Quantitative Approaches to Portfolio Optimisation

Summer Internship Project

Intern:
Shruti Sharma
MSc Data Science, CMI

Supervisor:
Shrish Trivedi

11 August, 2025

PROBLEM & OBJECTIVE

My Project's Objective: To systematically evaluate how different quantitative methods for estimating returns and risk impact portfolio performance.

The Problem: Traditional investing is often riddled with emotional bias, and struggles to scale effectively in the face of vast dataset.

- **Emotional bias** – investors make decisions driven by fear, greed, or herd mentality.
- **Scalability issues** – hard to process and interpret massive datasets efficiently.

In today’s markets, speed and accuracy in decision-making are critical, but human-driven processes struggle to keep up.

THE SOLUTION

The Solution: Quantitative investing

What is quantitative investing?
A data-driven, rule-based approach that uses that uses statistical and algorithmic models to guide investment decisions.

Why it works?

- Eliminates emotion from decisions.
- Scales efficiently — can analyse thousands of securities in seconds.
- Real-time capability — analyses live market data while incorporating historical trends.
- Uncovers hidden patterns — identifies statistical relationships missed by traditional methods.

DATA DESCRIPTION

Source:

- Yahoo Finance API via yfinance (free, reliable, well-suited for academic research)

Time Period:

- 20+ years (January 2003– January 2025).
- This window captures multiple market cycles — including the Global Financial Crisis (2008), COVID-19 crash and recovery (2020), Russia–Ukraine war (2022), and several general elections — making it ideal for evaluating long-term investment strategies across diverse market regimes.

Stock Universe:

- NIFTY 500 constituents: Large, liquid, representative sample of Indian equity market
- Filtered to stocks with ≥15 years continuous historical data
- Final dataset: ~300 stocks

FINAL DATA

Handling Missing Data:

- Forward-fill (ffill) to maintain continuity without artificial volatility.

Resampling:

- Weekly frequency (W-WED) to avoid start/end-of-week bias and capture typical trading patterns.

Return Calculation:

- Simple returns chosen over log returns.
- Weekly returns strike a balance between noise reduction and sample size.

Dataset Scope:

- Rolling 10-year training window (520 weeks) for each rebalance date.
- Backtesting period: 2014–2025.

Ready for: Return estimation, Covariance estimation, and Portfolio optimization

EXPECTED RETURNS ESTIMATION

In portfolio optimisation, the allocation of weights across assets depends greatly on our beliefs about future returns. A portfolio optimiser uses these return estimates to tilt the portfolio toward assets that are expected to outperform. But, returns are notoriously hard to estimate. A small error in estimating returns can lead to large and unstable portfolio weights . This is why robust and well-thought-out return models are essential. The quality of return forecasts can make or break portfolio performance.

Approaches tested in this study:

- Mean Historical Return
- Exponentially Weighted Mean (EWMA)
- Capital Asset Pricing Model (CAPM)

MEAN HISTORICAL RETURN

- **Definition:** Geometric average of past returns (equal weight to all).
- **Formula:**

$$\mu = \left(\prod_{t=1}^{T} (1 + r_t) \right)^{\frac{1}{T}} - 1$$

- **Strengths:**
  - Simple, easy to interpret (often used as a starting point in portfolio analysis).
  - Reflects long-term compounding.
- **Weaknesses:**
  - Assumes past performance will continue in the future.
  - Doesn’t account for market related explanatory factors.

EXPONENTIALLY WEIGHTED HISTORICAL MEAN

- **Definition:** Weighted average giving more importance to recent returns.
- **Formula:**

$$\text{EMA}_t = \alpha \cdot r_t + (1 - \alpha) \cdot \text{EMA}_{t-1} \qquad \alpha = \frac{2}{\text{span} + 1}$$

- **Strengths:**
  - Adapts quickly to market shifts and structural changes.
  - Caputres time varying trends better than simple historical average
- **Weaknesses:**
  - Choice of decay factor is subjective
  - Can overreact to short term market noise.
  - May miss long-term patterns if older data down weighted too much.

