UNIVERSITY OF TURKU
Turku School of Economics

Risk-Adjusted Performance of Volatility Targeting in the STOXX Europe 600

Evidence on Regime Dependence and Implementation Sensitivity, 2010–2025

Accounting and Finance
Bachelor's thesis

Author:
Tomi Kemppainen

Supervisor:
M.Sc. Ville Kukkonen

10.4.2026
Turku

Student’s statement regarding the use of Artificial Intelligence (AI) for preparing and/or writing this thesis:

☐ **I have not used any AI-based tools.**

☒ **I have used AI-based tools**. Their use is documented in the Appendix. The AI tools were used in a way that complies with academic integrity guidelines.

Bachelor's thesis

Subject: Accounting and Finance
Author: Tomi Kemppainen
Title: Risk-Adjusted Performance of Volatility Targeting in the STOXX Europe 600: Evidence on Regime Dependence and Implementation Sensitivity, 2010–2025
Supervisor: M.Sc. Ville Kukkonen
Number of pages: 37 pages + appendices 3 pages
Date: 10.4.2026

Abstract

This thesis studies whether a simple volatility targeting strategy can improve risk-adjusted returns in the STOXX Europe 600 Index relative to a static buy-and-hold benchmark. The motivation behind this thesis is the idea that market volatility is time-varying and at least partly predictable, which may allow investors to adjust market exposure in response to changing risk conditions. While earlier literature has documented the benefits of volatility management, the evidence is dominated by U.S. markets and remains more limited for broad European equities, particularly in the 2022–2025 environment.

The empirical analysis uses daily data from the STOXX Europe 600 Net Return Index over the period 2010–2025. The strategy uses a fixed 15% target volatility, a 21-day estimation window, and monthly rebalancing with a maximum leverage of 1.5x. The analysis covers three sub-periods, each of which reflects a distinct volatility regime. The periods are the pre-period (2010–2019), the crisis period (2020–2021), and the post-2022 period (2022–2025). In addition to standard performance measures, the thesis examines higher moments of the return distribution, whipsaw risk, parameter sensitivity, and cross-market robustness using the S&P 500 and OMX Helsinki Gross Index.

The results are mixed. Volatility targeting performs most favourably during the crisis period, where it improves the Sharpe ratio slightly and more clearly reduces maximum drawdown and excess kurtosis relative to buy-and-hold. Outside this period, the benefits are weaker and less consistent. In the pre-period, the strategy does not improve performance, while in the post-2022 period, it slightly increases CAGR but weakens risk-adjusted performance and slightly increases the maximum drawdown. The sensitivity analysis further shows that performance depends heavily on the volatility-estimation window, with a 63-day window performing considerably better in turbulent periods. The cross-market robustness analysis further shows that the strategy performs considerably better when applied to the S&P 500, whereas the OMX Helsinki Gross Index produces mixed results.

Overall, these findings suggest that volatility targeting is not a universally superior alternative to a buy-and-hold strategy in the STOXX Europe 600. Its usefulness appears conditional on market regime and implementation design rather than stable across all market environments.

Keywords: volatility targeting, realized volatility, risk-adjusted performance, STOXX Europe 600

Kandidaatintutkielma

Oppiaine: Laskentatoimi ja rahoitus
Tekijä: Tomi Kemppainen
Otsikko: Risk-Adjusted Performance of Volatility Targeting in the STOXX Europe 600: Evidence on Regime Dependence and Implementation Sensitivity, 2010–2025
Ohjaaja: KTM Ville Kukkonen
Sivumäärä: 37 sivua + liitteet 3 sivua
Päivämäärä: 10.4.2026

Tiivistelmä

Tässä kandidaatintutkielmassa selvitetään, voiko yksinkertainen volatiliteettitavoitteinen sijoitusstrategia parantaa riskikorjattua tuottoa STOXX Europe 600 -indeksissä verrattuna staattiseen osta ja pidä -strategiaan. Tutkielman taustalla on ajatus siitä, että markkinavolatiliteetti vaihtelee ajan myötä ja on ainakin osittain ennustettavissa, mikä voi mahdollistaa markkina-altistuksen sopeuttamisen muuttuvien riskiolosuhteiden mukaan. Vaikka aiempi kirjallisuus onkin osoittanut volatiliteettitavoitteisten strategioiden hyötyjä, näyttö painottuu Yhdysvaltojen markkinoille, kun taas laaja-alaisia Euroopan osakemarkkinoita koskeva näyttö on yhä vähäistä, erityisesti vuosien 2022–2025 aikana.

Empiirinen analyysi perustuu STOXX Europe 600 Net Return -indeksin päivittäiseen aineistoon vuosilta 2010–2025. Strategiassa käytetään kiinteää 15% tavoitevolatiliteettia, 21 päivän estimointi-ikkunaa sekä kuukausittaista salkun uudelleenpainotusta, jossa enimmäisvipu on 1,5x. Analyysi kattaa kolme ala-ajanjaksoa, joista jokainen edustaa erilaista volatiliteettiympäristöä. Nämä ajanjaksot ovat kriisiä edeltävä jakso (2010–2019), kriisijakso (2020–2021) ja vuoden 2022 jälkeinen jakso (2022–2025). Tavanomaisten suorituskykymittareiden lisäksi tutkielmassa tarkastellaan tuottojakauman korkeampia momentteja, whipsaw-riskiä, parametriherkkyyttä sekä tulosten luotettavuutta S&P 500- ja OMX Helsinki Gross -indeksien avulla.

Tutkimustulokset ovat ristiriitaisia. Volatiliteettitavoitteinen strategia toimii parhaiten kriisijaksolla, jolloin se parantaa hieman Sharpe-lukua ja pienentää selvemmin salkun maksimiarvonalenemaa sekä tuottojakauman huipukkuutta suhteessa osta ja pidä -strategiaan. Tämän aikajakson ulkopuoliset hyödyt ovat heikompia ja epäjohdonmukaisempia. Kriisiä edeltävänä jaksona strategia ei paranna suorituskykyä, kun taas vuoden 2022 jälkeisenä jaksona se nostaa hieman salkun keskimääräistä vuotuista kasvuvauhtia, mutta laskee samalla riskikorjattua tuottoa ja kasvattaa hieman salkun maksimiarvonalenemaa. Herkkyysanalyysi osoittaa lisäksi, että tulokset riippuvat voimakkaasti volatiliteetin estimointi-ikkunasta, sillä 63 päivän ikkuna toimii selvästi paremmin levottomassa markkinaympäristössä. Markkinoiden välisessä vertailuanalyysissä strategia toimii myös selvästi paremmin S&P 500 -indeksissä, kun taas OMX Helsinki Gross -indeksissä tulokset ovat vaihtelevampia.

Kaiken kaikkiaan tulokset viittaavat siihen, että volatiliteettitavoitteinen strategia STOXX Europe 600 -indeksissä ei ole yleisesti ottaen parempi vaihtoehto kuin osta ja pidä -strategia. Sen hyödyllisyys näyttää riippuvan pikemminkin markkinaregiimistä ja strategian toteutustavasta sen sijaan, että se toimisi yhtä hyvin kaikissa markkinaolosuhteissa.

Avainsanat: volatiliteettitavoittelu, realisoitunut volatiliteetti, riskikorjattu suorituskyky, STOXX Europe 600

TABLE OF CONTENTS

FIGURES

TABLES

1 Introduction

1.1 Background and Motivation

Market risk does not stay constant over time. An allocation that seems appropriate for a given level of risk may later become too aggressive or too conservative as market volatility changes. (Cont 2001; Moreira & Muir 2017.) Consequently, portfolio management is not only a question of asset selection, but also a question of whether risk exposure should stay fixed when the market's overall risk environment is unstable.

This question is especially relevant because financial market volatility is not random in a purely static sense. Previous research has established that volatility tends to cluster, meaning that periods of severe market turbulence are often followed by more turbulence, just as calm periods tend to persist over time. (Mandelbrot 1963; Cont 2001.) Moreover, market volatility has proven to be significantly more predictable than actual asset returns (Poon & Granger 2003). This creates the basic intuition for volatility targeting: if risk changes over time in an even slightly predictable way, investors may be able to improve portfolio performance by adjusting exposure in response to recent volatility rather than by maintaining a fixed allocation.

The academic literature has provided both supportive and cautious evidence on this idea. Moreira and Muir (2017) show that volatility-managed portfolios can improve risk-adjusted performance, and Harvey et al. (2018) further find that volatility targeting can also reduce the severity of tail events across a broad range of assets. However, later studies have taken a more conditional view, suggesting that these benefits are neither uniform across all markets nor robust to all implementation choices (Bongaerts et al. 2020; Cederburg et al. 2020). Based on these findings, the relevant question is no longer simply whether volatility targeting can work, but under which conditions it works, and how sensitive the results are to the design of the strategy itself.

This thesis is motivated by that narrower and more practical question. Although volatility targeting has been studied extensively, many of the most influential findings rely either on U.S. equity data or on broader cross-asset comparison (Moreira & Muir 2017; Harvey et al. 2018; Bongaerts et al. 2020). By contrast, evidence from broad European equity indices remains more limited, especially in the recent 2022–2025 environment. This period provides a useful setting for analysis because it combines monetary tightening, geopolitical disruption, inflationary pressures, and episodes of rapid market recovery (ECB 2025). Taken together, these factors create a clear motivation to assess whether the documented benefits of volatility targeting remain robust in European equity markets, and to provide new evidence on how much the strategy's overall success depends on the prevailing volatility regime.

