Constrained ETF Portfolio Optimization via Mixed-Integer Quadratic Programming

Examiner:	Dr. Sara Boni
Chair:	Chair of Entrepreneurial Finance
Submitted by:	Christian Baumann, Duc Trung Daniel Tran, Zhiqian Li
Submission Date:	15.01.2026

Abstract

The expanding universe of exchange-traded funds (ETFs) has increased the complexity of portfolio construction, rendering unconstrained mean–variance optimization impractical due to excessive concentration and limited implementability. This paper proposes a constrained ETF portfolio optimization framework formulated as a mixed-integer quadratic program (MIQP) that explicitly incorporates diversification, cost, and operational constraints. Building on the classical Markowitz mean–variance paradigm, the model integrates cardinality constraints, ETF-level weight bounds, country and industry exposure limits, total expense ratio penalties, and an explicit correlation penalty to promote robust diversification.

The empirical analysis is conducted on a universe of 120 highly liquid iShares ETFs, using a reproducible data pipeline based on web-scraped prices, holdings, and fund characteristics. The framework is calibrated to generate low-, medium-, and high-risk portfolios while maintaining an identical constraint structure. Results show that the optimized portfolios are sparse, interpretable, and consistently outperform an MSCI World benchmark in terms of risk-adjusted performance. The findings demonstrate that MIQP-based portfolio optimization offers a transparent and economically meaningful approach to ETF allocation under realistic investment constraints.

Contents

List of Figures

List of Tables

List of Abbreviations

AUM	assets under management
ETF	exchange-traded fund
MIQP	mixed-integer quadratic program
TER	Total Expense Ratio

1. Introduction

1.1 Motivation

Exchange-traded funds (ETFs) are investment vehicles that track underlying indices, asset classes, or investment strategies and are traded on exchanges similarly to individual stocks. Due to their diversification, transparency, liquidity, and relatively low costs, ETFs have become increasingly popular among both retail and institutional investors. At the same time, the growing number and variety of available ETFs has substantially increased the complexity of portfolio construction. Investors are therefore faced with the challenge of selecting and allocating capital across a large set of ETFs while balancing risk, return, and diversification objectives.

This development motivates the need for systematic and data-driven portfolio construction frameworks that go beyond ad hoc selection rules. In particular, realistic portfolio optimization approaches must account not only for risk and return trade-offs, but also for practical considerations such as diversification requirements, cost efficiency, and implementability constraints.

1.2 Research Question and Contribution

This study addresses the following research question: *What is the optimal distribution of investment capital across a diversified ETF portfolio when formulated as a mixed-integer quadratic program (MIQP)?*

More specifically, the study seeks to construct an optimal ETF portfolio that balances expected returns and risk while incorporating realistic selection and allocation constraints through an MIQP framework. To this end, the ETF portfolio selection problem is modeled as a mixed-integer quadratic program, enabling the simultaneous optimization of portfolio risk and return while explicitly controlling portfolio composition.

The proposed framework leverages historical price data as well as ETF-specific characteristics, such as expense ratios and sector exposures. In addition, an end-to-end pipeline for data retrieval, preprocessing, and portfolio optimization is developed. The pipeline is designed to be extensible and adaptable to different datasets, optimization objectives, and constraint specifications, thereby facilitating replication and future extensions. The effectiveness of the proposed approach is empirically assessed through a benchmark comparison with broad market indices.

1.3 Structure of the Paper

The remainder of this work is organized as follows. Chapter 2 reviews related literature on portfolio optimization and ETF-based investment strategies. Chapter 3 describes the research design, data collection, and preprocessing steps. Chapter 4 presents the portfolio optimization model and its formulation as a MIQP. Chapter 5 reports and discusses the empirical results, including a comparison with benchmark indices. Chapter 6 concludes the study and outlines potential directions for future research.

2. Related Work

2.1 Classical Portfolio Optimization

The foundation of modern portfolio optimization dates back to the work of Markowitz, 1952, who introduced the mean–variance framework for balancing expected returns and portfolio risk. In this classical setting, portfolio selection is an optimization problem based on expected returns and the covariance structure of asset returns. While the Markowitz framework establishes a theoretical basis for diversification, it is well known to be sensitive to estimation error and often produces highly concentrated portfolios when applied without additional constraints.

2.2 Machine Learning Approaches

More recently, ETF portfolio construction has increasingly adopted machine learning and artificial intelligence methods. These approaches typically rely on predictive models to forecast ETF returns or directly generate portfolio weights based on historical data. An example of this line of research is provided by Day and Lin, 2020, who apply artificial intelligence techniques to ETF market prediction and portfolio optimization. While such methods have demonstrated promising empirical performance, they often operate on relatively small asset universes and may offer limited transparency with respect to the economic drivers of portfolio allocations.

2.3 Positioning of the Present Study

In contrast to predictive approaches, the present study follows an optimization framework that emphasizes transparency and explicit control over portfolio characteristics. By formulating the ETF portfolio selection problem as a mixed-integer quadratic program, the framework builds on classical mean–variance theory while directly incorporating diversification, cost, and implementability constraints. This positions the proposed approach between traditional portfolio theory and modern machine learning-based methods, combining interpretability with practical relevance for ETF portfolio construction.