CAPITAL ASSET PRICING MODEL (CAPM) RETURN

- **Definition:** The expected return of an asset is determined by its sensitivity to market movements, captured through the asset's β, relative to the risk-free rate and the expected market return.
- **Formula:**

$$\mathbb{E}[R_i] = R_f + \beta_i \cdot (\mathbb{E}[R_m] - R_f) \qquad \beta_i = \frac{\mathrm{Cov}(R_i, R_m)}{\mathrm{Var}(R_m)}$$

- **Strengths:**
  - Incorporates market risk (systematic risk).
  - Theoretically grounded & widely used.
- **Weaknesses:**
  - Relies on perfect market assumptions.
  - Single-factor model; ignores other drivers.
  - Beta & market return hard to estimate accurately.

COVARIANCE MATRIX ESTIMATION

In portfolio optimisation, we must take up risk to pursue return. It is about understanding how much a portfolio's value can fluctuate and how different assets move together. We can't just rely on past data because market conditions change - periods of calm can be followed by sudden crises. Hence, we need stable and reliable risk estimates of the assets. So, optimisers can allocate weight among assets by weighing expected return against the uncertainty of those returns.

Approaches tested in this study:

- Sample Covariance
- Exponentially Weighted Covariance (EWCM)
- Ledoit–Wolf Shrinkage

SAMPLE COVARIANCE

- **Definition:** Historical measure of how pairs of assets move together.
- **Formula:**

$$\text{Sample Variance:} \quad \mathrm{Var}(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$\text{Sample Covariance:} \quad \mathrm{Cov}(x, y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

- **Strengths:**
  - Simple to compute & interpret.
  - Standard baseline for portfolio models.
- **Weaknesses:**
  - Assumes static relationships.
  - Poor performance in volatile markets.

EXPONENTIALLY WEIGHTED COVARIANCE

- **Definition:** More weight to recent data → better captures recent market shifts.
- **Formula:**

$$\text{Demeaning: } X_t' = X_t - \bar{X}, \quad Y_t' = Y_t - \bar{Y}$$

$$\text{Covariation Series: } C_t = (X_t - \bar{X})(Y_t - \bar{Y})$$

$$\text{EWMA}_t = \alpha C_t + (1 - \alpha)\text{EWMA}_{t-1}$$

$$\text{Cov}_{\text{EW}}(X, Y) = \text{EWMA}_T$$

- **Strengths:**
  - Responsive to structural changes.
  - Captures shifts in market volatility
  - Reduces noise from old data.
- **Weaknesses:**
  - Span/decay factor choice is subjective.
  - Can ignore long-term relationships.

COVARIANCE SHRINKAGE

- **Definition:** Blend noisy sample covariance with a structured target to reduce estimation error. Balance trade off between bias and variance.
- **Formula:**

$$\hat{\Sigma}_{\text{shrunk}} = (1 - \delta)S + \delta F$$

- The goal is to choose F and δ such that the mean squared error (MSE) between the estimated Σ and the true (but unknown) covariance matrix Σ is minimized.

- **Shrinkage targets:**
  - *Constant Variance Target:* All assets have the same variance, no correlations.
  - *Constant Correlation Target:* All asset pairs have the same correlation, variances preserved.
  - *Single-Factor Model Target:* Returns driven by one common factor (e.g., market index), like CAPM.

COVARIANCE SHRINKAGE

- **Strengths:**
  - Reduces the variance of estimates
  - Makes estimates more stable and robust in small-sample (more assets than data points) or high-dimensional settings
- **Weaknesses:**
  - Requires choosing good target & shrinkage intensity.
  - Introduces bias if target poorly chosen.

PORTFOLIO OPTIMISATION

- Portfolio construction is about balancing risk and return.
- People have different risk preferences, so the portfolio objective should match the investor's comfort with risk. Given the risk preference of the investor the goal is to find the best set of portfolio weights.
- Harry Markowitz (1952) introduced the **Mean-Variance Optimization model**.

$$\max_{\mathbf{w}} \ \boldsymbol{\mu}^\top \mathbf{w} - \lambda \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w} \quad \text{subject to} \quad \sum_{i=1}^{n} w_i = 1, \quad w_i \geq 0$$

- Here, $\lambda \geq 0$ is a risk-aversion parameter. Adjusting λ allows investors to generate different portfolios depending on their risk tolerance.
- By solving this optimization, investors can trace out a set of optimal portfolios offering the highest expected return for each level of risk.

