FMA4200 Final Report
Time-Series Modeling on Monthly Returns and Trading Strategies with 6 Portfolios

Jiaxing Wei
School of Data Science
The Chinese University of Hong Kong, Shenzhen
jiaxingwei@link.cuhk.edu.cn

1 Introduction

The econometric modeling of financial time series has undergone remarkable evolution since Bollerslev [1]'s groundbreaking development of the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) framework. This seminal work provided the foundation for modeling time-varying volatility, with subsequent extensions like Nelson's [2] Exponential GARCH (EGARCH) and Glosten, Jagannathan, and Runkle's [3] GJR-GARCH models introducing crucial innovations to capture the asymmetric impact of positive versus negative shocks on volatility processes. These methodological advances have revolutionized empirical finance by enabling precise quantification of key stylized facts in asset returns - particularly volatility clustering (where large changes tend to be followed by large changes of either sign) and the leverage effect (where negative returns increase future volatility more than positive returns of equal magnitude).

Parallel to these developments in time series analysis, Markowitz [4]'s mean-variance optimization framework established the cornerstone of modern portfolio theory, introducing the quantifiable trade-off between expected return and risk. However, subsequent research by Black and Litterman [5] exposed critical limitations in this approach, particularly the sensitivity to input parameters and the tendency to produce extreme portfolio weights when relying solely on historical return estimates. The ongoing debate about the practical utility of optimization techniques was further intensified by DeMiguel et.al. [6] influential study, which demonstrated that naive equally-weighted portfolios often outperform sophisticated optimization strategies out-of-sample, challenging conventional wisdom in asset allocation.

These two research trajectories - volatility modeling in time series and portfolio optimization - have largely developed independently despite their natural connections. The former provides sophisticated tools for estimating return distributions and their higher moments, while the latter offers frameworks for optimal capital allocation given these distributions. Recent empirical evidence suggests that bridging this disciplinary divide may yield significant improvements in both volatility forecasting and portfolio construction methodologies.

This study employs classical time series models to investigate the return characteristics of investment portfolios. Building upon relevant financial theories, we construct trading strategies and rigorously examine their empirical performance through comprehensive backtesting procedures [1, 4].

2 Dataset Description

Data Source: This dataset was generated by $CMPT\_ME\_BEME\_OP\_INV\_RETS$ using the 202412 *CRSP* database, providing both value-weighted and equal-weighted returns for portfolios sorted by market equity (ME) and operating profitability (OP). Portfolio formation occurs at the end of June each year, with OP defined as sales minus cost of goods sold minus selling expenses, measured at the prior fiscal year-end. Returns of all 6 portfolios in the original dataset are reported on an annual basis (January–December). Notably, the value-weighted averages of operating profitability for the extreme deciles (1 and 10) exhibit substantial outliers, attributable to extreme OP values among certain firms. Manual verification of the underlying accounting data confirms that these extreme values correctly reflect the financial statements of the respective firms. The breakpoints and portfolio compositions incorporate utilities and financial firms, ensuring comprehensive coverage across sectors. The portfolios formed from July of year t through June of t+1 comprise all NYSE, AMEX, and NASDAQ stocks that meet the following criteria:

① available market equity data as of June t

② positive book equity for year $t-1$

③ non-missing revenue data for $t-1$

④ available data for at least one of the following $t-1$ financial items: cost of goods sold, SG&A expenses, or interest expense.

Preprocessing:

① Missing observations are coded as -99.99 or -999. To mitigate the impact of outliers, the report recodes them as *NAN*.

② The monthly returns of the portfolio are converted into percentage terms and scaled to the range of $[-1, 1]$ to represent unleveraged price movements.

Variables: The dataset presented in Table 1 comprises monthly return series for six investment portfolios across 738 consecutive months, covering the period from July 1963 through December 2024. All portfolios exhibit positive mean returns, averaging approximately 1%. Their volatilities remain at relatively comparable levels, ranging between 4% and 6%, characterizing them as low-volatility instruments. Notably, all six portfolios demonstrate extreme monthly losses or gains reaching 17% or even 30%, indicating non-negligible tail risks. The six portfolios are constructed according to the following methodology:

① *SMALL LoOP*: Stocks with small market size and low operating profitability.

② *ME1 OP2*: Stocks with small market size and neutral operating profitability.

③ *SMALL HiOP*: Stocks with small market size and high operating profitability.

④ *BIG LoOP*: Stocks with large market size and low operating profitability.

⑤ *ME2 OP2*: Stocks with large market size and neutral operating profitability.

⑥ *BIG HiOP*: Stocks with large market size and high operating profitability.

Table 1: Portfolio Monthly Return Statistics

Statistic	SMALL LoOP	ME1 OP2	SMALL HiOP	BIG LoOP	ME2 OP2	BIG HiOP
count	738.000	738.000	738.000	738.000	738.000	738.000
mean	0.010	0.012	0.013	0.008	0.009	0.010
std	0.066	0.054	0.061	0.051	0.044	0.044
min	−0.320	−0.278	−0.307	−0.236	−0.203	−0.218
25%	−0.026	−0.018	−0.024	−0.019	−0.015	−0.016
50%	0.013	0.015	0.017	0.012	0.012	0.012
75%	0.049	0.045	0.049	0.040	0.035	0.039
max	0.310	0.260	0.277	0.177	0.182	0.171
skew	−0.219	−0.471	−0.473	−0.515	−0.373	−0.381
kurt	2.075	2.493	2.980	1.863	1.684	1.560

3 Distributional Modeling

The report examines the distributional properties of portfolio monthly returns through histogram analysis, which reveals distinct profitability patterns across investment strategies. The central location of return distributions, measured by sample means, serves as an indicator of expected return performance, with only portfolios showing statistically significant positive means qualifying as potentially economically viable investments.

The skewness coefficient of return distributions provides critical diagnostic information about performance asymmetry. larger skewed portfolios, characterized by frequent small gains and infrequent large losses, tend to produce more favorable investor experiences due to their higher probability of positive outcomes. Such distributional properties influence both investor satisfaction and long-term holding behavior, making skewness analysis particularly relevant for strategy evaluation.