1.2 Purpose of the Thesis and Research Questions

This thesis investigates whether a simple volatility-targeted strategy can outperform a traditional buy-and-hold approach in the STOXX Europe 600 in terms of risk-adjusted performance. It also evaluates whether the effectiveness of the strategy depends on the prevailing market environment and implementation choices. To examine this, the empirical analysis uses daily market data ranging from 2010 to 2025, divided into three sub-periods: the pre-period (2010–2019), the crisis period (2020–2021), and the post-2022 period (2022–2025).

This thesis is guided by one main research question and three supporting sub-questions. The main question is as follows:

- Does a volatility-targeting strategy improve risk-adjusted performance in the STOXX Europe 600 relative to a buy-and-hold benchmark?

To further examine volatility targeting and risk-adjusted performance, this thesis also includes three sub-questions, which are:

1. Does the effectiveness of the strategy vary across different market regimes, particularly between the pre-period, the crisis period, and the post-2022 environment?
2. Does volatility targeting reduce downside risk and tail heaviness in the return distribution?
3. How sensitive are the results to implementation choices such as the volatility-estimation window, target volatility, and market index?

1.3 Structure

The thesis is structured as follows. The second chapter introduces the required theoretical framework, which is needed to understand the concepts of this study. It will present the relevant theories about volatility, volatility clustering, and predictability, how the volatility-targeting system works, and the potential limitations of applying it in practice. The third chapter will then go over the methodology used in this thesis. It will explain the data used in this study, volatility estimation, strategy construction, and the robustness and sensitivity studies applied to the study. The fourth chapter presents the results of the study. Finally, the last chapter discusses and evaluates results and makes conclusions based on these.

2 Theoretical Background and Related Literature

2.1 Risk-Adjusted Returns and Portfolio Optimization

Introduced by Markowitz (1952), Modern Portfolio Theory proposes that rational investors systematically seek to maximize expected returns for the amount of risk they take. The theory explains the relationship between risk and return, showing that investors can build an optimal asset allocation, known as the efficient frontier, through strategic diversification. A crucial assumption of Modern Portfolio Theory is that investors are generally risk-averse and subsequently seek higher expected returns with lower return variance, which is treated as a measure of risk (Hull 2018, 5). Thereby, an investor who can choose between two portfolios with the same expected return will always prefer the one with lower volatility.

In this framework, portfolio performance cannot be evaluated based on return alone. A portfolio that generates high returns can be suboptimal if the underlying variance is disproportionately high. To address this, Sharpe (1966) introduced the Sharpe ratio, which measures excess returns per unit of total risk. A higher value points to a better risk-return trade-off, meaning that the underlying assets have been allocated more efficiently (Sharpe 1966). For this reason, the Sharpe ratio serves as the primary performance metric in this thesis.

This framework also raises a broader question: can investors improve risk-adjusted performance through dynamic risk management rather than simply accepting the risk profile of a static allocation? Fleming et al. (2001) provide early empirical support for this idea by estimating the economic value of volatility timing. They show that a risk-averse investor would be willing to pay an economically meaningful annual fee to replace an unconditionally efficient static portfolio with a volatility-timed conditional strategy. Their estimates suggest gains of roughly 50 to 200 basis points annually, depending on the investor's risk aversion. (Fleming et al. 2001.) These findings provide an early empirical justification for the idea that, if volatility is at least partly predictable, investors may improve portfolio performance by adjusting exposure over time. However, this argument depends on a critical assumption, that future volatility must be estimable with at least some accuracy from the past data. The following section, therefore, examines the empirical basis for volatility clustering and predictability.

2.2 Volatility Clustering and Predictability

A well-established feature in financial markets is the persistence of volatility, commonly referred to as volatility clustering (Mandelbrot 1963; Cont 2001). Mandelbrot (1963) was among the first to document this phenomenon, and Cont (2001) later described the same pattern, noting that autocorrelation of absolute or squared returns tends to remain positive and decay slowly over time. In practical terms, elevated volatility in the recent past tends to be followed by elevated volatility in the near future.

The establishment of formal econometric modelling for volatility clustering is attributed to Engle's (1982) Autoregressive Conditional Heteroskedasticity (ARCH) model, which linked the expected variance of returns to past squared returns. Building upon this foundation, Bollerslev (1986) introduced the Generalized ARCH (GARCH) framework, where the current conditional variance depends on both the latest market shock and the previous variance estimate. As Hull (2018, 234) illustrates, when squared returns are high in a given period, there is a tendency for subsequent squared returns to also be high, and vice versa. Furthermore, the GARCH model provides a tractable way to capture the autocorrelation structure observed in squared return series (Bollerslev 1986), indicating that volatility dynamics are structured rather than purely random.

Although GARCH models offer a parametric approach to volatility forecasting, simpler model-free alternatives have also demonstrated effectiveness. Early realized-volatility research showed that return variation can be measured directly from observed returns rather than only through a fully specified variance model. (Andersen & Bollerslev 1998.) Andersen et al. (2003) further show that, under suitable conditions, realized volatility provides an unbiased and highly efficient ex post estimator of return volatility. This makes realized-volatility-based approaches appealing in practice, since they rely directly on observed return data and they do not require the use of a fully specified parametric variance model from the beginning (Andersen & Bollerslev 1998; Andersen et al. 2003).

The predictability of volatility is also economically relevant. French et al. (1987) show that volatility is linked to expected returns, which suggests that volatility contains information that is relevant for investors. Moreover, this approach is supported by the forecasting literature, which shows that volatility is substantially more predictable than returns, even though no single forecasting model dominates in every setting (Poon & Granger 2003). This contrast between the predictability of volatility and the unpredictability of returns helps to explain why volatility targeting focuses on scaling exposure based on changes in risk rather than attempting to forecast returns directly.

Notably, changes in market risk are also asymmetric. In equity markets, volatility usually rises more after negative returns than positive returns of the same size. This phenomenon is known as the leverage effect. (Black 1976.) On a broader scale, financial return data displays several well-established characteristics, such as volatility clustering, non-normal distribution, and heavy tails (Cont 2001). Consistent with this, Hull (2018, 219) notes that the daily percentage changes for most market variables have much heavier tails than a standard normal distribution. In this thesis, this excess tail weight is evaluated using two higher-order moments of the return distribution: skewness and excess kurtosis.

These distributional features are directly relevant to the present study, because volatility targeting may affect not only average risk-adjusted returns but also the shape of the return distribution itself. Skewness measures the lack of symmetry in a return distribution, while kurtosis indicates the heaviness of the tails and how much data is concentrated around the centre compared to a standard normal distribution. Within the context of financial time series, negative skewness and excess kurtosis are especially important because they are associated with downside tail risk and a higher likelihood of extreme price movements. (Freeman et al. 2017, 60–61; DeCarlo 1997.) For this reason, changes in skewness and excess kurtosis are also relevant when assessing volatility-targeting strategies.

2.3 The Volatility Targeting Mechanism

The previous section established that volatility is persistent, at least partially predictable, and closely linked to the heavy tails observed in return distributions. This section presents the mechanism that exploits these facts: the volatility-targeting portfolio, which scales market exposure inversely with recent realized volatility so that the portfolio maintains an approximately constant level of risk over time.

The theoretical reasoning behind volatility targeting comes from an observed imbalance between risk and return in financial markets (Moreira & Muir 2017). Standard asset-pricing theory suggests that higher risk should be compensated by higher expected returns (Sharpe 1964). Therefore, if volatility increases were always matched by a proportional increase in expected returns, the risk-return ratio would remain relatively stable regardless of market conditions. In that case, volatility timing would have little room to improve portfolio efficiency.

However, actual market data suggests that this relationship between risk and return is much weaker than traditional theory would imply (Baker et al. 2011; Frazzini & Pedersen 2014). For instance, the well-documented low-volatility anomaly shows that lower-risk assets frequently produce better than expected returns, breaking the simple risk-return rule (Baker et al. 2011). A related explanation is that leverage-constrained investors may bid up high-beta assets, which lowers their future risk-adjusted returns relative to low-beta alternatives (Frazzini & Pedersen 2014). Moreover, periods of high market volatility are not typically matched by proportionally higher expected returns. This indicates that risk-adjusted performance often declines as volatility increases. As a result, investors may not be adequately compensated for the additional risk they take during these unstable periods. (Moreira & Muir 2017.) Volatility targeting aims to use this pattern to its advantage by lowering market exposure when volatility is high and raising it when volatility is low, ultimately seeking to improve overall risk-adjusted returns (Moreira & Muir 2017; Harvey et al. 2018).

In practice, one way to implement this mechanism is through a simple scaling rule. At the end of each month, the investor measures the realized volatility and adjusts next period's market exposure relative to a fixed volatility target. If the realized volatility exceeds the target, market exposure is reduced; if it is below the target, exposure is increased, subject to a leverage cap. In this form, the framework is simple and directly interpretable, since it requires only an estimate of recent realized volatility and a target level of risk, with no estimation of expected returns or correlations. (Moreira & Muir 2017.)