3. Methodology

3.1 Research Design and Data Collection

3.1.1 Research Methodology

This study adopts a quantitative, model-based research design aimed at constructing diversified ETF portfolios under realistic investment constraints. Building on the classical mean–variance framework of portfolio selection introduced by Markowitz, 1952, the analysis focuses on the allocation stage of the investment process rather than on return prediction. Expected returns and risk measures are therefore treated as model inputs, while the central objective is to analyze how different risk preferences translate into optimal portfolio compositions when diversification, cost, and implementability constraints are explicitly imposed.

Recent research on ETF portfolio construction frequently relies on machine learning and artificial intelligence methods to predict ETF returns and optimize portfolio weights, as illustrated by studies such as Day and Lin, 2020. While these approaches have demonstrated promising empirical performance, they often operate on relatively small asset universes and may provide limited economic interpretability of the resulting allocations. In contrast, the present study deliberately retains a convex optimization-based framework, allowing for transparent economic interpretation and direct control over portfolio characteristics through explicit constraints and penalty terms.

3.1.2 Asset Universe Construction and ETF Selection

The empirical analysis is conducted on a universe of exchange-traded funds (ETFs), which are particularly well suited for portfolio construction due to their transparency, diversification properties, and widespread adoption by both private and institutional investors. ETFs provide direct exposure to broad asset classes, regions, and industries, while offering detailed and regularly updated disclosures on fund holdings and characteristics.

To construct a broad and economically relevant investment universe, all available iShares ETFs were ranked according to their assets under managements (AUMs).

From this ranking, the top 120 ETFs were selected. This size-based filtering ensures that the asset universe consists of highly liquid and widely held instruments, thereby improving practical implementability and reducing the influence of niche or illiquid funds. Importantly, the selection procedure does not condition on past performance or thematic classifications, which mitigates ex-ante selection bias and preserves objectivity in the construction of the initial universe.

The choice of a relatively large initial universe follows common practice in institutional portfolio construction, where a broad investable set is defined prior to optimization. Starting from a large universe allows the optimization model to diversify across regions and industries, while final portfolio sparsity is achieved through explicit cardinality and concentration constraints. In this way, the model is required to select a limited number of ETFs based on the imposed objective function and constraints, rather than through ad hoc pre-selection.

3.2 Data preparation: Webscraping

Since we want to do the optimization based on most up-to-date and official data, we decided to use webscraping to collect price data from Yahoo Finance, ETF holdings data from iShares and Total Expense Ratio (TER) from justETF. The data retrieval process is implemented in the Python programming language, using the *requests* library (Python Software Foundation, 2023) for webpage interaction and the *pandas* library (pandas development team, 2020) for data parsing, transformation, and aggregation.

3.2.1 Data sources

Yahoo Finance provides detailed historical data on the evolution of ETF prices over a selected period of time. The platform offers daily time-series data on trading prices, including closing prices and adjusted closing prices. The adjusted closing price takes into account corporate actions such as stock splits and dividend distributions, and therefore provides a more accurate record of stock value over time.

iShares ETF Fund Navigator tool provides authoritative fund-level disclosures for iShares ETFs, including portfolio weights, sector allocations, and geographic exposure. This study primarily focused on iShares, a family of exchange-traded funds (ETFs) and index mutual funds managed by BlackRock. Disclosures from iShares ETF Fund Navigator tool represent one of the most official and up-to-date sources of information on iShares fund composition.

justETF serves as a complementary source of data in our study. While the platform provides comprehensive information on ETFs, including TER, holdings data and price data, its holdings data are less up-to-date than those from iShares. Therefore, we chose to rely on iShares for detailed holdings data, and used justETF exclusively for obtaining TER data.

3.2.2 Price Data

For downloading price data from Yahoo Finance, we used the Python package yfinance (Aroussi, 2023) and requested data using ticker symbols. However, for many assets, Yahoo Finance requires fully qualified ticker symbols, including the exchange suffix. A fully qualified ticker symbol for *SXR8* is *SXR8.DE*. Therefore, we first conduct a preliminary search for the fully qualified ticker symbols for each fund. We made use of the search functionality of yfinance and retrieved all matching ticker symbols associated with a given ticker abbreviation. Then, the most liquid ticker is selected from the returned tickers. From this preliminary search, we constructed a mapping of ticker abbreviations to fully qualified ticker symbols. With this mapping, we downloaded daily price data for the past six years. The raw data obtained from yfinance are in a pivoted format, with date and ticker symbol as row and column indices. To facilitate later modeling steps, we convert the data table into a long format.

3.2.3 ETF Holdings Data and Total Expense Ratio Data

Holdings information for iShares ETFs is obtained from the iShares online disclosure platform, as mentioned in Section 3.2.1. The collection of holdings data for each ETF in the dataset follows a structured procedure, as illustrated in Figure 1.

For each ETF, the data collection process begins with the identification of the corresponding product-specific webpage. From this page, the official holdings disclosure resource is located, and a download link to the underlying data file is generated. The resulting files are subsequently requested and retrieved for offline processing.