3.1 Distributional Properties

Figure 1 reveals that the return distributions of *SMALL LoOP* and *ME1 OP2* exhibit relatively normal characteristics, indicating these portfolios of small-to-medium capitalization stocks with limited operating profitability benefit from sufficient market information and balanced investor expectations. Their price movements approximate random Brownian motion, reflecting efficient market pricing where investor views are well-incorporated and trading activity is adequately competitive. This pattern suggests these firms face symmetric market sentiment without significant directional bias.

The analysis further shows that *SMALL HiOP*, *BIG HiOP*, and *ME2 OP2* display larger skewness in their return distributions, while *BIG LoOP* exhibits negative skewness. The large skewed portfolios, particularly small-cap high-profitability firms, demonstrate investors' strong preference for quality companies with superior operating performance, generating more frequent upside movements. Conversely, the negative skewness of *BIG LoOP*- large capitalization stocks with weak profitability - aligns with their lowest mean returns in Table 1, reflecting market skepticism about their growth prospects. These distributional properties confirm that profitability characteristics significantly influence both return magnitudes and their asymmetric patterns across different market capitalization segments.

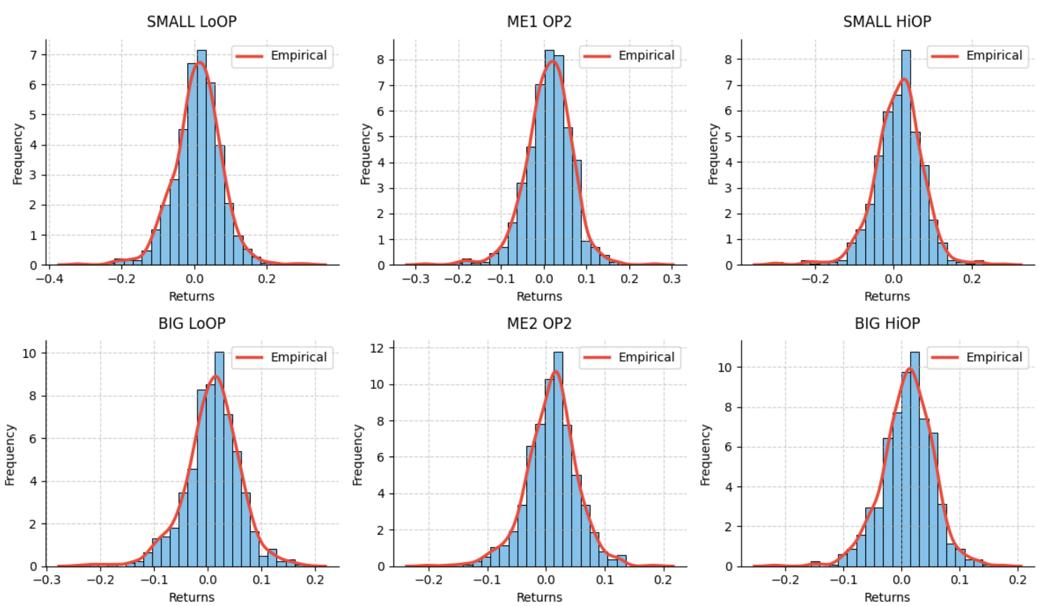


Figure 1: Distribution of monthly returns from 1963-07 to 2024-12

3.2 Distributional Variants

In this section, the report employs normal distribution and several common variant estimators to individually fit the 6 portfolios. The monthly returns sequence is denoted as $X_i, i \in \{1, 2, 3, 4, 5, 6\}$. These distributions will undergo parameter calibration via maximum likelihood estimation (MLE) across the sample space X_i, accompanied by corresponding statistical tests.

Normal Distribution: The normal distribution is characterized by its mean μ and variance σ^2. The probability density function (PDF) is represented as:

$$f_N(x) = \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{(x-\mu)^2}{2\sigma^2}), \forall x \in X_i$$

Student-t Distribution: can simultaneously model both fat-tailed behavior and volatility clustering effects. The PDF is given as:

$$f_t(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{(x-\mu)^2}{\nu s^2}\right)^{-\frac{\nu+1}{2}}, \forall x \in X_i$$

Normal Inverse Gaussian Distribution (NIG):will capture the leptokurtic and heavy-tailed characteristics of return distributions

$$f_{NIG}(x) = \frac{\alpha\delta}{\pi} \frac{K_1\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)}{\sqrt{\delta^2 + (x-\mu)^2}} e^{\delta\gamma + \beta(x-\mu)}, \forall x \in X_i$$

The parameters govern the distribution's properties as follows: α controls tail thickness, β adjusts distributional skewness, δ serves as the scale parameter, μ represents the location parameter, and $K_1()$ denotes the modified Bessel function of the second kind.

3.3 Hypothesis Test

The report employs the Shapiro-Wilk (SW) test to assess the normality of X_i, while utilizing the Kolmogorov-Smirnov (KS) test to evaluate whether X_i follows either the NIG or t-distribution at various confidence levels.

Shapiro-Wilk (SW): the test statistics is given by:

$$W = \frac{(\sum_{i=1}^n a_i x_i)^2}{(\sum_{i=1}^n (x_i - \bar{x})^2}$$

where x_i is the i-th order statistic of the sample X_i, while m_i, m_0 represent the i-th moment of standard normal distribution and the sample mean, respectively. Then a_i is the i-th element of the vector $(a_1, a_2, ..., a_n)$ given by:

$$a_i = \frac{m_i}{\sqrt{m_0 m_i}}$$

Then under the normal distribution with sample size n:

$$W \sim Beta(\frac{n}{2}, \frac{n}{2})$$

In this case, the p-value is the probability of making type I errors. Denoting w_{obs} as the value of the test statistic and $f_W(w)$ as the PDF of $Beta(\frac{n}{2}, \frac{n}{2})$:

$$p = P(W \leq w_{obs}|H_0) = \int_0^{w_{obs}} f_W(w)dw$$

Kolmogorov-Smirnov (KS): The KS test is implemented to test whether the sample X_i is drawn from certain distributions:

$$D_n = sup_{x\in\mathbb{R}}||F_n(x) - F(x)||$$

where $F_n(x)$ is the empirical distribution function of the sample and $F(x)$ is the PDF of the hypothesized distribution. Under the null hypothesis, the test statistic $\sqrt{n}D_n$ converges to the Kolmogorov distribution:

$$P(\sqrt{n}D_n > c) \approx 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2c^2}$$

3.4 Test Statistics

From Table 2, 3, 4, 5. 6, 7 it can be observed that, at the 95% and 99% confidence levels, the distribution of the monthly returns of the six portfolios can be assumed to follow a t-distribution. The null hypothesis that does not follow a normal distribution or an NIG distribution cannot be rejected, indicating that, in a statistical sense, no significant results are obtained in the fitting process.