That said, volatility targeting is not tied to a single implementation. Existing approaches range from simple rolling window rules based on realized volatility to more complex model-based specifications that use conditional volatility estimators such as GARCH. (Moreira & Muir 2017; Harvey et al. 2018; Doan et al. 2018.) Among these, the framework proposed by Moreira and Muir (2017) is deliberately simple. This simplicity is supported by the findings that volatility is far more persistent and predictable than returns (Poon & Granger 2003), which suggest that even a straightforward backward-looking estimate can provide a meaningful signal for adjusting portfolio exposure. Kirby and Ostdiek (2012) likewise emphasize that simple timing-based allocation rules can outperform naïve diversification while maintaining relatively low turnover, even in the presence of transaction costs. Their results show that the practical value of timing strategies depends not only on forecasting ability but also on implementation frictions. In this respect, the appeal of Moreira and Muir's (2017) framework lies in its reliance on volatility estimates rather than return forecasts, which helps avoid the substantial estimation error associated with predicting expected returns.

Beyond improving risk-adjusted performance, volatility targeting may also affect the shape of the return distribution. Harvey et al. (2018) studied the impact of volatility targeting across more than 60 assets and showed that the strategy tends to compress tail events. This is particularly relevant on the left tail, where the largest negative returns tend to occur during periods of elevated volatility. Because the strategy reduces exposure precisely in such states, the most extreme negative observations are compressed. This tends to result in less negative skewness and lower excess kurtosis for risk assets. (Harvey et al. 2018.)

2.4 Conditional Volatility Targeting

The discussion in the previous section may suggest that scaling portfolio exposure in response to recent volatility is generally beneficial. However, later research indicates that the efficacy of volatility targeting is linked to the underlying market environment. While Moreira and Muir (2017) show that volatility-managed portfolios can improve risk-adjusted performance, further research indicates that these benefits are not uniform across all assets, time periods, or market conditions. For instance, research shows that volatility targeting consistently reduced the likelihood of extreme returns, but on the other hand, the improvements in Sharpe ratios are mostly seen in risk assets such as equities and credit, rather than across every asset class. (Harvey et al. 2018.)

A more direct challenge to the unconditional application of volatility targeting is provided by Bongaerts et al. (2020). They show that a standard implementable volatility-targeting strategy does not reliably improve risk-adjusted performance across international equity markets. In some cases, it may also exceed the target volatility, which can contribute to deeper drawdowns and greater tail risk. Their central argument is that the informational value of recent volatility is not constant across states. In particular, volatility clustering is stronger in high-volatility environments, and this relation between realized volatility and subsequent returns becomes more negative in extreme states. This implies that the case for scaling risk is strongest when volatility is unusually high or unusually low, whereas intermediate volatility states may not contain a strong enough signal to justify moving away from a static benchmark allocation. (Bongaerts et al. 2020.)

To address this, Bongaerts et al. (2020) propose a conditional strategy that adjusts portfolio weights strictly during extreme market conditions. This strategy reduces market exposure in high-volatility states, increases it in low-volatility states subject to a leverage limit, and otherwise maintains an unscaled exposure. This selective targeting yields better and more reliable Sharpe ratios compared to the conventional approach. The model also limits maximum drawdowns, expected shortfall, overall trading turnover, and leverage. Rather than abandoning the core logic of volatility targeting, this conditional approach simply restricts its practical use to extreme market states where the signal is strong. (Bongaerts et al. 2020.) Related evidence suggests that this conditionality concerns not only the market state in which scaling is applied, but also the way risk itself is measured. Wang and Yan (2021) show that strategies based on downside volatility tend to perform better than those based on total volatility, which indicates that the effectiveness of volatility management may depend on the choice of volatility measure as well.

2.5 Limitations: Whipsaw, Transaction Costs, and Implementation

Despite its theoretical appeal, volatility targeting is subject to practical limitations that may weaken its realized benefits (Cederburg et al. 2020). Initial evidence suggests that volatility-managed portfolios successfully improve risk-adjusted returns (Moreira & Muir 2017). However, more recent research indicates that the actual benefit depends heavily on the prevailing market environment and specific execution choices (Bongaerts et al. 2020; Cederburg et al. 2020). Consequently, the strategy should be evaluated not only in theory but also in terms of its practical limitations.

One frequently discussed limitation of this approach is whipsaw risk. This occurs when a volatility spike triggers a reduction in market exposure just before the market quickly rebounds. Since the strategy relies on backward-looking volatility estimates, it always reacts with a delay. If the market recovers rapidly, the portfolio may remain underexposed and therefore lag a static buy-and-hold allocation. Ultimately, the same mechanism that is designed to limit severe downside losses may also restrict participation in sudden market rallies. (Harvey et al. 2018.)

Transaction costs form a second limitation. Volatility targeting requires constant portfolio adjustments and can therefore generate significant turnover. Earlier research, nevertheless, suggests that volatility-managed portfolios may still perform well when moderate transaction costs are considered (Moreira & Muir 2017; Barroso & Santa-Clara 2015). However, this view is not unchallenged. Barroso and Detzel (2021) argue that the theoretical excess returns of volatility-managed portfolios are closely tied to limits to arbitrage. They show that these strategies often require high turnover, leverage, and short selling in relatively expensive-to-trade stocks. In practice, this suggests that the strategy may need its largest adjustments precisely when trading conditions are less favourable, for example, because bid-ask spreads are wider, and large positions are more difficult to scale efficiently. As a result, the abnormal returns of many volatility-managed equity portfolios may be weaker once such frictions are taken into account. (Barroso & Detzel 2021.)

A further concern relates to real-time implementation. While volatility-managed portfolios frequently produce strong in-sample alphas, they do not systematically produce stronger out-of-sample performance in the real world. One reason is that the relationship underlying these strategies is not stable over time. More specifically, patterns that seem strong in-sample may weaken or disappear when applied in real time. Additionally, some portfolio combinations that appear attractive in sample are simply too difficult to implement in actual trading environments. (Cederburg et al. 2020.)

Recent evidence suggests that part of this implementation problem may arise from differences in how target volatility is specified. Research distinguishes between volatility timing and target timing, showing that time-varying target volatility introduces an additional timing component beyond volatility timing alone. If the target remains strictly constant, portfolio adjustments depend entirely on volatility timing. However, a dynamically calibrated target may rise alongside market volatility, which actively restricts the strategy's ability to deleverage during financial stress. This structural difference helps explain conflicting historical results across the literature. (Xu 2026.) The strong performance observed by Moreira and Muir (2017) and Harvey et al. (2018) contrasts with the weak real-time execution highlighted by Bongaerts et al. (2020) and Cederburg et al. (2020). These varying outcomes may reflect the chosen target-volatility specification rather than a fundamental flaw in timing itself (Xu 2026).

Taken together, the literature suggests that volatility targeting is neither a universally superior alternative to static allocation nor a failed concept. Its effectiveness depends on volatility predictability, market regime, implementation frictions, and strategy design. (Moreira & Muir 2017; Bongaerts et al. 2020; Cederburg et al. 2020; Barroso & Detzel 2021.) This design sensitivity includes both the way risk is measured and the way the volatility target itself is specified, which may also affect out-of-sample performance (Wang & Yan 2021; Xu 2026). The relevant empirical question is therefore not whether volatility timing always works, but whether a simple and implementable specification can improve outcomes in each market environment. For this reason, the present thesis examines a fixed, monthly rebalanced volatility-targeting strategy on the STOXX Europe 600 across distinct volatility regimes.

3 Data and Methodology

3.1 Data Description

The primary dataset consists of daily closing prices of the STOXX Europe 600 Net Return index (STOXXR), obtained from LSEG Refinitiv Workspace. The index contains large-, mid-, and small-cap companies across 17 European countries and therefore provides broad and diverse exposure to the European equity market. The net return version is used, since it reflects dividend reinvestment after withholding taxes, which represents the return experience of a typical institutional investor more accurately than a price-only index.

1-month Euribor, which is also obtained from LSEG Refinitiv, serves as the proxy for the risk-free rate. The main reason for this choice is the data availability, since the euro short-term rate (€STR), which is now the official euro-area overnight reference rate, has been published only since October 2019 and therefore does not cover the full 2010–2025 sample (ECB 2026). For this reason, 1-month Euribor provides a practical way to construct a short-term risk-free proxy for the whole period. Since Euribor is quoted as an annualized percentage rate, the daily risk-free rate is calculated using a simple linear approximation. The annual rate is converted into decimal form and divided by 252, which represents the assumed number of trading days per year.

The full sample includes market data from January 2010 to December 2025. The entire sample is divided into three different sub-periods to investigate the strategy's performance in different market conditions. These are the pre-crisis period (2010–2019), which captures a prolonged economic expansion following the global financial crisis; the crisis period (2020–2021), reflecting extreme volatility and the subsequent recovery related to the COVID-19 pandemic; and the post-2022 period (2022–2025), which is shaped by the European Central Bank's monetary tightening, the economic repercussions of Russia's war against Ukraine, and renewed global trade tensions toward the end of the sample (ECB 2025). This periodization is also consistent with Bongaerts et al. (2020), who argue that the effectiveness of volatility targeting depends on the prevailing market regime.

3.2 Volatility Estimation

The strategy requires an estimate of current market volatility to determine the appropriate leverage at each rebalancing date. Following the approach of Moreira and Muir (2017), this thesis employs realized volatility estimated from a backward-looking rolling window of daily returns. This estimator is chosen for its transparency and ease of implementation, since realized volatility requires only historical return data and no strong distributional assumptions. The annualized realized volatility on day t is computed as:

$$\hat{\sigma}_t = \sqrt{252} \times \sqrt{\frac{1}{W-1}\sum_{i=0}^{W-1}(r_{t-i} - \bar{r}_t)^2} \ , \tag{1}$$

where W is the rolling window length measured in trading days, r_{t-i} is the simple daily return observed i days before day t, and $\bar{r}_t$ is the average daily return over the estimation window. The expression inside the square root is the sample variance of the last W daily returns, and taking the square root gives the daily sample standard deviation. The daily standard deviation is annualized by the factor $\sqrt{252}$ so that the estimate is directly comparable to the target volatility used in the strategy.