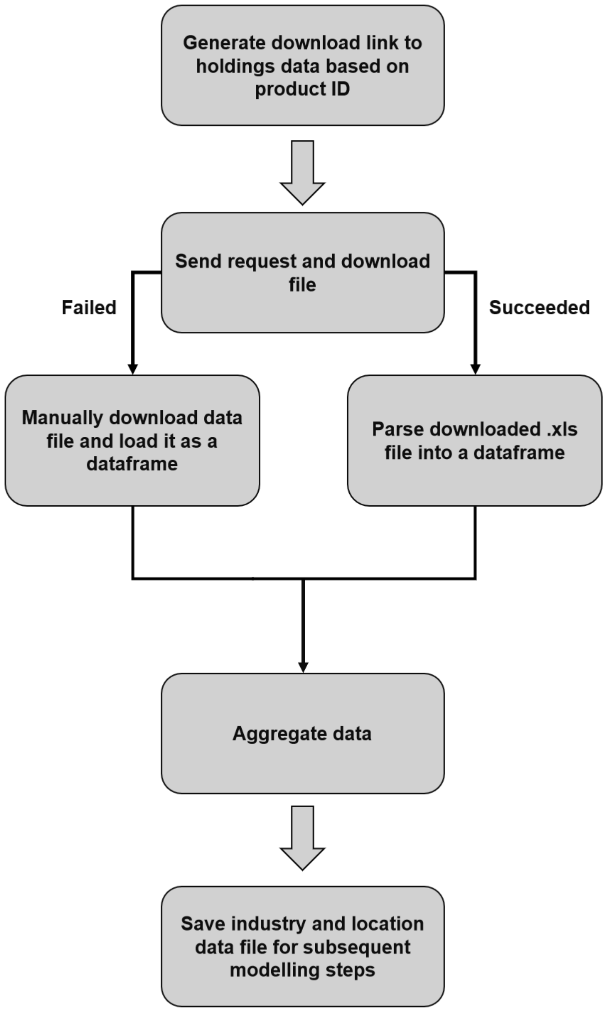
Table 1: Header rows for extracted ETF sector holdings data

Sector	Weight	Ticker
Cash / Derivatives	0.0052	IQQ0
Energy	0.0342	IQQ0
Financials	0.1426	IQQ0
Health Care	0.1514	IQQ0
Information Technology	0.2255	IQQ0

The downloaded holdings files are provided in a non-standard file format and includes header rows in addition to holdings data. Therefore, a parsing step is required to identify the relevant data region within each file. Header rows and auxiliary information are removed, and the remaining tabular data are parsed into a standardized pandas dataframe.

Once extracted, the holdings data for each ETF are aggregated at the industry and location level. This aggregation enables the incorporation of holdings-based characteristics into the portfolio optimization framework. The industry and location data of each ETF are then concatenated into two single dataframes and saved as csv files for later modelling steps. An exemplary table for sector holdings data resulting from the above process is presented in Table 1.

Figure 1: Data collection process for a single ETF



In a limited number of cases (only one ETF in our dataset), automated data retrieval is not possible due to access limitations or file availability issues. These exceptional cases are addressed through manual download from the issuer's website.

Finally, the TER information is obtained from justETF through a standard web-scraping process. The respective ETF webpages are requested and the retrieved HTML content is processed using Beautiful Soup package (Richardson, 2023) to extract the TER values.

The structured and reproducible process ensures data consistency and completeness.

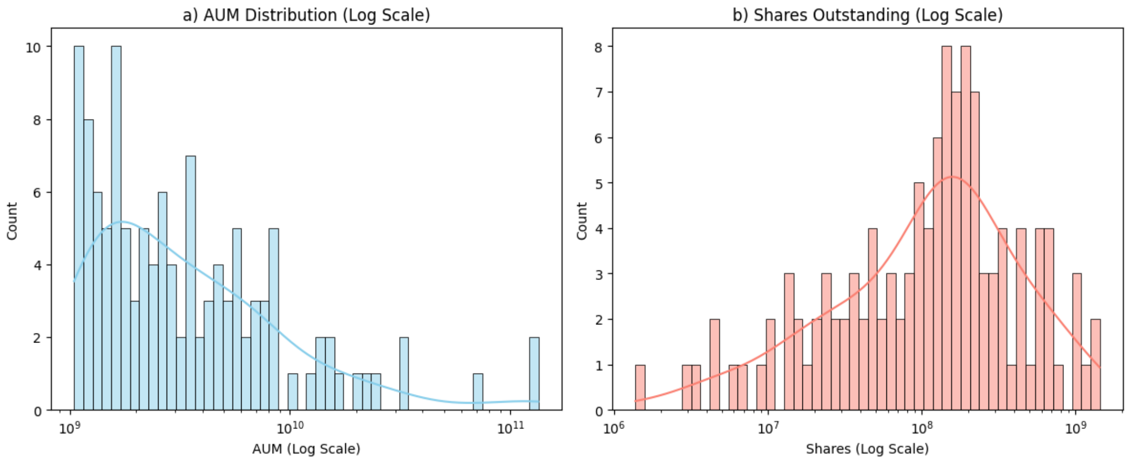
4. Data overview

In accordance with Section 3.1.2, we sorted the ETFs descendingly by their fund volumes and selected the top 120 funds. In this chapter, descriptive statistics and illustrative plots are provided to gain a clearer overview on the selected ETFs.