CHOICE OF OBJECTIVES

The 2 objectives we focused on:

Global Minimum Variance: Focuses primarily on minimising risk; Preferred by highly risk-averse investors who prioritise stability over returns.

Maximum Sharpe Ratio: Aims to maximise return per unit of risk (risk-adjusted return); Chosen by investors willing to take calculated risks for higher returns.

Why both?

- MVP = good baseline, robust
- Sharpe = explores performance potential
- Together = complete view of trade-offs for conservative & aggressive investors

Conservative investors → Prefer GMV.

Moderate/Aggressive investors → Lean toward Max Sharpe.

MINIMUM VARIANCE PORTFOLIO

- **Goal:** Minimize portfolio volatility
- **Formulation:**

$$\min_{w \in \mathbb{R}^N} \quad w^\top \hat{\Sigma}_h w + \lambda \sum_{i=1}^{N_h} c_{h,i} \left| w_{h,i} - w^*_{h-1,i} \right|$$

$$\text{Return target constraint: } \boldsymbol{\mu}_h^\top \mathbf{w} \geq b_h,$$

$$\text{Budget constraint: } \mathbf{1}^\top \mathbf{w} = 1,$$

$$\text{Gross exposure constraint: } \|\mathbf{w}\|_1 = \sum_{i=1}^{N} |w_i| \leq \kappa.$$

Pros:

- Stable returns
- Robust even if expected returns are uncertain
- Low volatility and smaller drawdowns — better downside protection.

Cons:

- May lag in bull markets
- Ignores expected returns, so can miss high-return opportunities.
- Can become concentrated in few assets (low volatility)

MAX SHARPE RATIO PORTFOLIO

- **Goal:** Minimize portfolio volatility
- **Formulation:**

$$\max_{\mathbf{w}\in\mathbb{R}^N} \quad \frac{\hat{\mu}_h^\top \mathbf{w}}{\sqrt{\mathbf{w}^\top \hat{\Sigma}_h \mathbf{w}}} - \lambda \sum_{i=1}^{N_h} c_{h,i} \left| w_{h,i} - w^*_{h-1,i} \right|$$

Constraints:

$$\text{Budget constraint: } \mathbf{1}^\top \mathbf{w} = 1,$$

$$\text{Gross exposure constraint: } \|\mathbf{w}\|_1 = \sum_{i=1}^{N} |w_i| \le \kappa.$$

Pros:

- Explicitly targets optimal risk-adjusted return.
- Can generate higher returns in favorable market conditions (bull markets).

Cons:

- Highly sensitive to expected return estimates, which are very noisy (can underperform out of sample if return estimates are wrong).
- Can produce extreme weights (over-concentration) if one asset's estimated Sharpe is high.

EVALUATION METRICS

We assess portfolios using multiple metrics to cover:

- **Return generation** — how much money it makes
- **Risk exposure** — how much uncertainty or downside
- **Risk-adjusted returns** — efficiency of returns for the risk taken
- Metrics Used:
  - CAGR
  - Annualized Volatility
  - Maximum Drawdown
  - Sharpe Ratio

EVALUATION METRICS

Compounded Annual Growth Rate:

Average annual growth rate over time assuming compounding.

$$\mathrm{CAGR} = \left(\frac{V_{\text{final}}}{V_{\text{initial}}}\right)^{\frac{1}{n}} - 1$$

Advantage:

- Smooths out year-to-year fluctuations in returns.
- Presents a single, steady annual growth rate for easier portfolio comparisons.

Limitation:

- Does not account for volatility or risk.
- Two portfolios can have the same CAGR but very different return patterns (stable vs. volatile).
- Must be paired with other metrics (Sharpe, Sortino, Max Drawdown) for a complete performance view.

EVALUATION METRICS

Volatility:

Calculated as the standard deviation of returns

- Measures risk by quantifying how much portfolio returns fluctuate over time.
- High volatility → returns vary widely from average, more uncertainty & risk.
- Low volatility → returns are more stable and consistent.
- Complements return metrics (e.g., CAGR) by showing consistency of growth.
- Key input for risk-adjusted performance measures like Sharpe & Sortino ratios.

