Systematic multi-factor trading strategy based on SGX market

Cen, Yu

2023

Project No: 1134-221

Systematic Multi-Factor Trading Strategy

Based on SGX Market

Submitted by: Cen Yu

Supervisor: A/P Wong Jia Yiing, Patricia

School of Electrical & Electronic Engineering

A final year project report presented to the Nanyang Technological University

in partial fulfilment of the requirements of the degree of

Bachelor of Engineering

2023

Table of Contents

Abstract i

Acknowledgement ii

Acronyms.. iii

Chapter 1 Introduction.. 1

1.1 Backgrounds.. 1

1.2 Objectives.. 2

1.3 Scope.. 4

Chapter 2 Literature Review.. 10

2.1 Modern Portfolio Theory.. 10

2.2 Capital Asset Pricing Model.. 13

2.3 Fama-French Three Factor Model.. 14

Chapter 3 Methodology.. 16

3.1 Fama-Macbeth Regression.. 16

3.2 LASSO Regression.. 18

Chapter 4 Data Collection and Preprocessing.. 20

4.1 Technical Signals.. 20

4.2 Fundamental Signals.. 21

Chapter 5 Statistical Analysis.. 23

5.1 Cross-Sectional Analysis....................23

5.2 Time-Series Analysis....................26

Chapter 6 Strategy Construction and Back Test....................31

6.1 Introduction & Preparation....................31

6.2 Simple Ranking Strategy....................35

6.3 Signal Weighting Strategy....................46

6.4 Sharpe Ratio Optimization....................58

6.5 Market Beta Hedging....................67

6.6 Volatility Targeting and Risk Management....................75

Chapter 7 Conclusion....................84

Chapter 8 Reflection on Learning Outcome Attainment....................85

Reference....................88

Appendix....................89

Abstract

Numerous studies have been conducted over the years in an attempt to identify profitable trading strategies for financial markets. However, the high level of efficiency in larger markets, such as the US stock market, has led to such strategies being quickly exploited by investors and arbitrageurs. In contrast, smaller markets, such as the SGX, are thought to contain more inefficiencies, providing opportunities for profitable trading strategies.

The objective of this project is to develop a systematic trading strategy for SGX stocks, utilizing a combination of traditional approaches such as fundamental and technical analysis, as well as quantitative approaches like statistical analysis. The aim is to create a robust and profitable strategy that performs well in both back tests and real market conditions.

Acknowledgement

First of all, I would like to convey my sincere thanks to my final year project supervisor Prof Patricia Wong for giving me the chance to conduct this research on the SGX stock market. As I am currently doing an internship as a quantitative researcher, I really appreciate the opportunity for me to have a deeper understanding on multi-factor trading strategies.

Secondly, I would like to express my great gratitude to my internship supervisors. They guide me through the world of financial market and systematic trading, patiently teach me the basic concepts and techniques used in this highly competitive field. Without them, many ideas of my final year project will not be realized.

Cen Yu

February 2023

Acronyms

CAPM	Capital Asset Pricing Model
DR	Daily Return
EMH	Efficient Market Hypothesis
SGX	Singapore Exchange
STI	Straits Times Index
MPT	Modern Portfolio Theory
LASSO	Least Absolute Shrinkage and Selection Operator
SMB	Small Minus Big
HML	High Minus Low
CAGR	Composite Annual Growth Rate
KCP	Keltner Channel Percentage
WR	Williams %R
PSAR	Parabolic Stop and Reversal
VPT	Volume Price Trend

Chapter 1
Introduction

1.1Backgrounds

On the financial markets, there are different types of participants — retailers, speculators, long-term investors, hedge funds, etc. However, most of them share only one common goal, that is making money and managing risk. Ever since the existence of financial markets, hundreds of thousands of studies have been done to try to find profitable trading strategies. Among all the studies, fundamental analysis and technical analysis are by far the two most widely used categories of methods to speculate the market.

Fundamental analysis (FA) is a broad type of analysis method, aimed to measure the intrinsic value of stocks, examining all kinds of factors that could possibly affect the company and the overall market condition. Macroeconomic factors, including high level statistics such as GDP and CPI, and microeconomic factors, including the financial ratios of a specific company, are commonly used in the studies of financial analysis.

Technical analysis (TA) is another type of trading discipline. Different from financial analysis, technical analysis gathers past data from market activities, usually price and volume, to speculate the price trend of stocks. Various technical factors have been developed by researchers over the history of the stock market, some of them are as simple as moving averages, while some of them are much more complicated. Nevertheless, most of them can be calculated by detailed data of price and volume.

Systematic trading, or quantitative trading, usually refers to a way of setting trading rules and goals to make trading decisions more robust against discretionary human behaviors. Systematic trading often combines both fundamental analysis and technical analysis, while the latter is more commonly used. This is mainly due to the sparsity of fundamental data, as most of the economic or financial data are published on a quarterly, if not yearly basis. Thus, high frequency systematic trading strategies usually only use technical factors as the main input.

1.2Objectives

The objective of this project is to develop a profitable systematic trading strategy by incorporating both fundamental and technical factors. One of the key challenges faced by quantitative researchers is identifying effective factors that have robust statistical relationships with asset returns, thereby generating alpha. Classical models, such as the Capital Asset Pricing Model [1] and Fama-French 3-factor model [2], typically rely on only a few factors to predict asset returns. However, with decades of development, hundreds of signals have been discovered by quantitative researchers worldwide. Thus, determining the explanatory power of a signal for asset pricing has become a fundamental problem [3].

To achieve our objectives, we have divided the project timeline into several steps. The first step involves confirming the investment universe and signal universe, and preprocessing the data. In the second step, we will test the statistical significance of candidate signals, including technical and fundamental signals, using both cross-sectional and time-series analyses. In the third step, we will develop predictive models using the selected signals from the first step and compare the performance of various models. Finally, we will construct trading strategies based on the selected signals and models, conduct backtests, and provide recommendations.

1.3Scope

In this section, we are going to introduce the scope of this project, namely our investment universe and signal universe. Before we explain further, it is important for us to introduce the efficient market hypothesis (EMH).

1.3.1 Efficient Market Hypothesis

In financial economics, the efficient market hypothesis (EMH) posits that asset prices reflect all available information, leaving no excess return for investors to exploit [4]. Excess return refers to the additional profit earned above the market return.

EMH comprises three forms of efficient markets: weak, semi-strong, and strong. In a weak-form efficient market, asset prices reflect all historical information. In a semi-strong-form efficient market, asset prices reflect all public information, including both historical and current. In a strong-form efficient market, asset prices reflect not only public information but also private or insider information.

If the real financial market were strong-form efficient, neither technical nor fundamental analysis would be effective, as technical analysis examines historical information and fundamental analysis studies public information. However, in reality, the financial market is far from ideal and offers numerous opportunities for investors to exploit.

1.3.2 Investment Universe

Following the rationale of EMH, we may conclude that a financial market will have more opportunities for gaining excess return if it is less efficient. Thus, how to find a less efficient market would be an important problem to address.

Usually, a market with lower liquidity, smaller scale and shorter history will be less efficient. This is why apart from the financial market in developed countries, hedge funds nowadays also focus more and more on markets in developing countries, such as the Chinese market, Indian market, Indonesia market, etc.

Singapore Exchange Limited (SGX) was founded on 1 December, 1999. As of February 2022, it has 672 listed securities in total, and a total market capitalization of around 900 billion Singapore dollars. In comparison, the Nasdaq stock market has a market capitalization of 19.4 trillion US dollars and over 3500 listed securities. In terms of turnover, SGX only has a daily average volume of 1.6 billion Singapore dollars, while Nasdaq has a daily volume of 273 billion US dollars. Under any standard, it is valid to say that SGX is a relatively less efficient market than the US stock market.

However, we also do not want our target investment universe to be too inefficient. If a financial market is extremely illiquid, the transaction cost involved in trading activities could be so high such that it will offset much of the excess return that our trading strategies produce. Furthermore, the volume of our strategies could also be seriously capped since we will not be able to trade too much on an inefficient market.

Hence, although we choose SGX to study in our project, we do not want our target investment universe to include all the listed securities on SGX. In fact, we want the stocks in our universe to be as liquid as possible, as the Singapore stock market is already relatively inefficient.

The Straits Time Index (STI) is a market capitalization weighted index, which tracks the performance of top 30 companies listed on SGX. A capitalization-weighted index is a stock market index weighted by the market value of its component stocks. Under this rule, a company with higher market capitalization will have a larger impact to the index, while a company with lower value will receive less weight. The movement of such indices, such as S&P 500, NASDAQ 100, STI, is usually a benchmark for the performance of a financial market. Since STI contains the 30 top companies listed on SGX, which has the highest market capitalization and at same time the highest volume, the component stocks in STI will be the ideal investment universe for us.

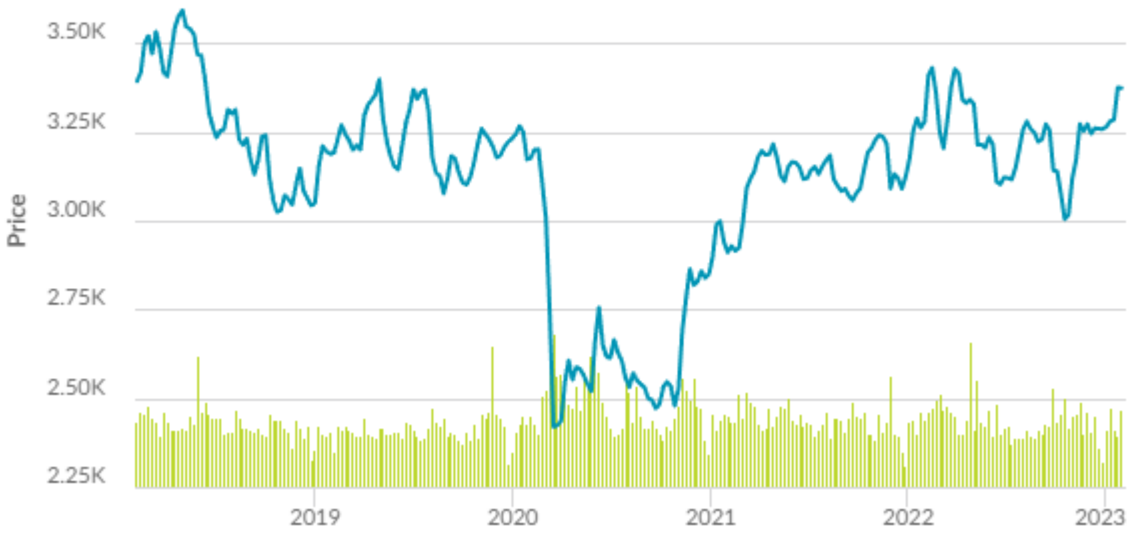


Figure 1.1 Recent Performance of STI

After careful examination, some of the components in STI either do not have enough historical data or have severe problems with liquidity. In the end, we decide to choose 25 of the 30 stocks as our investment universe.

1.3.3 Signal Universe

For the signal universe, we have both candidate signals from fundamental analysis and technical analysis. For the fundamental analysis part, we are going to use typical financial ratios retrieved from financial statements of these companies. The main data source we use here is Capital IQ. For the technical analysis part, we are going to use the python package TA. It will generate around 100 technical indicators based on the stocks open, high, low, close and volume data (OHLCV). These indicators are categorized into five types, momentum, trend, volume, volatility and other. Following is the introduction for some of the typical signals that we are going to test in this project.

1.3.3.1 Market Capitalization

Market capitalization refers to the total value of a company's outstanding shares of stock in the stock market. It is calculated by multiplying the current stock price of a company by the total number of its outstanding shares. Market capitalization is an important metric used by investors to evaluate a company's size and overall market value.

1.3.3.2 P/E Ratio

The P/E ratio, or price-to-earnings ratio, is a commonly used financial metric that compares the current market price of a company's stock to its earnings per share (EPS). The P/E ratio is calculated by dividing the current market price per share of a company's stock by its EPS over the previous 12 months.

1.3.3.3 P/B Ratio

The P/B ratio, or price-to-book ratio, is a financial metric used to compare a company's current market price to its book value per share. Book value per share is calculated by taking a company's total shareholder equity and dividing it by the number of outstanding shares. The P/B ratio is calculated by dividing a company's current market price per share by its book value per share. It is a measure of how much investors are willing to pay for a company's assets compared to their accounting value.

1.3.3.4 Williams %R

Williams %R is a technical indicator that measures overbought or oversold levels of an asset by comparing the current closing price to the high-low range over a specified period of time. It was developed by Larry Williams, a well-known trader and author.

Williams %R oscillates between 0 and -100, with readings between 0 and -20 indicating overbought conditions and readings between -80 and -100 indicating oversold conditions. The indicator is often used in conjunction with other technical indicators to confirm trend reversals or to identify potential entry and exit points for a trade.

Williams %R is calculated using the following formula:

Where:

Highest High is the highest price over the specified period

Lowest Low is the lowest price over the specified period

Close is the closing price of the current period

1.3.3.5 Keltner Channel Percentage

The Keltner Channel is a technical indicator that is used to identify trends and volatility in the market. It consists of three lines: a middle line, which is typically a simple moving average (SMA), and two outer bands, which are typically a certain number of average true range (ATR) values above and below the middle line.

The Keltner Channel Percentage is calculated using the following formula:

1.3.3.6 Volume Price Trend

The Volume Price Trend (VPT) is a technical indicator that is used to relate both price and volume. VPT is calculated based on cumulative volume and the trend of the stock price. Following is the formula to calculate VPT.

Chapter 2
Literature Review

2.1Modern Portfolio Theory

Modern Portfolio Theory (MPT) is an investment strategy that suggests that an investor can create a diversified portfolio with a mix of different assets that will generate the maximum expected return for a given level of risk [5]. The key idea behind MPT is that by combining different assets, an investor can reduce the overall risk of their portfolio while maintaining or increasing the expected return. This is because different assets tend to have different returns and risk levels, and combining them can lead to more stable overall portfolio performance. MPT also uses mathematical models to analyze and measure the risk and expected return of a portfolio, allowing investors to make informed decisions about the allocation of their assets.

The basic assumption of MPT is that investors are all risk-averse, which means that under a certain level of expected return, every investor would seek lower risk.

Expected return:

$$\mathrm{E}(R_p) = \sum_i w_i \, \mathrm{E}(R_i)$$

where is the return of the portfolio, E() is the expected return of asset i, is the proportion of asset i in the portfolio.

Portfolio return variance:

$$\sigma_p^2 = \sum_i w_i^2 \sigma_i^2 + \sum_i \sum_{j \neq i} w_i w_j \sigma_i \sigma_j \rho_{ij},$$

where is the standard deviation of the return of asset i in a specific period, is the correlation coefficient between the returns on asset i and asset j.

Since the correlation coefficient ranges between -1 to 1, it is clear that the more negatively correlated the returns of the assets, the lower their overall risk. This is the core rationale of portfolio diversification.

With the same asset universe and different weight, we can have a universe of portfolios all with different expected return and variance. In this universe, we can select all portfolios that have the highest expected return under every level of risk. These portfolios constitute the efficient frontier.

The Efficient Frontier is a concept in Modern Portfolio Theory that represents the boundary between portfolios that provide the maximum expected return for a given level of risk. In other words, it's a graphical representation of the optimal trade-off between risk and return for a set of investments. The efficient frontier shows the range of portfolios that an investor can choose from to maximize their expected return while minimizing their risk. Portfolios that fall on the efficient frontier are considered "efficient" because they offer the best possible return for a given level of risk. Portfolios that fall below the efficient frontier are considered sub-optimal because they have less return for the same amount of risk, while portfolios that fall above the efficient frontier are considered infeasible because they require an unrealistic level of risk for the expected return.

On the efficient frontier, there is one tangent portfolio, which has the highest ratio of expected return minus risk free rate versus standard deviation. This portfolio is usually considered as the optimal portfolio as it maximizes the expected risk-adjusted return. This ratio is the Sharpe ratio.

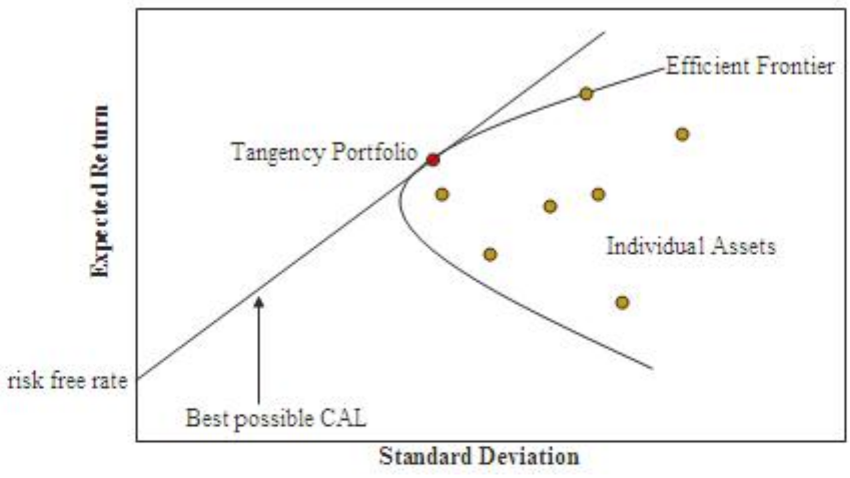


Figure 2.1 Graph Presentation of MPT

2.2Capital Asset Pricing Model

Following the proposition of MPT, later in the 1960s, many economists independently developed the famous Capital Asset Pricing Model (CAPM), including Jack Treynor, William F. Sharpe, John Lintner and Jan Mossin. CAPM is considered to be the foundation of most modern asset pricing models.

CAPM is a theoretical framework that calculates the expected return of an investment based on the risk-free rate, the expected market return, and the asset's systematic risk as measured by beta [6]. It is a widely used tool in finance for evaluating the expected performance of investments and for pricing securities. The same as MPT, the model assumes that investors are rational and risk-averse and that they require a higher expected return for taking on additional risk. It also assumes that all investors have the same view of the market and that the market is efficient. In simple terms, CAPM helps in determining the fair rate of return for an investment, taking into account the amount of risk involved.

Formula of CAPM:

$$\frac{E(R_i) - R_f}{\beta_i} = E(R_m) - R_f$$

or

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

where

$$\beta_i = \frac{\mathrm{Cov}(R_i, R_m)}{\mathrm{Var}(R_m)} = \rho_{i,m} \frac{\sigma_i}{\sigma_m}$$

The central idea of CAPM is that the expected return of every asset (or portfolio) can be explained by its sensitivity to the market excess return. If two assets have the same correlation with market return, the asset with higher risk will have the higher expected return. However, if an asset has a very high beta, it also has to tolerate higher risk when the market overall performs terribly.

2.3Fama-French Three Factor Model

Although the Capital Asset Pricing Model is widely recognized as the foundation of other modern asset pricing models, it only uses one factor to calculate the expected return, which is over-generalized and fails numerous empirical tests.

In 1993, Eugene Fama and Kenneth French designed the Fama-French three-factor model, which added two factors to the original CAPM[7]. The two factors are SMB and HML, respectively.

$$r = R_f + \beta(R_m - R_f) + b_s \cdot SMB + b_v \cdot HML + \alpha$$

In the model, SMB refers to “Small Minus Big”, which means the return difference between small cap stock and big cap stock; HML refers to “High Minus Low”, which means the return difference between high book-to-market ratio stock and low book-to-market ratio stock.

In their study, Fama and French discovered that two types of stocks tend to have better performance than the market, which is the stocks with small market capitalization and the stocks with higher book-to-market ratio. Stocks with higher book-to-market ratio are usually referred to as value stocks, compared to growth stocks. The three factor model explains over 90% of the asset expected return, which is much higher than the 70% explanatory power of the CAPM.

Later in 2015, Fama and French extended the model to a five-factor model[8]. The additional two factors are profitability and investment. Although the effectiveness of some of these factors is still being debated, the model that Fama and French proposed offers a benchmark for investors to construct their own multi-factor asset pricing models.

Chapter 3
Methodology

3.1Fama-Macbeth Regression

There are typically two perspectives of factor testing, cross-sectional test and time series test. For different types of factors, we will use different tests and give the corresponding explanations of doing so.

Previously, we have introduced that CAPM uses β to represent the market risk of an asset. Since then, β is widely used to represent the coefficient of any kind of general risk factors that are expected to have explanatory power to asset pricing. Fama-Macbeth regression [9] is a such a method to estimate the betas and risk premia of risk factors.

Fama-MacBeth regression is a statistical method used in finance to examine the relationship between asset returns and one or more factors that explain these returns. In this regression, a time-series of cross-sectional regressions is run to estimate the relationship between a dependent variable (e.g., stock returns) and one or more independent variables (e.g., market returns, size, value, etc.). This method allows for a more efficient estimation of the parameters than a traditional time-series regression and accounts for the potential cross-sectional dependence in the data.