The baseline specification uses W = 21 trading days, which is approximately one calendar month and is a common choice in the volatility-timing literature. In the sensitivity analysis, a longer window of W = 63 trading days is also assessed to see whether a smoother and slower-moving volatility estimate affects performance.

The choice of realized volatility over alternative volatility estimators, such as GARCH or implied volatility from option markets, is motivated by simplicity, transparency, and the absence of parametric assumptions. This is broadly consistent with Moreira and Muir (2017), who rely on simple backward-looking volatility measures in their empirical implementation. Although more complex estimators might respond differently to changing market conditions, realized volatility offers a straightforward and widely used measure that can be implemented in real time with minimal modelling choices. This makes it well-suited for the present thesis, which emphasizes a simple and implementable volatility-targeting framework.

3.3 Strategy Construction

The volatility targeting strategy actively adjusts market exposure to maintain a roughly constant target level of portfolio volatility. The basic mechanism is straightforward: when realized volatility is high relative to the target, exposure is reduced, whereas low realized volatility leads to higher exposure. The strategy follows the general volatility-targeting framework of Moreira and Muir (2017) and is implemented here in a target-volatility form consistent with Harvey et al. (2018).

At the end of each month, a leverage multiplier is formed using the most recent realized volatility estimate. The multiplier applied in month $t + 1$ is defined as

$$L_{t+1} = \min\left(\frac{\sigma^*}{\hat{\sigma}_t}, L_{max}\right), \tag{2}$$

where L_{t+1} is the leverage applied during month $t+1$, σ^* is the target volatility, $\hat{\sigma}_t$ is the realized volatility estimate at the end of month t, and L_{max} is the maximum leverage cap. This formulation is directly interpretable in a target-volatility setting, as exposure is reduced when estimated volatility exceeds the target and increased when it falls below it. Because both target volatility and realized volatility are positive, the leverage multiplier is always non-negative. This means that the strategy does not take short positions. The leverage cap prevents excessive leverage during low-volatility periods and keeps the model realistic.

The daily return of the volatility-targeted portfolio follows directly from the leverage rule defined in Equation (2). Since the strategy allocates a weight of L_t to the market portfolio and the remaining weight $1 - L_t$ to the risk-free asset, the portfolio return is calculated as

$$r_t^{VT} \; = \; (1 - L_t) r_{f,t} \; + \; L_t \times r_{m,t} \,, \tag{3}$$

where r_t^{VT} is the daily return of the strategy, $r_{f,t}$ is the daily risk-free rate, and $r_{m,t}$ is the daily market return. In economic terms, the strategy scales the market excess return by the chosen leverage multiplier. When $L_t < 1$, the remaining weight is invested in the risk-free asset, whereas when $L_t > 1$, the strategy takes a leveraged position in the market portfolio.

Rebalancing is implemented monthly. More specifically, the strategy uses the last available leverage signal from month t and applies it through month $t+1$. Leverage, therefore, remains constant within each calendar month, and no further rebalancing occurs within the month. This timing design choice ensures that only information available at the end of the previous month is used when determining the next month's exposure.

The baseline specification uses a target volatility of 15%, a 21-day volatility-estimation window, and a maximum leverage of 1.5. These values provide a reasonable benchmark for a broad equity index and are varied later in the sensitivity analysis in order to assess the robustness of the results.

3.4 Performance Metrics

The performance of the volatility-targeted strategy is assessed relative to a buy-and-hold benchmark using widely recognized risk-adjusted performance measures. Return performance is measured by the compound annual growth rate (CAGR), which reflects the constant annualized rate of return that would generate the observed cumulative growth over the sample period. Total risk is measured by annualized volatility, calculated from the standard deviation of daily returns, and then annualized.

Risk-adjusted performance is evaluated using the Sharpe and Sortino ratios. The Sharpe ratio, introduced by Sharpe (1966), measures mean excess return per unit of total risk and is defined as

$$Sharpe\ Ratio = \frac{E\left(r_t - r_{f,t}\right)}{\sigma\left(r_t - r_{f,t}\right)}\sqrt{252}\,, \tag{4}$$

where r_t is the daily portfolio return and $r_{f,t}$ the daily risk-free rate. A higher Sharpe ratio indicates a more favourable risk-adjusted return profile. (Sharpe 1966.) The Sortino ratio applies the same logic, except it replaces the total volatility with downside deviation, so that only negative excess returns contribute to the risk measure. In this thesis, the minimum acceptable return is set to zero. Sortino ratio is defined as

$$Sortino\ Ratio = \frac{E\left(r_t - r_{f,t}\right)}{\sigma_d\left(r_t - r_{f,t}\right)}\sqrt{252}\,, \tag{5}$$

where σ_d is the downside deviation of daily excess returns. This makes the Sortino ratio particularly relevant for strategies that aim to reduce downside risk. (Sortino & Van der Meer 1991.)

Maximum drawdown measures the largest decline in cumulative wealth within the sample period and is used here as a downside risk measure in portfolio evaluation (Bacon 2008, 68–69). It is calculated as

$$MDD = \ min_t\left(\frac{V_t}{max_{s\leq t}\ V_s} - 1\right), \tag{6}$$

where V_t represents cumulative wealth at time t, and $max_{s\leq t}$ is the highest cumulative wealth level reached up to time t. The Calmar ratio complements this by relating annualized return to maximum drawdown and, in this thesis, is calculated as

$$Calmar\ ratio = \frac{CAGR}{|MDD|}. \tag{7}$$

This approach is consistent with common applied portfolio-performance practice, although the original Calmar ratio was defined over a 36-month window (Young 1991). Together, these metrics capture different dimensions of performance: absolute returns through CAGR, risk-adjusted returns through the Sharpe and Sortino ratios, and tail risk exposure through maximum drawdown and the Calmar ratio.

In addition to standard performance metrics, the distributional properties of daily returns are examined through skewness and excess kurtosis. These higher moments are relevant because previous research suggests that volatility targeting may alter the shape of the return distribution by compressing extreme observations. In particular, Harvey et al. (2018) demonstrate that the strategy reduces the likelihood of extreme returns and tends to improve left-tail outcomes. Skewness and excess kurtosis are therefore reported as additional measures to assess whether similar distributional effects are also present in the STOXX Europe 600 sample.

3.5 Robustness Design

To assess whether the findings from the STOXX Europe 600 generalize to other equity markets, the same strategy is applied to two additional indices. The S&P 500 Total Return Index (SPXTR) serves as a natural benchmark, since both Harvey et al. (2018) and Moreira and Muir (2017) study volatility-targeting strategies primarily in a U.S. equity market. The OMX Helsinki Gross Index (OMXHGI) represents a smaller, more concentrated Nordic market with less diversification than the STOXX Europe 600 and S&P 500. This contrast helps assess whether volatility targeting remains effective in a less diversified index environment. Including these indices allows the strategy to be compared across markets with different breadth, concentration, and structure.

The 1-month U.S. Treasury bill rate is used as the risk-free rate for the S&P 500, while the OMXHGI analysis uses the same 1-month Euribor. All other strategy parameters remain constant to ensure comparability. The same three sub-periods are used for all indices. This design allows for a direct comparison of whether the volatility targeting yields different results in markets with distinct breadth, liquidity, and volatility characteristics.

The sensitivity analysis examines whether the baseline results depend on the chosen target volatility and volatility-estimation window. To do so, the target volatility is varied across five levels (10%, 12%, 15%, 18%, and 20%), while the estimation window is set to either 21 or 63 trading days. These alternative specifications are evaluated separately for each of the three sub-periods. This allows for a systematic assessment of whether the performance of volatility targeting is sensitive to parameter choices and whether the same specification performs similarly across different market environments.

4 Empirical Results

4.1 Baseline Performance

This section shows the primary results of the volatility targeting strategy applied to the STOXX Europe 600 Net Return index across the three sub-periods. The key performance results for both the buy-and-hold and volatility-targeted portfolio are summarized in Table 1.

Table 1. Performance metrics for buy-and-hold and volatility-targeted strategies

STOXX Europe 600 Net Return Index. CAGR = compound annual growth rate. Ann. Vol = annualized volatility. Max DD = maximum drawdown. Lev. = mean leverage. B&H = buy-and-hold. Vol-T = volatility-targeted strategy.

Period	Strategy	CAGR	Ann. Vol	Sharpe	Sortino	Max DD	Calmar	Lev.
Pre (2010–2019)	B&H	8.40%	15.8%	0.584	0.825	–25.2%	0.333	1.00
	Vol-T	8.17%	16.4%	0.557	0.779	–25.5%	0.320	1.15
Crisis (2020–2021)	B&H	11.54%	21.7%	0.638	0.836	–35.4%	0.326	1.00
	Vol-T	10.29%	18.0%	0.663	0.862	–30.3%	0.340	1.01
Post (2022–2025)	B&H	9.96%	13.7%	0.590	0.817	–16.2%	0.615	1.00
	Vol-T	10.11%	15.8%	0.538	0.735	–18.4%	0.550	1.24

Table 1 shows that during the pre-crisis period, the buy-and-hold strategy slightly outperformed the volatility-targeted portfolio, achieving an 8.40% CAGR compared to 8.17%. In addition, the benchmark's Sharpe ratio (0.584) also exceeded that of the managed strategy (0.557). Maximum drawdowns were nearly identical (–25.2% vs. –25.5%), suggesting that the strategy did not provide meaningful downside protection during this prolonged, relatively low-volatility bull market period.

