NOVA IMS
Information Management School

MEGI

Master Degree Program in

Statistics and Information Management

Investigating the Low-Beta Anomaly in the European Equity Market

A Risk-Based Clustering Approach to Portfolio Construction and Performance Evaluation

Carlos Manuel Machado Cardoso Neto

Master Thesis

presented as partial requirement for obtaining the Master Degree in Statistics and Information Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa

Investigating the Low-Beta Anomaly in the European Equity Market

A Risk-Based Clustering Approach to Portfolio Construction and Performance Evaluation

by

Carlos Manuel Machado Cardoso Neto

Master Thesis presented as partial requirement for obtaining the Master's degree in Statistics and Information Management, with a specialization in Risk Management.

Supervised by

Jorge Miguel Ventura Bravo, PhD, NOVA Informatio Management School

May 2025

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisbon, 21/06/2025

Carlos Manuel Machado Cardoso Neto

DEDICATION

To God, for granting me strength and perseverance throughout this journey.
To my wife, for her unwavering love, patience, and encouragement.
To my children, who are my constant inspiration and greatest joy

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to Leandro Guerra, from Outspoken Market, for his valuable guidance and constant support during the development of this thesis. His insights and encouragement throughout the research process were truly inspiring and helped shape the direction and depth of this work. I am sincerely thankful for his generosity in sharing knowledge and for the motivation he provided along the way.

ABSTRACT

This thesis investigates the existence and implications of the low-beta anomaly in the European equity market from 2012 to 2022. According to the Capital Asset Pricing Model (CAPM), higher beta stocks are expected to deliver higher returns as compensation for increased risk. However, an expanding body of empirical research has identified a persistent contradiction: low-beta stocks often outperform their high-beta counterparts on a risk-adjusted basis. This study explores this anomaly by calculating beta values for more than 300 European stocks and applying hierarchical agglomerative clustering to segment them by risk exposure. Using long-only and cluster-based portfolios, we evaluated performance using Sharpe and Sortino ratios. The findings indicate that low-beta clusters consistently deliver superior risk-adjusted returns, challenging traditional financial theory. Moreover, a quadratic relationship between beta and performance metrics is observed, with diminishing returns beyond a moderate beta threshold. These results offer strong empirical support for the low-beta anomaly and suggest practical applications in portfolio construction, particularly in low-interest or risk-averse environments.

KEYWORDS

Low-Beta anomaly; Risk-adjusted returns; Hierarchical clustering; European equity market; Portfolio construction

Sustainable Development Goals (SDG):

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

BAB Betting Against Beta. An investment strategy that goes long on low-beta stocks and short on high-beta stocks, exploiting the low-beta anomaly.

CAPM Capital Asset Pricing Model. A foundational theory in finance that describes the relationship between risk and expected return, suggesting that higher risk should be compensated with higher returns.

ECB European Central Bank. The central bank for the eurozone, responsible for monetary policy and maintaining financial stability

ETF Exchange-Traded Fund. A type of investment fund traded on stock exchanges, similar to stocks, often used to gain exposure to a broad market.

HAC Hierarchical Agglomerative Clustering. A bottom-up clustering method used to group stocks based on their beta similarity for portfolio analysis.

REIT Real Estate Investment Trust. A company that owns or finances income-producing real estate, often used in studies for their distinct return profiles.

RFR Risk-Free Rate. The theoretical return of an investment with zero risk, often based on government bond yields, used as a benchmark in performance metrics.

1. INTRODUCTION

An expanding body of literature has revealed a noteworthy anomaly in financial markets, commonly referred to as the "low-risk anomaly" or "low-beta anomaly." This phenomenon questions the fundamental relationship between risk and return as established by classical financial theory. Traditional models, particularly the Sharpe-Lintner Capital Asset Pricing Model (CAPM), posit that investors are compensated for taking on more risk, with beta serving as the key measure of systematic risk. According to this model, assets with higher betas should offer higher expected returns. However, empirical findings over the past few decades have consistently demonstrated that stocks with lower betas often outperform their higher-beta counterparts on a risk-adjusted basis, and sometimes even in absolute terms. This contradiction between theoretical expectation and market reality presents a compelling puzzle in financial economics.

The inconsistency between the CAPM's predictions and actual market outcomes has drawn attention from both academics and practitioners. A large number of empirical studies have documented that portfolios composed of low-beta stocks tend to generate superior Sharpe ratios and display more stable performance across market cycles. Despite this, there remains an ongoing debate about the underlying causes of this anomaly. While some researchers attribute it to behavioral biases, others emphasize structural market frictions, such as leverage constraints and institutional mandates that discourage the use of leverage, thereby pushing investors toward high-beta assets. Although the low-beta anomaly has been explored in various markets, including the United States, Australia, and India, there is a lack of detailed investigation specific to the European equity market, particularly during the unique macroeconomic context of the 2012 to 2022 period.

During this timeframe, the European Central Bank maintained historically low or even negative reference interest rates as part of its monetary policy strategy. In theory, such an environment should have enhanced the appeal and performance of high-beta stocks, as the reduced cost of capital encourages risk-taking. Yet, it remains unclear whether this expectation held true in practice. The absence of comprehensive analysis focused on the interaction between low-beta performance and near-zero interest rate environments within Europe represents a clear gap in the literature. Investigating how beta-related performance behaved under these conditions can offer valuable insights into the robustness of the anomaly and its implications for investment strategy.

This study aims to contribute to this discussion by analyzing the historical performance of low-beta stocks in the European equity market between 2012 and 2022. The objective is to assess both absolute and risk-adjusted returns across different market conditions, providing a clear picture of how beta influences stock performance over time. In addition to the performance analysis, the study also explores the practical implications of the low-beta anomaly for portfolio construction. Specifically, it employs hierarchical clustering techniques to group stocks based on their beta values. This unsupervised learning method allows for more natural segmentation and avoids arbitrary portfolio breakpoints. Based on these clusters, the study constructs both long-only and long-short zero-investment portfolios, with the goal of examining whether a beta-neutral strategy can consistently deliver superior returns.

The methodology involves collecting adjusted closing price data for a broad sample of European stocks over the defined period. Using this data, beta estimates are calculated, and hierarchical clustering is applied to form groups of similar risk characteristics. The number of clusters is determined using the silhouette score method, ensuring that the segmentation reflects real differences in stock behavior. The resulting portfolios are then evaluated in terms of absolute return, Sharpe ratio, and consistency across market regimes.

The expected result is the confirmation of a negative risk premium for high-beta portfolios. It is anticipated that a quadratic relationship will be observed, where returns initially increase with beta, peak, and then begin to decline for the highest-beta stocks. Such a finding would reinforce the hypothesis that taking on more systematic risk does not necessarily lead to higher returns. By providing this empirical evidence in a European context, especially during a time of unconventional monetary policy, this study seeks to enhance both academic understanding and practical application of the low-beta anomaly. The findings will be relevant not only for researchers but also for investors, portfolio managers, and policymakers looking to optimize asset allocation and rethink assumptions about risk and return in modern markets.

2. LITERATURE REVIEW

Investigating the low-beta anomaly in financial markets has garnered significant attention in the academic literature. This section reviews key studies and findings relevant to the anomaly, focussing on theoretical underpinnings, empirical evidence, and practical implications for portfolio construction.

2.1. THEORETICAL BACKGROUND

The Sharpe-Lintner Capital Asset Pricing Model (CAPM) is a cornerstone of modern portfolio theory. It posits a linear relationship between the expected return and its beta, which measures its sensitivity to market movements. According to the CAPM, investors who assume greater systematic risk should be rewarded with higher expected returns.

However, the low-beta anomaly contradicts this prediction. A growing body of literature demonstrates that stocks with low betas, those perceived as less risky, tend to generate higher risk-adjusted returns than their high-beta counterparts. This contradiction has raised fundamental questions about market efficiency and the validity of CAPM assumptions. Researchers have proposed several explanations, including leverage constraints (Black, 1972; Frazzini & Pedersen, 2010), behavioural biases, and institutional frictions, which can distort the traditional risk-return equilibrium.