All 6 portfolios are well characterized by the Student-t distribution, indicating that portfolios sorted by size and operating profitability share similar distributional properties, particularly the presence of statistically significant fat tails that cannot be ignored.

Table 2: Hypothesis Test on SMALL LoOP

Distribution	Statistic	p-value ($\alpha = 0.05$)	p-value ($\alpha = 0.01$)
Normal	0.980	0.000	0.000
Student-t	0.029	0.560	0.560
NIG	4.683	0.000	0.000

Table 3: Hypothesis Test on ME1 OP2

Distribution	Statistic	p-value ($\alpha = 0.05$)	p-value ($\alpha = 0.01$)
Normal	0.974	0.000	0.000
Student-t	0.023	0.806	0.806
NIG	5.712	0.000	0.000

Table 4: Hypothesis Test on SMALL HiOP

Distribution	Statistic	p-value ($\alpha = 0.05$)	p-value ($\alpha = 0.01$)
Normal	0.967	0.000	0.000
Student-t	0.024	0.797	0.797
NIG	5.113	0.000	0.000

Table 5: Hypothesis Test on BIG LoOP

Distribution	Statistic	p-value ($\alpha = 0.05$)	p-value ($\alpha = 0.01$)
Normal	0.976	0.000	0.000
Student-t	0.024	0.778	0.778
NIG	6.101	0.000	0.000

Table 6: Hypothesis Test on ME2 OP2

Distribution	Statistic	p-value ($\alpha = 0.05$)	p-value ($\alpha = 0.01$)
Normal	0.981	0.000	0.000
Student-t	0.027	0.660	0.660
NIG	7.159	0.000	0.000

Table 7: Hypothesis Test on BIG HiOP

Distribution	Statistic	p-value ($\alpha = 0.05$)	p-value ($\alpha = 0.01$)
Normal	0.985	0.000	0.000
Student-t	0.023	0.816	0.816
NIG	-	-	-

4 Time-Series Modeling

As shown in Figure 2, 3, 4, portfolios with small market capitalization and OP have the best long position value, followed by portfolios with medium size and OP. Although the OP of these stocks is not as high as that of the former, the size advantage makes their valuation relatively stable. When OP is low, large size ensures that stock prices do not experience long-term declines, as investors naturally trust large companies. The profitability of the remaining portfolios, however, is relatively mediocre. Nonetheless, all portfolios have been profitable in the long run over the past 60 years, especially when compound interest has grown thousands of times compared to simple interest. As Buffett famously said [], "The best thing you can do is to be a friend of time."

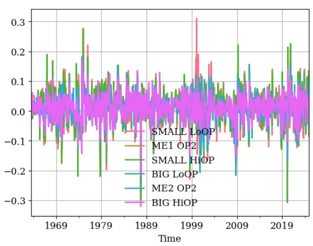


Figure 2: Monthly Returns

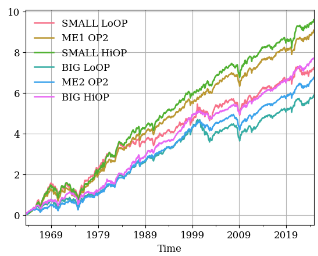


Figure 3: Simple Interest

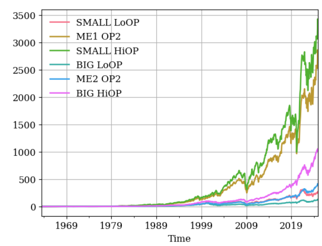


Figure 4: Compound Interest

4.1 Stationary Test

The monthly return series of the 6 portfolios undergo pre-fitting tests in the following order:

① *Augmented Dickey-Fuller (ADF) test*: In this mission, the report utilizes constant and trend ADF to examine the stationary of returns:

$$lnS_t = \alpha lnS_{t-1} + \beta + \gamma t + \sum_{i=1}^{p} \beta_i \Delta lnS_{t-i} + \epsilon_t$$

The test statistic is $DF = \frac{\hat{\alpha}-1}{SE(\hat{\alpha})}$ with null hypothesis $\alpha = 1$. If the test is rejected, then the time series is stationary, i.e., it does not have a unit root. This indicates constant mean and variance over time, which is a key assumption for time series models like ARIMA.

② *Ljung-Box test*: Denotes $\hat{\rho}(k)$ as the sample autocorrelation function (ACF) of returns and m a prespecified positive integer. The statistics can be written as:

$$Q(m) = T(T+2) \sum_{k=1}^{m} \frac{\hat{\rho}(k)}{T-k}$$

If $Q(m) > \chi^2_{1-\alpha}(m)$, where $\chi^2_{1-\alpha}(m)$ is the $(1-\alpha)$ quantile of the chi-square distribution with m degrees of freedom, it rejects the null hypothesis of white noise at the significance level α, i.e., returns series contain predictable structure and may benefit from further modeling (e.g., AR, MA, ARIMA)

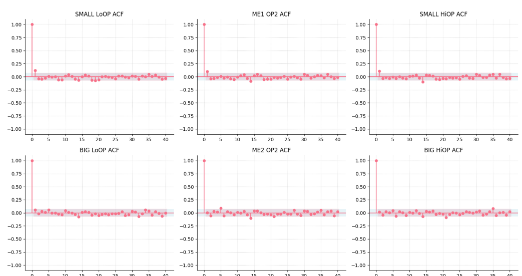
Referring to Table 8, at the 5% significance level, the ADF test indicates that all six portfolio return series are stationary. For *SMALL LoOP*, *ME1 OP2*, and *SMALL HiOP*, the Ljung-Box test further reveals significant autocorrelation, suggesting that these series are suitable for basic AR or ARMA models. In contrast, *BIG LoOP*, *ME2 OP2*, and *BIG HiOP* show no significant autocorrelation, indicating that further testing is required to detect volatility clustering before applying a GARCH model, or alternatively, external variables could be introduced for modeling. Figure 5a also reflects the complex nonlinear dependencies between the monthly return series of the last three portfolios. When the order is less than 5, the ACF is not significant, indicating that ARMA modeling is not suitable. The report will examine their annualized volatility changes in the next subsection to ultimately determine the appropriate time series model.

