Ca' Foscari
University
of Venice

Master's Degree programme
in Global Development and Entrepreneurship

Curriculum: Entrepreneurship

Final Thesis

Preference-Based Portfolio Optimization:

A Multiple-Criteria Decision Analysis Integrating Return, Risk and ESG

Supervisor
Ch. Prof. Maria Barbati

Graduand
Riccardo Trevisan
Matriculation Number 881602

Academic Year
2024/2025

Table of Contents

1. Introduction 1

2. Literature Review 3
   2.1 The Early Foundations of Portfolio Optimization 3
   2.2 The Evolution of Portfolio Optimization Theory 8
   2.3 The Rise of Sustainability in Portfolio Optimization 12
   2.4 MCDA Application in Sustainable Portfolio Optimization 16
   2.5 Balancing ESG Objectives and Financial Performance 25

3. Theoretical Methodology: Decision Framework 33
   3.1 Introduction 33
   3.2 Methodology 36
      3.2.1 Structuring the Problem and Data Collection 36
      3.2.2 Simulation of Investor Preferences through Dirichlet Distribution 41
      3.2.3 Segmentation of Investor Preferences via K-Means Clustering 45
         3.2.3.1 Investor Preferences Clustering 45
         3.2.3.2 Internal Validation of Clustering Solution – Silhouette Score 48
         3.2.3.3 Centroid Extraction and Preference Parametrization 50
      3.2.4 Aggregated Utility Scores Computation 52
      3.2.5 Quadratic Programming for Portfolio Optimization and Capital Distribution 54
      3.2.6 Three-Dimensional Utility Space Representation 57

4. Framework Implementation for Tri-Criterion Portfolio Optimization 61
   4.1 Introduction 61
   4.2 Practical Analysis 63
      4.2.1 Structuring the Problem and Data Collection 63
      4.2.2 Simulation of Investor Preferences through Dirichlet Distribution 68
      4.2.3 Segmentation of Investor Preferences via K-Means Clustering 70
         4.2.3.1 Investor Preferences Clustering 70
         4.2.3.2 Internal Validation of Clustering Solution - Silhouette Score 72
         4.2.3.3 Centroid Extraction and Preference Parametrization 74
      4.2.4 Aggregated Utility Scores Computation 76
      4.2.5 Quadratic Programming for Portfolio Optimization and Capital Distribution 78
      4.2.6 Three-Dimensional Utility Space Representation 84

5. Results Interpretation, Sensitivity Analysis and Future Developments 89
5.1 Empirical Results and Structural Properties of the Optimized Portfolios 90
5.2 Economic and Behavioral Interpretation of Portfolio Allocation Patterns 93
5.3 Sensitivity Analysis: Alternative Preference Weight Configurations 100
5.3.1 Structuring the Problem and Data Collection 100
5.3.2 Simulation of Investor Preferences through Dirichlet Distribution 103
5.3.3 Segmentation of Investor Preferences via K-Means Clustering 104
5.3.3.1 Investor Preferences Clustering 104
5.3.3.2 Internal Validation of Clustering Solution - Silhouette Score 105
5.3.3.3 Centroid Extraction and Preference Parametrization 106
5.3.4 Aggregated Utility Scores Computation 108
5.3.5 Quadratic Programming for Portfolio Optimization and Capital Distribution 109
5.3.6 Three-Dimensional Utility Space Representation 111
5.3.7 Cross-Scenario Comparative and Interpretative Analysis 114
5.4 Structural Extensions and Methodological Refinements 118

6. Conclusion 123

Bibliography 125

Appendix A – Data Sources and Descriptive Statistics 131

Appendix B – RStudio Script for Portfolio Optimization 145

Summary

English Version

The research develops a preference-based portfolio optimization framework grounded in a Multi-Criteria Decision Analysis (MCDA) architecture that integrates financial return, volatility risk and Environmental, Social and Governance (ESG) dimensions as explicit and structurally equivalent evaluative criteria. The framework is designed to respond to the structural inadequacy of forecast-dependent optimization approaches by configuring portfolio construction as an endogenous outcome of explicit preference alignment.

Investor heterogeneity is formally represented through compositional preference vectors defined over the Return–Risk–ESG simplex and generated via Dirichlet distributions, ensuring internal coherence, strict positivity and adherence to unit-sum constraints within a fully data-driven modeling environment. These preference structures are sequentially segmented using unsupervised K-Means clustering, yielding three stable investor groups whose centroids provide the criterion-specific priority weights. The resulting weights are synthesized into Aggregated Utility Scores (AUS), producing a transparent and traceable metric of the attractiveness of each Exchange-Traded Fund (ETF), constituting the investment universe, according to investor preferences. The portfolio construction process is formalized through a constrained Quadratic Programming (QP) optimization procedure, which delineates non-negative, fully invested and feasibility-compliant allocation vectors by minimizing the squared Euclidean distance between portfolio allocations and the preference-based utility structure of each investor segment. Capital allocation outcomes are accordingly derived endogenously from the interaction between normalized asset performance, explicit preference weights and optimization constraints, without imposing predetermined investment orientations or market assumptions.

The empirical implementation demonstrates that, under balanced preference configurations, the framework generates diversified portfolio compositions sustained by assets exhibiting robust cross-dimensional performance. Conversely, under polarized preference environments explored through sensitivity analysis, the optimization process yields more specialized portfolios that maintain internal consistency and constraint compliance, confirming the robustness and interpretability of the framework across divergent preference structures.

In synthesis, the analysis establishes a replicable, scalable and inherently transparent decision-support model for preference-driven portfolio optimization. Through the formalization of preference heterogeneity as the core structuring mechanism of capital allocation, the framework offers a methodologically grounded model with clear operational relevance for portfolio construction within multi-dimensional, sustainability-oriented and structurally evolving investment landscapes.

Italian Version

La ricerca delinea un modello di ottimizzazione di portafoglio fondato su un'architettura Multi-Criteria Decision Analysis (MCDA), la quale integra il rendimento finanziario, il rischio di volatilità e le dimensioni Environmental, Social and Governance (ESG) quali criteri valutativi espliciti e strutturalmente equivalenti. Il modello si pone come risposta all'inadeguatezza strutturale degli approcci dipendenti dalle previsioni, configurando la distribuzione del capitale come un prodotto endogeno del sistema di preferenze.

L'eterogeneità degli investitori viene formalizzata in vettori di preferenza composizionali, definiti sul simplesso Rendimento–Rischio–ESG e generati mediante distribuzioni Dirichlet. Questa imposizione assicura coerenza interna, positività e il pieno rispetto dei vincoli di somma unitaria in un ecosistema di modellazione interamente basato sull'evidenza dei dati. Tali strutture di preferenza sono sottoposte ad una segmentazione tramite clustering non supervisionato K-Means, da cui derivano tre gruppi stabili di investitori i cui centroidi forniscono i pesi prioritari specifici per ciascun criterio. I pesi ottenuti sono sintetizzati in Aggregated Utility Scores (AUS), i quali costituiscono una metrica trasparente e tracciabile dell'attrattività di ogni Exchange-Traded Fund (ETF), che compongono l'universo d'investimento, in base alle preferenze degli investitori. Il processo di costruzione dei portafogli è formalizzato tramite una procedura di ottimizzazione di Programmazione Quadratica (QP) vincolata, la quale identifica vettori di allocazione non negativi, interamente investiti e conformi ai requisiti di ammissibilità, minimizzando la distanza Euclidea quadratica tra le allocazioni di portafoglio e la struttura di utilità basata sulle preferenze di ciascun segmento. Gli esiti della distribuzione del capitale derivano pertanto intrinsecamente dall'interazione tra le performance normalizzate degli asset, i pesi espliciti delle preferenze e i vincoli di ottimizzazione, senza l'imposizione di predefiniti orientamenti d'investimento o assunzioni di mercato arbitrarie.

L'implementazione empirica evidenzia come, in presenza di preferenze bilanciate, il modello generi composizioni di portafoglio diversificate, basate su asset che esibiscono performance robuste in tutte le dimensioni analizzate. Al contrario, in scenari polarizzati, l'ottimizzazione produce portafogli più concentrati che mantengono comunque la coerenza strutturale e il rispetto dei vincoli, confermando la robustezza del framework anche con preferenze divergenti e concentrate.

In sintesi, questa analisi definisce un modello di supporto alle decisioni replicabile, scalabile ed intrinsecamente trasparente per l'ottimizzazione di portafogli in funzione dell'eterogeneità delle preferenze. Attraverso la formalizzazione dell'eterogeneità come perno strutturale della distribuzione del capitale, la ricerca propone un approccio rigoroso e operativamente rilevante per la gestione dei processi decisionali multicriterio in contesti d'investimento multidimensionali, orientati alla sostenibilità ed in costante evoluzione.

1. Introduction

The acceleration of the global sustainable development trajectory and the compounding effects of climate transition dynamics have decisively reconfigured the structural drivers and the systemic mechanisms of capital markets. Traditionally anchored in the quantitative trade-off between expected return and financial risk, portfolio optimization frameworks conventionally rely on analytical structures designed to model investor preferences within bi-dimensional utility representations. Within this evolving landscape, the Environmental, Social and Governance (ESG) dimensions emerge as distinct evaluative criteria that redefine the prioritization of investment alternatives, exposing the intrinsic limitations of traditional paradigms in the formalization of such complex preference structures. To address these requirements, the Multiple-Criteria Decision Analysis (MCDA) supports the synthesis of qualitative and quantitative preferences into specific allocation weights through the integration of multifaceted dimensions and the explicit modeling of the priority assigned to each criterion.

Building upon these theoretical foundations, the present research develops a structured and data-driven decision framework to examine the relationship between heterogeneous investor preferences and portfolio optimization. Specifically, the architecture integrates a set of interconnected analytical stages to define multidimensional utility structures across return, risk and ESG criteria, and evaluate the impact of these preference configurations on the selection of the investment alternatives within the optimized portfolio compositions.

In this research, the literature review establishes the theoretical foundation of the research through a comprehensive examination of the evolution of portfolio optimization theory, mapping the developments from early principles to modern frameworks addressing the limitations of traditional risk–return models. Subsequently, the research introduces the proposed methodological architecture, formalizing the components of the model and establishing the basis for the empirical implementation that operationalizes the framework within a transparent computational environment.

Ultimately, the concluding section presents the results of the optimization process and examines the structural properties of the resulting portfolios. Beyond the empirical evaluation, the research provides an interpretative perspective that relates the generated allocation patterns with established investment strategies and, successively, verifies the robustness of the model through a sensitivity analysis conducted under alternative preference specifications.

In summary, this research aims to contribute to the literature on portfolio management through a structured framework that integrates heterogeneous preferences within a multidimensional decision environment to ensure the transparent optimization of targeted portfolio allocations.

2. Literature Review

2.1 The Early Foundations of Portfolio Optimization

The discipline of portfolio optimization finds its foundations in early attempts to describe market behavior through intuitive reasoning and empirical observation, preceding the institutionalization of systematic formal methodologies. At this initial stage, the study of financial markets remained largely descriptive, relying on qualitative assessments of the dynamics of security prices, observable market patterns and prevailing investor expectations.

These approaches were notably constrained by the absence of formal mathematical structures capable of representing uncertainty and, more broadly, of formalizing the relationship between risk and return. The lack of a unified analytical framework, coupled with a conceptual treatment of financial outcomes as static events, precluding their status as stochastic processes governed by measurable probability distributions, restricted investment decisions to the domain of experience and subjective judgment. The transition toward a scientific discipline necessitated the gradual integration of mathematical reasoning and statistical inference, whereby market behavior articulation through quantifiable mutual influence established the necessary preconditions for decision optimization under uncertainty.

The work of Bernoulli (1738) in the Theory of Utility introduced the concept of risk preferences, proposing that investors seek to maximize their expected utility rather than simply the expected wealth. It was demonstrated that the decreasing marginal utility of wealth, a core assumption of risk aversion, could be explained by concave utility functions, such as the logarithmic one:

$$U(W) = ln(W)$$

In this formulation $U(W)$ refers to the level of utility or subjective satisfaction derived from the wealth (W). The concave function inherent to the natural logarithmic form implies diminishing marginal utility, whereby each additional unit of wealth provides a proportionally smaller increment of benefit as total wealth increases. While the

intuitions about expected utility maximization and risk aversion provided the theoretical basis for decision-making in stochastic environments, their application to the valuation of financial assets necessitated a rigorous definition of the objective sources of risk.

One of the first applications of statistical reasoning to financial markets appears in the work of Bachelier (1900) in the *Theory of Speculation*, which modeled price changes as independent and normally distributed, laying the groundwork for what would later be formalized as the random walk of asset prices. This assumption, which treats prices as essentially unpredictable, provided the basis for the development of stochastic models and pricing of derivatives. Lacking the ability to explicitly balance risk and return, the early financial models failed to provide a mechanism for a systematic approach to portfolio selection.

The integration of market dynamics with economic principles of capital allocation was facilitated by key macroeconomic and microeconomic contributions. Specifically, the intellectual trajectory of portfolio theory was significantly shaped by the contributions of Irving Fisher (1930) who formalized the time value of money through discounting, and by John Maynard Keynes (1936) who emphasized important concepts about uncertainty and introduced the theory of liquidity preference, framing interest rates as the reward for parting with liquidity.

Collectively, their contributions shaped later developments in financial economics and influenced subsequent approaches to asset valuation. Building on the principles of valuation formalized by Fisher, the influential Dividend Discount Model (DDM) proposed by John Burr Williams (1938) described how the value of a stock is primarily determined by the present value of its future dividends:

$$P_0 = \sum_{t=1}^{\infty} \frac{D_t}{(1+r)^t}$$

In this formulation, P_0 is the intrinsic value of the stock in the current period, derived from its anticipated future dividends, in which D_t represents the expected dividend per share in each subsequent period t. All dividends are discounted back to the present using r, which denotes the minimum expected return required for the risk assumed.

These models offered methods for asset valuation, but they failed to define the axiomatic rules governing how rational individuals should choose among multiple risky assets or uncertain prospects, highlighting the necessity of establishing a rigorous decision criterion for agents operating in non-deterministic environments (Samuelson, 1937).

Within this context, the development of axiomatic decision theory by Von Neumann and Morgenstern (1944) formally addressed this gap, establishing that rational investors maximize their expected utility, formalized as:

$$E[u(x)] = \sum_i p_i u(x_i)$$

Under this specific model, the expected utility $E[u(x)]$ is calculated as the sum of products obtained by multiplying the utility $u(x_i)$ of the individual outcome by its respective probability of occurrence (p_i). This structure demonstrates the fundamental principle by which the rational agent evaluates a risky prospect based on the subjective value or utility derived from each potential outcome. These theories provided the conceptual basis for rational decision-making, acting as the foundations of risk preferences and expected utility, thus constituting the essential explanation of individual asset behavior. However, these models did not extend to practical portfolio selection, and their lack of structured management approaches left the field largely intuitive and subjective, without a quantitative basis for risk-return optimization.

Harry Markowitz's (1952) Modern Portfolio Theory (MPT) transformed the field of investment management by providing the intellectual basis for the discipline and by directly addressing the shortcomings of prior models, specifically regarding the absence of rigorous mathematical frameworks for multi-asset portfolio construction.

In a clear departure from earlier approaches relying on general utility functions, this model precisely quantified portfolio risk as the variance (σ^2) of portfolio returns. This advancement redefined portfolio management by demonstrating that investment decisions must satisfy the dual objective of maximizing returns and minimizing risk.

This analysis established the concept of the efficient frontier, a graphical representation of the set of optimal portfolios offering the highest return for a given level of risk, or conversely, the lowest risk for a predetermined level of expected return. The selection of a specific portfolio along this frontier is determined by the individual utility function, often approximated by a quadratic utility function reflecting risk aversion:

$$U = E(R) - \frac{1}{2}\lambda\sigma^2$$

This expression formally models the trade-off between the expected return $E(R)$ and the risk penalty, $-\frac{1}{2}\lambda\sigma^2$, where the utility ($U$) is maximized when the rational investor balances the anticipated gain against the variance σ^2, scaled by the risk aversion coefficient (λ).

By formalizing the Mean-Variance trade-off and offering the prescriptive method for portfolio selection, this approach redefined the basis for portfolio management. Furthermore, the lack of a rigorous quantitative basis for the intuitively understood benefits of diversification was addressed by establishing the covariance between asset returns as the primary determinant of portfolio risk. This theoretical insight proved that combining assets with low or negative correlation could reduce the overall portfolio risk more effectively than simply selecting individual low-risk holdings, providing the rationale for the identification of non-dominated portfolios.

Despite the revolutionary contributions to the discipline of efficient portfolio construction through the optimization of the risk-return trade-off, the MPT intrinsically exhibits significant limitations, particularly concerning the practical application and the high sensitivity of input parameters. The computational requirements of the theory present a considerable limitation related to the necessity

of the estimation of expected return and variance for each asset, as well as the covariance for every possible asset pair. For a portfolio of N assets, this requires the estimation of N expected returns, N variances and $\frac{N(N-1)}{2}$ covariances. This approach presented a significant computational challenge, especially in the 1950s and 1960s when computational resources were limited. As a result, the necessary calculations were computationally infeasible and increasingly difficult as the number of assets grew (Black & Litterman, 1992). This limitation was further complicated by the problem of estimation error. The model necessitates precise estimated values, typically derived from historical data, with the result that even minor inaccuracies or volatility can lead to a highly divergent portfolio composition (Michaud, 1989).

Another important limitation of the model was its theoretical framework for risk. By reducing all risk exclusively to the standard deviation (σ) of the portfolio returns, calculated as a combination of variances and covariances among the assets held, the theory fundamentally failed to distinguish between distinct typologies of risk. While the MPT successfully established the benefit of diversification for mitigating portfolio risk, it operated as a partial equilibrium model, failing to specify the requisite market equilibrium price or the appropriate compensation for assuming systematic risk.

Conclusively, the MPT laid the initial rigorous quantitative basis for portfolio construction processes by transitioning investment management into a discipline grounded in formal mathematical optimization. Despite these seminal contributions, its inherent structure as a partial equilibrium model left unaddressed the critical question of how markets determine equilibrium asset prices and the essential compensation for a given level of systematic risk. This specific limitation, though, did not undermine the influence of the theory, but rather it drove a new stream of research focused on extending the framework beyond individual portfolio selection toward a broader general equilibrium theory of capital market behavior.

2.2 The Evolution of Portfolio Optimization Theory

The intellectual development of financial theory, deriving from the inherent limitations of early portfolio optimization methodologies, fostered a crucial transition toward a comprehensive understanding of how equilibrium returns are determined in the aggregate market. This decisive shift transcended the initial focus on individual portfolio construction by addressing the critical challenge of effectively integrating the behavior of investors with the collective dynamics of capital markets, particularly regarding the mechanism through which aggregate risk exposure translates into systematic patterns of asset returns.

The introduction of the concept of a risk-free asset by James Tobin (1958) provided an important extension to preliminary portfolio optimization methodologies. A result later referred to as the Two-Fund Separation Theorem demonstrated that the investment decision can be divided into two phases, regardless of the risk tolerance of the investor. The first involves selecting a single optimal portfolio of risky assets, whereas the second involves the allocation of capital between this risky portfolio and a risk-free asset. This can be expressed as:

$$W = xP + (1 - x)F$$

In this formulation, W represents the complete portfolio, P the optimal risky portfolio and F the risk-free asset. The work of Tobin constituted a crucial bridge in the evolution of the MPT into the Capital Asset Pricing Model (CAPM) and the broader field of modern capital market theory. The principles of the Mean-Variance model (Markowitz, 1952) and the Two-Fund Separation Theorem (Tobin, 1958) emerged as the theoretical precursors to the development of the CAPM by Sharpe (1964), Lintner (1965) and Mossin (1966). The CAPM formalized a market equilibrium perspective, representing a significant epistemological shift through the disaggregation and quantification of risk. Specifically, the model defined unsystematic risk as the variability of returns stemming from idiosyncratic factors, asserting that this diversifiable risk could be eliminated and would therefore not be compensated by the market.

Conversely, systematic or market risk, the variability of returns from aggregated factors, was described as non-diversifiable and as the exclusive source of risk that requires a return premium. Additionally, the CAPM formalized the concept of beta (β) as the standard metric for systematic risk, quantifying the sensitivity of an asset to aggregate market fluctuations by expressing its covariance relative to the market portfolio. This allowed for the establishment of an equilibrium relationship between the expected return and the exposure to market risk of a given asset.

This approach to asset valuation and equity cost estimation simultaneously provides the conceptual justification for the market risk premium ($E(Rm) - Rf$), which is formally defined as the market portfolio's expected excess return above the risk-free rate. The Capital Market Line (CML), a fundamental element of the CAPM framework, graphically depicts the relationship between expected return and total risk (σ) of efficient portfolios. With its y-intercept fixed at the risk-free rate (Rf), the CML defines the optimal risk-return trade-off and the set of all efficient portfolios constructed via combinations of the risk-free asset and the market portfolio. In contrast, the Security Market Line (SML) represents the equilibrium relationship between an asset's expected return and its systematic risk (β). Unlike the CML, the SML applies to all portfolios, irrespective of their diversification level, and works as the primary tool for asset valuation. Assets plotting above the SML are considered undervalued, while those falling below it are considered overvalued.

The formula that establishes the relationship between the expected return and systematic risk of an asset, according to the CAPM, is:

$$E(R_i) = R_f + \beta_i\big(E(R_m) - R_f\big)$$

In this equation, $E(R_i)$ is the expected return, R_f the risk-free rate and $\beta_i\ \big(E(R_m) - R_f\big)$ the risk premium for assuming systematic risk. The theoretical implications of the CAPM extend to several practical applications in corporate finance and investment management, primarily as a mechanism for estimating the cost of equity capital K_e.

The equation for estimating the cost of equity capital specifies the minimum return demanded by shareholders for an organization, constituting an essential element in the calculation of the Weighted Average Cost of Capital (WACC) for corporate valuation and assessment of the viability of investment projects.

The simplifying assumptions underlying the CAPM, significantly divergent from empirical market realities, leave the model vulnerable to several notable limitations despite the theoretical elegance and widespread application. The model is based on a set of idealized assumptions, including perfect rationality and risk aversion, frictionless markets free from transaction costs or taxes, the capacity to lend and borrow capital at the risk-free rate and homogeneity of expectations, according to which all investors are assumed to share the same beliefs about expected returns, variances and covariances of assets.

In addition to these idealistic premises, a significant practical challenge of the model arises from the measurement of its core metric (β), as values derived from past data are not consistently reliable indicators of future risk. The beta coefficient of an asset exhibits temporal instability by evolving in response to shifts in business and market conditions, decreasing the reliability of the model for long-term forecasting.

Several academic analyses have demonstrated significant discrepancies between the theoretical formulations of the CAPM and its empirical performance, revealing an attenuated relationship between beta and expected returns. Specifically, the size effect, first documented by Rolf Banz (1981), showed that in US data from 1936 to 1975, small–cap stocks tended to generate higher returns than large–cap stocks, even after accounting for systematic risk. Similarly, empirical studies showed that, in specific markets and periods, value stocks with a low price-to-book ratio earned superior average returns compared to growth stocks with a high price-to-book ratio. The presence of these effects indicates that the market compensates for additional risk factors associated with firm size and companies' valuation, highlighting the inadequacy of CAPM in explaining the superior performance using beta as the exclusive explanatory variable.

The lasting contribution of CAPM to financial theory is its seminal role as the first model to systematically quantify the relationship between risk and return, a foundational paradigm that, despite the empirical limitations, fostered the development of more robust asset pricing models.

Building upon the empirical limitations of the CAPM, the research of Fama and French (1992) marked a significant evolution in asset pricing theory. Their work transformed the size and book-to-market effects into measurable risk factors through the construction of mimicking portfolios, facilitating the introduction of the more sophisticated Three-Factor model for capturing the risk dimensions consistently ignored by preceding frameworks and incorporating the original market factor alongside the two factors of size (SMB) and value (HML).

The Fama and French (1993) model became a foundational benchmark for empirical return analysis, especially in evaluating fund performance, testing market efficiency and assessing active investment management. This research provided the basis for an extensive body of work aimed at identifying additional factors, such as momentum, profitability and asset investment over the years. In financial practices, the model has provided investors and portfolio managers with more sophisticated instruments for diversification and factor-based portfolio construction, moving beyond the limitations of the traditional beta-only framework. The 'Common Risk Factors in the Returns on Stocks and Bonds' (Fama & French, 1993) reoriented the conceptual emphasis from single-factor paradigms to multifactor models through market anomaly reclassification as systematic risk manifestations.

These contributions integrated previously unexplained return patterns into a risk-based framework, establishing the conceptual basis for an extensive stream of research and driving financial theory toward non-economic variables inclusion, such as Environmental, Social and Governance (ESG) factors, in asset pricing methodologies.

2.3 The Rise of Sustainability in Portfolio Optimization

The intensifying focus on sustainability and climate change as determinants of economic and financial decision-making has driven a significant evolution within the discipline of portfolio optimization. The growing recognition of climate change as a systemic risk to global economic stability required a broader approach to investment strategies, leading to increased attention toward ESG principles.

The Stern Review (2006) offered a detailed macroeconomic evaluation that significantly reshaped both public and academic perceptions of climate change, with the central and most influential contribution characterizing climate change as the greatest market failure in history, resulting from the systematic failure to internalize the negative externalities associated with greenhouse gas emissions. This analysis provided a solid economic justification for large-scale governmental intervention and repositioned the subject from a purely environmental concern to an economic and financial risk.

The ethical divergence from conventional economic practices constituted the most radical feature of the research, stemming primarily from the adoption of a low discount rate (1.4% per year) to estimate the costs and benefits of climate change over long horizons. The adoption of this methodological approach generated unprecedented estimates of economic damage, ranging from 5% to 20% of GDP, and effectively translated future costs into present economic terms to establish climate change as a tangible threat to global prosperity.

Beyond quantifying the costs of inaction, the report highlighted the financial risk of stranded assets, explaining that fossil fuel reserves could lose value due to regulatory and market shifts. These insights, directly linked to the principles of risk and return management, provided both theoretical and practical foundations for subsequent studies on ESG integration and fostered the development of new financial instruments such as green bonds and carbon markets. The reclassification of climate risk as an economic concern raised a fundamental question for the financial landscape about whether integrating ESG factors into investment decisions would incur a financial cost.

With the debate shifting from theoretical justification to practical implementation, a new research agenda emerged to address the central question for investors concerning the extent to which ethical investing required a trade-off in returns.

One of the cautious early positions on the link between ESG criteria and corporate financial performance was established by Renneboog et al. (2008). Their study, through the analysis of the performance of Socially Responsible Investment (SRI) funds, examined the possibility that ethical restrictions might create a negative trade-off. This assertion stems from the economic principles explaining that the diversification of SRI funds is reduced when their investment universe is deliberately restricted, such as by excluding sectors like tobacco, alcohol and weapons. From a theoretical perspective, this reduced diversification could expose the portfolio to greater specific risk, potentially leading to underperformance compared to unconstrained portfolios. Their empirical evidence demonstrated that while SRI funds perform similarly to conventional funds, they often slightly underperform after fees.

Friede et al. (2015) contributed significantly to the debate, directly addressing the concerns raised by Renneboog et al. (2008). Their extensive meta-analysis, encompassing nearly 2,200 academic studies and reports published from 1970 to 2014, classified each contribution according to: the type of relationship (positive, negative, mixed, or non-significant), the specific ESG component, the financial performance measure (accounting-based, market-based or risk-based) and the geographic or sector context. After collecting the results, the researchers employed statistical techniques to compute an aggregate mean effect and conducted subgroup analyses to determine how the relationship varied across specific factors, such as industry sector or financial metrics.

The central finding of the meta-analysis was that ESG integration is not generally linked to lower returns as most of the studies (almost 90%) found a neutral or positive relationship with corporate performance, with most of them falling into the positive category.

The study concurrently uncovered a differentiated relationship between the ESG components showing that their impact varied. Specifically, it was demonstrated that evidence from Governance (G) and Environmental (E) factors tends to be more consistent, while results for the Social (S) dimension remain less conclusive, leading investors to allocate particular attention to how companies manage risks and opportunities associated with these aspects of business.

By providing robust and quantitative evidence drawn from decades of research, this analysis presented a detailed case for sustainable investing. The study elucidated that ESG represents a rational and potentially profitable component of financial analysis. Therefore, the debate shifted from whether sustainability and returns might coexist to the mechanisms of their synergy, evidencing that these factors function as genuine value drivers capable of improving risk management, operational efficiency and innovation over the long term. The findings provided a scientific basis for the explosive growth of the sustainable investing movement. By legitimizing ESG and demonstrating that aligning investment portfolios with sustainability objectives does not require a sacrifice of financial performance, this work created a critical benchmark for studies on the relationship between ESG and corporate financial performance.

The validation of the interplay between ESG performance and positive financial returns created a new area of research centered on capturing why and how sustainability integration generates value. Khan et al. (2016) provided a clear explanation by introducing and empirically confirming the concept of ESG materiality.

The study of Khan et al. (2016) extended beyond the undifferentiated approach of previous research, which assumed the universal importance of all environmental, social and governance factors for every company, independently from the industry, and effectively treating ESG criteria as a universal checklist. As a result, these earlier models created ESG ratings that did not accurately reflect the specific financial risks of the companies, leading to clear disconnections between the ratings and actual performance.

These analyses aimed to find a correlation between returns and a heterogeneous range of factors, overlooking the critical insight that only a portion of these factors were truly relevant to value creation.

To overcome this, Khan et al. (2016) found that the importance of the financial impact of ESG factors varies significantly across the companies. This insight arose from the empirical validation of ESG materiality, which demonstrated that firms excelling in material ESG issues significantly outperformed peers with weak performance, while no such effect was found for immaterial ESG issues. This research transformed the perception of ESG, shifting the focus from ethical obligations or costs to a strategic tool for value creation.

Their study challenged the undifferentiated treatment approach and demonstrated that monetary and financial benefits are the results of a targeted and selective allocation of resources towards elements that are significant to a particular industry or company. This concept provided a clear explanation for the origin of excess returns by showing that companies that actively target these material aspects tend to outperform competitors and reframing ESG as a strategic lever for enhancing competitiveness and long-term performance.