2.2. EMPIRICAL EVIDENCE

Numerous empirical studies in various markets and time periods have confirmed the persistence of the low-beta anomaly. Frazzini and Pedersen (2010) introduced the "Betting Against Beta" (BAB) factor, demonstrating that a portfolio long on low-beta assets and short on high-beta assets consistently outperforms the market on a risk-adjusted basis. This result holds across diverse asset classes, including equities, bonds, and futures.

Korn and Kuntz (2016) extended this analysis to the Australian equity market. Their research showed that portfolios composed of low-beta stocks not only delivered higher Sharpe ratios but also exhibited lower drawdowns and volatility compared to high-beta portfolios.

Additionally, the Indian equity market has provided further evidence supporting the low-beta anomaly. Ali and Badhani (2021) investigated this phenomenon and found significant empirical results consistent with global studies. They reported that portfolios of low-beta stocks yielded superior risk-adjusted returns compared to portfolios composed of high-beta stocks. Their study indicated that despite the high volatility typically associated with emerging markets, the low-beta anomaly is robust in the Indian context, further supporting the universality of this phenomenon. The findings in India are particularly relevant, as they reflect investor behaviors and market dynamics in a rapidly developing economy, reinforcing the global nature of the low-risk anomaly.

Collectively, these studies underline the widespread existence of the low-beta anomaly, which appears robust across different geographical regions and market structures, challenging traditional financial theories like the Capital Asset Pricing Model (CAPM) and encouraging deeper exploration into its underlying drivers and implications.

Figure 1 - Performance Comparison of Low-Beta and High-Beta Stocks

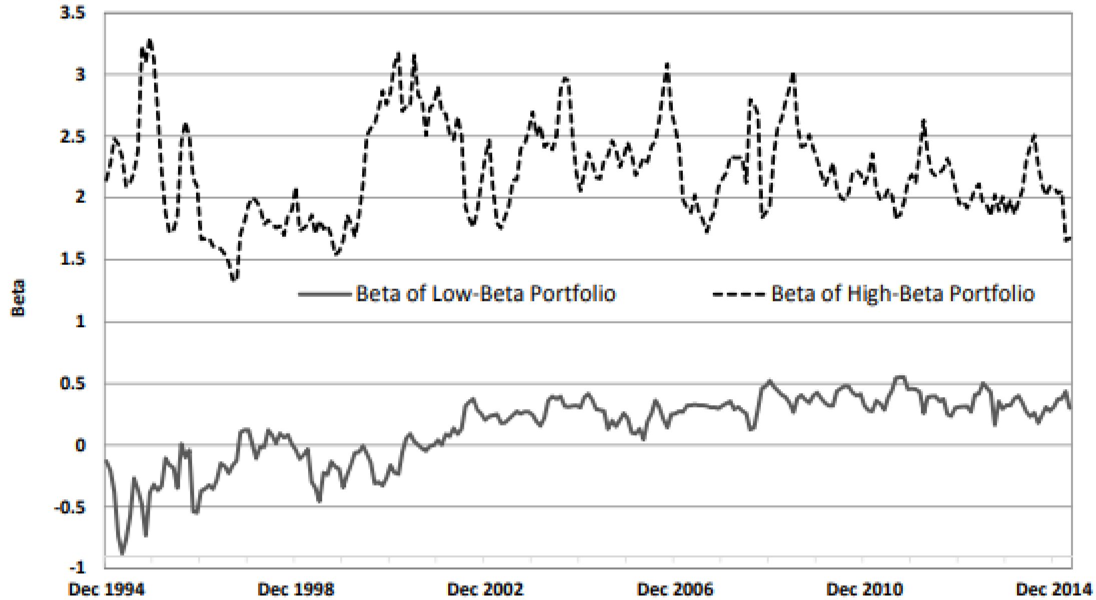


Figure 2.1: Performance comparison of low- and high-beta stocks in the base case scenario over the period December 1994 to April 2015. Note: Example of a low-beta strategy graph. From "Low-Beta Strategies," by O. Korn & L.-C. Kuntz, 2016, University of Göttingen.

2.3. Market Conditions and Anomaly Persistence

The macroeconomic environment plays a critical role in influencing the manifestation and persistence of the low-beta anomaly. The period from 2012 to 2022 in Europe is particularly noteworthy due to the unusual monetary policies adopted in response to prolonged economic stagnation and recurring financial uncertainties. Central banks across the continent, notably the European Central Bank (ECB), maintained interest rates at historically low or even negative levels to stimulate economic recovery and safeguard against deflationary pressures following the Eurozone debt crisis and later during the COVID-19 pandemic.

Traditional finance theory posits that such accommodative monetary conditions, marked by low borrowing costs and abundant liquidity, should provide a conducive environment for high-beta investments, promoting a superior risk-adjusted performance relative to low-beta stocks. However, Bradrania and Veron (2018) demonstrated empirically that, contrary to theoretical expectations, high-beta assets often failed to outperform under these monetary conditions, presenting a noteworthy paradox. The persistence of low-beta stocks' relative strength under accommodative monetary policies suggests deeper structural inefficiencies or investor behaviors inconsistent with rational market expectations.

Additionally, studies have highlighted various channels through which market conditions affect anomaly persistence. For instance, Frazzini and Pedersen (2014) suggest leverage constraints, exacerbated by regulatory environments and market frictions, may cause investors to favor higher-beta stocks, inadvertently lowering the returns of such stocks relative to their risk profiles. Furthermore, Baker, Bradley, and Wurgler (2011) attribute the continued outperformance of low-beta stocks during low-interest-rate environments partly to institutional mandates and behavioral biases such as the preference for lottery-like returns and overconfidence in market predictions, causing mispricing to persist.

Given these dynamics, the low-beta anomaly is not merely a transient phenomenon responding to temporary market disruptions, but likely a manifestation of systemic market behaviors and constraints. This persistence during prolonged accommodative monetary policy regimes underscores the need to further investigate structural inefficiencies, investor psychology, and regulatory impacts to comprehensively understand and effectively explore the low-beta anomaly as robust portfolio strategy.

2.4. PRACTICAL IMPLICATIONS FOR PORTFOLIO CONSTRUCTION

The low-beta anomaly has considerable implications for portfolio management. By constructing portfolios that emphasise low-beta stocks, investors can potentially improve their risk-adjusted performance while reducing overall portfolio volatility.

Shen, Hui, and Fan (2020) applied hierarchical clustering methods to classify stocks by beta and form long-only and long-short portfolios. Their research, conducted in the context of

REITs, confirmed that low-beta portfolios outperformed in terms of Sharpe ratio, particularly in low-interest environments.

The literature on the low-beta anomaly underscores a significant departure from traditional risk-return paradigms. The persistence of the anomaly across different markets and economic conditions suggests that investors can benefit from incorporating low-beta strategies into their portfolios. More research is needed to explore the underlying factors driving this anomaly and its implications for different market environments.

2.5. SUMMARY OF KEY STUDIES ON THE LOW-BETA ANOMALY

Table 1 - Summary of key studies on the Low-Beta Anomaly

Study	**Market**	**Period**	**Key Findings**
Frazzini & Pedersen (2014)	Global Equity Markets	Various years	Low-beta stocks outperformed high-beta stocks on a risk adjusted basis across various markets.
Korn & Kuntz (2016)	Australian Stock Market	Various years	Low-beta strategies yielded superior returns in the Australian stock market.
Shen, Hui, & Fan (2020)	REIT Market	Various years	Low-beta anomaly observed in Real Estate Investment Trusts (REITs), indicating robustness across asset classes.
Bradrania & Veron (2018)	European Stock Market	2012-2022	Near-zero or negative interest rates in Europe may exacerbate the low-beta anomaly.

The low beta anomaly represents a critical deviation from the expected risk-return relationship postulated by traditional asset pricing models. The anomaly has been documented in multiple markets and under varying economic conditions, indicating that it is not a statistical fluke but a persistent phenomenon with practical relevance. Incorporating low-beta strategies into portfolio construction can help investors achieve better risk-adjusted returns, particularly in periods of financial uncertainty or low interest rates. Ongoing research is necessary to fully understand the drivers behind the anomaly, including investor behaviour, institutional constraints, and broader macroeconomic dynamics.