There are two steps of regressions included in the Fama-Macbeth regression. The first regresses each asset’s return against multiple factors cross-sectionally to estimate each asset's exposure.

$$
\begin{aligned}
R_{1,t} &= \alpha_1 + \beta_{1,F_1} F_{1,t} + \beta_{1,F_2} F_{2,t} + \cdots + \beta_{1,F_m} F_{m,t} + \epsilon_{1,t} \\
R_{2,t} &= \alpha_2 + \beta_{2,F_1} F_{1,t} + \beta_{2,F_2} F_{2,t} + \cdots + \beta_{2,F_m} F_{m,t} + \epsilon_{2,t} \\
&\vdots \\
R_{n,t} &= \alpha_n + \beta_{n,F_1} F_{1,t} + \beta_{n,F_2} F_{2,t} + \cdots + \beta_{n,F_m} F_{m,t} + \epsilon_{n,t}
\end{aligned}
$$

The second regresses all asset returns for each time period against the previously estimated betas. The resulting matrix is then averaged to get the final estimation of risk premia.

$$
\begin{aligned}
R_{i,1} &= \gamma_{1,0} + \gamma_{1,1}\hat{\beta}_{i,F_1} + \gamma_{1,2}\hat{\beta}_{i,F_2} + \cdots + \gamma_{1,m}\hat{\beta}_{i,F_m} + \epsilon_{i,1} \\
R_{i,2} &= \gamma_{2,0} + \gamma_{2,1}\hat{\beta}_{i,F_1} + \gamma_{2,2}\hat{\beta}_{i,F_2} + \cdots + \gamma_{2,m}\hat{\beta}_{i,F_m} + \epsilon_{i,2} \\
&\vdots \\
R_{i,T} &= \gamma_{T,0} + \gamma_{T,1}\hat{\beta}_{i,F_1} + \gamma_{T,2}\hat{\beta}_{i,F_2} + \cdots + \gamma_{T,m}\hat{\beta}_{i,F_m} + \epsilon_{i,T}
\end{aligned}
$$

The Fama-MacBeth regression is widely used in empirical finance to study the factors that explain cross-sectional differences in stock returns and to test theories related to asset pricing. In our case, since we are not sure whether our candidate signals have explanatory power to the asset return, we need to first use Fama-Macbeth regression to test their statistical significance.

3.2LASSO Regression

LASSO (Least Absolute Shrinkage and Selection Operator) regression is a statistical method used for linear regression analysis that helps to address the problem of overfitting. In traditional linear regression, Ordinary Least Squares (OLS), the goal is to find the best-fitting line that minimizes the sum of squared residuals between the observed values and the predicted values. However, this method can lead to a complex model that fits the data too closely, capturing noise as well as signal.

$$y_i = \beta_1 \, x_{i1} + \beta_2 \, x_{i2} + \cdots + \beta_p \, x_{ip} + \varepsilon_i \, ,$$

LASSO regression solves this problem by adding a penalty term to the traditional linear regression objective function. This penalty term shrinks the magnitude of the coefficients of the less important variables towards zero, effectively excluding them from the model. The degree of shrinkage is controlled by a tuning parameter, commonly referred to as λ.

$$\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{N} \|y - X\beta\|_2^2 + \lambda \|\beta\|_0 \right\}$$

LASSO regression is useful for feature selection as well as for reducing the complexity of the model and improving its interpretability. It is also commonly used in high-dimensional data, where there are many potential independent variables and it is important to identify the most important predictors. In our case, this feature of LASSO regression can help us to filter the insignificant signals to avoid an over-complex model.

In this project, we will first use Fama-Macbeth regression to test the statistical significance of each candidate signal. Then, we will use LASSO regression to run time series regression on the selected signals to calculate the expected return for each asset on each day.

Chapter 4

Data Collection and Preprocessing

4.1 Technical Signals

First, we use python package yfinance to retrieve daily OHLCV (open, high, low, close, volume) data of the 25 stocks in Straits Times Index.

	Open	High	Low	Close	Volume
Date					
2012-01-03	7.461178	7.602683	7.454746	7.525498	2117000
2012-01-04	7.628410	7.647706	7.557658	7.628410	3115000
2012-01-05	7.634844	7.647708	7.583387	7.589819	1963000
2012-01-06	7.525497	7.660570	7.506201	7.628410	2583000
2012-01-09	7.583386	7.634843	7.506202	7.564090	2180000
...	...	...	...	...	...
2023-02-02	35.599998	35.720001	34.950001	35.080002	6961348
2023-02-03	35.360001	35.549999	35.060001	35.500000	4137400
2023-02-06	35.799999	35.900002	35.549999	35.840000	3581500
2023-02-07	36.000000	36.279999	35.959999	35.990002	4104700
2023-02-08	36.270000	36.270000	36.060001	36.169998	993300

Figure 4.1 Daily Data of DBS

Secondly, python package ta is used to generate 86 technical signals based on the OHLCV data. There are five types of technical signals: volume, volatility, trend, momentum and others. All the signals are generated automatically, the rationale and formula for calculating these signals will be introduced in the following chapter.

4.2 Fundamental Signals

Fundamental signals are retrieved from S&P Capital IQ database, including income sheet, balance sheet, market capitalization, and financial ratios on a quarterly basis.

DBS Group Holdings Ltd (SGX:D05) Financials > Historical Capitalization

Download This Page | Download Financials | Latest Annual | Latest Interim | FAQs | Create Activity | Add to Binder | 0 Items

Key Stats | Income Statement | Balance Sheet | Cash Flow | Multiples | Cap. Structure | Ratios | Supplemental | Industry Specific | Pension/OPEB | Segments

Frequency: Quarterly | Order: Latest on Right | Enable Freeze Panes

Currency: Trading Currency | Conversion: Historical | Go | More Options >>

'98 '99 '00 '01 '02 '03 '04 '05 '06 '07 '08 '09 '10 '11 '12 '13 '14 '15 '16 '17 '18 '20 '21 View All

In Millions of the trading currency, except per share items.

Balance Sheet as of:	Dec-31-2019	Jun-30-2020	Dec-31-2020	Jun-30-2021	Dec-31-2021	Jun-30-2022
Pricing as of*	Mar-09-2020	Aug-05-2020	Mar-08-2021	Aug-04-2021	Mar-08-2022	Aug-03-2022
Currency	*SGD*	*SGD*	*SGD*	*SGD*	*SGD*	*SGD*
Capitalization Detail						
Share Price	21.15	19.83	28.40	30.58	31.23	32.33
Shares Out.	2,555.6	2,539.1	2,556.2	2,568.2	2,573.3	2,573.5
Market Capitalization	54,050.3	50,349.7	72,595.5	78,534.5	80,364.5	83,201.3
Book Value of Common Equity	50,981.0	53,438.0	54,626.0	57,594.0	57,526.0	55,875.0
+ Book Value of Pref. Equity	-	-	-	-	-	-
= Total Equity	51,799.0	54,259.0	54,643.0	57,771.0	57,714.0	56,066.0
Tier 1 Capital	45,460.0	48,051.0	48,188.0	51,315.0	51,640.0	51,842.0
Tier 2 Capital	5,233.0	-	5,749.0	-	6,567.0	-
Total Capital	50,693.0	53,482.0	53,937.0	57,895.0	58,207.0	58,213.0
Tier 1 Capital Ratio %	15.0%	14.9%	15.0%	15.5%	15.1%	14.9%
Total Capital Ratio %	16.7%	16.6%	16.8%	17.5%	17.0%	16.7%
Leverage Ratio %	-	-	-	-	-	-

Figure 4.2 Historical Capitalization Data of DBS

Since the frequency of technical factors is on a daily basis, while the frequency of fundamental factors is on a quarterly (or yearly) basis, the data cleaning process and statistical inference process are different for the two categories.

	2012-12-31	2013-12-31	2014-12-31	2015-12-31	2017-01-03	2018-01-02	2018-12-31	2019-12-31	2020-12-31	2021-12-31
A17U.SI	5.883097	8.652858	8.170367	9.415011	12.840556	14.007218	12.711351	16.064806	16.603257	17.383548
BN4.SI	5.751533	7.657881	6.195232	5.924137	10.806948	56.708706	9.612896	15.310068	100	8.681677
BS6.SI	3.477157	5.221149	4.953716	6.648052	7.245986	8.167274	7.051463	6.434318	7.00068	6.840071
C07.SI	9.064809	7.495538	10.447191	11.025752	13.683755	10.325022	20.880203	8.731608	9.978168	8.866782
C09.SI	16.197436	12.086396	11.659333	8.714514	11.272451	21.424159	13.185836	17.752466	100	72.117339
C38U.SI	8.287201	7.467949	7.73169	8.228773	10.652611	9.094121	10.033885	11.423944	24.04543	11.919787
C6L.SI	32.55546	28.090703	33.691749	32.122641	12.975172	32.887939	8.245886	15.461319	100	100
D01.SI	24.104279	19.109115	18.226249	15.162788	16.741449	21.876688	122.543902	21.259603	19.522718	36.434718
D05.SI	6.361118	7.954585	9.12712	6.930201	7.970191	11.623112	9.049847	9.09775	12.913339	11.975118
F34.SI	9.424939	9.339759	10.317729	10.805071	14.101247	10.089354	10.93808	13.495069	13.660336	9.868654
G13.SI	23.14703	25.101949	20.875966	102.061466	34.838742	22.931212	14.009635	15.067813	144.657388	50.053344
H78.SI	8.020584	8.416542	8.882792	6.227237	3.489337	2.361063	5.022081	58.546381	100	100
J36.SI	9.956099	9.242957	10.144873	7.969583	6.674205	4.871523	13.335524	6.631578	100	8.798082
M44U.SI	6.187599	7.684839	6.403878	6.854366	9.690935	13.906384	6.445348	11.731024	17.984136	16.245696
ME8U.SI	5.298871	4.572757	4.950243	4.703539	7.719041	10.457026	9.53271	15.758381	14.954513	36.029353
N2IU.SI	2.943859	4.465919	5.545513	6.012892	7.36937	9.518305	6.921548	10.333002	10.903423	91.370022
O39.SI	5.727449	9.22886	7.477448	7.005479	8.566212	10.644184	8.950358	8.638541	11.800561	10.17866
S58.SI	13.791157	14.353363	15.911075	19.122536	22.108947	20.997848	19.324737	22.792895	26.8	100
S63.SI	13.440673	14.377404	13.951628	12.791276	15.452558	15.784136	17.966516	18.035724	21.16839	19.530129
S68.SI	17.719407	16.576593	19.488138	18.315019	17.693562	19.630693	18.419693	22.003114	19.900976	21.862527
U11.SI	8.628834	8.721218	9.512034	7.991254	9.008035	11.252346	9.107212	9.241275	12.511324	10.841548
U14.SI	4.530823	4.939826	6.486936	10.556826	14.322432	7.218294	11.122658	13.377997	480.405861	19.067606
U96.SI	9.307656	9.24826	7.83299	8.465764	12.001161	13.649214	12.987785	13.420124	20.756671	13.110929
V03.SI	9.324519	10.067336	19.02204	10.482025	11.526699	12.38029	8.966541	11.058635	17.90286	16.410248
Z74.SI	8.506035	11.218908	11.983087	11.381087	11.544421	11.868965	7.520002	16.108039	33.415011	66.782504

Figure 4.3 PE Ratio Data for 25 stocks

Chapter 5

Statistical Analysis

5.1 Cross-sectional Analysis

In previous chapters, we have obtained all the candidate signals we are planning to test. In this chapter we will first test the statistical significance of these signals cross-sectionally using Fama-Macbeth regression.

For every single day, we regress each asset return against all the signals. Thus, we have 25 data points each day. This will give us very rough estimations about the real value of the beta for each signal. However, since we are using 10 years data, we have over 2000 this kind of beta estimates for each of them. Thus, we can take the mean value and standard deviation to calculate the t-statistics for every signal.

Here, we will only show the signal that has an absolute t-statistics larger than 1.65, which means that the signal is at least 90% statistically significant. The full testing result will be shown in the appendix.

	mean	std_error	tstat
const	7.155305e-07	3.197605e-07	2.237708
volume_adi	3.173622e-12	1.260145e-12	2.518458
volume_obv	-1.514741e-12	7.045296e-13	-2.150004
volume_vpt	-4.975933e-09	2.484339e-09	-2.002920
volatility_bbp	-6.054552e-06	1.506918e-06	-4.017836
volatility_kcp	-1.249597e-05	4.319628e-06	-2.892833
volatility_dcl	-4.875365e-06	2.903431e-06	-1.679174
trend_vortex_ind_neg	3.762517e-06	2.107820e-06	1.785028
trend_dpo	-6.062420e-06	2.516366e-06	-2.409196
trend_adx_neg	-7.185507e-05	2.918650e-05	-2.461928
trend_cci	9.457750e-06	5.158484e-06	1.833436
trend_psar_down_indicator	-1.241840e-05	7.246310e-06	-1.713754
momentum_rsi	-8.944417e-05	2.651041e-05	-3.373926
momentum_stoch_signal	5.621210e-05	1.946618e-05	2.887680
momentum_wr	-7.758708e-05	1.561972e-05	-4.967251
momentum_pvo	2.874176e-05	1.240984e-05	2.316046
others_dr	-2.970545e-05	1.599878e-05	-1.856732
others_dlr	-2.878745e-05	1.595720e-05	-1.804042

Table 5.1 Signals with significant t-stats

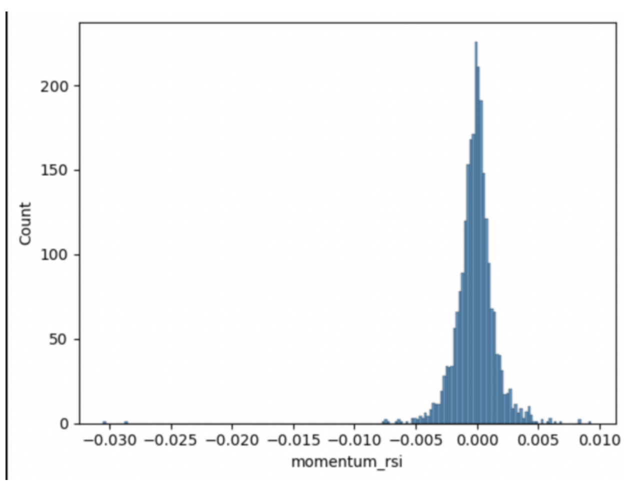


Figure 5.1 RSI (t-stats = -3.37)

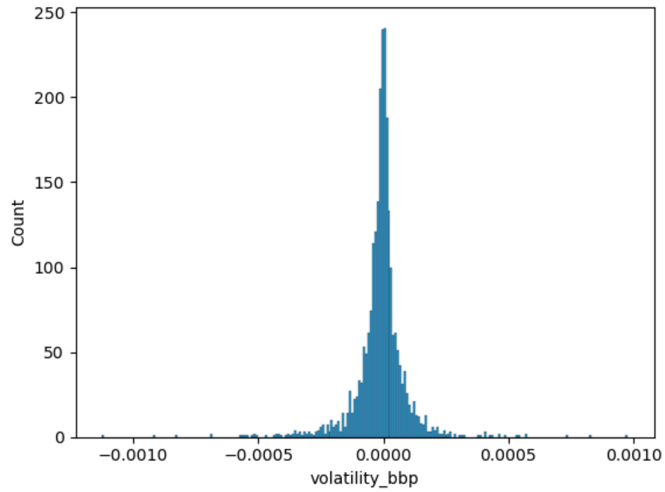


Figure 5.2 Bollinger Band Percentage,BBP (t-stats = -4.02)

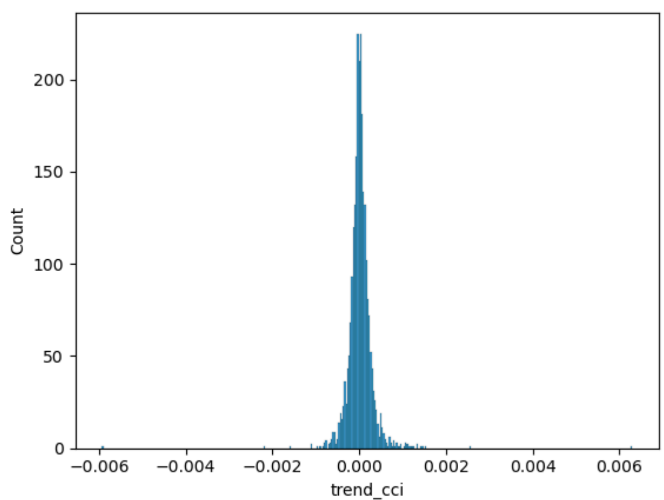


Figure 5.3 Commodity Channel Index, CCI (t-stats = 1.83)

From the histograms, it seems that the mean values of the coefficients of these signals are very close to zero, even though they are all statistically significant based on their t-stats.

In order to utilize this kind of tiny edge, we will use multiple techniques in the strategy construction period.

5.2 Time-series Analysis

In the previous section we have tested our candidate signals cross-sectionally. The signals that are statistically significant were selected. However, the Fama-Macbeth regression only tells us the overall significance of the signals on the whole investment universe, we still do not know how these signals will perform on each individual stock. Thus, we need to use time-series regression to get a more detailed view.

Firstly, we run OLS on each individual signal for the period between 2012 and 2022. In each regression, there are two independent variables, market return and the signal being tested. The reason is that market return usually has major explanatory power on the return of individual stock. If we run regression using only the candidate signals, the testing result could be distorted and the p-values are more likely to be significant.

Another important problem to address is that all the signals may have very different scales, especially the signals related to trading volume. Hence, normalizing the signals on a time-series basis before regression is usually a good practice.

After the first step, we have a summary dataframe including the time-series p-values for every signal on each stock. It is quite obvious that not all signals are significant for every stock. In fact, only a few signals are significant on every individual stock. This implies that we cannot use all the candidate signals to construct our strategy, even if they are valid in the Fama-Macbeth regression.

Ticker	volume_adi	volume_obv	volatility_bbp	volatility_kcp	volatility_dcp	trend_dpo	trend_adx_neg	trend_cci	trend_psar_up_indicator	momentum_rsi
A17U.SI	0.679	0.887	0.000	0.000	0.000	0.526	0.006	0.001	0.524	0.000
BN4.SI	0.106	0.969	0.050	0.068	0.307	0.013	0.194	0.038	0.338	0.309
BS6.SI	0.308	0.556	0.865	0.373	0.562	0.578	0.587	0.750	0.094	0.757
C07.SI	0.954	0.001	0.643	0.589	0.398	0.512	0.631	0.647	0.442	0.270
C09.SI	0.387	0.445	0.435	0.365	0.275	0.647	0.448	0.689	0.690	0.231
C38U.SI	0.515	0.875	0.000	0.000	0.000	0.065	0.162	0.016	0.043	0.001
C6L.SI	0.184	0.039	0.075	0.031	0.012	0.033	0.965	0.567	0.832	0.055
D01.SI	NaN	0.732	0.410	0.065	0.906	0.888	0.202	0.127	0.391	0.257
D05.SI	0.646	0.425	0.073	0.059	0.188	0.476	0.090	0.093	0.091	0.160
F34.SI	0.961	0.768	0.574	0.212	0.386	0.029	0.666	0.930	0.023	0.421
G13.SI	0.841	0.807	0.344	0.091	0.522	0.321	0.837	0.641	0.231	0.474
H78.SI	0.679	0.059	0.073	0.011	0.036	0.861	0.426	0.151	0.555	0.103
J36.SI	0.460	0.387	0.052	0.003	0.046	0.879	0.222	0.199	0.941	0.017
M44U.SI	0.557	0.260	0.000	0.000	0.000	0.185	0.647	0.041	0.126	0.001
ME8U.SI	0.773	0.589	0.000	0.000	0.000	0.072	0.007	0.000	0.329	0.000
N2IU.SI	0.804	0.532	0.000	0.000	0.000	0.031	0.265	0.001	0.970	0.000
O39.SI	0.358	0.374	0.735	0.800	0.444	0.377	0.566	0.853	0.283	0.393
S58.SI	0.227	0.296	0.023	0.002	0.003	0.859	0.801	0.337	0.218	0.008
S63.SI	0.900	0.653	0.029	0.004	0.077	0.853	0.162	0.151	0.016	0.010
S68.SI	0.567	0.866	0.043	0.097	0.011	0.529	0.138	0.104	0.544	0.015
U11.SI	0.536	0.545	0.997	0.530	0.599	0.653	0.399	0.925	0.314	0.457
U14.SI	0.486	0.736	0.486	0.189	0.542	0.789	0.333	0.963	0.018	0.355
U96.SI	0.389	0.648	0.498	0.094	0.340	0.531	0.406	0.735	0.715	0.253
V03.SI	0.293	0.950	0.546	0.818	0.229	0.379	0.762	0.964	0.252	0.813
Z74.SI	0.151	0.263	0.014	0.003	0.016	0.935	0.039	0.100	0.721	0.016

Table 5.2 Time-series p-values for every signal(not fully shown) on each stock

For the second step, we will employ LASSO regression that we introduced before to select the most significant signals for each category. The rationale behind this step is that the correlation between signals under the same category is usually high. Thus, their explanatory power will be similar. If we use multiple similar signals in one strategy, the strategy will be biased toward that specific type of signal. The drawback of this problem is that, in certain market periods where this type of signal does not work very well, the performance of the strategy will be affected seriously.