Table 8: Stationary and Autocorrelation p-Values

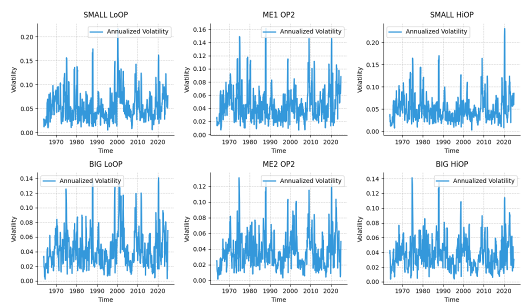
Portfolio	ADF	Ljung-Box
SMALL LoOP	0.000	0.001
ME1 OP2	0.000	0.008
SMALL HiOP	0.000	0.003
BIG LoOP	0.000	0.136
ME2 OP2	0.000	0.965
BIG HiOP	0.000	0.596

4.2 Volatility Clustering

Another common phenomenon in financial time series is volatility clustering, where high or low volatility may persist for a continuous period of time. This phenomenon suggests that investors' expectations of the market tend to have a certain degree of inertia, meaning that even if a shift occurs, it takes time for the change to be fully transmitted.



(a) ACF of 6 Portfolios



(b) Annualized Volatility of 6 Portfolios

The pronounced volatility clustering observed in Figure 5b indicates that GARCH-family models represent a feasible modeling direction for these return series where ARMA specifications prove inadequate. Consistent with the findings from Section 4.1 and the current analysis, the return dynamics of the three small-cap portfolios will be modeled using ARIMA-GARCH specifications, whereas standard GARCH models will be implemented for the large-cap portfolios.

4.3 ARIMA-GARCH

According to Engle [7], the ARIMA-GARCH specification simultaneously accommodates the mean dynamics through trend components and volatility clustering patterns in residual terms, with explicit parametric treatment for fat-tailed distributional properties.

Let y_i be the observed value at time t of returns, L be the lag operator ($L^k y_t = y_{t-k}$). Furthermore, σ_t^2 refers to the conditional variance, and α_i, β_i are ARCH coefficients and GARCH coefficients, respectively. It expects that the stationary condition is satisfied, i.e., roots of $1 - \sum_{i=1}^{p} \phi_i L^i = 0$ lie outside unit circle, and $\sum_{i=1}^{m} \alpha_i + \sum_{i=1}^{m} \beta_j < 1$. There exists positively constraints: $\omega > 0, \alpha_i \geq 0, \beta_j \geq 0$. The innovation distribution is required as: $z_t \sim \mathcal{N}(0, 1)$. The model consists of two components:

① *Mean Equation (ARIMA)*:

$$(1 - \sum_{i=1}^{p} \phi_i L^i)(1 - L)^d y_t = c + (1 + \sum_{j=1}^{q} \theta_j L^j)\epsilon_t, \epsilon_t = \sigma_t z_t$$

② *Volatility Equation (GARCH)*:

$$\sigma_t^2 = \omega + \sum_{i=1}^{m} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{n} \beta_i \sigma_{t-j}^2$$

This unified representation highlights the model's dual capability to capture: (i) Mean dynamics via ARIMA (trends/autocorrelation) (ii) Volatility clustering via GARCH (heteroskedasticity)

4.4 GARCH Family

The presence of volatility clustering in financial time series warrants the application of GARCH-type models. Below we specify the canonical GARCH formulation and derive 2 economically meaningful extensions:

① *Exponential GARCH (EGARCH)*: The EGARCH model [2] captures asymmetric volatility responses through its logarithmic formulation, allowing for negative coefficients while guaranteeing positive conditional variance without parameter constraints:

$$ln\sigma_t^2 = \omega + \sum_{i=1}^{m}[\alpha_i \frac{|\epsilon_{t-i}|}{\sigma_{t-i}} + \gamma_i \frac{|\epsilon_{t-i}|}{\sigma_{t-i}}] + \sum_{j=1}^{n} \beta_j ln\sigma_{t-j}^2$$

② *GJR-GARCH*: Proposed by [3], this variant introduces an indicator function to model differential impacts of positive/negative shocks, making it particularly effective for analyzing "leverage effects" in equity markets:

$$\sigma_t^2 = \omega + \sum_{i=1}^{m}[\alpha_i \epsilon_{t-i}^2 + \gamma_i \epsilon_{t-i}^2 \mathbb{1}_{\epsilon_{t-i}<0}] + \sum_{j=1}^{n} \beta_j \sigma_{t-j}^2$$

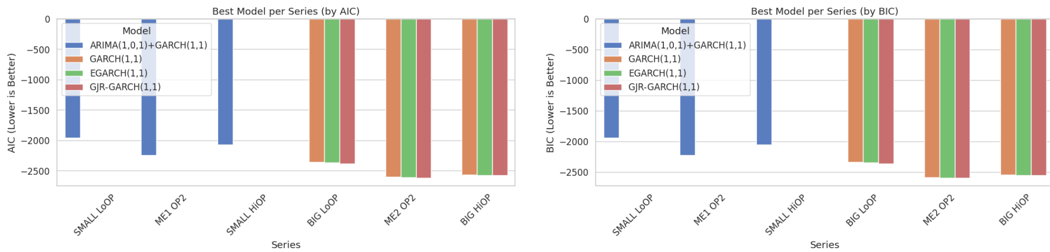
4.5 Fitting Results

As shown in Table 9, and based on the modeling frameworks described in Sections 4.3 and 4.4, the GJR-GARCH model provides the best in-sample fit—measured by both AIC and BIC—for the monthly returns of large-cap stock portfolios, followed by the E-GARCH specification. This suggests that large-cap stocks exhibit stronger volatility in response to negative shocks compared to positive ones. Although the E-GARCH model also captures this asymmetry, it performs slightly worse than the GJR-GARCH model, confirming that large-cap stocks' volatility is best modeled with a model that accounts for asymmetric effects.