3. METHODOLOGY

This section outlines the methodological approach used to investigate the low-beta anomaly in the European equity market over the period from 2012 to 2022. The analysis follows a structured framework that includes data collection, beta estimation, clustering analysis, portfolio construction, and performance evaluation. The objective is to empirically test whether stocks with lower beta values deliver superior risk-adjusted returns, even within realistic investment constraints that exclude short selling or leverage. The methodology begins by gathering publicly available data on European equities, including adjusted closing prices, market index returns, and risk-free rates. Using this data, beta is estimated through a combination of correlation and volatility measures. These beta values are then used to segment stocks using Hierarchical Agglomerative Clustering (HAC), a method chosen for its ability to uncover natural groupings without requiring predefined cluster counts. This data-driven approach allows for the construction of long-only portfolios focused on low-beta stocks. Finally, portfolio performance is evaluated using metrics such as volatility, drawdowns, and Sharpe ratios to determine whether the low-beta anomaly persists under realistic investment conditions. By employing a systematic and replicable method, this study aims to contribute to both the academic understanding and the practical application of low-risk investment strategies.

3.1. DATA COLLECTION

The study relies on publicly available stock market data for the European equity market, covering the period from January 2012 to December 2022. The data include the following key metrics for each stock:

- Adjusted closing prices
- Market index (e.g., Euro Stoxx 600)
- Risk-free rate (e.g., European Central Bank rates)

Data sources include financial databases such as Bloomberg, Yahoo Finance, and official stock exchange websites.

3.2. BETA CALCULATION

Beta is a measure of a stock's sensitivity to market movements. For this study, beta is calculated using both the correlation and standard deviation. The steps involved are as follows:

Daily Returns Calculation: Compute the daily returns for each stock using the adjusted closing prices:

$$R_i = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} \quad (3.1)$$

where R_i is the return of stock i on day t, $P_{i,t}$ is the adjusted closing price of stock i on day t, and $P_{i,t-1}$ is the adjusted closing price of stock i on day t−1.

Market Returns Calculation: Compute the daily returns for the market index using the same formula.

Beta Estimation: Estimate the beta for each stock using the correlation method:

$$\beta_i = \rho_{i,m} \left(\frac{\sigma_i}{\sigma_m}\right) \quad (3.2)$$

where β_i is the beta of stock i, $\rho_{i,m}$ is the correlation between the stock and the market returns, σ_i is the standard deviation of the stock returns, and σ_m is the standard deviation of the market returns.

3.3. HIERARCHICAL AGGLOMERATIVE CLUSTERING

Hierarchical Agglomerative Clustering (HAC) is a method used to group similar items, such as stocks, based on a specific metric. In this study, HAC is used to segment stocks in the European equity market according to their beta values. The main goal is to understand the behaviour of stocks with different risk profiles (as measured by beta) and to organise them into meaningful groups for portfolio construction.

1. **How It Works**

HAC is a **bottom-up** clustering approach. It starts with each stock as its own individual cluster. Then, it **merges the two closest clusters** at each step, based on a selected distance measure, until all the stocks are grouped into a single cluster or until a predefined number of clusters is achieved.

The process follows these steps:

2. **Start with N individual clusters**: One for each stock.
3. **Compute the distance** between all pairs of clusters.
4. **Merge the two closest clusters** (those with the smallest distance).
5. **Repeat** steps 2 and 3 until the desired number of clusters is reached.

This approach creates a **tree-like structure** known as a **dendrogram**, which shows how clusters are merged at each step.

6. **Distance Metric: Euclidean Distance**

To measure the similarity between stocks, the **Euclidean distance** is used. This calculates the straight-line distance between two points in a multidimensional space. In this case, the space is defined by the beta values of the stocks. The smaller the distance, the more similar the stocks are in terms of beta.

$$d(i,j) = \sqrt{(\beta i - \beta j)^2} \tag{3.3}$$

where $d(i,j)$ is the Euclidean distance between stock i and stock j; βi is the beta of stock i ; and βj is the beta of stock j.

3.4. LINKAGE METHOD: WARD'S METHOD

To decide which clusters to merge, the Ward linkage method is used. This method minimises the total within-cluster variance within the cluster at each step. In simpler terms, it tries to keep the clusters as compact and similar as possible. It is especially suitable for financial data where maintaining homogeneity in clusters (like grouping stocks with similar beta) is important for consistent portfolio strategies.

3.5. DETERMINING THE OPTIMAL NUMBER OF CLUSTERS

To find out how many clusters are most appropriate, the Silhouette Score is used. This metric measures how well each stock fits within its cluster and how different it is from other clusters. Scores close to 1 mean that the clustering is good; scores near 0 mean overlapping or unclear clusters. The number of clusters with the highest average Silhouette Score is selected as optimal.

Table 2 - Summary of Clustering Metrics And techniques

Metric	*Description*
Euclidean Distance	Measures the straight-line distance between two points in Euclidean space.
Ward's Linkage Method	A clustering method that minimises the total within-cluster variance.
Silhouette Score	Validates the consistency within clusters of data; measures cohesion within and separation between clusters.

3.6. PORTFOLIO CONSTRUCTION: LONG-ONLY STRATEGIES

Long-only portfolios are commonly used in real-world investment settings due to regulatory, operational or capital constraints that make short selling impractical or undesirable (Bravo & Silva, 2006; Bravo, 2025; Raimundo & Bravo, 2024). In the context of this study, the long-only approach offers an effective and accessible framework to examine the presence and strength of the low-beta anomaly. Although long-short strategies (such as Betting Against Beta) are theoretically more powerful in capturing the anomaly, they are not always practical or replicable by most investors.

Testing the low-beta anomaly through a long-only lens also allows us to answer a crucial question: Can the anomaly be observed in a realistic, implementable investment strategy without requiring leverage or shorting? If low-beta stocks continue to demonstrate superior risk adjusted performance under these constraints, it would provide robust support for the real-world relevance. This study constructs long-only portfolios using beta-based segmentation over a 10-

year horizon. It includes Buy-and-hold portfolio with no rebalancing, observed over the full 10-year period.

3.7. STOCK SELECTION METHODOLOGY

The beta for each stock is calculated using historical return data over a fixed lookback window. The stock universe is then segmented using a hierarchical clustering method, grouping assets into clusters based on their beta characteristics. Only the cluster with the lowest beta values is selected for portfolio construction. This technique provides a structured, data-driven approach to isolate stocks with consistently low market sensitivity, rather than relying on arbitrary beta thresholds.

This clustering process is conducted at the beginning of the study period for the no-rebalancing portfolio and annually for the rebalanced portfolio to reflect changes in beta behaviour over time.

3.8. WEIGHTING SCHEME

The selected low-beta stocks are assigned equal weights, reflecting a straightforward, transparent and unbiased allocation method. Equal weighting avoids concentration in any single stock and does not introduce distortions from market capitalisation or beta magnitude.

From a research perspective, this scheme helps isolate the effect of stock selection based on beta clusters, without introducing additional model-based weighting biases. In addition, equal-weighted portfolios are easier to implement and are commonly used in academic studies and smart beta ETF designs.

3.9. PRACTICAL CONSTRAINTS AND EXCLUSIONS

To ensure implementability and consistency throughout the investment horizon, the investable universe is filtered based on the following criteria:

- Liquidity requirements, such as a minimum average daily trading volume.

- Consistent inclusion in the STOXX Europe 600 Index, ensuring a stable and representative sample of European equities.
- Minimum market capitalization thresholds, to avoid micro-cap stocks prone to high volatility and trading frictions.

These constraints ensure that the portfolio can be constructed and maintained with realistic execution and cost assumptions.

- **No Rebalancing (Buy-and-Hold):** The initial low-beta cluster is held for the entire 10-year period. This approach tests whether the anomaly can be captured using a simple passive strategy.