EVALUATION METRICS

Max Drawdown:

Largest observed loss from peak to trough during a given period.

$$\text{Max Drawdown} = \frac{\text{Trough Value} - \text{Peak Value}}{\text{Peak Value}}$$

Advantage:

- Clearly measures worst-case historical decline in portfolio value.
- Useful for assessing downside risk and recovery periods.

Limitation:

- Focuses only on the single worst decline
- Ignores frequency of smaller drawdowns.
- Needs to be paired with other risk metrics for a complete view.

EVALUATION METRICS

Sharpe Ratio:

Excess return over the risk-free rate per unit of total risk (volatility).

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p}$$

Advantage:

- Measures risk-adjusted return, making different portfolios directly comparable.
- Useful for capital allocation and portfolio selection decisions.

Limitation:

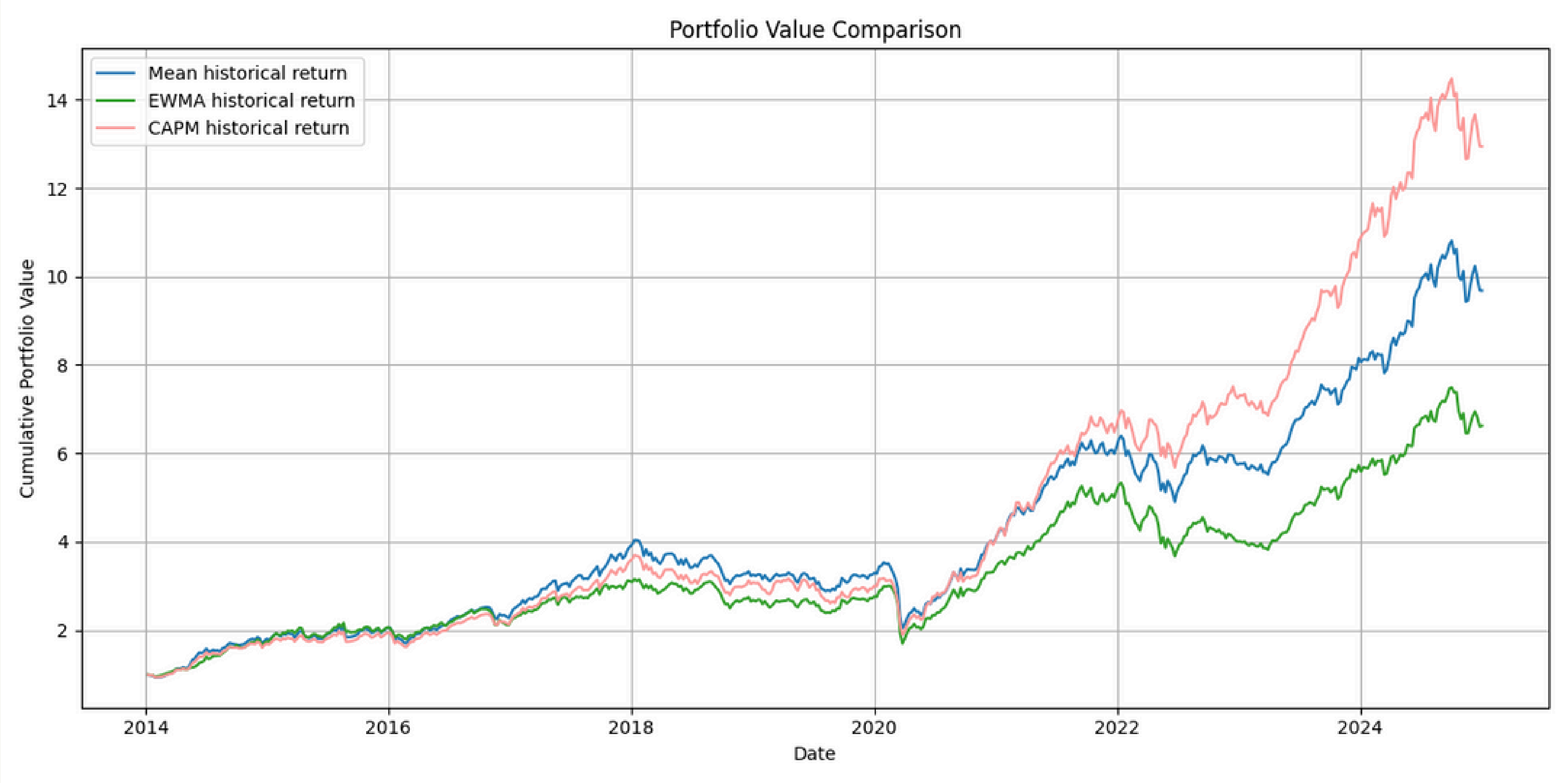
- Treats upside and downside volatility equally.
- May understate the appeal of investments with asymmetric return profiles.
- Sensitive to assumptions about the risk-free rate.