The economic significance of this finding is that the use of a standard GARCH model may underestimate tail risk, particularly in the presence of strong negative news or market shocks affecting large-cap stocks. Positioning strategies based on volatility and return predictions from traditional GARCH models may fail under such circumstances.

Table 9: Model Fit Results (AIC and BIC)

Portfolios	Model	AIC	BIC
SMALL LoOP	ARIMA(1,0,1)+GARCH(1,1)	-1959.699	-1941.283
ME1 OP2	ARIMA(1,0,1)+GARCH(1,1)	-2244.775	-2226.359
SMALL HiOP	ARIMA(1,0,1)+GARCH(1,1)	-2072.672	-2054.256
BIG LoOP	GARCH(1,1)	-2360.000	-2341.584
BIG LoOP	EGARCH(1,1)	-2368.822	-2350.406
BIG LoOP	GJR-GARCH(1,1)	-2383.927	-2360.907
ME2 OP2	GARCH(1,1)	-2603.927	-2585.512
ME2 OP2	EGARCH(1,1)	-2613.220	-2594.804
ME2 OP2	GJR-GARCH(1,1)	-2616.492	-2593.472
BIG HiOP	GARCH(1,1)	-2562.339	-2543.923
BIG HiOP	EGARCH(1,1)	-2571.065	-2552.650
BIG HiOP	GJR-GARCH(1,1)	-2574.567	-2551.548



(a) AIC Value of Fitting

(b) BIC Value of Fitting

Figures 6a and 6b further demonstrate that the volatility of small-cap stocks is influenced by more market microstructure factors or non-linear effects. This complexity makes the ARIMA-GARCH model less effective than the pure GARCH model in capturing their volatility characteristics. Additionally, due to the presence of time-series dependencies, forecasting for small-cap stocks is more challenging than for large-cap stocks.

5 Trading Strategies

5.1 Joint Multivariate Time-Series Modeling

Now, by treating a subset of the 6 assets as a single, larger portfolio, then $\mathbf{X}_t = (X_{1,t}, X_{2,t}, ..., X_{d,t})$ can be taken as a d-dimensional vector process. Then the Vector Autoregressive (VAR) model:

$$\mathbf{X}_t = \mathbf{c} + \sum_{i=1}^{p} \mathbf{A}_i \mathbf{X_{t-i}} + \epsilon_\mathbf{t}$$

where $\epsilon_t \sim WN(\mathbf{0}, \mathbf{\Sigma}_\epsilon)$, $\mathbf{c}$ is a d-dimensional vector, $\mathbf{A}_i$ is a $d \times d$ matrix, and p is the order.

The optimal order p can be obtained by optimizing:

① *AIC*:

$$AIC(p) = log(||\hat{\mathbf{\Sigma}}_\epsilon(p)|| + \frac{2d^2p}{T})$$

② *BIC*:

$$BIC(p) = log(||\hat{\mathbf{\Sigma}}_\epsilon(p)|| + \frac{d^2plog(T)}{T})$$

③ *HQIC*:

$$HQIC(p) = log(||\hat{\mathbf{\Sigma}}_\epsilon(p)|| + \frac{2d^2plog(log(T))}{T})$$

5.2 Cointegration Test

The Johansen test reformulates the VAR model into a Vector Error Correction Model (VECM) to test for cointegration with the long-run matrix Π:

$$\Delta\mathbf{Y}_t = \Pi\mathbf{Y}_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta\mathbf{Y}_{t-i} + \epsilon_t$$

Then the hypothesis testing on determining the rank of the matrix Π: if $rank(\Pi) = 0$, then no cointegration; if $0 < rank(\Pi) = r < n$, then r cointegration relationships exist; otherwise all series are stationary. Two statistics are implemented into the test:

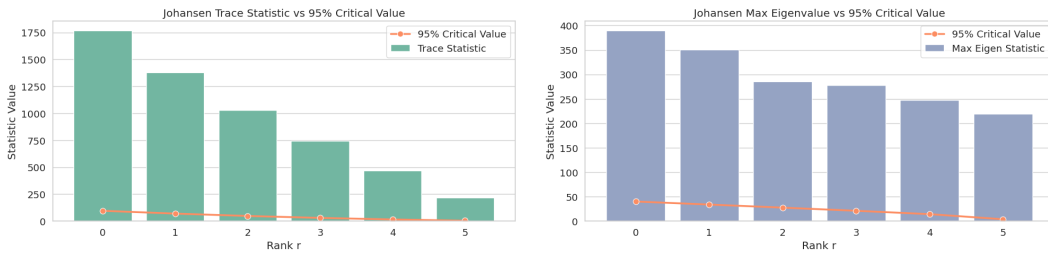
① *Trace (Null: smaller than r)*:

$$Trace(r) = -T \sum_{i=r+1}^{n} ln(1 - \lambda_i)$$

② *Maximum eigenvalue (Null: equals to r)*:

$$MaxEig(r, r+1) = -T ln(1 - \lambda_{r+1})$$

where λ_i are the ordered eigenvalues from the estimated Π.



(a) Johansen Trace

(b) Johansen Max Eigenvalue

In Figure 7a and 7b, all Johansen test statistics exceed the corresponding 95% confidence critical values, revealing strong cointegration among the monthly returns of the six portfolios. Therefore, they can be treated as a unified system for constructing statistical arbitrage portfolios.

5.3 Statistical Arbitrage

Algorithm 1 Stat-Arb Strategy

Input:
  β: Cointegration vector
  $\mathbf{X}_t$: Returns matrix
  σ_1: Short-sell threshold
  σ_2: Long-buy threshold
Output:
  W: Target positions

procedure STATARB($\beta, \mathbf{X}_t, \sigma_1, \sigma_2$)
  $S_t \leftarrow \beta \top \mathbf{X_t}$ ▷ Calculate spread
  if $S_t \sim \mathbb{I}(0)$ **then**
    return $S_t = \beta_1 x_{1t} + ... + \beta_n x_{nt}$ ▷ Expect for mean-reversion
  end if
  if $S_t > \sigma_1$ **then**
    $W_t \leftarrow -1$ ▷ Exceed upper threshold, expect falling
  else if $S_t < \sigma_2$ **then**
    $W_t \leftarrow 1$ ▷ Exceed lower threshold, expect rising
  else
    $W_t \leftarrow 0$ ▷ Within thresholds, neutral position
  end if
  return W ▷ Generate Positions
end procedure

The statistical arbitrage guiding by 1 results for six portfolios are shown in Figure 8, using a 24-month rolling window, 2-month holding period, and trading thresholds of ±1.5 standard deviations within each window. Table 10 clearly demonstrates that this simple statistical arbitrage strategy fails to generate consistent profits. Over the past fifty-plus years, the strategy only executed 354 trades, delivering meager annualized returns below 1% with a Sharpe ratio of just 0.2, while suffering a maximum drawdown as high as 21%. Its overall performance significantly underperformed a long-only buy-and-hold approach. These results prove that arbitrage strategies relying solely on basic entry/exit rules are unsuitable for asset management products - the absence of proper position management leads to excessive risk exposure, while the oversimplified long-short logic renders the core concept fundamentally ineffective.