These contributions offered practical guidance to investors and companies for identifying ESG factors that increase financial performance within each sector. This work also inspired a new generation of ESG rating tools and methodologies built on the concept of materiality, such as Sustainability Accounting Standards Board (SASB). This evidence demonstrated that companies acting responsibly on material ESG factors are not merely making choices, but strategic economic decisions for durable value creation.

To incorporate ESG insights and implement sustainability principles, MCDA models offer structured and sophisticated approaches to balance financial performance with various criteria in sustainable portfolio optimization processes.

2.4 MCDA Application in Sustainable Portfolio Optimization

The advancements in financial research have established that the integration of ESG factors into investment decision-making processes meets ethical considerations, fosters value and improve long-term risk management. This evidence redefines the nature of climate risk, transforming it from a purely environmental challenge to a systematic economic threat.

Despite the solid conceptual foundations, the practical application of these principles into portfolio construction processes remains a significant challenge, as the simple adoption of these criteria is insufficient to ensure consistency, transparency and robustness. The effective conversion of sustainable finance principles into operational models requires structured methodologies and quantitative approaches that balance multiple financial and non-financial variables. Within this context, MCDA models offer robust mechanisms for the explicit incorporation of ESG factors into portfolio management.

Traditional optimization approaches, exclusively focused on the risk-return trade-off, preclude the integration of ethical considerations or non-financial objectives required by investor profiles seeking to combine financial performance with sustainability goals. To overcome this inadequacy, Ballestero et al. (2012) implemented a multi-objective framework that elevates ethical objectives from simple constraints to optimization criteria, placing them on the same level as risk and return.

Their central contribution lies in the application of MCDA to create a bi-criteria model for portfolio selection, which extends the classical expected utility by incorporating ethical objectives as quantifiable goals.

The bi-criteria utility function proposed aims to minimize the composite index of variability of two distinct utility objectives:

· The financial utility, which refers to the traditional goal of maximizing portfolio returns without considering ethical factors, is expressed as:

$$Eu_1(\widehat{R_1}) \rightarrow u_1(\overline{R}_1)$$

This equation describes how the expected utility Eu_1 of the entire portfolio return $(\widehat{R_1})$, computed as the sum of the return of all assets multiplied by their respective weights, is equivalent to the utility of the expected return $u_1(\overline{R}_1)$, consistent with utility theory.

· The ethical utility represents the objective of an investor who values SRI and tends toward the utility of the expected return of a subset of ethical assets, expressed as:

$$Eu_2(\widehat{R_2}) \rightarrow u_2(\overline{R}_2)$$

In this formulation, analogous to that of financial utility, the primary objective is to maximize the expected return $(\widehat{R_2})$ of the ethical portfolio, which includes only ethical assets, whereas the term $(\overline{R}_2)$ denotes the return of a subset of ethical assets that contribute to this objective.

· The objective function aims to find the combination of assets that minimizes the overall risk of the portfolio, formulated as:

$$v = \min X\, V X^T$$

This equation includes the vector X, representing the portfolio weights, the composite variable matrix V, defined as a linear combination of the return covariance matrices for the financial and ethical objectives weighted by their respective Absolute Risk Aversion (ARA) coefficients, and the transpose of the weight vector X^T for computing the total portfolio variability. The model minimizes the variability index to find a solution that incorporates the investor's risk aversion to both financial and ethical components, making it suitable for a heterogeneous range of profiles.

The key conceptual point is the use of ARA coefficients for quantifying the investor's willingness to sacrifice financial returns for ethical purposes and measure the actual risk tolerance, moving beyond treating sustainability as a binary outcome determined by positive or negative screening. Their application permits the identification of distinct investor profiles, demonstrating that an investor with high ethical risk aversion (r_2) will be inclined to sacrifice a portion of financial return to achieve ethical goals, whereas one with low ethical risk aversion will align more with the traditional profile. The model defines three distinct investor profiles through the selection of coefficients (r_1) and (r_2):

· Traditional investors are mathematically characterized by setting $r_2 = 0$, which reduces the problem to maximize financial utility and aligns with the classical optimization model

· Strong green investors assign a dominant weight to r_2, focusing exclusively on achieving ethical objectives without considering financial returns

· Weak green investors, who represent the profile the model is designed to assist, are risk-averse toward both financial and ethical goals and seek an optimal compromise between the two, whereby both coefficients assume positive values

The main contribution of Ballestero et al. (2012) is the exploitation of ARA coefficients to model financial risk while translating the subjective preferences of investors into quantifiable variables. This approach reframed the focus from a simple ethical or unethical binary choice to optimally ethical investments, transforming the MCDA into a powerful tool for capturing investor preferences, selecting sustainable portfolios and developing more personalized investment strategies.

Drawing from the solid and innovative foundations established by Ballestero et al. (2012), subsequent research evolved the adoption of multi-criteria models to address new market dynamics and to fully capture the multidimensional nature of sustainability. While primarily focused on complex business management and project selection problems, the paper by Tinoco et al. (2018) represents a key point for sustainable optimization processes.

The Integrated Model for Probabilistic Optimization of Project Portfolios represents a holistic framework for addressing the complexity and the uncertainty of corporate projects. Its mathematical core lies in the unification of three dimensions into a single probabilistic measure, known as the Consolidated Return (CR), and in the use of Monte Carlo simulation to effectively manage the variability of measures. The Consolidated Return (CR_j) for a single project is computed as the sum of economic return (R_j), environmental impact (E_j) and social impact (S_j) of the project j, modeled as random variables and expressed in monetary terms. The model assumes these three variables are described by probability distributions, such as normal or triangular, instead of fixed values.

The non-linear optimization problem for project portfolio selection, based on implementing criteria that cover ESG dimensions, is formulated as:

$$\max(CR) = \sum_{i=1}^{n} P_i(CR_i)$$

Here, CR_i represents the benefit or success rate for the project i and P_i represents a binary variable assuming value one if the project is selected, or zero otherwise, with the entire formulation designed to maximize the cumulative success rate of selected projects. This maximization objective is subject to a mutual exclusion constraint to mitigate potential conflicts, a budget constraint to define the maximum economic threshold for financial viability and a man-hours constraint where the aggregate labor requirement for the project must not exceed the total capacity of available human resources.

By utilizing the Monte Carlo simulation, the model is repeatedly run on many samples, with the values for the impacts extracted from their respective probability distributions. The process yields an entire distribution of possible CR values, resulting in a final histogram of optimal solutions that illustrates the probability of projects being included in an optimal portfolio across a wide range of future scenarios.

The work of Tinoco et al. (2018) introduced a robust model for the management of uncertainty, but the evolution of research required more sophisticated approaches to overcome the limitations posed by linear assessments. Within this context, Chen et al. (2021) provided a more realistic model, advancing the practical validity of MCDA models.

The initial contribution of this framework criticizes existing rating models such as those leveraging Thomson Reuters scores, positing that linear weighted average approaches are incapable of accurately reflecting reality due to the assumption of mutual independence among ESG factors. For example, effective corporate governance is often a prerequisite for implementing efficient environmental and social policies, but linear models fail to account for these complex synergies and interactions. Furthermore, a linear model allows an exceptionally high score in one dimension, such as social, to offset a critically weak score in another, like environmental impact. This enables a company with a poor performance in one criterion to still receive a positive overall score due to a high score in another, contradicting the principle of sustainability which requires solid performance across all dimensions.

To mitigate the drawbacks of linear models, Chen et al. (2021) proposed a non-linear Data Envelopment Analysis (DEA) model for assessing the efficiency of each company (Decision-Making unit) in transforming its ESG attributes into a single performance measure. Unlike linear models that rely on weighted sums, this model incorporates quadratic and cubic terms into the performance measure to capture the interaction effects between ESG dimensions and to reflect the concept of evidence enhancement. This approach produces a non-compensatory and more realistic efficiency score, preventing strong performance in one dimension from compensating for weaknesses in another one.

The performance measure provided by the model is expressed as:

$$\max S_0 = w_{E0}E_0 + w_{S0}S_0 + w_{G0}G_0 + \sum_{i,j=E,S,G} w_{ij}x_{i0}x_{j0} + \sum_{i,j,k=E,S,G} w_{ijk}x_{i0}x_{j0}x_{k0}$$

The primary goal of the non-linear DEA model is to maximize the Relative Efficiency Score of the firm under examination (DMU_0). The Relative Efficiency Score formulation is a composite sum structured around a standard linear base, which includes the raw ESG scores (E_0, S_0, and G_0) weighted by the optimal coefficients (w). This linear foundation is complemented by non-linear terms designed to capture complex ESG interactions, including the quadratic terms ($w_{ij}x_{i0}x_{j0}$) for the pairwise interaction between two distinct dimensions and the cubic term ($w_{ijk}x_{i0}x_{j0}x_{k0}$) for capturing the higher-order synergistic effect among all three. The objective function is solved individually for each firm, seeking to make that specific firm maximally efficient relative to all the others. This non-linear maximization generates a score ranging from 0 to 1, providing a robust and realistic efficiency assessment that prevents exceptional performance in one area from obscuring a critical deficiency in another.

The research presents a three-phase framework for the practical selection of sustainable portfolios, structured as follows:

· ESG evaluation phase where the non-linear DEA model is applied to compute the ESG score for each asset, representing an efficiency index that incorporates non-linear interactions

· Asset screening phase in which the DEA cross-efficiency technique benchmarks each asset against its competitors' strategies, creating a more robust asset pool by eliminating units that are seemingly strong but inefficient when compared to rivals

· Portfolio optimization phase that leverages a modified Mean-Variance model, which extends to include the ESG score for enabling the construction of portfolios that satisfy financial constraints while achieving high levels of sustainability performance

Despite the significant contributions of Chen et al. (2021), the methodological architecture developed herein constitutes a structural departure driven primarily by the specific requirements for multi-criteria capital allocation. The core objective of Chen et al. (2021) is the estimation of relative ESG efficiency scores through a non-compensatory DEA formulation, whereas the current model is explicitly designed for portfolio structuring, positioning sustainability as an intrinsic decision criterion rather than as an intermediate output.

Moreover, the two methodologies diverge significantly in their treatment of investor rationale, with Chen et al. (2021) assuming homogeneous preferences implicitly embedded within the efficient frontier, while the architecture proposed advances the approach by explicitly modeling investor heterogeneity via robust behavioral segmentation and centroid-based aggregation.

This difference is concurrently reflected in the graphical representation. Consistent with their objective of relative performance measurement, the efficiency-based visualizations employed in DEA studies omit the incorporation of explicit preference integration. In contrast, the three-dimensional utility space constitutes an explicit decision arena for the direct mapping of internal coherence between asset characteristics, preference-weighted utilities and the optimal asset deployment.

The studies by Tinoco et al. (2018) and Chen et al. (2021) offered two significant contributions to the evolution of multi-criteria sustainability valuation models. While the first one introduced a comprehensive framework that extended corporate risk assessment to include ESG factors, the second one developed a non-linear mathematical approach for capturing their intricate relationships.

Although these works refined models for the construction of sustainable portfolios, the next step has involved addressing the challenges of converting theoretical frameworks into practical applications, particularly concerning the phenomenon of greenwashing, a misleading practice that poses significant threats to market credibility given the growing demand for ethical investments.

Greenwashing refers to the deceptive communication by a company of its ethical or sustainable commitments, contrary to its actual actions or investments. In a financial context, this translates into the creation of portfolios promoted for their ESG attributes, despite containing investments in companies with demonstrable negative impacts. This practice has shifted the scientific focus from the means of optimization to the validation and measurement of actual impact. Therefore, rather than exclusively providing more sophisticated updates of previous models, Steuer and Utz (2023) offered a direct response to this central market challenge through a reorientation from theoretical optimizations to the quantitative verification and transparency of results.

The research developed the three-criteria efficient surface, defined within the three-dimensional space of risk, return and ESG, composed of $\max_x \mu_p = \mu^T x$ for the return maximization, $\min_x \sigma_p^2 = x^T \Sigma x$ for risk minimization, and $\max_x \gamma_p = \gamma^T x$ for the ESG component maximization. This formulation is subject to a budget constraint for ensuring a fully invested portfolio and to a non-negativity constraint for preventing short selling by requiring all weights to be non-negative.

Each point on this surface represents a Pareto-efficient portfolio that is not dominated by any other that might offer higher returns, lower risks or superior ESG scores. The primary innovation of Steuer and Utz (2023) lies in making the complex three-dimensional efficient surface visually interpretable by developing the Non-Contour Efficient Fronts (NCEFs) concept. Instead of visualizing the entire surface, sectioning it at specific ESG score levels enables the model to compute and plot the corresponding two-dimensional efficient frontier.

From a mathematical perspective, this requires solving a two-dimensional optimization problem and adding an equality constraint for the ESG objective. Therefore, in addition to budget and non-negativity constraints, the optimization for return maximization and risk minimization includes an ESG scoring constraint. This ethical filter requires the portfolio to meet or exceed a minimum sustainability level determined by the investor, consequently excluding all asset combinations that fall below this threshold, regardless of their financial performance.

Concurrently, this approach offers a tool for detecting greenwashing, as any portfolio situated within the NCEF shows inconsistencies between its declared ESG commitment and actual performance, failing to achieve the best possible risk-return trade-off for its sustainability level. As a result, if a fund promotes itself as sustainable but its portfolio is not situated on the efficient frontier corresponding to its declared ESG level, it provides investors with an objective indication that the fund is not offering the best possible risk-return trade-off.

The methodological evolution of sustainable portfolio construction has advanced from traditional optimization to comprehensive multi-criteria frameworks, leading to robust tools that address market complexities and the necessity for transparency in the face of risks like greenwashing. However, with the theoretical foundation for integrating non-financial goals largely established, contemporary research has necessarily moved the focus from the methodological challenge of how to achieve portfolio optimality to the empirical consequences of this integration, quantifying the direct financial implications and performance trade-offs of sustainable investing.

2.5 Balancing ESG Objectives and Financial Performance

The necessity of incorporating sustainability dimensions into financial decision-making processes has fundamentally reshaped modern portfolio theory. While prior analyses successfully established the theoretical foundations of mathematical modeling for portfolio optimization and analyzed the statistical principles managing market equilibrium theory, the contemporary research agenda has decisively converged upon the empirical validation of the ESG-financial performance link. This stream of research is primarily focused on the determination of whether the incorporation of sustainability criteria necessitates an ethical sacrifice or, conversely, whether such inclusion fosters enhanced risk-adjusted efficiency within investment portfolios.

In line with this focus, the research by Gasser et al. (2017) introduces a three-criteria optimization architecture that integrates a social responsibility dimension (θ) combined with expected return (μ) and risk (σ) of an asset. This approach extends beyond the traditional application of binary ESG constraints to allow the explicit maximization of social responsibility, regulated by a preference parameter (γ), and transforms the problem into the determination of optimality within the three-dimensional space (μ, σ, θ).

To visualize and manage the complexity derived from the non-financial dimension, the model introduces the Capital Allocation Plane (CAP) to map the entire non-dominated set of optimal portfolios. This plane provides a comprehensive view of all possible trade-offs, enabling investors to select the allocation suited to their preferences, particularly between a μ/σ-optimal portfolio and a θ/σ-optimal portfolio. The empirical application of the extensive ASSET4 (Thomson Reuters) dataset reveals that an absolute maximization of ethical commitment (θ) necessitates a statistically significant contraction in both expected returns and the Sharpe ratio relative to the purely μ/σ-optimal frontier. Simultaneously, the research demonstrates that portfolios characterized by elevated, but not maximal, social responsibility score are achievable with a marginal performance deficit, pointing to a non-linear relationship between ethical utility and financial optimization.

Irrespective of the performance penalty, the inclusion of social responsibility scores radically alters the structure of the portfolios. The θ/σ-optimal portfolio produces a significantly higher sustainability score than the conventional counterpart, confirming that the model effectively achieves the social objective when maximized. This evidence establishes that social responsibility represents a cost variable in terms of financial performance and demonstrates that the exploitation of CAP is fundamental for identifying the allocation that optimally balances the three objectives under examination.

The empirical evaluation conducted by Cesarone et al. (2022) examines an extended Mean-Variance framework across different markets and temporal horizons, addressing the established premise that the incorporation of sustainability parameters transforms portfolio structure and implies a measurable trade-off between financial performance and portfolio composition. The research exploits a three-criteria (Return-Variance-ESG) model, formalizing the optimization objective through Convex Quadratic Programming (QP) to minimize portfolio variance, while maintaining fixed thresholds for expected returns and sustainability metrics, for identifying the Pareto-optimal surface within the three-dimensional space. The study employed an out-of-sample analysis across five real stock market datasets, utilizing two distinct time periods strategically to capture the potential normative variations stemming from global regulatory events.

The research indicates that the financial utility derived from ESG integration is strictly dependent on temporal and geographical factors, evidenced by the correlation between stringent ESG targets and superior portfolio performance in recent US market cycles, which contrasts with the lack of demonstrable improvement in profitability within the European context.

Extending the multi-objective approach focused on ESG integration with portfolio optimization models, Xidonas and Essner (2024) shift the focus toward the absolute maximization of sustainability performance by introducing an innovative approach for sustainable management centering on multi-objective optimization exclusively focused on ESG risk performance.

The methodological principles of this framework refer to the adoption of the min-max criterion to minimize the weighted maximum deviation of the portfolio from the achievable maximal target value across the three ESG pillars, specifically defined as Environmental Risk Performance (ERP), Social Risk Performance (SRP) and Governance Risk Performance (GRP).

From a managerial perspective, the key contribution lies in the flexibility afforded by customized weights (W_{ERP}, W_{SRP}, W_{GRP}), which enables the decision-maker to calibrate and reflect the specific priorities among the different ESG objectives. Moreover, the model incorporates the Controversy Performance (CP), imposing an explicit minimum threshold constraint (CP^{LB}) to limit the exposure of the portfolio to highly controversial companies. Empirical analyses conducted through backtesting confirmed the model's validity, demonstrating that the ESG-optimal portfolios generate systematically superior risk-adjusted returns (Sharpe ratio) compared to market benchmarks. This evidence supports the hypothesis of a positive relationship between effective ESG risk management and long-term financial performance.

The contribution by Abate et al. (2024) concentrates on a purely empirical analysis of European mutual funds, analyzing their actual performance relative to their ESG ratings. This research transitions from the theoretical definition of ESG efficiency to the concrete evaluation of its capacity to generate financial value, assessed using DEA efficiency measures and risk-adjusted performance metrics.

This analysis focuses on whether ESG integration leads to an ethical sacrifice or, conversely, to superior financial performance, measuring overall efficiency through the application of the DEA model, which assigns a relative efficiency score to each fund within the sample. The model assesses the net efficiency of the fund by measuring the net accumulation factor (U_j), accounting for the cumulative return net of ongoing charges and front-end loads, against the Standard Deviation, the Downside Deviation and the Kurtosis. The study is based on a sample of 634 European equity mutual funds which, to maximize the differentiation and limit potential bias, was deliberately

divided into two extreme subgroups consisting of high ESG funds (top 10% of the rating distribution) and low ESG funds (bottom 10%).

The investigative conclusions, supported by statistical tests such as the t-test and Kolmogorov-Smirnov, indicate the superiority of high ESG-rated funds, which achieve an average DEA efficiency score of 0.661, significantly higher than the 0.562 of low ESG-rated funds. Simultaneously, the analysis of the DEA variables identified that high ESG funds exhibit a superior average net accumulation factor (1.361 vs 1.268), significantly lower percentages of funds generating a negative cumulative return (3.85% vs 17.19%) and lower average operating costs (1.53% vs 1.92%), which contradicts the assumption of increased operational expenditures. Conclusively, this research rejects the ethical sacrifice hypothesis, demonstrating how portfolios with high ESG ratings offer significantly more efficient combinations of return, risk and cost.

Following the empirical demonstration that ESG criteria integration can yield a tangible competitive advantage in fund efficiency and performance, recent literature introduces a more critical and profound examination driven by concerns regarding the heterogeneity and the non-comparability of ratings across different agencies.

Cesarone et al. (2024) address these discrepancies by redefining the rating anomaly as an explicit and manageable model risk, subsequently developing the k-sum model grounded in QP for instability mitigation. The k-Worst Non-ESG score aggregates the k-worst (highest) scores through a worst-case analysis to guarantee resilience against the most negative ESG assessments and excessive rating variance. The results demonstrate the efficacy of the model, revealing that portfolios optimized using the k-sum, with k=2, achieve systematically superior values for the Sharpe ratio compared to those constructed on mean or single agency ratings. This observed robustness yields greater out-of-sample resilience during periods of high uncertainty and provides lower allocation volatility, resulting in reduced transaction costs.

In summary, the k-sum model proves to be the most robust methodology for isolating the most reliable sustainability signals, while simultaneously improving the portfolio risk-adjusted financial performance and operating stability.

The necessity to refine the technical structure of optimization for maximizing portfolio efficiency and stability has led asset allocation problems to become an effort to achieve coherence between economic performance and social responsibility, moving beyond the technical neutrality of preceding models. Within this context, the article by Varmaz et al. (2024) proposes a unified framework, composed of a threefold methodological innovation, to reconcile and empirically test the diverse theoretical positions concerning the connection among ESG, risk and return.

The first core contribution resolves the difficulty in specifying abstract investor preference parameters (α, λ, ε) for traditional utility maximization. They demonstrated the mathematical equivalence between the preference-based model and a reformulated optimization program that minimizes portfolio variance subject to explicit constraints on the desired levels of portfolio return $\left(\mu_p^*\right)$ and ESG rating $\left(\theta_p^*\right)$. This substitution of abstract coefficients with concrete and achievable targets facilitates the practical implementation of M-V-ESG optimization.

The second significant contribution refers to the development of a formal empirical test to resolve the theoretical ambiguity concerning the role of ESG in asset pricing. The field debates whether ESG acts as a characteristic providing an exclusive non-risk-related premium, known as ESG reward (Strand 1), or works as a systematic risk factor impacting both expected return and covariance, the ESG risk factor loading (Strand 2). The resulting test establishes the empirical foundation for allocation logic, helping investors determine the dominant explanatory impact of the ESG characteristic or the factor loading. Importantly, Strand 1 implies the preferential overweighting of high-ESG assets based purely on their return potential, without a corresponding increase in portfolio risk, whereas Strand 2 necessitates an explicit return-risk trade-off due to the associated increase in portfolio risk.

Finally, the methodology addresses the technical limitation of large-scale covariance matrix estimation, which is susceptible to singularity and unreliability in large asset universes. By adopting a Multi-Index Model (MIM), the optimization problem is notably simplified, avoiding the direct estimation of the full covariance matrix and allowing for large and unbalanced investment portfolio optimization.

Extending the analytical solutions from Varmaz et al. (2024), which offered a closed-form optimization with explicit ESG preferences, Bertelli and Torricelli (2024) developed a methodological benchmarking framework, based on the residual risk minimization model, for sustainable optimization that aims to compare three distinct approaches for constructing ESG-compliant and optimal portfolios.

The first employs a traditional Negative Screening Method, where the optimization is performed only on a pre-selected sample of assets, excluding low-ESG companies. The second Constraint Method is implemented across an investment universe subject to a minimum threshold on the average ESG score of the portfolio. The third approach, the Optimal Compromise, integrates the exclusion of companies with extremely low ESG scores with the imposition of a target constraint on the desired ESG value, combining the benefits of ethical screening and financial optimization. The study evaluates the robustness of these methodologies through key performance metrics, such as Sharpe ratio, Sortino ratio and Calmar ratio, each of which explicitly incorporates measures of downside risk.

The portfolios derived from the screening strategy failed to demonstrate significantly superior risk-adjusted efficiency compared to benchmarks, suggesting that simple upfront ethical filters are not sufficient to generate financial value and supporting the conclusion that ESG integration requires an active optimization process. By contrast, the incorporation of an explicit ESG constraint generates portfolios with superior risk-return performance, as evidenced by the consistently dominant efficacy of the third strategy across evaluation metrics, implying a non-linear association between average ESG scores and portfolio efficiency.

More precisely, superior Sharpe ratios tend to emerge at intermediate and high ESG levels, suggesting that efficiency is maximized when these criteria work as active constraints rather than passive exclusion filters. This finding is supported by the evidence that attempting to achieve the maximal ESG rating might, in fact, incur a return penalty.

The strategies employing MCDA for portfolio optimization problems have provided evidence that the full value of sustainability is realized when it works as an active constraint within a structured quantitative framework. Nevertheless, the utility of these models is inherently compromised by their dependency on the quality, granularity and timeliness of data. These structural limitations are further increased by the necessity of relying on the static and often inconsistent ratings provided by specialized agencies.

This structural gap between methodological sophistication and persistent data limitations constitutes the critical friction point that justifies and motivates the future potential of MCDA, contingent upon their capacity to interact constructively with Artificial Intelligence (AI). Through the exploitation of machine learning architectural properties, AI can enhance the analytical capabilities of MCDA by improving data collection and analysis, for instance, by deploying Natural Language Processing (NLP) algorithms to analyze sustainability reports, corporate news and social media communications to help collect detailed and up-to-date ESG information, overcoming the reliance on static agency ratings. Additionally, these algorithms can improve the identification of complex patterns and correlations between financial and ESG factors that would otherwise be overlooked with traditional instruments, leading to more detailed predictive assessments of sustainability performance.

Conclusively, MCDA has laid the foundation, but its definitive future potential will be realized through the complete synergistic integration with AI technologies to create more transparent and sophisticated sustainable investment frameworks.

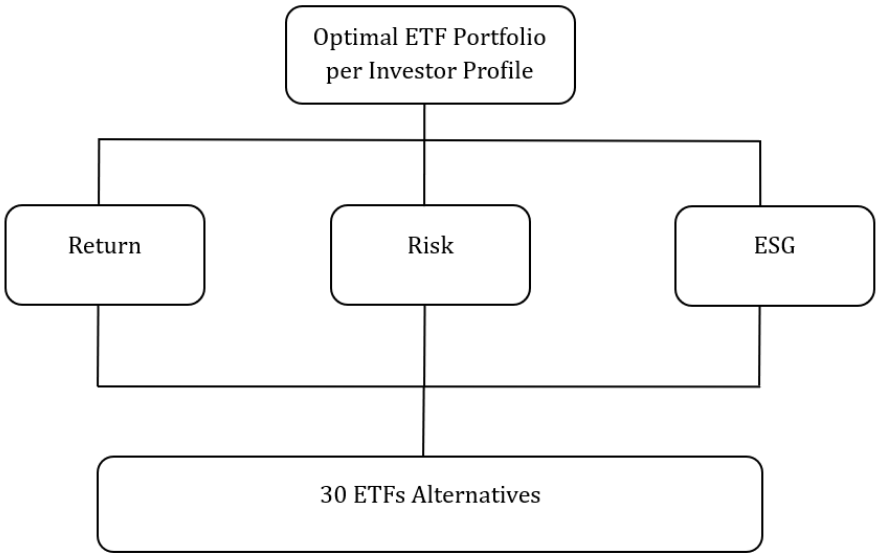
3. Theoretical Methodology: Decision Framework

3.1 Introduction

This section provides a detailed exposition of the proposed theoretical framework, structured as a six-step sequential and interdependent model to support the construction of tailored investment portfolios. The methodology addresses the complexities of heterogeneous investor preferences through the integration of multiple decision-making criteria, specifically three-year return, risk volatility and ESG performance.

The composition of this decision-making mechanism is detailed in *Figure 3.1*, which maps the transition from the primary investment objective to the evaluative criteria and their operational implementation through the set of investment alternatives.

Figure 3.1 – Structure of the Tri-Criterion Decision Framework



This structure provides the analytical foundation for the quantitative component of the model and enables the transition from qualitative preference representation to utility-based prioritization of the investment alternatives.

The Multi–Criteria Decision Analysis (MCDA) constitutes the central pillar of the model, establishing the structure for the decision problem and the quantification of the relative importance assigned by the investors to the evaluative criteria. To account for investor heterogeneity, the model utilizes the Dirichlet distribution to generate a dataset of preference vectors, where each vector defines a unique investor archetype with a specific priority distribution across the selected dimensions. This stage leads to the application of K-Means clustering to identify homogeneous investor segments, where cluster centroids act as representative profiles for the empirical assessment of preferences.

With the cluster-specific weights derived, the framework evaluates a fixed and normalized sample of Exchange-Traded Funds (ETFs) pre-selected for their diverse exposures in sustainability, risk and return performance. These normalized performance measures are combined with the corresponding preference weights to compute the Aggregated Utility Scores (AUS) that quantify the suitability of each alternative according to the investor priorities.

The final stage leverages Quadratic Programming (QP) to derive capital allocations for each segment by minimizing the deviation between AUS and portfolio weights, operating under a budget identity constraint and predefined exposure limits. Ultimately, both portfolios and investment alternatives are jointly projected within the tri-criterion space to provide a comprehensive representation of the utility landscape. The subsequent sections will delineate each phase of the framework, providing the rigorous mathematical foundations, procedural clarity and theoretical justification required to validate the conceptual integrity of the decision-making model.

Figure 3.2 – Six-Step Framework

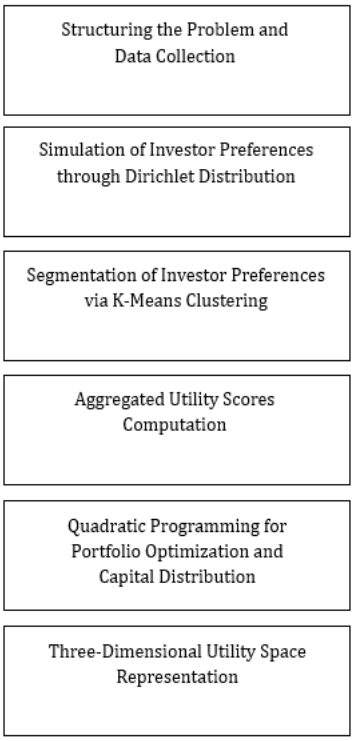


Figure 3.2 maps the operational sequence of the framework, with each modular component designed to perform specific analytical functions to ensure methodological transparency and replicability.