BACKTEST DESIGN

- **Period:** 2013–2025
- **Training windows:**
  - Returns → 10 years (520 weeks)
  - Covariance → 1 year (52 weeks)
- **Rebalancing frequency:** Weekly
- **Decide Investor Objective:** Conservative → Global Minimum Variance (GMV). Aggressive → Maximum Sharpe Ratio.
- **Estimate Key Inputs:**
  - Expected Returns: Historical Mean, EW Mean, CAPM, or ML-based signals.
  - Risk (Covariance Matrix): Sample, EWMA, or Shrinkage.
- **Portfolio Optimisation:** Compute optimal weights given chosen objective and constraints.
- **Evaluation & Validation:** Assess performance metrics (CAGR, volatility, Sharpe, drawdowns).

RESULTS

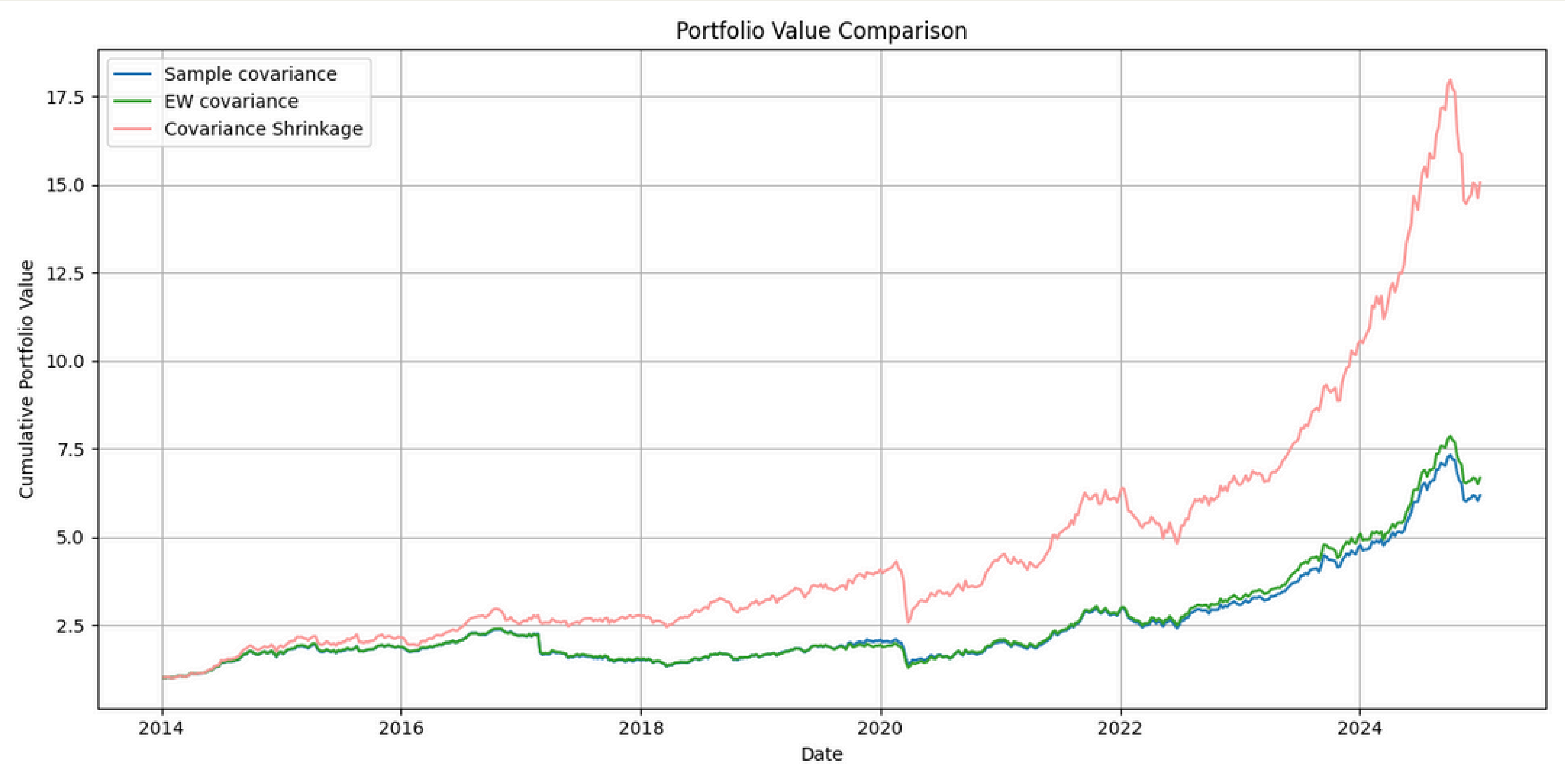
Among the three methods of return estimation the method that considers market behaviour produced the most stable and highest long-term growth.