From a portfolio construction perspective, while the natural cointegration exists across the six portfolios segmented by market capitalization and profitability dimensions, our empirical results demonstrate that traditional time-series predictive models with cointegration testing fail to generate statistically significant arbitrage opportunities in this asset basket. This suggests the necessity of incorporating more sophisticated cross-sectional long-short frameworks to effectively capture price divergences. The intuition suggests that mispricing investors are limited and short-lived, causing arbitrage opportunities to vanish quickly with narrow profit margins - substantially increasing strategy difficulty.

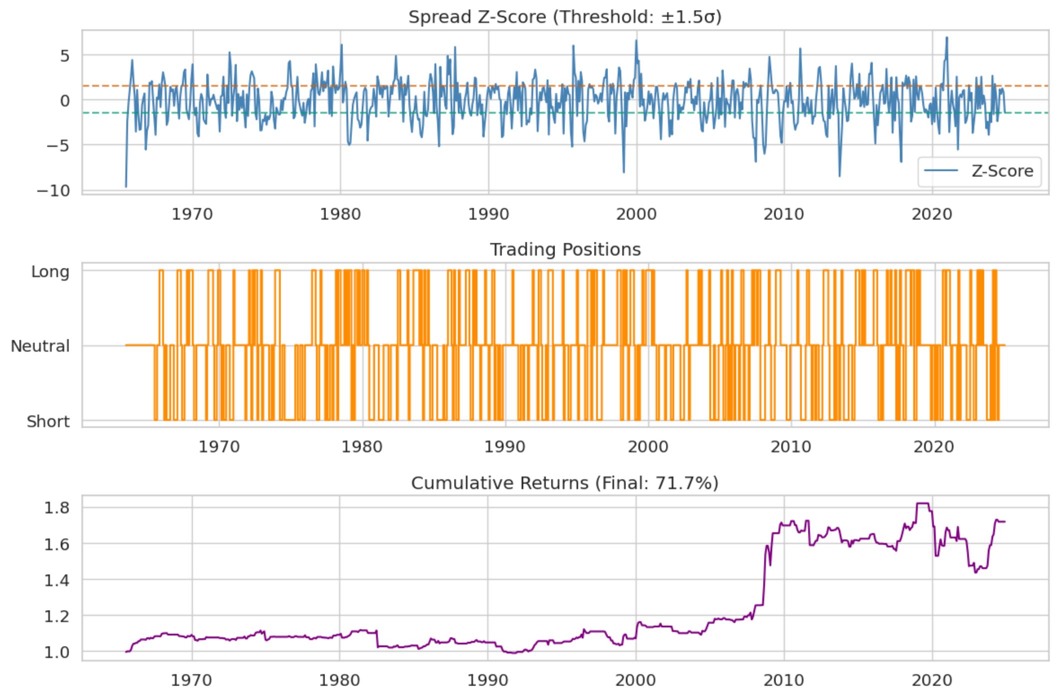


Figure 8: Star-Arb From 1963-07 to 2024-12

Table 10: Backtest Performance

Statistic	Ann-Ret	Ann-Vol	Ann-SR	Max Drawdown	Total Ret	Trading Times
strategy	0.91%	4.56%	0.200	-21.09%	71.71%	354

5.4 Optimal Mean-Variance Portfolio

The mean-variance portfolio optimization can be represented as:

$$\sigma^2 = \min_{w} \quad \frac{1}{2} w \top \Sigma w$$

subject to:

$$w \top \mu + (1 - w \top e) r_f = z, w \top e = 1$$

Where z is the target return level and w is the vector of portfolio weights. The constraints of the optimal problem provides an efficient frontier of such portfolios. Let $\mu_p = w \top \mu$ be the expected returns of the portfolio, and $\sigma_p^2 = w \top \Sigma w$ be the variance of the portfolio, the Sharpe ratio could be estimated as: $SR = \frac{\mu_p - r_f}{\sigma_p}$ with risk-free rate r_f.

To be specific with the weight vector w, the solution can be obtained by solving the Lagrangian optimal problem:

$$\mathcal{L}(w, \lambda) = \frac{1}{2} w \top \Sigma w - \lambda (w \top \mu + (1 - w \top e) r_f - z)$$

With the first order condition $\Sigma w - \lambda(\mu - er_f) = 0$ and budget constraints $(1 - w\top e)r_f = 1$ it derives:

$$w = \lambda\Sigma^{-1}(\mu - er_f)$$

In this section of the report, the original 6 portfolios are treated as independent assets, and Algorithm 2 is employed to solve for the optimal weights of this expanded portfolio. Since short-selling constraints do not apply in the U.S. stock market, the constraint in Line 14 of Algorithm 2 is disabled.

Algorithm 2 Mean-Variance Portfolio Optimization

```
 1: Input:
 2:     μ: Expected asset returns (vector)
 3:     Σ: Covariance matrix of returns
 4:     γ: Risk aversion parameter
 5:     r_f: Risk-free rate (optional)
 6: Output:
 7:     W: Optimal portfolio weights
 8:
 9: procedure MEANVARIANCE(μ, Σ, γ, r_f)
10:     Define objective function:
11:         max_w  w^⊤μ − (γ/2) w^⊤Σw                 ▷ Mean-variance utility
12:     Subject to:
13:         Σ w_i = 1                                  ▷ Full investment
14:         w_i ≥ 0                                    ▷ No short-selling (optional)
15:     Solve quadratic programming problem:
16:         W ← argmax (μ^⊤w − (γ/2) w^⊤Σw)
17:     return W                                       ▷ Optimal weights
18: end procedure
```

Perform backtesting according to the following settings:

① *Risk-free rate (r_f)*: Since one of the theoretical foundations of mean-variance portfolios is the classic CAPM model, the 10-year U.S. Treasury yield is adopted as an approximation of the risk-free rate for the U.S. stock market. According to FERD [8], this value is set at 4% in this report.