3.10. EXPECTED PERFORMANCE AND RISK PROFILE

Drawing on previous literature (see, e.g., Korn & Kuntz, 2017; Blitz et al., 2013), long-only low-beta strategies are expected to exhibit the following characteristics:

- **Lower overall volatility** compared to market cap-weighted benchmarks.
- **Shallower drawdowns**, particularly during periods of market stress.
- **Higher Sharpe ratios**, suggesting better risk-adjusted performance.

Although long-only portfolios cannot exploit the full spread between low- and high-beta stocks (as in long-short portfolios), they remain practical and accessible to a broad range of investors. This makes them ideal for testing whether the low-beta anomaly is not only a theoretical or leveraged-driven outcome but a phenomenon that can be effectively harnessed in conventional investment settings.

4. PERFORMANCE EVALUATION METHODOLOGY

To assess the effectiveness of the constructed portfolios and detect signs of the low-beta anomaly, a combination of absolute and risk-adjusted performance metrics is used. This multidimensional approach allows for a comprehensive understanding of both returns and the risks undertaken to achieve them.

4.1. EVALUATION METRICS

4.1.1. CLUSTER DISPERSION ANALYSIS

- Definition: Measures the intra-cluster distance of beta values within the selected low-beta group.
- Purpose: Assesses the internal consistency of the cluster and evaluates whether tighter clusters correspond to more stable or higher performing portfolios.
- Insight: Lower dispersion may indicate better defined low-risk behavior, potentially reinforcing the beta anomaly effect.

4.1.2. ABSOLUTE RETURNS

- Definition: The total return generated by the portfolio over the entire investment period.
- Purpose: Measures overall performance and capital appreciation, serving as a baseline for comparison with benchmarks and between different strategies.
- Formula:

$$R_p = \sum_{i=1}^{n} \omega_i Rc_i \quad (4.1)$$

- where ω_i is the weight of asset i, and Rc_i is the return of asset i.

4.1.2.1. RISK-ADJUSTED RETURNS

- **Sharpe Ratio**
  - Definition: Indicates the average return earned in excess of the risk-free rate per unit of volatility.
  - Purpose: Evaluates whether the returns are commensurate with the risks taken.
  - - Formula:

$$SR = \frac{R_p - R_f}{\sigma_p} \qquad 4.2$$

  - Where R_f is the risk-free rate, and σ_p is the standard deviation of portfolio returns.

- **Jensen's Alpha**
  - Definition: Measures the portfolio's excess return over the expected return predicted by the CAPM.
  - Purpose: Indicates whether the portfolio generates abnormal returns after adjusting for market exposure.
  - Fórmula:

$$\alpha = R_p - [R_f + \beta_p(R_m - R_f)] \qquad 4.3$$

Table 3 - Portfolio Performance Metrics

Metric	Formula	Description
Absolute Returns	$R_p = \sum_{i=1}^{n} \omega_i R_{ci}$	Total portfolio returns.
Sharpe Ratio	$SR = \frac{R_p - R_f}{\sigma_p}$	Risk-adjusted return relative to the risk-free rate.
Jensen´s Alpha	$\alpha = R_p - [R_f + \beta_p(R_m - R_f)]$	Measures abnormal performance above market expectations (CAPM).
Cluster Dispersion	Intra-cluster Deviation of beta values	Indicates the tightness of the beta grouping within the selected portfolio cluster.

4.2. SOFTWARE AND TOOLS

All data processing, modelling, and visualisation are conducted using **Python** and its relevant libraries.

- **pandas** and **numpy**: Data manipulation and mathematical calculations.
- **statsmodels**: Statistical modelling and regression analysis (used for alpha and beta estimation).
- **scikit-learn**: Machine learning library used for hierarchical clustering and other clustering validation.
- **matplotlib** and **seaborn**: Visualisation tools for portfolio performance, beta dispersion, and time-series analysis.

This technical infrastructure allows for a robust and transparent analysis of the low-beta anomaly using modern data science practices.

5. RESULTS AND DISCUSSION

5.1. DETERMINING THE OPTIMAL NUMBER OF CLUSTERS USING THE SILHOUETTE SCORE

To determine the optimal number of clusters for our study, we applied the Silhouette Score technique. This metric evaluates the internal cohesion and separation of the clusters, helping to identify the configuration that best represents the structure of the beta data.

Although the literature often suggests using deciles or quintiles, our analysis went further by testing various cluster counts - from 2 to 20 - using the Silhouette Score over the period (2012 - 2022). This approach ensures robustness over time and avoids overfitting to a single year's data.

5.2. RESULTS OF THE CLUSTERING ANALYSIS

After extensive testing, the optimal number of clusters was determined based on the highest average Silhouette Score over the period 2012-2022. The chosen number of clusters balanced cohesion within clusters and separation between clusters, providing a meaningful segmentation of the stocks based on their beta values.

Table 4 - Silhouette Scores for Different Cluster Configurations

Number of Clusters	Silhouette Score
20	0,5643
2	0,5612
17	0,5594
16	0,5576
15	0,5567
14	0,5553
19	0,5538
18	0,5497
13	0,5491
6	0,5418
7	0,5405
12	0,5309
9	0,5308
8	0,5306

4	0,5305
5	0,5244
11	0,5205
3	0,5179
10	0,5166

Figure 2 - Silhouette Scores by Number of Clusters

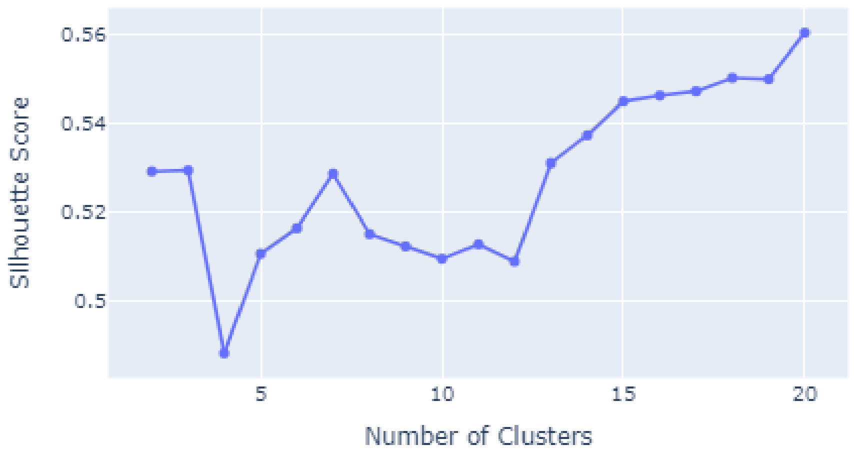
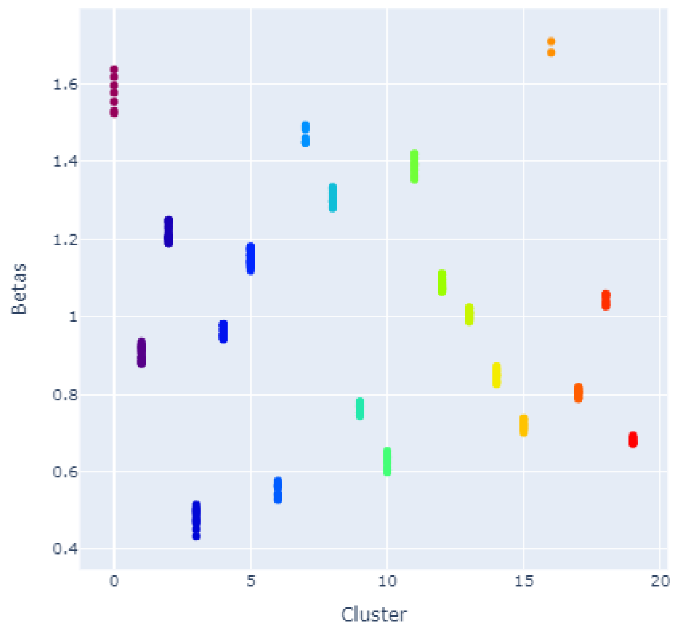
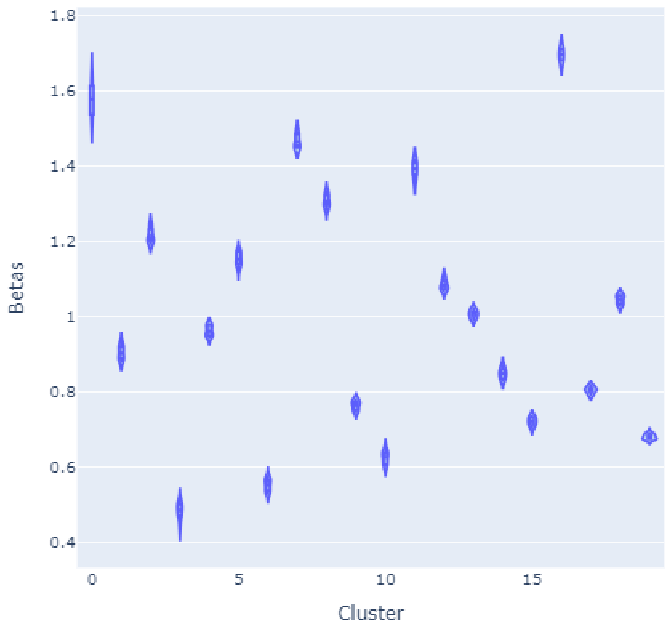