3.2 Methodology

3.2.1 Structuring the Problem and Data Collection

The expansion of modern finance has progressively transitioned the field of sustainable portfolio optimization from an exclusive focus on financial dimensions to addressing more multifaceted decision problems for corporate and institutional stakeholders. This multidimensionality necessitates the adoption of robust methodologies for integrating broader considerations and investor-specific preferences into capital allocation strategies.

Among the principal methodologies, Multi–Criteria Decision Analysis (MCDA) offers a highly systematic method for preference measurement and quantification. These frameworks facilitate the transparent and replicable assessment of priorities across competing objectives by decomposing highly-dimensional decision problems into analytical frameworks comprising goals, criteria and alternatives (Salo et al., 2024). The formal synthesis of qualitative judgments and quantitative metrics allows these approaches to reconcile divergent objectives within complex market environments, effectively transposing conflicting priorities into a rigorous decision-making process (Belton & Stewart, 2002).

The initial application of this framework involves simulating a comprehensive dataset of investor preferences, expressed as priority vectors designed to capture a broad spectrum of realistic attitudes, which are categorized into three distinct preference clusters:

· The Return-oriented segment focuses on financial performance by seeking portfolios with the highest expected returns, thereby relegating risk and ESG to secondary considerations

· The Risk-oriented segment prioritizes volatility attributes as the primary determinant for portfolio construction, manifesting a deliberate focus on market sensitivity. This framework operationalizes such behavior through normalized risk (σ_j^{norm}), where higher values denote a more pronounced exposure to price fluctuations.

By modeling risk as a central decision-making dimension, the model accounts for a wide range of preferences, spanning from conservative to aggressive, while supporting diversification across instruments with differing fluctuation profiles and enabling adaptive allocation strategies in dynamic market environments.

· The ESG-oriented segment focuses on sustainability and ethical considerations, designating the integration of Environmental, Social and Governance dimensions as the primary objective, even when this necessitates a reduced emphasis on conventional financial trade-offs (Amel-Zadeh & Serafeim, 2017). Although the ESG dimension incorporates an extensive plurality of thematic determinants, its formalization as a singular aggregate variable encompasses the necessary dimensions to constitute the essential foundation for sustainability-driven strategic orientations.

At the conceptual level, the environmental pillar captures and quantifies investment contributions to ecosystem preservation and natural resource stewardship, addressing the mechanics of the climate transition and the deployment of circular production models, ranging from biodiversity protection and decarbonization to the optimization of material life cycles. The social dimension reflects aspects related to the interaction between economic activity and societal structures, accounting for labor conditions, employee well-being, community engagement and inclusive economic participation. Together, these dimensions foster an interdependent relationship between allocative strategies and outcomes related to social cohesion, workforce stability and equitable value distribution, while the governance pillar defines the institutional structures, including board independence and transparency, that fundamentally regulates the management of operational and reputational risks.

To formalize these heterogeneous preference structures and map multidimensional qualitative pillars into quantitative segments, the requirement for granular insight extraction from unlabeled financial data positions clustering as the primary unsupervised learning technique (Jain et al., 1999).

Among various algorithmic techniques, K-Means clustering (MacQueen, 1967) optimizes computational efficiency through the minimization of the within-cluster sum of squares (WCSS), employing an iterative process that assigns observations to the proximal centroid. In this multidimensional feature space, the Euclidean distance quantifies the geometric separation required to drive the system toward a state of convergence.

The adoption of Euclidean distance as the primary metric of geometric proximity aligns with the intrinsic structure of the preference data and the analytical requirements for statistically salient partitioning in a normalized multidimensional space. Represented as continuous vectors over a common scale, investor preferences are articulated whereby each dimension contributes proportionally to the aggregate configuration of the observation. Within this setting, the adopted geometric representation supports a stable and interpretable partition of the preference space, allowing differences among profiles to be assessed consistently while preserving analytical tractability and coherence with the operation role assigned to clustering within the framework (Everitt et al, 2011).

The identification of analytically robust preference clusters facilitates the transition from descriptive profile mapping toward the formal synthesis of optimal portfolios, where mathematical programming functions as the mechanism for formalizing multidimensional qualitative orientations into deterministic capital allocation weights. Fundamentally, mathematical programming frameworks in portfolio management integrate objective functions with operational constraints to transform qualitative goals into solvable mathematical structures. The objective function defines the optimization target by prioritizing diverse performance metrics and strategic objectives, while the constraints establish the feasible region through the formalization of budget identities and multi-level allocation restrictions.

The architectural specification of these optimization problems typically categorizes into two principal computational paradigms:

· Linear Programming (LP) denotes optimization frameworks defined by strictly linear objective functions and constraints. Although the Mean-Variance model (Markowitz, 1952) traditionally relies on quadratic risk measures, LP formulations adopt linear proxies to improve computational scalability for large-scale portfolio configurations and ensure global optimality within convex environments in which linear approximations reflect the underlying risk preferences.

· Quadratic Programming (QP) encompasses optimization frameworks that include quadratic objective functions subject to linear constraints. Unlike LP, which leverages linear proxies, QP explicitly internalizes the second-order nature of financial risk by expressing portfolio variance as a quadratic measure of asset weights. This mathematical structure allows for the precise delineation of the efficient frontier, enabling the determination of optimal configurations that satisfy the multi-criteria trade-offs inherent to different investor profiles (Boyd and Vandenberghe, 2004).

The methodological implementation of these multi-criteria trade-offs is dependent on the employment of efficient market vehicles capable of supporting the dynamic requirements of the optimized portfolios.

In this context, the Exchange-Traded Funds (ETFs) provide the operational vehicles for implementing data-driven portfolio strategies, with structural features, including liquidity, transparency and cost efficiency, supporting dynamic portfolio rebalancing while ensuring faithful replication of the target asset allocation, particularly when navigating socially responsible structures. The in-kind creation and redemption mechanism contributes to price efficiency by maintaining alignment between market price and underlying net asset values, reinforcing the suitability of these instruments for systematic portfolio implementation (Madhavan and Sobczyk, 2014).

The evolution from unidimensional financial optimization toward multidimensional analytical frameworks reflects the necessity of addressing heterogeneous investor objectives within contemporary decision-making. While traditional models provide a robust foundation for asset allocation, extending these structures to incorporate multiple criteria enables a flexible representation of return, risk and sustainability within a unified architecture. Through the synergistic interplay between preference segmentation and operational efficiency of modern investment instruments, the proposed framework facilitates the systematic alignment of different priorities with feasible portfolio configurations, offering a transparent and scalable approach for integrating financial targets with broader considerations and preserving the analytical rigor for practical implementation (Utz et al., 2014).

3.2.2 Simulation of Investor Preferences through Dirichlet Distribution

Accounting for the inherent variability of investor priorities represents a structural prerequisite for personalized portfolio construction, as undifferentiated frameworks tend to obscure critical latent distinctions within investor utility functions. Due to the unobservable nature of individual preference weights within structured datasets, the model employs the Dirichlet distribution to generate preference vectors that satisfy the axiomatically required constraints of non-negativity and unit sum, ensuring a mathematically consistent representation of priorities without introducing arbitrary biases.

The Dirichlet distribution represents a continuous multivariate probability distribution conventionally adopted to model proportion or weights subjected to specific structural constraints. The broad utility in modeling compositional data stems from a flexible mathematical structure, making it particularly suitable for applications involving proportion-constrained multivariate representations (Aitchison, 1982).

Formally, the distribution is defined over the standard simplex, a geometric space encompassing vectors of non-negative real numbers whose collective components sum to unity. In the case of D dimensions, the $(D-1)$ dimensional simplex is given by:

$$T^{D-1} = \{p_i \in R^D \mid p_i \geq 0 \text{ for all } i, \sum_{i=1}^{D} p_i = 1\}$$

The geometric configuration of the simplex is determined by its dimensionality, representing a line segment at $D = 2$, expanding into a triangular surface at $D = 3$, and defining a three-dimensional tetrahedron at $D = 4$, rendering the space both closed and bounded.

The probability density function (PDF) of the Dirichlet distribution defines the likelihood of a specific composition vector $p = (p_1, p_2, \dots, p_D)$ occurring within the simplex, where, for a given parameter α, the density is formalized as:

$$f(p; \alpha) = \frac{1}{B(\alpha)} \prod_{i=1}^{D} p_i^{\alpha_i - 1}$$

In this formulation, p_i represents the components of the preference vector, α_i the concentration parameters that shape the distribution and $B(\alpha)$ the multivariate Beta function for ensuring normalization. The function assigns higher probability density to vectors whose components align with the expected proportions implied by the α parameters, such that if $\alpha_i > \alpha_j$, vectors displaying a larger p_i component will exhibit a higher likelihood than those with a larger p_j.

The distribution exhibits analytically tractable statistical properties contingent upon the parameter vector α:

· $E[p_i] = \frac{\alpha_i}{\alpha_0}$ where $\alpha_0 = \sum_{j=1}^{D} \alpha_j$

· $\mathrm{Var}(p_i) = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0+1)}$

· $\mathrm{Cov}(p_i, p_j) = -\frac{\alpha_i \alpha_j}{\alpha_0^2(\alpha_0+1)}$

These expressions, reflecting respectively the expected value, variance and covariance between different components p_i and p_j $for (i \neq j)$, delineate the fundamental features of the distribution. From a Bayesian perspective, the Dirichlet distribution functions as the conjugate prior for the multinomial distribution (Ferguson, 1973), while simultaneously constituting the multivariate generalization of the Beta distribution, which is recovered in the bivariate case where $D = 2$ (Connor & Mosimann, 1969).

The standard approach for sampling from a Dirichlet distribution entails a two-stage process through Gamma distributions:

· The initial stage focuses on the generation of Gamma variables: for each component $i = 1, \ldots, D$, a random value $G_i \sim \mathrm{Gamma}(\alpha_i, 1)$ is sampled from a Gamma distribution parametrized by a shape of α_i and rate of 1.

· The second stage involves the normalization of the vector $G = (G_1, \ldots, G_D)$ by dividing each G_i by the sum of all Gamma values $\sum_{j=1}^{D} G_j$. This operation yields the elements of the Dirichlet-distributed vector p as $p_i = \frac{G_i}{\sum_{j=1}^{D} G_j}$.

Furthermore, each marginal component p_i follows a Beta distribution $Beta(\alpha_i, \alpha_0 - \alpha_i)$ where $\alpha_0 = \sum_{j=1}^{D} \alpha_j$, implying an intrinsic negative correlation with the remaining components due to the unit sum constraint. The resulting vector $p = (p_1, \dots, p_D)$ lies within the $(D-1)$-dimensional standard simplex, ensuring that generated vectors constitute valid probability distributions whose central tendency and dispersion are governed by the concentration vector α.

Beyond the formal properties, the Dirichlet distribution maintains a distinct epistemological status as the maximum non-committal distribution on the simplex when constrained by a known expected value. In this sense, it facilitates preference simulation while precluding exogenous assumptions, ensuring that probability mass is allocated exclusively according to the specified constraints to derive unbiased profiles reflecting structural requirements instead of arbitrary specifications. From a practical perspective, the sampling of the vectors that quantify the relative importance that each investor assigns to the three criteria is conducted without a fixed random seed. This seed-free configuration prevents artificial determinism in the generated data and ensures each execution produces a genuinely stochastic and statistically solid sample of preferences.

Mathematically, these vectors are sampled from the 2-dimensional standard simplex T_{D-1}, a subset of R^3 characterized by the previously established constraints. The sampling process starts with the specification of a symmetric distribution parameterized by the concentration vector $\alpha = (\alpha, \alpha, \alpha)$, where $\alpha > 0$. For each simulated investor, a random sample from the Dirichlet distribution generates a unique preference vector: $p_i \sim \text{Dirichlet}(\alpha)$ for $i = 1, \dots, N$.

The full set of preference vectors is collected into a matrix $P \in R^{N\,x\,3}$, where each row corresponds to a single investor, representing the input for the clustering process and allowing for the identification of latent decision logics shared across groups of investors.

The employment of a symmetric Dirichlet distribution, where all components of the concentration vector are equal, establishes an epistemologically neutral starting points through the parametrization $\alpha = (2,2,2)$, implying that:

· The Expected Value of each component p_i of a Dirichlet-distributed vector is given by $E[p_i] = \frac{\alpha_i}{\alpha_0}$ with $\alpha_0 = \sum_{j=1}^{3} \alpha_j$. Substituting the chosen symmetric parametrization of $\alpha_i = 2$ for all i, this yields $E[p_i] = \frac{2}{6} = \frac{1}{3}$, implying that, on average, the generated preference vectors are centered around equal weighting across the three criteria.

· The total concentration parameter $\alpha_0 = 2 + 2 + 2 = 6$ directly influences the variance $\mathrm{Var}(p_i) = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0+1)} = \frac{2\cdot(6-2)}{6^2\cdot(6+1)} = \frac{2\cdot4}{36\cdot7} = \frac{8}{252} \approx 0.0317$ and ensures a high level of dispersion. This level enables diversity in simulated investor profiles and, despite the average preference being balanced due to the symmetric parametrization, individual vectors may still reflect investors with stronger leanings toward criteria for preserving heterogeneity.

The simulated vectors maintain a neutral orientation through a symmetric parametrization governed by α_0, yielding sufficient diversity and obviating the need for additional normalization. Following the synthesis of preference population, the deployment of a clustering architecture facilitates the aggregation of the generated observations, transposing individual-level heterogeneity into a reduced set of representative archetypal profiles.

3.2.3 Segmentation of Investor Preferences via K-Means Clustering

3.2.3.1 Investor Preferences Clustering

The multivariate dimensionality and inherent heterogeneity of the simulated investor preference population necessitate a systematic mechanism for dimensionality reduction to ensure analytical tractability. The adoption of the clustering algorithm addresses this requirement by partitioning data into homogeneous segments through an unsupervised learning technique that operates independently of predefined group knowledge, diverging from the categorical dependency of supervised architectures. By exploiting the internal structure of the dataset, this approach allows the identification of latent configurations that would otherwise remain unobserved within an unlabeled preference space.

The primary objective concerns the assignment of observations to distinct groups wherein within-cluster similarity exceeds between-group dissimilarity, converting individual-level heterogeneity into a refined set of interpretable aggregate profiles. To ensure methodological integrity, the resulting segmentation requires semantic interpretability and structural robustness, rendering each cluster a conceptually coherent category suitable for subsequent analytical modeling.

In line with these principles, K-means operates as a partitioning algorithm designed to organize observations into mutually exclusive clusters, ensuring each observation is associated with the most proximal centroid. Formally, the algorithm minimizes the objective function defined as the total within-cluster sum of squares (WCSS):

$$\min \sum_{k=1}^{K} \sum_{p_i \in C_k} |p_i - c_k|^2$$

in which p_i represents the observation assigned to cluster C_k, while c_k denotes the centroid, with the function measuring the total squared Euclidean distance between points and their respective cluster centers. The algorithm iteratively refines the assignment of observations by continually updating cluster centroid positions and reallocating data points based on relative geometric proximity.

The procedure entails the initialization of a predetermined number of centroids, representing preliminary cluster centers, typically derived from a random sampling mechanism or more robust initialization architectures. Upon this configuration, observation assignment is managed by a geometric proximity defined by the Euclidean distance.

$$\text{dist}(p_i, C_k) = \sqrt{\sum_{d=1}^{D} (p_i - C_k)^2}$$

By preserving the geometric structure imposed by unit-sum constraints, the Euclidean metric establishes a consistent proxy for proximity within the bounded preference space.

Within the proposed framework, the iterative reassignment of observations and subsequent centroid updates progressively minimize intra-cluster dispersion, driving the optimization toward locally stable configurations. The process reaches convergence when additional iterations yield negligible variations in cluster membership, indicating the attainment of a stationary point in the variance-based objective function consistent with the assumed convex and isotropic structure of the space. Consequently, this stability ensures that the resulting partition reflects the interpretable groupings grounded in the intrinsic geometry of the investor preferences.

Each investor is associated with a preference vector $(p_1, p_2, p_3) \in R^3$, where each component represents the relative importance assigned to the specific decision criterion. The entire population of simulated investors is mapped into a matrix $P \in R^{N\times 3}$, where rows represent individual investors and columns correspond to the respective decision criteria. The compositional nature of these vectors, evidenced by non-negativity and unit-sum requirements, geometrically constrains the dataset to the 2-dimensional simplex T_{D-1}, with $D = 3 \Rightarrow T_2 \subset R^3$. Although each preference vector comprises three dimensions, one for each criterion, the constraints restrict the data to a 2-dimensional simplex embedded in R^3, resulting in two intrinsic degrees of freedom.

To mitigate biased results stemming from initialization sensitivity, the algorithm executes multiple runs with randomized starting points, leveraging the 'nstart' parameter. Each investor is assigned to the cluster whose centroid minimizes the squared Euclidean distance within the constrained three-dimensional space.

The selection of $k = 3$ for the clustering algorithm represents the minimum partition required to preserve semantic coherence within the decision space, deriving from its consistent emergence as an empirically robust configuration across simulations (Hennig, 2015). The configurations where $k < 3$ resulted in the collapse of distinct evaluative dimensions, whereas higher values produced clusters that were either unstable across iterations or lacked semantic differentiation, effectively fragmenting the dominant preference logic without increasing the explanatory power of the model. Therefore, this segmentation enables the identification of three coherent investor profiles, each reflecting a unique prioritization logic and aligning with the established metrics for structural stability. The output of the K-means clustering includes:

· A cluster assignment vector $z = (z_1, z_2, \dots, z_n)$, $z_i \in \{1,2,3\}$, $z \in \{1,2,3\}^N$ wherein each element z_i indicates the cluster membership of the investor i

· A set of three centroids $c_{k1}, c_{k2}, c_{k3} \in R^3$, each representing the average preference configuration of the investor in its corresponding cluster C_k

Alternative clustering paradigms, such as hierarchical and density-based frameworks, proved analytically incongruent, evincing a specific reliance on distance thresholds or distinct linkage criteria incompatible with the geometric scale and with the compositional structure of the simplex data. The segmentation primarily classifies individuals by observable attributes to uncover underlying evaluative patterns that guide how distinct groups prioritize financial return, risk and ESG considerations. The compositional and unlabeled nature of the input vectors ensures that the resultant clusters are composed exclusively by the intrinsic configuration of preferences, representing an authentic logic independent of imposed classifications.

3.2.3.2 Internal Validation of Clustering Solution – Silhouette Score

The silhouette score (Rousseeuw, 1987), a widely recognized internal validation method, assesses the interpretability and the empirical soundness of the clustering structure by evaluating the observation's intra-cluster cohesion relative to its separation from the nearest alternative segment.

For a given observation i, the silhouette coefficient (h_i) is computed as:

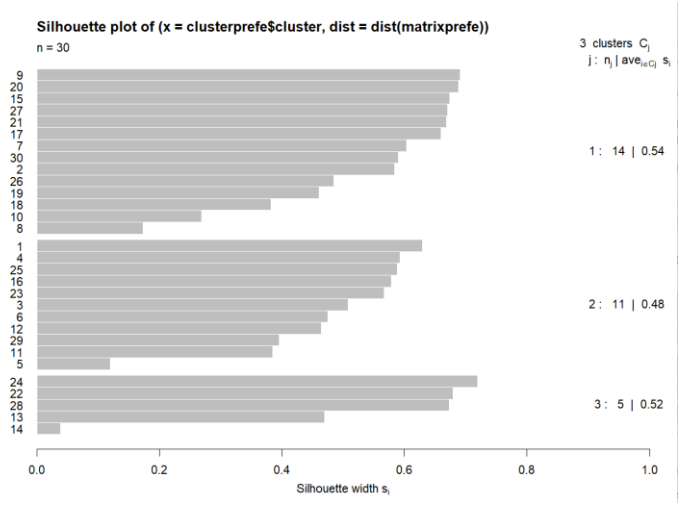
$$h_i = \frac{b_i - a_i}{\max(a_i, b_i)}$$

In this formulation, b_i represents the nearest-cluster dissimilarity, while a_i defines the average dissimilarity of observation i with all other points in the same cluster, thereby quantifying the intra-cluster cohesion.

This score, bounded within the interval [-1,1], represents a granular metric for evaluating the structural integrity of the segmentation. The values close to zero denote spatial positioning proximal to the decision boundaries of the data space, whereas negative values highlight potential misclassifications where the dissimilarity of the assigned cluster exceeds the distance to the nearest alternative. The global average across the aggregate provides a pragmatic assessment of clustering efficacy, where values exceeding 0.5 generally attest to the emergence of a robust underlying structure.

Despite the broad adoption, the silhouette score fundamentally relies on the assumption of reasonably convex and isotropic cluster structures, which may limit its interpretability when applied to datasets with irregular geometries or significant size imbalances. The corresponding plot acts as a graphical diagnostic tool, enabling a clear visual representation of clustering structure quality by displaying the individual silhouette scores for each observation, organized by cluster.

Figure 3.3 – Silhouette Score Plot



The individual observations are represented by horizontal bars, with their length directly reflecting their silhouette value (h_i), with longer bars indicating a higher degree of cohesion within the assigned cluster. The plot organizes observations vertically by cluster, creating distinct blocks for clear visual assessments of within-cluster consistency. Adjacent to each cluster, the plot reports the observation count and the average silhouette score for the cluster for an effective summary of its quality.

This internal data-driven mechanism for evaluating the clustering structure adequacy independently of external labels, functioning as a comparative benchmarking tool for alternative solutions rather than an absolute performance measure.

3.2.3.3 Centroid Extraction and Preference Parametrization

The validation of the internal clustering structure provides the methodological foundation for interpreting the cluster centroids, which, as arithmetic averages of investor preference vectors, act as archetypal representations of collective decision logic. The analysis of these compositions identifies the dominant strategic orientations within each segment and enables the construction of preference-aligned structures. From the mathematical perspective, the computation of each centroid corresponds to the arithmetic mean of the Dirichlet-simulated preference vectors within that cluster (Hartigan and Wong, 1979).

The centroid c_k is calculated as:

$$c_k = \frac{1}{n_k} \sum_{p_i \in C_k} p_i$$

In this formulation, p_i represents the preference vector for each observation, and n_k indicates the number of observations within the cluster C_k. The centroid c_k constitutes a three-dimensional vector representing the aggregate prioritization of a group, where the components c_{k1}, c_{k2} *and* c_{k3} correspond to the preference weights assigned to return, risk and ESG dimensions, respectively.

Therefore, the resulting vectors, denoted as $c_k = (c_{k1}, c_{k2}, c_{k3})$, are subsequently repurposed from descriptive statistics into operational preference weights for the following analytical stage of the model. More precisely, these compositional weights provide the quantitative foundation for the multi-criteria aggregation process, acting as behavioral parameters that express the relative importance assigned by each investor cluster to three decision criteria. This framework obviates the limitations of subjective preference elicitation by exploiting a neutrally parametrized Dirichlet distribution for statistical data generation. Such a design ensures a scalable and data-driven architecture where investor weights are established through robust empirical simulation, effectively maintaining systemic consistency and precluding the need for external ad hoc assumptions.

From a structural perspective, each centroid acts as both a formalized decision logic and a representative behavioral profile, bridging the data-driven segmentation stage with the multi-criteria evaluation process. This duality establishes a principled framework that integrates unsupervised preference discovery with structured decision modeling, yielding a unified model that seamlessly connects clustering, preference estimation and multi-criteria portfolio optimization. In the subsequent analysis, these centroid-based weights are integrated into an additive utility aggregation framework to derive the Aggregated Utility Scores (AUS) for each investment alternative, reflecting its suitability relative to the preference structure of the corresponding investor cluster.

3.2.4 Aggregated Utility Scores Computation

The transition from behavioral preference modeling to utility-based evaluation is formalized through the introduction of the Aggregated Utility Scores (AUS), which constitute the composite attractiveness measures for each investment alternative, reflecting the quantitative alignment between an ETF and the distinct prioritization logic of a specific investor segment. Aligned with the Multi–Attribute Utility Theory paradigm (Keeney & Raiffa, 1976), the overall utility assigned to each alternative is defined as a function of performance across individual criteria and the relative weights attributed to those dimensions within the preference structure (Abbas, 2018).

Let each investor cluster k be represented by the centroid vector of preference weights:

$$c_k = (c_{k1}, c_{k2}, c_{k3}),$$

such that $c_{ki} \geq 0$ and $\sum_{i=1}^{3} c_{ki} = 1$, which expresses the preference weights assigned to return, risk and ESG, respectively. Analogously, let each investment alternative j be represented by its normalized performance vector, with each dimension normalized to a benefit-focused scale to ensure comparability across heterogeneous metrics:

$$g_j = \left(R_j^{\text{norm}}, \sigma_j^{\text{norm}}, ESG_j^{\text{norm}}\right) \in [0,1]^3$$

The Aggregated Utility Score for each alternative j, evaluated under the preference structure of the cluster k, is then defined as:

$$U_{kj} = \sum_{i=1}^{3} c_{ki} g_{ji} = c_k^{\top} g_j$$

or, expressed explicitly for the three decision criteria:

$$U_{kj} = c_{k1} R_j^{\text{norm}} + c_{k2} \sigma_j^{\text{norm}} + c_{k3} ESG_j^{\text{norm}}$$

Geometrically, the AUS induce a continuous utility mapping over the tri-criterion space, whereby each ETF is represented as a discrete point whose scalar projection onto the cluster-specific preference vector determines its aggregated utility, with higher values indicating a more robust congruence with the weighting structure of the segment.

The restriction of preference weights and performance metrics to the unit interval ensures that the AUS remain bounded within [0,1]. This analytical synthesis converts the preference information embedded in each cluster centroid into a continuous utility measure, formalized through an additive aggregation of marginal utilities under the assumption of mutual preferential independence.

The integration of financial and non-financial dimensions within a unified numerical domain establishes formal equivalence between heterogeneous criteria, allowing sustainability to align symmetrically with economic performance to reflect the intricate and compensatory trade-off mechanisms inherent in professional investor behaviors. Additionally, in contrast to subjective elicitation approaches, this model employs an empirical and data-driven estimation of weights derived directly from cluster centroids, ensuring that additive structure maintains the convexity required for the seamless integration of utility values into subsequent stages without behavioral distortions. The capacity to integrate these objectives through utility-based architectures facilitates the subsequent derivation of concrete portfolio allocations, through Quadratic Programming (QP), and ensures simultaneous mathematical consistency and behavioral interpretability across all multi-dimensional investment decisions.

3.2.5 Quadratic Programming for Portfolio Optimization and Capital Distribution

The operationalization of preference-based utility metrics into implementable portfolio configurations necessitates an explicit optimization layer to formalize investor priorities subject to operational and regulatory constraints. This computational stage moves beyond unconstrained allocations derived from raw utility scores, which typically lack the structural robustness essential for portfolio stability, by adopting a constrained optimization framework to reconcile mathematical rigor with the practical requirements of asset management.

The general formulation of Quadratic Programming (QP) problems for portfolio optimization is expressed as:

$$\min_{x} \ \frac{1}{2} x^{\top} Q x + \vec{r}^{\top} x$$

subject to the linear constraints $Ax \ \leq b$ and $A_{eq} x = b_{eq}$ alongside the boundary limits $l_c \leq x \leq u_c$, employs an allocation vector x comprising the decision variables x_i that correspond to the percentage of capital assigned to the i-th constituent asset in the portfolio.

Under the proposed framework, the objective function is formally designed to minimize the structural Euclidean distance between the portfolio weight vector x and the utility architecture, defined by the AUS vector $u = (u_1, u_2, \dots, u_n)^{\top}$, which reflects the relative utility of each alternative for a specific investor cluster:

$$\min_{x} (x - u)^{\top} (x - u)$$

By algebraically expanding the expression:

$$(x - u)^{\top}(x - u) = x^{\top} x - 2u^{\top} x + u^{\top} u$$

This formulation represents a Quadratic Programming structure, in which the quadratic term $x^{\top} x$ is defined by the identity matrix $Q = 2I$, while the linear vector $\vec{r}$ is defined as $-2u$. The linear term $-2u^{\top} x$ acts as the directional driver toward the reference utility vector, ensuring the optimization remains contingent upon segment-specific preferences, providing the requisite target bias preventing convergence toward a trivial

zero-vector solution. The constant term $u^{\top}u$ remains analytically neutral in the optimization process, as its independence from the decision variable x maintains the coordinates of the minimum point unaffected across the three portfolios. This mathematical reformulation shifts the focus from variance-minimization objectives prioritizing risk-based interactions to a preference-alignment approach, leveraging a distance minimization approach to preserve the integrity of the multidimensional utility scores while ensuring that final weights represent a direct reflection of the predefined priorities.

The objective function of the optimization problem is formally expressed as:

$$\min_{x} \sum_{i=1}^{n} (x_i - u_i)^2$$

Here, $x = (x_1, x_2, \dots, x_n)^T \in R^n$ constitutes the vector of portfolio weights, representing the portion of capital allocated to each of the n assets, while $u = (u_1, u_2, \dots, u_n)^T \in R^n$ the reference vector of AUS, specific to the investor cluster under consideration. By treating each asset-level deviation $(x_i - u_i)$ symmetrically and independently, the model imposes a uniform penalty structure that avoids implicit prioritization across assets and preserves neutrality with respect to the dimensional composition of the decision space.

Reflecting the quadratic structure of its components, the strictly convex and differentiable nature of the objective function across R^n ensures a unique global solution within the feasible region (Mencarelli & D'Ambrosio, 2019), driving the portfolio toward the target vector while linear constraints enforce the financial and regulatory pragmatism necessary for the consistency of the final allocations.

The budget identity constraint enforces a full capital allocation by requiring the asset weights to sum to unity, mathematically formalized as $\sum_{i=1}^{n} x_i = 1$, or $1^{\top}x = 1$ in vector notation. The short-selling restriction, expressed as $x_i \geq 0 \ \ \forall i \in \{1, \dots, n\}$, ensures all asset weights remain non-negative to satisfy institutional directives where short exposure is disallowed or strategically irrelevant (Jagannathan & Ma, 2003).

