(returns,), method='bounded')
    return result.x


# Compute lambda for each return series
lambdas = np.apply_along_axis(estimar_lambda, 0, returns_matrix)

# Serie with indices names and respective lambdas
lambdas_series = pd.Series(lambdas, index=indices)

print(lambdas_series)
```

Annex D: Other relevant tables

Historical Simulation No. Non-Violations	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Before Covid-19 Crisis	516	516	516	516	**516**
Covid-19 Crisis	515	515	515	515	**515**
After Covid-19 Crisis	505	503	521	521	**508**

Table 27: Backtesting, number of non-violations per index and portfolio for HS method

BRW No. Non-Violations	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Before Covid-19 Crisis	515	519	517	519	**518**
Covid-19 Crisis	501	507	508	507	**506**
After Covid-19 Crisis	504	503	521	521	**508**

Table 28: Backtesting, number of non-violations per index and portfolio for BRW method

HW No. Non-Violations	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Before Covid-19 Crisis	522	519	514	513	**521**
Covid-19 Crisis	521	521	521	521	**521**
After Covid-19 Crisis	516	507	523	523	**517**

Table 29: Backtesting, number of non-violations per index and portfolio for HW method

Historical Simulation Violations (%)	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Before Covid-19 Crisis	1.15%	1.15%	1.15%	1.15%	1.15%
Covid-19 Crisis	1.15%	1.15%	1.15%	1.15%	1.15%
After Covid-19 Crisis	3.44%	3.82%	0.38%	0.38%	2.87%

Table 30: Backtesting, percentage of violations per index and portfolio for HS method

BRW Violations (%)	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Before Covid-19 Crisis	1.34%	0.57%	0.96%	0.57%	0.77%
Covid-19 Crisis	3.84%	2.69%	2.50%	2.69%	2.88%
After Covid-19 Crisis	3.63%	3.82%	0.38%	0.38%	2.87%

Table 31: Backtesting, percentage of violations per index and portfolio for BRW method

HW Violations (%)	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Before Covid-19 Crisis	0.00%	0.57%	1.53%	1.72%	0.19%
Covid-19 Crisis	0.00%	0.00%	0.00%	0.00%	0.00%
After Covid-19 Crisis	1.34%	3.06%	0.00%	0.00%	1.15%

Table 32: Backtesting, percentage of violations per index and portfolio for HW method

Historical Simulation Non-Violations (%)	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Before Covid-19 Crisis	98.85%	98.85%	98.85%	98.85%	98.85%
Covid-19 Crisis	98.85%	98.85%	98.85%	98.85%	98.85%
After Covid-19 Crisis	96.56%	96.18%	99.62%	99.62%	97.13%

Table 33: Backtesting, percentage of non-violations per index and portfolio for HS method

BRW Non-Violations (%)	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Before Covid-19 Crisis	98.66%	99.43%	99.04%	99.43%	99.23%
Covid-19 Crisis	96.16%	97.31%	97.50%	97.31%	97.12%
After Covid-19 Crisis	96.37%	96.18%	99.62%	99.62%	97.13%

Table 34: Backtesting, percentage of non-violations per index and portfolio for BRW method

HW Non-Violations (%)	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Before Covid-19 Crisis	100.00%	99.43%	98.47%	98.28%	99.81%
Covid-19 Crisis	100.00%	100.00%	100.00%	100.00%	100.00%
After Covid-19 Crisis	98.66%	96.94%	100.00%	100.00%	98.85%

Table 35: Backtesting, percentage of non-violations per index and portfolio for HW method

HS - before Covid-19 crisis	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
π	0.01149	0.01149	0.01149	0.01149	0.01149
π0	0.01163	0.01163	0.01163	0.01163	0.00969
π1	0.00000	0.00000	0.00000	0.00000	0.16667
Ln	0.13954	0.13954	0.13954	0.13954	3.79685
p-value	0.708740708	0.708740708	0.708740708	0.708740708	0.0513491
Null hypothesis	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*

Table 36: Parameters used for the independence test – HS, before COVID-19 Crisis

BRW - before Covid-19 crisis	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
π	0.01341	0.00575	0.00958	0.00575	0.00766
π0	0.01165	0.00578	0.00967	0.00578	0.00772
π1	0.14286	0.00000	0.00000	0.00000	0.00000
Ln	3.16997	0.03468	0.09671	0.03468	0.06178
p-value	0.075003809	0.852263008	0.755810026	0.852263008	0.803709554
Null hypothesis	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*

Table 37: Parameters used for the independence test – BRW, before COVID-19 Crisis

HW - before Covid-19 crisis	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
π	0.00000	0.00575	0.01533	0.01724	0.00192
π0	0.00000	0.00578	0.01556	0.01559	0.00192
π1	0.00000	0.00000	0.00000	0.11111	0.00000
Ln	0.00000	0.03468	0.24904	2.20499	0.00384
p-value	1.00000000	0.852263008	0.617753768	0.137565181	0.950596407
Null hypothesis	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*

Table 38: Parameters used for the independence test – HW, before COVID-19 Crisis