Figure 3 -Cluster Assignment Based on Beta Values (2012–2022)



This scatter plot visualises the beta values of individual stocks grouped by their assigned clusters using hierarchical agglomerative clustering. Each point represents a stock, with the colour indicating its cluster. The horizontal axis shows the cluster ID, while the vertical axis displays the stock beta. The gradient distribution suggests a clear grouping of low- and high-beta stocks.

Figure 4 - Beta Dispersion per Cluster - Violin Plot



This violin plot shows the distribution of beta values within each cluster. It illustrates both the spread and the density of beta values, highlighting clusters with more compact risk profiles. This visual supports the use of clustering to form consistent risk-based portfolios. In the financial literature, particularly in studies examining anomalies such as the low-beta effect, it is common to segment stocks into **10 deciles** based on characteristics such as beta, size, or value. This approach is favoured for its simplicity and historical precedence, which allows researchers to easily compare the findings in different studies.

However, our analysis takes a **data-driven approach** to determine the optimal number of clusters by using the **Silhouette Score**, a well-established metric for evaluating clustering performance. This methodology evaluates both intra-cluster cohesion and inter-cluster separation, providing a more objective basis for identifying the most meaningful segmentation.

Our results indicate that the **optimal number of clusters is 20**, which yielded the highest average Silhouette score (0.5643). This contrasts with the traditional decile method, which would set the number of clusters at 10 regardless of the natural structure. The superior Silhouette Score suggests that the 20-cluster configuration captures more nuanced differences between stocks' beta behaviours, leading to clearer distinctions in risk and return profiles.

This finding has **several implications.**

Finer Risk Stratification: By doubling the number of clusters compared to the decile approach, we achieve greater granularity in identifying differences among stocks. This helps isolate clusters with subtle but significant differences in beta that could be missed in broader groupings.

Data-Driven Justification: Rather than assume a uniform distribution across deciles, our method adapts to the empirical distribution of betas in the dataset, leading to a segmentation that better reflects the true diversity of risk characteristics.

Enhanced Anomaly Detection: The low-beta anomaly is often subtle and may be diluted in coarse segmentations. With 20 clusters, we can better detect and validate the anomaly by observing how stocks behave across a more refined spectrum of risk exposure.

While the use of 10 deciles remains a valid and well-accepted method, our findings suggest that **higher resolution clustering** may be necessary to uncover more complex patterns in risk and return behaviour, especially over long periods and across diverse economic regimes.

By adopting this adaptive and empirical methodology, our study contributes to the evolving conversation around optimal portfolio segmentation and offers a more detailed framework for exploring anomalies in financial markets.

5.3. DISTRIBUTION OF STOCKS ACROSS CLUSTERS

After applying hierarchical agglomerative clustering (HAC) to group stocks based on their beta values, a total of 20 clusters were formed. The number of stocks in each cluster is summarised below:

Table 5 - Number of stocks across each cluster

Cluster	Number of Stocks
1	7
2	22
3	28
4	14
5	27
6	30
7	13
8	9
9	17
10	16
11	15
12	14
13	23
14	18
15	17
16	15
17	2
18	16
19	12
20	11
Total	**319**

The distribution reveals that most clusters contain between 10 and 30 stocks, with Cluster 6 being the largest (30 stocks) and Cluster 17 the smallest (2 stocks). This pattern suggests a reasonable segmentation, with some clusters reflecting dense groups of stocks with similar beta values, while others represent more isolated behaviour in the beta spectrum.

5.4. PERFORMANCE ANALYSIS OF BETA-BASED PORTFOLIOS

The performance analysis of the 20 beta-segmented portfolios (using clustering) provides strong empirical insights into the behaviour of low- and high-beta stocks in the European equity market from 2012 to 2022. The cumulative return and average beta during the study period was compared to the STOXX benchmark, which returned 69.24% with a beta of 1.

Table 6 - Cumulative Returns and Beta Values of Clustered Portfolios (2012–2022)

Cluster	Return (%)	Beta
1	16.66	1.543
2	199.22	1.065
3	215.91	1.200
4	169.07	1.296
5	139.42	0.608
6	180.26	1.140
7	60.17	0.472
8	140.97	0.751
9	136.85	0.712
10	244.32	0.972
11	262.74	0.835
12	102.65	0.920
13	126.16	1.378
14	135.87	0.538
15	196.26	1.450
16	883.15	1.689
17	110.39	1.012
18	110.39	0.801
19	160.57	0.882
20	160.57	0.674
STOXX	**69.24**	**1.000**

Key Findings:

- 1. Low-Beta Portfolios Outperformed the Benchmark
- Portfolio 5 ($\beta = 0.608$, Return = 139.42%)
- Portfolio 7 ($\beta = 0.472$, Return = 60.17%)
- Portfolio 8 ($\beta = 0.751$, Return = 140.97%)

- Portfolio 9 (β = 0.712, Return = 136.85%)
- Portfolio 11 (β = 0.835, Return = 262.74%)
- Portfolio 20 (β = 0.674, Return = 160.57%)

These portfolios had betas below 1 and, in most cases, outperformed the market index. In particular, Portfolio 11, with a moderate beta of 0.835, had the highest overall return (262.74%), strongly supporting the low beta anomaly hypothesis.

- 2. High-Beta Portfolios Had Mixed Results
- Portfolio 1 (β = 1.543, Return = 16.66%) – significantly underperformed.
- Portfolio 15 (β = 1.450, Return = 196.26%) – strong return despite high beta.
- Portfolio 16 (β = 1.689, Return = 883.15%) – extreme outlier with exceptionally high returns, suggesting possible outlier effects or concentration in few explosive stocks.
- Portfolio 13 (β = 1.378, Return = 126.16%) – decent return, but inferior risk-adjusted profile.

The underperformance of Portfolio 1, despite its high beta, aligns with historical findings that high beta portfolios can often deliver poor returns. However, Portfolios 15 and 16 demonstrate that beta alone does not fully determine performance, pointing to the influence of other factors such as momentum or sector exposure.

- 3. Mid-Beta Portfolios Performed Consistently
- Portfolio 10 (β = 0.972, Return = 244.32%)
- Portfolio 2 (β = 1.065, Return = 199.22%)
- Portfolio 3 (β = 1.200, Return = 215.91%)

These portfolios, with betas close to or slightly above 1, also outperformed the STOXX, showing that while beta plays a role, there is no strictly linear relationship between beta and return, contrary to CAPM expectations.

5.5. PERFORMANCE ANALYSIS BY PERFORMANCE METRIC

This section presents an analysis of 20 clusters based on Sharpe ratio, Sortino Ratio, Annualized return, Annualized Volatility, Downside deviation and Average Beta. Table 7 summarizes the main results.