Metrics	Mean Historical Return	Exponentially Weighted Mean Historical Return	CAPM Return
Compounded Annual Growth Rate (CAGR)	20.17%	15.30%	23.94%
Annualised Volatility	17.65%	16.16%	18.37%
Max Drawdown/volatility	2.53	2.38	2.48
Sharpe Ratio	1.13	0.97	1.26

RESULTS

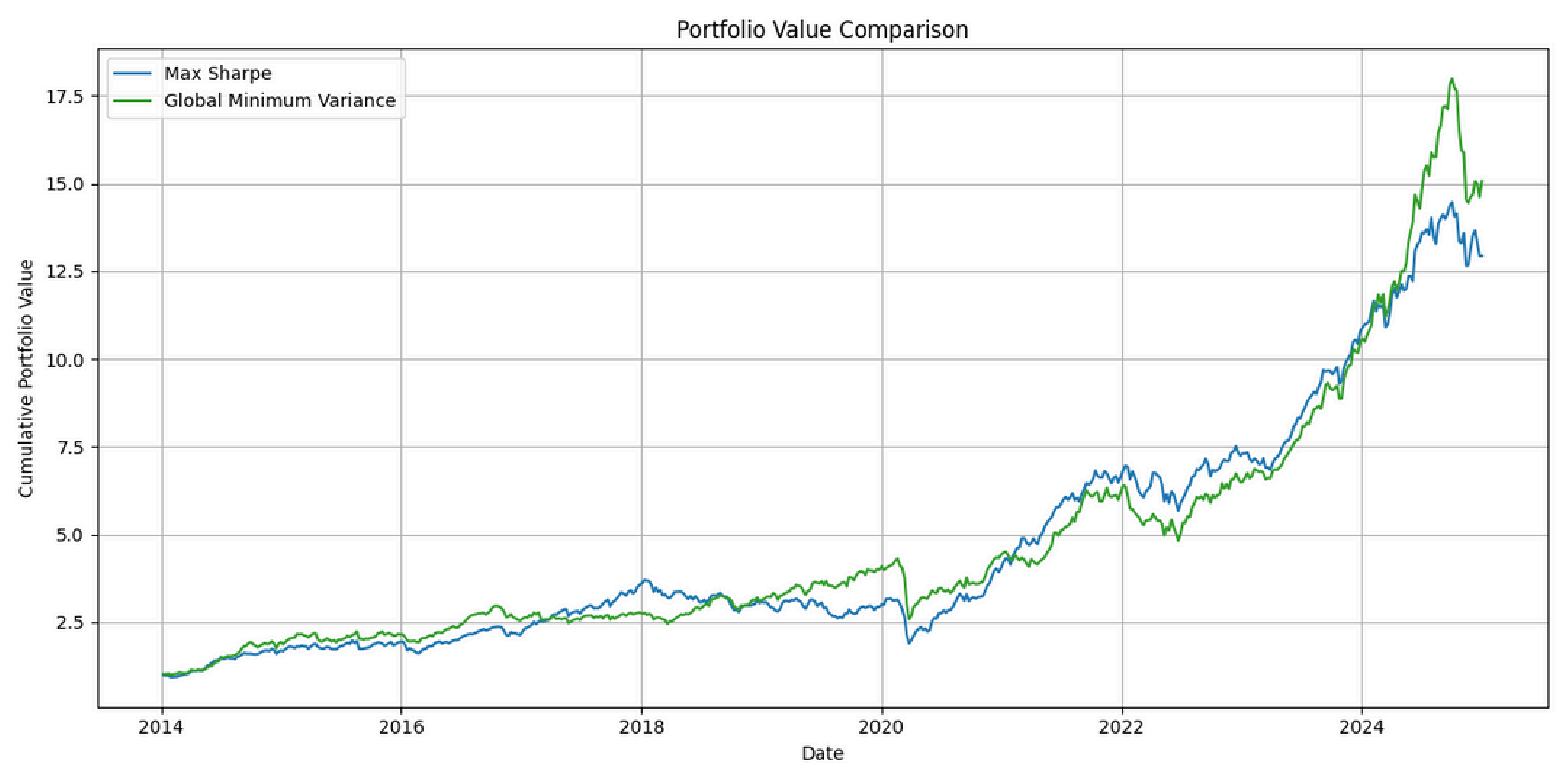
Among the three methods of risk estimation the method that regularizes the estimates by introducing a stable target delivered the best results.