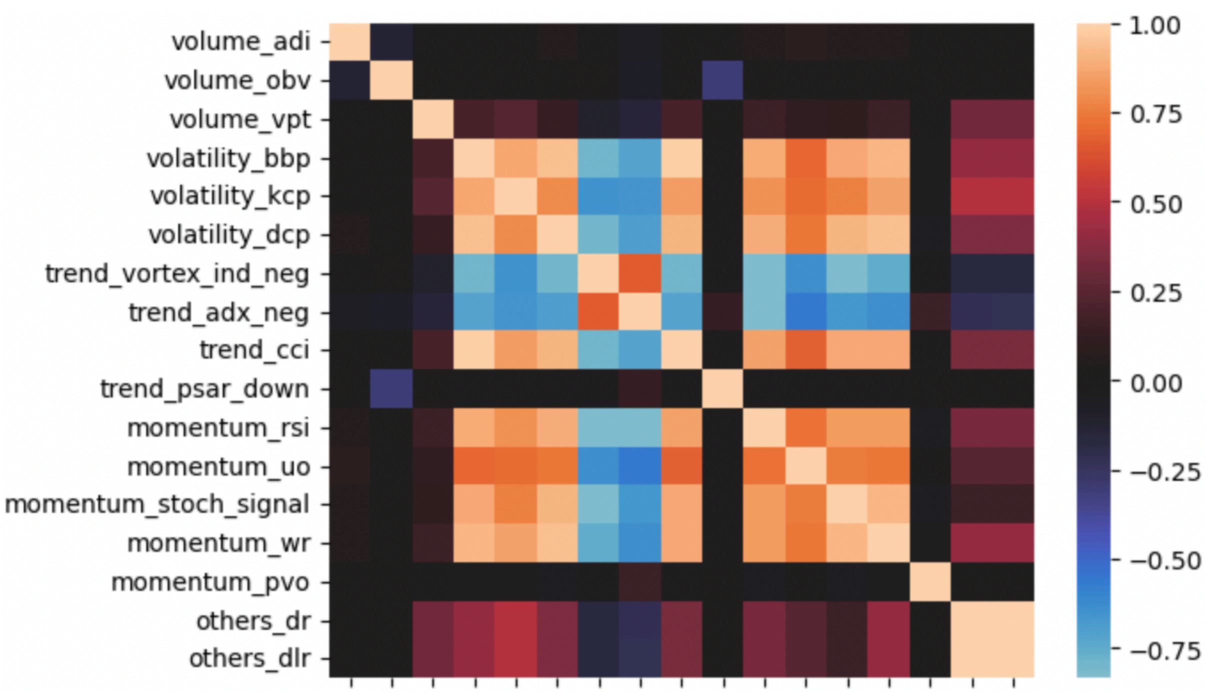


Figure 5.4 The heatmap of correlation between signals

Signal	Coefficient
volatility_bbp	0.012
volatility_kcp	0.510
volatility_dcp	0.160
momentum_rsi	0.028
momentum_uo	0.316
momentum_stoch_signal	0.000
momentum_wr	0.439
momentum_pvo	0.116
trend_vortex_ind_neg	0.029
trend_adx_neg	0.005
trend_cci	0.178
trend_psar_down	0.280
volume_adi	0.063
volume_obv	0.084
volume_vpt	0.533

Table 5.3: Lasso regression coefficients (alpha = 0.03)

From the regression results, the strongest signals from each category are volatility_kcp, momentum_wr, volume_vpt and trend_psar_down. Since trend_psar_down is paired with trend_psar_up, we will present it as trend_psar in the next chapter.

Chapter 6

Strategy Construction and Back Test

6.1 Introduction & Preparation

In this chapter, we will use the candidate signals that we have tested statistically to construct systematic trading strategies.

To begin constructing and backtesting our strategies, we must first establish a benchmark for comparison. As we focused on a subset of stocks that exhibit certain characteristics, we decided to use an equally weighted buy-and-hold portfolio as our benchmark, since we did not include all the stocks inside the Straits Times Index. This benchmark portfolio achieved a cumulative return of 100% over the entire backtesting period. However, during the COVID-19 outbreak in 2020, the portfolio experienced a substantial drawdown, with a maximum drawdown of over 30%.

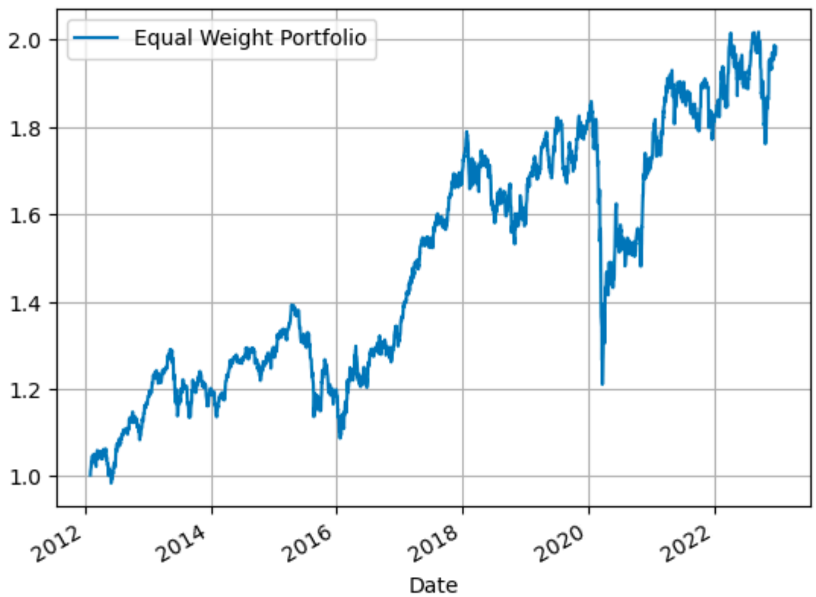


Figure 6.1 10-year performance of the Equal Weight Buy-and Hold Portfolio

To further quantify and measure the performance of this benchmark as well as our later strategies, we will introduce three indicators, namely Sharpe Ratio, Compound Annual Growth Rate and Max Drawdown.

Here, the "Return on Investment" is the average return earned by an investor on a particular investment, the "Risk-Free Rate" is the rate of return that an investor can earn with no risk, such as a Treasury bill or a savings account, and the "Standard Deviation of Investment" is a measure of the volatility or risk of the investment. For convenience, we usually take the Risk-Free Rate as 0.

A higher Sharpe Ratio indicates that an investment has provided a better return for the amount of risk taken, whereas a lower Sharpe Ratio indicates that an investment has provided a lower return for the amount of risk taken. The Sharpe Ratio can be used to compare different investment opportunities or to evaluate the performance of a particular investment over time.

In this formula, "Ending Value" represents the total value of the investment at the end of the period, "Beginning Value" represents the total value of the investment at the beginning of the period, and "Number of Years" represents the length of the investment period.

CAGR takes into account the compounding effect of investment returns over time, which means that it considers the fact that investment returns earned in one year will themselves earn returns in subsequent years. This compounding effect can have a significant impact on the overall return of an investment over time.

In this formula, "Peak Value" represents the highest value of the investment during the period under consideration, and "Trough Value" represents the lowest value of the investment during the same period. Max drawdown is often used by investors and analysts to evaluate the risk of an investment or a portfolio, as it provides an estimate of the potential maximum loss that an investment could experience during a market downturn.

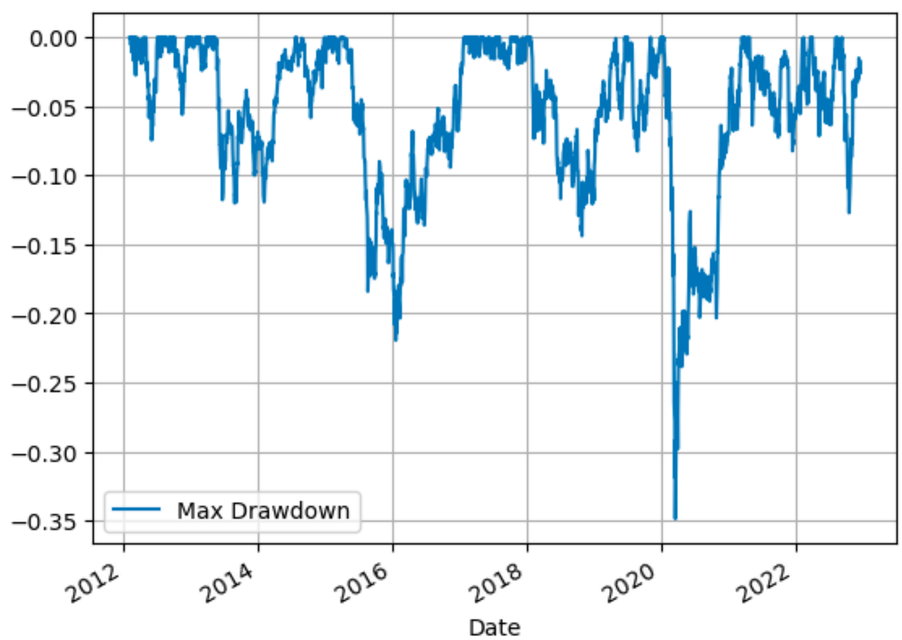


Figure 6.2 Max Drawdown of benchmark portfolio

By the definitions above, we can calculate the three performance indicators for the benchmark portfolio.

Portfolio	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
Benchmark	0.545	6.24	34.87

Table 6.1 Performance indicators for benchmark portfolio

6.2 Simple Ranking Strategy

In this section, we will start to use the signals that we have selected in the previous chapter to construct trading strategies. The first and also the simplest type of strategy that we are going to test is the ranking strategy, which means that we are going to long (short) the stocks with top (bottom) n rank on the signal, while their weights will be evenly distributed.

6.2.1 Volatility_kcp

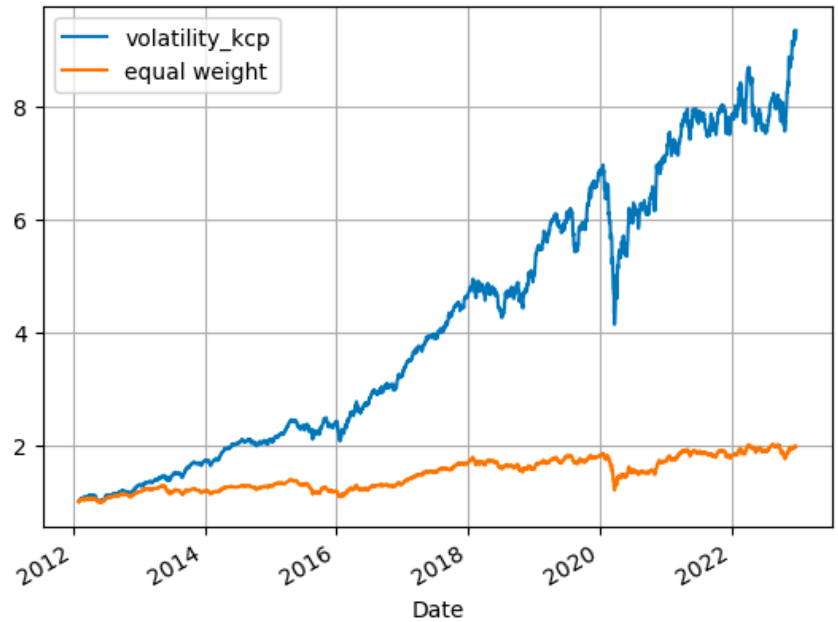


Figure 6.3 Cumulative Return, KCP, rank=5

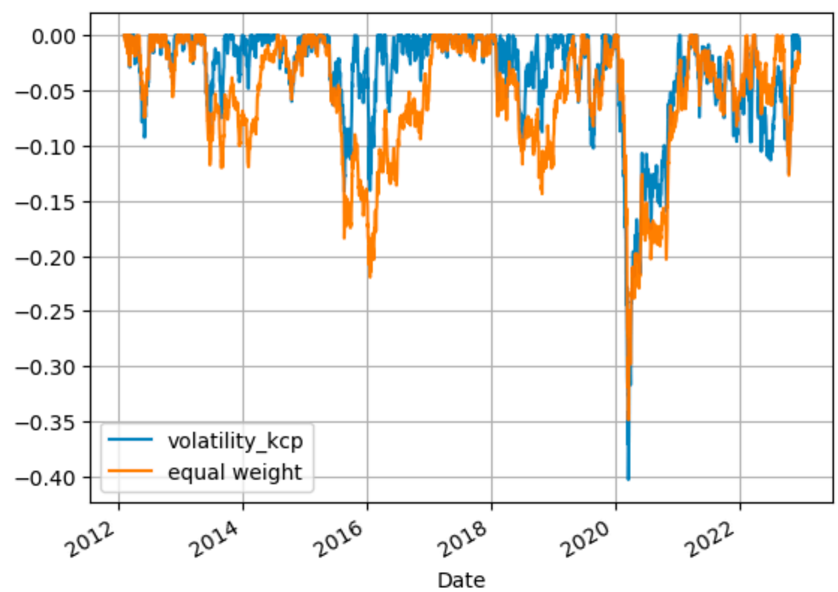


Figure 6.4 Max Drawdown, KCP, rank = 5

Rank	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
3	0.942	17.62	47.26
4	1.111	19.98	44.61
5	1.161	19.83	42.46
6	1.248	20.68	42.90
7	1.405	22.84	40.53
8	1.399	22.15	40.56
9	1.333	20.39	39.44
10	1.290	19.40	40.31

Table 6.2 Performance metrics of KCP, simple ranking

6.2.2 Momentum_wr

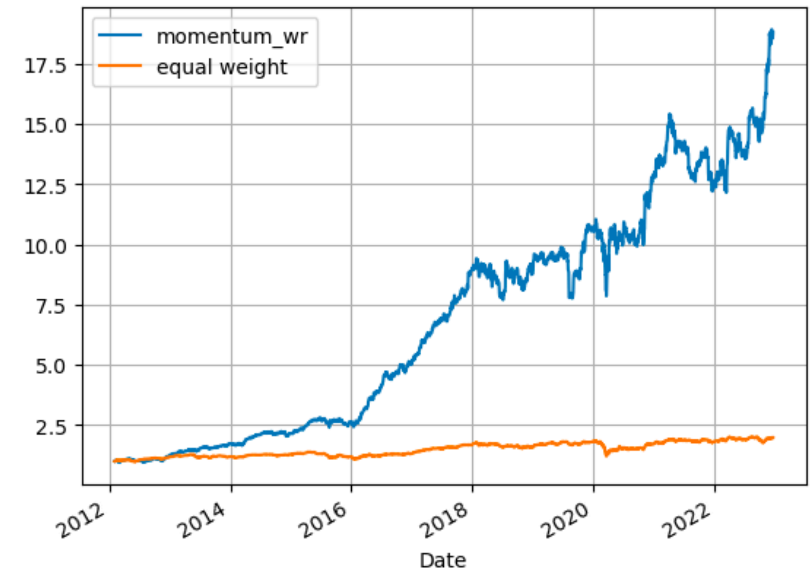


Figure 6.5 Cumulative Return, WR, rank=2

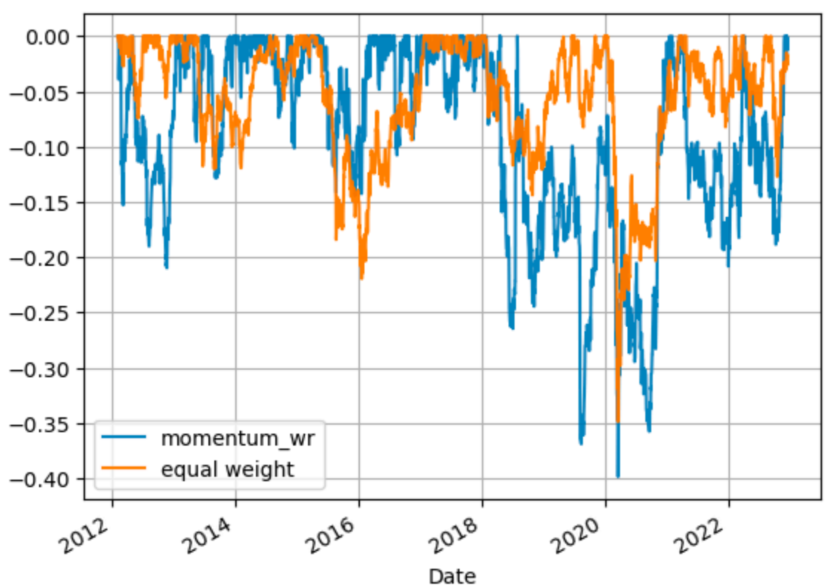


Figure 6.6 Max Drawdown , WR, rank=2

Rank	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
1	1.109	26.92	39.85
2	1.470	31.02	28.83
3	1.387	27.30	42.57
4	1.365	24.26	39.04
5	1.348	22.53	38.33
6	1.316	21.16	38.33
7	1.187	18.20	40.07

Table 6.3 Performance metrics of WR, simple ranking

6.2.3 Volume_vpt

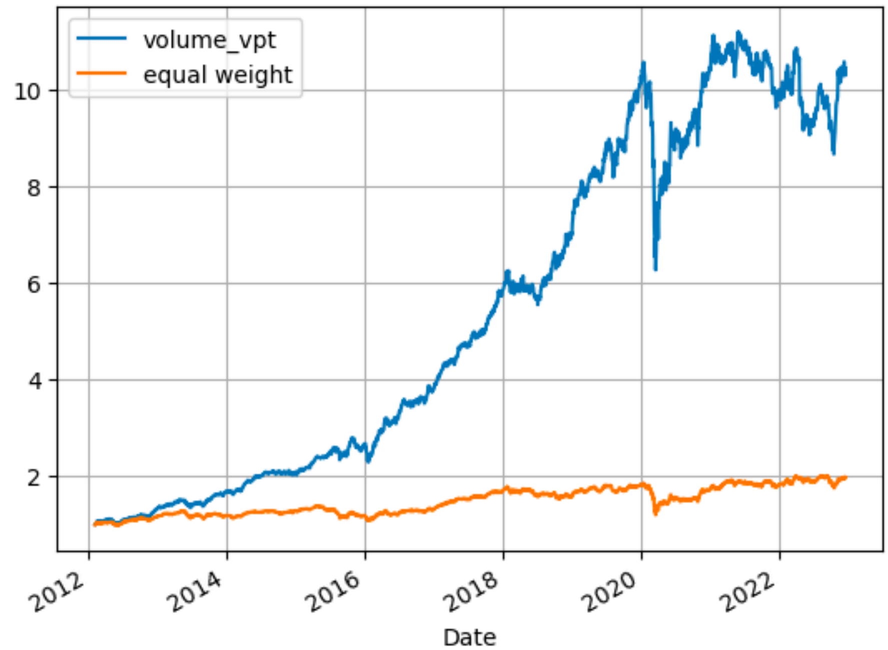


Figure 6.7 Cumulative Return, VPT, rank=6

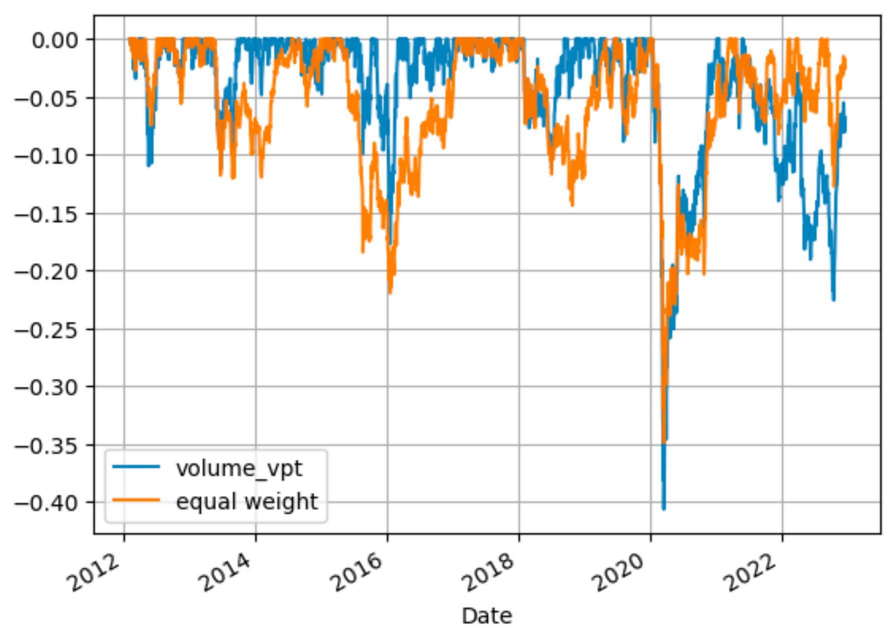


Figure 6.8 Max Drawdown , VPT, rank=6

Rank	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
3	1.165	22.03	38.56
4	1.336	23.84	40.24
5	1.383	23.88	42.34
6	1.462	24.08	40.63
7	1.336	20.98	42.47
8	1.330	20.23	41.51