4.1 Characteristics of Fund Size and Issuance

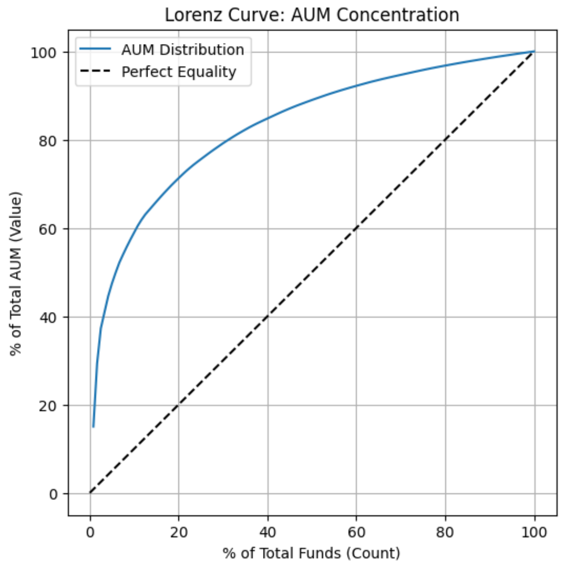
Both distribution plots in Figure 2 adopt a logarithmic scale for clearer illustration of values spanning several orders of magnitude. From subplot a), it can be observed that the distribution of Asset under management among the top 120 selected ETFs is skewed to the right. This indicates that a small number of funds in our dataset attract large amounts of capital while the remaining funds are significantly smaller and individually attract substantially less capital. Subplot b) shows that shares outstanding of the selected ETFs is more evenly distributed, exhibiting a slightly longer left tail than the right. This indicates a healthy diversity of funds offering in our dataset, i.e., shares outstanding are not concentrated in a small number of top funds, even though the asset under management varies drastically among all ETFs.

Figure 2: Asset Under Management and Shares Outstanding Distribution



Distribution of assets under management and shares outstanding for the top 120 selected ETFs, plotted on a logarithmic scale.

Figure 3: Lorenz Curve for Asset Under Management



The horizontal axis represents the cumulative percentage count of funds in the dataset, ordered from highest to lowest AUM, and the vertical axis represents the corresponding cumulative percentage of total AUM value.

In Figure 3, a simple concentration analysis is conducted with the use of Lorenz curve. It can be observed from the plot that the total asset under management is largely concentrated in the top funds in our dataset, which aligns with our observation from Figure 2 subplot (a).

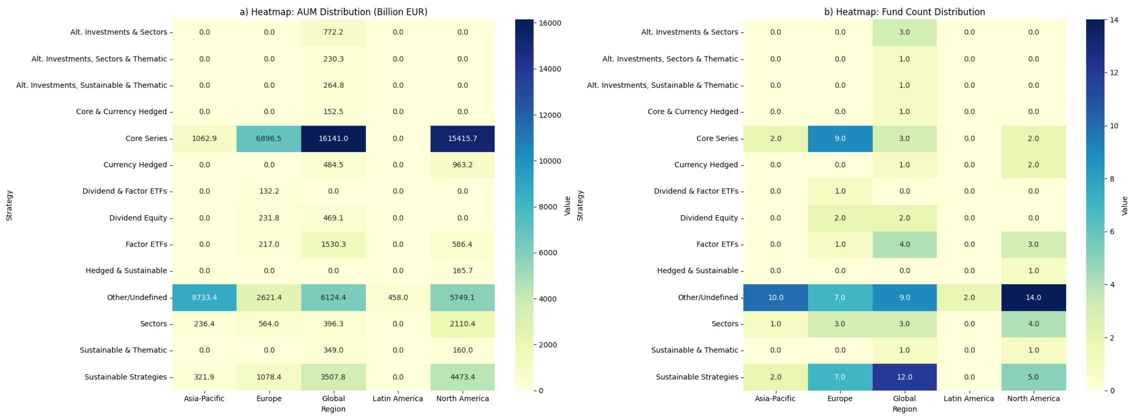
4.2 Regional Segmentation and Strategic Composition

This section examines the regional segmentation and strategic composition of ETFs in the dataset using heatmap analysis. The two subplots in Figure 4 display the absolute distribution of strategies across regions. Subplot (a) displays the distribution in absolute AUM values to represent capital scale, while subplot (b) presents the corresponding fund counts. Additionally, the two subplots in Figure 5 illustrate the same categorical distribution, but are normalized by row. Those normalized plots highlight the relative concentration of assets for each strategy, enabling the comparison of the regional focus of different strategies regardless of the total market share of strategies.

The absolute distribution plots reveal a strong contrast between product volume and capital allocation. As shown in Figure 4(a), the Core Series strategy dominates other strategies in terms of AUM, with the vast majority of capital concentrated in the Global and North America regions.