Metrics	Sample Covariance	Exponential Covariance	Covariance Shrinkage
Compounded Annual Growth Rate (CAGR)	11.64%	11.98%	15.84%
Annualised Volatility	12.75%	12.57%	11.5%
Max Drawdown/volatility	2.36	2.43	2.16
Sharpe Ratio	0.92	0.96	1.33

RESULTS

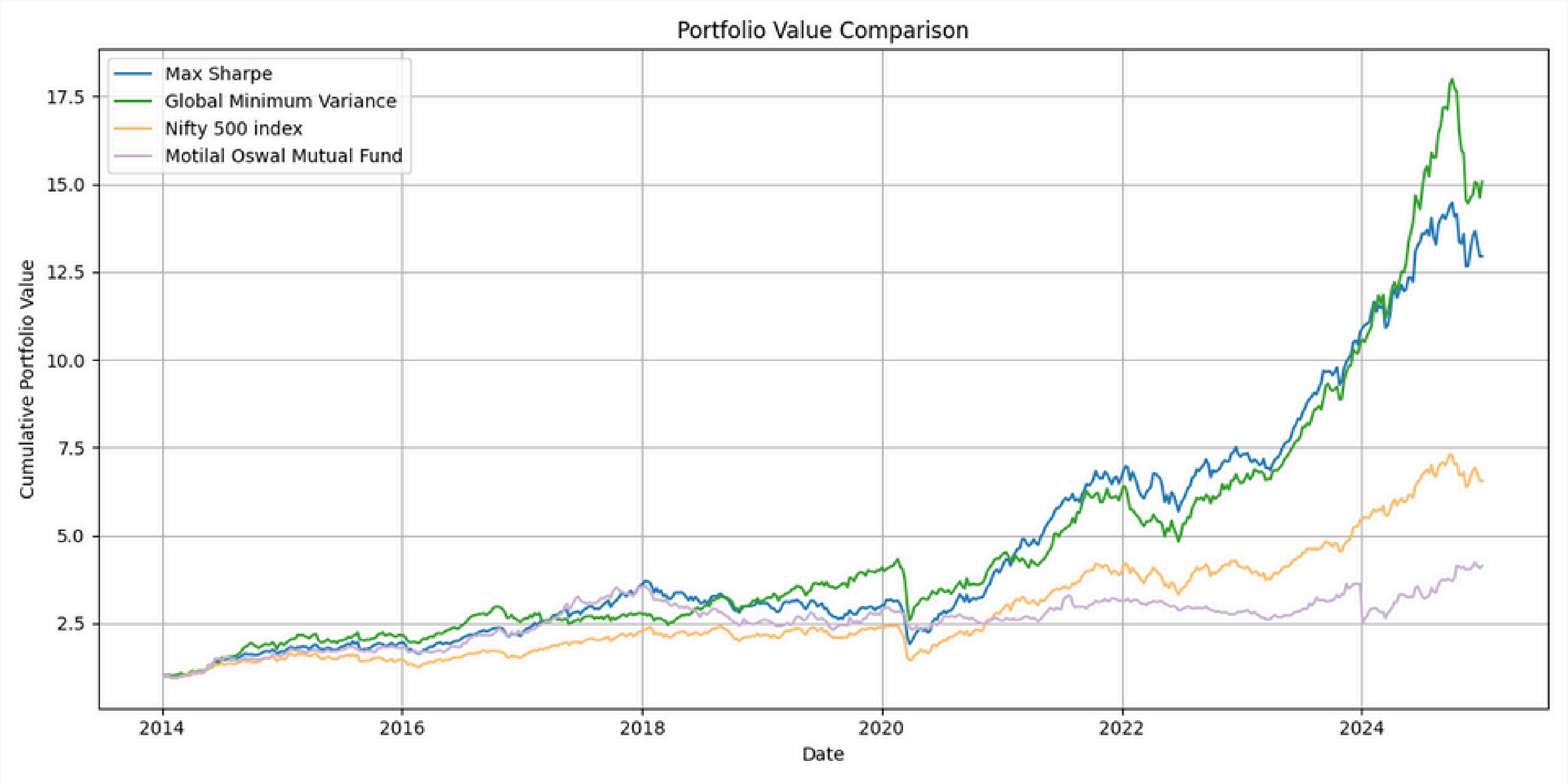
1. The return-focused objective was highly sensitive to errors in return forecasts, which lead to unstable allocations.
2. In contrast, the risk-focused strategies performed better because it depended less on noisy return estimates and focused more on building stable, low-volatility portfolios.
3. Portfolios built with reliable return and risk estimates performed better across multiple metrics.