Table 6.4 Performance metrics of VPT, simple ranking

6.2.4 Trend_psar

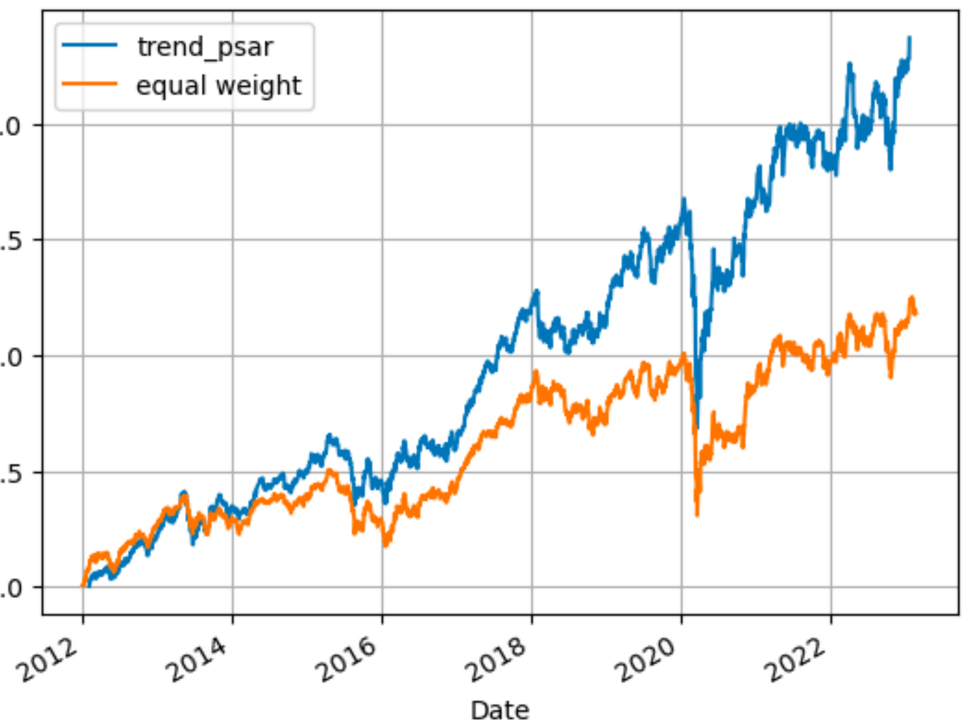


Figure 6.9 Cumulative Return, PSAR, rank=5

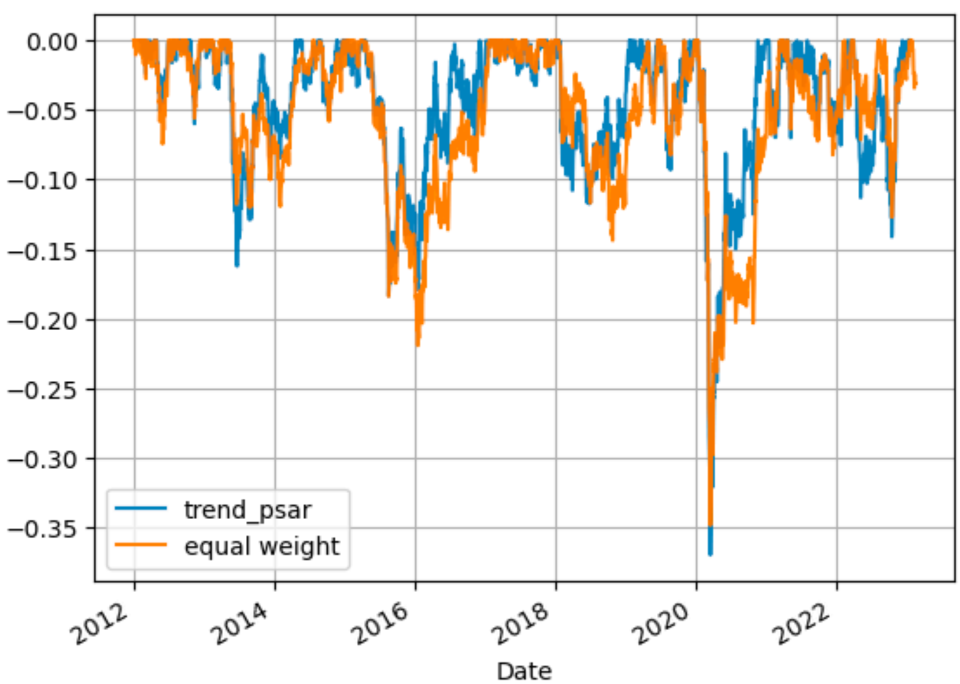


Figure 6.10 Max Drawdown , PSAR, rank=5

Rank	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
3	0.86	15.43	31.76
4	0.819	13.42	42.87
5	0.809	12.24	40.8
6	0.786	11.26	40.34
7	0.824	11.55	38.06
8	0.858	11.75	37.02

Table 6.5 Performance metrics of PSAR, simple ranking

6.2.5 others_dr

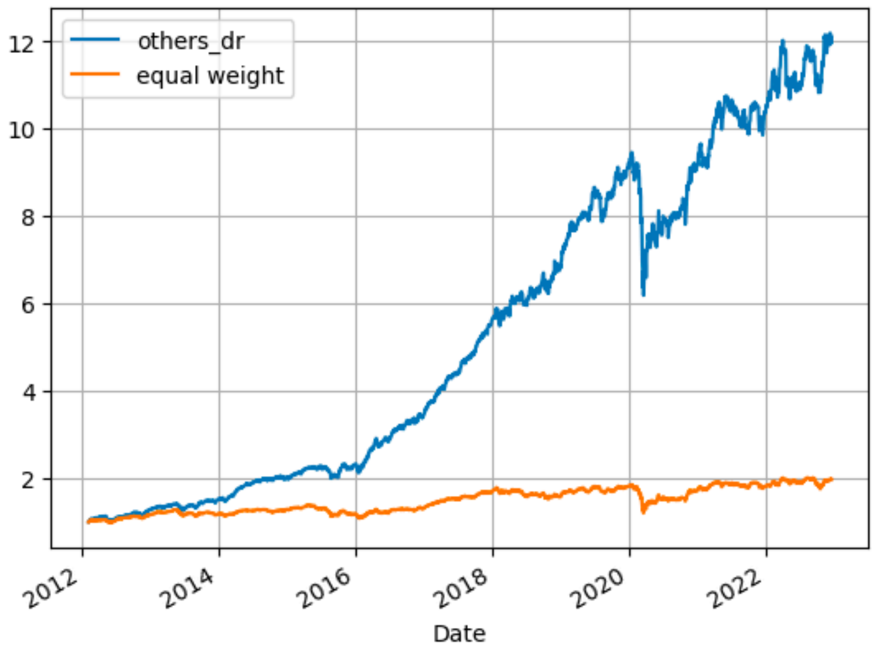


Figure 6.11 Cumulative Return, DR, rank=8

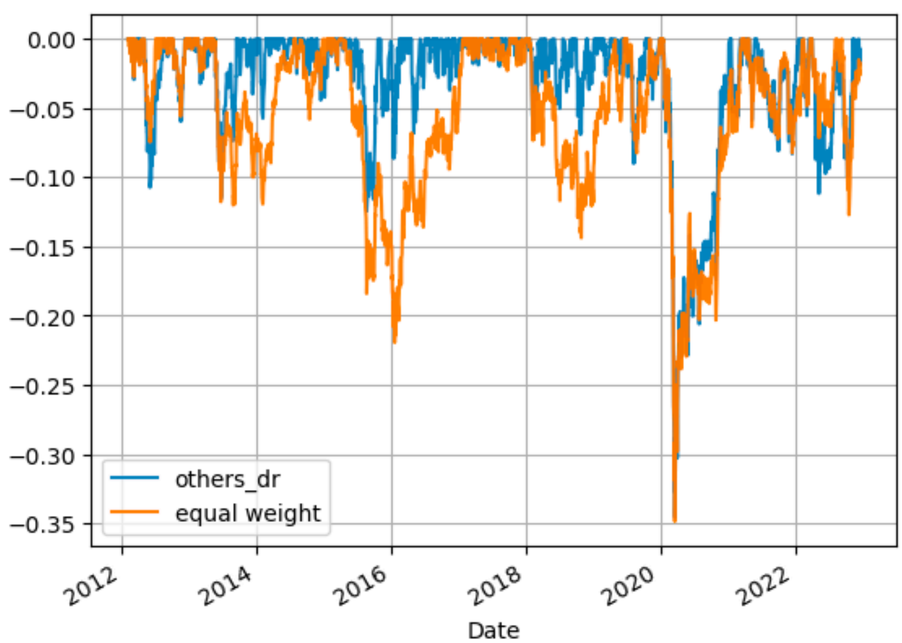


Figure 6.12 Max Drawdown , DR, rank=8

Rank	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
1	1.292	38.83	45.81
2	1.478	34.61	38.44
3	1.540	30.80	34.59
4	1.594	29.11	34.84
5	1.647	28.29	33.81
6	1.659	27.19	34.79
7	1.662	26.16	34.27
8	1.681	25.81	34.58
9	1.676	25.08	34.05
10	1.599	23.34	34.18

Table 6.6 Performance metrics of DR, simple ranking

6.2.6 Section Summary

From the testing result, we can get an initial conclusion that volatility_kcp, momentum_wr, volume_vpt and others_dr are effective signals. With reasonable rank selected, their performance is much better than the original equal weight strategy. However, trend_psar is not as effective in this simple ranking strategy. From the cumulative return plot, we can find that the trend signal seems to strengthen the trend of the original strategy. Thus, its CAGR and Max Drawdown both increased proportionally.

On the other hand, the testing result also shows that, although Sharpe Ratio and CAGR increase, it is relatively hard to improve Max Drawdown by this simple ranking strategy. This means that the risk is not well controlled. A major reason is that since we are trading stocks based on their signal ranking, we literally only hold a portion of our investment universe. This will inevitably increase the risk of our strategy. In the following sections, we will try to solve this problem.

6.3 Signal Weighting Strategy

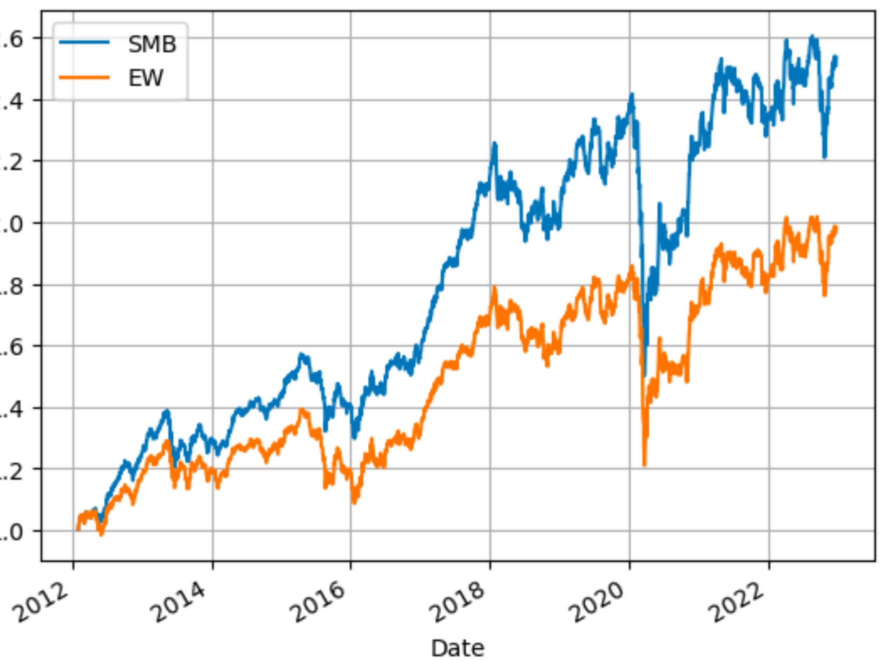
In this section we are going to weight the stocks by the value of signals. However, since different signals have different values, we will need to transform the signals accordingly.

6.3.1 Profit/Earnings Ratio (P/E) and Market Capitalization (MC)

In the previous chapter we introduced the Fama-French three-factor model. In the formula there are two fundamental signals, HML and SMB, which are corresponding to value stocks and small-cap stocks. In this project, we will use Profit/Earnings Ratio (P/E) and Market Capitalization (MC) to represent these two signals.

Based on the conclusion of the Fama-French three-factor model, stocks with smaller MC and lower PE usually have higher return. Hence, we will first take the reciprocal of the signals, then weight the stocks proportionally.


HML
EW
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
2012
2014
2016
2018
2020
2022
Date


SMB
EW
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
2012
2014
2016
2018
2020
2022
Date

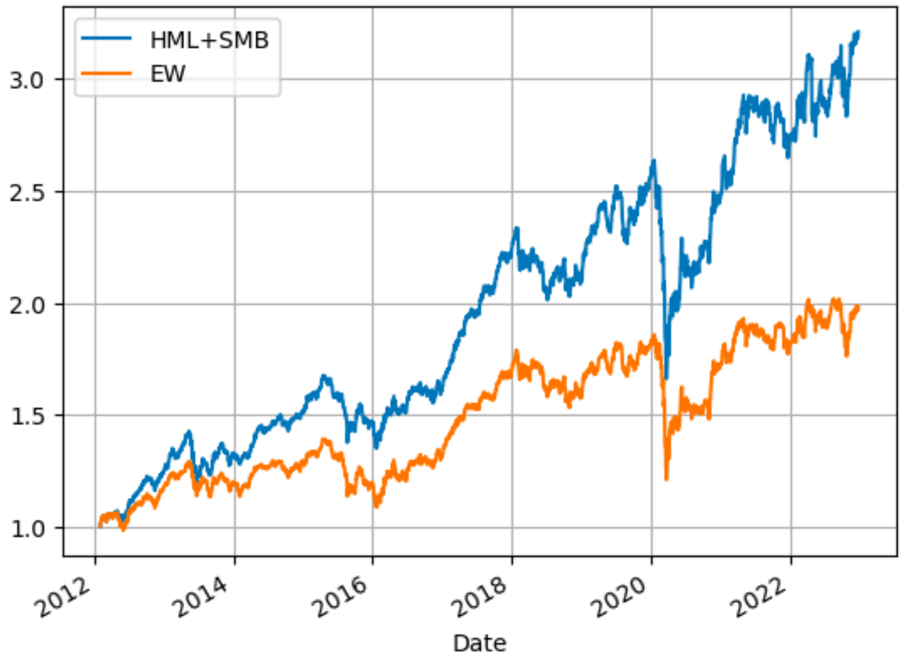


Figure 6.13 Cumulative Return , SMB,HML,signal weighting

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
HML	0.710	8.50	34.15
SMB	0.734	8.94	37.88
HML+SMB	0.88	11.34	37.00

Table 6.7 Performance metrics of HML,SMB, signal weighting

Based on the result, we can find that these two fundamental signals are both effective. However, compared to technical signals, their pre-cost performance is not as good. This is because the frequency of these two types of signals is very different. Fundamental signals are updated quarterly or yearly, while technical signals are updated daily or even more frequently. When combined both fundamental signals together, they collectively perform better.

6.3.2 Volatility_kcp

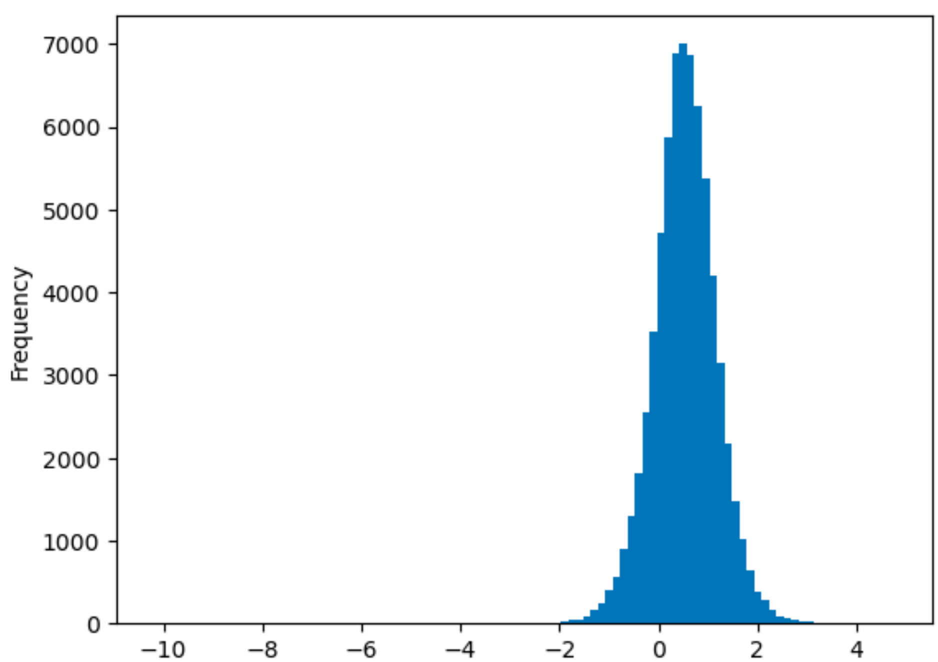


Figure 6.14: Distribution of KCP

Signal transformation:

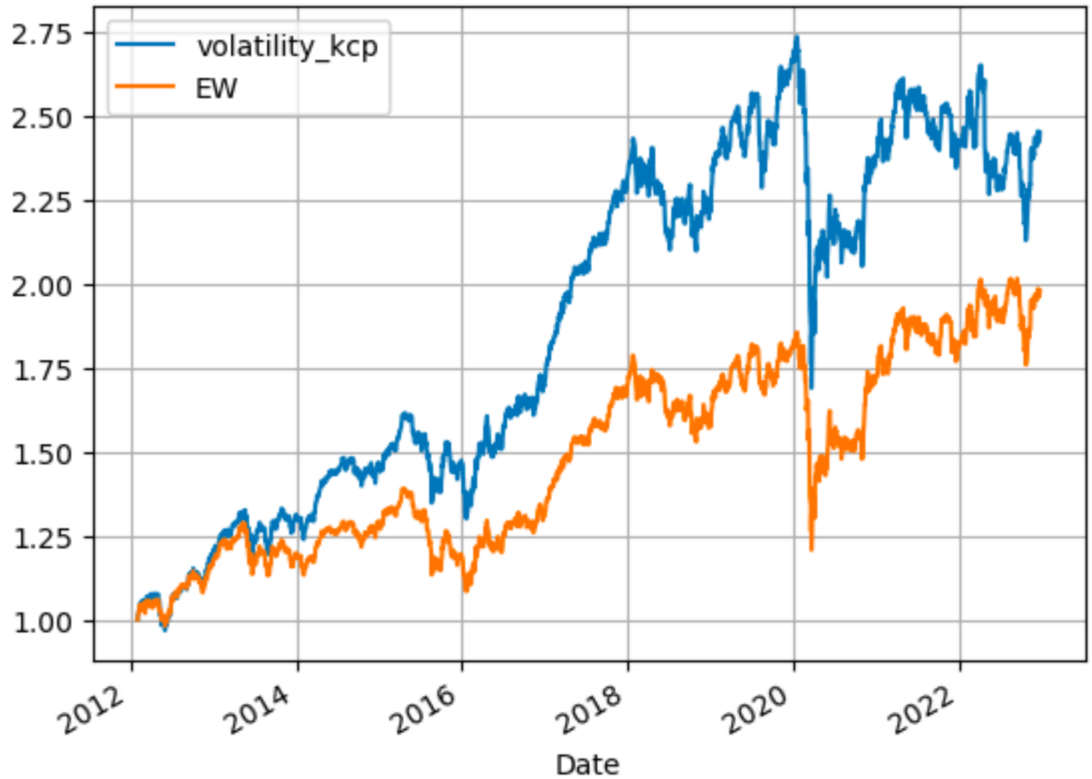


Figure 6.15 Cumulative Return, KCP, signal weighting

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
volatility_kcp	0.677	8.61	38.19