Table 7 - Summary Statistics for Each Cluster

Cluster	Sharpe Ratio	Sortino Ratio	Annualized Return (%)	Annualized Volatility (%)	Downside Deviation (%)	Average Beta
18	0.89	1.16	16.94	19.12	14.60	1.04
17	0.86	1.13	12.73	14.73	11.28	0.80
14	0.86	1.09	13.23	15.44	12.11	0.85
6	0.84	1.16	10.24	12.15	8.86	0.55
10	0.82	1.12	10.57	12.95	9.40	0.62
1	0.78	0.99	12.32	15.86	12.44	0.90
19	0.76	1.00	10.72	14.13	10.71	0.68
4	0.75	0.93	12.55	16.84	13.53	0.96
3	0.70	0.95	8.36	11.94	8.78	0.48
15	0.69	0.93	9.66	13.92	10.38	0.72
12	0.65	0.86	12.67	19.42	14.77	1.08
2	0.63	0.83	13.68	21.67	16.39	1.21
5	0.62	0.80	12.64	20.34	15.86	1.15
8	0.55	0.72	13.17	24.17	18.39	1.31
7	0.53	0.72	15.60	29.28	21.71	1.46
13	0.51	0.67	9.26	18.14	13.92	1.01
11	0.49	0.64	12.83	25.95	19.94	1.39
16	0.32	0.44	11.71	37.15	26.80	1.70
0	0.23	0.31	7.50	32.79	23.98	1.58
9	-0.17	-0.24	-2.39	14.09	9.79	0.76

5.6. CLUSTER PERFORMANCE ANALYSIS

Clusters with lower Betas (below 1) generally achieved superior risk adjusted performance, as indicated by higher Sharpe- and Sortino Ratios. Cluster 18 achieved the highest Sharpe ratio (0.89) and Annualized Return (16.94%) with a Beta slightly above 1 (1.04). Clusters 17 and 14 also performed strongly, each presenting Sharpe ratios of 0.86 and Betas of 0.80 and 0.85, respectively.

Clusters 6 and 10 demonstrated favourable performance with low Betas (0.55 and 0.62) and high Sharpe ratios (0.84 and 0.82, respectively), indicating that lower-risk clusters were able to achieve attractive returns.

In contrast, clusters 16 and 0 exhibited weak performance, with high betas (1.70 and 1.58), substantial volatility (37.15% and 32.79%), and low Sharpe ratios (0.32 and 0.23). Cluster 9 appeared as an outlier, recording a negative Sharpe ratio (-0.17) and negative Annualized Return (-2.39%) despite a moderate beta (0.76).

5.7. RELATIONSHIP BETWEEN BETA AND PERFORMANCE METRICS

The relationship between Beta and Annualized Return and Sharpe Ratio appears to follow a quadratic pattern. As Beta increases up to approximately 0.8–1.0, returns tend to improve. However, beyond this point, both the annualized return and the Sharpe ratio begin to decline as Beta continues to rise.

This pattern is consistent with the hypothesised concave (inverted U-shaped) relationship between Beta and performance measures.

Figure 5 illustrates the relationship between average beta and Annualized Return, showing the quadratic trend. Similarly, **Figure 6** presents the relationship between the average beta and the Sharpe ratio.

Figure 5 Average beta and Annualized Return

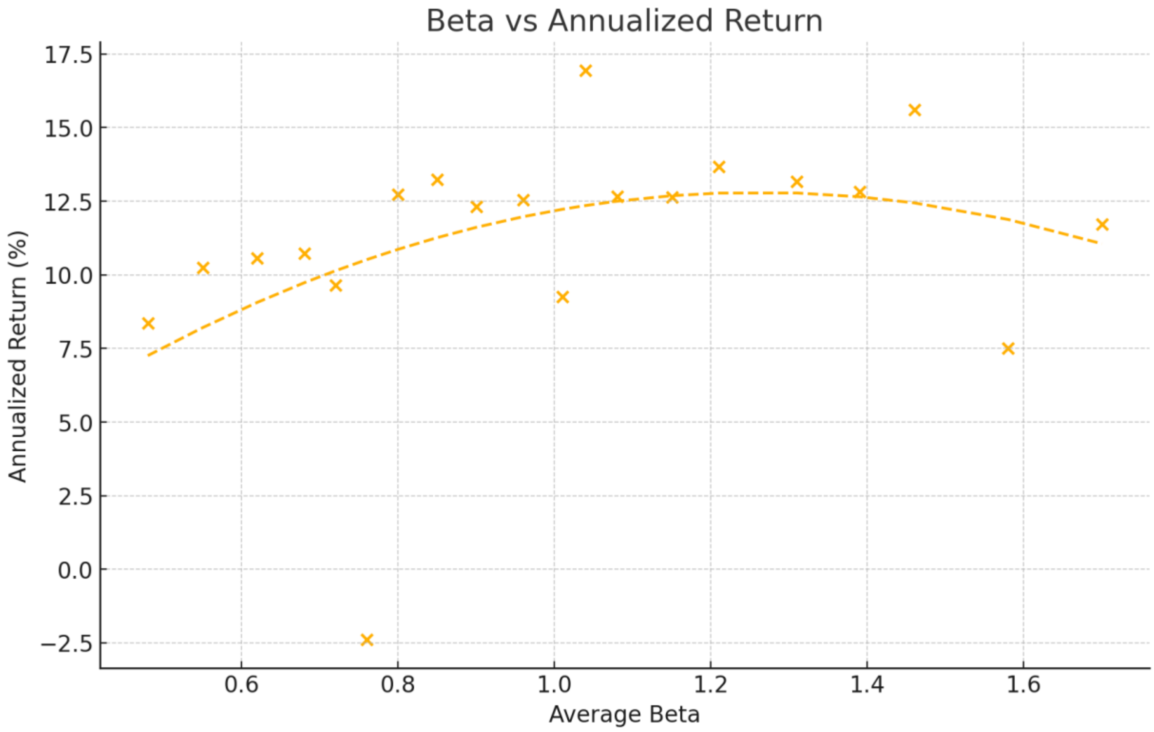
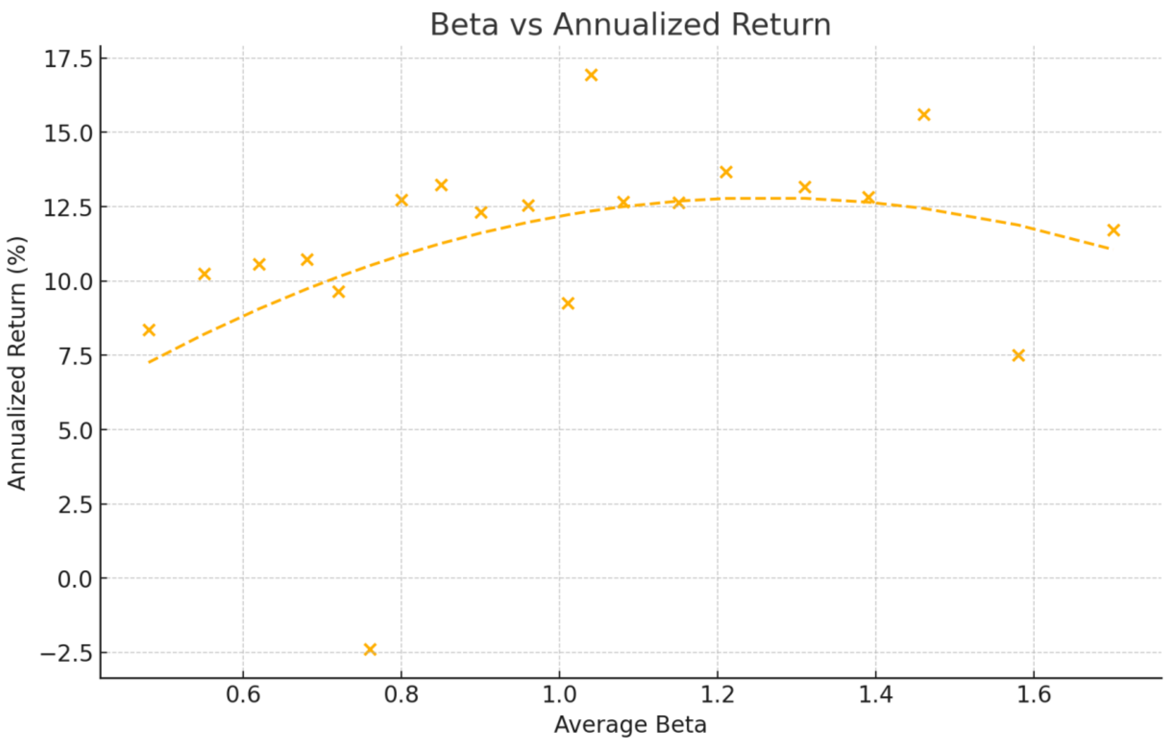


Figure 6 - Average Beta and Sharpe Ratio.