The crisis regime reveals a slightly more nuanced picture. While buy-and-hold reached a higher CAGR (11.54% vs. 10.29%), the volatility-targeted strategy delivered a superior Sharpe ratio (0.663 vs. 0.638) and Sortino ratio (0.862 vs. 0.836). In addition, the maximum drawdown was reduced from –35.4% to –30.3%, which is a 5.1-percentage point improvement in tail risk protection. This finding is consistent with the logic of Moreira and Muir (2017), who show that volatility targeting tends to be most beneficial when volatility is high, since expected returns do not rise proportionally with risk. The mean leverage of 1.01 implies that the strategy rotated between de-risking during volatility spikes and leveraging during calmer months, resulting in an average market exposure close to one.

The volatility-targeted portfolio reached a slightly higher CAGR (10.11% vs. 9.96%) in the post-2022 period, but the improved CAGR came at the expense of increased volatility (15.8% vs. 13.7%).

Furthermore, the maximum drawdowns also became slightly worse (–18.4% vs. –16.2%). During this period, the mean leverage was 1.24, which suggests that the strategy was mostly leveraged above one and both gains and losses were amplified.

Figure 1 illustrates these dynamics visually. The cumulative wealth paths of the two strategies are closely intertwined throughout the sample, with the most visible divergence occurring during the COVID-19 drawdown in early 2020, where the volatility-targeted strategy's shallower decline is apparent.

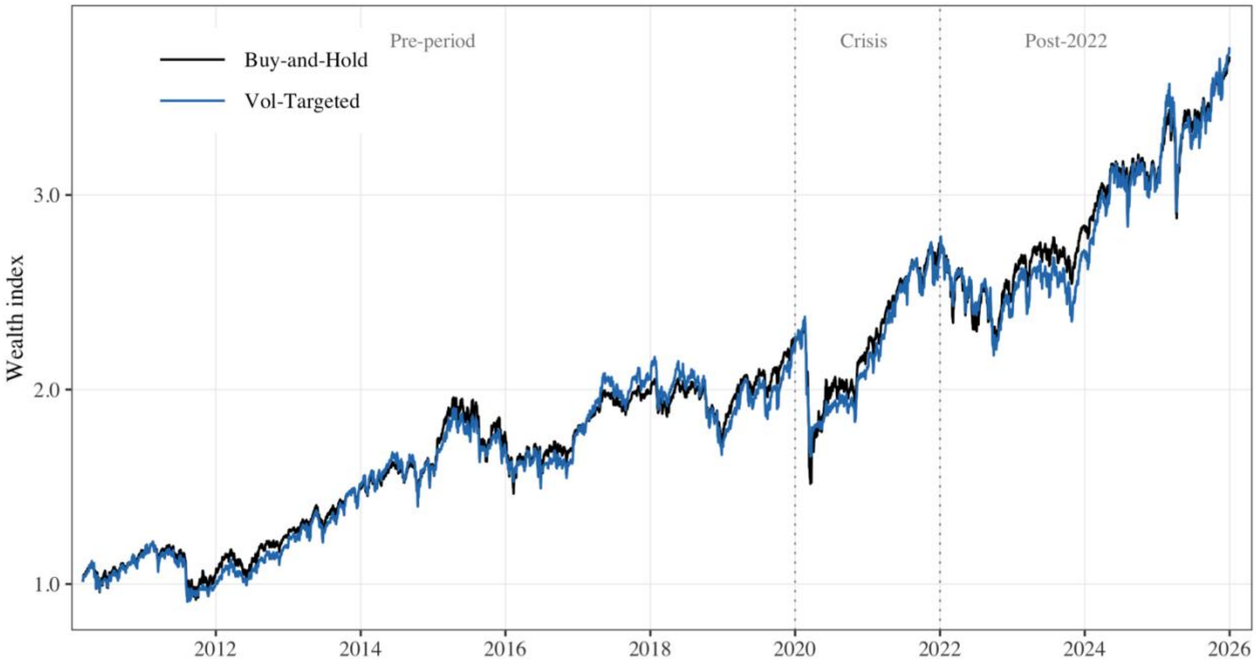


Figure 1. Wealth paths of buy-and-hold and volatility-targeted strategies

STOXX Europe 600 Net Return Index. At the beginning of each sub-period, the volatility-targeted series is re-anchored to the contemporaneous buy-and-hold wealth level to enable within-period comparison.

All in all, the baseline results remain mixed. The crisis period provides some support to the usefulness of volatility targeting in terms of drawdown reduction and a slight improvement in risk-adjusted returns. However, the strategy does not deliver equally clear benefits in the other sub-periods. This indicates that the benefits of volatility targeting may depend on the prevailing market environment rather than arising consistently across all periods. This interpretation is also consistent to some extent with Bongaerts et al. (2020), who argue that conventional volatility targeting is least effective in situations where volatility is not extremely high, as recent volatility may not provide a sufficiently strong signal to deviate from a static allocation.

4.2 Volatility Regime Analysis

To examine the conditional performance of volatility targeting more directly, each sub-period is divided into high- and low-volatility regimes. The classification is based on the median realized volatility within each sub-period and follows the general framework of Bongaerts et al. (2020). The corresponding results are shown in Table 2.

Table 2. Sharpe ratios and annualized excess returns by volatility regime

The regime split is based on the within-period median realized volatility. ΔSR = Sharpe(Vol-T) − Sharpe(B&H). Excess returns are annualized. The p-value refers to a t-test of the difference in daily excess returns between the two strategies within each regime.

Period	Regime	SR B&H	SR Vol-T	ΔSR	Ann. excess return B&H	Ann. excess return Vol-T	p-value
Pre (2010–2019)	High vol	0.696	0.699	+0.003	13.8%	12.9%	0.762
	Low vol	0.454	0.382	–0.071	4.6%	5.3%	0.703
Crisis (2020–2021)	High vol	0.504	0.502	–0.002	14.3%	10.2%	0.737
	Low vol	1.123	0.884	–0.238	13.4%	13.8%	0.939
Post (2022–2025)	High vol	0.431	0.190	–0.241	7.3%	3.5%	0.314
	Low vol	0.957	1.058	+0.101	8.8%	13.5%	0.101

A consistent pattern emerges across all three sub-periods. In times of low volatility, when the strategy applies leverage above 1, the managed portfolio tends to generate higher excess returns. This is most apparent in the post-2022 low-volatility regime, where annualized excess returns reach 13.5% compared to 8.8% for the benchmark. On the other hand, in high-volatility regimes, the strategy's deleveraging decreases returns, which is most noticeable in the post-2022 high-volatility regime, where the Sharpe ratio decreases from 0.431 to 0.190.

This asymmetry explains why the aggregate results in Table 1 are mixed. The strategy reduces exposure when volatility rises, but this de-risking can also limit participation in the following market recoveries. However, none of the p-values for the difference in daily excess returns reached the 5% significance level. This lack of statistical power is at least partially attributable to the limited number of observations within each regime, since the sample is split by both sub-period and volatility state. To put it another way, the observed differences are suggestive but not statistically strong enough to support firm conclusions about regime-level excess return differences between the two approaches.

Figure 2 illustrates the strategy's dynamic behaviour. The most notable observation is the sharp and deep drop in leverage during the COVID-19 crisis in early 2020 (reaching approximately 0.25), suggesting aggressive de-risking. It subsequently recovers during the second half of 2020 and alternates between leveraged and de-risked positions throughout the post-2022 period.

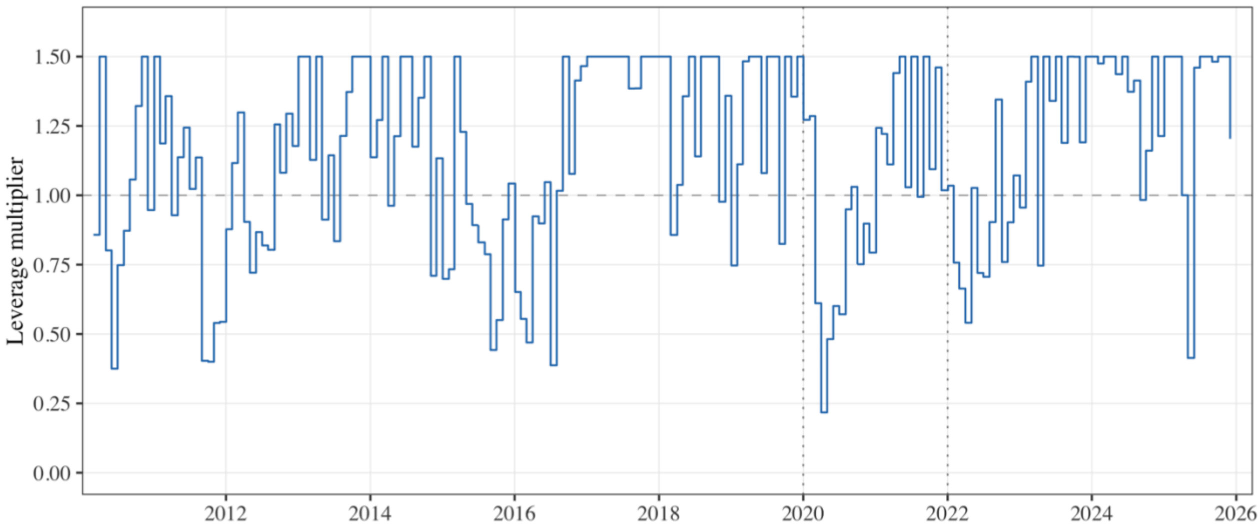


Figure 2. Monthly applied leverage of the volatility-targeted strategy

4.3 Tail Risk and Higher Moments

Beyond Sharpe ratios and drawdown, volatility targeting may also affect the distributional shape of returns. This is relevant because the strategy is designed to reduce exposure during high-volatility periods, when extreme return realizations are more likely to occur. Table 3 presents skewness and excess kurtosis for the buy-and-hold and volatility-targeted portfolios across the three sub-periods.