Table 6.8 Performance metrics of KCP, signal weighting

6.3.3 Momentum_wr

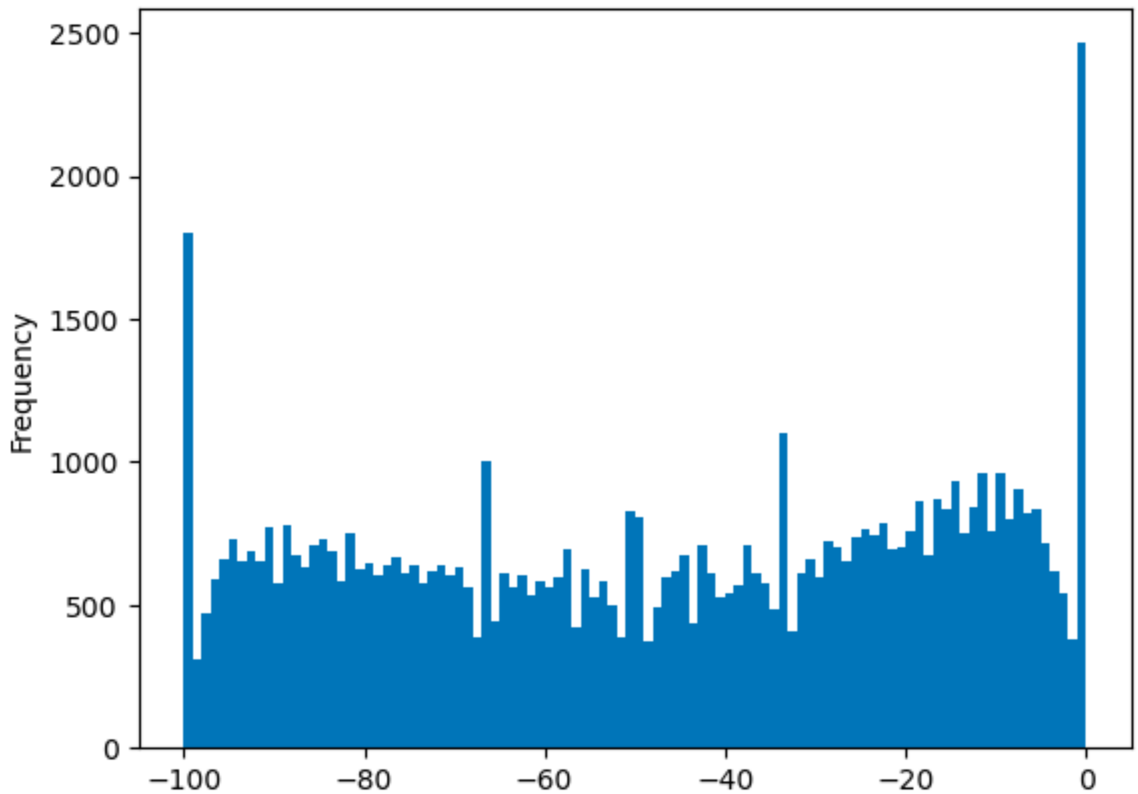


Figure 6.16: Distribution of WR

Signal transformation:

Figure 6.17 Cumulative Return, WR, signal weighting

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
momentum_wr	0.995	13.03	38.47

Table 6.9 Performance metrics of WR, signal weighting

6.3.4 Trend_psar

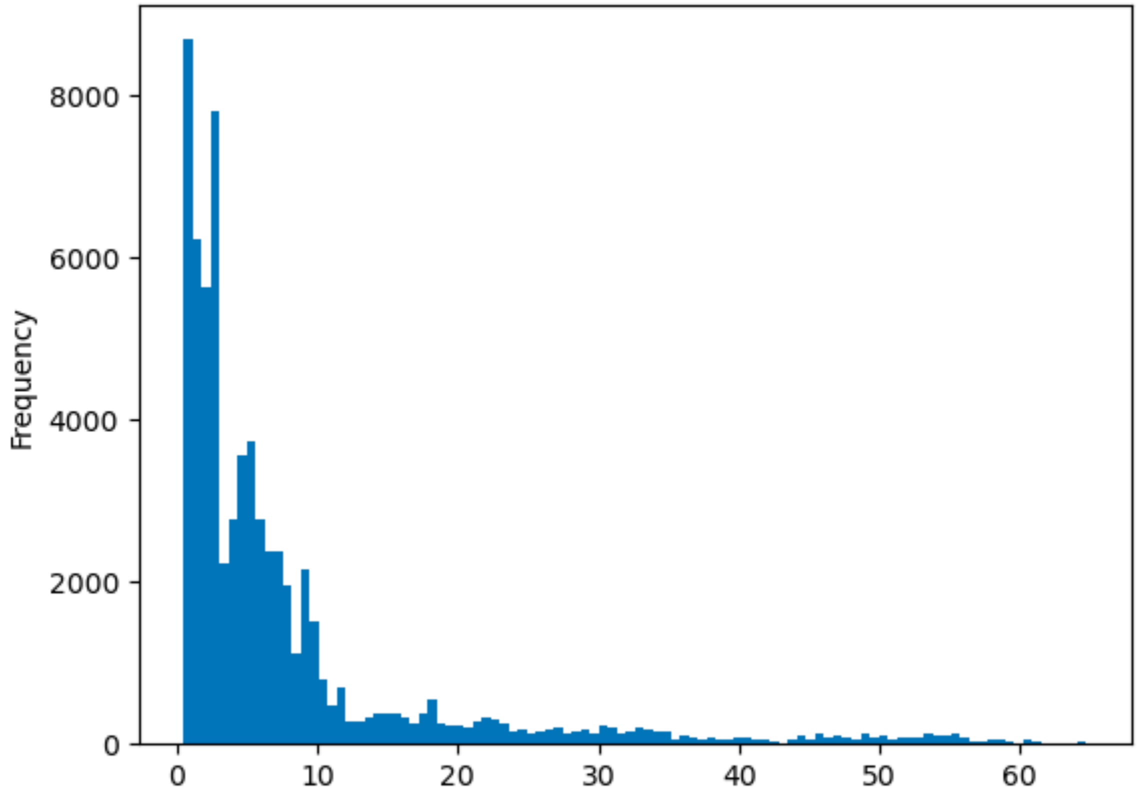


Figure 6.18: Distribution of PSAR

Signal_transformation:

Figure 6.19 Cumulative Return, PSAR, signal weighting

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
trend_psar	0.888	12.52	36.14

Table 6.10 Performance metrics of PSAR, signal weighting

6.3.5 Others_dr

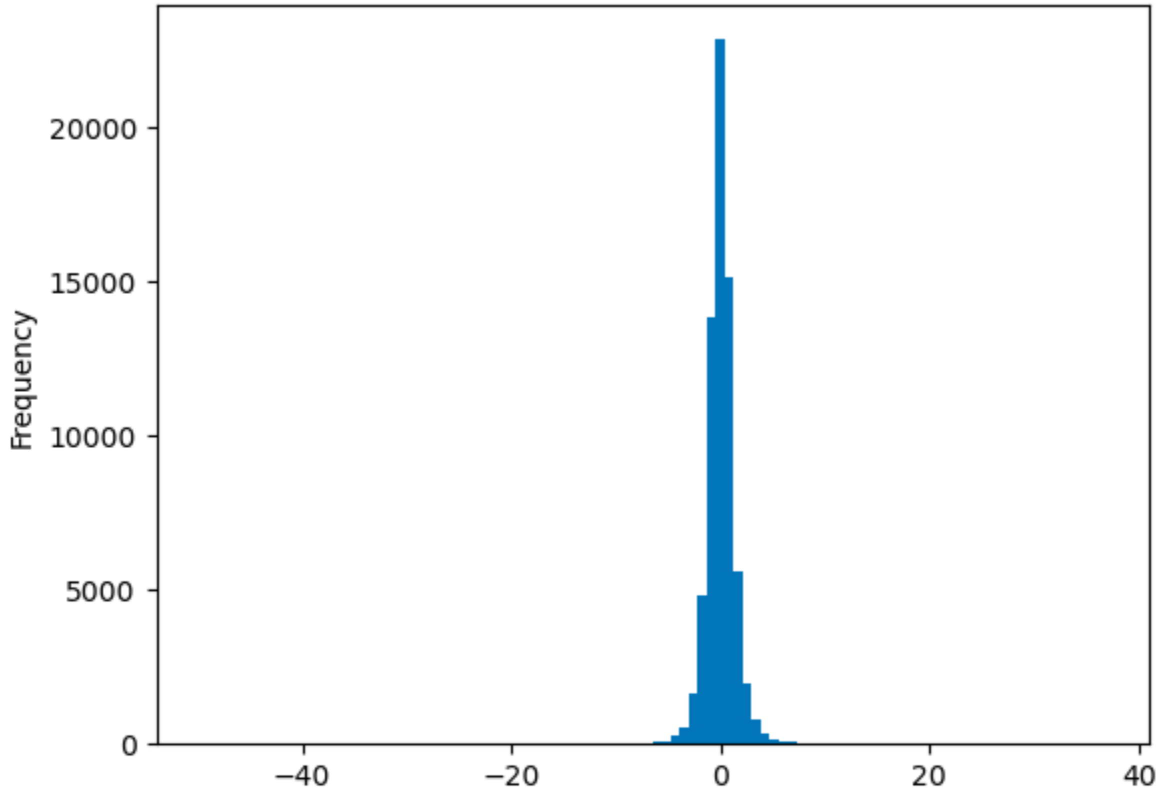


Figure 6.20: Distribution of DR

Signal_transformation:

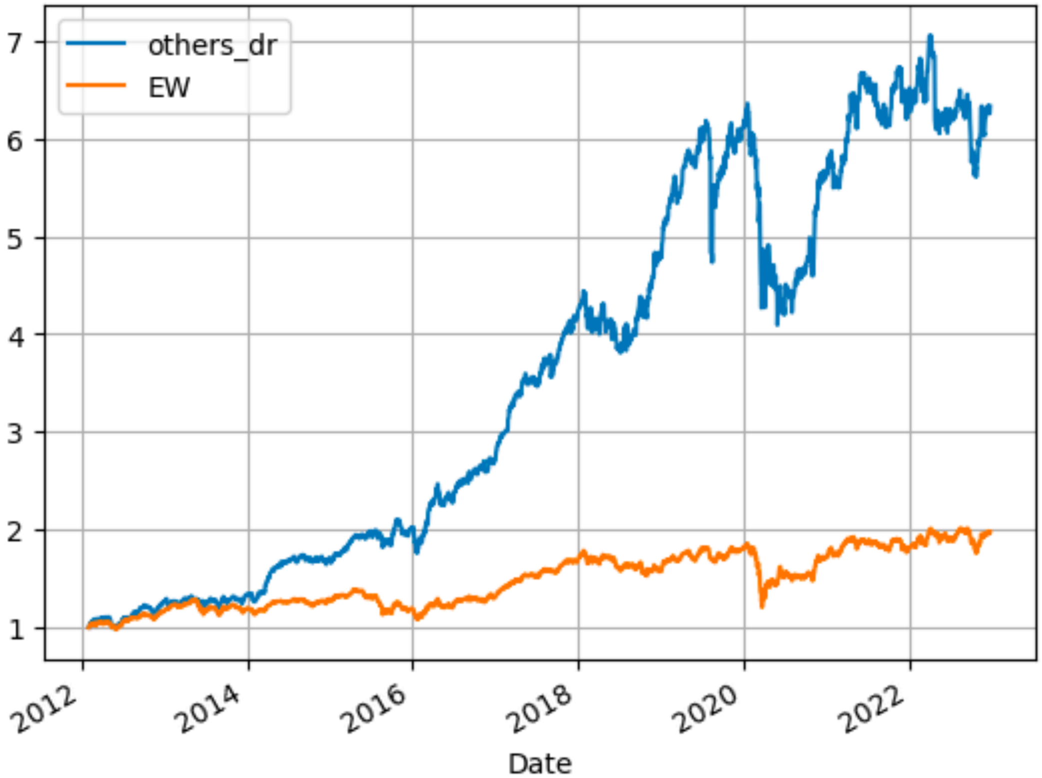


Figure 6.21 Cumulative Return, DR, signal weighting

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
others_dr	1.035	18.52	35.59

Table 6.11 Performance metrics of DR, signal weighting

6.3.6 Section Summary

In this section, we weight the stocks based on their signals cross-sectionally, using reciprocal method and exponential method. However, the performance of the signal weighting strategy seems to be worse than the simple ranking strategy according to the testing results. This is because the weight of the stocks now is not as extreme as before. The merit of doing this is to decrease the turnover of the strategy and therefore save transaction cost.

6.4 Sharpe Ratio Optimization

Sharpe Ratio is almost the most important metric to measure how successful a trading strategy is. It is also named as the “risk adjusted return”, which shows its ability to reflect two most important things that investors focus on - risk and return. Hence, in order to construct a more precise strategy, one reasonable way is to try to maximize the Sharpe ratio of our strategy.

The Sharpe ratio itself is actually quite easy to compute. We just need two numbers - the average and the standard deviation of the returns over a period. However, when we want to optimize the Sharpe ratio of our strategy, we actually want to maximize our expected Sharpe ratio - which is computed by expected return and expected volatility.

Where

The expected return will be estimated by our signals. For every trading day, we will estimate the beta of our signals based on a look back period. The covariance matrix will also be calculated based on historical data. We use Scipy in python to calculate the position that will maximize the expected Sharpe ratio. The position every day is calculated iteratively.

6.4.1 Volatility_kcp

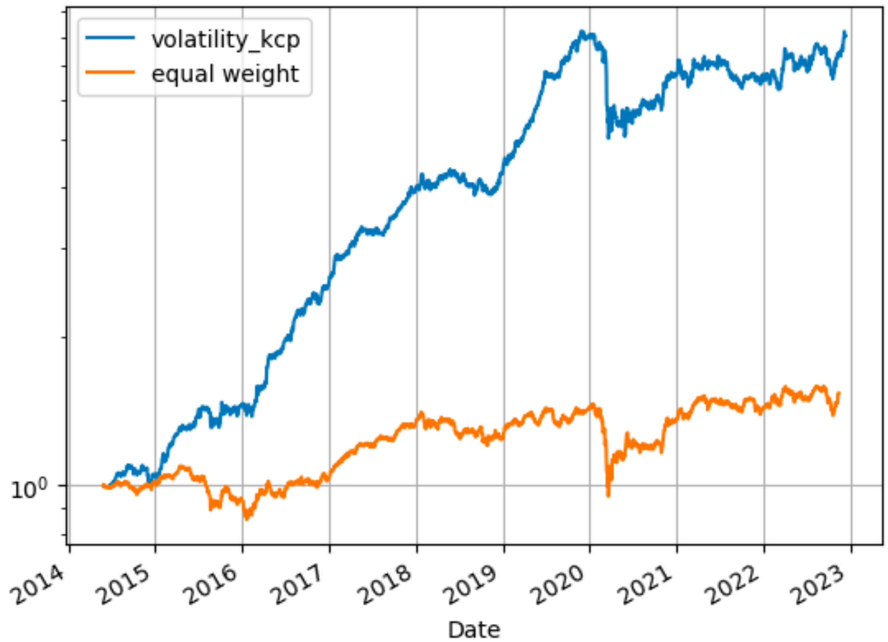


Figure 6.22 Cumulative Return, KCP, SR optimization

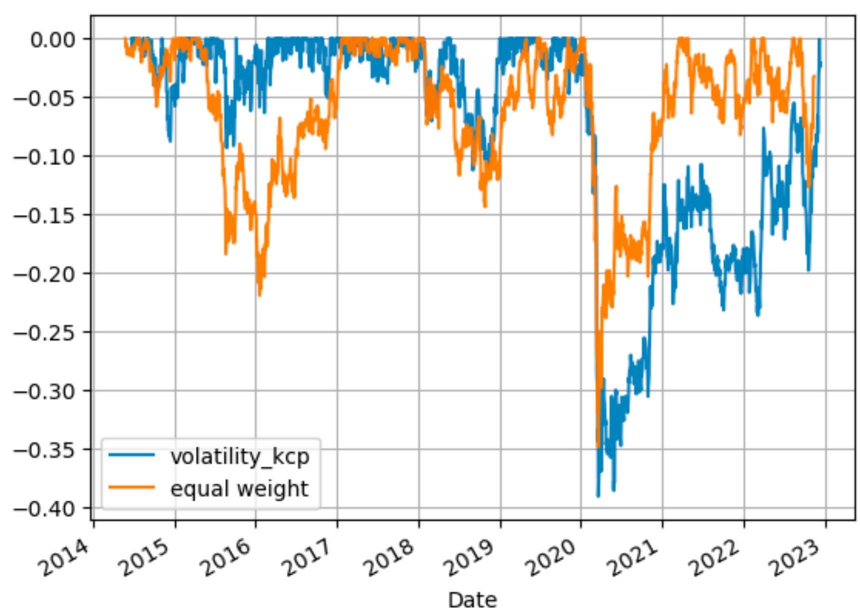


Figure 6.23 Max Drawdown, KCP, SR optimization

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
volatility_kcp	1.438	27.99	39.42

Table 6.12 Performance metrics of KCP, SR optimization

6.4.2 Momentum_wr

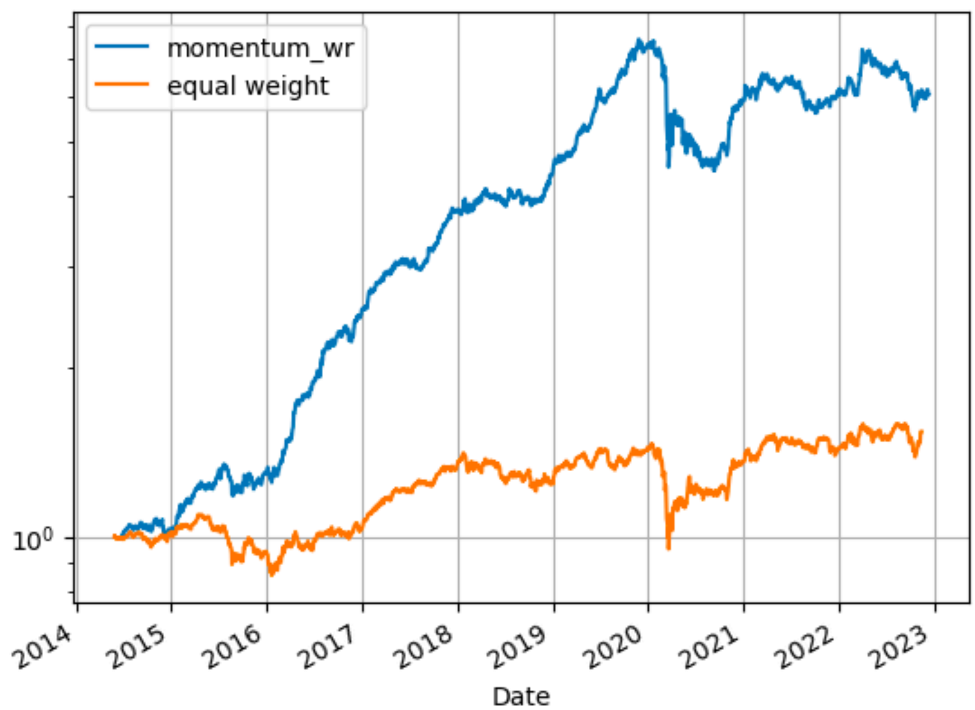


Figure 6.24 Cumulative Return, WR, SR optimization

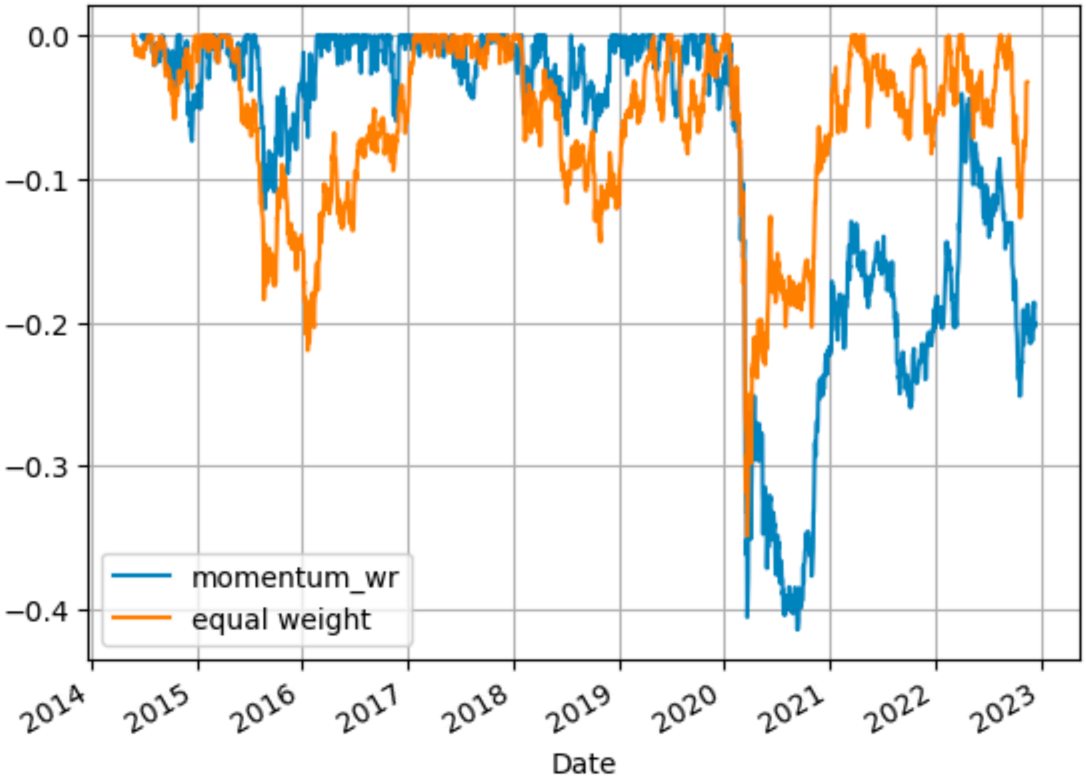


Figure 6.25 Max Drawdown, WR, SR optimization

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
momentum_wr	1.345	23.78	41.42