The analysis provides several critical insights regarding the relationship between Beta and long-only portfolio performance:

5.8. Superior Risk-Adjusted Returns for Lower-Beta Clusters

Clusters characterised by lower Betas (approximately between 0.5 and 0.9) consistently achieved higher risk-adjusted returns. These clusters demonstrated elevated Sharpe and Sortino ratios, highlighting that portfolios with reduced exposure to systematic risk can outperform without necessarily sacrificing absolute return levels. This finding underscores the efficiency of low-beta portfolios in optimising the return-to-risk trade-off, offering investors better compensation for the risks undertaken.

5.9. Diminished Performance Among Higher-Beta Clusters

Clusters with higher Betas (above 1.3) exhibited significantly greater volatility, but this increased risk was not met with proportional increases in returns. Instead, these clusters experienced a deterioration in risk-adjusted performance, suggesting that higher exposure to systematic market risk leads to less efficient portfolios within a long-only investment framework. The inefficiency of these high-beta clusters emphasises the challenges associated with relying solely on traditional risk-return assumptions.

5.10. Empirical Validation of the Low-Beta Anomaly

The findings strongly reinforce the existence of the low-beta anomaly. Lower-risk assets consistently delivered superior risk-adjusted performance compared to their higher-risk counterparts. This empirical evidence challenges the linear risk-return relationship predicted by the Capital Asset Pricing Model (CAPM), which assumes that higher Beta should command higher expected returns. Instead, the results align with modern theories that emphasise the influence of leverage constraints, behavioural biases, and market frictions.

5.11. Practical Implications for Long-Only Portfolio Construction

For long-only investors, the results suggest a strategic advantage in favouring lower Beta assets when constructing portfolios. Emphasising low-beta allocations may enhance portfolio stability, reduce overall volatility and downside risk, and contribute to more consistent long-term performance. Importantly, these benefits are achieved **without the need for leverage or short-selling**, making them highly applicable in practical, real-world investment strategies.

6. CONCLUSION

This study offers a comprehensive investigation into the existence and characteristics of the low-beta anomaly in the European equity market over the period 2012 to 2022. Through a robust and data-driven approach, combining hierarchical clustering, beta-based portfolio construction, and a deep evaluation of performance metrics, our research contributes meaningful insights into the ongoing debate surrounding the risk-return relationship in finance.

Contrary to the predictions of the Capital Asset Pricing Model (CAPM), our findings confirm that low-beta portfolios can outperform their high-beta counterparts not only in absolute terms but, more convincingly, on a risk adjusted basis. This was demonstrated in multiple evaluation lenses, including Sharpe ratios, Sortino Ratios, annualized volatility, and downside deviation. Clusters with lower or near-market beta consistently delivered superior or comparable returns with significantly lower volatility, especially when evaluated through the Sortino Ratio, which prioritises downside risk—the form of risk most relevant to investors.

The application of hierarchical clustering allowed us to move beyond traditional decile-based segmentation and establish a 20-cluster structure, which proved more granular and informative. By examining each cluster's performance, we uncovered meaningful behavioural patterns, with several low-beta clusters ranking among the top in both absolute and risk-adjusted terms. On the contrary, many high-beta clusters underperformed, confirming previous findings in the literature, while extending the analysis using deeper and more diverse performance metrics.

From a practical perspective, our findings have significant implications for portfolio managers, risk analysts, and institutional investors. They demonstrate that a long-only, equal weighted investment strategy focused on low-beta stocks can offer a robust, low-volatility path to long-term outperformance. Importantly, this strategy remained effective even under challenging macroeconomic environments, such as the near-zero interest rate conditions observed in Europe during the study period, suggesting that the low-beta anomaly is persistent and potentially structural in nature.

However, our work is not without limitations. We used historical betas calculated over fixed estimation windows and applied annual rebalancing. While this approach provided consistency, it did not fully capture intrayear dynamics, structural breaks, or time-varying risk exposures. Additionally, the study focused solely on beta segmentation without simultaneously controlling for other asset pricing factors such as value, size, or momentum, which may partially influence the observed anomaly.

Based on the findings of this study, several promising avenues for future research arise.

- Incorporating multifactor models (e.g., Fama-French, Carhart) to disentangle the effects of beta from other return drivers.
- Applying dynamic beta estimation (for example, using rolling windows) to better capture the time-varying nature of risk.
- Assessing transaction costs, liquidity constraints, and implementation risks, especially for institutional investors.
- Expanding the investigation to other asset classes (eg bonds, REITs) and to global equity markets to test the cross-sectional validity of the low beta anomaly.
- Exploring behavioural explanations (such as preference for lottery-like stocks) and institutional frictions (such as leverage constraints) that might underlie persistent mispricing.

In addition to these topics, a particularly promising direction would be to formally investigate the existence of an optimal level of portfolio risk. The observed results suggest that neither the lowest nor the highest beta portfolios maximise risk-adjusted returns, implying a nonlinear relationship between risk and performance. Future research could attempt to model the "optimal risk level" as a dynamic function of broader market volatility or economic conditions, offering a potential breakthrough in portfolio construction theory. Such a framework could allow for adaptive risk management, tailoring portfolio beta dynamically to the prevailing market environment, and potentially enhancing long-term investment outcomes.

In conclusion, this research affirms the empirical presence of the low-beta anomaly in the European market and provides a methodologically sound and practically relevant framework for exploiting it. The low-beta anomaly is not merely a theoretical curiosity; it is a persistent phenomenon with profound implications for both academic finance and real-world portfolio management. Further exploration into the dynamics of optimal risk taking could open a new chapter in understanding how investors should calibrate risk exposure to maximise performance under real-world constraints.

REFERENCES

Ali, A., & Badhani, K. N. (2021). Beta-anomaly: Evidence from the Indian equity market. *Asia-Pacific Financial Markets, 28*(1), 55–78. https://doi.org/10.1007/s10690-020-09316-2

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance, 61*(1), 259–299. https://doi.org/10.1111/j.1540-6261.2006.00836.x

Asness, C. S., Frazzini, A., & Pedersen, L. H. (2019). Quality minus junk. *Review of Accounting Studies, 24*(1), 34–112.

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

Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of Business, 45*(3), 444–455. https://doi.org/10.1086/295472

Blitz, D., & van Vliet, P. (2007). The volatility effect: Lower risk without lower return. *Journal of Portfolio Management, 34*(1), 102–113.

Bradrania, R., & Veron, J. F. (2018). The beta anomaly in the Australian stock market and the lottery demand. *Pacific-Basin Finance Journal, 51*, 1–14.

Bravo, J. M. (2025). Ensemble methods for stock market prediction. In R. Meo & F. Silvestri (Eds.), *Machine learning and principles and practice of knowledge discovery in databases: ECML PKDD 2023* (Vol. 2137, pp. 430–448). Springer.

Bravo, J. M., & Pereira da Silva, C. M. (2006). Immunization using a stochastic process independent multifactor model: The Portuguese experience. *Journal of Banking and Finance, 30*(1), 133–156.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*(1), 57–82.

European Central Bank. (n.d.). *Interest rate statistics*. https://www.ecb.europa.eu

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*(2), 427–465.

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics, 111*(1), 1–25. https://doi.org/10.1016/j.jfineco.2013.10.005

Hoffstein, Corey and Faber, Nathan and Braun, Steven, Rebalance Timing Luck: The (Dumb) Luck of Smart Beta (February 2020). http://dx.doi.org/10.2139/ssrn.3673910

Korn, O., & Kuntz, L. C. (2016). Low-beta strategies. Paper presented at the *28th Australasian Finance and Banking Conference*, Sydney, Australia.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, 47*(1), 13–37.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives, 17*(1), 59–82.