Ultimately, to promote diversification and mitigate concentration risk, the model establishes an upper-bound constraint on the capital weight allocations, $x_i \leq \overline{x_i}\ \forall i \in \{1, \dots, n\}$, in which the parameter x_i represents the maximum allowable exposure to asset i to prevent disproportionate concentration and foster portfolio stability.

The intrinsic modularity of the framework allows for significant divergence in asset allocation, as an alternative prioritized by return-oriented cluster, for instance, might receive a negligible weight within the ESG-oriented portfolio, ensuring model scalability for the seamless integration of the evolving preferences within dynamic advisory settings. The optimized configurations define specific coordinates within the three-dimensional utility space, where each position rigorously reflects the weighted averages of the underlying constituent ETFs to provide a geometric utility landscape, confirming the structural coherence and internal consistency of the preference-driven framework.

3.2.6 Three-Dimensional Utility Space Representation

The theoretical framework for visualizing the decision environment, structured across the three evaluation criteria, represents a retrospective geometric synthesis of the relationship between normalized performance metrics, aggregated utilities and optimized portfolio configurations. The projection of the optimization results into a unified reference space organizes the allocations within a geometrically interpretable architecture, reinforcing the analytical integrity of the model and illustrating the strategic positioning of the optimal solutions.

In this generalized representation, the three-dimensional utility visualization defines each investment alternative *j* as a discrete coordinate within the performance space, as defined by its normalized vector:

$$g_j = \left(R_j^{\text{norm}}, \sigma_j^{\text{norm}}, ESG_j^{\text{norm}}\right) \in [0,1]^3$$

This geometric projection maps each alternative to a point p_j in the normalized tri-criterion space according to:

$$p_j = (x_j, y_j, z_j) = g_j$$

The resulting set $\{p_1, p_2, \dots, p_n\}$ constitutes a point cloud within the tri-criterion cube whose spatial dispersion describes the structure of the available investment opportunities across the three dimensions.

For each investor cluster k the additive model defines the aggregated utility of an investment alternative $U_{kj} \in [0,1]$ as the weighted sum $U_{kj} = \sum_{i=1}^{3} c_{ki} \cdot p_{ji}$, where the preference weights satisfy $c_{ki} \geq 0$ and $\sum_{i=1}^{3} c_{ki} = 1$ (Ishizaka & Nemery, 2013). The chromatic distribution is parametrically mapped to the aggregated utility as:

$$\text{Color}(p_j) = f(U_{kj})$$

where the resulting gradient architecture $\nabla U_k = (c_{k1}, c_{k2}, c_{k3})$ quantifies the utility levels relative to the selected investor cluster. Within this tri-criterion space, increasing depth in tone denotes higher tiers of perceived utility, ensuring ordinal consistency throughout the performance domain.

Let $x = (x_1, x_2, \dots, x_n)$ denote the feasible portfolio weight vector derived from the optimization, with $x_i \geq 0$ and $\sum_{i=1}^{n} x_i = 1$. The location of the portfolio associated with cluster k is expressed as a barycenter P_k of the points p_j, weighted by the allocation vector x, specifically as:

$$P_k = \left(\sum_{j=1}^{n} x_j R_j^{\text{norm}}, \sum_{j=1}^{n} x_j \sigma_j^{\text{norm}}, \sum_{j=1}^{n} x_j ESG_j^{\text{norm}}\right) \text{ or } P_k = f\left(U_{kj}, g_j\right)$$

These coordinates represent the geometric projection of the optimal allocation vector within the normalized decision environment, in which the position of P_k reflects the equilibrium between the behavioral priorities and the feasibility constraints governing the portfolio space.

The additive formulation defines the iso-utility surfaces $S_k(u)$ as planes within the tri-criterion space representing the set of all points (r, s, e) that yield a constant utility level $u \in [0,1]$ for a given cluster:

$$S_k(u) = \{(r, s, e) \in [0,1]^3 : c_{k1} r + c_{k2} s + c_{k3} e = u\}$$

Therefore, each coordinate on such a plane identifies a distinct configuration of return, risk and ESG scores perceived as equally desirable according to the specific weighting structure of the investor segment. The visualization preserves strict monotonicity to ensure that incremental performance improvements generate non-decreasing utility outcomes, while the chromatic scale regulates how tonal variations reflect proportional utility differentials. Accordingly, the optimized portfolios reside within regions of maximal utility density, although concentration limits and budget identity constraints enforce a displacement from the unconstrained optimum to ensure practical investability.

Empirically, the return-oriented gradient aligns with the x-axis, whereas the risk-oriented cluster exhibits a diagonal trajectory and the ESG-oriented segment displays a vertical alignment. In all instances, the positioning of the portfolio markers validates the optimization stage by mapping cluster-specific prioritization logic onto the feasible set with mathematical fidelity, ensuring that the resulting allocations function as spatially consistent representations of the underlying utility gradients.

Figure 3.4 – Geometric Representation of Investment Universe and Optimized Portfolio

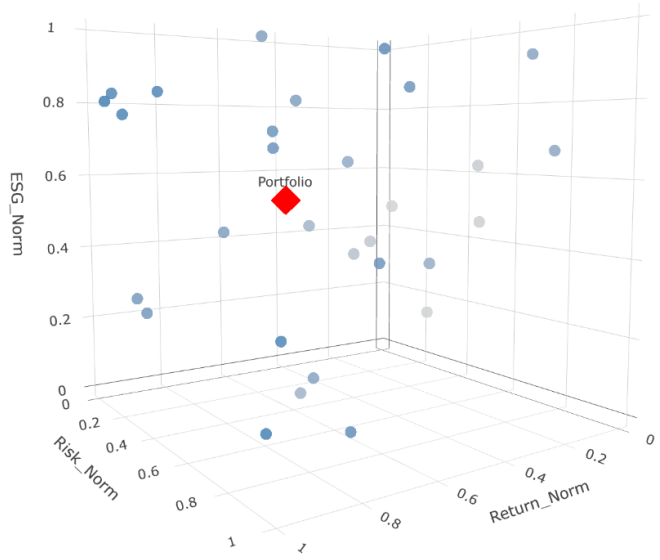


*Figure 3.*4 validates the coexistence of financial and sustainability attributes within an analytical space regulated by the principles of additive utility, where the blue spheres, defined by the vectors of normalized performances, constitute the feasible opportunity set whose spatial dispersion characterizes the structural diversity of the universe.

In that framework, chromatic intensity acts as a visual proxy for scalar utility, as darker hues denote higher AUS and lighter tones indicate lower composite attractiveness relative to the specific preference structure. The red diamond identifies the portfolio barycenter with coordinates defined by the weighted means of the normalized criteria, symbolizing the equilibrium of an allocation that maximizes coherence with the investor's utility landscape.

The integration of individual investment alternatives and optimized portfolios within a unified reference system formalizes the preference utility architecture into a tangible spatial representation, prioritizing the realization of specific behavioral profiles over the construction of a generalized efficient frontier. This visualization provides a behavioral map of aggregated utility, where the spatial convergence of the portfolio barycenter toward regions of peak chromatic intensity facilitates a robust transition from theoretical formulation to empirical implementation.

4. Framework Implementation for Tri-Criterion Portfolio Optimization

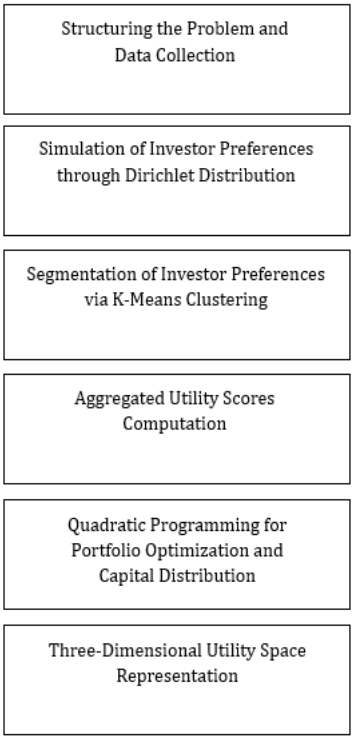
4.1 Introduction

The contemporary financial environment is currently being reshaped by structural transformation driven by the convergence of ecological transitions, the rapid expansion of digital technologies and a fundamental realignment of societal expectations regarding corporate accountability. Collectively, these forces maintain an ongoing reconfiguration of global economic dynamics and necessitate a precise refinement of investor opportunity assessments. Simultaneously, the exponential advancement in data availability and computational intelligence accelerates the diffusion of advanced decision-support systems, characterized by the institutionalization of machine learning and automated frameworks.

The increasing accessibility of sophisticated financial tools empowers a wider range of investors to navigate multifaceted information environments and synthesize strategies tailored to specific individual objectives. This structural shift necessitates a fundamental realignment of the operational logic within financial markets, where the supplementation of traditional metrics with alternative indicators transforms investor rationale into an insight-oriented process. Within this context, the capacity of analytical models to unify various performance criteria emerges as a critical advantage for addressing contemporary investment complexities.

Building upon these environmental shifts, this section provides an empirical implementation of a fully integrated decision-making architecture. The framework synthesizes simulated investor preferences, data segmentation and multi-criteria analysis to manage the final portfolio optimization process. Beyond the primary application, the analysis emphasizes methodological transparency and replicability through the exploitation of the R statistical programming environment to ensure both computational rigor and the integrity of the data processing.

Figure 4.1 – Six-Step Framework



The application of the architecture to a dataset of Exchange-Traded Funds (ETFs) operationalizes the preference-based portfolio construction through quantitative procedures and converts the theoretical methodology into actionable empirical results. The implementation of this framework generates three distinct portfolios reflecting heterogeneous investor preferences, with these profiles converted into capital allocations, ensuring transparent and data-driven results through a structured and replicable process.

4.2 Practical Analysis

4.2.1 Structuring the Problem and Data Collection

The selection of the ETF dataset is driven by an extensive representational scope across diverse asset classes, geographic exposures and thematic orientations, with the key performance indicators sourced from recognized open-access platforms, specifically Yahoo Finance, JustETF and ETFdb, to ensure the empirical validity of the model. The data granularity and consistency provided by these platforms facilitate the construction of a reproducible decision-making environment, culminating in a comprehensive dataset that reflects contemporary market metrics and supports the cross-criteria performance evaluation essential to the proposed optimization framework.

The variables act as descriptive proxies to support the methodological framework, prioritizing their structural role over granular historical performance benchmarking. Within this context, these metrics represent coherent orders of proportions and structurally consistent correlations between return, risk and ESG dimensions, in alignment with observable market dynamics. Consequently, this dataset provides a representative calibration of the fundamental trade-offs required to evaluate the three strategic portfolio profiles.

The investment alternatives are characterized by a predefined set of parameters, fully documented in *Table A.1*, encompassing the ETF name, Ticker symbol, Year of Launch and the Asset Under Management. This data architecture integrates specific performance and sustainability metrics including one and three-year return profiles and volatility risk alongside a comprehensive ESG evaluation. To ensure analytical transparency, the ESG dimension is also disaggregated into its constituent dimensions within the descriptive dataset, therefore providing the necessary granularity to validate the composite ESG score exploited in the multi-criteria optimization process. Their integration establishes a balanced framework for capturing both traditional and emerging investor priorities, while facilitating consistent normalization and aggregation techniques.

While the database encompasses a comprehensive range of financial and structural metrics, the analysis deliberately constrains the multi-criteria evaluation to three primary dimensions: three-year return, volatility risk and the aggregate ESG score. By prioritizing these core evaluative axes over other factors, the model ensures a high degree of analytical focus and internal coherence. This strategic selection aligns with the objectives of the model to establish a representative financial-sustainability system, where the numerical consistency and bounded scale of these variables facilitate robust normalization and integration within the analytical procedure.

The cross-sectional configuration of the database provides a static representation of the material performance metrics for each ETF across a consistent three-year temporal horizon. While the investigative strategy limits the observation of temporal dynamics, it effectively mitigates distortions attributable to short-term variations, ensures congruence with the underlying rationale behind preference-driven rankings and enables the deployment of clustering algorithms by eliminating structural noise arising from asynchronous and overlapping return periods.

Table 4.1 depicts a representative sample of five ETFs extracted from the dataset, delineating the structural characteristics of the input data prior to processing. For a comprehensive overview of the non-normalized performance of all 30 ETFs, refer to *Table A.1* in the Appendix.

Table 4.1 – Sample of Five ETFs

ETF Name	Ticker	Launch	AUM ($M)	Return	Risk	ESG Score
ISHARES EA	ESGD	2016	10,890	25.4	13.2	8.1
XTRACKER	USSG	2019	501,86	29.5	13.9	8.6
SPDR INT	DWFI	2019	80,160	23.4	12.7	8.2
CXSE CHINA	CXSE	2012	504,4	15.4	22.4	7.85
VANGUARD S	VSGX	2018	5,630	22.8	13.0	8.0

This example provides an empirical illustration of the unprocessed numerical composition of the dataset, underscoring the extended diversity of asset profiles prior to the normalization process.

The identified ETFs exhibit significant divergences in strategic orientation, risk-return dynamics and sustainability profiles, providing the necessary variability to inform the Multi-Criteria Decision Analysis (MCDA). Sourced directly from the originating platforms, these quantitative metrics initially feature disparate units of measure and varying scales that preclude direct comparison. To prevent methodological bias and ensure the integrity of the weighting procedure, a standardized normalization process is applied to harmonize the inherent disparities between indicator levels, facilitating a proportional assessment of performance across all dimensions.

The model leverages the Min-Max linear scaling technique to ensure the homogeneity of the ETF performance metrics, systematically transforming each numerical data point into a dimensionless value within the closed interval [0,1]. By calibrating the results to the specific range observed for each variable within the dataset, this approach guarantees ordinal consistency and preserves the relative positional relationship among all values for a given criterion. The normalized score for any individual observation within the dataset is derived as:

$$x_{ij\ norm} = \frac{\left(x_{ij} - x_{ij\ min}\right)}{\left(x_{ij\ max} - x_{ij\ min}\right)}$$

where x_{ij} represents the raw value of the criterion for a given ETF, while $x_{ij\ min}$ and $x_{ij\ max}$ denote the minimum and maximum observed values of the criterion across the entire column in the dataset.

As an empirical validation of this process, the *ISHARES EA (ESGD)* is evaluated through its raw return, risk and ESG metrics of 25.4, 13.2 and 8.1 against the respective dataset ranges of 15.4 to 35.2, 11.8 to 22.4 and 7.66 to 9.57.

· Normalized Return $\frac{25.4-15.4}{35.2-15.4} \approx 0.5051$

· Normalized Risk $\frac{13.2-11.8}{22.4-11.8} \approx 0.1321$

· Normalized ESG $\frac{8.10-7.66}{9.57-7.66} \approx 0.2304$

The impact of Min-Max normalization is further illustrated in *Table 4.2*, which provides the transformed values for the previously detailed set of five ETFs, enabling a direct assessment of the harmonization achieved through the scaling procedure. A detailed summary of the normalized values of the ETF performance is provided in *Table A.2* of the Appendix.

Table 4.2 – Sample of Five ETFs after Normalization (Min-Max Scaling)

ETF Name	Ticker	Launch	AUM ($M)	Norm_Return	Norm_Risk	Norm_ESG
ISHARES EA	ESGD	2016	10,890	0.5051	0.1321	0.2304
XTRACKER	USSG	2019	501,86	0.7121	0.1981	0.4922
SPDR INT	DWFI	2019	80,160	0.4040	0.0849	0.2827
CXSE CHINA	CXSE	2012	504,4	0.0	1.0	0.0995
VANGUARD S	VSGX	2018	5,630	0.3737	0.1132	0.1780

The adoption of column-wise Min-Max scaling reflects the requirement for bounded outcomes and the preservation of ordinal consistency within the multi-criteria synthesis. By independently rescaling each criterion, the procedure ensures that transformed values preserve the relative proportional distances of the original distributions, while avoiding both the unbounded results resulting from alternative procedures and the loss of distributional variance inherent in ranking-centric approaches. Although other scaling procedures generate different numerical ranges, qualitative allocation patterns and interpretative insights remain invariant to the specific normalization scheme, with Min-Max normalization ensuring cross-criteria commensurability without imposing distributional assumptions on the underlying data. Nevertheless, the rescaling process remains sensitive to the presence of outliers, as extreme observations might exert a disproportionate influence that compresses the scale and reduces the discriminative power of the indicators. Moreover, the assumption of linear comparability across the transformed range may not fully encapsulate non-linear shifts in investor preferences, necessitating a cautious interpretation of these standardized outputs within the final optimization stages.

The empirical implementation of this framework was conducted within the RStudio environment, which, despite the lower prioritization of interactive visualization tools, was selected for its robust computational architecture and extensive library support for multi-criteria decision analysis.

For transparency, the complete algorithmic structure is documented in *Script B.1* of the Appendix, with the key functional components presented within the text to delineate their specific roles in the architecture. To balance stochastic variety and methodological rigor, non-deterministic sampling was utilized for the preference generation, while a fixed initialization parameter is applied exclusively to the clustering stage. Collectively, these foundations provide the necessary prerequisites for a coherent transition from input structuring to model execution.

4.2.2 Simulation of Investor Preferences through Dirichlet Distribution

The practical framework begins with the simulation of investor preferences utilizing sampling from a Dirichlet distribution, which facilitates the representation of compositional data and the modeling of latent trade-offs among the three competing priorities. The resultant vectors formally capture the multi-dimensional investor rationale by assigning weights to each decision criterion, consequently establishing a structured preference input for the subsequent optimization stages.

The R-based implementation generates a matrix of thirty simulated investors comprising stochastic vectors constrained to unity, utilizing a computational process that deliberately omits a fixed random seed to maintain non-deterministic generation.

```
Code 4.1 – Simulation of Investor Preferences

preferences=function(n_investors=30,alpha=c(2,2,2)){

 prefs=as.data.frame(rdirichlet(n_investors,alpha))

 colnames(prefs)=c("Return_Pref","Risk_Pref","ESG_Pref")

 prefs$Investors=paste0("Investor",1:n_investors)

 prefs=prefs[,c("Investors","Return_Pref","Risk_Pref","ESG_Pref")]

 return(prefs)}
```

The `rdirichlet()` function, part of the `MCMCpack` package, is used to generate random samples from a Dirichlet distribution. While the `MCMCpack` is a specialized library for Bayesian analysis that extensively exploits Markov Chain Monte Carlo (MCMC) methods, it is important to note that the `rdirichlet()` does not employ these algorithms, rather it produces random vectors by specifying the number of observations and the corresponding vector of concentration parameters. The concentration vector, denoted as `alpha=c(2,2,2),` is selected to reflect moderate variation without favoring any specific dimensions.

The output is a matrix in which each row represents a simulated investor and each column a normalized preference weight. To illustrate the structure of the simulated preference data, *Table 4.3* presents a sample of five investors and their respective weights across the three dimensions. For a full analysis of the simulated preferences of the investors, including their assigned weights, refer to *Table A.3*.

Table 4.3 – Sample of Five Investor Preferences from Dirichlet Distribution

Investor	Return Preference	Risk Preference	ESG Preference
Investor 1	0.1350	0.6484	0.2166
Investor 2	0.7748	0.0253	0.1999
Investor 3	0.1971	0.1320	0.6710
Investor 4	0.5512	0.1715	0.2773
Investor 5	0.1475	0.1330	0.7195

Table 4.3 exemplifies and illustrates the distributional heterogeneity of the simulated investor sample. For instance, the first investor prioritizes the risk dimension by allocating a dominant 64.84% weight to that criterion, whereas the second profile exhibits a markedly divergent strategic logic by concentrating 77.48% of its preference weight on return. This variance highlights the capacity of the distribution to simulate a broad spectrum of investor rationales, ranging from highly conservative to aggressively performance-oriented.

The compositional structure of these profiles, with each vector constrained to a unit sum condition, ensures both logical coherence and mathematical compatibility with the unsupervised learning stage. By populating the decision-making environment with these internally consistent yet diverse preference structures, the framework establishes a robust foundation for the clustering algorithm to identify and formalize the three segments that guide the stages of the framework.

4.2.3 Segmentation of Investor Preferences via K-Means Clustering

4.2.3.1 Investor Preferences Clustering

To address the interpretative complexity of the simulated investor preference profiles, the framework identifies internally homogeneous segments of investors with convergent preference structures. This process facilitates the derivation of a representative profile for each segment, providing a stylized foundation for the subsequent allocation logic. For this purpose, the clustering procedure is executed within a three-dimensional space to partition the 30 simulated profiles into distinct clusters.

The segmentation is implemented through the `kmeans()` algorithm in R, utilizing a fixed configuration of three centers. To mitigate the risk of convergence to suboptimal solutions, the algorithm selects the most robust outcome from 25 random initializations through an iterative process that maximizes inter-cluster variance while simultaneously enhancing intra-cluster cohesion. Notwithstanding the development of specialized metrics for the analysis of compositional data, the Euclidean distance is adopted within this framework, as the normalization of the preference vectors ensures a more interpretable and geometrically consistent solution for the optimization space.

```
Code 4.2 – Clustering of Investor Preferences

matrix_preferences=preferences[,c("Return_Pref","Risk_Pref","ESG_Pref")]

set.seed(123)

cluster_preferences=kmeans(matrix_preferences,centers=3,nstart=25)

preferences$Cluster=cluster_preferences$cluster
```

The practical application of the clustering technique is illustrated in *Table 4.4*, which presents the normalized preferences and the resulting cluster assignments for the same five-investor sample previously examined. This tabular representation demonstrates how the distance-based partitioning converts into concrete groupings, effectively bridging the gap between the initial simulation of preferences and the formal identification of the three strategic profiles.

Table 4.4 – Sample of Five Investor Preferences from Dirichlet Distribution with the Assigned Cluster

Investor	Return Preference	Risk Preference	ESG Preference	Cluster
Investor 1	0.1350	0.6484	0.2166	2
Investor 2	0.7748	0.0253	0.1999	1
Investor 3	0.1971	0.1320	0.6710	3
Investor 4	0.5512	0.1715	0.2773	1
Investor 5	0.1475	0.1330	0.7195	3

The unsupervised clustering analysis yielded a robust segmentation of the simulated investor profiles, resulting in a three-cluster configuration composed of 14 investors in the first cluster, 11 in the second and 5 in the third. This distributional structure emerged from the intrinsic stochastic patterns within the data, whereby each cluster identifies a distinct preference configuration that serves as a representative archetype for the optimization process.

As reflected in the results, the first cluster is characterized by a dominant preference for return and a marginal emphasis on risk and ESG dimensions, identifying a performance-driven investor segment. The second cluster exhibits a moderate orientation on return combined with a significant weight allocated to risk, indicating a risk-prioritizing group with limited ESG sensitivity. Finally, the third cluster displays a more balanced structure with moderate preferences for both return and risk, while demonstrating the highest average ESG priority.

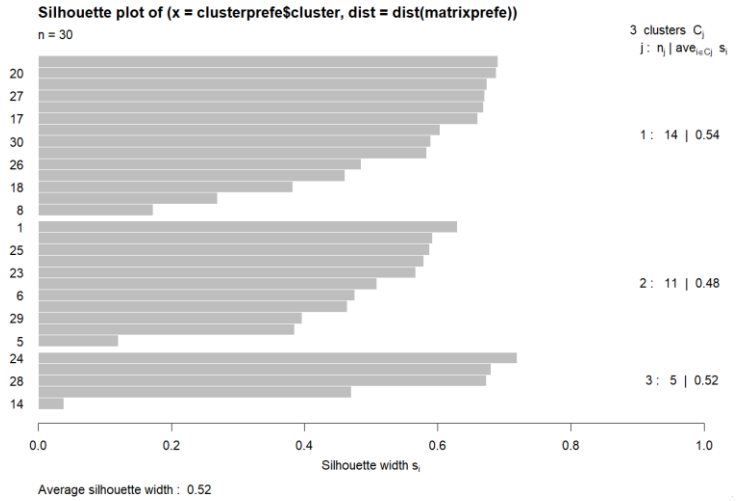
4.2.3.2 Internal Validation of Clustering Solution - Silhouette Score

To enhance the diagnostic assessment and validate the structural integrity of the clustering solution, the Silhouette score is leveraged to quantify the cohesion of each observation to its assigned cluster relative to its separation from the others. The `Cluster` package is employed to compute this score, while the `Factoextra` includes the functions for visualizing and interpreting the outputs.

```
Code 4.3 – Silhouette Score

sil=silhouette(cluster_preferences$cluster,dist(matrix_preferences))

mean(sil[,3]) #To compute the average Silhouette width
```

This data-driven approach provides empirical support for the three-cluster configuration, confirming that the solution reflects the intrinsic structure of the data rather than being the result of an arbitrary selection.

Figure 4.2 – Silhouette Score



The Silhouette analysis yielded a score of 0.516, providing a quantitative validation of the classification quality. This metric, which acts as a diagnostic value for assessing the structural integrity of the clustering, indicates a substantial degree of intra-cluster cohesion and inter-cluster separation. Considering that scores exceeding the 0.5 threshold are generally interpreted as evidence of a significant underlying structure, this result provides empirical support for the validity of the three-cluster segmentation.

To further validate this configuration, additional exploratory trials with alternative cluster numbers were performed, consistently resulting in lower average coefficients. This comparative assessment reinforces the empirical soundness of the solution and confirms its alignment with the intrinsic distribution of the data.

4.2.3.3 Centroid Extraction and Preference Parametrization

Leveraging this validated segmentation, the extraction of cluster centroids facilitates the transition from the conceptual identification of investor archetypes to the formal quantitative inputs required for the multi-criteria aggregation stage. Each centroid represents the mean normalized preference vector of its respective cluster, encapsulating a distinct collective prioritization logic that provides the foundation for the utility score computations.

```
Code 4.4 – Centroids extraction

centroids_pref=aggregate(. ~ Cluster, data=preferences[,-1], FUN=mean)

centroids_pref
```

The `aggregate()` function, operating on the preference dataset with investor identifiers excluded, summarizes the normalized preference dimensions by cluster, producing `centroids_pref` output for a concise and accurate representation of the group-level logic.

Table 4.5 – Centroids of the Three Clusters

	Return Preference	Risk Preference	ESG Preference
Cluster 1	0.5537	0.2011	0.2452
Cluster 2	0.2169	0.5390	0.2441
Cluster 3	0.1484	0.2064	0.6452

The first cluster, which comprises almost half of the sampled investors, exhibits a pronounced orientation toward financial performance. This is evidenced by a return weight exceeding 55%, while the risk and the ESG dimensions account for approximately 20% and 24% of the total allocation respectively, suggesting a tolerance for moderate risk and subordinating ESG considerations to the main goal of maximizing return.

The second cluster is defined by the centrality of the risk dimension, which represents the primary determinant of its collective investment logic. This significant focus might reflect either an opportunistic approach leveraging volatility for market gain or an acute sensitivity to risk exposure.

While the model remains neutral regarding these underlying motivations, the final portfolio configuration will reflect this role of risk, positioning risk as a factor to be either leveraged for opportunity or mitigated for protection.

The third one, which allocated nearly 65% of its decision weight to the ESG criterion, displays a profound deviation from conventional investment strategies that reflects a values-based rationale where non-financial considerations prevail over traditional metrics.

4.2.4 Aggregated Utility Scores Computation

The centroid vectors represent the inherent valuation frameworks that drive the investment preferences of each identified investor segment. These vectors represent the essential components for generating the Aggregated Utility Scores (AUS) and shaping portfolio compositions, acting as the analytical bridge that converts qualitative behavioral archetypes into the quantitative inputs required for portfolio optimization. By transforming cluster-specific priorities into empirical weight sets, this stage ensures that the resulting financial decisions are rigorously aligned with the data-driven logic of each group.

Specifically, the AUS for each ETF is derived through a weighted summation of its normalized performance across the three evaluation criteria. By leveraging the specific weights incorporated within the centroid vectors, this process quantifies the overall alignment of each investment alternative with the strategic objective of the cluster. This aggregation facilitates a rigorous comparative assessment of the alternatives while establishing the requisite ranking mechanism to identify the optimal instruments tailored to each unique investor profile.

The AUS for each investment alternative is formally expressed as:

$$U_{kj} = c_{k1}R_j^{\text{norm}} + c_{k2}\sigma_j^{\text{norm}} + c_{k3}ESG_j^{\text{norm}}$$

This formula employs a standard linear aggregation method to systematically combine normalized performances with cluster-specific decision weights, producing a single score for each investment alternative.

To illustrate the computational process, the AUS for the *ESGD* is calculated using its normalized performance metrics of Return at 0.5051, Risk at 0.1321 and ESG at 0.2304, in conjunction with the priority vector the of return-oriented cluster of 0.553728 for Return, 0.201097 for Risk and 0.245176 for ESG:

$$U_{kj} = (0.553728 \cdot 0.5051) + (0.201097 \cdot 0.1321) + (0.245176 \cdot 0.2304)$$
$$\approx 0.279688 + 0.026565 + 0.056489 = 0.362742$$

This score quantifies the congruence of the *ESGD* with the preference structure of the return-oriented cluster. *Table 4.6* exemplifies the AUS for the same representative sample of ETFs for an accurate comparison of their performance. The complete set of AUS for all the investment alternatives, organized by clusters, is reported in *Table A.4* in the Appendix.

Table 4.6 – Sample of Five ETFs with the relative Aggregated Utility Scores

ETF Name	Norm_Return	Norm_Risk	Norm_ESG	AUS_C1	AUS_C2	AUS_C3
ISHARES EA	0.5051	0.1321	0.2304	0.3627	0.2370	0.2509
XTRACKER	0.7121	0.1981	0.4922	0.5548	0.3813	0.4641
SPDR INT	0.4040	0.0849	0.2827	0.3101	0.2024	0.2599
CXSE CHINA	0.0	1.0	0.0995	0.2255	0.5634	0.2706
VANGUARD S	0.3737	0.1132	0.1780	0.2733	0.1855	0.1937

The *XTRACKER (USSG)*, for instance, is identified as the preferred asset for both return-oriented and ESG-oriented investors among this sample, highlighting its robust performance across the criteria. Conversely, the *CXSE CHINA (CXSE)* receives the highest valuation from the risk-oriented group, whereas its comparatively higher volatility and weaker ESG profile result in a lower valuation from the other two segments. The remaining assets, namely the *ESGE* and *SPDR INT (DWFI)*, indicate a moderate degree of alignment with the return-oriented cluster, while the *VANGUARD S (VSGX)* exhibits a generally low valuation across all three investor profiles.