Table 6.13 Performance metrics of WR, SR optimization

6.4.3 Volume_vpt

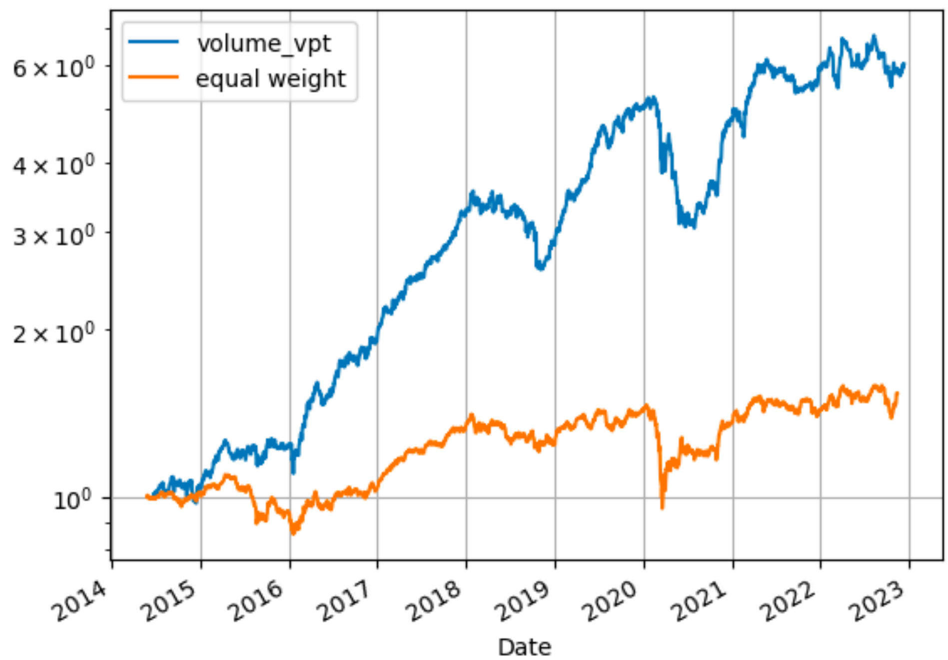


Figure 6.25 Cumulative Return, VPT, SR optimization

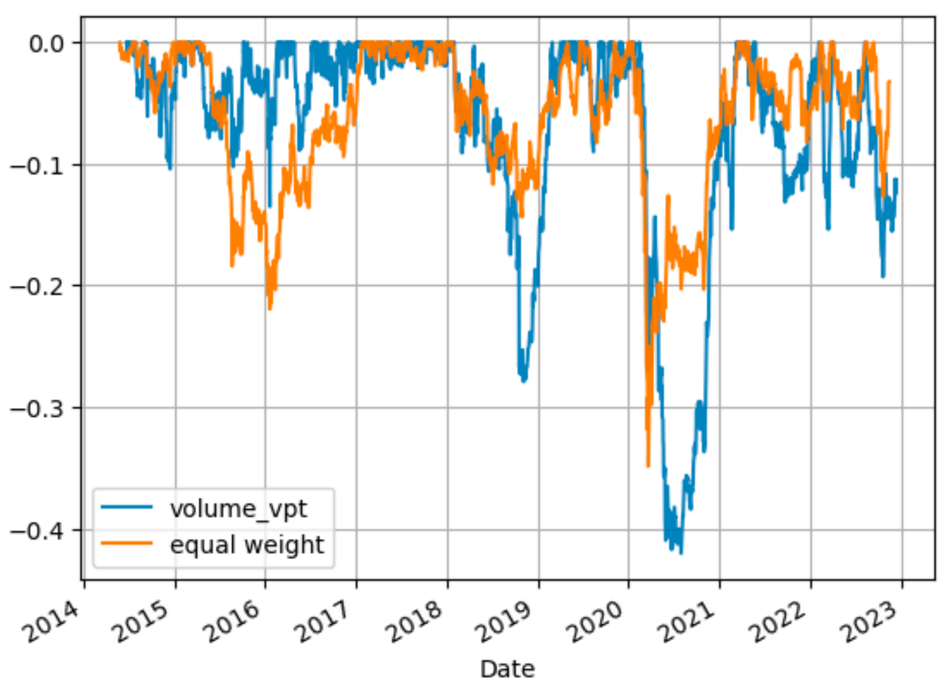


Figure 6.26 Max Drawdown, VPT, SR optimization

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
volume_vpt	1.123	23.66	42.04

Table 6.13 Performance metrics of VPT, SR optimization

6.4.4 Trend_psar

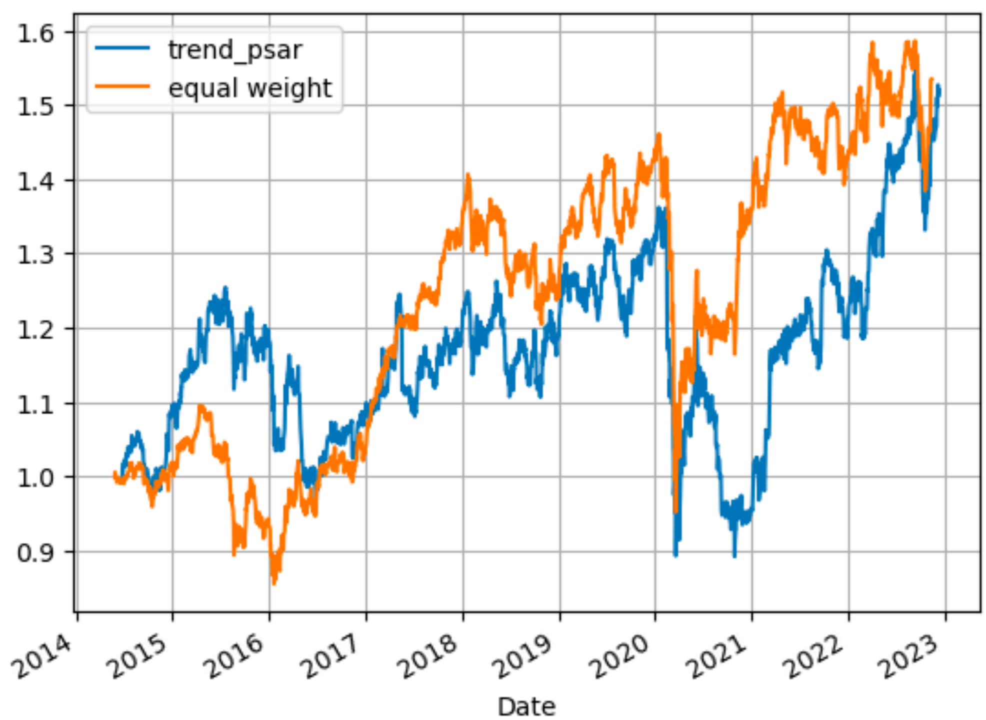


Figure 6.27 Cumulative Return, PSAR, SR optimization

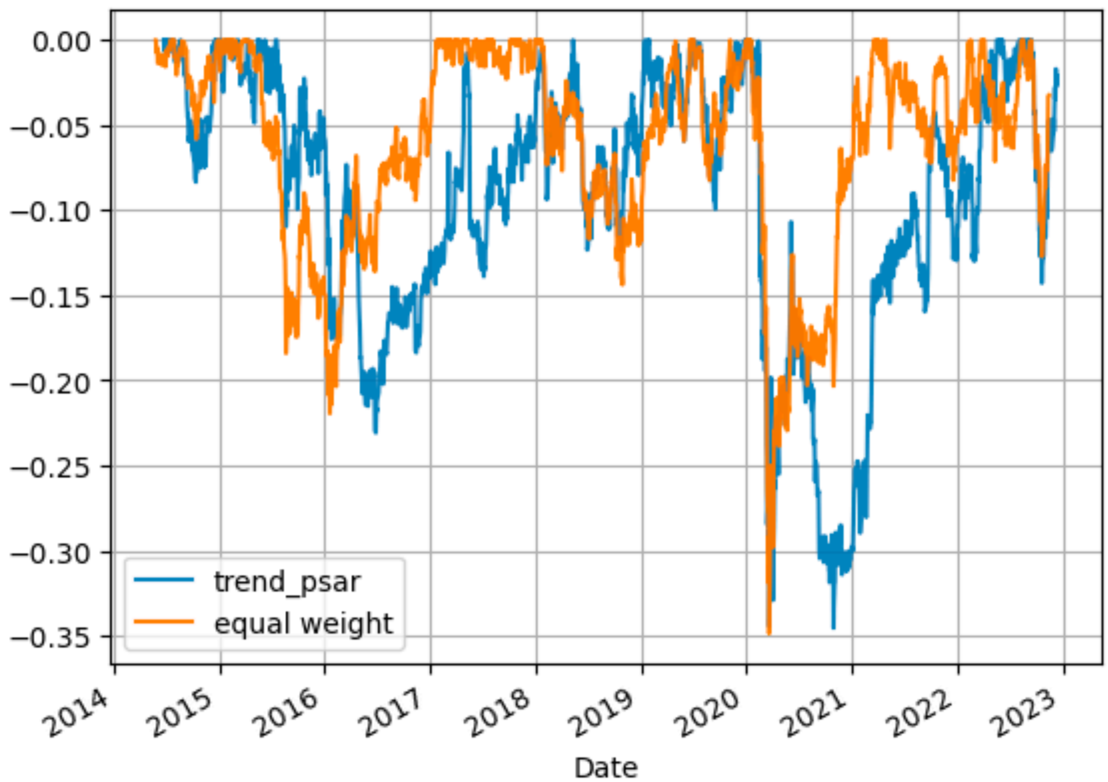


Figure 6.26 Max Drawdown, PSAR, SR optimization

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
trend_psar	0.373	5.08	34.55

Table 6.14 Performance metrics of PSAR, SR optimization

6.4.5 others_dr

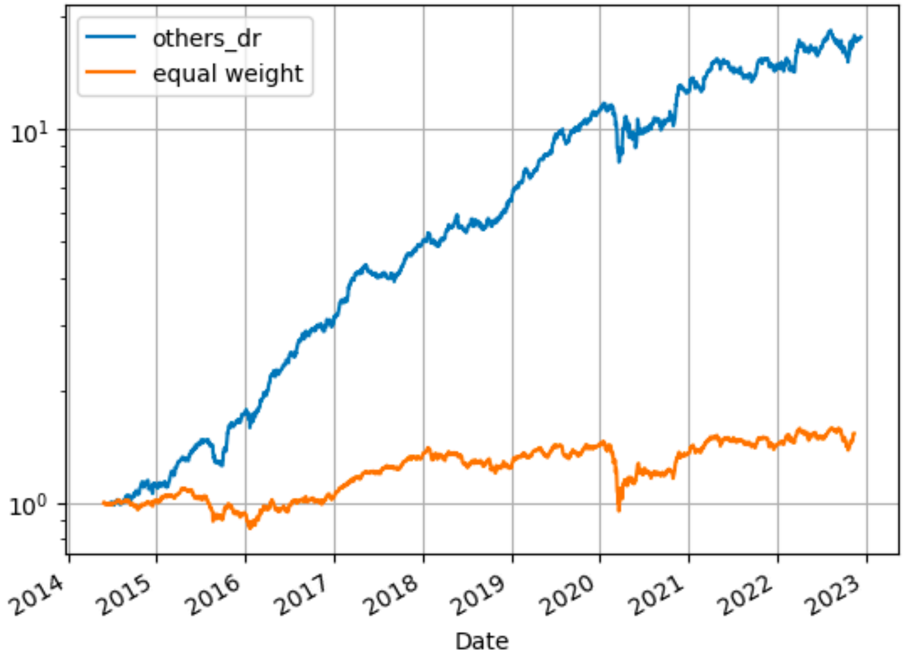


Figure 6.27 Cumulative Return, DR, SR optimization

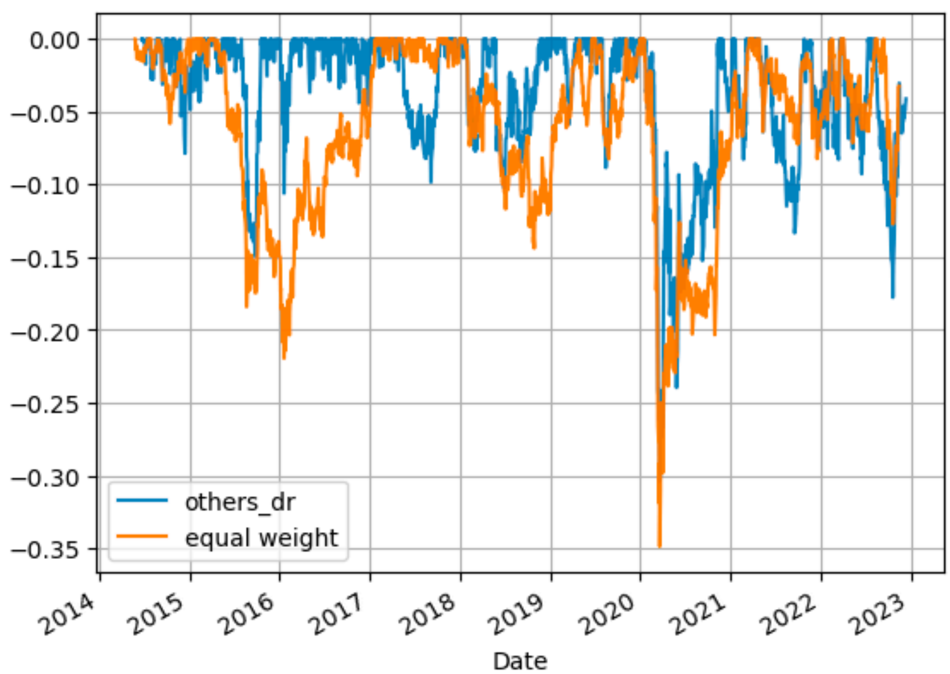


Figure 6.28 Max Drawdown, DR, SR optimization

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
others_dr	1.899	40.45	30.42

Table 6.15 Performance metrics of DR, SR optimization

6.4.6 Section Summary

In this section we use an iterative method to optimize the expected Sharpe ratio. The effect of this strategy is not consistent on each signal. For signal others_dr, this method increases its performance on every metric, while for trend_psar, its performance becomes unreasonably bad. There are some potential reasons behind this. The first possible reason is that, in this section, since we do not allow short trading, the weight of a stock can never be smaller than zero. However, the optimal position could require a negative weight on certain stock. Secondly, trend_psar is the combination of two similar but separate signals. In order to preserve its predictability, we might need to construct two different strategies.

6.5 Market Beta Hedging

In previous sections, we have successfully constructed strategies which have better Sharpe ratio and CAGR than our benchmark portfolio. However, almost all the strategies still have the same or even higher volatility and max draw down. The main reason is that we did not hedge market risk on these strategies. Hence, if we want to control the risk of our portfolio, the first thing we can do is to hedge the market beta on the assets.

Previously, we have set the boundary of weight for each stock to be 0 to 1, which means that we will never short any stocks. In reality, short selling is a very common action, especially in trend following strategies. In our case, we will loosen the boundary of weight to -1 to 1.

6.5.1 Volatility_kcp

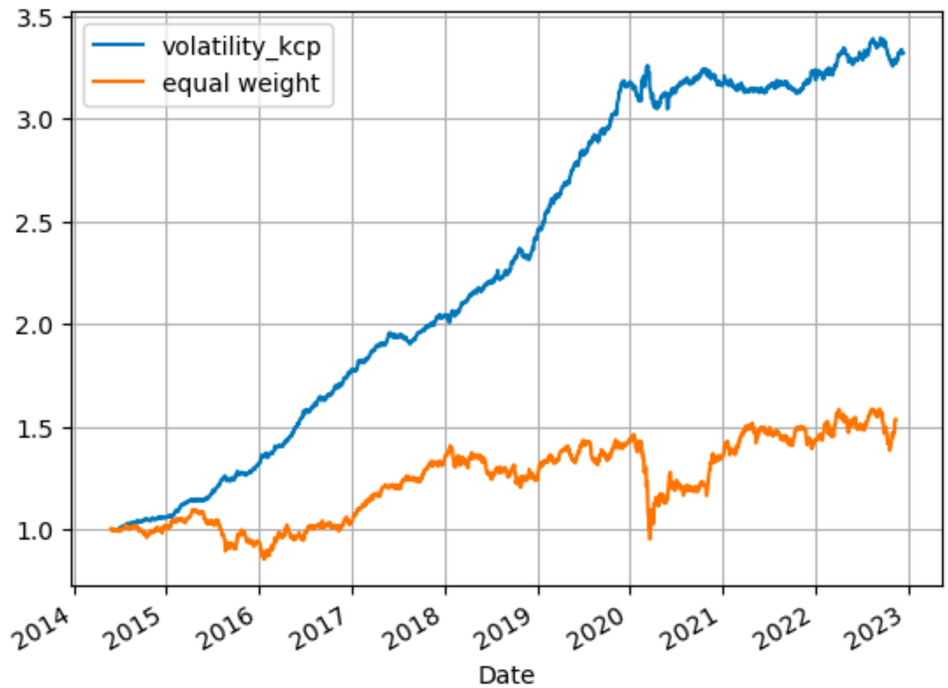


Figure 6.29 Cumulative Return, KCP, SR optimization with hedge

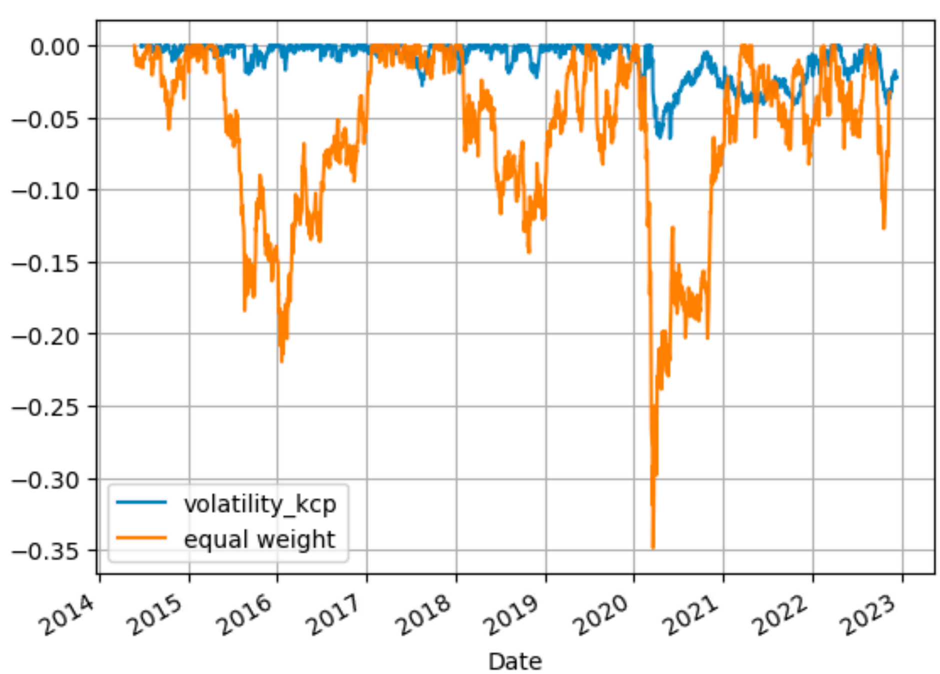


Figure 6.30 Max Drawdown, KCP, SR optimization with hedge

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
volatility_kcp	3.084	15.26	6.45