Table 3. Skewness and excess kurtosis of daily returns

Period	Strategy	Skewness	Excess Kurtosis
Pre (2010–2019)	B&H	–0.247	4.30
	Vol-T	–0.390	3.39
Crisis (2020–2021)	B&H	–1.455	15.04
	Vol-T	–1.276	6.58
Post (2022–2025)	B&H	–0.449	4.14
	Vol-T	–0.594	2.33

The most notable finding comes from the crisis regime. In this period, the buy-and-hold strategy has an excess kurtosis of 15.04, meaning exceptionally fat tails, while the volatility-targeted strategy lowers this to 6.58, which represents a 56% reduction. This considerable decrease in tail risk implies that the probability of extreme daily losses is substantially lower under volatility targeting. The skewness also improves slightly from –1.455 to –1.276.

In the post-2022 period, excess kurtosis falls from 4.14 to 2.33, indicating a lower degree of tail heaviness under volatility targeting. However, skewness becomes more negative (–0.449 vs. –0.594). This suggests that the effect of volatility targeting on higher moments is not uniform across all dimensions of the return distribution. The evidence is more consistent for kurtosis than for skewness.

In the pre-period, the reduction in excess kurtosis from 4.30 to 3.39 again points to lower tail heaviness under volatility targeting, although skewness becomes slightly more negative. Overall, these findings are broadly in line with Harvey et al. (2018), who show that volatility targeting tends to compress tail events by reducing exposure during periods of elevated risk. In the present study, this effect appears most clearly in the reduction of excess kurtosis, while the impact on skewness remains more mixed.

One possible explanation for this is that volatility targeting affects kurtosis and skewness through different channels. While deleveraging during high-volatility periods compresses the most extreme observations, leverage above one in calmer periods may amplify return fluctuation more broadly. This may help explain why the strategy improves kurtosis more consistently than skewness.

4.4 Whipsaw Analysis

As discussed earlier in this thesis, a practical concern in volatility targeting is the whipsaw effect. To examine this mechanism, deleveraging episodes are defined as months in which the leverage multiplier falls by more than 20%. For each such episode, the cumulative STOXX Europe 600 market return over the following 21 trading days is then calculated. A positive forward return refers to the share of deleveraging episodes followed by a positive cumulative market return over the next 21 trading days. A higher share therefore suggests that deleveraging was often followed by a rebound, which is consistent with the whipsaw mechanism. The results are reported in Table 4.

Table 4. Whipsaw analysis of forward returns after deleveraging episodes

Period	Episodes	Positive 21-day market return (%)	Mean 21-day market return	t-statistic (p-value)
Pre (2010–2019)	27	55.6%	+0.72%	0.993 (0.165)
Crisis (2020–2021)	7	71.4%	–0.27%	–0.098 (0.538)
Post (2022–2025)	8	50.0%	+1.09%	0.690 (0.256)

The results provide only limited support for the whipsaw hypothesis. In the pre-period, 27 deleveraging episodes are identified, and 55.6% of them are followed by a positive cumulative market return over the following 21 trading days. The mean 21-day market return is slightly positive at +0.72%, but the result is not statistically significant at conventional levels ($p = 0.165$).

In the crisis period, 71.4% of the seven episodes are followed by positive returns, which is directly consistent with the whipsaw mechanism. However, the mean forward return remains negative at –0.27%, and the very small number of episodes means that this result should be interpreted with caution.

In the post-2022 period, the evidence is also mixed. The forward return distribution is even at 50.0%, while the mean market return is +1.09%. Although this suggests that rebounds did occur after some deleveraging episodes, the result is again statistically insignificant. Taken together, these findings suggest that whipsaw may occur in individual episodes, but the evidence is not strong enough to conclude that it was a systematic source of underperformance in the STOXX Europe 600.

4.5 Sensitivity Analysis

To assess the study's sensitivity to chosen parameters, Table 5 reports Sharpe ratios across several target volatilities and estimation windows. The main comparison is between the 21-day and 63-day volatility-estimation windows.

Table 5. Sensitivity analysis: Sharpe ratios across parameter combinations

Period	Window	σ*=10%	σ*=12%	σ*=15%	σ*=18%	σ*=20%
Pre (2010–2019)	21d	0.512	0.526	0.557	0.568	0.571
	63d	0.482	0.483	0.513	0.513	0.508
Crisis (2020–2021)	21d	0.608	0.629	0.663	0.692	0.741
	63d	1.704	1.704	1.704	1.764	1.781
Post (2022–2025)	21d	0.537	0.538	0.538	0.530	0.526
	63d	0.734	0.732	0.725	0.677	0.668

The results indicate that the choice of volatility-estimator window materially influences performance. In the crisis period, extending the window from 21 to 63 days raises the Sharpe ratio from 0.663 to 1.704 at 15% target volatility. The same direction appears in the post-2022 period, where the Sharpe ratio increases from 0.538 to 0.725. By contrast, in the pre-period, the longer window slightly weakens performance, suggesting that window length becomes particularly important in more turbulent market environments. However, this finding should be treated with caution. The strong crisis-period performance of the 63-day window is likely partly specific to the sample, as the COVID-19 shock produced a sustained period of elevated volatility that favoured a slower-moving signal. Whether the same specification would perform equally well in shorter and sharper dislocations remains uncertain.

A possible explanation is that the longer window smooths out the volatility signal, which helps prevent the strategy from premature deleveraging during short-term volatility spikes. In calmer market conditions, however, this smoothing appears less beneficial. The sensitivity analysis therefore suggests that the volatility-estimation window is not just a minor technical choice, but an important component of the strategy design.

The relationship between target volatility and performance also varies across periods. In the pre-period, and especially during the crisis, higher target volatilities are associated with higher Sharpe ratios. In the post-2022 period, by contrast, the relationship is flat or slightly negative. This suggests that taking on more exposure was not systematically rewarded in the recent market environment.

4.6 Robustness: S&P 500 and OMXHGI

To evaluate the generalizability of the study, the same strategy is applied to the S&P 500 Total Return Index and the OMX Helsinki Gross Index. Table 6 provides a cross-market comparison based on key performance metrics.

Table 6. Cross-market comparison of buy-and-hold and volatility-targeted strategies

CAGR = compound annual growth rate. SR = Sharpe ratio. MDD = maximum drawdown. Sort. = Sortino ratio.

Market	Period	CAGR B&H	CAGR Vol-T	SR B&H	SR Vol-T	MDD B&H	MDD Vol-T	Sort B&H	Sort Vol-T
S&P 500	Pre	13.88%	16.22%	0.919	0.971	–19.4%	–20.5%	1.295	1.352
	Crisis	31.66%	22.29%	1.168	1.098	–28.4%	–18.2%	1.643	1.498
	Post	14.11%	15.49%	0.600	0.731	–22.1%	–17.1%	0.863	1.042
OMXHGI	Pre	8.81%	10.19%	0.553	0.656	–33.8%	–28.3%	0.784	0.929
	Crisis	23.29%	18.38%	1.093	1.078	–29.0%	–19.6%	1.492	1.472
	Post	7.58%	6.22%	0.396	0.307	–18.0%	–23.4%	0.554	0.428

The robustness results for the S&P 500 are clearly more favourable than those for the STOXX Europe 600. The Sharpe ratio improves in both the pre-period and the post-2022 period, while maximum drawdown is reduced materially during both the crisis regime and the post-2022 period. Overall, the findings suggest that volatility targeting performs more consistently in the S&P 500 than in the main European benchmark.

Another notable finding in the robustness study is that the strategy applied to the S&P 500 yields significantly higher excess returns than the buy-and-hold strategy in the pre-period low-volatility regime. This is also the only case in the thesis in which the difference in excess returns is statistically significant at conventional levels ($t = 3.065$, $p = 0.0022$). The result is broadly consistent with Moreira and Muir (2017), since the strategy takes its highest exposure in relatively calm market conditions. One possible explanation for the stronger results in the S&P 500 is that volatility timing may work more effectively in the U.S. equity markets than in a broad European equity index. It is also possible that backward-looking volatility measures provide a more informative input for portfolio scaling in the S&P 500. This interpretation is broadly consistent with the empirical findings of Fleming et al. (2001), who showed that volatility timing provides economically valuable results in a U.S. market setting.

The OMXHGI, by contrast, produces mixed results across the three sub-periods. In the pre-period, the strategy delivers a clear improvement, as both risk-adjusted returns and maximum drawdown show clear gains relative to buy-and-hold. During the crisis period, the main benefits come from downside protection, since the strategy substantially reduced maximum drawdown, while the Sharpe and Sortino ratios remain essentially unchanged. The post-2022 period, however, reveals a deterioration. Both risk-adjusted performance and CAGR weaken, and maximum drawdown worsens from –18.0% to –23.4%, which represents the most severe drawdown amplification across all tested market-period combinations.