The ordering of investment alternatives according to their descending AUS captures the behavioral heterogeneity embedded within the investor segments, accounting for the observation that the same instrument may rank differently across clusters as a consequence of distinct decision-weight structures. Although these scores are derived from a shared evaluative framework that enables quantitative comparability across different clusters, their interpretation remains cluster-specific, in the sense that they express the degree of alignment between an investment alternative and a given prioritization structure rather than constituting an absolute measure of performance.

4.2.5 Quadratic Programming for Portfolio Optimization and Capital Distribution

The cluster-specific AUS provide the foundation for the constraint-bound Quadratic Programming (QP) framework, which operationalizes the formalization of the investment strategies. The primary objective of this stage involves determining the optimal capital allocation for each investor group by identifying an allocation vector, representing the percentage exposure of each asset, which minimizes its squared Euclidean distance from the ideal AUS vector.

The model architecture is modified to reorient the optimization objective from conventional risk mitigation toward an alignment-based focus. From a mathematical standpoint, this transition is achieved by replacing the general covariance matrix, traditionally employed to capture asset risk and correlation, with a scaled identity matrix. This reconfiguration of the core optimization logic drives the maximization of alignment with the AUS while simultaneously satisfying all predefined portfolio constraints.

The optimization formulation is structured to align with the canonical form required by the `solve.QP` function:

$$\min_x \ \frac{1}{2} x^\top D x - d^\top x$$

subject to $1^\top x = 1,\ x_i \geq lc,\ \ x_i \leq uc$

The optimization objective aims to minimize the squared Euclidean distance between the portfolio allocation vector x and the target AUS score vector u, formally defined as $\sum_i (x_i - u_i)^2$.

To implement this objective within the optimization process, the model parameters necessitate a formal mathematical restatement to align with the required numerical inputs of the solver. This reparameterization involves the definition of the matrix D as a scaled identity matrix in which all diagonal elements equal to $2 (D = 2I)$, reflecting the quadratic component of the expanded objective function. Correspondingly, the linear coefficient vector d is defined as $2u$, in which u represents the target AUS vector.

To ensure computational efficiency and algorithmic alignment, the optimization framework incorporates a $\frac{1}{2}$ scalar multiplier within its objective function. Through the elimination of the differential coefficient when deriving the gradient Qx from the quadratic form $x^{\top}Qx$, this refinement removes redundant coefficients during the iterative identification of stationary points. Rather than being merely notional, this simplification significantly enhances computational tractability and promotes numerical stability while effectively accelerating algorithmic convergence toward the global minimum.

Beyond the associated mathematical advantages, most Quadratic Programming (QP) solvers, including the `Quadprog` package, operate under canonical objective-function specifications. The inclusion of this positive scalar factor does not affect the location of the optimal solution, as the multiplier acts uniformly across the objective function, leaving the underlying functional geometry and feasible region invariant. This preservation of the optimization landscape ensures that the resulting minimizer, defined as the optimal allocation vector, remains unchanged, with the influence of the factor limited exclusively to the absolute value of the objective function. In the context of portfolio selection, this absolute value is of secondary importance compared to the allocation vector, which constitutes the primary decision-making output. This flexibility allows for a seamless transition between optimization paradigms, distinguishing the traditional Markowitz model, which employs the covariance matrix as the Q operator to minimize variance, from the proposed formulation that restructures these parameters to emphasize the alignment with the Aggregated Utility Scores.

The portfolio allocation is subject to a budget constraint that prescribes the sum of all allocations equal to one, $(\sum_{i=1}^{n} x_i = 1)$, to a non-negativity constraint $x_i \geq 0$ for preventing short selling, and an upper-bound constraint, $x_i \leq u_c$, to avoid excessive concentration in individual assets. To satisfy the requirements of the `solve.QP` function, all the constraints must conform to the format of linear inequality constraints, specifically $Ax \geq b$.

This necessitates transforming the upper-bound constraint $x_i \leq u_c$ into the required format. By negating the inequality, the constraint is rewritten as $-x_i \geq -u_c$ to include a negative identity block within the constraint matrix. Accordingly, the lower-bound constraint $x_i \geq l_c$ does not require any reformulation because its structure is already consistent with $Ax \geq b$. The empirical implementation leverages the solve.QP function from the Quadprog package to process strictly convex quadratic programs.

Code 4.5 – Quadratic Programming procedure

```
optimize_alloc=function(scores,lc=0.005,uc=0.15){
 s=scores/sum(scores)            #Normalization
 n=length(s)
 Dmat=2*diag(n)
 dvec=2*s
 #Constraints Definition
 Aeq=matrix(1,n,1); beq=1        #sum(weights)=1
 A_lc=diag(n); b_lc=rep(lc,n)    #lower-bound constraint
 A_uc=-diag(n); b_uc=rep(-uc,n)  #upper-bound constraint
 Amat=cbind(Aeq,A_lc,A_uc)
 bvec=c(beq,b_lc,b_uc)
 solve.QP(Dmat,dvec,Amat,bvec,meq=1)$solution }
```

In this formulation, the constraints are encoded within the `Amat` and `bvec` vectors, in accordance with the required canonical form. The budget constraint, $\sum_{i=1}^{n} x_i = 1$, is handled as an equality by the first column of the constraint matrix `Aeq=matrix(1,n,1)` and the corresponding right-hand side `beq=1`. The `meq=1` parameter specifies the number of constraints in the system that should be treated as exact equalities.

The lower-bound constraint, $x_i \geq l_c$, is included in the identity block matrix `A_lc=diag(n)` and the `b_lc=rep(lc,n)` vector of lower-bound elements, while the restated upper-bound constraint, $-x_i \geq -u_c$ in `A_uc=-diag(n)` and `b_uc=rep(-uc,n)`.

By column-binding the individual constraint matrices, the final matrix `Amat` is constructed as `Amat=cbind(Aeq,A_lc,A_uc)` and the corresponding right-hand side is formed by concatenating their associated elements using `bvec=c(beq,b_lc,b_uc)`. This process ensures that the constraints are collectively combined into a single matrix and vector, as required by the solver.

The `solve.QP(Dmat,dvec,Amat,bvec,meq=1)$solution` function computes the unique global allocation vector that satisfies the budget identity and the inequality constraints while minimizing the squared Euclidean distance to the target vector, generating a cluster-specific set of optimal capital allocations for each investor profile.

```
Code 4.6 – Cluster-specific Portfolio Optimization

res_C1=optimize_alloc(etf_data$AUS_C1,lc=0.005,uc=0.15)

res_C2=optimize_alloc(etf_data$AUS_C2,lc=0.005,uc=0.15)

res_C3=optimize_alloc(etf_data$AUS_C3,lc=0.005,uc=0.15)
```

By generating an independent set of optimal portfolio weights for each segment, the algorithm produces three distinct portfolios consistent with the strategic priorities of each preference structure. To provide a representative illustration of the optimization results, *Table 4.7, 4.8 and 4.9* present the investment alternatives that received the largest allocation shares within the Return, Risk and ESG driven investor profiles, for an immediate visual comparison aiming to highlight the differences in strategic emphasis across investor segments.

These partial representations do not reflect the full capital allocation, therefore the presented shares do not sum to unity**.** For the complete overview of the three final portfolios, refer to *Table A.5, A.6 and A.7* in the Appendix.

Table 4.7 – Top Three ETFs for Return-Oriented Cluster

ETF Name	Capital Allocation
USXF ETF	0.0569
NUVEEN GRO	0.0482
ISHARES ADV	0.0463

Table 4.8 – Top Three ETFs for Risk-Oriented Cluster

ETF Name	Capital Allocation
USXF ETF	0.0571
CXSE CHINA	0.0528
LDEM EM	0.0449

Table 4.9 – Top Three ETFs for ESG-Oriented Cluster

ETF Name	Capital Allocation
USXF ETF	0.0624
RBND ESG	0.0579
WBIG ESG	0.0550

The resulting portfolios, whose allocation patterns reflect the prioritization structure, exhibit balanced diversification and robust compliance with the constraint limits. By effectively distributing allocations across a greater number of assets, the upper-bound constraint promotes diversification, reduces the incidence of zero-weighted holdings and prevents the excessive concentration in individual alternatives. The optimization process produces diverse allocation strategies for each segment, confirming the ability to account for the variations in preferences.

The recurrent prioritization of a high-performance multi-dimensional asset across the three cluster-specific portfolios is independent of localized dominance within a single evaluative dimension or any inherent bias within the optimization framework. This phenomenon is fundamentally attributable to the structural mechanism of the AUS formulation, which, in the presence of balanced or moderately differentiated preference configurations, identifies assets that optimize the collective objective function.

Such instruments demonstrate consistently superior normalized performance across the dimensions, functioning as a convergent point of stability within the decisional space. In contexts characterized by relatively balanced or limited polarization, the additive and compensatory nature of the utility aggregation promotes the selection of alternatives that achieve high overall coherence over those offering peak performance in a single isolated criterion. Accordingly, assets displaying consistent cross-criterion stability tend to emerge naturally as optimal selections even when segment-specific priorities differ.

This outcome aligns the mathematical output with institutional investment frameworks, where diversified instruments exhibiting multi-dimensional stability are strategically deployed within integrated capital allocation landscapes. In this setting, the model operationalizes the preference for assets that maintain structural integrity across moderate or mixed objective environments, mirroring the selection criteria employed in professional multi-objective portfolio management. The resulting allocations therefore reflect the internal logic of the preference-driven aggregation mechanism, whereby the model prioritizes assets that maintain alignment with complex multi-objective structures to ensure the behavioral consistency of the final allocations.

4.2.6 Three-Dimensional Utility Space Representation

The results of the optimization procedure are spatially formalized through a tridimensional mapping of the investment universe, establishing a rigorous geometric representation of the core architecture of the model. This spatialization facilitates an immediate interpretation of the alignment between the aggregated utilities and the resulting optimized portfolio configurations.

Each investment alternative is projected within a three-dimensional Euclidean space defined by normalized return, risk and ESG axes, where the coordinates of each asset correspond to its relative performance metrics. Consequently, this spatial distribution defines a continuous and bounded geometric domain, providing a structural lens through which the convergence of investor preferences and capital allocation can be analyzed.

Code 4.7 – Visualization of the Tri-Criterion Utility Space

```
cmin=min(etf_data[,c("AUS_C1", "AUS_C2", "AUS_C3")])
cmax=max(etf_data[,c("AUS_C1", "AUS_C2", "AUS_C3")])
plot_cluster=function(df, portfolios_row ,aus_col ,title_text){
 vals=df[[aus_col]]
 plot_ly(data=df,x=~Return,y=~Risk,z=~ESG,type="scatter3d",mode="markers",
 marker=list(size=5,opacity=0.85, color=vals, cmin=cmin, cmax=cmax,
 colorscale=list(c(0,"#d3d3d3"),c(1,"#1f77b4")),
 colorbar=list(title=list(text="AUS"))) %>%
 add_trace(data=portfolios_row,x=~Return,y=~Risk,z=~ESG,
 type="scatter3d",mode="markers+text",
 marker=list(size=10, color="red",symbol="diamond"),
 text=~Cluster, textposition="top center",inherit=FALSE) %>%
 layout(title=title_text,scene=(list(xaxis=list(title="Normalized Return"),
 yaxis=list(title="Normalized Risk"), zaxis=list(title="Normalized ESG")))}
```

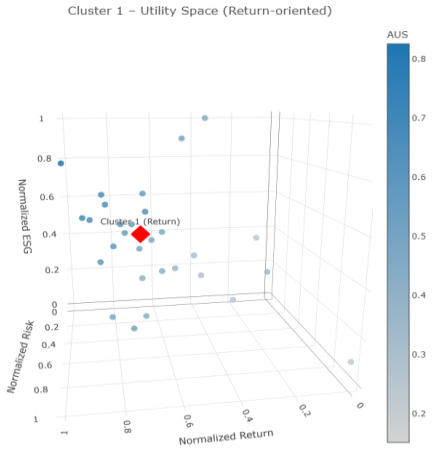
The interactive three-dimensional representation is executed through the `plot_ly()` function, which projects each investment alternative into the normalized decision space.

The color parameter applies a continuous tonal gradient based on the AUS associated with the selected investor cluster, effectively encoding the intensity of the perceived utility within the visual domain. The optimized portfolio barycenter is integrated through the `add_trace()` function and appears as a distinct red diamond marker. This point is positioned according to the weighted mean of the constituent assets across the three dimensions, representing the mathematical centroids of the capital distribution. Moreover, the `layout()` function organizes the visualization space to ensure consistent axis labeling and portfolio scaling across the geometric cube to maintain a uniform framework for comparative analysis.

```
Code 4.8 – Graphical Output for Cluster-specific Utility Space

plot_cluster(etf_data, portfolios[1,,drop=FALSE],"AUS_C1","Cluster 1:")

plot_cluster(etf_data, portfolios[2,,drop=FALSE],"AUS_C2","Cluster 2:")

plot_cluster(etf_data, portfolios[3,,drop=FALSE],"AUS_C3","Cluster 3:")
```

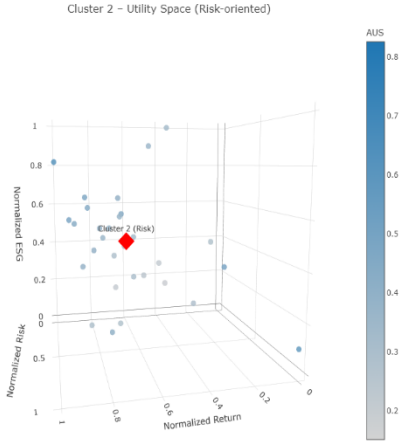
The iterative execution of the function generates three independent visualizations corresponding to the identified investor segments, providing a sequential basis for a comparative assessment of the spatial displacement across the portfolios. This approach demonstrates how distinct preference weights reconfigure the underlying utility landscape to drive differentiated capital allocations within a unified evaluative framework. By observing the shift in the portfolio barycenter relative to the asset universe, these visualizations illustrate the sensitivity of the optimization model to the specific strategic profiles of each investor segment.

Figure 4.3 – Return-Oriented 3D Visualization



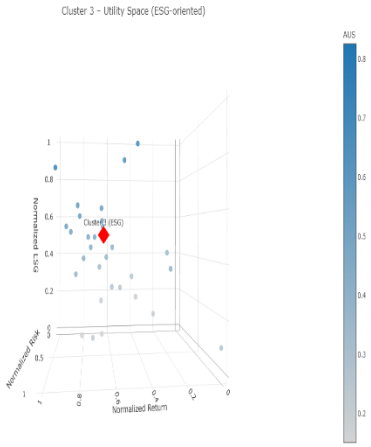
The pronounced alignment of high-utility markers along the Return axis provides the empirical confirmation that financial performance represents the primary leading criterion for this segment. This localized spatial density directly informs the optimization trajectory, ensuring that the resulting portfolio barycenter is positioned precisely within the high-utility zone. Consequently, the proximity of the optimal solution to the most relevant asset clusters demonstrates a rigorous consistency between the underlying preference structure of the segment and the final quantitative output. While the point cloud exhibits moderate dispersion, the overall geometric configuration maintains a systematic directional bias toward the Return dimension. The observed distribution validates the capacity of the model for precise asset prioritization, as it effectively captures the strategic emphasis of the investor profile without compromising the structural integrity of the broader investment universe.

Figure 4.4 – Risk-Oriented 3D Visualization



The color gradient within the risk-oriented cluster reveals a significant alignment along the Risk axis, where the utility surface increases monotonically in direct proportion to normalized risk scores. This orientation suggests that the objective function is primarily sensitive to risk dimension, while the return and the ESG parameters retain only a marginal influence on the final utility distribution. Consequently, the positioning of the optimized portfolio barycenter within the high-risk quadrant of the geometric domain validates the dominance of risk-seeking preference structures. This spatial convergence confirms that the optimization framework effectively converts the strategic priorities of the segment into a definitive capital allocation.

Figure 4.5 – ESG-Oriented 3D Visualization



The vertical dominance of the utility gradient along the ESG axis, coupled with the positioning of the portfolio marker near the upper boundary of the coordinate space, underscores the methodological emphasis on sustainability over financial attributes. This configuration reveals a localized preference for investment vehicles characterized by consistently high ESG scores, as evidenced by the visual concentration within the utility cloud.

The barycentric coordinates of the optimized portfolio offer a high-density numerical representation of the segment's position within the tri-dimensional criterion space. Unlike individual assets, whose spatial coordinates are fixed by normalized performance metrics, the portfolio's location represents a convex combination of the eligible investment alternatives obtained via weighted aggregation using the cluster-specific allocation vector derived from the optimization stage.

The coordinates (0.676, 0.218, 0.446) place the return-oriented portfolio within a region dominated by high-return exposure. In this configuration, the secondary yet non-negligible influence of risk and the ESG performance provides empirical evidence for the direct transformation of the preference system into a precise spatial outcome.

Similarly, the risk-oriented portfolio, which demonstrates the highest risk coordinate among the three configurations at (0.628, 0.263, 0.434), reflects a systematic shift toward assets with elevated normalized volatility in accordance with the specific emphasis of the cluster on the risk dimension.

The ESG-oriented portfolio, located at coordinates (0.642, 0.219, 0.503), presents a pronounced elevation along the vertical axis relative to the other clusters. The moderate return and risk values illustrate the compensatory function of financial attributes when sustainability acts as the primary decision criterion, highlighting the structural weight assigned to the ESG dimension while maintaining essential diversification across the alternative axes.

The adoption of uniform spatial scaling and consistent color encoding across the three visualizations establishes a basis for mathematical comparability between segments. This standardized representation enables a rigorous interpretation of the geometric relationships among the criteria by ensuring that spatial shifts are attributed to preference variations rather than to coordinate inconsistencies. Accordingly, the visualizations offer a definitive mapping of how heterogeneous investor preferences, quantified through aggregated utility functions, converge into the final composition and spatial positioning of the portfolios within the tri-criterion decision space.

5. Results Interpretation, Sensitivity Analysis and Future Developments

The concluding chapter of the research details the optimization results through a set of complementary analytical perspectives, providing a structured analytical and descriptive foundation for the final reflection on the methodological contributions and empirical validity of the proposed framework.

The initial section summarizes the structural properties of the resulting portfolio configurations, focusing on the empirical allocation structures and their compliance with the imposed constraints to define the fundamental quantitative parameters required for the subsequent interpretative analysis. Subsequently, the second stage relates the observed allocation patterns to established investment benchmarks and strategic rationales, providing an economically grounded reading of the portfolio configurations.

The implementation of a sensitivity analysis examines the stability of the proposed framework under an alternative preference structure to assess how variations in input priorities manifest within the optimization process and influence the resulting allocation structures without altering the underlying model architecture. Conclusively, the closing section delineates potential trajectories for future developments by identifying structural extensions and methodological refinements, including the expansion of the evaluative criteria, the broadening and diversification of the investment universe and the integration of temporal dynamics within the established research environment.

5.1 Empirical Results and Structural Properties of the Optimized Portfolios

This section reports the findings of the optimization process and outlines the significant structural properties of the portfolio compositions, focusing on the resulting capital allocations generated under the operational constraints and on the specific investment alternatives prioritized within each cluster-specific configuration.

The return-oriented portfolio structure emerges from the explicit prioritization of profitability, as evidenced by the substantial allocations to instruments such as *USXF ETF (USXF)*, *NUVEEN GRO (NULG)* and *ISHARES ADV (DMXF)*. These investment alternatives capture the primary objective of the segment by concentrating capital toward high-growth potential targets, while preserving sufficient allocative breadth to comply with the predefined diversification constraints.

A markedly different allocation pattern is observed within the risk-oriented portfolio, in which the market exposures are articulated through high-volatility instruments, specifically *CXSE CHINA (CXSE)* and *LDEM EM (LDEM)*, aligning the investment universe with the specific inclinations of the investor segment.

The ESG-oriented portfolio exhibits the most pronounced thematic consistency, assigning dominant distributions to funds characterized by superior sustainability performance, such as *RBND ESG (RBND)* and *WBIG ESG (WBIG)*. The prioritization of these alternatives, even in instances where their return performance remains moderate relative to competing instruments, confirms the seamless integration of extra-financial considerations into tangible portfolio allocations.

Extending the descriptive overview of the optimization results, the analysis of the composition of the portfolio structures reveals the consistent centrality of a specific subset of generalist investment alternatives. Instruments such as *USXF*, *DMXF* and *ISHARES KLD (DSI)* maintain high-ranking positions across all the configurations, exhibiting a notable resilience to shifts in the optimization objective. Analytically, this outcome transcends single-objective dominance to denote a state of structural robustness across the integrated decision space.

These alternatives, defined by high normalized performance across the entire evaluative set, demonstrate an intrinsic allocative invariance, functioning as stochastic anchors of the model that persist across heterogeneous preference configurations. Consequently, these assets act as a stabilizing mechanism, ensuring that concurrently with the reallocation toward specific-cluster priorities, the foundational architecture of the portfolios remains rooted in vehicles that provide multi-dimensional utility.

Complementing the foundational stability of generalist assets, the specific priorities of each cluster determine the final concentration of specialized assets, ensuring a precise alignment between allocations and strategic objectives. In particular, the 5.27% weighting of *CXSE* within the risk-oriented configuration exemplifies this allocation principle, confirming the utility of the asset for investor who explicitly favor elevated volatility profiles over moderate return and sustainability performance. Similarly, the superior sustainability profile of *RBND* accounts for its significant allocation within the ESG-oriented portfolio, whereas the moderate return and limited volatility restrict the instrument to a marginal presence in the alternative structures.

This divergence validates the effectiveness of the segmentation in aligning allocations with the dominant objectives of the clusters. Whereas a unidimensional architecture would exclude these instruments as non-optimal outliers, the proposed framework mechanically maps specialized asset value to the specific evaluative structure, rendering them instrumental for fulfilling the distinct requirements of each archetype.

Beyond the descriptive comparison between the preference structures and the portfolio configurations, the Sum of Squared Errors (SSE) and the numerical stability across iterations validate the cohesion of the identified clusters and the reproducibility of the results. From a geometric perspective, the comparative analysis of the three-cluster utility environments indicates that portfolio barycenters occupy distinct regions of the normalized decision space, consistent with the relative emphasis assigned to each criterion.

Conclusively, the proposed framework preserves internal consistency across the three configurations, where the stability of broadly diversified generalist combined with the sensitivity of specialized investment alternatives ensures that each allocation remains consistent with the underlying objective, providing the empirical basis for the subsequent interpretative analysis.

5.2 Economic and Behavioral Interpretation of Portfolio Allocation Patterns

The analytical scope of this section provides an interpretative reading of the portfolio configurations by relating their structural composition to established investment practices. In this context, the analysis remain bounded to the operational structure of the framework and examines how heterogeneous preferences, upon being mapped onto constrained allocation vectors, generate differentiated portfolios that mirror recognizable real-world asset allocation strategies.

The distribution patterns within the three distinct portfolio structures demonstrate that a concentrated range of investment alternatives attracts significant capital allocations, irrespective of the prevailing optimization objectives. From a market-based perspective, this stability aligns with broader economic strategies in which portfolios are structured upon a stable core of broadly diversified instruments acting as common denominators to build the primary exposure, while differentiation is achieved through tactical modulations and auxiliary concentrations that reorient the configurations toward specific thematic priorities.

Accordingly, the persistent emergence of significant allocations across heterogeneous architectures, specifically exemplified by *USXF* and *DMXF*, suggests that these assets exhibit an invariant structural relevance, maintaining a foundational indispensability that function independently of specific optimization objectives.

The interpretative utility of the leading allocations that define the constitutive base of each configuration is integrated with the insights derived from persistent marginal capital weights, which effectively delineate the structural boundaries of the decision landscape. In this context, the recurrent marginal allocations to instruments such as *ISHARES EM (ESGE)* across all portfolios provide a strategic perspective, as these instruments offer exposure to asset classes that would otherwise remain unrepresented.

These consistent low weightings reflect a deliberate trade-off between primary optimization goals and the requirement for market access, revealing that such vehicles function as complementary diversifiers for secondary exposures, although the persistent residual positioning indicates a limited marginal contribution toward the realization of targeted priorities.

The integrated and differentiated nature of the allocation framework is revealed in the configuration of the return-oriented portfolio structure. In common investment practices, the prioritization of the return dimension is often associated with the methodological and operational assumption that superior expected returns derive from a predominant exposure to a specific investment style, typically concretized through the selection of either growth-focused or value-focused strategies. Such interpretations, although prevalent within institutional allocation frameworks, tend to simplify the relationship between return objectives and portfolio composition by assuming that maximizing profitability requires a narrow specialization within a restricted typology of investment.

Such a premise is challenged by the observation that, despite the potential for extended outperformance within specific investment styles, cyclical efficacy remains subject to market conditions, which precludes the constant maximization of portfolio returns.

Within the proposed framework, the return-oriented portfolio does not converge into the concentration of a single investment style, representing instead an allocation structure that combines exposures to assets characterized by robust growth expectations with instruments exhibiting broader diversification properties. Specifically, the capital distribution maintains significant allocations across multiple sources of return and asset classes, as evidenced, for instance, by the weighting of *NULG*, which ensures exposure to high-growth equities without assuming a dominant role in the configuration. This principle is reinforced by the complementary inclusion of assets with distinct sectoral and thematic exposures, such as *DMXF* and *DSI*, which track broad-market and core-equity dimensions.

Conclusively, from an interpretative perspective, this allocation confirms that return prioritization under optimization constraints does not necessarily imply stylistic exclusivity, as the return-oriented portfolio emerges through a weight distribution that prioritizes performance-contributing assets while maintaining structural diversification.

The central insight derived from the composition of the risk-oriented portfolio stems from the alignment of the capital distribution with established investment strategies, whereby institutional entities, such as multi-strategy portfolios and opportunistic global investors, maintain a deliberate exposure to market segments characterized by pronounced performance fluctuations and structural uncertainty. This positioning materializes through capital exposures toward emerging markets, cyclically sensitive sectors and high-volatility equity segments, in which investors embrace greater variability in exchange for exposure to a broader range of potential outcomes, including scenarios associated with asymmetric payoffs and repricing opportunities.

Within the structure of the risk-oriented portfolio, the allocation to *CXSE* aligns with exposures to market segments subject to elevated regulatory and policy uncertainty, where return distributions exhibit higher dispersion relative to more stable asset classes. Similarly, the allocations to *LDEM* and *SPDR EM (SPYX)* represent consistent exposures to dimensions of increased sensitivity to macroeconomic conditions of developing economies, in which factors like balance of payments crises, sudden capital flow reversals and the disintegration of correlation dynamics during systemic stress conditions align the portfolio with the specific dynamics of these market segments.

Therefore, the observed allocation patterns reflect a configuration in which higher-volatility exposures assume a more prominent role within the portfolio structure, rather than functioning as a cost parameter to be minimized, resulting in greater sensitivity to market movements and broader performance outcome variance across macroeconomic landscapes.

The dimension of sustainability within the ESG-oriented portfolio operates, from a purely interpretative perspective, dualistically as a theme, constituted as an ethical factor, and as an evaluative dimension by performing, in a partial and indirect manner, risk management functions analogous to traditional oversight. More specifically, as demonstrated within capital markets, the ESG factor acts as a layer of soft risk governance, deriving from the indirect inclusion of information regarding operational robustness, governance quality and vulnerability to reputational shocks, which are often not fully captured by standard risk metrics. This arrangement does not presuppose that the sustainability dimension ensures lower risk, nor does it imply a deliberate mechanism for risk mitigation. On the contrary, it enables the alignment of the portfolio with modern investment landscapes where the ESG factor, as the dominant variable, directs allocation to instruments whose methodologies reflect a specific sensitivity toward non-financial risks. In this context, ESG serves to redefine the frontier of attractive alternatives through a qualitative reorganization of allocative priorities.

Transcending the contingent and debatable narrative that attributes an automatic stabilizing capacity to these criteria, the construction of ESG-oriented portfolio might be framed as the adoption of an alternative perspective in which risk factors capable of undermining business value and continuity without immediately translating into volatility are intercepted, while only to a limited extent, by qualitative and structural filters focused on legal and governance aspects.

Within the ESG-oriented portfolio, the allocations toward *RBND* and *WBIG* reflect a strategy prioritizing alternatives designed to exclude or minimize exposure to entities operating in controversial industries, specifically tobacco, alcohol and fossil fuels, or those involving violations of United Nations principles, while simultaneously providing diversified coverage oriented to enterprises characterized by enhanced informational transparency and superior capacity for adaptation to evolving landscapes.

A second relevant observation concerns the systematic under-allocations for alternatives such as *FNCL ESG (FNCL)*, *FNDB ESG (FNDB)* and *IVE VALUE (IVE),* whose weights result significantly lower than those observed in the other portfolios. Such reductions reflect the result of a selective reallocation induced by the priority assigned to sustainability, where the ESG dimension functions as a qualitative discriminator against assets incongruent with a more stringent and integrated conception of non-financial trade-offs. Specifically, the *FNCL*, which is intensively concentrated in financial services, banking and intermediaries, maintains a structurally cyclical profile that, despite the inclusion of an ESG filter, is highly dependent on macroeconomic conditions, financial leverage and systemic risk. Therefore, this exposure diverges with the objective of a portfolio in which sustainability assumes a dominant role, attributable to the heightened sensitivity of the sector to regulatory shifts and systemic shocks.

Similarly, the *FNDB* and *IVE*, which target fundamental factors and traditional models predominantly comprised of mature, capital-intensive legacy firms, are relatively marginalized due to their reliance on financial metrics and remain less selective regarding institutional governance factors. In such context, such assets appear less effective in filtering qualitative vulnerabilities that contemporary ESG investors consider essential within a broader strategy of soft risk management.