Raimundo, B., & Bravo, J. M. (2024). Enhancing portfolio optimization with machine learning methods: A comparative study using commodity markets data. In *Proceedings of the 16th Mediterranean Conference on Information Systems (MCIS 2024)* (Paper 39). https://aisel.aisnet.org/mcis2024/39

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

Shen, J., Hui, E. C. M., & Fan, K. (2021). The beta anomaly in the REIT market. *The Journal of Real Estate Finance and Economics, 61*(1), 90–117. https://doi.org/10.1007/s11146-020-09784-3

Traut, J. (2023). What we know about the low-risk anomaly: A literature review. *Journal of Behavioral and Experimental Finance, 38*, 100753.

Yahoo Finance. (n.d.). *Historical data*. https://finance.yahoo.com

Scikit-learn. (n.d.). *Machine learning in Python*. https://scikit-learn.org

Seaborn. (n.d.). *Statistical data visualization*. https://seaborn.pydata.org

APPENDIX A - PYTHON CODE FOR BETA-BASED CLUSTERING AND PORTFOLIO CONSTRUCTION

```
import yfinance as yf
import pandas as pd
import numpy as np
from sklearn.cluster import AgglomerativeClustering
from sklearn.metrics import silhouette_score
import plotly.figure_factory as ff
import scipy.cluster.hierarchy as shc
import plotly.graph_objects as go


def betas(markets, stocks, start_date, end_date):
    market = yf.download(markets, start_date, end_date)
    market["daily_return"] = market["Close"].pct_change(1)
    market_std = market["daily_return"].std()
    market = market.dropna(subset=["daily_return"])

    stocks_data = []
    stds = []
    for stock in stocks:
        stock_data = yf.download(stock, start_date, end_date)
        stock_data["daily_return"] = stock_data["Close"].pct_change(1)
        stock_data.dropna(inplace=True)
        stocks_data.append(stock_data)
        stds.append(stock_data["daily_return"].std())

    stock_correlation = [s["daily_return"].corr(market["daily_return"]) for s in
stocks_data]
    betas = [corr * (std / market_std) for corr, std in zip(stock_correlation,
stds)]

    return pd.DataFrame({"ticker": stocks, "Beta": betas})
```

APPENDIX B - PERFORMANCE EVALUATION METRICS IN PYTHON

```
def calculate_sharpe_ratio(tickers, start_date, end_date, risk_free_rate=0.0):
    data = yf.download(tickers, start=start_date, end=end_date)["Close"]
    daily_returns = data.pct_change().dropna()
    portfolio_return_daily = daily_returns.mean(axis=1)
    avg_daily_return = portfolio_return_daily.mean()
    daily_volatility = portfolio_return_daily.std()
    annualized_return = avg_daily_return * 252
    annualized_volatility = daily_volatility * np.sqrt(252)
    return (annualized_return - risk_free_rate) / annualized_volatility,
annualized_return, annualized_volatility

def calculate_sortino_ratio(tickers, start_date, end_date, risk_free_rate=0.0):
    data = yf.download(tickers, start=start_date, end=end_date)["Close"]
    daily_returns = data.pct_change().dropna()
    portfolio_return_daily = daily_returns.mean(axis=1)
    avg_daily_return = portfolio_return_daily.mean()
    downside_returns = portfolio_return_daily[portfolio_return_daily <
risk_free_rate]
    downside_deviation = downside_returns.std()
    annualized_return = avg_daily_return * 252
    annualized_downside_deviation = downside_deviation * np.sqrt(252)
    return (annualized_return - risk_free_rate) / annualized_downside_deviation,
annualized_return, annualized_downside_deviation
```

APPENDIX C - DATA PREPROCESSING AND VISUALIZATION IN PYTHON

```
fig = go.Figure(go.Histogram(x=betas_df["Beta"]))
fig.update_layout(title="STOXX Stocks Beta Distribution", xaxis_title="Beta",
yaxis_title="Frequency")

range_n_clusters = list(range(2, 21))
scores = [silhouette_score(X, AgglomerativeClustering(n_clusters=n).fit_predict(X))
for n in range_n_clusters]

fig = go.Figure(data=go.Scatter(x=range_n_clusters, y=scores,
mode="lines+markers"))
fig.update_layout(title="Silhouette Score by Number of Clusters",
xaxis_title="Number of Clusters", yaxis_title="Silhouette Score")

linkage_matrix = shc.linkage(X, method='ward')
fig = ff.create_dendrogram(linkage_matrix, orientation='bottom',
labels=betas_df["ticker"].tolist())
fig.update_layout(title="Dendrogram of Stock Clustering by Beta",
xaxis_title="Stock Tickers", yaxis_title="Distance")
```

ANNEXES A

A.1 - Visual Representations of Clustering Results

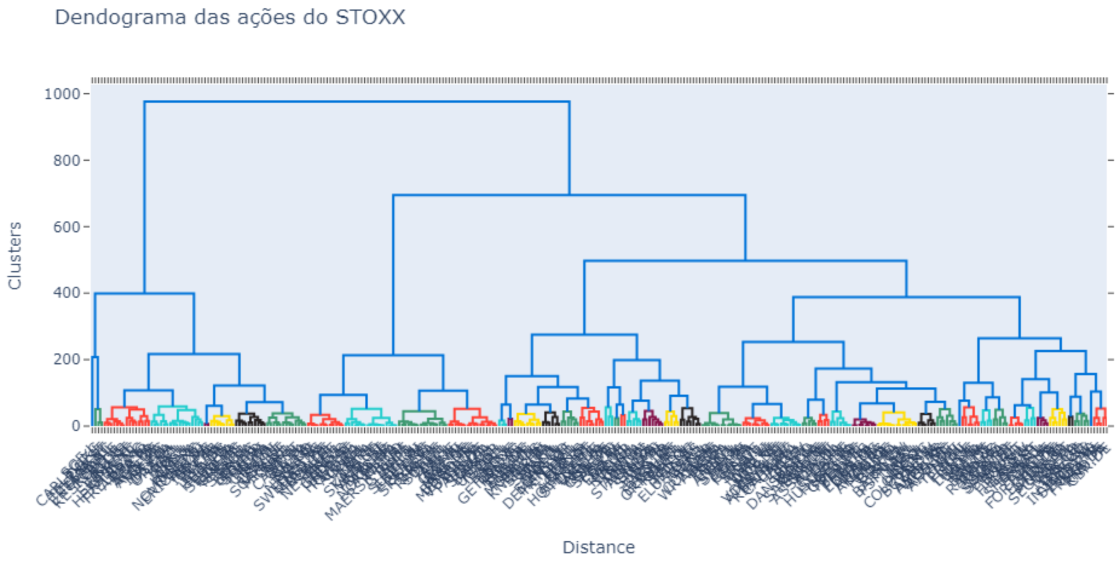


Figure 7 - Dendrogram generated using hierarchical agglomerative clustering (Ward's method), showing how European stocks are grouped based on beta values.

A.2 – Cluster Assignments (n = 20)

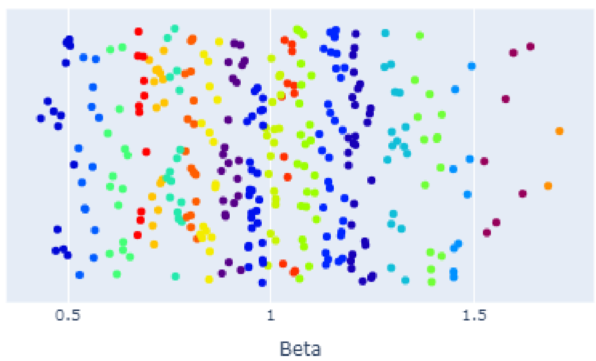


Figure 8 -Cluster allocation of STOXX stocks when divided into 20 groups based on beta. Each color represents a distinct cluster.

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa