Optimization Based Index Enhancement Method For SnP 500 Stock Selection

Zhehao Zhang
Yiling Wang

Cornell University
SYSEN 6800 Final Project Report

November 25, 2025

Abstract

This project proposes an optimization-based index-enhancement model for the S&P 500 universe. The model follows the structure of practical Chinese index-enhancement frameworks, but adapts it to U.S. market conditions where all index constituents are tradable and daily price limits do not apply. At each rebalancing date, the optimizer takes as inputs: current S&P 500 benchmark weights, a cross-sectional factor (alpha) score, the previous-period portfolio, and sector exposure data. The objective is to maximize factor-implied expected return net of linear transaction costs. To indirectly control tracking error, we impose a full-investment constraint, stock-level active-weight bounds around the benchmark, sector-neutrality constraints with a $\pm 2\%$ tolerance, and a turnover cap of 20%. The resulting formulation is realistic, implementable in standard solvers, and can be later backtested under different parameter settings.

1 Introduction

Index enhancement aims to generate moderate, persistent excess returns over a broad market benchmark while keeping the portfolio close to the index in terms of risk and sector exposures. In this project, we adapt an optimization-based index-enhancement framework that was originally designed for the Chinese A-share market to the U.S. market, specifically to the S&P 500 universe.

Unlike the original setting, the U.S. large-cap universe does not suffer from daily price limits or long suspensions, so the optimization can focus on the classic building blocks: active-weight bounds, sector-neutrality constraints, and turnover control. From a practical standpoint, even small improvements over a broad benchmark matter. Large asset owners and quantitative equity desks often run mandates that are benchmarked to the S&P 500 and are evaluated on information ratio rather than raw return. In such settings, delivering an additional 30–80 bps per year with controlled tracking error can be economically meaningful when the capital base is large. At the same time, naively tilting toward factor signals may create unintended sector bets or excessive turnover, which either increases implementation cost or violates portfolio guidelines. An optimization-based construction is therefore important because it lets us express alpha views while explicitly enforcing the classic real-world constraints (full investment, stock-level active-weight bounds, sector neutrality, turnover control). In other words, the optimizer is the layer that makes a factor idea implementable on the S&P 500 universe. The final model will take as inputs the current index weights, a cross-sectional factor signal, and the previous-period portfolio, and will output a new set of portfolio weights.

2 Background

This project addresses a practical investing task: start from a broad U.S. stock index and aim to earn a bit more than it does, while keeping the portfolio clearly recognizable as the same index. At each scheduled review, we choose how much to hold of each stock for the next period. We face everyday limits that real investors care about: we should not drift far from the index as a whole, we should avoid putting too much into any single company, we should keep the mix across industries similar to the index, and we should keep trading modestly so that costs do not eat into results.

The hard part is uncertainty: no one knows which stocks will lead next month. To cope with that, we use a simple scoring rule to nudge the portfolio slightly toward more promising stocks and away from less promising ones. A computer program searches for a set of small, careful adjustments that obey all of the guardrails above. We accept a proposed portfolio only if it stays close to the index and the expected after-cost outcome improves by a sensible margin. We judge success in plain terms: over time, after costs, does the portfolio earn a little more than the index, while its ups and downs remain very similar to the index's? If so, the approach meets its goal of "index enhancement" without taking on obvious additional risk.

The intended impact is practical and measurable. For investors who must track the market closely (such as retirement plans or index-based funds), even small, steady improvements compound into meaningful gains over the years, without requiring complex or opaque strategies. Because the adjustments are small, transparent, and cost-aware, the method is straightforward to operate, easy to explain to non-specialists, and suitable for real-world constraints like trading costs, oversight, and risk limits. In short, the project offers a disciplined way to seek a modest but reliable improvement while preserving the familiar behavior of a broad market index.

3 Literature Review

The theoretical foundation of modern portfolio optimization originates from the expected return/variance of returns (E–V) rule proposed by Markowitz in 1952[1], which formulates the trade-off between expected return and risk. In the Markowitz framework, investors would tend to choose the efficient portfolio that minimizes variance for the given expected return or maximize expected return for a given level of risk.

Building upon this foundation, Fabozzi, Markowitz, and Gupta (2008)[2] further the mean-variance portfolio selection theory by proposing a systematic way to measure and estimate the key parameters in the original model. In the revised framework, a portfolio's expected return is expressed as the weighted average of the expected returns of individual assets, where each weight reflects the asset's share of the portfolio's total market value. In contrast, the portfolio's overall risk depends not only on the assets' individual volatility but also on the covariance or correlations among them.

Based on these relationships, the portfolio selection problem can be framed as a quadratic programming model with respect to weights $\mathbf{w}$ that explicitly balances expected return and portfolio risk. The general mean–variance optimization problem can be expressed as:

$$\max_{\mathbf{w}} \ \boldsymbol{\mu}^\top \mathbf{w} - \lambda \mathbf{w}^\top \Sigma \mathbf{w}, \quad \text{s.t.} \quad \sum_i w_i = 1, \text{for } i = 1, \ldots, n \tag{1}$$

where $\boldsymbol{\mu}$ represents the vector of expected returns, Σ denotes the covariance matrix of asset returns, and λ is the investor's risk-aversion parameter. This quadratic objective function models the trade-off between expected return and risk, and its solution yields the set of efficient portfolios.

However, implementing this model in practice involves several challenges. First, constructing the covariance matrix requires estimating $\frac{n(n+1)}{2}$ covariance terms σ_{ij} from historical data. For large asset universes, this results in a dense covariance matrix in which most entries are nonzero, making the associated quadratic program computationally demanding to solve[3]. Although factor models and sparse-matrix techniques[4] can ease this burden to some extent, obtaining an exact optimal solution remains difficult.

Second, the traditional formulation relies primarily on a single market factor to explain asset returns, overlooking the contribution of other systematic factors that drive the portfolio performance. Moreover, the classical mean-variance model does not account for transaction costs, turnover, or benchmark-relative constraints, which limits its direct applicability in real-world financial markets.

These limitations have boosted the development of multi-factor frameworks that take a broader set of systematic return drivers into consideration. A major advancement addressing this limitation is the multi-factor asset-pricing framework introduced by Fama and French[5], which demonstrates that cross-sectional stock returns are related not only to market exposure but also to size and value factors. The Fama–French model provides a foundation for modern multi-factor investing and motivates the use of factor-based signals for creating expected-return estimations.

In parallel, Grinold and Kahn's fundamental law of active management [6] provides theoretical justification for factor-based index enhancement. Their concept formalizes the relationship between forecast accuracy, opportunity breadth, and the achievable information ratio and highlights that active returns are produced by matching portfolio exposures with predictive signals while eliminating unwanted bets. This perspective closely relates to our optimization model, which translates a cross-sectional factor score into benchmark-relative tilts under practical constraints.

Motivated by these limitations and theoretical developments, recent industry research has focused on developing alternative linear formulations that incorporate multi-factor models to improve scalability and practicality. The quantitative research team at BOC International [7] developed an index-enhanced optimizer for the Chinese A-share market. This linear optimizer accounts for transaction costs with constraints on stock-level weights, portfolio factor exposure, and turnover. Empirical backtesting on the CSI-300 and CSI-500 indices demonstrates that the optimizer effectively controls tracking error and yields stable excess returns under appropriate parameter configurations.

Building on this framework, our project adapts and extends the BOCI model to the U.S. market by tailoring its structure to the S&P 500 universe and evaluating how constraint tightness and transaction-cost assumptions influence the portfolio performance.

4 Problem Statement

4.1 Given

- N: number of constituent stocks in the S&P 500 at the rebalancing date.
- $\mathbf{b} \in \mathbb{R}^N$: benchmark weight vector of the S&P 500.
- $\mathbf{f} \in \mathbb{R}^N$: cross-sectional factor (alpha) scores, standardized each period.
- $\mathbf{w}^{\text{prev}} \in \mathbb{R}^N$: portfolio weights before rebalancing.
- $S \in \mathbb{R}^{N\times 11}$: sector exposure matrix (GICS 11 sectors).
- $c = 0.0005$: per-unit transaction-cost coefficient (5 bps).
- $\delta = 0.01$: stock-level active-weight tolerance.
- $\tau = 0.20$: turnover upper bound.
- $\gamma^{\text{sector}} = 0.02$: sector exposure tolerance.
- $\alpha = 0.0004$: scaling from factor score to expected return.

4.2 Determine

A post-trade portfolio weight vector $\mathbf{w} \in \mathbb{R}^N$.

4.3 Objective

Define period expected return as $\boldsymbol{\mu} = \alpha\mathbf{f}$. We maximize expected return net of transaction cost:

$$\max_{\mathbf{w}} \quad \boldsymbol{\mu}^\top \mathbf{w} \;-\; c\sum_{i=1}^{N} |w_i - w_i^{\text{prev}}| \,. \tag{2}$$

4.4 Constraints

Full investment

$$\sum_{i=1}^{N} w_i = 1. \tag{3}$$

Stock-level active-weight bounds

$$\max(0,\, b_i - \delta) \le w_i \le \min(0.10,\, b_i + \delta), \quad i = 1, \ldots, N. \tag{4}$$

Sector-exposure neutrality Let $\mathbf{s}^{\text{bmk}} = S^\top \mathbf{b}$ be the benchmark sector weights. Then

$$\left| S^\top \mathbf{w} - \mathbf{s}^{\text{bmk}} \right| \le \gamma^{\text{sector}} \mathbf{1}_{11}. \tag{5}$$

Turnover constraint

$$\sum_{i=1}^{N} |w_i - w_i^{\text{prev}}| \le \tau. \tag{6}$$

5 Model Formulation and Solution Algorithm

This section specifies the optimization model used by our index enhancement strategy and the solution workflow used at each rebalance. The goal is to stay close to the benchmark while making small, cost-aware tilts toward higher scored names.

5.1 Nomenclature

Sets and indices

- $i \in \mathcal{N} = \{1, \ldots, N\}$: stocks in the investable universe.
- $j \in \mathcal{J} = \{1, \ldots, 11\}$: GICS sectors.

Parameters

- $b_i \in [0,1]$: benchmark weight of stock i (S&P 500).
- $\mu_i \in \mathbb{R}$: cross-sectional score for stock i (higher is better).
- $c \geq 0$: per-unit transaction-cost penalty in the objective.
- $\tau \geq 0$: turnover cap, i.e., allowed ℓ_1 deviation from previous weights.
- $\delta \geq 0$: per-name deviation band around the benchmark ($|w_i - b_i| \leq \delta$).
- $\bar{w} \in (0,1]$: single-name weight cap (we use $\bar{w} = 10\%$).
- $\gamma^{\text{sector}} \geq 0$: sector-exposure deviation band.
- $S \in \{0,1\}^{N \times 11}$: sector-membership matrix; column S_j is sector j.
- $\mathbf{w}^{\text{prev}} \in [0,1]^N$: portfolio weights before the rebalance.
- $\mathbf{1}$: vector of ones of conformable dimension.

Decision variables

- $w_i \geq 0$: new portfolio weight of stock i.
- $t_i \geq 0$: auxiliary turnover variable enforcing $t_i = |w_i - w_i^{\text{prev}}|$.

5.2 Mathematical Formulation

We maximize a simple linear proxy for expected improvement (score-tilt) net of linear trading penalties, subject to tight benchmark-tracking constraints:

$$
\begin{aligned}
\max_{\mathbf{w},\mathbf{t}} \quad & \sum_{i=1}^{N} \mu_i w_i \;-\; c\sum_{i=1}^{N} t_i && (7)\\
\text{s.t.} \quad & \sum_{i=1}^{N} w_i = 1 \qquad \text{(fully invested)}, && (8)\\
& 0 \le w_i \le \bar{w}, \qquad i = 1,\dots,N \qquad \text{(single-name cap)}, && (9)\\
& -\delta \le w_i - b_i \le \delta, \qquad i = 1,\dots,N \qquad \text{(band around benchmark)}, && (10)\\
& -\gamma^{\text{sector}}\mathbf{1}_{11} \le S^\top \mathbf{w} - S^\top \mathbf{b} \le \gamma^{\text{sector}}\mathbf{1}_{11} \qquad \text{(sector neutrality)}, && (11)\\
& \sum_{i=1}^{N} t_i \le \tau \qquad \text{(turnover cap)}, && (12)\\
& t_i \ge w_i - w_i^{\text{prev}}, \quad t_i \ge -(w_i - w_i^{\text{prev}}), \qquad i = 1,\dots,N \qquad \text{(turnover definition)}, && (13)\\
& w_i \ge 0,\; t_i \ge 0, \qquad i = 1,\dots,N.
\end{aligned}
$$

Discussion. Constraint (10) limits name-level drift; (11) keeps sector tilts small (measured as the difference between portfolio and benchmark weights per sector); (12) guards against excessive trading; (13) linearizes the absolute turnover. With these choices the model is a pure linear program (LP).

Linearization and Solver-Readiness (Pyomo & GAMS)

The formulation contains absolute-value terms $|w_i - w_i^{\text{prev}}|$. Standard LP solvers do not accept absolute values directly, but they are linearized by introducing $t_i \ge 0$ and enforcing (13). The original turnover cap $\sum_i |w_i - w_i^{\text{prev}}| \le \tau$ becomes $\sum_i t_i \le \tau$, and the objective $\max \boldsymbol{\mu}^\top \mathbf{w} - c\sum_i |w_i - w_i^{\text{prev}}|$ becomes (7): $\max \boldsymbol{\mu}^\top \mathbf{w} - c\sum_i t_i$. The sector-exposure constraints are already linear: $-\gamma^{\text{sector}}\mathbf{1} \le S^\top(\mathbf{w} - \mathbf{b}) \le \gamma^{\text{sector}}\mathbf{1}$.

Pyomo implementation. Declare `w[i] >= 0` and `t[i] >= 0`. Add (13) using a `ConstraintList`, encode (8)–(11) directly, and build the objective as `sum(mu[i]*w[i]) - c*sum(t[i])`.

The model solves with any LP solver, e.g. `SolverFactory("gurobi")`, `SolverFactory("appsi_hig` or `SolverFactory("cbc")`.

GAMS implementation. Use two equations per asset: `t(i) =g= w(i) - wprev(i);` and `t(i) =g= -(w(i) - wprev(i));`. The objective is `sum(i, mu(i)*w(i)) - c*sum(i, t(i` All constraints are linear, so standard LP engines (CPLEX, Gurobi, XPRESS) apply without modification.

5.3 Solution Algorithm

We implement a monthly, trade-day aligned backtest that repeatedly solves the LP at each rebalance date and applies the resulting weights until the next date (similar to the concept of recursion). The procedure below follows the code (`rebalance_days`, `holdto_days`, `solve_lp`, arrays for weights, returns, and turnover).

Inputs.

1. Trading calendar with monthly endpoints: $(d, d_{\text{next}}) \in \{\texttt{rebalance\_days}\} \times \{\texttt{holdto\_days}\}$.
2. Price panel `close[date, ticker]` for the universe `universe`.
3. Benchmark weights $\mathbf{b}$ on each rebalance date.
4. Sector matrix S (one-hot per stock across 11 sectors).
5. Score panel `mu_panel[date, ticker]` giving μ_i.
6. Hyperparameters $(\delta, \tau, \gamma^{\text{sector}}, c)$ and solver name (default `"appsi_highs"`).

Initialization.

1. Set previous weights to the benchmark: $\mathbf{w}^{\text{prev}} \leftarrow \mathbf{b}$ (code: `w_prev = b.copy()`).
2. Create collectors: `weights_rows`, `port_rets`, `turnovers`.

Loop over months. For each pair (d, d_{next}):

1. **Build the signal.** Extract $\boldsymbol{\mu}_d$ for `universe` and fill missing scores with zero for feasibility: `mu = mu_panel.loc[d, universe].fillna(0.0).to_numpy()`.
2. **Solve the LP.** Call

```
w, t, status = solve_lp(mu, b, S, w_prev,
                        delta=0.01, tau=0.20,
                        gamma_sector=0.02, c=0.0005,
                        solver_name="appsi_highs")
```

   which constructs (7)–(13) in Pyomo, adds the name cap ($\bar{w} = 10\%$), and solves. The function returns the optimal weights $\mathbf{w}$, the auxiliary variables $\mathbf{t}$ (turnover by name), and a solver status flag.
3. **Numerical guards.** Clip tiny negatives to zero and renormalize so that $\sum_i w_i = 1$; ensure sector and band constraints are satisfied within solver tolerance.
4. **Account for realized turnover.** $T_d \leftarrow \sum_i |w_i - w_i^{\text{prev}}|$; append to `turnovers`.
5. **Hold to d_{next}.** Compute next-period simple returns with a price mask:

```
r_next_vec = close.loc[d_next, universe] / close.loc[d, universe] - 1.0
```

   If a price is missing at either date, exclude that name from the dot product (masking) so the realized return is computed on available assets only.
6. **Realized portfolio return.** $r^{\text{port}}_{d\to d_{\text{next}}} \leftarrow \sum_i w_i \, r_{i,d\to d_{\text{next}}}$; append to `port_rets`.
7. **Roll forward.** Save the weight row for reporting (`weights_rows.append(...)`), then set $\mathbf{w}^{\text{prev}} \leftarrow \mathbf{w}$.

Post-processing and evaluation.

1. Concatenate `weights_rows` into a DataFrame indexed by rebalance dates.
2. From `port_rets` compute cumulative return, annualized return/volatility, Sharpe (with a zero cash rate), information ratio vs. the benchmark, and tracking error (using the benchmark's realized path).
3. Summarize turnover statistics (mean, median, 95th percentile) and average dollar trading cost using $c \cdot T_d$ as the per-period penalty used in the optimization.

Hyperparameter tuning. We run a small grid over $(\delta, \gamma^{\text{sector}}, \tau, c)$ and repeat the monthly loop for each setting. The grid focuses on four economically meaningful hyperparameters that directly appear in the constraints:

- Active-weight tolerance $\delta \in \{0.005,\ 0.010,\ 0.020\}$,
- Turnover cap $\tau \in \{0.10,\ 0.20,\ 0.30\}$,
- Sector-exposure tolerance $\gamma^{\text{sector}} \in \{0.01,\ 0.02\}$,
- Transaction-cost coefficient $c \in \{0.0003,\ 0.0005\}$.

This yields $3 \times 3 \times 2 \times 2 = 36$ configurations, which is computationally tractable for a linear program.

For every grid point, we record the after-cost information ratio and tracking error. The final choice targets a stable IR improvement with controlled drift (small sector/name deviations) and moderate realized turnover. The code enables easy substitution of the LP backend via the `solver_name` argument (`"appsi_highs"` by default; `"gurobi"` if available).

Complexity and runtime. Each monthly optimization is a medium-scale LP with $O(N)$ variables ($\mathbf{w}$ and $\mathbf{t}$) and $O(N + J)$ linear constraints; HiGHS solves these instances in sub-seconds to seconds on our machine. The backtest therefore scales linearly with the number of rebalance periods and grid size.

Reproducibility notes (as in code). Scores are `fillna(0.0)` at the rebalance to avoid infeasibility; realized returns exclude symbols lacking either d or d_{next} prices; starting weights are set to the benchmark at the first date; all weights are renormalized after minor numerical adjustments.

6 Results and Discussion

This section presents empirical results for the index-enhancement strategy implemented using the linear optimization model described earlier. Using daily data of stocks from **January 2, 2020 to December 31, 2024**, we conduct a monthly backtest inside a structured grid search over the major hyperparameters governing active-weight bounds δ, turnover limits τ, sector neutrality γ^{sector}, and transaction-cost penalties c. The purpose of this grid search is to determine how different choices of parameters affect the strategy's ability to convert factor information into excess returns, and ultimately to identify the parameter combinations that maximize the index-enhancement performance while maintaining risk control and trading stability.

For each parameter tuple, we compute evaluation metrics including annualized information ratio, monthly tracking error, turnover, sector deviations, and the number of binding active-weight constraints. Since index-enhancement portfolios are designed to remain close to the benchmark while generating modest excess returns, the information ratio is our primary evaluation metric. It captures risk-adjusted active performance and is defined as

$$\mathrm{IR} = \frac{\text{annualized active return}}{\text{annualized tracking error}},$$

where active return $=$ portfolio return - benchmark return, and the tracking error is calculated as the standard deviation of the time series of active returns.

6.1 Optimal parameter configuration

Among all 36 configurations in the grid, the parameter set that achieves the highest annualized information ratio is

$$(\delta, \tau, \gamma^{\text{sector}}, c) = (0.020,\ 0.300,\ 0.010,\ 0.0003).$$

This configuration generates an annualized information ratio of approximately 1.21, with an annualized return of 23.4% and volatility of 19.3% over the sample period. However, this configuration also exhibits relatively high average monthly turnover (17.7%) and a large number of binding active-weight constraints (approximates 445), suggesting that its superior performance is achieved by taking on substantial flexibility relative to the benchmark. Although it maximizes the information ratio, such a configuration may be less suitable in settings where turnover stability or benchmark resemblance are significant operational considerations.

Therefore, to identify a more balanced parameter set, we impose empirical bounds based on the distribution of all grid outcomes: the annualized information ratio must lie above the grid's 75th percentile, while both the average monthly turnover and average maximum sector deviation must fall below their respective 75th percentiles. Applying these limitations yields the configuration

$$(\delta, \tau, \gamma^{\text{sector}}, c) = (0.020,\ 0.100,\ 0.010,\ 0.0003),$$

which achieves an annualized information ratio of 1.15 with a substantially lower average turnover of 9.9%, an annualized return of 23.6%, and a maximum sector deviation of only 1%. In addition, this configuration exhibits a moderate number of binding active-weight constraints (on average 326 bounds), indicating that it maintains benchmark discipline without being overly constrained.

Notably, this configuration also yields the highest annualized return among all grid points, indicating that it strikes a better balance between risk-adjusted performance, trading cost efficiency, and benchmark tracking discipline than the unconstrained configuration.

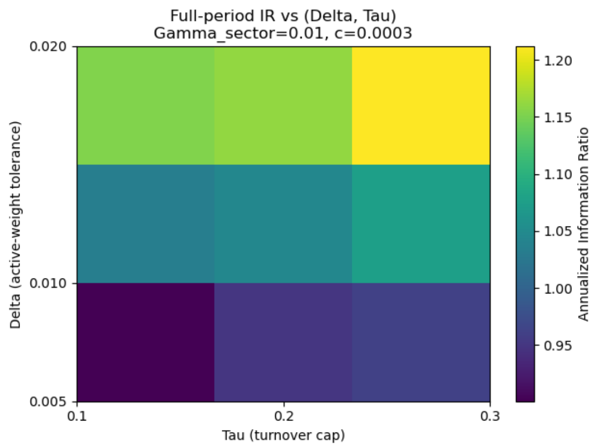
6.2 Visualization of Grid Search Results

After determining the optimal parameter configuration, we use a series of visual diagnostics to help explain how performance differs throughout the entire grid. These visualizations illustrate the interaction between model flexibility, turnover costs, and benchmark drift, and they provide an intuitive understanding of why certain parameter regions outperform others.

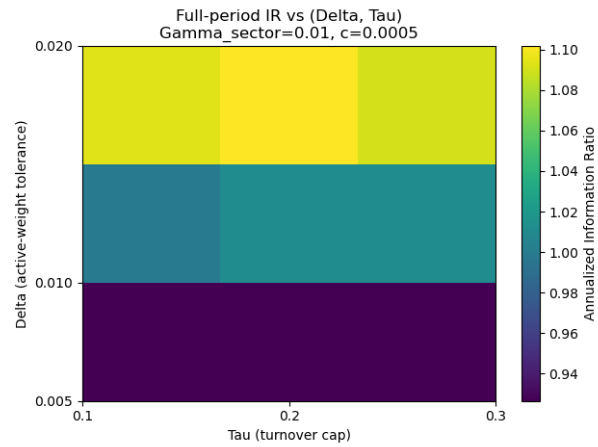
6.2.1 Heatmaps of Information Ratio Across (δ, τ)

Since the information ratio is our primary performance metric, we visualize its behavior across the (δ, τ) grid using heatmaps to examine how different levels of model flexibility influence risk-adjusted active returns. The sector-exposure tolerance γ^{sector} controls benchmark drift and the transaction cost c governs trading intensity, both of which define the broader risk and cost environment of the optimizer. Moreover, since the optimal parameter configuration identified earlier lies in the regime $\gamma^{\text{sector}} = 0.01$ and $c = 0.0003$, their corresponding heatmap is informative. To gain a complete understanding of how the information ratio surface responds across different risk–cost environments, we include heatmaps for all combinations of $\gamma^{\text{sector}} \in \{0.01, 0.02\}$ and $c \in \{0.0003, 0.0005\}$, shown in Figure 1.

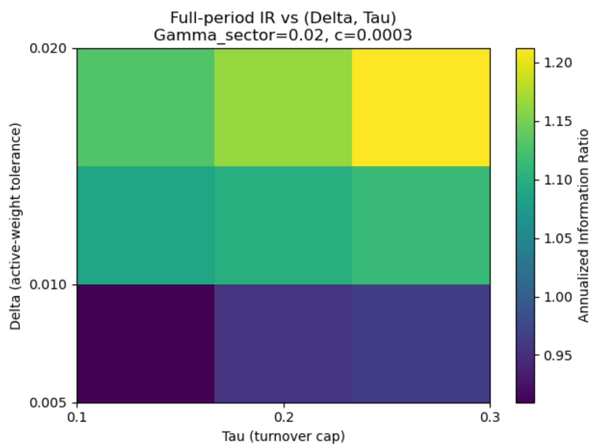
These plots provide a clear view of how active-weight tolerance and turnover caps interact: the highest information ratios mainly arise in the upper-right region of the grid, where both the active-weight tolerance δ and turnover cap τ take moderate-to-large values. This suggests that when constraints are overly tight, the optimizer cannot meaningfully tilt away from benchmark weight, whereas excessively loose constraints lead to frequent rebalancing and greater exposure drift. The heatmaps therefore highlight the parameter region that achieves a favorable balance between factor-expression flexibility and effective control of trading intensity and benchmark deviation.



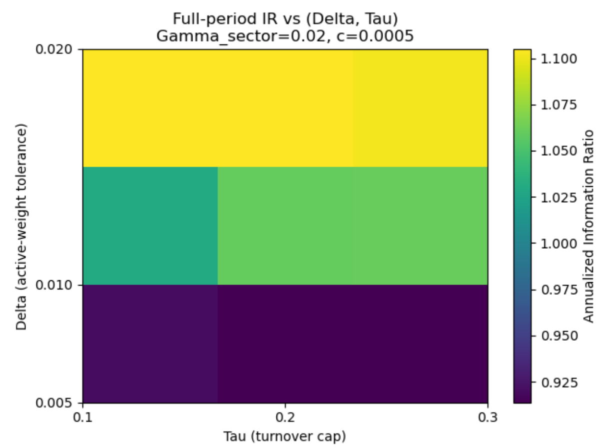
(a) $\gamma^{\text{sector}} = 0.01,\ c = 0.0003$



(b) $\gamma^{\text{sector}} = 0.01,\ c = 0.0005$



(c) $\gamma^{\text{sector}} = 0.02,\ c = 0.0003$



(d) $\gamma^{\text{sector}} = 0.02,\ c = 0.0005$

Figure 1: Heatmaps of annualized information ratio across (δ, τ) for all combinations of sector-neutrality tolerance $\gamma^{\text{sector}} \in \{0.01, 0.02\}$ and cost coefficient $c \in \{0.0003, 0.0005\}$.

6.2.2 Comparison Plots Across Parameter Dimensions

To further analyze how performance changes with specific parameter choices, we generate comparison plots that fix two hyperparameters while varying a third. Although the best-performing configuration identified above corresponds to $\gamma^{\text{sector}} = 0.01$ and $c = 0.0003$, we illustrate the turnover and active-weight effects using the setting $\gamma^{\text{sector}} = 0.02$ and $c = 0.0005$ to highlight the general patterns observed across the grid.

As shown in Figure 2, the information ratio exhibits different sensitivities to the turnover cap depending on the active-weight tolerance δ. When $\delta = 0.005$, the information ratio remains below

0.925 and initially declines as τ increases before rising again at higher turnover levels, indicating that excessively tight name-level constraints restrict the optimizer's ability to benefit from additional rebalancing. For $\delta = 0.010$, the information ratio increases with τ and then stabilizes, suggesting that a moderate level of flexibility is sufficient for the optimizer to express most of the factor signal. When $\delta = 0.020$, the information ratio is the highest across all settings and remains largely stable for small-to-moderate turnover caps before eventually decreasing at high τ, reflecting that overly frequent rebalancing becomes counterproductive once transaction costs dominate.

Similarly, Figure 3 plots the information ratio against the active-weight tolerance δ for fixed turnover caps and transaction cost coefficient. Across all values of τ, the information ratio rises as δ increases, with a steeper improvement when moving from very tight to moderately flexible active-weight bounds and a smaller marginal gain as δ becomes larger. Moreover, the curves for $\tau = 0.20$ and $\tau = 0.30$ are nearly indistinguishable at first, indicating that once the turnover cap is sufficiently permissive, further increases in τ have little additional effect on performance.

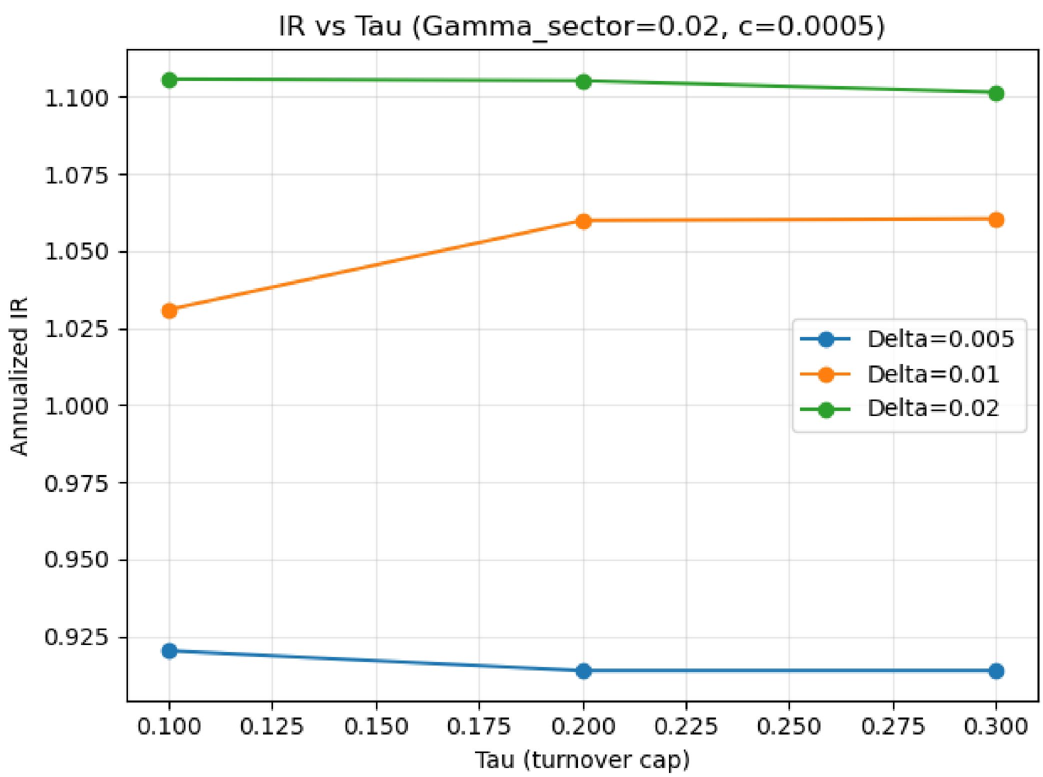


Figure 2: Information ratio as a function of turnover cap τ for fixed $\gamma^{\text{sector}} = 0.02$ and $c = 0.0005$.

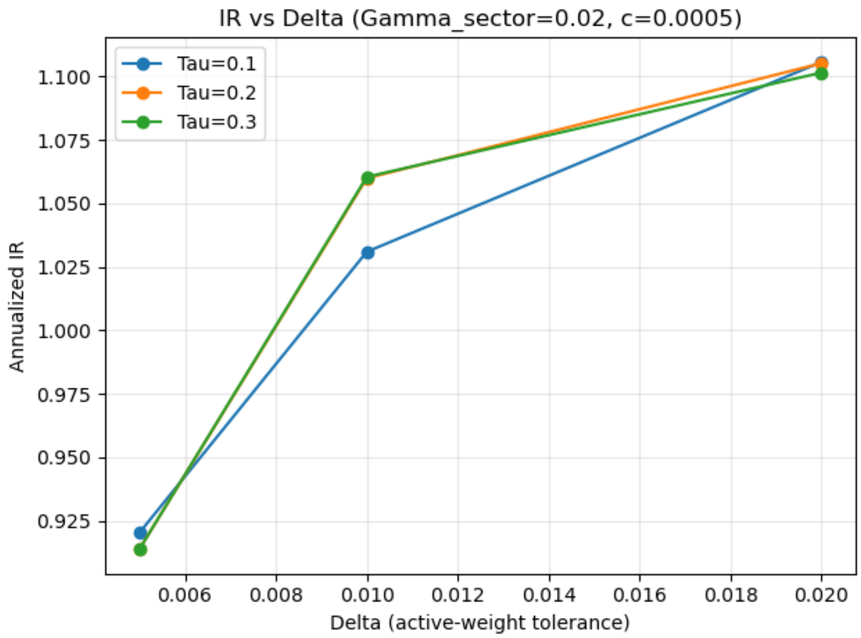


Figure 3: Information ratio as a function of active-weight tolerance δ for fixed $\gamma^{\text{sector}} = 0.02$ and $c = 0.0005$.

Finally, we examine the relationship between turnover and performance. Figure 4 plots the annualized information ratio against the average monthly turnover across all grid configurations. The scatter shows that portfolios with turnover around 10% per month already achieve reasonably strong information ratios, indicating that a moderate amount of rebalancing is sufficient for the optimizer to capture most of the available factor signal. The information ratio improves further only when turnover rises to around 18% per month, reflecting that stronger performance is obtained at the cost of greater trading activity. This pattern reinforces the turnover bounds used in our selection procedure: moderate turnover yields robust performance, whereas extremely high turnover offers better performance gains at the expense of higher trading intensity.

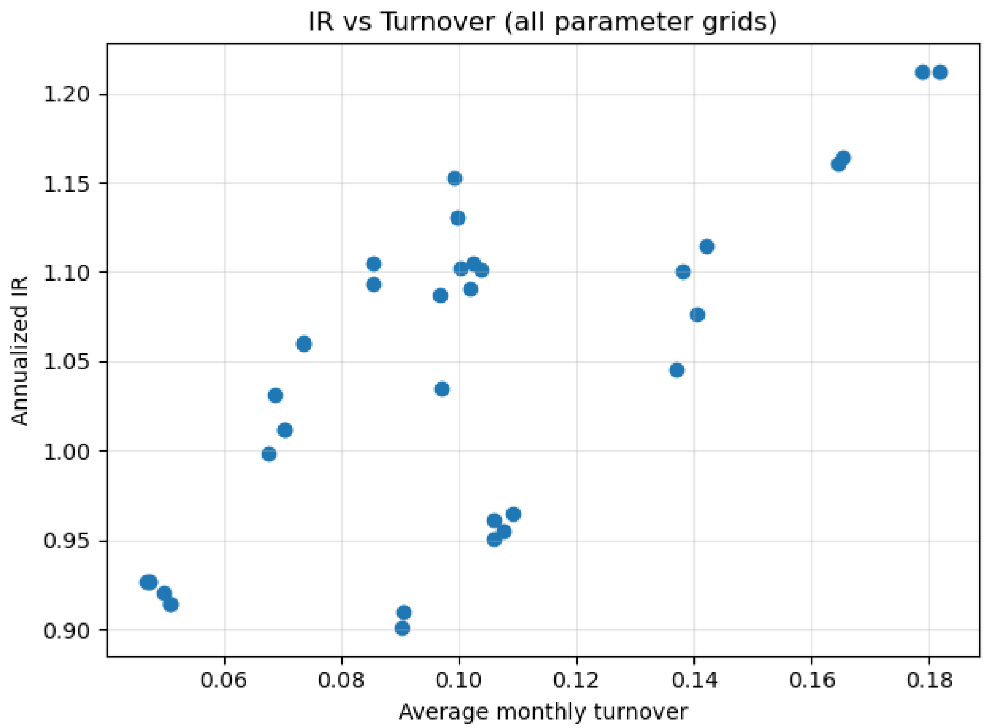


Figure 4: Information ratio versus average monthly turnover across all grid configurations.

6.3 Cumulative Return Comparison Across Parameter Configurations

From the investors' perspective, the cumulative performance of their portfolios is a key consideration when evaluating a strategy. Thus, we compare the realized cumulative returns of the best- and worst-performing parameter configurations from the grid search. These plots provide an intuitive view of how different levels of constraint tightness and trading flexibility translate into portfolio outcomes relative to the S&P 500 benchmark.

Figure 5 displays the cumulative returns of the top three configurations relative to the S&P 500 benchmark. Prior to 2024, the enhanced portfolios moved very closely with the benchmark and at times even fell slightly behind it. Beginning in early 2024, however, the cumulative returns of the enhanced portfolios begin to exceed the benchmark more significantly, although their overall trajectory continues to track the benchmark's broad market movements. This pattern indicates that while the model maintains tight tracking behavior, it is able to convert factor information into meaningful excess performance when market conditions become more favorable.

Several features of the U.S. equity market during 2024-2025 likely contributed to this outcome.

Market leadership became increasingly concentrated among a small group of large technology and AI-related firms, creating wider cross-sectional dispersion in stock returns. At the same time, trend persistence strengthened and idiosyncratic volatility declined, raising the efficiency of the momentum factor used by the optimizer.

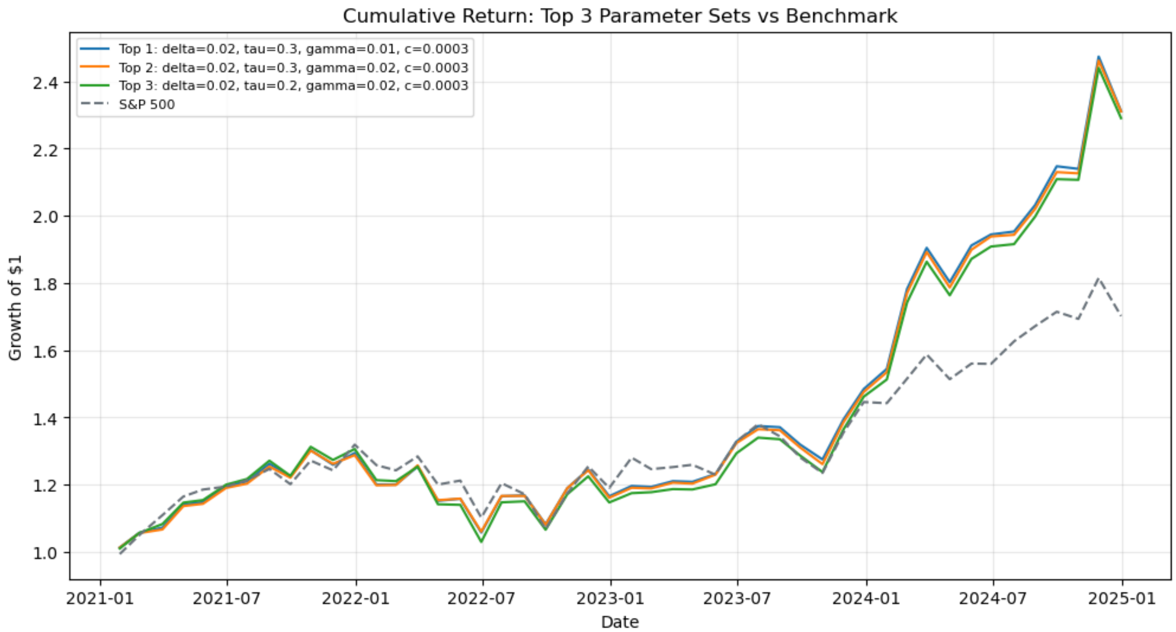


Figure 5: Cumulative returns of the top-three configurations vs. benchmark.

In contrast, Figure 6 presents the cumulative returns of the bottom three IR configurations. Similar to the top models, these portfolios begin to generate noticeable excess returns over the benchmark after January 2024. However, the magnitude of their outperformance is much smaller than that of the top configurations. Because these parameter settings contain tighter constraints or insufficient flexibility, the optimizer is less able to express the factor signal effectively. As a result, the portfolios either track the benchmark closely throughout much of the sample or deliver only modest outperformance even during periods when the factor information becomes more valuable.

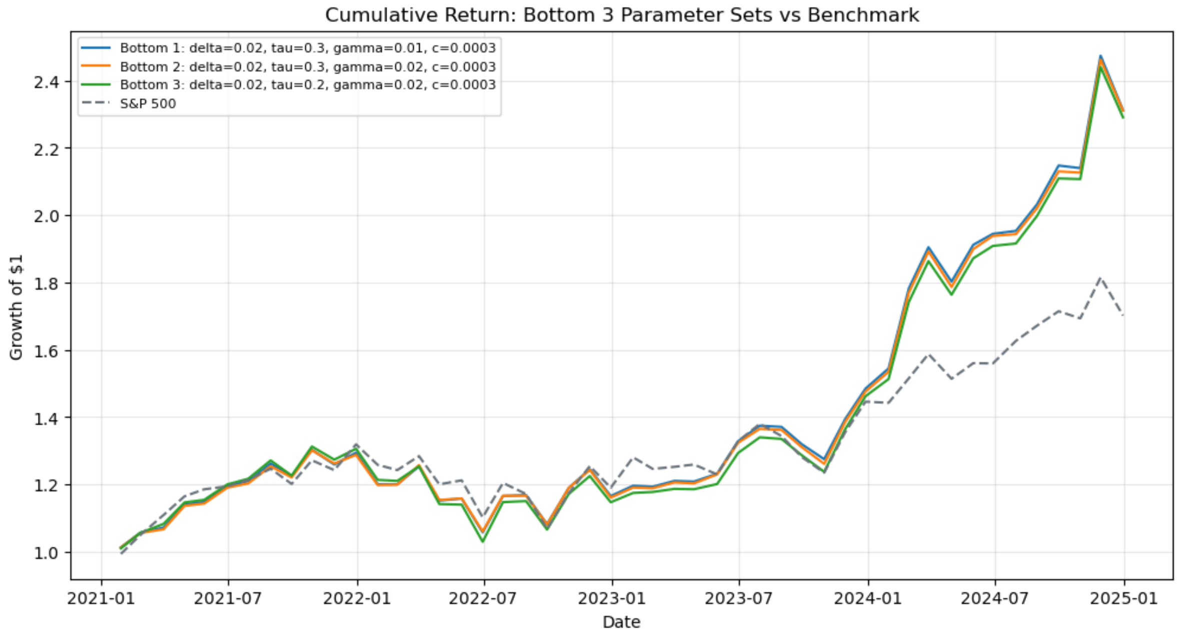


Figure 6: Cumulative returns of the bottom-three configurations vs. benchmark.

These two plots illustrate that the appropriate choice of hyperparameters enables stable index enhancement. Moreover, the results demonstrate that, despite being originally developed for the Chinese equity market, the adjusted model can still generate meaningful excess returns in the U.S. market when selecting suitable parameters, reflecting the robustness of the underlying optimization framework.

7 Conclusion and Recommendations

This project develops and evaluates an optimization-based index-enhancement framework for the S&P 500 universe, adapting an approach originally designed for the Chinese A-share market to U.S. equity conditions. The model combines a linear expected-return objective with transaction costs and a series of practical constraints, including full investment, stock-level active-weight bands, sector neutrality, and turnover control. Through a comprehensive grid search, we investigate how key hyperparameters influence portfolio performance and identify a robust parameter set where the strategy consistently generates meaningful excess returns.

The empirical results suggest that appropriate choices of active-weight tolerance, turnover caps, and sector deviation limits enable stable and persistent index enhancement. The top-performing configurations achieve information ratios above 1.10 and annualized returns exceeding 22.5%, while still tracking the benchmark closely. The enhanced portfolio began to generate noticeable excess returns after early 2024, probably caused by stronger trend persistence and reduced idiosyncratic noise, which increased the effectiveness of cross-sectional factor signals. These findings demonstrate that this optimization framework is robust and transferable across markets, providing a disciplined way to exploit predictable structure in equity returns.

Recommendations and Future research

Given that the current level of excess return remains relatively modest, several extensions could further strengthen the portfolio's performance and broaden the model's practical applicability:

- **Incorporate multiple factors.** The current model relies on a single cross-sectional score. Integrating additional signals, such as quality, value, or low-risk, is likely to enhance robustness and reduce dependence on a single style factor.
- **Dynamic transaction costs.** Real-world trading costs vary with market depth, volatility, and order size. For further work, we can replace the fixed linear cost with a piecewise-linear or liquidity-sensitive model, allowing the optimizer to more accurately penalize high-impact trades.
- **Evaluate alternative rebalance frequencies.** Monthly rebalancing works well in our tests, but biweekly or quarterly schedules may reduce turnover while maintaining performance, especially in high-cost or capacity-constrained settings.
- **Extend analysis to a longer time period.** Our backtest covers a five-year lookback window, which provides only a limited view of the model's behavior across market conditions. Extending the evaluation over a longer historical period or conducting an out-of-sample test would offer a more robust assessment of the model's stability across diverse market regimes.
- **Incorporate optimizer diagnostics or regularization.** Techniques such as shrinkage, ridge penalties, or robust optimization can help stabilize the portfolio weights by reducing sensitivity to noise in the factor signal or to stocks that cluster near their constraint boundaries. These methods effectively smooth the optimization landscape, preventing highly concentrated solutions and improving the model's robustness in environments with volatile or weak predictive signals.

Limitations

- **Constituent survivorship.** The backtest implicitly assumes a fixed S&P 500 membership. Using point-in-time constituent sets (adds/drops) would reduce survivorship bias and better reflect investability.
- **Simplified trading costs.** A single linear cost cannot capture liquidity tiers, volatility spikes, or market-impact nonlinearity. Results may be optimistic for small-cap tails or fast markets.
- **Capacity and frictions.** Turnover and sector caps control risk, but we do not model borrowing constraints, locate fees, taxes, or internal crossing. Capacity limits could bind for large AUM.
- **Signal fragility.** A one-factor specification increases regime sensitivity. Weak or crowded signals can push solutions to constraint boundaries and raise turnover.

Operational considerations for deployment

- **Data hygiene.** Use point-in-time constituents and free-float/DR share lines; reconcile corporate actions; forward-fill only for trading holidays, never across suspensions.
- **Rebalance mechanics.** Freeze weights at close t, trade at open $t{+}1$ or VWAP window; enforce a minimum trade size to avoid dust orders; defer trades failing liquidity checks.
- **Solver robustness.** Warm-start with w_{t-1}; add small ridge/shrinkage on active weights if solutions repeatedly hit bounds; set explicit failsafes when the LP is infeasible (fallback to benchmark).
- **Risk monitoring.** Track realized TE, sector/industry drifts, top name concentrations, and turnover budget usage; alert on rule breaches and trigger claw-back trades if limits are exceeded.
- **Change management.** Log parameter versions ($\delta,\ \tau,\ \gamma_{\text{sector}},\ c,\ \alpha$), data hashes, and solver versions to ensure reproducibility and auditability.

Overall, this project demonstrates that an appropriately constructed linear optimization framework can deliver meaningful index enhancement in the U.S. equity market, even when adapted from a setting with different institutional features. With further refinements, the approach may serve as a practical component in quantitative equity strategies aimed at improving benchmark returns under realistic constraints.

References

[1] H. Markowitz, “Modern portfolio theory,” *Journal of Finance*, vol. 7, no. 11, pp. 77–91, 1952.

[2] F. J. Fabozzi, H. M. Markowitz, and F. Gupta, “Portfolio selection,” *Handbook of finance*, vol. 2, pp. 3–13, 2008.

[3] H. Konno and H. Yamazaki, “Mean-absolute deviation portfolio optimization model and its applications to tokyo stock market,” *Management science*, vol. 37, no. 5, pp. 519–531, 1991.

[4] A. F. Perold, “Large-scale portfolio optimization,” *Management science*, vol. 30, no. 10, pp. 1143–1160, 1984.

[5] E. F. Fama and K. R. French, “The cross-section of expected stock returns,” *the Journal of Finance*, vol. 47, no. 2, pp. 427–465, 1992.

[6] R. C. Grinold and R. N. Kahn, “Active portfolio management,” 2000.

[7] Guo, Ce and Song, Kundi, “Index-enhancement portfolio optimizer: A comprehensive framework built from the ground up,” BOC International Securities Co., Ltd., Tech. Rep., March 2024, Quantitative Multi-Factor Stock Selection Series (VII).