Aligned with the institutional strategies of pension funds, insurance companies and sovereign-like allocators, this framework exploits ESG dimensions as a supplementary layer of due diligence or an analytical diagnostic validation on risk, specifically designed to mitigate exposure to sanctions, non-compliance and governance failures. Within this perspective, the logic of ESG as a partial substitute remains distinct from explicit volatility constraints yet serves to fulfill institutional requirements for reputational resilience and stakeholder accountability within increasingly stringent regulatory landscapes.

Ultimately, the economic and behavioral interpretation of the configurations demonstrates how the proposed framework maps heterogeneous preference structures into outcomes consistent with institutional practices without imposing predefined assumptions on strategy or asset semantics. These results emerge endogenously from the interaction between normalized performance metrics and optimization constraints, bypassing the need for embedded economic knowledge. The resulting portfolios thus represent structurally coherent architectures reflecting the internal logic of the framework while remaining strictly descriptive, providing an economically grounded reading of the observed patterns whose robustness is further examined through sensitivity analysis.

5.3 Sensitivity Analysis: Alternative Preference Weight Configurations

5.3.1 Structuring the Problem and Data Collection

The structural robustness of an analytical framework constitutes a central metric of its functional utility and methodological validity within non-stationary financial markets characterized by systemic stochastic volatility and parametric instability. In this context, the effectiveness of decision-support systems moves beyond the derivation of a singular optimal solution, requiring instead a capacity for interpretative stability across divergent preference architectures.

The sensitivity analysis evaluates the resilience of the model by subjecting the optimization framework to configurational shifts in the input data, utilizing the more pronounced preference concentrations of the alternative scenario as a rigorous stress test to validate its capacity for generating consistent outputs under polarized conditions (Saltelli et al., 2008). This approach accounts for the inherent evolutionary dynamics of investor preference structures, which exhibit continuous reorientation in response to market trajectories, shifting regulatory requirements or realignments of strategic financial objectives.

Across both principal and alternative experimental settings, investor preferences are operationalized as compositional vectors defined over the three evaluative dimensions of the framework, mapping realizations onto a standard simplex embedded in $\mathbb{R}^3$ to ensure strict positivity and structural satisfaction of the unit-sum constraint. The divergence between the two scenarios refers not to the dimensionality of the preference space, the distributional domain or the stochastic nature of the generation process, but to the specific parametrization of the concentration regime.

The framework preserves a unified modeling architecture, ensuring internal coherence of the preference data while retaining full comparability across the settings to comprehensively explore the model behavior under systematically differing preference regimes.

In the principal scenario, investor preferences are formalized through a symmetric Dirichlet distribution defined by moderate concentration parameters Dir(2,2,2), producing preference vectors characterized by mixed and relatively balanced weights across the dimensions. This parametrization yields realizations principally located centrally within the simplex to generate heterogeneous but equilibrated preference profiles, precluding the dominance of any individual evaluative dimension within the decision space and guaranteeing that investor structures exhibit multifaceted priorities and explicit trade-offs consistent with a diversified environment where multiple objectives jointly influence the portfolio construction process.

In contrast, the alternative scenario reorients the distributional density toward a low-concentration specification by modifying the regime of the distribution, while maintaining the dimensionality and symmetry of the space to represent environments where investor priorities are sharply polarized. Conceptually, this setting is defined by Dirichlet parametrizations with concentration values in the interval $\alpha \in [0.1,0.4]$, which are mathematically calibrated to favor the emergence of polarized realizations concentrated near the vertices of the simplex. Under this specification, the empirical results reported in the analysis are obtained from the specification Dir(0.3,0.3,0.3), which reduces the aggregate concentration to reallocate the probabilistic mass toward more unbalanced vectors. To isolate a distinct dominance structure for the sensitivity analysis, the framework incorporates an intensification of the inherent density concentrations independent of exogenous inputs or predetermined preference archetypes. Reflecting the theoretical properties of the distribution, the resulting profiles demonstrate an increased degree of specialization to reflect distinctly defined and structurally polarized configurations.

This design exploits the properties of the Dirichlet distribution, in which the intensity of the concentration parameters governs the dispersion of the generated weights. By intentionally modifying the initial investor priorities within a completely data-driven process, the analysis seeks to isolate and quantify the model response, ensuring that observed variations in portfolio configurations are strictly attributable to the

preference regime, allowing the alternative scenario to function as a coherent stress test of structural robustness under polarized environments.

The methodological framework established in the preceding chapters underpins the current shift in focus exclusively toward the presentation of the empirical results of the analysis, with the specific R code for the alternative scenario documented in *Script B.2* of the Appendix.

5.3.2 Simulation of Investor Preferences through Dirichlet Distribution

The alternative preference environment is established through a deliberate realignment of the dispersion parameters within the preference-generation process, while maintaining strict adherence to the defined compositional structure. This modulation facilitates an evaluation of framework stability under conditions of structural polarization, isolating the response to divergent investor configurations without compromising the stochastic dimensionality of the architecture.

Table 5.1 – Sample of Five Investor Preferences in the Alternative Scenario

Investor	Return Preference	Risk Preference	ESG Preference
Investor 1	0.9688	0.0103	0.0209
Investor 2	0.0152	0.9586	0.0263
Investor 3	0.0466	0.0088	0.9447
Investor 4	0.9102	0.0183	0.0716
Investor 5	0.0424	0.9281	0.0296

Table 5.1 illustrates the preference weights for a representative sample of the simulated investors under the alternative scenario. The first observation reflects a primary concentration of 96.88% on the Return dimension, with marginal allocation of 1.03% and 2.09% to Risk and ESG respectively. In contrast, the second investor profile exhibits a divergent configuration by assigning 95.85% to the Risk criterion, while Return and ESG factors receive 1.51% and 2.63% of the preference weight. This comparison provides a valuable preliminary overview of the preference distributions within the alternative environment, while the comprehensive dataset for all simulated observations is documented in *Table A.8* in the Appendix.

These preference configurations provide the fundamental input vectors for the subsequent multi-criteria synthesis, establishing the necessary quantitative diversity required to evaluate the robustness of the allocation logic. The distinct weighting patterns documented in this analytical sequence ensure the derivation of a broad spectrum of investor priorities, consequently validating the structural flexibility of the optimization framework under non-uniform conditions.

5.3.3 Segmentation of Investor Preferences via K-Means Clustering

5.3.3.1 Investor Preferences Clustering

The intensification of preference polarization within the alternative environment requires a rigorous assessment of the clustering framework to determine its functional resilience. This analysis examines the ability of the grouping mechanism, implemented through the K-means algorithm utilizing Euclidean distance, to identify coherent and interpretable investor segments when subjected to pronounced dispersion settings. The investigation validates the robustness of the classification process against these preference vectors, confirming the inferential integrity of categorical boundaries despite the systemic increase in stochastic volatility and outlier incidence.

Table 5.2 – Sample of Five Investor Preferences in the Alternative Scenario with the Assigned Cluster

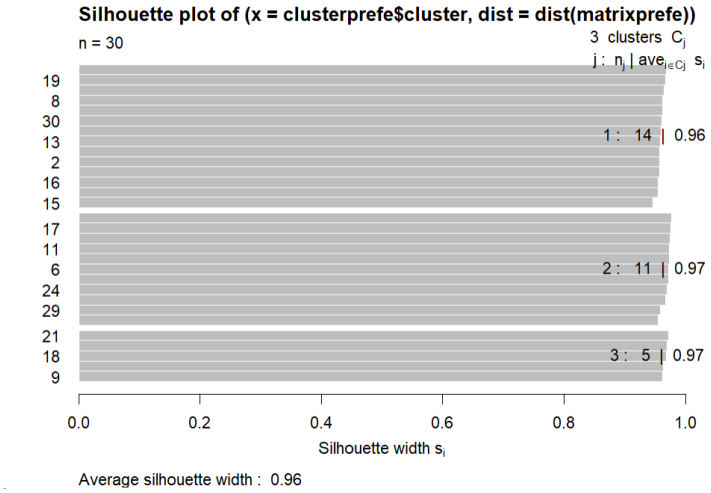
Investor	Return Preference	Risk Preference	ESG Preference	Cluster
Investor 1	0.9688	0.0103	0.0209	1
Investor 2	0.0152	0.9586	0.0263	2
Investor 3	0.0466	0.0088	0.9447	3
Investor 4	0.9102	0.0183	0.0716	1
Investor 5	0.0424	0.9281	0.0296	2

The unsupervised classification process partitioned the simulated profiles into a three-cluster architecture, establishing distinct groupings that reflect the concentrations of the alternative environment. This segmentation logic identifies specific densities within the distribution, with each group representing a stable behavioral archetype defined by a unique vector of criteria weights. The benchmarking of the segmentation against multiple iterations yielded systematically diminished performance, which confirms that the selected architecture optimally captures the stochastic distribution of the preference data. Notably, the alternative scenario preserves the fundamental thematic structure of the original clusters, maintaining a dominant inclination toward Return within the first cluster, Risk in the second and ESG in the third segment.

5.3.3.2 Internal Validation of Clustering Solution - Silhouette Score

To assess the structural integrity of the clustering process, the Silhouette score quantifies the internal cohesion and spatial separation of each observation, validating the alignment of individual priorities within their assigned segments.

Figure 5.1 – Silhouette Score in the Alternative Scenario



The observed Silhouette coefficient of 0.9647 establishes a significant degree of cluster cohesion and separation, providing robust statistical validation for the intra-cluster homogeneity and demonstrating that investors within each segment manifest highly convergent preference architectures. While the high Silhouette score is influenced by the significant polarization of preferences, this outcome is substantively informative, as it confirms that the clustering procedure maintains interpretability and stability under extreme parametric boundaries.

This result validates the structural robustness of the model, as the observed cluster distinctness identifies as an endogenous property of the underlying data distribution, appearing invariant to the specific preference configuration employed.

5.3.3.3 Centroid Extraction and Preference Parametrization

Following the empirical validation of the unsupervised identification of investor segments, the derivation of cluster centroids provides the essential quantitative bridge between identified investor archetypes and the multi-criteria aggregation stage, ensuring that each collective preference structure is precisely mapped for the final utility score evaluations.

Table 5.3 – Centroids of the Three Clusters in the Alternative Scenario

	Return Preference	Risk Preference	ESG Preference
Cluster 1	0.9306	0.0412	0.0282
Cluster 2	0.0359	0.9312	0.0329
Cluster 3	0.0368	0.0361	0.9271

The first segment reflects an absolute prioritization of capital growth, as the return dimension accounts for 93% of the preference structure, effectively indicating an orientation in which the attainment of high returns represents the central benchmark for this group.

The second segment displays a pronounced risk-centric orientation, with risk criterion dominating the preference vector and establishing an investment logic governed primarily by volatility-related considerations. This configuration subordinates both Return and ESG considerations to the overarching objective of exposure management, manifesting a strategic framework where the systematic and methodological integration of risk represents the foundational pillar for the final portfolio architecture.

The third segment is characterized by a significant commitment to extra-financial impact, with nearly 92% of the priority vector allocated to the ESG dimension, suggesting a profile that diverges from traditional financial trade-offs and positions sustainability as the determinant factor of asset selection.

These clustering outcomes confirm the robustness of the segmentation procedure under conditions of extreme dominance within individual preference vectors, while simultaneously revealing the capacity of the unsupervised learning stage to extract stable collective archetypes from structurally polarized preference distributions, grounding the subsequent portfolio optimization phase in mathematically sound and internally consistent investor profiles.

5.3.4 Aggregated Utility Scores Computation

The synthesis of the Aggregated Utility Scores (AUS) for each investment asset is operationalized through the multi-criteria aggregation mechanism, wherein the normalized performance metrics are weighted according to the previously derived centroid vectors. This procedure ensures a precise mapping of the investment universe onto the specific preference structure of each segment as it formalizes the transition from collective archetypes to individual asset evaluations. The application of these polarized preference structures illustrates a subsequent reconfiguration of asset valuations, determining, in the process, the extent to which modified decision logics influence the relative attractiveness of the investment alternatives.

Table 5.4 – Sample of Five ETFs with the relative Aggregated Utility Scores in the Alternative Scenario

ETF Name	Norm_Return	Norm_Risk	Norm_ESG	AUS_C1	AUS_C2	AUS_C3
ISHARES EA	0.5051	0.1321	0.2304	0.4820	0.1487	0.2370
XTRACKER	0.7121	0.1981	0.4922	0.6847	0.2262	0.4896
SPDR INT	0.4040	0.0849	0.2827	0.3874	0.1029	0.2800
CXSE CHINA	0.0	1.0	0.0995	0.0440	0.9345	0.1283
VANGUARD S	0.3737	0.1132	0.1780	0.3574	0.1247	0.1829

These numerical outputs reflect the degree of alignment between each alternative and the underlying strategic objectives of the cluster. *Table 5.4* shows a selection of these results for the purpose of immediate performance benchmarking, although the full distribution of scores across all segments remains available for reference within *Table A.9* in the Appendix.

The emergent utility profiles indicate that the aggregation mechanism preserves structural consistency while it maintains a high degree of sensitivity to fluctuations in preference intensity, with the observed shifts in rankings across different clusters reflecting empirical reconfigurations of the evaluative priorities.

5.3.5 Quadratic Programming for Portfolio Optimization and Capital Distribution

The optimization of capital distribution for each investor segment, required for the construction of a distinct portfolio for each cluster, relies on the identification of an allocation vector that minimizes the squared Euclidean distance relative to the ideal utility benchmark. By determining the specific percentage exposure for each financial instrument, this procedure aligns the portfolio composition with the theoretical maximum utility of the cluster and formalizes the strategic preferences into concrete investment allocations. Additionally, this analysis evaluates the sensitivity of the constrained optimization process to changes in the underlying utility structure, confirming the feasibility, stability and interpretability of the Quadratic Programming (QP) stage under conditions of extreme preference configurations.

Table 5.5 – Top Three ETFs for Return-Oriented Cluster in the Alternative Scenario

ETF Name	Capital Allocation
USXF ETF	0.0547
NUVEEN GRO	0.0503
SUSL LEAD	0.0485

Table 5.6 – Top Three ETFs for Risk-Oriented Cluster in the Alternative Scenario

ETF Name	Capital Allocation
CXSE CHINA	0.1248
LDEM EM	0.0824
FNCL ESG	0.0552

Table 5.7 – Top Three ETFs for ESG-Oriented Cluster in the Alternative Scenario

ETF Name	Capital Allocation
RBND ESG	0.0767
WBIG ESG	0.0697
USXF ETF	0.0659

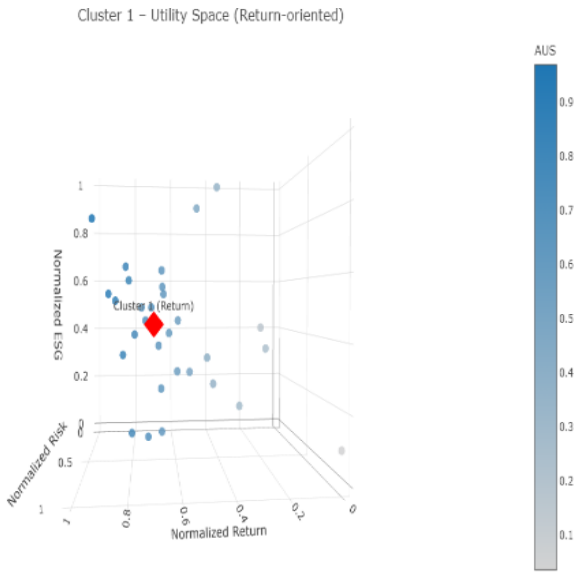
To illustrate the practical outcomes of the optimization process, *Tables 5.5*, *5.6* and *5.7* delineate the specific assets with the highest concentration of capital within the return, risk and ESG segments. This focused selection facilitates a comparative analysis of the diverging strategic priorities that characterize each investor profile. As these tables represent only the primary elements of each strategy, the reported weights do not constitute the entire portfolio and therefore do not aggregate to unity. The comprehensive documentation of the three final portfolios, including the full distribution of capital across all the investment alternatives, is provided in *Tables A.10*, *A.11* and *A.12* within the Appendix.

The optimization outputs confirm the constraint architecture's effectiveness in governing allocation behavior across polarized environments, functioning as a stabilizing force to prevent preference extremity amplification while ensuring portfolio compositions remain diversified and compliant with feasibility bounds.

5.3.6 Three-Dimensional Utility Space Representation

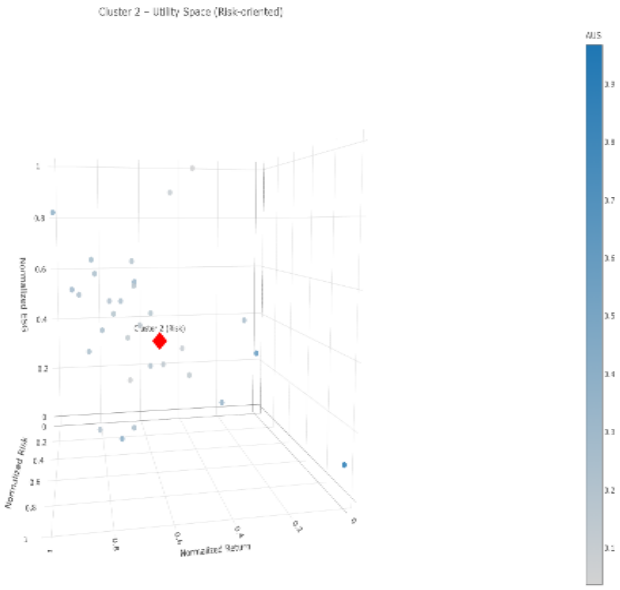
The outcomes of the optimization process are projected through a graphical transposition into a three-dimensional coordinate system, providing a structural visualization of the core decision framework, allowing for a clear assessment of the spatial relationship between the utility scores and the portfolio weights. Therefore, the geometric distribution of the investment alternatives serves to validate the internal consistency of the model by illustrating how specific strategic orientations determine the spatial arrangement of assets within the optimization space.

Figure 5.2 – Return-Oriented 3D Visualization in the Alternative Scenario



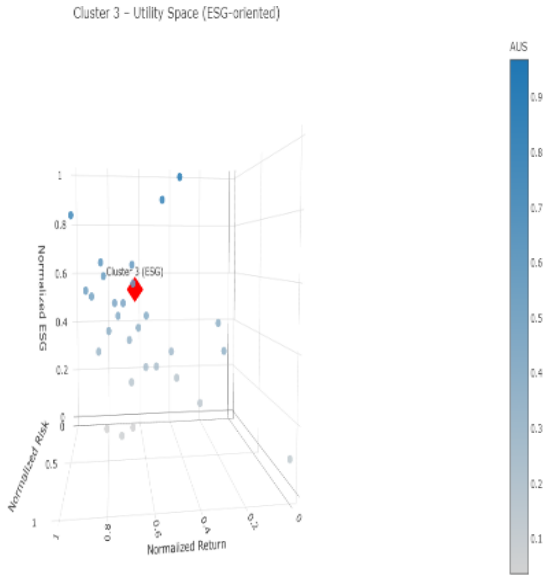
In the return-oriented segment, the primary configuration produces a broad distribution across the return dimension, associated with a moderate dispersion along the secondary axes. Upon the application of the alternative preference structure, this distribution experiences a structural compression and moves closer to the frontier of maximum normalized return. Such geometric progression demonstrates how the intensified concentration of weights restricts the selection of assets recognized as viable by the model, producing a more selective and performance-intensive allocation structure.

Figure 5.3 – Risk-Oriented 3D Visualization in the Alternative Scenario



The risk-oriented segment exhibits the most significant reorientation, as the alternative configuration reveals a pronounced inclination toward the high-volatility region with the distribution centered stabilizing at elevated normalized risk values. While the primary scenario displayed a marginal oblique alignment between return and volatility, the subsequent shift corresponds precisely with the modified preference structure. This transformation demonstrates how the model systematically converts risk prioritization into a vertically extended and volatility-dominant arrangement of assets.

Figure 5.4 – ESG-Oriented 3D Visualization in the Alternative Scenario



In the ESG-oriented segment, the principal configuration exhibited a broad vertical distribution with moderate dispersion, whereas the alternative configuration compresses this range into a concentrated ESG-centric alignment where the portfolio center attains a higher coordinate along the sustainability axis. The visual convergence of the utility distribution provides evidence of the more rigorous exclusion of assets with suboptimal sustainability performance, which in turn establishes a demanding selection environment where capital is strictly allocated to the highest-rated investment alternatives.

The cross-scenario comparison reveals that portfolio coordinates shift systematically in proportion to modifications applied to the core investor preferences. These spatial displacements maintain the geometry of the tri-criterion space, as a deterministic transformation driven exclusively by the reconfigured preference weights.

Specifically, the return-oriented cluster moves from (0.676, 0.218, 0.446) to (0.699, 0.206, 0.426), illustrating a coherent progression along the return axis to capture the influence of the intensified financial priorities. In contrast, the risk-oriented coordinates shift from (0.628, 0.263, 0.434) to (0.556, 0.357, 0.364), representing a clear convergence toward high-volatility regions. Ultimately, the ESG-oriented cluster transitions from (0.642, 0.219, 0.503) to (0.642, 0.200, 0.545), where the elevation along the sustainability axis demonstrates a non-stochastic evolution of the portfolio position.

Across all the segments, the spatial movements remain directionally consistent with preference modifications, as the utility space retains its structural integrity while the distribution centers shift in response to the modified weights. This consistency confirms the internal robustness of the model, which yields diversified and coherent allocations from a common architectural foundation. Finally, the optimization outputs exhibit technical compliance with the feasibility requirements of the framework, while demonstrating a precise responsiveness to heterogeneous investor profiles, through the consistent application of reconfigured preference structures.

5.3.7 Cross-Scenario Comparative and Interpretative Analysis

The capital distribution of the return-oriented portfolio in the principal configuration exhibits a relatively balanced composition across the investment universe, where *USXF* retains the maximum relative weight (5.69%) within a progressively descending series of capital assignments, exemplified by the allocations to *NULG* (4.82%) and *DMXF* (4.63%). In the alternative scenario, while the *USXF* allocation remains the principal structural component (5.47%), the weighting of low-priority alternatives, notably *FGDL GOLD (FGDL)* and *CXSE*, converges toward the lower boundary of the feasibility constraint. This divergence underscores a systematic penalization of misaligned assets driven by the intensified prioritization for capital growth, which yields a more pronounced capital concentration within a restricted subset of superior performers.

The second cluster portfolio, defined by a proportionally homogeneous allocation structure in the principal scenario where *USXF* (5.71%) constitutes the primary allocation, demonstrates a significantly intensified concentration within the alternative scenario, evidencing a fundamental sensitivity to polarized preference regime. This configuration, assigning the highest weight to *CXSE* (12.48%), reflects a deliberate orientation toward investment alternatives perceived as highly volatile or possessing significant specific risk, inducing allocations in assets such as *WBIG* and *RBND* to retract toward the lower bound of the allocation range due to the limited congruence with the governing portfolio objectives.

The third cluster portfolio features a distinct and pronounced prioritization toward assets with an explicit focus on ESG considerations, evidenced by the principal scenario allocations to *RBND* (5.79%) and *WBIG* (5.50%). This preference for intrinsic sustainability exposures is further intensified within the alternative configuration, suggesting that alternatives with robust sustainability characteristics are favored over those demonstrating a moderately balanced performance profile across the evaluative criteria. This structural realignment leads to the marginalization of low-priority alternatives such as *FNCL*, *IVE* and *FNDB*, resulting in their allocations being significantly reduced toward the lower allocative threshold.

The comparative analysis between the scenarios reveals salient variations in the final portfolio configurations and the optimization-derived allocation structures, which establish a distinct interpretative domain for the evaluation of the portfolio compositions.

The primary insight derived from the sensitivity analysis concerns the fundamental divergence in investment strategies, which shifts the focus from an equilibrium-oriented approach toward an environment that emphasizes capital concentration, where consistent allocations stem from relatively balanced preferences. Conversely, the strategy observed within the alternative scenario exhibits an alignment of capital allocations with the specific strategic objectives of the three segments, where a substantial aggregation of capital across a limited portion of investment alternatives is associated with the systematic marginalization of the residual asset universe.

Under this framework, the principal scenario prioritizes integrative instruments that operate as structural pillars through balanced performance across all criteria, whereas the alternative configuration displaces this versatility with a functional mapping of specialized assets. This transition leverages the alternatives with high single-criterion efficiency to ensure that each allocation contributes optimally to the singular objective of the respective segment and calibrates each alternative in relation to the specific ability to maximize the marginal utility of the target criterion.

The comparative analysis delineates a specialized management approach toward thematic investment alternatives, where the principal configuration utilizes funds with a specific thematic focus, such as *BKWO WOM (BKWO)*, as equilibrium-driven components to facilitate structural diversification across all portfolios. This inclination undergoes a substantial transformation within the alternative scenario, as the distributional density shifts toward an allocation tail with a less dispersed profile. Unlike the granularity observed in the principal configuration, these marginal assignments adhere to the lower feasibility bounds, confirming the compliance of the model with Quadratic Programming (QP) constraints and the efficiency benefit attributed to the primary investment selections.

The results derived from the sensitivity analysis facilitate the identification of specific strategic drivers for each investor segment, as the shift toward specialized assets reveals the distinct logic governing each portfolio composition. These observations constitute strictly retrospective interpretations, where the focus remains exclusively on the delineation of structural consistencies and allocation patterns that emerge from the constructed decision environment.

The interpretative insight derived from the return-oriented portfolio reveals the persistence of allocations toward large-cap stock index replicators, where the presence of instruments such as *SPDR S&P500 (EFIV)* and *VANGUARD S (VSGX)* reflects a selective role in response to a highly concentrated and performance-driven environment. Specifically, these assets provide exposure to highly liquid and operationally stable firms through the replication of large companies, subject to either stringent ESG filters or broad market representativeness. Within this polarized preference structure, such an inclusion facilitates an aggressive growth approach that leverages scalable return generation, aligning with the financial rationale that identifies large-cap equity exposures as a reliable source of economic stability and capital appreciation.

The analytical focus regarding the second cluster portfolio reveals a specific risk exposure profile, in which the intensified allocations toward instruments such as *CXSE*, *ESGE* or *LDEM* reflect a mechanical preference for elevated volatility and strategic exposure to assets with a structurally wide dispersion of potential outcomes. Under an asset-pricing lens, these instruments reside within environments where return distributions arise from fragmented and weakly stabilizing drivers, such as macroeconomic imbalances, asymmetric information flows and limited market depth. These conditions facilitate payoff structures characterized by compounded average volatility, which remain markedly disproportionate and context-specific as realizations often diverge substantially from central tendencies, a dynamic that, from an interpretative perspective, operationalizes the risk dimension as structural dispersion.

An absolute prioritization of the Risk dimension induces the optimization process to reallocate capital toward assets defined by a broad breadth of possibilities, diverging from a focus on instruments that primarily amplify fluctuations around a stable mean.

This orientation prioritizes wide and discontinuous future scenarios over probabilistically concentrated outcomes, ensuring a deliberate integration of structurally dispersed payoffs as an alternative to a defensive or protective strategy.

The systematic reduction of concentrated thematic exposures reveals the intensified selectivity of the third cluster portfolio under polarized preference structures. This vertical displacement occurs through the marked subordination of specific asset classes, such as *CXSE*, driven by a structural reallocation toward investment vehicles with robust sustainability characteristics over those with moderately balanced performance profiles.

The association of this outcome with the general investment rationale and the specific exposure to emerging markets underscores a structural risk liability inherent to these geographic domains, which arise fundamentally from restricted corporate reporting, the extensive influence of public sector stakeholders and conflicting oversight requirements within governance standards. Within this configuration, the absolute priority assigned to ESG considerations, coupled with the restrictive determinants of the allocation protocol, imposes a substantial penalty on any instruments whose ethical consistency is compromised by endogenous uncertainties. Therefore, the inability of marginal ESG benefits to offset the high penalty scores inherent to these governance risks induces the mechanical downgrade of these assets, confirming their status as inefficient allocation drivers for the achievement of the main objective within the portfolio.

The comparative analysis validates the effectiveness of the model, as a common architectural foundation produces diversified and consistent allocations in direct response to shifting preference structures. Concurrently, the logical sensitivity across the portfolios maintains technical compliance with heterogeneous investor profiles, establishing a robust basis for future methodological refinements.

5.4 Structural Extensions and Methodological Refinements

The established framework, while internally consistent and computationally robust, relies on a set of assumptions that ensure methodological transparency but simultaneously delineate the explanatory scope of the model. This equilibrium between clarity and scope identifies a definitive trajectory for future developments, where specific methodological advancements will expand the interpretative domain of the architecture along complementary dimensions.

This developmental trajectory follows a three-dimensional progression across orthogonal axes to enhance the analytical depth, geographic breadth and temporal orientation of the model. The first dimension encompasses a structural expansion of the evaluative criteria to internalize supplementary financial parameters that reinforce the foundational pillars of the methodology. The second component addresses the extension of the investable universe toward broader market exposures and factor orientations, while the final transitions the methodology from a static and cross-sectional format toward a dynamic configuration that facilitates period recalibration through historical data.

The methodology integrates three evaluative dimensions into a linear multi-criteria aggregation framework to achieve solution efficiency and preserve clarity in portfolio outcomes. This configuration facilitates the transparent construction of preference weights while it simultaneously reduces the cognitive complexity of weight elicitation and enables the direct derivation of normalized priority vectors for the computation of Aggregated Utility Scores.

The first pillar of methodological expansion involves the enrichment of the evaluative criteria, a structural refinement that integrates supplementary financial parameters to capture market realities and ensure the pragmatic viability of the portfolio configurations. This alignment bridges the theoretical model with the specific requirements of the current investment landscape, where each additional parameter enhances the marginal utility of the selection process.