Metrics	Global Minimum Variance	Max Sharpe Ratio
Compounded Annual Growth Rate (CAGR)	23.94%	15.84%
Annualised Volatility	18.37%	11.5%
Max Drawdown/Volatility	2.48	2.16
Sharpe Ratio	1.26	1.33

RESULTS

1. Optimized portfolios, especially GMV, offered superior risk-adjusted returns and stability, while benchmarks and high-volatility strategies achieved higher growth at the cost of efficiency.
2. Performance differences reflect each strategy's approach to risk and diversification.



Metrics	Global Minimum Variance	Max Sharpe Ratio	Nifty 500 index	Motilal Oswal Mutual Fund
Compounded Annual Growth Rate (CAGR)	23.94%	15.84%	14.68%	24%
Annualised Volatility	18.37%	11.5%	15.70%	47.94%
Max Drawdown/Volatility	2.48	2.16	2.12	1.65
Sharpe Ratio	1.26	1.33	0.95	0.74

CONCLUSION & TAKEAWAYS

Main Findings:

- CAPM returns + Shrinkage covariance = best risk-adjusted performance
- GMV strategy more robust out-of-sample than Max Sharpe
- Optimized portfolios beat Nifty 500 index, and passive strategies like equal weight, and mutual fund benchmarks

Key Takeaways:

1. Estimator Choice Matters — CAPM outperforms mean returns & EMA
2. Stable Risk Models Win — Shrinkage (Ledoit–Wolf) > sample covariance
3. Risk Focus Beats Return Chase — GMV > Max Sharpe in stability & efficiency
4. Quant Models Can Outperform Benchmarks — especially in risk-adjusted terms
5. Mind the Limitations — survivorship bias, simplified costs, short-selling assumptions

LIMITATIONS

1. Survivorship Bias

- Used current NIFTY 500 constituents only → excludes delisted/merged/failed firms
- Inflates performance estimates
- Bias-free datasets are usually paid

2. Fixed Transaction Costs

- Assumed constant 1% per trade
- Ignores liquidity, volume, and market impact

3. Short Selling Restrictions

- Assumed unrestricted short selling
- Ignores borrowing fees, margin rules, and stock availability

4. Constant Risk-Free Rate

- Fixed at 7% (long-term govt. securities)
- Real rates vary with macroeconomic conditions

THANK YOU