Compared with the STOXX Europe 600, where the post-2022 deterioration in drawdown is more modest, the OMXHGI result suggests that volatility targeting may be less robust in a narrower market index. One possible explanation is that the OMXHGI, with only 130 constituents, offers substantially less diversification than the STOXX Europe 600. Its narrower composition may therefore generate volatility dynamics that are less stable and less compatible with the assumptions of a simple volatility-targeting rule.

These cross-market results help to identify the boundaries of volatility targeting. The strategy appears to perform most favourably in the U.S. market setting represented by the S&P 500. By contrast, the results are more mixed in the STOXX Europe 600 and may even turn adverse in a narrower market such as the OMXHGI during certain periods. This has direct practical relevance, since the U.S.-focused evidence that dominates the literature may give too favourable an impression of the strategy's usefulness for European investors. The full-period cumulative return paths of the volatility-targeted portfolios for both markets are presented in Appendix 1 and Appendix 2.

5 Conclusion and Discussion

This thesis evaluated how a simple volatility-targeting strategy performed on the STOXX Europe 600 compared to a static buy-and-hold approach. The analysis also aimed to identify whether its effectiveness depends on the underlying market conditions and implementation choices. The results suggest that volatility targeting was not a consistently superior alternative to buy-and-hold. Instead, its effectiveness depended strongly on the prevailing market environment. The most apparent support for the strategy came from the crisis period of 2020–2021, when the volatility-targeted portfolio achieved a slightly higher Sharpe ratio and a clearly smaller maximum drawdown than the benchmark. Outside of this extreme volatility period, the benefits largely disappeared. The strategy failed to improve risk-adjusted returns or provide meaningful downside protection during the quieter pre-period. Performance post-2022 was equally mixed: while the compound annual growth rate (CAGR) improved slightly, the portfolio's volatility increased, the Sharpe ratio decreased, and the drawdown grew slightly larger. Taken together, these findings suggest that volatility targeting appears conditionally useful rather than consistently superior.

This evidence becomes more convincing when attention is extended beyond average risk-adjusted returns alone. In the crisis period, volatility targeting not only reduced maximum drawdowns but also substantially lowered excess kurtosis, cutting it by more than half from 15.04 to 6.58. This indicates that the strategy compressed the most extreme return realizations during the most turbulent part of the sample. The effect on skewness was weaker and more mixed across periods, but the kurtosis results provide clearer support for the view that volatility targeting can improve the shape of the return distribution by reducing tail heaviness. In this respect, the findings are broadly consistent with Harvey et al. (2018), who argue that volatility targeting compresses tail events by cutting exposure during periods of elevated risk. The results therefore suggest that the strategy may be more convincing as a downside risk management tool than a universally superior return improvement strategy.

The findings also align reasonably well with the wider literature on regime dependence. Moreira and Muir (2017) show that volatility-managed portfolios can improve risk-adjusted performance when volatility rises without a proportional increase in expected returns, and the crisis-period results are consistent with that mechanism. At the same time, the larger pattern of this thesis is closer to the more conditional interpretation of Bongaerts et al. (2020). The results do not support the view that simple volatility scaling improves performance in every environment. Instead, the strategy seems to become most relevant when market conditions are sufficiently stressed for volatility to contain a stronger signal about future risk. This interpretation is further supported by the regime analysis, where low-volatility states often benefited from leverage, while high-volatility states frequently involved weaker returns due to de-risking. At the same time, none of the regime-level return differences for the STOXX Europe 600 reached conventional significance levels. The evidence should therefore be interpreted as suggestive of a regime-dependent pattern rather than as definitive proof of systematic return differences.

One of the most important findings of the thesis concerns the sensitivity of performance to the volatility-estimation window. Extending the window from 21 to 63 trading days materially improved performance in turbulent periods. In the crisis period, the Sharpe ratio rose from 0.663 to 1.704 at a 15% target volatility, while the post-2022 period also showed a clear improvement under the longer specification. This difference is too large to be treated as a minor robustness detail. Instead, it suggests that the volatility-estimation window is a central design choice. A plausible interpretation is that the longer window smooths the volatility signal and therefore reduces overly reactive deleveraging during short-lived volatility spikes. This suggests that balancing between speed of adjustment and signal stability is an important part of volatility-targeting design rather than a secondary technical detail.

This observation also motivates a broader methodological reflection. The use of backward-looking realized volatility is transparent, simple, and easy to implement, which makes it well-suited to the framework of this thesis. However, it is still only one possible way to estimate risk. A GARCH-based conditional volatility model could, in principle, have generated different leverage paths, especially during periods of rapidly changing volatility. This does not mean that the current approach is inappropriate. Rather, it means that the present findings should be interpreted as evidence on one specific and deliberately simple implementation of volatility targeting, not as a universal statement about every possible volatility-managed strategy. A similar point also applies to the choice of a fixed 15% target volatility. Xu (2026) argues that mixed out-of-sample evidence may partly reflect differences in how target volatility itself is specified rather than volatility timing alone. From that perspective, the fixed target used in the present thesis is not only a simple implementation but also a design choice, although the thesis does not directly test whether alternative target specifications would have been more effective.

The cross-market robustness analysis provides further evidence on the boundaries of volatility-targeting. The S&P 500 results were clearly more favourable than those for the STOXX Europe 600, while the OMXHGI results were more mixed and, in the post-2022 period, clearly weaker. One possible explanation is that volatility targeting works more effectively in the S&P 500 because it reflects a more unified and less fragmented market environment, whereas the STOXX Europe 600 aggregates firms from multiple national markets with potentially more heterogeneous volatility dynamics. In such an environment, backward-looking realized volatility may provide a less uniform signal for volatility scaling. This interpretation is in line with wider evidence showing that European markets are strongly exposed to cross-market volatility transmission (Kaeck & Alexander 2013). This is an important practical finding, since much of the literature is based on U.S. equity data. The present results suggest that U.S.-based evidence may give too positive an impression of the strategy's usefulness for European investors.

Several limitations should nevertheless be acknowledged. First, although whipsaw is a theoretically important limitation of volatility targeting, the direct whipsaw analysis in the thesis provides only limited support for the view that it was a systematic driver of underperformance in the STOXX Europe 600. This suggests that the weaker results observed outside the crisis period are more likely to reflect broader issues of regime dependence and implementation design than whipsaw alone. Second, the empirical analysis does not incorporate transaction costs, taxes, or execution frictions at a detailed level, even though these may matter in practice. Third, the post-2022 period remains relatively short for making strong conclusions. Fourth, although the results are clearly regime-dependent, the thesis does not directly estimate a conditional volatility-targeting strategy in the spirit of Bongaerts et al. (2020).

Overall, the results of this thesis suggest that volatility targeting is best understood as a conditional and design-sensitive risk-management strategy rather than as a universally superior alternative to buy-and-hold. Its clearest value in the STOXX Europe 600 appears in crisis conditions, where it modestly improves risk-adjusted performance and more clearly reduces drawdowns and tail heaviness. Outside such environments, the benefits seem weaker and less consistent. The main contribution of this thesis is therefore not to show that volatility targeting always works, but to show that its usefulness depends on market regime, index structure, and implementation design. Future research could extend this analysis by testing a direct conditional volatility-targeting strategy, comparing rolling-window and GARCH-based volatility estimators, and examining whether the strong sensitivity to estimation-window length also appears in other European equity markets.

References

Andersen, T. G. – Bollerslev, T. (1998) Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, Vol. 39 (4), 885–905.

Andersen, T. G. – Bollerslev, T. – Diebold, F. X. – Labys, P. (2003) Modeling and forecasting realized volatility. *Econometrica*, Vol. 71 (2), 579–625.

Bacon, C. R. (2008) *Practical portfolio performance: measurement and attribution*. 2nd ed. John Wiley & Sons, Chichester.

Baker, M. – Bradley, B. – Wurgler, J. (2011) Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly. *Financial Analysts Journal*, Vol. 67 (1), 40–54.

Barroso, P. – Detzel, A. (2021) Do limits to arbitrage explain the benefits of volatility-managed portfolios? *Journal of Financial Economics*, Vol. 140 (3), 744–767.

Barroso, P. – Santa-Clara, P. (2015) Momentum has its moments. *Journal of Financial Economics*, Vol. 116 (1), 111–120.

Black, F. (1976) Studies of stock price volatility changes. *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section*, 177–181.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, Vol. 31 (3), 307–327.

Bongaerts, D. – Kang, X. – van Dijk, M. A. (2020) Conditional volatility targeting. *Financial Analysts Journal*, Vol. 76 (4), 68–85.

Cederburg, S. – O'Doherty, M. S. – Wang, F. – Yan, X. S. (2020) On the performance of volatility-managed portfolios. *Journal of Financial Economics*, Vol. 138 (1), 95–117.

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

DeCarlo, L. T. (1997) On the meaning of kurtosis: Standardized moments, residuals, and the meaning of kurtosis. *Psychological Methods*, Vol. 2 (3), 292–307.

Doan, B. – Papageorgiou, N. – Reeves, J. J. – Sherris, M. (2018) Portfolio management with targeted constant market volatility. *Insurance: Mathematics and Economics*, Vol. 83, 134–147.

Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, Vol. 50 (4), 987–1007.