The liquidity dimension reflects the capacity to execute rapid transactions without the induction of material price distortions, providing the critical boundary for portfolio implementability, as insufficient trading efficiency in the underlying securities might compromise the feasibility of theoretically optimal allocations. This dimension is approximated through several market-based indicators, most notably the Average Daily Trading Volume (ADTV), which proxies secondary-market depth and exhibits a positive association with liquidity, with higher values indicating greater absorptive capacity. In parallel, the turnover ratio relates traded volume to outstanding shares, where elevated turnover denotes enhance liquidity, together transforming market absorption into quantifiable determinants of executable optimized allocations.

Within the prospective analytical framework, the cost-efficiency dimension captures the ability of an investment alternative to provide benchmark exposure while limiting return erosion arising from both explicit and implicit costs, encompassing management fees and the accuracy of index replication. In this context, the Total Expense Ratio (TER), defined as the annual percentage of asset under management, is inversely normalized such that lower cost structures correspond to higher evaluative scores. Collectively, these indicators operationalize efficiency as a measurable constraint governing the feasibility of portfolio allocations under real-world market execution conditions.

The dividend stability parameter, instead, describes the degree of regularity and predictability within the income stream of a financial instrument. This dimension complements the dual evaluation of Risk and Return through the incorporation of an income-focused metric, particularly critical for Multi Criteria Decision Analysis (MCDA).

The distributive flow operates as a utility factor to determine annualized distributions, whereas the framework simultaneously incorporates payout variance as a cost proxy to account for large fluctuations that might undermine the structural stability of the income stream. Furthermore, the frequency of dividend suspensions or reductions functions as a supplementary risk proxy, where a lower incidence of distributive contractions determines a superior normalized score. Collectively, these metrics establish dividend stability as an evaluative criterion linking portfolio composition to both capital appreciation and cash flow consistency.

The six-criterion structure provides a more granular approximation of the complex decision environment through an enhanced framework that facilitates a multi-dimensional assessment, where liquidity determines implementation feasibility, cost efficiency internalizes the frictions that affect net outcomes and the dividend stability captures the income regularity unaccounted for within the price-based risk and return metrics.

Beyond the integration of supplementary evaluative criteria, the second component of the methodological refinement encompasses the expansion of the investment universe, which, in the proposed framework, is deliberately composed of a restricted range of investment alternatives to maintain analytical interpretability and ensure computational tractability.

The geographic scope expansion denotes the systematic inclusion of instruments providing exposure to markets currently underrepresented within the baseline scenario. Although the existing framework integrates assets linked to United States, Europe and specific emerging markets, the incorporation of sustainability-aligned alternatives providing exposure to the Asian-Pacific, Latin American, and African developed regions functions to enhance geographic diversification and capture heterogeneous growth trajectories.

Similarly, the sectoral and thematic expansion facilitates the extension of the financial instrument universe across specific industrial and topical domains. While the present configuration relies on broad-market sustainability funds, a logic refinement involves incorporating industry-specific alternatives, such as technology, healthcare or utilities, characterized by distinct financial and ESG profiles. Furthermore, thematic products targeting renewable energy or social inclusion provide a deeper layer of granularity by addressing structural trends that transcend traditional sector classifications. These instruments establish the infrastructure necessary to determine whether the optimization process systematically allocates capital toward technological innovation, economic transitions or defined sustainability objectives.

Additionally, the integration of style exposures captures the divergence in investment rationale and factor orientation within the financial environment, where the incorporation of sustainability-aligned strategies focused on size, volatility or quality expands the decision space to include instruments with risk-return dynamics that deviate systematically from market benchmarks. This expansion reveals latent preferences for specific investment styles, ensuring that the resulting outcomes achieve greater representativeness of realistic investable sets while enhancing the external validity of the capital distribution analysis.

Complementing the evaluative criteria and investment universe expansion, the third domain of internal developments involves the dynamic extension of the framework via the application of historical time series to recalibrate instrument scores across rolling windows. Although the model operates cross-sectionally by design to ensure methodological transparency and facilitate the validation of the decision construction model, the integration of temporal data facilitates a transition toward longitudinal robustness.

Evaluating return and risk dimensions through historical series extends the analysis beyond unidimensional point-in-time assessments, as cumulative average growth rates computed across short, medium and long windows reveal short-term momentum and longer-term persistence. The volatility and downside-focused metrics, such as semi-deviation or maximum drawdown, provide an intertemporal contextualization of the risk dimension. This integration acts as a temporal filter to separate sustained performance trends from temporary market volatility, ensuring that the optimization output reflects the underlying resilience of the instrument profiles across fluctuating environments. Similarly, a systematic examination of the longitudinal evolution of ESG scores across multiple windows identifies whether an instrument exhibits temporal stability or progressive improvement in the sustainability profile. The structural incorporation of these trajectories into a dynamic framework supports the categorization of assets characterized by consistent ESG orientations from those with fluctuating metrics, which enhances the ex-post calibration of sustainability as a primary decision criterion.

Despite the significant expansion of the methodological architecture across three orthogonal axes, specifically the incrementation of analytical depth through additional criteria, the extension of universe scope via an increased investment universe and the integration of temporal dynamics through historical series, the internal refinements strictly adhere to the established methodological boundaries and preserve the foundational logic sequence of the multi-criteria decision analysis model.

6. Conclusion

The proposed research developed a structured and data-driven decision-making framework combining Multiple-Criteria Decision Analysis (MCDA), Dirichlet-based preference simulation, K-means clustering, Quadratic Programming (QP) and visual utility representations for deriving optimized portfolio configurations from multidimensional preference structures. The analytical results demonstrated that differentiated preference structures generated distinct allocation patterns, while preserving internal consistency under the operational constraints, confirming the capacity to align with the strategic prioritization across return, risk and ESG criteria.

From a financial perspective, the empirical evidence illustrates how the explicit integration of ESG factors enables sustainability to operate as a structural evaluative dimension within the portfolio construction process. This analytical approach ensures that return, risk and ESG contribute to capital allocation in accordance with investor preference structures, supporting a scalable tri-criterion framework capable of incorporating heterogeneous evaluative dimensions within a consistent optimization model.

Methodologically, the framework introduces a formalization of investor preferences through explicitly defined vectors operating as direct inputs for the optimization process, establishing a transparent relationship with the final portfolio allocations that differs from traditional approaches in which preference assumptions remain implicitly incorporated within the objective functions. Within this context, the integration of an additional evaluative dimension highlights the flexibility of the architecture, suggesting the potential for future strategic inclusion of extra-financial variables to examine how these drivers influence the dynamics of portfolio optimization processes.

Additionally, these findings revealed that portfolio optimization operates as a preference-driven process, where allocation outcomes emerge as a function of the underlying evaluative structures as opposed to universally defined solutions, resulting in differentiated portfolio configurations across distinct preference profiles.

Nevertheless, the framework operates within a static cross-sectional setting, without addressing the temporal evolution of financial and extra-financial dimensions, and relies on a targeted investment universe and specific evaluative criteria to prioritize interpretability over the representation of global market dynamics. Although these structural assumptions preserve internal consistency, the resulting framework boundaries delineate a trajectory toward future dynamic evaluation, expanded decision spaces and methodological refinements.

To conclude, the research aims to contribute to the academic context focusing on the integration of portfolio optimization with multi-criteria decision-making frameworks, providing a replicable, transparent and scalable environment for the reflection of heterogeneous utility structures within optimized portfolio compositions.

Bibliography

Abate, G., Basile, I., & Ferrari, P. (2024). The integration of environmental, social and governance criteria in portfolio optimization: An empirical analysis. *Corporate Social Responsibility and Environmental Management*, *31*(3), 2054–2065. https://doi.org/10.1002/csr.2682

Abbas, A. E. (2018). *Foundations of multiattribute utility* (1st Ed.). Cambridge University Press. https://doi.org/10.1017/9781316596739

Aitchison, J. (1982). The Statistical Analysis of Compositional Data. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *44*(2), 139–160. https://doi.org/10.1111/j.2517-6161.1982.tb01195.x

Amel-Zadeh, A., & Serafeim, G. (2017). *Why and How Investors Use ESG Information: Evidence from a Global Survey*. Said Business School, University of Oxford, Harvard Business School. https://ssrn.com/abstract=2925310

Bachelier, L. (1900). Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure*, *17*, 21–86. https://doi.org/10.24033/asens.476

Ballestero, E., Bravo, M., Pérez-Gladish, B., Arenas-Parra, M., & Plà-Santamaria, D. (2012). Socially Responsible Investment: A multicriteria approach to portfolio selection combining ethical and financial objectives. *European Journal of Operational Research*, *216*(2), 487–494. https://doi.org/10.1016/j.ejor.2011.07.011

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3–18. https://doi.org/10.1016/0304-405X(81)90018-0

Belton, V., & Stewart, T. J. (2002). *Multiple Criteria Decision Analysis*. Springer US. https://doi.org/10.1007/978-1-4615-1495-4

Bernoulli, D. (1954). Exposition of a New Theory on the Measurement of Risk. *Econometrica*, *22*(1), 23–36. https://doi.org/10.2307/1909831 (Original Work Published 1738)

Bertelli, B., & Torricelli, C. (2024). The trade-off between ESG screening and portfolio diversification in the short and in the long run. *Journal of Economics and Finance*, *48*(2), 298–322. https://doi.org/10.1007/s12197-023-09652-9

Black, F., & Litterman, R. (1992). Global Portfolio Optimization. *Financial Analysts Journal*, *48*(5), 28–43. https://doi.org/10.2469/faj.v48.n5.28

Boyd, S. P., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press. https://web.stanford.edu/~boyd/cvxbook/

Cesarone, F., Martino, M. L., Ricca, F., & Scozzari, A. (2024). Managing ESG ratings disagreement in sustainable portfolio selection. *Computers & Operations Research*, *170*, 106676. https://doi.org/10.1016/j.cor.2024.106766

Cesarone, F., Martino, M. L., & Carleo, A. (2022). Does ESG Impact Really Enhance Portfolio Profitability? *Sustainability*, *14*(4), 2050. https://doi.org/10.3390/su14042050

Chen, L., Zhang, L., Huang, J., Xiao, H., & Zhou, Z. (2021). Social responsibility portfolio optimization incorporating ESG criteria. *Journal of Management Science and Engineering*, *6*(1), 75–85. https://doi.org/10.1016/j.jmse.2021.02.005

Connor, R. J., & Mosimann, J. E. (1969). Concepts of Independence for Proportions with a Generalization of the Dirichlet Distribution. *Journal of the American Statistical Association*, *64*(325), 194–206. https://doi.org/10.1080/01621459.1969.10500963

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster Analysis* (1st ed.). Wiley. https://doi.org/10.1002/9780470977811

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, *47*(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Ferguson, T. S. (1973). A Bayesian Analysis of Some Nonparametric Problems. *The Annals of Statistics*, *1*(2), 209–230. https://doi.org/10.1214/aos/1176342360

Fisher, I. (1930). *The Theory of Interest: As Determined by Impatience to Spend Income and Opportunity to Invest It*. Macmillan Company.

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, *5*(4), 210–233. https://doi.org/10.1080/20430795.2015.1118917

Gasser, T., Peters, G. P., Fuglestvedt, J. S., Collins, W. J., Shindell, D. T., & Ciais, P. (2017). Accounting for the climate–carbon feedback in emission metrics. *Earth System Dynamics*, *8*(2), 235–263.https://doi.org/10.5194/esd-8-235-2017

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-means clustering algorithm. *Applied Statistics*, *28*(1), 100–108. https://doi.org/10.2307/2346830

Hennig, C. (2015). What are the true clusters? *Pattern Recognition Letters*, *64*, 53–62. https://doi.org/10.1016/j.patrec.2015.04.009

Ishizaka, A., & Nemery, P. (2013). *Multi-Criteria Decision Analysis: Methods and Software* (1st Ed.). Wiley. https://doi.org/10.1002/9781118644898

Jagannathan, R., & Ma, T. (2003). Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *The Journal of Finance*, *58*(4), 1651–1683. https://doi.org/10.1111/1540-6261.00580

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys*, *31*(3), 264–323. https://doi.org/10.1145/331499.331504

Keeney, R. L., Raiffa, H., & Rajala, D. W. (1979). Decisions with Multiple Objectives: Preferences and Value Trade-Offs. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(7), 403–413.https://doi.org/10.1109/TSMC.1979.4310245

Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money*. Macmillan and Co., Ltd.

Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate Sustainability: First Evidence on Materiality. *The Accounting Review*, *91*(6), 1697–1724. https://doi.org/10.2308/accr-51383

Lintner, J. (1965a). Security Prices, Risk, and Maximal Gains from Diversification. *The Journal of Finance*, *20*(4), 587–615. https://doi.org/10.1111/j.1540-6261.1965.tb02930.x

Lintner, J. (1965b). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, *47*(1), 13–37. https://doi.org/10.2307/1924119

MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley Symposium on Mathematical Statistics and Probability*, *1*, 281–297. https://projecteuclid.org/euclid.bsmsp/1200512992

Madhavan, A., & Sobczyk, A. (2014). Price Dynamics and Liquidity of Exchange-Traded Funds. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2429509

Markowitz, H. M. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification of investments*. John Wiley & Sons, Inc.

Mencarelli, L., & D'Ambrosio, C. (2019). Complex portfolio selection via convex mixed-integer quadratic programming: A survey. *International Transactions in Operational Research*, *26*(2), 389–414. https://doi.org/10.1111/itor.12541

Michaud, R. O. (1989). The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analysts Journal*, *45*(1), 31–42. https://doi.org/10.2469/faj.v45.n1.31

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, *34*(4), 768–783. https://doi.org/10.2307/1910098

Renneboog, L., Ter Horst, J., & Zhang, C. (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance*, *14*(3), 302–322. https://doi.org/10.1016/j.jcorpfin.2008.03.009

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65.https://doi.org/10.1016/0377-0427(87)90125-7

Salo, A., Doumpos, M., Liesiö, J., & Zopounidis, C. (2024). Fifty years of portfolio optimization. *European Journal of Operational Research*, *318*(1), 1–18. https://doi.org/10.1016/j.ejor.2023.12.031

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Global Sensitivity Analysis: The Primer* . Wiley. https://doi.org/10.1002/9780470725184

Samuelson, P. A. (1937). A Note on Measurement of Utility. *The Review of Economic Studies*, *4*(2), 155–161.https://doi.org/10.2307/2967612

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *The Journal of Finance*, *19*(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Stern, N. (Ed.). (2006). *The Economics of Climate Change: The Stern Review*. HM Treasury. https://www.lse.ac.uk/granthaminstitute/publication/the-economics-of-climate-change-the-stern-review/

Steuer, R. E., & Utz, S. (2023). Non-contour efficient fronts for identifying most preferred portfolios in sustainability investing. *European Journal of Operational Research*, *306*(2), 742–753. https://doi.org/10.1016/j.ejor.2022.08.007

Tinoco, M. A. C., Dutra, C. C., Ribeiro, J. L. D., Miorando, R. F., & Caten, C. S. T. (2018). An integrated model for evaluation and optimisation of business project portfolios. *European Journal of Industrial Engineering*, *12*(3), 442–463. https://doi.org/10.1504/EJIE.2018.092010

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

Utz, S., Wimmer, M., Hirschberger, M., & Steuer, R. E. (2014). Tri-criterion inverse portfolio optimization with application to socially responsible mutual funds. *European Journal of Operational Research*, *234*(2), 491–498. https://doi.org/10.1016/j.ejor.2013.07.024

Varmaz, A., Fieberg, C., & Poddig, T. (2024). Portfolio optimization for sustainable investments. *Annals of Operations Research*, *341*(2–3), 1151–1176. https://doi.org/10.1007/s10479-024-06189-w

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.

Williams, J. B. (1938). *The Theory of Investment Value*. Harvard University Press.

Xidonas, P., & Essner, E. (2024). On ESG Portfolio Construction: A Multi-Objective Optimization Approach. *Computational Economics*, *63*(1), 21–45. https://doi.org/10.1007/s10614-022-10327-6

Appendix A – Data Sources and Descriptive Statistics

Table A.1 – Investment Universe Raw Data

ETF Name	Ticker	Launch Year	AUM (USD Millions)	ESG Score	E Score	S Score	G Score	1-Y Return	3-Year Return	Volatility Risk
WBIG ESG	WBIG	2014	29.3	9.38	9.45	9.32	9.37	8.2%	24.5%	12.4%
RBND ESG	RBND	2020	22.9	9.57	9.62	9.51	9.58	7.8%	22.1%	11.8%
USXF ETF	USXF	2020	57.6	9.22	9.31	9.15	9.20	12.4%	35.2%	15.6%
IVE VALUE	IVE	2000	47,970	7.68	7.72	7.65	7.67	9.8%	28.4%	14.2%
FNDB ESG	FNDB	2013	1,050	7.66	7.71	7.62	7.65	11.2%	31.5%	13.8%
FNCL ESG	FNCL	2013	2,420	7.72	7.75	7.70	7.71	10.5%	29.8%	16.2%
FGDL GOLD	FGDL	2022	464.8	8.45	8.92	8.12	8.31	6.6%	18.2%	14.5%
CXSE CHINA	CXSE	2012	504.4	7.85	7.92	7.81	7.82	-5.2%	15.4%	22.4%
RAYD DEV	RAYD	2021	92.8	8.12	8.15	8.10	8.11	9.1%	26.8%	13.9%
BKWO WOM	BKWO	2023	41.6	8.75	8.42	9.15	8.68	10.8%	28.4%	14.8%
UDI DIV	UDI	2022	6.3	7.95	7.88	7.98	7.99	11.5%	28.4%	12.2%
LDEM EM	LDEM	2020	32.1	8.35	8.42	8.31	8.32	4.2%	19.5%	18.6%
FEUS CLIM	FEUS	2021	119.05	8.88	9.15	8.72	8.77	9.8%	28.4%	13.5%
ISHARES OPT	SUSA	2005	3,760	8.25	8.32	8.21	8.22	11.8%	32.4%	14.1%
SUSL LEAD	SUSL	2019	1,030	8.65	8.72	8.61	8.62	12.1%	33.2%	13.8%
ISHARES AW	ESGU	2016	15,510	8.5	8.6	8.4	8.5	10.2%	30.1%	13.5%
VANGUARD E	ESGV	2018	11,850	8.3	8.4	8.2	8.3	9.8%	28.7%	12.8%
SPDR S&P500	EFIV	2020	1,140	8.4	8.5	8.3	8.4	11.1%	31.2%	14.1%
ISHARES EA	ESGD	2016	10,890	8.1	8.2	8.0	8.1	8.7%	25.4%	13.2%
XTRACKER	USSG	2019	501.86	8.6	8.7	8.5	8.6	10.5%	29.5%	13.9%
NUVEEN GRO	NULG	2016	1,950	8.7	8.8	8.6	8.7	12.3%	33.8%	14.5%
SPDR INT	DWFI	2016	80,160	8.2	8.3	8.1	8.2	7.9%	23.4%	12.7%
ISHARES ADV	DMXF	2020	648.8	8.9	9.0	8.8	8.9	11.8%	32.1%	14.2%
ISHARES EM	ESGE	2016	5,600	7.9	8.0	7.8	7.9	6.5%	20.8%	15.6%
VANGUARD S	VSGX	2018	5,630	8.0	8.1	7.9	8.0	7.5%	22.8%	13.0%
NUVEEN VAL	NULV	2016	1,920	8.4	8.5	8.3	8.4	9.7%	27.6%	13.1%
ISHARES KLD	DSI	2006	5,100	8.8	8.9	8.7	8.8	11.5%	31.8%	14.0%
SPDR EM	SPYX	2015	2,500	8.6	8.7	8.5	8.6	10.9%	30.5%	13.8%
ISHARES LC	CRBN	2014	990.4	8.5	8.6	8.4	8.5	8.9%	26.7%	13.6%
SPDR GEN. D.	SHE	2016	283.6	8.7	8.8	8.6	8.7	9.6%	28.2%	13.4%

Note. Data collected from financial platforms: Yahoo Finance, JustETF and ETFDB

Table A.2 – Investment Universe Normalized Data

ETF Name	Ticker	Launch Year	AUM (USD Millions)	Normalized Return	Normalized Risk	Normalized ESG
WBIG ESG	WBIG	2014	29.3	0.4595959596	0.0566037736	0.9005235602
RBND ESG	RBND	2020	22.9	0.3383838384	0.0	1.0
USXF ETF	USXF	2020	57.6	1.0	0.3584905660	0.8167539267
IVE VALUE	IVE	2000	47,970	0.6565656566	0.2264150943	0.0104712042
FNDB ESG	FNDB	2013	1,050	0.8131313131	0.1886792453	0.0
FNCL ESG	FNCL	2013	2,420	0.7272727273	0.4150943396	0.0314136126
FGDL GOLD	FGDL	2022	464.8	0.1414141414	0.2547169811	0.4136125655
CXSE CHINA	CXSE	2012	504.4	0.0	1.0	0.0994764398
RAYD DEV	RAYD	2021	92.8	0.5757575758	0.1981132076	0.2408376963
BKWO WOM	BKWO	2023	41.6	0.6565656566	0.2830188679	0.5706806283
UDI DIV	UDI	2022	6.3	0.6565656565	0.0377358491	0.1518324607
LDEM EM	LDEM	2020	32.1	0.2070707071	0.6415094340	0.3612565445
FEUS CLIM	FEUS	2021	119.05	0.6565656566	0.1603773585	0.6387434555
ISHARES OPT	SUSA	2005	3,760	0.8585858586	0.2169811321	0.3089005236
SUSL LEAD	SUSL	2019	1,030	0.8989898990	0.1886792453	0.5183246073
ISHARES AW	ESGU	2016	15,510	0.7424242424	0.1603773585	0.4397905759
VANGUARD E	ESGV	2018	11,850	0.6717171717	0.0943396226	0.3350785340
SPDR S&P500	EFIV	2020	1,140	0.7979797980	0.2169811321	0.3874345550
ISHARES EA	ESGD	2016	10,890	0.5050505051	0.1320754717	0.2303664922
XTRACKER	USSG	2019	501.86	0.7121212121	0.1981132076	0.4921465969
NUVEEN GRO	NULG	2016	1,950	0.9292929293	0.2547169811	0.5445026178
SPDR INT	DWFI	2016	80,160	0.4040404040	0.0849056604	0.2827225131
ISHARES ADV	DMXF	2020	648.8	0.8434343435	0.2264150943	0.6492146597
ISHARES EM	ESGE	2016	5,600	0.2727272727	0.3584905660	0.1256544503
VANGUARD S	VSGX	2018	5,630	0.3737373737	0.1132075472	0.1780104712
NUVEEN VAL	NULV	2016	1,920	0.6161616162	0.1226415094	0.3874345550
ISHARES KLD	DSI	2006	5,100	0.8282828283	0.2075471698	0.5968586387
SPDR EM	SPYX	2015	2,500	0.7626262626	0.1886792453	0.4921465969
ISHARES LC	CRBN	2014	990.4	0.5707070707	0.1698113208	0.4397905759
SPDR GEN. D.	SHE	2016	283.6	0.6464646465	0.1509433962	0.5445026178

Note. The Return, Risk and ESG scores are standardized using Min-Max scaling to ensure cross-dimensional comparability and to prevent scale-driven bias in the subsequent optimization stages. While the values reported in the table are rounded for expositional clarity, all empirical computations rely on a full-precision normalized vectors to preserve numerical accuracy and integrity throughout the analysis

Table A.3 – Investor Preferences Principal Scenario

Investor	Return Preference	Risk Preference	ESG Preference	Sum
Investor 1	0.134998	0.648432	0.216570	1
Investor 2	0.774832	0.025287	0.199881	1
Investor 3	0.266924	0.438515	0.294561	1
Investor 4	0.190742	0.500223	0.309035	1
Investor 5	0.087654	0.468155	0.444191	1
Investor 6	0.247856	0.730112	0.021132	1
Investor 7	0.597412	0.256470	0.146118	1
Investor 8	0.369847	0.219823	0.410330	1
Investor 9	0.551238	0.171499	0.277263	1
Investor 10	0.512115	0.389102	0.098783	1
Investor 11	0.353102	0.550219	0.096679	1
Investor 12	0.277654	0.425834	0.296512	1
Investor 13	0.080123	0.326002	0.594395	1
Investor 14	0.177332	0.353343	0.468325	1
Investor 15	0.644501	0.168395	0.187104	1
Investor 16	0.090842	0.721993	0.187165	1
Investor 17	0.684315	0.080452	0.234137	1
Investor 18	0.477189	0.340498	0.182313	1
Investor 19	0.458126	0.297122	0.244752	1
Investor 20	0.636327	0.108329	0.255344	1
Investor 21	0.564110	0.121310	0.314579	1
Investor 22	0.197045	0.132010	0.670945	1
Investor 23	0.216466	0.466112	0.317422	1
Investor 24	0.147532	0.132993	0.719475	1
Investor 25	0.254026	0.583415	0.162559	1
Investor 26	0.448305	0.268382	0.283313	1
Investor 27	0.537128	0.213112	0.249760	1
Investor 28	0.140128	0.087643	0.772229	1
Investor 29	0.265224	0.396112	0.338264	1
Investor 30	0.496258	0.155312	0.348430	1

Note. Preference weights are drawn from the completely data-driven Dirichlet distribution process and capture the relative importance assigned by each investor to Return, Risk and ESG within the principal scenario. By construction, this process imposes a unit-sum constraint, such that preference weights sum to one for each investor.

Table A.4 - Investment Universe Aggregated Utility Scores in Principal Scenario

ETF Name	Ticker	AUS C1	AUS C2	AUS C3
WBIG ESG	WBIG	0.486655	0.349949	0.660854
RBND ESG	RBND	0.432554	0.317432	0.695370
USXF ETF	USXF	0.826083	0.609476	0.749394
IVE VALUE	IVE	0.411685	0.267018	0.150977
FNDB ESG	FNDB	0.488190	0.278075	0.159653
FNCL ESG	FNCL	0.493905	0.389176	0.213906
FGDL GOLD	FGDL	0.230919	0.268904	0.340393
CXSE CHINA	CXSE	0.225488	0.563357	0.270609
RAYD DEV	RAYD	0.417715	0.290439	0.281715
BKWO WOM	BKWO	0.560411	0.434239	0.524065
UDI DIV	UDI	0.408381	0.199777	0.203183
LDEM EM	LDEM	0.332260	0.478905	0.396248
FEUS CLIM	FEUS	0.552429	0.384743	0.542628
ISHARES OPT	SUSA	0.594808	0.378583	0.371531
SUSL LEAD	SUSL	0.662827	0.423190	0.506778
ISHARES AW	ESGU	0.551175	0.354813	0.427047
VANGUARD E	ESGV	0.473064	0.278295	0.335362
SPDR S&P500	EFIV	0.580498	0.384596	0.413178
ISHARES EA	ESGD	0.362744	0.236988	0.250887
XTRACKER	USSG	0.554800	0.381327	0.464071
NUVEEN GRO	NULG	0.699301	0.471735	0.541802
SPDR INT	DWFI	0.310092	0.202380	0.259877
ISHARES ADV	DMXF	0.671713	0.463399	0.590755
ISHARES EM	ESGE	0.253914	0.283079	0.195576
VANGUARD S	VSGX	0.273335	0.185513	0.193675
NUVEEN VAL	NULV	0.460845	0.294277	0.366705
ISHARES KLD	DSI	0.646729	0.437172	0.550871
SPDR EM	SPYX	0.580874	0.387213	0.469628
ISHARES LC	CRBN	0.457989	0.322641	0.403499
SPDR GEN. D.	SHE	0.521831	0.354443	0.478395

Note. The computational variables are processed employing full numerical precision to preserve the stochastic integrity of the optimization process, while the values reported within the table are rounded to ensure formal conciseness.

Table A.5 – Optimal Asset Allocation for the Return-Oriented Cluster in Principal Scenario

ETF Name	Ticker	Capital Allocation
USXF ETF	USXF	0.056880
NUVEEN GRO	NULG	0.048151
ISHARES ADV	DMXF	0.046251
SUSL LEAD	SUSL	0.045639
ISHARES KLD	DSI	0.044531
ISHARES OPT	SUSA	0.040956
SPDR EM	SPYX	0.039996
SPDR S&P500	EFIV	0.039970
BKWO WOM	BKWO	0.038587
XTRACKER	USSG	0.038201
FEUS CLIM	FEUS	0.038038
ISHARES AW	ESGU	0.037951
SPDR GEN. D.	SHE	0.035931
FNCL ESG	FNCL	0.034008
FNDB ESG	FNDB	0.033614
WBIG ESG	WBIG	0.033509
VANGUARD E	ESGV	0.032573
NUVEEN VAL	NULV	0.031732
ISHARES LC	CRBN	0.031535
RBND ESG	RBND	0.029784
RAYD DEV	RAYD	0.028762
IVE VALUE	IVE	0.028347
UDI DIV	UDI	0.028119
ISHARES EA	ESGD	0.024977
LDEM EM	LDEM	0.022878
SPDR INT	DWFI	0.021351
VANGUARD S	VSGX	0.018821
ISHARES EM	ESGE	0.017483
FGDL GOLD	FGDL	0.015900
CXSE CHINA	CXSE	0.015526

Note. The portfolio weights are derived through the Quadratic Programming (QP) stage subject to a budget identity constraint to ensure full capital allocation, combined with predefined lower and upper bound constraints to regulate the individual asset exposure.