Table 6.16 Performance metrics of KCP, SR optimization with hedge

6.5.2 Momentum_wr

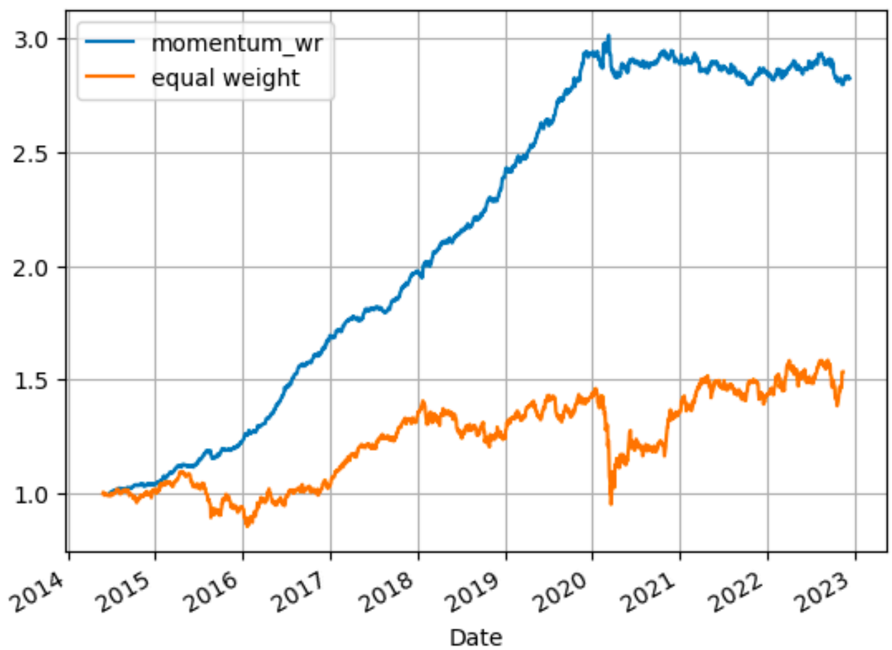


Figure 6.31 Cumulative Return, WR, SR optimization with hedge

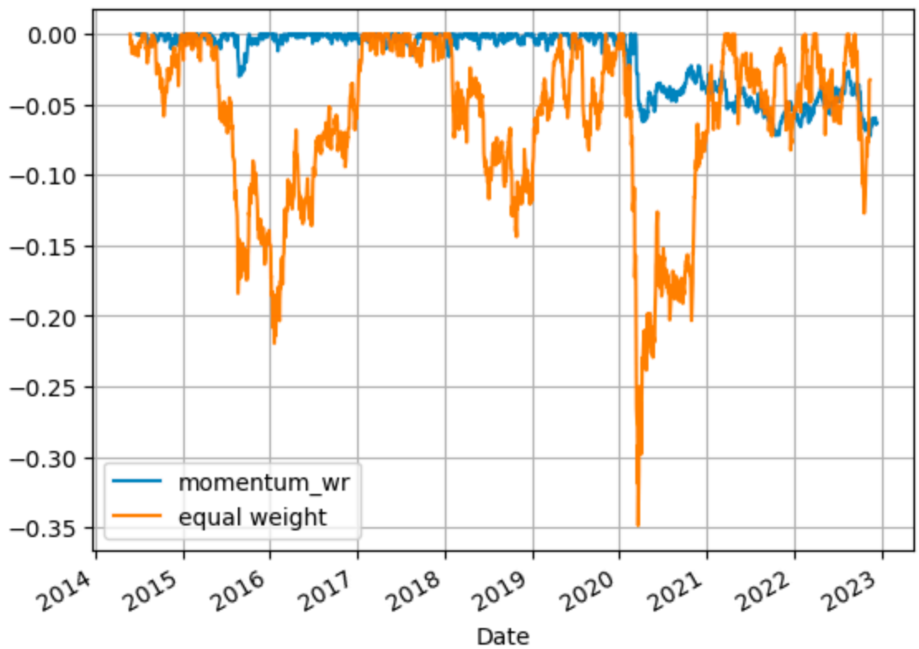


Figure 6.32 Max Drawdown, WR, SR optimization with hedge

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
volatility_kcp	2.859	13.09	7.24

Table 6.17 Performance metrics of WR, SR optimization with hedge

6.5.3 volume_vpt

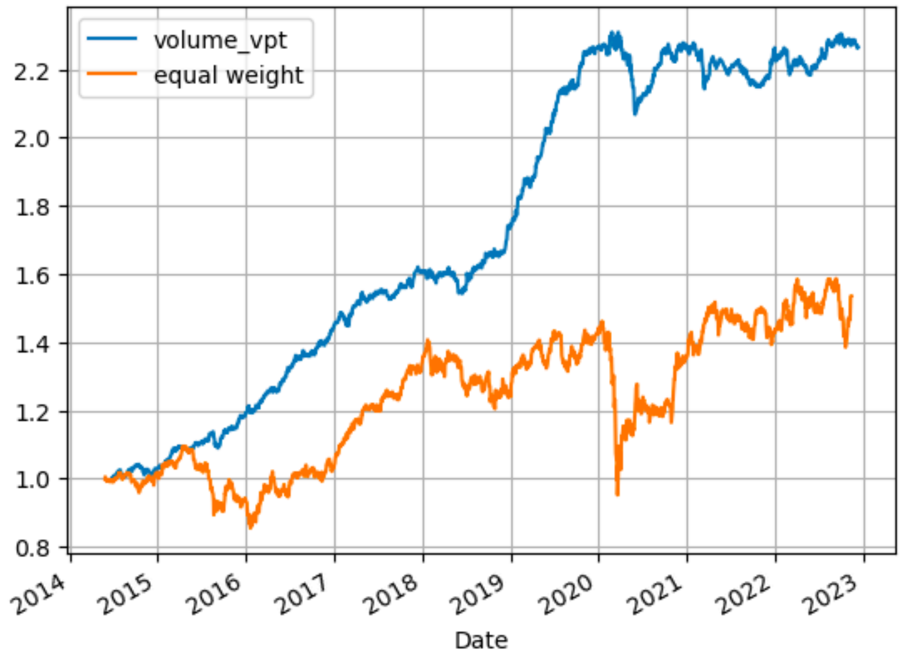


Figure 6.29 Cumulative Return, VPT, SR optimization with hedge

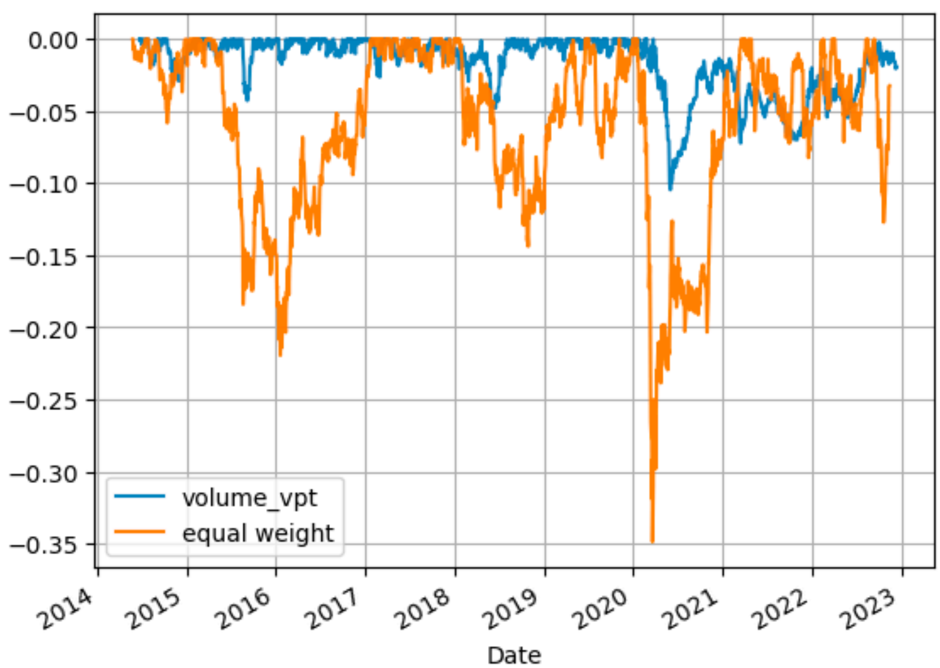


Figure 6.32 Max Drawdown, VPT, SR optimization with hedge

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
volume_vpt	1.88	10.15	10.45

Table 6.18 Performance metrics of VPT, SR optimization with hedge

6.5.4 Others_dr

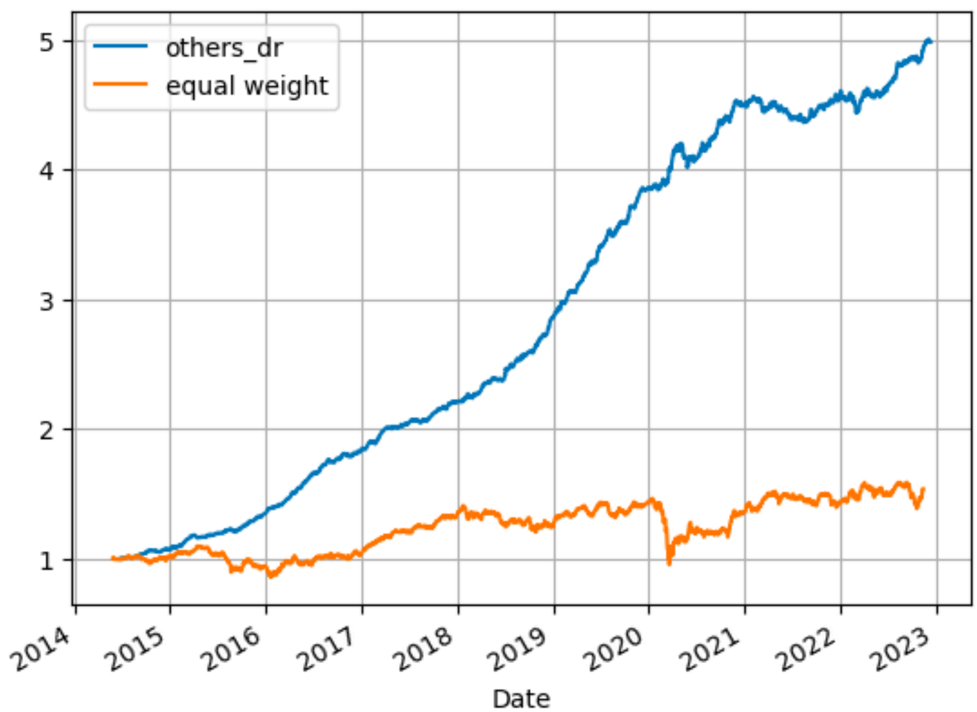


Figure 6.33 Cumulative Return, DR, SR optimization with hedge

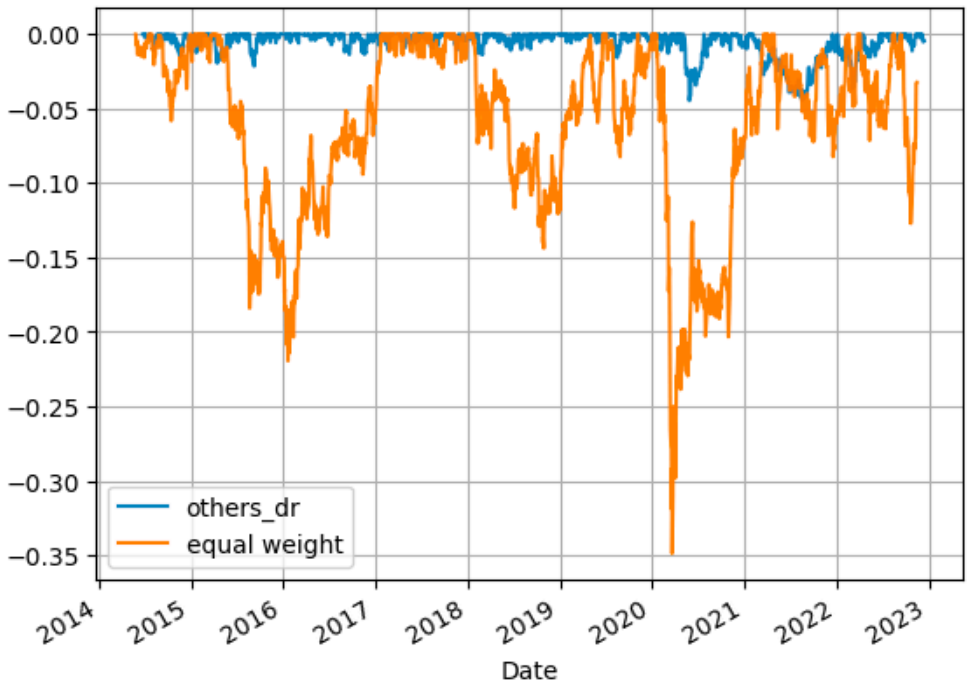


Figure 6.34 Max Drawdown, DR, SR optimization with hedge

Signal	Sharpe Ratio	CAGR(%)	Max Drawdown(%)
others_dr	4.231	20.93	4.46

Table 6.18 Performance metrics of DR, SR optimization with hedge

6.5.5 Section Summary

In this section, we have successfully hedged the market risk in our portfolio by ensuring that the sum of the market beta is zero. As a result, we can observe a smoother cumulative return curve, indicating that our strategies are less volatile than long-only strategies. Moreover, our Sharpe ratio has significantly increased.

However, it is important to note that most of the signals do not perform well in the years following 2020. This could be due to various reasons. Firstly, the Covid-19 outbreak in 2020 had a severe impact on the global economy and financial markets worldwide, which could have altered the SGX's environment significantly. Secondly, the increasing number of investors and funds using these signals might have caused them to lose their effectiveness gradually.

Despite this, the signal "others_dr" remains effective even after 2020, suggesting that the SGX is more of a mean-reverting market than a trendy one. This is because "others_dr" favors stocks that have a recent daily return below the average.

6.6 Volatility Targeting and Risk Management

To enhance the Sharpe ratio of our strategies, it is crucial to manage our portfolio's risk meticulously. Typically, an asset's risk in finance is determined by its return volatility. When an asset exhibits high volatility, it is advisable to reduce our position on the asset to mitigate the overall portfolio risk. This practice is commonly known as volatility targeting.

6.6.1 Asset level volatility targeting

In practice, we begin by establishing a target volatility level for our portfolio. Next, we calculate the ratio of this target volatility to the recent volatility of each stock. This ratio yields a leverage ratio for each stock. If the recent volatility of a stock is high, we will reduce its position, while increasing it for stocks with lower recent volatility. To prevent overexposure to any one stock, we cap the leverage ratio.

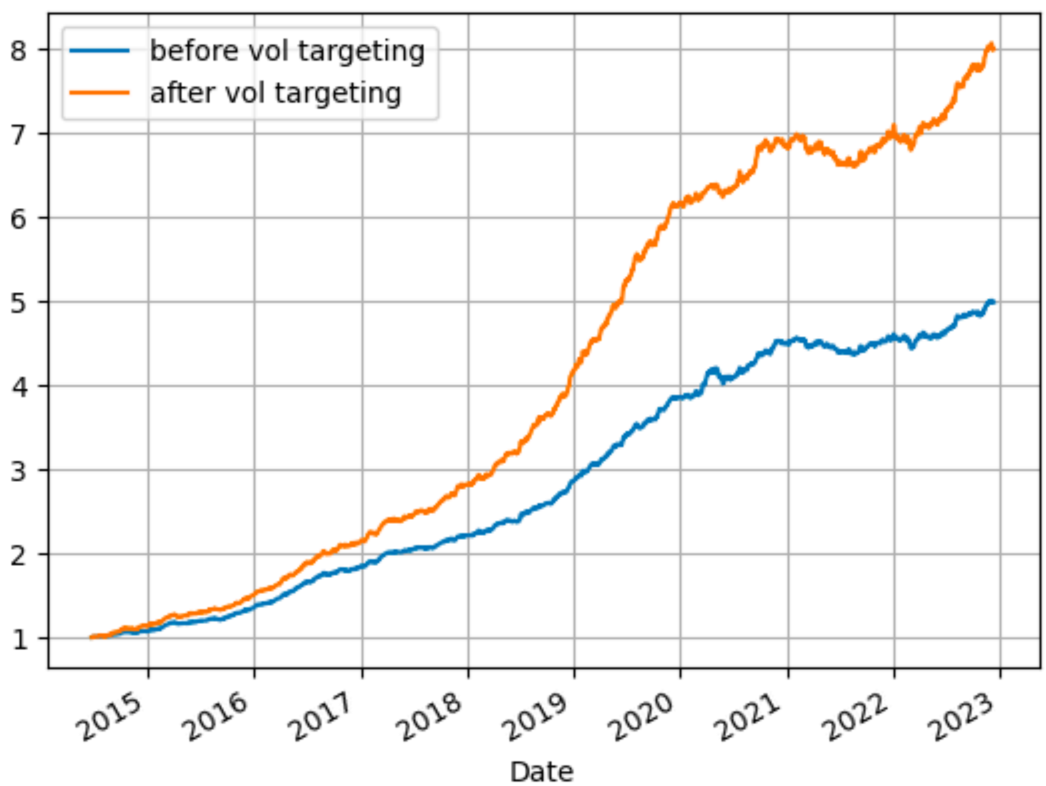


Figure 6.33 Comparison of Cumulative Return, DR

	Before Volatility Targeting	After Volatility Targeting
Sharpe Ratio	4.231	4.617

Table 6.19 Sharpe Ratio Comparison

Take the signal others_dr as an example, when we set the volatility target as 1.5% and cap the leverage ratio to 2, the sharpe ratio increases significantly from 4.231 to 4.617.

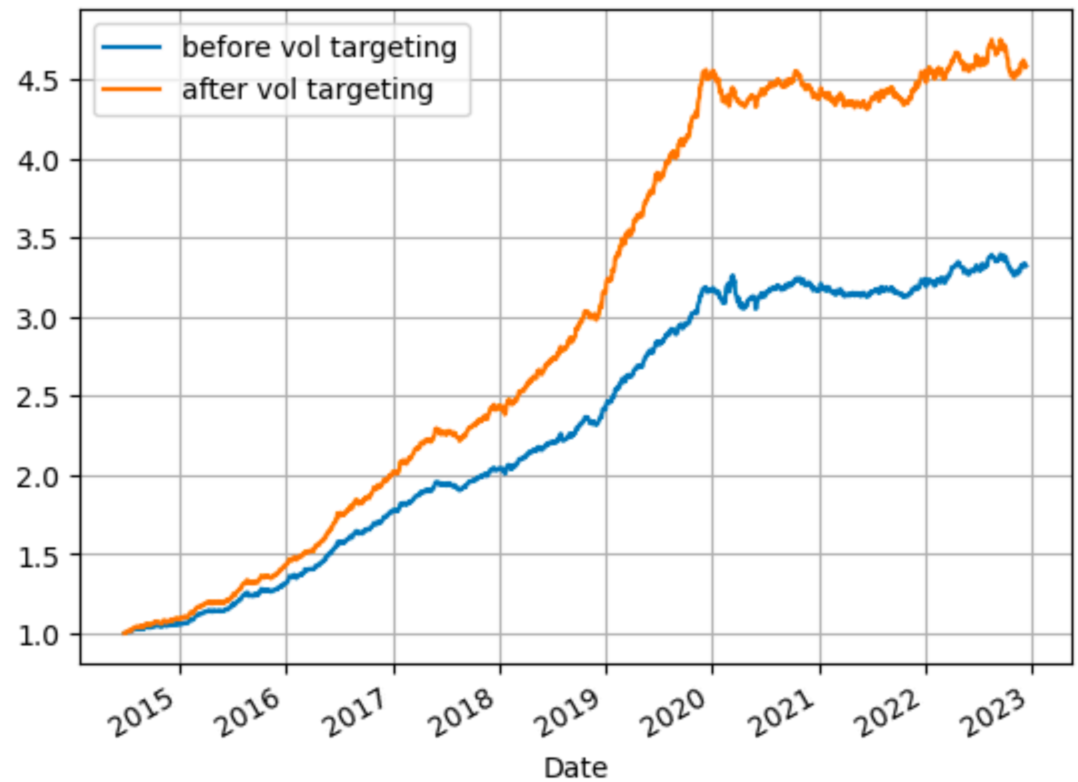


Figure 6.34 Comparison of Cumulative Return, KCP

	Before Volatility Targeting	After Volatility Targeting
Sharpe Ratio	3.084	3.520

Table 6.20 Sharpe Ratio Comparison

We found that for volatility_kcp, the Sharpe ratio increases from 3.084 to 3.52. Similar effects can be observed on other signals as well.

6.6.2 Strategy level volatility targeting

In the last section, we adjust our position based on the volatility of each asset. Beyond that, a further level of volatility targeting on the strategy level is also possible. We still set a target volatility for our strategy, then calculate a leverage ratio based on the recent volatility of our strategy.