② *Weights constraints*: The backtest employs no leverage, with individual asset weights constrained between [-1, 1].

③ *Transaction costs*: The simulation excludes transaction costs, aiming to focus on how the Sharpe-optimized portfolio improves profitability and risk resilience compared to the equal-weighted benchmark portfolio.

④ *Profits*: Backtesting uses simple returns as the profitability metric, aiming to reflect the strategy's absolute profitability by excluding the impact of entry timing and holding periods on total returns. This setup also eliminates the effect of market timing on performance. If a strategy consistently delivers stable simple returns, its compounded returns will inherently be stronger. Conversely, such an approach also makes it less likely for compounded investments to experience significant drawdowns.

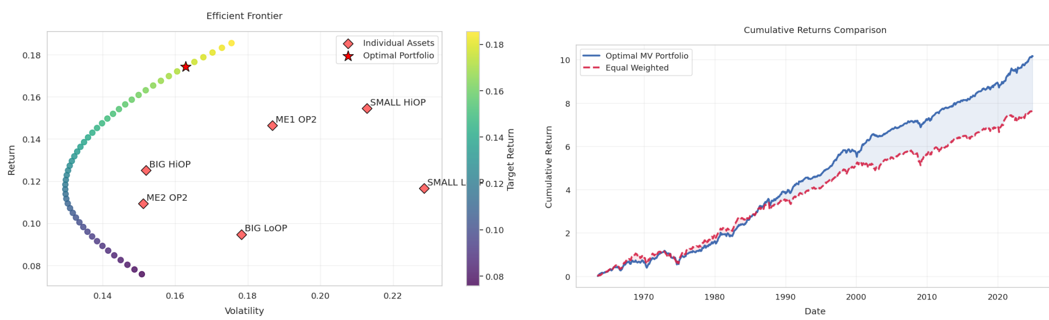
Table 11: Optimal Mean-Variance VS Equal Weight

strategy	Ann-Ret	Ann-Vol	Ann-SR	Max Drawdown	Total Ret
Mean-Variance	18.69%	17.72%	0.830	-75.98%	1149.72%
Equal Weight	12.45%	17.42%	0.490	-69.37%	765.63%

Table 10 reveals that the mean-variance optimized portfolio significantly outperforms the equal-weighted benchmark across all key metrics: delivering higher returns, while maintaining lower volatility and reduced maximum drawdown, which collectively contribute to its enhanced Sharpe ratio of 25%.

The optimal portfolio, denoted by the red star marker on the efficient frontier, demonstrates close proximity to assets including *BIG HiOP* and *ME2 OP2*, indicating these likely constitute substantial allocations given their superior risk-adjusted characteristics. Conversely, highly volatile assets such as *SMALL HiOP* and *SMALL LoOP* occupy distant positions from the efficiency frontier 9a, suggesting either negligible inclusion or complete exclusion from the optimized allocation owing to their disproportionate risk exposure. *BIG LoOP*, presenting neither competitive returns nor effective risk mitigation, is similarly anticipated to receive marginal weighting. *ME1 OP2*, while demonstrating intermediate risk-return properties, likely plays only a secondary role in the portfolio composition. This configuration reflects a rigorous optimization process that selectively incorporates assets demonstrating optimal risk-return efficiency while systematically minimizing exposure to suboptimal, high-volatility components.

The optimal mean-variance portfolio demonstrates consistent outperformance relative to the equally weighted benchmark; however, as evidenced in Figure 9b, it exhibits transient periods of underperformance. FERD [8] reveals that these episodes coincide with instances where the 3-year U.S. Treasury yield substantially exceeded the report's assumed risk-free rate. This interest rate anomaly created a pronounced dividend yield premium that adversely affected the portfolio optimization framework, introducing temporary distortions to the mean-variance efficiency.



(a) Efficient Frontier of Mean-Variance Portfolio

(b) Backtest

5.5 Improvement

The report introduces the Black-Litterman framework to calibrate the mean-variance portfolio with distributional assumptions:

① *Prior Distribution (Market Equilibrium)*:

$$r \sim \mathcal{N}(\Pi, \tau\Sigma), \Pi = \gamma\Sigma_{w_{mkt}}$$

where γ is the risk aversion coefficient, w_{mkt} is the market capitalization weights.

② *Views Distribution (Investor Opinions)*:

$$Q = Pr + \sigma, \sigma \sim \mathcal{N}(0, \Omega)$$

where Ω is a diagonal matrix of view confidences.

③ *Posterior Distribution*:

$$r|Q \sim \mathcal{N}(\hat{\mu}, \hat{\Sigma})$$

The report employs the framework provided in Algorithm 3 to incorporate posterior return estimates for adjusting the covariance matrix in Algorithm 2, and conduct playback testing and comparative analysis under identical configurations and constraints.

Algorithm 3 Black-Litterman Portfolio Calibration

Input:
  μ_{hist}: Historical asset returns (vector)
  Σ: Historical covariance matrix
  w_{mkt}: Market capitalization weights
  P: View pick matrix ($k \times n$)

Q: View return vector ($k \times 1$)
Ω: View confidence matrix ($k \times k$)
τ: Scaling constant (default=0.05)
γ: Risk aversion parameter
Output:
W: Optimal portfolio weights

procedure BLACKLITTERMAN($\mu_{hist}, \Sigma, w_{mkt}, P, Q, \Omega, \tau, \gamma$)
  Compute implied equilibrium returns:
    $\Pi \leftarrow \gamma \Sigma w_{mkt}$ ▷ Reverse optimization
  Compute posterior returns:
    $M \leftarrow \left[(\tau\Sigma)^{-1} + P^\top \Omega^{-1} P\right]^{-1}$
    $\hat{\mu} \leftarrow M \left[(\tau\Sigma)^{-1}\Pi + P^\top \Omega^{-1} Q\right]$
  Compute posterior covariance:
    $\hat{\Sigma} \leftarrow \Sigma + M$
  Solve mean-variance optimization:
    $W \leftarrow \text{argmax} \left(\hat{\mu}^\top w - \frac{\gamma}{2} w^\top \hat{\Sigma} w\right)$
    s.t. $\sum w_i = 1,\ w_i \geq 0$
  return W ▷ BL-optimized weights
end procedure