Table A.6 – Optimal Asset Allocation for the Risk-Oriented Cluster in Principal Scenario

ETF Name	Ticker	Capital Allocation
USXF ETF	USXF	0.057114
CXSE CHINA	CXSE	0.052793
LDEM EM	LDEM	0.044879
NUVEEN GRO	NULG	0.044207
ISHARES ADV	DMXF	0.043425
ISHARES KLD	DSI	0.040968
BKWO WOM	BKWO	0.040693
SUSL LEAD	SUSL	0.039657
FNCL ESG	FNCL	0.036470
SPDR EM	SPYX	0.036286
FEUS CLIM	FEUS	0.036055
SPDR S&P500	EFIV	0.036041
XTRACKER	USSG	0.035734
ISHARES OPT	SUSA	0.035477
ISHARES AW	ESGU	0.033250
SPDR GEN. D.	SHE	0.033215
WBIG ESG	WBIG	0.032794
ISHARES LC	CRBN	0.030235
RBND ESG	RBND	0.029747
NUVEEN VAL	NULV	0.027577
RAYD DEV	RAYD	0.027217
ISHARES EM	ESGE	0.026528
VANGUARD E	ESGV	0.026079
FNDB ESG	FNDB	0.026059
FGDL GOLD	FGDL	0.025199
IVE VALUE	IVE	0.025022
ISHARES EA	ESGD	0.022208
SPDR INT	DWFI	0.018965
UDI DIV	UDI	0.018721
VANGUARD S	VSGX	0.017385

Table A.7 – Optimal Asset Allocation for the ESG-Oriented Cluster in Principal Scenario

ETF Name	Ticker	Capital Allocation
USXF ETF	USXF	0.062405
RBND ESG	RBND	0.057906
WBIG ESG	WBIG	0.055032
ISHARES ADV	DMXF	0.049194
ISHARES KLD	DSI	0.045873
FEUS CLIM	FEUS	0.045186
NUVEEN GRO	NULG	0.045118
BKWO WOM	BKWO	0.043641
SUSL LEAD	SUSL	0.042201
SPDR GEN. D.	SHE	0.039838
SPDR EM	SPYX	0.039107
XTRACKER	USSG	0.038645
ISHARES AW	ESGU	0.035562
SPDR S&P500	EFIV	0.034407
ISHARES LC	CRBN	0.033601
LDEM EM	LDEM	0.032997
ISHARES OPT	SUSA	0.030939
NUVEEN VAL	NULV	0.030537
FGDL GOLD	FGDL	0.028346
VANGUARD E	ESGV	0.027927
RAYD DEV	RAYD	0.023459
CXSE CHINA	CXSE	0.022535
SPDR INT	DWFI	0.021641
ISHARES EA	ESGD	0.020892
FNCL ESG	FNCL	0.017813
UDI DIV	UDI	0.016920
ISHARES EM	ESGE	0.016286
VANGUARD S	VSGX	0.016128
FNDB ESG	FNDB	0.013295
IVE VALUE	IVE	0.012572

Table A.8 – Investor Preferences Alternative Scenario

Investor	Return Preference	Risk Preference	ESG Preference	Sum
Investor 1	0.945373	0.048364	0.006263	1
Investor 2	0.046560	0.008784	0.944656	1
Investor 3	0.959952	0.015014	0.025034	1
Investor 4	0.915293	0.046012	0.038695	1
Investor 5	0.015148	0.958549	0.026303	1
Investor 6	0.907918	0.056785	0.035297	1
Investor 7	0.968806	0.010269	0.020924	1
Investor 8	0.040736	0.052392	0.906872	1
Investor 9	0.030468	0.900731	0.068801	1
Investor 10	0.032671	0.951800	0.015529	1
Investor 11	0.902145	0.073009	0.024846	1
Investor 12	0.910000	0.078573	0.011427	1
Investor 13	0.032713	0.914312	0.052975	1
Investor 14	0.020978	0.944230	0.034792	1
Investor 15	0.906229	0.068335	0.025436	1
Investor 16	0.042357	0.928095	0.029548	1
Investor 17	0.027865	0.057397	0.914738	1
Investor 18	0.048615	0.918514	0.032871	1
Investor 19	0.960374	0.008184	0.031442	1
Investor 20	0.027445	0.929923	0.042632	1
Investor 21	0.043236	0.944658	0.012106	1
Investor 22	0.916039	0.068571	0.015390	1
Investor 23	0.927324	0.051023	0.021653	1
Investor 24	0.048862	0.034906	0.916232	1
Investor 25	0.966712	0.019080	0.014208	1
Investor 26	0.932107	0.015936	0.051957	1
Investor 27	0.019981	0.027061	0.952958	1
Investor 28	0.034488	0.947092	0.018420	1
Investor 29	0.910148	0.018307	0.071545	1
Investor 30	0.066592	0.905314	0.028094	1

Note. Preference weights are drawn from the completely data-driven Dirichlet distribution process and capture the relative importance assigned by each investor to Return, Risk and ESG within the alternative scenario. By construction, this process imposes a unit-sum constraint, such that preference weights sum to one for each investor.

Table A.9 - Investment Universe Normalized Data and Aggregated Utility Scores in Alternative Scenario

ETF Name	Ticker	AUS C1	AUS C2	AUS C3
WBIG ESG	WBIG	0.455389	0.098838	0.853803
RBND ESG	RBND	0.343067	0.045058	0.939545
USXF ETF	USXF	0.968383	0.396604	0.806994
IVE VALUE	IVE	0.620667	0.234730	0.042073
FNDB ESG	FNDB	0.764455	0.204894	0.036736
FNCL ESG	FNCL	0.694832	0.413673	0.070864
FGDL GOLD	FGDL	0.153736	0.255865	0.397845
CXSE CHINA	CXSE	0.044048	0.934477	0.128354
RAYD DEV	RAYD	0.550790	0.213059	0.251587
BKWO WOM	BKWO	0.638772	0.305876	0.563473
UDI DIV	UDI	0.616861	0.063664	0.166257
LDEM EM	LDEM	0.229359	0.616690	0.365743
FEUS CLIM	FEUS	0.635629	0.193949	0.622089
ISHARES OPT	SUSA	0.816661	0.243048	0.325811
SUSL LEAD	SUSL	0.858985	0.225037	0.520409
ISHARES AW	ESGU	0.709876	0.190480	0.440847
VANGUARD E	ESGV	0.63840	0.122945	0.338792
SPDR S&P500	EFIV	0.762476	0.243457	0.396358
ISHARES EA	ESGD	0.481982	0.148720	0.236960
XTRACKER	USSG	0.684706	0.226221	0.489580
NUVEEN GRO	NULG	0.890642	0.288446	0.548197
SPDR INT	DWFI	0.387423	0.102861	0.280023
ISHARES ADV	DMXF	0.812484	0.262457	0.641080
ISHARES EM	ESGE	0.272101	0.347759	0.139516
VANGUARD S	VSGX	0.357446	0.124680	0.182862
NUVEEN VAL	NULV	0.589399	0.149028	0.386259
ISHARES KLD	DSI	0.796180	0.242593	0.591355
SPDR EM	SPYX	0.731313	0.219280	0.491100
ISHARES LC	CRBN	0.550479	0.193073	0.434868
SPDR GEN. D.	SHE	0.623186	0.181639	0.534042

Note. The computational variables are processed employing full numerical precision to preserve the stochastic integrity of the optimization process, while the values reported within the table are rounded to ensure formal conciseness.

Table A.10 – Optimal Asset Allocation for the Return-Oriented Cluster in Alternative Scenario

ETF Name	Ticker	Capital Allocation
USXF ETF	USFX	0.054687
NUVEEN GRO	NULG	0.050290
SUSL LEAD	SUSL	0.048499
ISHARES OPT	SUSA	0.046105
ISHARES ADV	DMXF	0.045869
ISHARES KLD	DSI	0.044947
FNDB ESG	FNDB	0.043153
SPDR S&P500	EFIV	0.043041
SPDR EM	SPYX	0.041278
ISHARES AW	ESGU	0.040065
FNCL ESG	FNCL	0.039215
XTRACKER	USSG	0.038642
BKWO WOM	BKWO	0.036044
VANGUARD E	ESGV	0.036023
FEUS CLIM	FEUS	0.035866
SPDR GEN. D.	SHE	0.035162
IVE VALUE	IVE	0.035020
UDI DIV	UDI	0.034804
NUVEEN VAL	NULV	0.033251
RAYD DEV	RAYD	0.031067
ISHARES LC	CRBN	0.031050
ISHARES EA	ESGD	0.027175
WBIG ESG	WBIG	0.025671
SPDR INT	DWFI	0.021827
VANGUARD S	VSGX	0.020131
RBND ESG	RBND	0.019318
ISHARES EM	ESGE	0.015304
LDEM EM	LDEM	0.012886
FGDL GOLD	FGDL	0.008609
CXSE CHINA	CXSE	0.005000

Table A.11 – Optimal Asset Allocation for the Risk-Oriented Cluster in Alternative Scenario

ETF Name	Ticker	Capital Allocation
CXSE CHINA	CXSE	0.124778
LDEM EM	LDEM	0.082345
FNCL ESG	FNCL	0.055237
USXF ETF	USXF	0.052958
ISHARES EM	ESGE	0.046435
BKWO WOM	BKWO	0.040843
NUVEEN GRO	NULG	0.038515
ISHARES ADV	DMXF	0.035045
FGDL GOLD	FGDL	0.034165
SPDR S&P500	EFIV	0.032508
ISHARES OPT	SUSA	0.032453
ISHARES KLD	DSI	0.032393
IVE VALUE	IVE	0.031343
XTRACKER	USSG	0.030207
SUSL LEAD	SUSL	0.030049
SPDR EM	SPYX	0.029280
RAYD DEV	RAYD	0.028449
FNDB ESG	FNDB	0.027359
FEUS CLIM	FEUS	0.025897
ISHARES LC	CRBN	0.025780
ISHARES AW	ESGU	0.025434
SPDR GEN. D.	SHE	0.024254
NUVEEN VAL	NULV	0.019899
ISHARES EA	ESGD	0.019858
VANGUARD S	VSGX	0.016648
VANGUARD E	ESGV	0.016417
SPDR INT	DWFI	0.013735
WBIG ESG	WBIG	0.013198
UDI DIV	UDI	0.008501
RBND ESG	RBND	0.006017

Table A.12 – Optimal Asset Allocation for the ESG-Oriented Cluster in Alternative Scenario

ETF Name	Ticker	Capital Allocation
RBND ESG	RBND	0.076737
WBIG ESG	WBIG	0.069723
USXF ETF	USXF	0.065893
ISHARES ADV	DMXF	0.052320
FEUS CLIM	FEUS	0.050766
ISHARES KLD	DSI	0.048252
BKWO WOM	BKWO	0.045971
NUVEEN GRO	NULG	0.044721
SPDR GEN. D.	SHE	0.043563
SUSL LEAD	SUSL	0.042448
SPDR EM	SPYX	0.040050
XTRACKER	USSG	0.039926
ISHARES AW	ESGU	0.035939
ISHARES LC	CRBN	0.035450
FGDL GOLD	FGDL	0.032421
SPDR S&P500	EFIV	0.032299
NUVEEN VAL	NULV	0.031473
LDEM EM	LDEM	0.029795
VANGUARD E	ESGV	0.027590
ISHARES OPT	SUSA	0.026528
SPDR INT	DWFI	0.022782
RAYD DEV	RAYD	0.020455
ISHARES EA	ESGD	0.019259
VANGUARD S	VSGX	0.014833
UDI DIV	UDI	0.013475
ISHARES EM	ESGE	0.011287
CXSE CHINA	CXSE	0.010374
FNCL ESG	FNCL	0.005671
IVE VALUE	IVE	0.005000
FNDB ESG	FNDB	0.005000

Appendix B – RStudio Script for Portfolio Optimization

Script B.1 – Complete RStudio Code Principal Scenario

```
#Necessary Packages
install.packages("MCMCpack")
install.packages("cluster")
install.packages("factoextra")
install.packages("quadprog")
install.packages("dplyr")
install.packages("plotly")

#Necessary Libraries
library(MCMCpack)
library(cluster)
library(factoextra)
library(quadprog)
library(dplyr)
library(plotly)
```

```
#Uploading Investment Universe Normalized Performances

etf_data <- data.frame(

ETF = c("WBIG_ESG", "RBND_ESG", "USXF_ETF", "IVE_VALUE", "FNDB_ESG", "FNCL_ESG", "FGDL_GOLD",
"CXSE_CHINA","RAYD_DEV", "BKWO_WOM", "UDI_DIV", "LDEM_EM", "FEUS_CLIM", "ISHARES_OPT",
"SUSL_LEAD", "ISHARES_AW", "VANGUARD_E", "SPDR_S&P500", "ISHARES_EA", "ISHARES_EM",
"XTRACKER", "NUVEEN_GRO", "NUVEEN_V", "SPDR_INT", "ISHARES_ADV", "VANGUARD_S", "SPDR_EM",
"ISHARES_KLD", "SPDR_GEN_D", "ISHARES_LC"),

Return = c(0.4595959596, 0.3383838384, 1.0, 0.6565656566, 0.8131313131, 0.7272727273,
0.1414141414, 0.0, 0.5757575758, 0.6565656566, 0.6565656565, 0.2070707071, 0.6565656566,
0.8585858586, 0.8989898990, 0.7424242424, 0.6717171717, 0.7979797980, 0.5050505051,
0.2727272727, 0.7121212121, 0.9292929293, 0.6161616162, 0.4040404040, 0.8434343435,
0.3737373737, 0.7626262626, 0.8282828283, 0.6464646465, 0.5707070707),

Risk = c(0.0566037736, 0.0, 0.3584905660, 0.2264150943, 0.1886792453, 0.4150943396,
0.2547169811, 1.0, 0.1981132076, 0.2830188679, 0.0377358491, 0.6415094340, 0.1603773585,
0.2169811321, 0.1886792453, 0.1603773585, 0.0943396226, 0.2169811321, 0.1320754717,
0.3584905660, 0.1981132076, 0.2547169811, 0.1226415094, 0.0849056604, 0.2264150943,
0.1132075472, 0.1886792453, 0.2075471698, 0.1509433962, 0.1698113208),

ESG = c(0.9005235602, 1.0, 0.8167539267, 0.0104712042, 0.0, 0.0314136126, 0.4136125655,
0.0994764398, 0.2408376963, 0.5706806283, 0.1518324607, 0.3612565445, 0.6387434555,
0.3089005236, 0.5183246073, 0.4397905759, 0.3350785340, 0.3874345550, 0.2303664922,
0.1256544503, 0.4921465969, 0.5445026178, 0.3874345550, 0.2827225131, 0.6492146597,
0.1780104712, 0.4921465969, 0.5968586387, 0.5445026178, 0.4397905759) )


#Dirichlet Distribution

preferences = function(n_investors=30, alpha=c(2,2,2)){

 prefs = as.data.frame(rdirichlet(n_investors, alpha))

 colnames(prefs) = c("Return_Pref", "Risk_Pref", "ESG_Pref")

 prefs$Investors = paste0("Investor", 1:n_investors)

 prefs = prefs[, c("Investors", "Return_Pref", "Risk_Pref", "ESG_Pref")]

 return(prefs) }

preferences = preferences(n_investors = 30, alpha = c(2,2,2))


#K-Means Clustering

matrix_preferences = preferences [, c("Return_Pref", "Risk_Pref", "ESG_Pref")]

set.seed(123)

cluster_preferences = kmeans(matrix_preferences, centers = 3, nstart = 25)

preferences$Cluster = cluster_preferences$cluster
```

```
#Silhouette Score
sil = silhouette(cluster_preferences$cluster, dist(matrix_preferences))
mean_sil = mean(sil[,3])
plot(sil)

#Centroids Computation
centroids_pref = aggregate(. ~ Cluster, data = preferences[,-1], FUN=mean)
centroids_pref$Dominant=c("Return", "Risk", "ESG")[apply(centroids_pref
  [,c("Return_Pref", "Risk_Pref", "ESG_Pref")], 1, which.max)]
print(centroids_pref)

#Extraction of Cluster weights from centroids table
w1 = centroids_pref [centroids_pref$Cluster ==1, c("Return_Pref", "Risk_Pref", "ESG_Pref")]
w2 = centroids_pref [centroids_pref$Cluster ==2, c("Return_Pref", "Risk_Pref", "ESG_Pref")]
w3 = centroids_pref [centroids_pref$Cluster ==3, c("Return_Pref", "Risk_Pref", "ESG_Pref")]

#Aggregated Utility Scores
etf_data$AUS_C1 = with(etf_data, Return*w1$Return_Pref + Risk*w1$Risk_Pref + ESG*w1$ESG_Pref)
etf_data$AUS_C2 = with(etf_data, Return*w2$Return_Pref + Risk*w2$Risk_Pref + ESG*w2$ESG_Pref)
etf_data$AUS_C3 = with(etf_data, Return*w3$Return_Pref + Risk*w3$Risk_Pref + ESG*w3$ESG_Pref)
stopifnot(all(c("AUS_C1", "AUS_C2", "AUS_C3") %in% names(etf_data)))

#Quadratic Programming
optimize_alloc = function(scores, lc= 0.005, uc = 0.15){
 scores = as.numeric(scores)
 stopifnot(all(is.finite(scores)),sum(scores)>0)
 s = scores / sum(scores) #Normalization
 n=length(s)
 Dmat = 2 * diag(n)
 dvec = 2 * s
```

```
 #Constraints Definition
#sum(weights) = 1
 Aeq = matrix(1,n,1)
 beq = 1
 #Lower – bound constraint
 A_lc = diag(n)
 b_lc = rep(lc, n)
 #Upper – bound constraint
 A_uc = -diag(n)
 b_uc = rep(-uc, n)
 Amat = cbind(Aeq, A_lc, A_uc)
 bvec = c(beq, b_lc, b_uc)
 solve.QP(Dmat, dvec, Amat, bvec, meq=1)$solution }

#Cluster Portfolio Optimization
res_C1 = optimize_alloc(etf_data$AUS_C1, lc = 0.005, uc = 0.15)
res_C2 = optimize_alloc(etf_data$AUS_C2, lc = 0.005, uc = 0.15)
res_C3 = optimize_alloc(etf_data$AUS_C3, lc = 0.005, uc = 0.15)

#Cluster-specific optimal portfolio compositions
portfolio_tab = function(etf_labels, weights, cluster_name){
 ord = order(weights, decreasing = TRUE)
 data.frame( Cluster = cluster_name, ETF = etf_labels[ord], Weight = round(weights[ord],6)) }

#Optimal weight vectors for each cluster
portfolio_C1 = portfolio_tab(etf_data$ETF, res_C1, "Cluster 1 (Return-Oriented)")
portfolio_C2 = portfolio_tab(etf_data$ETF, res_C2, "Cluster 2 (Risk-Oriented)")
portfolio_C3 = portfolio_tab(etf_data$ETF, res_C3, "Cluster 3 (ESG-Oriented)")
```

```
#Portfolios Barycenters
portfolios = data.frame(
 Cluster= c("Cluster 1: ", "Cluster 2: ", "Cluster 3: "),
 Return = c(sum(etf_data$Return * res_C1), sum(etf_data$Return * res_C2),
 sum(etf_data$Return *   res_C3)),
 Risk = c(sum(etf_data$Risk * res_C1), sum(etf_data$Risk * res_C2), sum(etf_data$Risk * res_C3)),
 ESG = c(sum(etf_data$ESG * res_C1), sum(etf_data$ESG * res_C2), sum(etf_data$ESG * res_C3)))

#Common color scale for AUS across clusters
cmin = min(etf_data[, c("AUS_C1", "AUS_C2", "AUS_C3")])
cmax = max(etf_data[, c("AUS_C1", "AUS_C2", "AUS_C3")])

#Visualization of the Tri-Criterion Utility Space
plot_cluster = function(df, portfolios_row, aus_col, title_text){
 vals=df[[aus_col]]
 plot_ly(data = df, x = ~ Return, y = ~ Risk, z = ~ ESG, type ="scatter3d", mode = "markers",
 showlegend = FALSE,
 marker = list(size = 5, color = vals, cmin = cmin, cmax = cmax, colorscale = list(c(0, "#d3d3d3"),
 c(1, "#1f77b4")),
 colorbar = list(title = list(text = "AUS"))
 )) %>%
 add_trace(data = portfolios_row, x = ~ Return, y = ~ Risk, z = ~ ESG, type ="scatter3d",
 mode = "markers+text",
 showlegend = FALSE, marker = list(size = 10, color = "red", symbol = "diamond"),
 text = ~ Cluster, textposition = "top center", inherit = FALSE) %>%
 layout(title = title_text, scene = list(xaxis = list(title = "Normalized Return"),
 yaxis = list(title = "Normalized Risk"),
 zaxis = list(title = "Normalized ESG") ))}
```

```
#Generate the three cluster-specific AUS-colored utility spaces and portfolio barycenters
plot_cluster(etf_data, portfolios[1,, drop=FALSE], "AUS_C1",
"Cluster 1 – Utility Space (Return-oriented)")
plot_cluster(etf_data, portfolios[2,, drop=FALSE], "AUS_C2", "Cluster 2 – Utility Space (Risk-oriented)")
plot_cluster(etf_data, portfolios[3,, drop=FALSE], "AUS_C3", "Cluster 3 – Utility Space (ESG-oriented)")
```

Script B.2 – Complete RStudio Code Alternative Scenario

```
#Necessary Packages
install.packages("MCMCpack")
install.packages("cluster")
install.packages("factoextra")
install.packages("quadprog")
install.packages("dplyr")
install.packages("plotly")

#Necessary Libraries
library(MCMCpack)
library(cluster)
library(factoextra)
library(quadprog)
library(dplyr)
library(plotly)
```

```
#Uploading Investment Universe Normalized Performances

etf_data <- data.frame(

ETF = c("WBIG_ESG", "RBND_ESG", "USXF_ETF", "IVE_VALUE", "FNDB_ESG", "FNCL_ESG", "FGDL_GOLD",
"CXSE_CHINA","RAYD_DEV", "BKWO_WOM", "UDI_DIV", "LDEM_EM", "FEUS_CLIM", "ISHARES_OPT",
"SUSL_LEAD", "ISHARES_AW", "VANGUARD_E", "SPDR_S&P500", "ISHARES_EA", "ISHARES_EM",
"XTRACKER", "NUVEEN_GRO", "NUVEEN_V", "SPDR_INT", "ISHARES_ADV", "VANGUARD_S", "SPDR_EM",
"ISHARES_KLD", "SPDR_GEN_D", "ISHARES_LC"),

Return = c(0.4595959596, 0.3383838384, 1.0, 0.6565656566, 0.8131313131, 0.7272727273,
0.1414141414, 0.0, 0.5757575758, 0.6565656566, 0.6565656565, 0.2070707071, 0.6565656566,
0.8585858586, 0.8989898990, 0.7424242424, 0.6717171717, 0.7979797980, 0.5050505051,
0.2727272727, 0.7121212121, 0.9292929293, 0.6161616162, 0.4040404040, 0.8434343435,
0.3737373737, 0.7626262626, 0.8282828283, 0.6464646465, 0.5707070707),

Risk = c(0.0566037736, 0.0, 0.3584905660, 0.2264150943, 0.1886792453, 0.4150943396,
0.2547169811, 1.0, 0.1981132076, 0.2830188679, 0.0377358491, 0.6415094340, 0.1603773585,
0.2169811321, 0.1886792453, 0.1603773585, 0.0943396226, 0.2169811321, 0.1320754717,
0.3584905660, 0.1981132076, 0.2547169811, 0.1226415094, 0.0849056604, 0.2264150943,
0.1132075472, 0.1886792453, 0.2075471698, 0.1509433962, 0.1698113208),

ESG = c(0.9005235602, 1.0, 0.8167539267, 0.0104712042, 0.0, 0.0314136126, 0.4136125655,
0.0994764398, 0.2408376963, 0.5706806283, 0.1518324607, 0.3612565445, 0.6387434555,
0.3089005236, 0.5183246073, 0.4397905759, 0.3350785340, 0.3874345550, 0.2303664922,
0.1256544503, 0.4921465969, 0.5445026178, 0.3874345550, 0.2827225131, 0.6492146597,
0.1780104712, 0.4921465969, 0.5968586387, 0.5445026178, 0.4397905759) )


#Dirichlet Distribution (Low-Concentration Regime for Alternative Scenario)

preferences = function(n_investors=30, alpha=c(0.3,0.3,0.3)){

 prefs = as.data.frame(rdirichlet(n_investors, alpha))

 colnames(prefs) = c("Return_Pref", "Risk_Pref", "ESG_Pref")

 prefs$Investors = paste0("Investor", 1:n_investors)

 prefs = prefs[, c("Investors", "Return_Pref", "Risk_Pref", "ESG_Pref")]

 return(prefs) }

preferences = preferences(n_investors = 30, alpha = c(0.3,0.3,0.3))


#K-Means Clustering

matrix_preferences = preferences [, c("Return_Pref", "Risk_Pref", "ESG_Pref")]

set.seed(123)

cluster_preferences = kmeans(matrix_preferences, centers = 3, nstart = 25)

preferences$Cluster = cluster_preferences$cluster
```

```
#Silhouette Score
sil = silhouette(cluster_preferences$cluster, dist(matrix_preferences))
mean_sil = mean(sil[,3])
plot(sil)

#Centroids Computation
centroids_pref = aggregate(. ~ Cluster, data = preferences[,-1], FUN=mean)
centroids_pref$Dominant=c("Return", "Risk", "ESG")[apply(centroids_pref
  [,c("Return_Pref", "Risk_Pref", "ESG_Pref")], 1, which.max)]
print(centroids_pref)

#Extraction of Cluster weights from centroids table
w1 = centroids_pref [centroids_pref$Cluster ==1, c("Return_Pref", "Risk_Pref", "ESG_Pref")]
w2 = centroids_pref [centroids_pref$Cluster ==2, c("Return_Pref", "Risk_Pref", "ESG_Pref")]
w3 = centroids_pref [centroids_pref$Cluster ==3, c("Return_Pref", "Risk_Pref", "ESG_Pref")]

#Aggregated Utility Scores
etf_data$AUS_C1 = with(etf_data, Return*w1$Return_Pref + Risk*w1$Risk_Pref + ESG*w1$ESG_Pref)
etf_data$AUS_C2 = with(etf_data, Return*w2$Return_Pref + Risk*w2$Risk_Pref + ESG*w2$ESG_Pref)
etf_data$AUS_C3 = with(etf_data, Return*w3$Return_Pref + Risk*w3$Risk_Pref + ESG*w3$ESG_Pref)
stopifnot(all(c("AUS_C1", "AUS_C2", "AUS_C3") %in% names(etf_data)))

#Quadratic Programming
optimize_alloc = function(scores, lc= 0.005, uc = 0.15){
 scores = as.numeric(scores)
 stopifnot(all(is.finite(scores)),sum(scores)>0)
 s = scores / sum(scores) #Normalization
 n=length(s)
 Dmat = 2 * diag(n)
 dvec = 2 * s
```

```
 #Constraints Definition
#sum(weights) = 1
 Aeq = matrix(1,n,1)
 beq = 1
 #Lower – bound constraint
 A_lc = diag(n)
 b_lc = rep(lc, n)
 #Upper – bound constraint
 A_uc = -diag(n)
 b_uc = rep(-uc, n)
 Amat = cbind(Aeq, A_lc, A_uc)
 bvec = c(beq, b_lc, b_uc)
 solve.QP(Dmat, dvec, Amat, bvec, meq=1)$solution }

#Cluster Portfolio Optimization
res_C1 = optimize_alloc(etf_data$AUS_C1, lc = 0.005, uc = 0.15)
res_C2 = optimize_alloc(etf_data$AUS_C2, lc = 0.005, uc = 0.15)
res_C3 = optimize_alloc(etf_data$AUS_C3, lc = 0.005, uc = 0.15)
#Cluster-specific optimal portfolio compositions
portfolio_tab = function(etf_labels, weights, cluster_name){
 ord = order(weights, decreasing = TRUE)
 data.frame( Cluster = cluster_name, ETF = etf_labels[ord], Weight = round(weights[ord],6)) }

#Optimal weight vectors for each cluster
portfolio_C1 = portfolio_tab(etf_data$ETF, res_C1, "Cluster 1 (Return-Oriented)")
portfolio_C2 = portfolio_tab(etf_data$ETF, res_C2, "Cluster 2 (Risk-Oriented)")
portfolio_C3 = portfolio_tab(etf_data$ETF, res_C3, "Cluster 3 (ESG-Oriented)")
```

```
#Portfolios Barycenters
portfolios = data.frame(
 Cluster= c("Cluster 1: ", "Cluster 2: ", "Cluster 3: "),
 Return = c(sum(etf_data$Return * res_C1), sum(etf_data$Return * res_C2),
 sum(etf_data$Return * res_C3)),
 Risk = c(sum(etf_data$Risk * res_C1), sum(etf_data$Risk * res_C2), sum(etf_data$Risk * res_C3)),
 ESG = c(sum(etf_data$ESG * res_C1), sum(etf_data$ESG * res_C2), sum(etf_data$ESG * res_C3)))

#Common color scale for AUS across clusters
cmin = min(etf_data[, c("AUS_C1", "AUS_C2", "AUS_C3")])
cmax = max(etf_data[, c("AUS_C1", "AUS_C2", "AUS_C3")])

#Visualization of the Tri-Criterion Utility Space
plot_cluster = function(df, portfolios_row, aus_col, title_text){
 vals=df[[aus_col]]
 plot_ly(data = df, x = ~ Return, y = ~ Risk, z = ~ ESG, type ="scatter3d", mode = "markers",
 showlegend = FALSE,  marker = list(size = 5, color = vals, cmin = cmin, cmax = cmax,
 colorscale = list(c(0, "#d3d3d3"), c(1, "#1f77b4")),
 colorbar = list(title = list(text = "AUS"))
 )) %>%

 add_trace(data = portfolios_row, x = ~ Return, y = ~ Risk, z = ~ ESG, type ="scatter3d",
 mode = "markers+text",
 showlegend = FALSE, marker = list(size = 10, color = "red", symbol = "diamond"),
 text = ~ Cluster, textposition = "top center", inherit = FALSE) %>%
 layout(title = title_text, scene = list(xaxis = list(title = "Normalized Return"),
 yaxis = list(title = "Normalized Risk"),
 zaxis = list(title = "Normalized ESG") ))}
```

```
#Generate the three cluster-specific AUS-colored utility spaces and portfolio barycenters
plot_cluster(etf_data, portfolios[1,, drop=FALSE], "AUS_C1",
"Cluster 1 – Utility Space (Return-oriented) ")
plot_cluster(etf_data, portfolios[2,, drop=FALSE], "AUS_C2", "Cluster 2 – Utility Space (Risk-oriented) ")
plot_cluster(etf_data, portfolios[3,, drop=FALSE], "AUS_C3", "Cluster 3 – Utility Space (ESG-oriented) ")
```