European Central Bank (2025) An overview of the ECB's monetary policy strategy. <https://www.ecb.europa.eu/mopo/strategy/strategy-review/ecb.strategyreview202506_strategy_overview.en.html>, retrieved 13 March 2026.

European Central Bank (2026) Euro short-term rate (€STR). <https://www.ecb.europa.eu/stats/financial_markets_and_interest_rates/euro_short-term_rate/html/index.en.html>, retrieved 13 March 2026.

Fleming, J. – Kirby, C. – Ostdiek, B. (2001) The economic value of volatility timing. *The Journal of Finance*, Vol. 56 (1), 329–352.

Frazzini, A. – Pedersen, L. H. (2014) Betting against beta. *Journal of Financial Economics*, Vol. 111 (1), 1–25.

Freeman, J. – Anderson, D. – Sweeney, D. – Williams, T. – Shoesmith, E. (2017) *Statistics for business and economics*. 4th EMEA ed. Cengage Learning, London.

French, K. R. – Schwert, G. W. – Stambaugh, R. F. (1987) Expected stock returns and volatility. *Journal of Financial Economics*, Vol. 19 (1), 3–29.

Harvey, C. R. – Hoyle, E. – Korgaonkar, R. – Sargaison, S. – Taylor, M. (2018) The impact of volatility targeting. *The Journal of Portfolio Management*, Vol. 45 (1), 14–33.

Hull, J. C. (2018) *Risk management and financial institutions*. 5th ed. John Wiley & Sons, Hoboken, NJ.

Kaeck, A. – Alexander, C. (2013) Stochastic volatility jump-diffusions for European equity index dynamics. *European Financial Management*, Vol. 19 (3), 470–496.

Kirby, C. – Ostdiek, B. (2012) It's all in the timing: simple active portfolio strategies that outperform naïve diversification. *Journal of Financial and Quantitative Analysis*, Vol. 47 (2), 437–467.

Mandelbrot, B. (1963) The variation of certain speculative prices. *The Journal of Business*, Vol. 36 (4), 394–419.

Markowitz, H. (1952) Portfolio selection. *The Journal of Finance*, Vol. 7 (1), 77–91.

Moreira, A. – Muir, T. (2017) Volatility-managed portfolios. *The Journal of Finance*, Vol. 72 (4), 1611–1644.

Poon, S. H. – Granger, C. W. (2003) Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, Vol. 41 (2), 478–539.

Sharpe, W. F. (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, Vol. 19 (3), 425–442.

Sharpe, W. F. (1966) Mutual fund performance. *The Journal of Business*, Vol. 39 (1), 119–138.

Sortino, F. A. – Van der Meer, R. (1991) Downside risk. *The Journal of Portfolio Management*, Vol. 17 (4), 27–31.

Wang, F. – Yan, X. S. (2021) Downside risk and the performance of volatility-managed portfolios. *Journal of Banking & Finance*, Vol. 131, 106198.

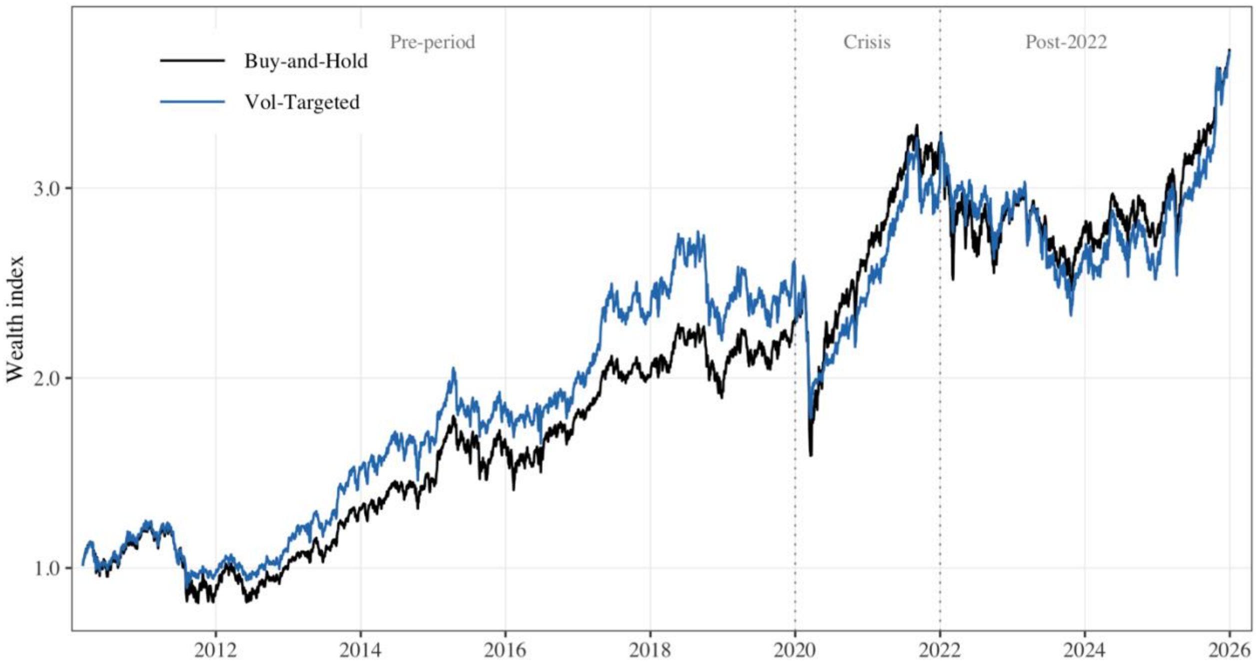
Xu, X. (2026) Beyond volatility timing: The hidden power of target timing in portfolio management. *The Journal of Portfolio Management*, Vol. 52 (2), 122–147.

Young, T. W. (1991) Calmar ratio: A smoother tool. *Futures*, Vol. 23 (1), 40.

Appendices

Appendix 1. Wealth paths in the S&P 500, with volatility-targeted series re-anchored at period breaks

Appendix 2. Wealth paths in OMXHGI, with volatility-targeted series re-anchored at period breaks



Appendix 3. Declaration of the Use of Artificial Intelligence (AI)

In the preparation of this thesis, artificial intelligence was used in a limited and supportive role. The tools were mainly used to speed up preliminary literature screening, assist with parts of the R code used on the empirical analysis, support language editing and translation, and improve clarity of selected technical expressions. I confirm that I used these tools with caution, critically evaluated all outputs, and take full responsibility for the content, methods, interpretations, conclusions, and compliance of this thesis with the requirements of the Turku School of Economics.

1. **Tool: OpenAI ChatGPT 5.2**

Stage of use: ideation, literature search, composition/editing
Purpose of use:
ChatGPT 5.2 was used to brainstorm and refine the structure of the thesis and to help identify potentially relevant academic articles more efficiently. In practice, it was used to suggest search terms, propose potentially relevant papers, and summarize individual academic articles at a preliminary level in order to assess whether they were worth reading in full. It was also used occasionally to improve the wording of selected sentences, captions, and short passages written by me. Representative example prompts: “Suggest relevant search terms and article types for a thesis on volatility targeting in European equity markets”, “Summarize the main contributions of this article and explain whether it may be relevant for volatility targeting”, and “Rewrite this figure caption in clearer academic English without changing the meaning”.

2. **Tool: Claude Sonnet 4.6**

Stage of use: data analysis and composing/editing
Purpose of use:
Claude was used primarily to support parts of the R coding process in the empirical analysis. This included help with debugging code, improving code structure, and explaining what individual parts of the script were doing. Some of the R code used in the thesis was developed with the assistance of AI tools, but the final code was tested, modified, and interpreted by me (the author). Claude was also used occasionally to suggest clearer wording for some phrases, and it also helped alongside ChatGPT with brainstorming and outlining the thesis structure. Representative example prompts: “Help identify why this R code produces unstable leverage values” and “Explain what this section of R code does step by step”.

3. **Tool: Google Gemini 3 Pro**

Stage of use: literature search, composing/editing
Purpose of use:
Gemini was used in a limited support role to speed up the initial screening of relevant literature, to assist with language-level improvements in selected passages, to help alphabetize the reference list, and to support the translation of the Finnish version of the abstract text from the English original. It was also used in some cases to help phrase short technical expressions more clearly. Representative example prompts: "Summarize this article briefly and explain whether it is relevant for volatility targeting" and "Translate this abstract text into Finnish and keep the academic meaning unchanged".

4. **Tool: DeepL**

Stage of use: composing/editing
Purpose of use:
DeepL was used for translation support in some parts during the writing process, especially when translating more complex expressions between Finnish and English. It was also used in the preparation of the Finnish version of the abstract from the English original. Representative example prompts: "Translate this expression from Finnish into English".

5. **Additional AI-assisted language support**

Stage of use: composing/editing
Purpose of use:
AI tools were also used in a limited way for proofreading and language refinement. This included correcting spelling mistakes, improving grammar, and helping phrase a few specific expressions more clearly. One example was the wording of a figure caption, that explained how the volatility-targeted wealth series was re-anchored to the contemporaneous level of the buy-and-hold wealth index at period breaks.

Overall, AI tools were used only as support tools in the preparation of this thesis. They were not used to generate the core arguments, final conclusions, or the analytical contribution of the thesis. I emphasize and confirm that I have used AI with caution, and I take full responsibility for the whole bachelor's thesis. The author holds on to full responsibility for the content and its alignment with the thesis requirements and guidelines of the Turku School of Economics.