In the discussion of Section 4, it has been demonstrated that small-cap, high-OP portfolios are most likely to achieve long-term positive returns, followed by large-cap, high-OP portfolios. Therefore, it is reasonable to believe that they can significantly outperform other portfolios. Accordingly, one may set the investor's views as follows:

$$P = \begin{bmatrix} -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 0.05 \\ 0.03 \\ 0.02 \\ 0.01 \\ 0.04 \\ 0.03 \\ 0.025 \\ 0.02 \end{bmatrix}$$

As evidenced by the backtesting results presented in Figure 10 and Table 12, the Black-Litterman portfolio outperformed the mean-variance portfolio in terms of profitability through more substantial and aggressive weight allocations, thereby enhancing the original returns. However, correspondingly, during the high risk-free rate period in the early backtesting phase, it also exhibited larger drawdowns than the mean-variance portfolio. Consequently, while the Black-Litterman approach may not necessarily optimize the Sharpe ratio, employing a lower risk penalty coefficient can indeed improve annualized returns.

The results of this section demonstrate that successful market timing is also a crucial factor in improving the strategy's Sharpe ratio. Although the prevailing view in the industry is that the success rate of market timing is low, entering the market at a local peak of the strategy may lead to a prolonged period of continuous drawdown. This would significantly harm the investment experience and severely undermine the strategy's performance metrics, potentially misleading investors into underestimating the strategy's long-term potential.

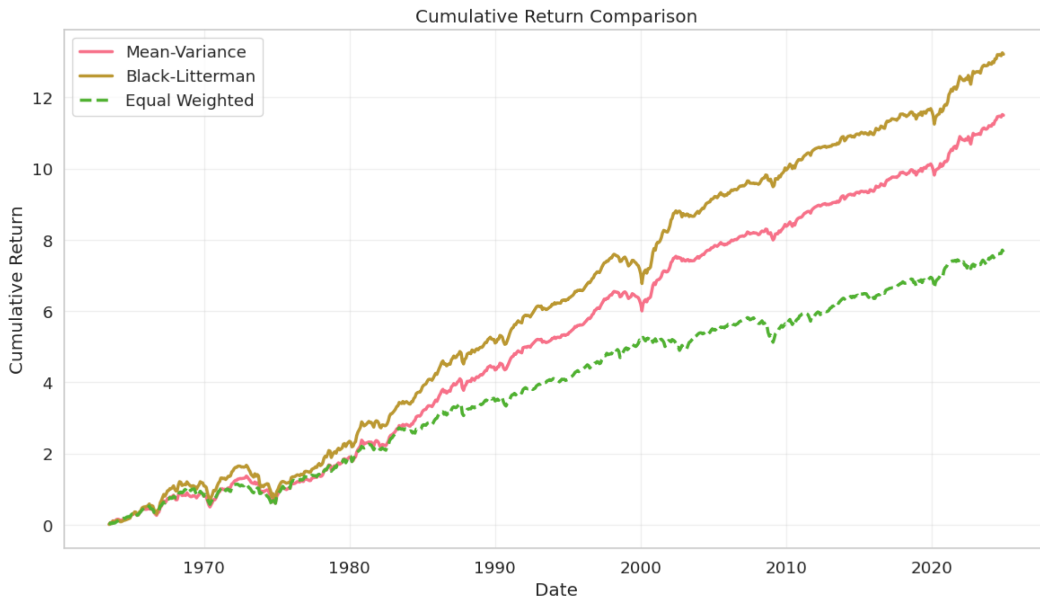


Figure 10: Black-Litterman Calibration

Table 12: Three Strategies

strategy	Ann-Ret	Ann-Vol	Ann-SR	Max Drawdown	Total Ret
Black-Litterman	21.48%	21.48%	0.810	-98.50%	1321.12%
Mean-Variance	18.69%	17.72%	0.830	-75.98%	1149.72%
Equal Weight	12.45%	17.42%	0.490	-69.37%	765.63%

6 Conclusions

The report begins by examining the monthly return characteristics of six investment portfolios categorized by market capitalization and operating profitability (OP). It conducts descriptive modeling and fitting of the return distributions, confirming that their monthly returns exhibit fat-tailed distributions and volatility clustering. Building upon this analysis, the report applies ARIMA-GARCH and several GARCH variants to model the time series properties of each portfolio's return series, specifically testing for stationarity and autocorrelation. The best-fitting results are achieved using GJR-GARCH, indicating that these return series are significantly affected by leverage effects. Notably, large-cap portfolios demonstrate greater susceptibility to downward movements in response to negative shocks.

Subsequently, this study examines the cointegration relationships among the six investment portfolios and constructs statistical arbitrage strategies. The results indicate that a simple rolling threshold-based signal framework for position entry/exit fails to deliver satisfactory profitability. Finally, this report investigates the long-term performance of mean-variance optimization compared to equal-weighted models. The results confirm that the Sharpe ratio serves as a valid objective function that can significantly enhance returns. However, this strategy also slightly amplifies the maximum drawdown, indicating that substantial improvements remain necessary before practical implementation.

Notably, even a simple constant-view Black-Litterman modification outperformed the mean-variance portfolio, demonstrating that incorporating posterior information has a positive effect on enhancing investment performance. Furthermore, when incorporating the risk-free rate into the optimization objective, dynamic adjustment of this parameter could be considered to achieve more efficient Sharpe ratio optimization.

References

[1] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

[2] Daniel B Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347–370, 1991.

[3] Lawrence R Glosten, Ravi Jagannathan, and David E Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801, 1993.

[4] Harry Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.

[5] Fischer Black and Robert Litterman. Global portfolio optimization. *Financial Analysts Journal*, 48(5):28–43, 1992.

[6] Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953, 2009.

[7] Robert Engle. Dynamic conditional correlation: A simple class of multivariate garch models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

[8] Federal Reserve Economic Data (FRED). 10-year treasury constant maturity rate (dgs10), 2024. Accessed: 2024-07-01.