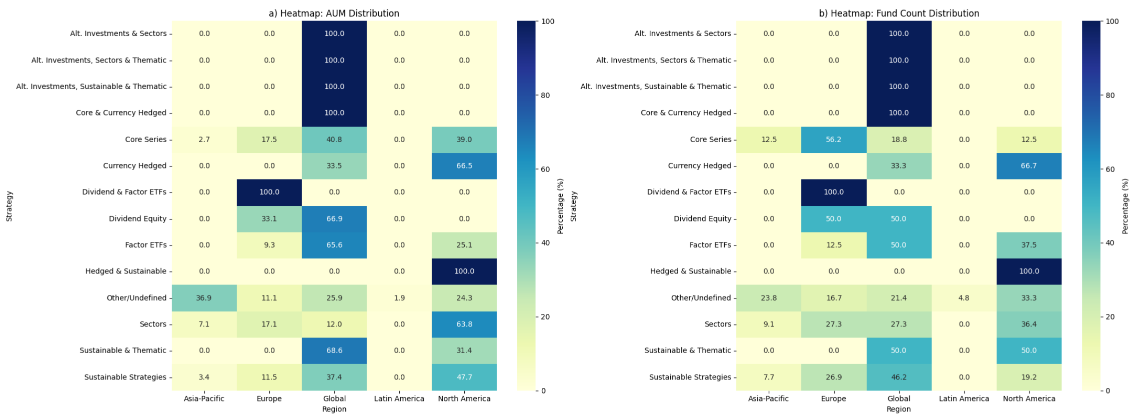
In contrast, Figure 4(b) indicates that the highest numerical density of funds (fund counts) lies in the Sustainable Strategies and Other/Undefined categories. This suggests that these market segments are relatively highly fragmented, and issuers are

Figure 4: Heatmap for AUM and Fund Distribution in Absolute Value



launching new products though these funds are currently managing significantly lower capital volumes compared to the established Core Series.

Figure 5: Heatmap for Normalized AUM Distribution



For each row, proportion values are normalized such that their sum equals one.

A comparative analysis of the heatmaps in Figure 5 reveals distinct regional specializations for specific investment strategies. Alternative Investment strategies are heavily concentrated in the Global region, suggesting these niche assets are primarily packaged as broad-market products rather than regional-specific ones. Similarly, Hedged & Sustainable strategies are found exclusively in North America, while Dividend & Factor ETFs appear unique to Europe within this dataset.

A notable discrepancy is observed in the Dividend Equity segment. While the fund count (Figure 5 subplot b) is distributed evenly between Europe and the Global region, the AUM-weighted distribution (Figure 5 subplot a) shows a significant bias toward the Global region. This indicates that while Europe offers a similar variety of dividend-focused products, the capital concentration is vastly superior in Global dividend funds.

5. Portfolio Optimization Model

This study considers the construction of a diversified exchange-traded fund (ETF) portfolio using a constrained mean–variance optimization framework, following the classical formulation of portfolio selection introduced by Markowitz, 1952 augmented with explicit diversification, cost, and concentration controls. The objective is to identify an allocation that balances expected return against multiple dimensions of risk while respecting practical investment constraints commonly faced by institutional and private investors.

5.1 Mathematical Model

5.1.1 Asset Universe and Return Estimation

Let $i = 1, \ldots, N$ index a universe of N ETFs. Historical daily prices are used to compute logarithmic returns, which are subsequently annualized. Based on these returns, we estimate:

- the vector of expected annual returns $\boldsymbol{\mu} \in \mathbb{R}^N$,
- the annualized return covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{N \times N}$,
- the correlation matrix $\boldsymbol{R} \in \mathbb{R}^{N \times N}$.

While the covariance matrix captures overall portfolio variance, correlations play a key role in diversification. To explicitly penalize excessive co-movement between ETFs, we define the off-diagonal correlation matrix

$$\boldsymbol{C}_{\text{off}} = \boldsymbol{R} - \boldsymbol{I},$$

which excludes self-correlation terms and isolates cross-asset dependence.

5.1.2 Portfolio Weights and Auxiliary Variables

Let $w_i \geq 0$ denote the portfolio weight invested in ETF i. Short-selling is excluded, reflecting the long-only nature of most ETF portfolios. To control portfolio sparsity, we introduce binary selection variables $y_i \in \{0,1\}$ indicating whether ETF i is included in the final portfolio.

5.1.3 Objective Function

The portfolio is obtained by minimizing the following objective function:

$$\begin{aligned} \min_{\boldsymbol{w},\boldsymbol{y}} \quad & \alpha\, \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w} + \beta \sum_{i=1}^{N} \sum_{j=1}^{N} C_{\mathsf{off},ij} w_i w_j - \gamma\, \boldsymbol{\mu}^\top \boldsymbol{w} \\ & + \delta \sum_{i=1}^{N} \mathsf{TER}_i w_i + \lambda \sum_{i=1}^{N} w_i^2. \end{aligned} \tag{5.1}$$

Each term has a financial interpretation:

- The first term represents the classical mean–variance risk penalty, capturing total portfolio volatility in the sense of Markowitz, 1952.
- The second term explicitly penalizes correlated ETF exposures, promoting diversification beyond variance minimization alone, thereby achieving lower portfolio volatility.
- The third term rewards higher expected portfolio returns (based on historical data).
- The fourth term incorporates ongoing management costs through weighted total expense ratios (TERs).
- The final ℓ_2 regularization term discourages excessive concentration in individual ETFs, leading to more stable and implementable allocations.

The scalar parameters $\alpha, \beta, \gamma, \delta,$ and λ control the investor's risk aversion, diversification preference, return orientation, cost sensitivity, and concentration aversion, respectively.

5.1.4 Budget and Weight Constraints

The portfolio is fully invested:

$$\sum_{i=1}^{N} w_i = 1.$$

Individual ETF weights are bounded to prevent overexposure to a single fund:

$$0 \le w_i \le w_i^{\max}, \qquad \forall i = 1, \ldots, N.$$

5.1.5 Geographic and Sector Diversification Constraints

To ensure adequate diversification across regions and industries, the portfolio is constrained with respect to country and sector exposures.

Let A_{ci} denote the fraction of ETF i invested in country c, for $c = 1, \ldots, C$. Country exposure constraints are given by

$$\sum_{i=1}^{N} A_{ci} w_i \le E_c^{\max}, \qquad \forall c.$$

Similarly, let B_{ki} denote the exposure of ETF i to industry k, for $k = 1, \ldots, I$. Industry diversification is enforced via

$$\sum_{i=1}^{N} B_{ki} w_i \le I_k^{\max}, \qquad \forall k.$$

5.1.6 Cardinality Constraint

To limit operational complexity and transaction costs, the number of ETFs in the portfolio is restricted. This is achieved using a mixed-integer formulation:

$$w_i \le M y_i, \qquad \forall i,$$

$$\sum_{i=1}^{N} y_i \le K,$$

where K is the maximum number of ETFs allowed and M is a sufficiently large constant. This constraint promotes sparse, interpretable portfolios while preserving diversification benefits.

5.1.7 Optimization Problem Class

The resulting model is a **mixed-integer quadratic program (MIQP)** with a convex quadratic objective and linear constraints, solved by the Gurobi Solver (Gurobi Optimization, LLC, 2024) implemented with the programming language Python. This formulation integrates classical mean–variance theory with modern portfolio construction considerations, including explicit correlation control, cost awareness, and real-world implementability constraints.

As such, the framework provides a flexible and economically interpretable tool for ETF portfolio construction under realistic investment guidelines.

5.2 Constraint Design and Parameter Calibration

The portfolio optimization model incorporates a set of structural constraints and penalty parameters reflecting practical considerations in ETF portfolio construction. This section provides the economic and operational rationale for the chosen bounds and parameter values.

5.2.1 Maximum Country Exposure

Country-level exposure is capped at

$$E_c^{\max} = 30\%, \qquad \forall c.$$

This upper bound serves multiple purposes. First, it prevents excessive concentration in a single macroeconomic region, thereby reducing vulnerability to country-specific risks such as monetary policy shocks, political instability, or regulatory changes. Second, a 30% cap aligns with common diversification guidelines used in multi-asset and global equity mandates, where single-country exposures above one-third of the portfolio are typically viewed as concentrated unless explicitly benchmark-driven.

From a practical standpoint, this constraint also mitigates the tendency of unconstrained mean–variance optimization to overweight large and highly liquid markets (e.g. the United States) due to their historically favorable risk–return profiles. By imposing a cap, the model is forced to allocate capital across multiple regions, improving global diversification and robustness.

5.2.2 Maximum Industry Exposure

Industry exposures are similarly constrained by

$$I_k^{\max} = 30\%, \qquad \forall k.$$

This restriction limits sector concentration and reduces exposure to industry-specific cycles, technological disruptions, and regulatory changes. In particular, sectors such as technology or financials often dominate capitalization-weighted indices and may receive disproportionate weights in unconstrained optimization settings. The 30% threshold ensures that no single industry becomes the dominant driver of portfolio performance or risk.

5.2.3 ETF-Level Weight Constraints

Individual ETF weights are bounded by

$$0 \leq w_i \leq 50\%, \qquad \forall i.$$

This constraint prevents the portfolio from degenerating into a quasi-single-fund solution, even when one ETF appears highly attractive on a risk-adjusted basis. From an implementation perspective, it reduces idiosyncratic fund risk, including tracking error, replication method risk, and provider-specific operational risk.

5.2.4 Cardinality Constraint

The number of ETFs in the portfolio is limited to

$$\sum_{i=1}^{N} y_i \leq K, \quad \text{with } K = 10.$$

This cardinality constraint reflects real-world considerations such as transaction costs, monitoring effort, and portfolio transparency. A limited number of holdings simplifies portfolio management, improves interpretability for investors, and reduces rebalancing complexity.

From an economic perspective, this constraint enforces a trade-off between diversification and simplicity, consistent with the observation that a relatively small number of broad ETFs can capture a substantial portion of global market risk premia.

5.2.5 Risk Profile Parameterization and Customization

Rather than relying on a single fixed calibration, the penalty and reward parameters α, β, γ, δ, and λ are used to generate a set of predefined portfolio solutions corresponding to different investor risk preferences. Specifically, the model is calibrated to produce three representative risk profiles: *low risk*, *medium risk*, and *high risk*.

The core idea is to adjust the relative importance of risk-related penalties and return-oriented rewards while keeping the constraint structure unchanged. This approach ensures that all resulting portfolios satisfy the same diversification, cost, and implementability requirements, thereby making the outcomes directly comparable.

Risk Profiles

Three portfolio configurations are generated by calibrating the objective function parameters to reflect different risk preferences.

The *low-risk* profile assigns greater weight to variance, correlation, and concentration penalties relative to expected return, resulting in stable and well-diversified allocations. The *medium-risk* profile balances risk penalties and return incentives, representing a diversified long-term investment strategy. The *high-risk* profile increases the emphasis on expected returns, allowing more aggressive allocations within the imposed diversification and cardinality constraints.

Each profile serves as a baseline that can be further fine-tuned to accommodate individual investor preferences. Differences across portfolios therefore reflect risk appetite rather than changes in the underlying constraint structure.

6. Results

This chapter presents the optimal ETF portfolios obtained using a constrained mixed-integer quadratic program (MIQP) optimization framework. To determine the optimal distribution of investment capital, the model was solved under three distinct risk profiles—High, Medium, and Low—defined by variations in the objective-function parameters $(\alpha, \beta, \gamma, \delta, \lambda)$. All portfolios satisfy identical feasibility, diversification, and exposure constraints. All reported performance metrics are in-sample and intended for structural comparison across risk profiles rather than for performance forecasting.

6.1 Portfolio-Level Comparison

Table 2 reports the key portfolio performance metrics across the three risk profiles, evaluated over a six-year historical window and compared against the MSCI World benchmark.

Table 2: Portfolio Performance compared to the MSCI World Benchmark

	High-Risk	**Medium-Risk**	**Low-Risk**	**MSCI World**
α (variance penalty)	0.5	1.0	2.0	–
β (correlation penalty)	0.3	1.0	4.0	–
γ (return reward)	4.0	1.5	0.8	–
δ (TER penalty)	0.1	0.2	0.3	–
λ (concentration penalty)	0.05	0.2	0.5	–
Number of ETFs	6	5	4	1
Expected return	18.39%	16.72%	10.03%	11.26%
Volatility	19.43%	18.25%	16.04%	16.98%
Sharpe ratio	0.95	0.92	0.63	0.66
Weighted average TER	0.43%	0.41%	0.30%	0.20%

The results show that the optimized portfolios offer competitive risk-adjusted returns. Specifically, the high- and medium-risk profiles outperform the MSCI World benchmark in terms of Sharpe ratio, while the low-risk profile successfully minimizes absolute volatility. The **High-Risk** profile yields the most attractive risk-adjusted outcome, achieving the highest Sharpe ratio of 0.95 and an expected return of 18.39%.

As the optimization parameters shift toward risk aversion (increasing α, β, and λ), the **Low-Risk** profile successfully minimizes volatility to 16.04%, lower than the benchmark's 16.98%, while still maintaining a competitive Sharpe ratio of 0.63 compared to the benchmark's 0.66. The cost of this optimization is a higher weighted average TER (0.30%) compared to the benchmark (0.20%), which is offset by the risk reduction.

6.2 Optimal ETF Allocations

To understand the optimal distribution of capital, we analyze the specific asset selection in Table 3. The MIQP solver identifies a "core" allocation common to all profiles while rotating satellite assets based on risk tolerance.

Table 3: ETF Portfolio Weights by Risk Profile

ETF	High-Risk (%)	Medium-Risk (%)	Low-Risk (%)
EURO STOXX Banks 30–15 (EXX1.DE)	29.68	25.62	–
Gold Producers (Acc) (IS0E.DE)	29.35	27.60	25.20
S&P 500 Information Technology Sector (QDVE.DE)	25.66	25.29	–
Nikkei 225 (Acc) (SXRZ.DE)	15.21	–	–
MSCI China A (Acc) (36BZ.DE)	–	20.78	24.80
MSCI Europe CTB Enhanced ESG (DIST) (EMNU.DE)	–	–	50.00
S&P 500 Health Care Sector (QDVG.DE)	–	0.70	–
S&P 500 Communication Sector (IUCM.L)	0.09	–	–
Core MSCI Pacific ex Japan (SXR1.DE)	0.01	–	–
STOXX Global Select Dividend 100 (ISPA.DE)	–	–	–

Across the High and Medium profiles, the optimal distribution relies heavily on **Financials** (EURO STOXX Banks) and **Materials** (Gold Producers). However, the Low-Risk profile introduces a structural shift to broad Eurozone exposure.

Significant rotation occurs in the remaining capital:

- The **High-Risk** portfolio directs capital toward growth, allocating 25.66% to US Information Technology and 15.21% to Japanese Equities (Nikkei 225).
- The **Medium-Risk** portfolio replaces Japanese exposure with Emerging Markets (MSCI China A) at 20.78%, likely to manage correlation while maintaining growth through US Tech (25.29%).
- The **Low-Risk** portfolio completely eliminates the volatile Information Technology and Banking sector ETFs. Instead, it reallocates that capital to broad Eurozone Equities (MSCI EMU) at 50.00%, while increasing the exposure to China (24.80%) and maintaining a solid base in Gold Producers (25.20%).

6.3 Geographic and Industry Exposure

The distribution of capital is tightly bounded by the predefined feasibility constraints, specifically the 30% cap on any single country or industry.

6.3.1 Country Exposure

Table 4 demonstrates how the optimization hits the upper bounds of geographic constraints.

Table 4: Top 10 Country Exposures (%)

Rank	High-Risk		Medium-Risk		Low-Risk	
	Country	**%**	**Country**	**%**	**Country**	**%**
1	United States	30.00	United States	30.00	China	26.36
2	Canada	16.31	China	22.50	Canada	14.00
3	Japan	15.27	Canada	15.34	United Kingdom	11.05
4	Spain	8.83	Spain	9.00	France	7.70
5	Italy	7.13	Italy	7.30	Switzerland	7.46
6	France	4.24	Switzerland	5.51	Germany	7.33
7	South Africa	3.21	South Africa	3.28	Netherlands	3.76
8	Australia	3.09	Australia	2.90	United States	3.67
9	Netherlands	2.40	Germany	2.04	Spain	2.97
10	Germany	2.36	Ireland	0.80	South Africa	2.75

The **United States** exposure is capped at the maximum 30.00% for High and Medium portfolios, indicating it is the primary driver of expected returns. However, in the **Low-Risk** profile, US exposure drops significantly to 3.67%. Diversification in this conservative profile is instead achieved through significant allocations to **China** (26.36%) and a broad mix of European nations including the **United Kingdom** 11.05%, **France** 7.70%, and **Switzerland** 7.46%.

6.3.2 Industry Exposure

Table 5 highlights the sectoral structural breaks required to minimize volatility.

Table 5: Top 10 Industry Exposures (%)

Rank	High-Risk		Medium-Risk		Low-Risk	
	Industry	**%**	**Industry**	**%**	**Industry**	**%**
1	Financials	30.00	Financials	30.00	Materials	30.00
2	Information Technology	30.00	Information Technology	30.00	Financials	17.95
3	Materials	30.00	Materials	30.00	Industrials	12.71
4	Industrials	2.80	Industrials	3.25	Information Technology	9.69
5	Consumer Discretionary	2.78	Consumer Staples	1.75	Health Care	8.48
6	Communication Services	1.91	Health Care	1.74	Consumer Staples	6.47
7	Health Care	1.23	Consumer Discretionary	1.22	Consumer Discretionary	5.05
8	Consumer Staples	0.60	Utilities	0.74	Utilities	3.40
9	Cash / Derivatives	0.35	Energy	0.60	Energy	2.53
10	Real Estate	0.26	Communication Services	0.27	Communication Services	1.93

The **High-Risk** and **Medium-Risk** portfolios maximize aggressive sector bets, hitting the 30.00% cap simultaneously in **Financials**, **Information Technology**, and **Materials**.

To achieve the optimal Low-Risk distribution, the model enforces a pivotal shift. While **Materials** remain saturated at 30.00%, exposure to Financials drops to 17.95% and Information Technology falls to 9.69%. The capital is redistributed towards **Industrials** 12.71% and **Health Care** 8.48%, confirming that within this framework, the optimal method to reduce portfolio variance is to pivot away from pure Tech and Banking concentration into a broader mix of industrial and defensive sectors.

7. Outlook and Potential Extensions

Several ways exist to extend and improve the proposed portfolio optimization framework.

First, the estimation of expected returns could be enhanced by incorporating machine learning or factor-based forecasting models. Instead of relying on historical average returns over a fixed window, predictive models could exploit macroeconomic variables, valuation metrics, or momentum signals to generate forward-looking return estimates, thereby reducing estimation error and improving out-of-sample performance.

Second, the current constraint set could be generalized to allow for adaptive or scenario-dependent bounds. For instance, country or sector exposure limits could vary across market regimes or be relaxed for regions with lower volatility or strong diversification benefits. Such flexibility would enable the model to respond more dynamically to changing market conditions while preserving risk controls.

Third, a key limitation of the present framework is its static, single-period formulation. In practice, portfolio management is inherently dynamic: asset returns, risk characteristics, and optimal allocations evolve over time, and portfolios are periodically rebalanced in response to new information.

A promising extension is therefore the formulation of a multi-period optimization problem in which portfolio weights are updated across discrete time steps. In such a setting, the objective function could be augmented with explicit rebalancing costs or turnover penalties to discourage excessive trading. This would allow the model to balance expected performance improvements against transaction costs and operational frictions. Formally, a dynamic version of the model could introduce a time-indexed portfolio weights and link consecutive periods through rebalancing constraints. This extension would transform the framework into a true portfolio management model, capable of capturing path dependency and trading dynamics while retaining the core convex quadratic structure within each period.

Fourth, to assess the economic relevance of the proposed optimization framework, future work should incorporate systematic backtesting. An in-sample backtest can be implemented by repeatedly solving the optimization problem on a rolling or expanding historical window and evaluating realized portfolio performance over subsequent holding periods. This approach enables the evaluation of portfolio stability, turnover, and sensitivity to estimation error using standard performance metrics such as returns, volatility, Sharpe ratios, and drawdowns, as well as comparisons with benchmark indices. Moreover, turnover statistics provide insight into practical implementability under cardinality and transaction cost constraints. Overall, in-sample backtesting constitutes an essential step toward validating the robustness of the framework before extending the analysis to out-of-sample or live settings.

Overall, these extensions would further improve the economic realism and practical applicability of the optimization framework while maintaining its core convex quadratic structure.

Bibliography

Aroussi, R. (2023). Yfinance: Yahoo finance market data downloader [Accessed: 01 2026].

Day, M.-Y., & Lin, J.-T. (2020). Artificial intelligence for etf market prediction and portfolio optimization. *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 1026–1033. https://doi.org/10.1145/3341161.3344822

Gurobi Optimization, LLC. (2024). *Gurobi optimizer* (Version 10.0) [Accessed: 01 2026]. https://www.gurobi.com

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91. Retrieved January 11, 2026, from http://www.jstor.org/stable/2975974

pandas development team, T. (2020, February). *Pandas-dev/pandas: Pandas* (Version latest). Zenodo. https://doi.org/10.5281/zenodo.3509134

Python Software Foundation. (2023). Requests: Http for humans [Accessed: 01 2026].

Richardson, L. (2023). Beautiful soup documentation [Accessed: 01 2026].