HS - Covid-19 crisis	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
π	0.01152	0.01152	0.01152	0.01152	0.01152
π0	0.01165	0.00971	0.01165	0.01165	0.01165
π1	0.00000	0.16667	0.00000	0.00000	0.00000
Ln	0.13981	3.79320	0.13981	0.13981	0.13981
p-value	0.708470979	0.051461196	0.708470979	0.708470979	0.708470979
Null hypothesis	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*

Table 39: Parameters used for the independence test – HS, COVID-19 Crisis

BRW - Covid-19 crisis	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
π	0.03839	0.02687	0.02495	0.02687	0.02879
π0	0.03593	0.02367	0.02362	0.02564	0.02569
π1	0.10000	0.14286	0.07692	0.07143	0.13333
Ln	1.52978	3.84431	0.97471	0.76648	3.35031
p-value	0.216145801	0.049915037	0.323507343	0.381309649	0.067192207
Null hypothesis	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*

Table 40: Parameters used for the independence test – BRW, COVID-19 Crisis

HW - Covid-19 crisis	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
π	0.00000	0.00000	0.00000	0.00000	0.00000
π0	0.00000	0.00000	0.00000	0.00000	0.00000
π1	0.00000	0.00000	0.00000	0.00000	0.00000
Ln	0.00000	0.00000	0.00000	0.00000	0.00000
p-value	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000
Null hypothesis	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*

Table 41: Parameters used for the independence test – HW, COVID-19 Crisis

HS - after Covid-19 crisis	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
π	0.03442	0.04000	0.00382	0.00382	0.03048
π0	0.03564	0.04167	0.00384	0.00384	0.03143
π1	0.00000	0.00000	0.00000	0.00000	0.00000
Ln	1.28344	1.75051	0.01536	0.01536	1.00606
p-value	0.257260375	0.185813044	0.901381949	0.901381949	0.315848685
Null hypothesis	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*

Table 42: Parameters used for the independence test – HS, after COVID-19 Crisis

BRW - after Covid-19 crisis	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
π	0.03633	0.04000	0.00382	0.00382	0.03048
π0	0.03770	0.04167	0.00384	0.00384	0.03143
π1	0.00000	0.00000	0.00000	0.00000	0.00000
Ln	1.43288	1.75051	0.01536	0.01536	1.00606
p-value	0.231295128	0.185813044	0.901381949	0.901381949	0.315848685
Null hypothesis	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*

Table 43: Parameters used for the independence test – BRW, after COVID-19 Crisis

HW - after Covid-19 crisis	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
π	0.01338	0.03238	0.00000	0.00000	0.01147
π0	0.01357	0.03346	0.00000	0.00000	0.01161
π1	0.00000	0.00000	0.00000	0.00000	0.00000
Ln	0.18993	1.13801	0.00000	0.00000	0.13927
p-value	0.662976295	0.286073698	1.0000000	1.0000000	0.709009691
Null hypothesis	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*	*Do not reject*

Table 44: Parameters used for the independence test – HW, after COVID-19 Crisis

value in EUR

Historical Simulation	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	0.73%	0.53%	0.52%	0.50%	1.95%
FX VaR	0.24%	0.30%			0.47%

Table 45: VaR as percentage of value invested – HS before COVID-19 Crisis

value in EUR

BRW	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	0.68%	0.65%	0.53%	0.53%	2.04%
FX VaR	0.22%	0.31%			0.47%

Table 46: VaR as percentage of value invested – BRW before COVID-19 Crisis

value in EUR

HW	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	1.31%	0.62%	0.49%	0.48%	2.23%
FX VaR	0.44%	0.57%			0.82%

Table 47: VaR as percentage of value invested – HW before COVID-19 Crisis

value in EUR

Historical Simulation	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	1.13%	1.07%	1.06%	1.10%	4.30%
FX VaR	0.25%	0.31%			0.43%

Table 48: VaR as percentage of value invested – HS COVID-19 Crisis

value in EUR

BRW	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	0.71%	0.86%	0.93%	0.86%	3.09%
FX VaR	0.22%	0.28%			0.40%

Table 49: VaR as percentage of value invested – BRW COVID-19 Crisis

value in EUR

HW	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	5.72%	5.31%	4.37%	5.86%	21.27%
FX VaR	0.32%	0.62%			0.69%

Table 50: VaR as percentage of value invested – HW COVID-19 Crisis

value in EUR

Historical Simulation	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	0.82%	0.78%	0.89%	0.88%	2.83%
FX VaR	0.33%	0.31%			0.47%

Table 51: VaR as percentage of value invested – HS after COVID-19 Crisis

value in EUR

BRW	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	0.80%	0.78%	0.89%	0.91%	2.86%
FX VaR	0.38%	0.28%			0.51%

Table 52: VaR as percentage of value invested – BRW after COVID-19 Crisis

value in EUR

HW	S&P 500	FTSE 100	EUROSTOXX 50	CAC 40	TOTAL
Total VaR	1.61%	1.13%	1.17%	1.05%	4.52%
FX VaR	0.52%	0.57%			0.74%

Table 53: VaR as percentage of value invested – HW after COVID-19 Crisis