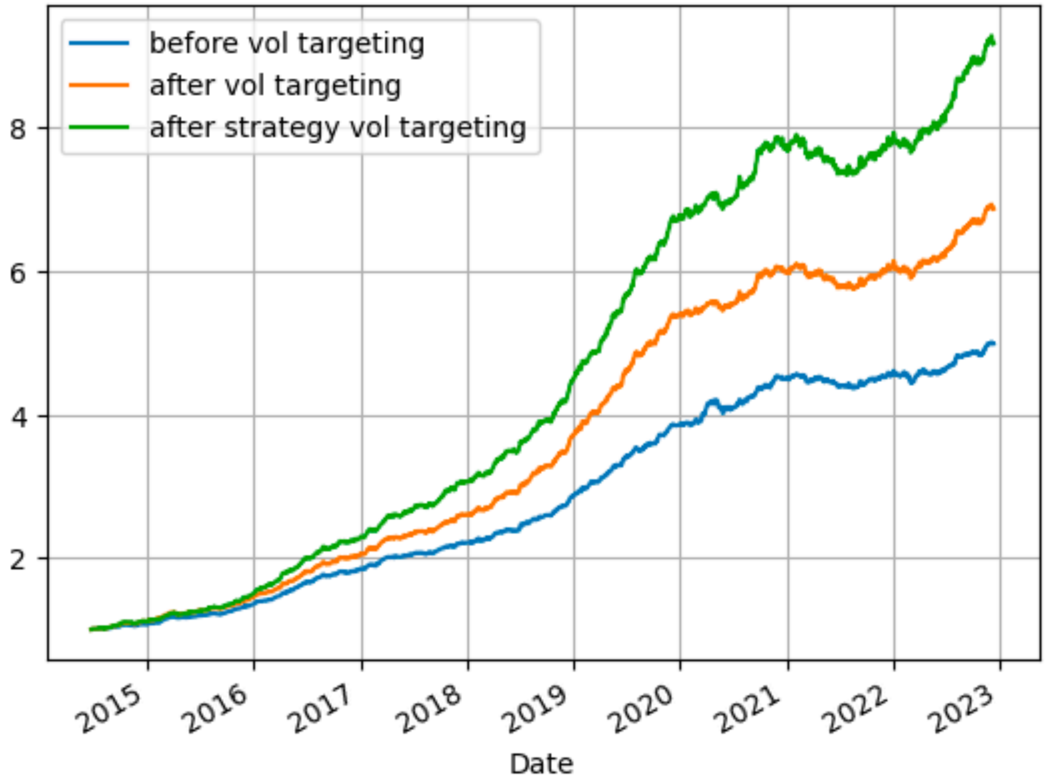


Figure 6.35 Comparison of Cumulative Return, DR, with hedge

	Asset level only	Strategy level
Sharpe Ratio	4.617	4.675

Table 6.21 Sharpe Ratio Comparison

For signal others_dr, since the strategy level volatility is already very low, we set the volatility target to 0.4%. Although the sharpe ratio does not improve a lot, the strategy level volatility targeting does help to increase sharpe ratio as well as CAGR.

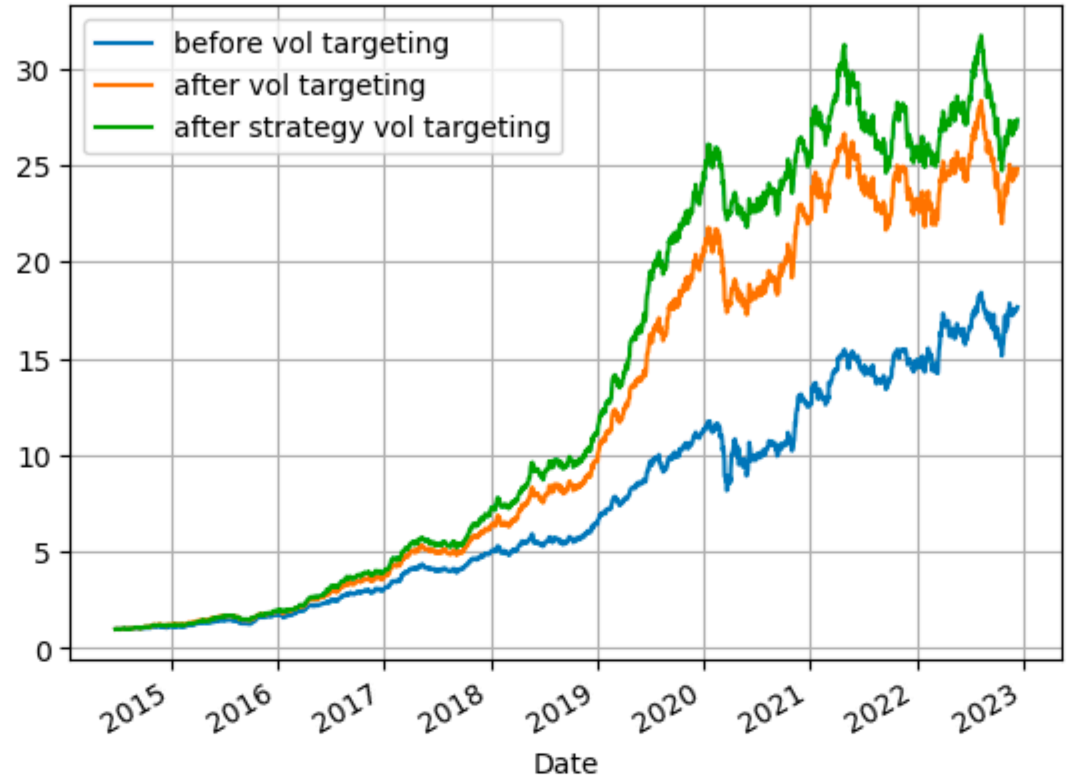


Figure 6.35 Comparison of Cumulative Return, DR, without hedge

	Asset level only	Strategy level
Sharpe Ratio	2.324	2.503

Table 6.22 Sharpe Ratio Comparison

However, for other strategies which have higher strategy level volatility, the effect of strategy level volatility targeting is more significant. For example, for the same signal others_dr, if we do not hedge market risk, the volatility will be much higher. In this case, the sharpe ratio increases from 2.3 to 2.5, which is a larger progress compared to the previous case.

6.6.3 Signal Combination

In addition to volatility targeting, another prevalent risk management technique is to combine multiple strategies. This approach is based on the premise that different signals capture diverse profitable trading opportunities. Consequently, a strategy that integrates various signals is likely to have lower risk since the risk of each signal can be partially hedged by the others.

	DR	KCP	WR	VPT
DR	1.000000	0.347450	0.318401	0.427581
KCP	0.347450	1.000000	0.810738	0.448537
WR	0.318401	0.810738	1.000000	0.362171
VPT	0.427581	0.448537	0.362171	1.000000

Table 6.23 Correlation between strategies

From the correlation chart, we can find that the return of volatility_kcp and momentum_wr has a correlation of 0.81, indicating that these two signals are likely to be similar signals, while others_dr and volume_vpt are relatively uncorrelated to each other. Hence, we can combine these signals based on their individual performance and correlation with each other.

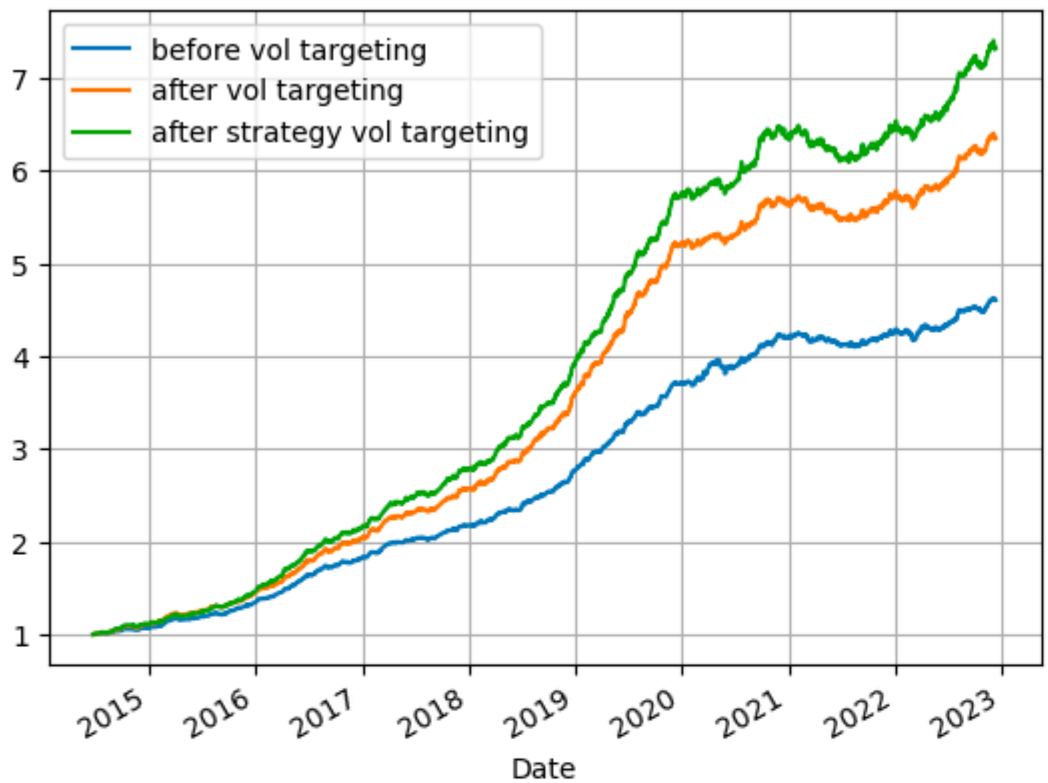


Figure 6.36 Comparison of Cumulative Return, DR, without hedge

	No Vol Targeting	Asset level only	Strategy level
Combined	4.503	4.883	4.915
Others_dr	4.231	4.617	4.675
Vol_kcp	3.084	3.520	3.552
Momentum_wr	2.859	3.164	3.209
Volume_vpt	1.880	2.199	2.162

Table 6.24 Performance Comparison between Strategies

It is clear that, although others_dr has better Sharpe ratio than any other signal, the Sharpe ratio of the combined strategy is higher than the Sharpe ratio of every individual strategy. This shows that the orthogonality of signals does offer another way for us to hedge intrinsic risk of our strategies.

Chapter 7

Conclusion

This project has successfully constructed several multi-factor trading strategies using classical financial theories and statistical models. The testing results demonstrate the feasibility of using both fundamental and technical signals to generate profitable trades in the SGX market.

Initially, we selected our signal universe and investment universe based on the efficient market hypothesis. The signals we selected includes fundamental factors like P/E ratio, P/B ratio and market capitalization, which measures the profitability, market evaluation of the stocks, as well as technical factors like daily return, volatility_kcp, momentum_wr, etc., which measures the short-term price and volume characteristic of the stocks. We then evaluated various classic models and statistical methods, including Fama-French Three Factor model, Fama-Macbeth Regression and Lasso Regression, testing our signals on both cross-sectional and time-series basis. Subsequently, we construct strategies trading the significant signals and adjust them in diverse ways to improve our strategies' performance, such as optimizing the Sharpe ratio, hedging market risk, targeting volatility, combining signals, etc. Ultimately, the results indicate that our strategies significantly outperform the market. Among all the signals, "others_dr" (daily return) has the best performance.

Several key takeaways emerged from our analysis. Firstly, the signal "others_dr" performed better than both "volatility_kcp" and "momentum_wr," indicating that SGX is more of a mean-reverting market than a trendy one. Secondly, all of the signals we utilized experienced reduced effectiveness following the COVID-19 outbreak in 2020, potentially indicating a fundamental shift in market conditions. Lastly, technical signals tend to exhibit greater potential for generating excess returns than fundamental signals.

However, implementing our strategy in the real market poses certain challenges. One such challenge is transaction costs, which may include fees, commissions, bid-ask spreads, taxes, and other expenses incurred during the transaction process. Investors and hedge funds may encounter different costs, which could impact their preference for low versus high turnover strategies. Additionally, SGX is not a very liquid market, so liquidity problems could arise if the portfolio size is too large, and orders may become difficult to fill. Therefore, we may need to refine our trading system to overcome these challenges when executing the strategy.

Chapter 8

Reflection on Learning Outcome Attainment

Based on the successful completion of this project, I have gained a wealth of engineering knowledge that has helped me develop a deeper understanding of trading strategies and financial theories. Through the process of selecting signal and investment universes based on the efficient market hypothesis, I learned about market efficiency and the various factors that can influence stock prices.

Moreover, by evaluating classic models and statistical methods such as Fama-French Three Factor model, Fama-Macbeth Regression, and Lasso Regression, I gained a strong understanding of how statistical models can be used to generate profitable trades in financial markets.

In addition to this, I developed robust problem analysis skills through the construction of trading strategies and the adjustment of those strategies to improve their performance. I learned how to use techniques such as optimization of Sharpe ratio, hedging market risk, targeting volatility, and combining signals to improve the effectiveness of the strategies.

Finally, the ability to develop effective solutions that can be used in real-world trading environments has been a significant outcome of this project. I have gained valuable insights into the challenges faced by investors and hedge funds when implementing trading strategies in the market, such as transaction costs and liquidity issues. As a result, I have learned the importance of refining trading systems to overcome these challenges and make the strategies more effective.

In conclusion, this project has provided me with a valuable learning experience, helping me to develop a wide range of engineering knowledge, problem analysis skills, investigation techniques, and the ability to develop effective solutions that can be applied in real-world trading environments. These skills will be invaluable in my future endeavors in the financial industry.

References

[1] Sharpe, W. F. (1964, Sept.). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. The Journal of Finance, 19, 425-442.

[2] Fama, E. F., & French, K. R. (1996). Multifactor Explanations of Asset Pricing Anomalies. The Journal of Finance, 51, 55-84.

[3] Feng, G., Giglio, S., & Xiu, D. (2020, June). Taming the Factor Zoo: A Test of New Factors. The Journal of Finance, 75, 1327-1370.

[4] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383-417.

[5] Markowitz, H.M. (March 1952). "Portfolio Selection". The Journal of Finance. 7 (1): 77–91.

[6] French, Craig W. (2003). "The Treynor Capital Asset Pricing Model". Journal of Investment Management. 1 (2): 60–72. SSRN 447580.

[7] Fama, E. F.; French, K. R. (1993). "Common risk factors in the returns on stocks and bonds". Journal of Financial Economics. 33: 3–56.

[8] Fama, E. F., & French, K. R. (2015, April). A five-factor asset pricing model. Journal of Financial Economics, 116, 1-22.

[9] Fama, E. F., & Macbeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. The Journal of political economy, 81, 607-636.

Appendix

Fama-Macbeth Regression Result

Signal	Mean	Std. Error	t-stats
volume_adi	2.80E-12	1.07E-12	2.608912165
volume_obv	-1.38E-12	6.31E-13	-2.183668602
volume_cmf	-6.41E-07	4.14E-06	-0.154657054
volume_fi	6.95E-10	3.16E-09	0.220112723
volume_em	2.15E-06	1.47E-05	0.146227968
volume_sma_em	-4.84E-06	2.32E-05	-0.208892177
volume_vpt	-2.49E-09	2.31E-09	-1.07806811
volume_vwap	1.15E-06	1.89E-06	0.607969622
volume_mfi	1.26E-05	1.43E-05	0.87718731
volume_nvi	8.12E-07	8.22E-07	0.987466451
volatility_bbm	1.27E-06	1.70E-06	0.751466833
volatility_bbh	2.51E-06	2.73E-06	0.919591083
volatility_bbl	4.13E-08	2.95E-06	0.013982052
volatility_bbw	1.24E-07	2.27E-05	0.005472049

volatility_bbp	-5.45E-06	1.29E-06	-4.227194079
volatility_bbhi	-5.99E-06	6.10E-06	-0.982767405
volatility_bbli	2.62E-06	4.37E-06	0.598948895
volatility_kcc	1.71E-06	2.00E-06	0.855791644
volatility_kch	2.00E-06	2.06E-06	0.973988171
volatility_kcl	1.42E-06	2.20E-06	0.646272719
volatility_kcw	-1.01E-05	1.55E-05	-0.650334249
volatility_kcp	-1.59E-05	3.89E-06	-4.087320757
volatility_kchi	-9.37E-06	8.49E-06	-1.103140104
volatility_kcli	-7.51E-07	7.43E-06	-0.101070028
volatility_dcl	-2.36E-06	2.79E-06	-0.84693675
volatility_dch	1.30E-06	2.56E-06	0.508650212
volatility_dcm	-5.31E-07	1.78E-06	-0.298793012
volatility_dcw	-1.80E-05	2.45E-05	-0.73645445
volatility_dcp	-4.11E-06	2.02E-06	-2.038293901
volatility_atr	4.22E-07	7.17E-07	0.587654829
volatility_ui	5.79E-06	1.52E-05	0.379991325

trend_macd	1.22E-06	1.22E-06	1.002650651
trend_macd_signal	1.32E-06	1.23E-06	1.076146443
trend_macd_diff	-9.74E-08	4.00E-07	-0.243597822
trend_sma_fast	1.12E-06	1.93E-06	0.581830967
trend_sma_slow	3.61E-07	1.66E-06	0.218350616
trend_ema_fast	1.45E-06	1.87E-06	0.777598357
trend_ema_slow	2.30E-07	1.40E-06	0.163861752
trend_vortex_ind_pos	1.51E-06	1.78E-06	0.847078093
trend_vortex_ind_neg	2.94E-06	1.86E-06	1.577862508
trend_vortex_ind_diff	-1.43E-06	2.71E-06	-0.528773046
trend_trix	9.24E-09	3.76E-07	0.024581334
trend_mass_index	1.29E-05	1.95E-05	0.663683285
trend_dpo	-4.30E-06	2.26E-06	-1.905887352
trend_kst	-5.36E-06	1.04E-05	-0.513061612
trend_kst_sig	-4.29E-06	8.25E-06	-0.519854702
trend_kst_diff	-1.07E-06	1.50E-05	-0.071642681
trend_ichimoku_conv	1.37E-06	2.10E-06	0.653591074

trend_ichimoku_base	-1.34E-06	2.01E-06	-0.669081236
trend_ichimoku_a	1.48E-08	1.60E-06	0.00928513
trend_ichimoku_b	-3.60E-06	3.61E-06	-0.998572273
trend_stc	1.18E-05	7.74E-06	1.526768818
trend_adx	-3.10E-06	1.91E-05	-0.162592956
trend_adx_pos	-1.92E-08	2.70E-05	-0.000711514
trend_adx_neg	-6.34E-05	2.72E-05	-2.327881593
trend_cci	9.38E-06	4.75E-06	1.972961755
trend_visual_ichimoku_a	-7.10E-06	4.84E-06	-1.465994694
trend_visual_ichimoku_b	-1.17E-05	8.13E-06	-1.441916471
trend_aroon_up	-2.06E-07	5.95E-06	-0.034595274
trend_aroon_down	2.48E-06	5.95E-06	0.41622781
trend_aroon_ind	-2.68E-06	4.02E-06	-0.668271526
trend_psar_up	6.44E-07	4.13E-06	0.155939695
trend_psar_down	1.07E-07	3.98E-06	0.026797517
trend_psar_up_indicator	1.27E-05	5.72E-06	2.210876943
trend_psar_down_indicat	-4.25E-06	6.21E-06	-0.683915968

or			
momentum_rsi	-0.000112427	2.43E-05	-4.624821678
momentum_stoch_rsi	-6.65E-06	3.88E-06	-1.712137459
momentum_stoch_rsi_k	-2.76E-06	3.29E-06	-0.839757242
momentum_stoch_rsi_d	-1.14E-06	3.37E-06	-0.339156786
momentum_tsi	2.91E-05	2.54E-05	1.147800795
momentum_uo	-4.21E-05	1.88E-05	-2.235484555
momentum_stoch	-6.35E-06	1.83E-05	-0.346705502
momentum_stoch_signal	6.42E-05	1.83E-05	3.508501272
momentum_wr	-8.27E-05	1.42E-05	-5.839730017
momentum_ao	3.59E-06	3.07E-06	1.168311442
momentum_roc	-2.85E-05	2.51E-05	-1.131884872
momentum_ppo	-3.57E-06	3.76E-06	-0.94765286
momentum_ppo_signal	7.56E-08	3.34E-06	0.02263543
momentum_ppo_hist	-3.64E-06	2.26E-06	-1.608593674
momentum_pvo	2.35E-05	1.23E-05	1.913339107
momentum_pvo_signal	1.01E-05	1.62E-05	0.624232892

momentum_pvo_hist	1.34E-05	1.87E-05	0.715584482
momentum_kama	2.95E-06	2.18E-06	1.349302122
others_dr	-3.46E-05	1.51E-05	-2.295734943
others_dlr	-3.37E-05	1.50E-05	-2.246721777
others_cr	-3.48E-06	7.66E-06	-0.455085068