Proceedings of the 2nd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/92/20231225

Construction of a Scoring and Ranking System for Multi-Factor Quantitative Stock Selection Strategy with Daily Rebalancing: An Analysis of SSE A-shares Component Stocks

Yihan Ren[1,a], Yulei Lin[2,b,*], Jiatong Li[3,c], Linxuan Guo[4,d], Wangdi Chen[5,e]

[1]Yihan Ren, School of Mathematics and Statistics, Northeastern University at Qinhuangdao, Qinhuangdao,066004, China

[2]Yulei Lin, Beijing 21st Century International School, Beijing,100142, China

[3]Jinan University-University of Birmingham Joint Institute, Jinan University & University of Birmingham, Guangdong, 511400, China

[4]Beijing-Dublin International College, Beijing University of Technology & University College Dublin, Beijing,100124, China

[5]College of Liberal Arts & Sciences, University of Illinois Urbana-Champaign, Urbana, 61801, US

a. 13819575627@163.com, b. linyl15033769663@163.com, c. 1368396658@qq.com, d. 2212810284@qq.com, e. wangdic2@illinois.edu

Yihan Ren and Yulei Lin contributed equally to this work and should be considered as co-first author.

Jiatong Li, Linxuan Guo and Wangdi Chen contributed equally to this work and should be considered as co-second author.

**corresponding author*

Abstract: In this study, we introduce a four factor equal-weight scoring model, utilizing SSE A-shares from 2019 to 2023, aimed at identifying stocks of high investment potential. We innovatively incorporate the E/I factor to assess the intrinsic investment value of stocks, integrating it with PEG, RSI, and beta to ensure a comprehensive information capture. While assuming the feasibility of short-selling in the A-share market, our proposed strategy offers insights for future statistical arbitrage strategies in this market, underscoring the significance of risk hedging. Our preliminary analysis suggests that the influence of the four factors on a stock's intrinsic value is not strictly linear. Consequently, we employ a factor rating approach to enhance their explanatory power regarding a stock's intrinsic investment value. Initial results indicate that while our original strategy effectively manages risk, it compromises on returns. However, post-adjustment, the refined scoring model demonstrates robust profitability, yielding an annual return of 23.967078% and a Sharpe ratio of 1.146165. These findings validate the efficacy of our proposed strategies, offering traders a novel investment direction.

Keywords: Multi-factor scoring model, SSE A-shares, risk hedging, E/I factor

DOI: 10.54254/2754-1169/92/20231225

1. Introduction

1.1. Idea of Strategies

Our research team has developed a comprehensive scoring model, incorporating four key factors (PEG, RSI, beta, and E/I), to predict and rank the future return trends of stocks in SSE A-Share. This model is further enhanced by daily adjustments to our portfolio and a strategic weight allocation to individual stocks using the Total Risk Contribution (TRC) concept. The underpinnings of our approach can be traced back to Fama-French's three-factor model (1993) and Ray Dalio's Risk Parity Theory. However, our model extends and diverges in certain aspects, showcasing a novel approach in the multi-factor scoring landscape.

1.2. Highlight

1.2.1. Economic Intuition

Our strategy capitalizes on the intrinsic value of a company, its industry dynamics, and real-time stock performance. By integrating four key indicators - PEG, RSI, beta, and E/I - we tap into the company's fundamentals, market momentum, and systemic risk. The E/I, a ratio of external to internal trading volumes, serves as a barometer for market demand. A surge in external volume underscores heightened market demand, signaling potential trading value and arbitrage opportunities.

1.2.2. Signal Generation

The signal for our investment strategy is derived from the stock scores computed by our multi factor scoring model. We rank these scores in descending order. A higher score suggests a higher probability of the stock yielding higher future returns, indicating a strong upward trend, thereby prompting a buy decision. Conversely, a lower score signifies a likely downward trend in future returns, thus leading to a decision to short-sell the respective stock.

Besides, we employ a combination methodology that processes these raw scores into a time series of position vectors. This methodology takes into account the historical performance, volatility, and correlation of each stock, ensuring that our signals are not only based on current scores but also on the broader market context. By transforming our raw signals into position vectors, we can better gauge the potential movement of each stock in the context of our entire portfolio.

This enhanced scoring and combination system serves as a robust signal generator in our investment strategy, directing our buy and short-sell actions more effectively.

1.2.3. Portfolio Construction

Our investment strategy is anchored on a daily trading and rebalancing frequency, ensuring real-time responsiveness to market dynamics. We manage a dynamically adjusted portfolio comprising 200 stocks from the SSE A-shares. Utilizing our multi-factor scoring model, we rank all stocks at the end of each trading day based on the most recent stock price data.

For the long positions, we select the 100 highest-scoring stocks, each allocated an equal proportion of the portfolio. The purchase is executed at the opening of the market on the following day, adhering to the Total Risk Contribution (TRC) method for capital allocation. These stocks are held throughout the day and are entirely sold off at the opening of the market on the subsequent day.

Concurrently, under our assumption of permissible short selling in the A-share market, we choose the 100 lowest-scoring stocks for short positions, each also allocated an equal proportion of the portfolio. The short selling is conducted at the opening of the market on the following day, using an equal-weight method for capital allocation. These short positions are maintained throughout the day and are entirely bought back to close the positions at the beginning of the next day.

Our strategy is designed to achieve market neutrality through a balanced long/short hedge. This ensures that we not only harness the predictive advantages of the multi-factor model but also mitigate systematic market risk. Through this daily trading and rebalancing approach, combined with our meticulous stock selection, we aim to achieve excess returns while effectively managing risk.

1.2.4. Performance Estimate

We project an annualized return of approximately 15.5%. This estimation is derived by considering the integration of four key indicators, the real-time scoring model, and historical performance of multi factor strategies in similar market conditions.

Based on our multi-factor scoring model and the market dynamics of the SSE A-shares, we anticipate the standard deviation of returns to be around 11%. This takes into account the market's inherent volatility and the balancing effect of our long/short hedge.

Considering the daily rebalancing, real-time responsiveness to market dynamics, and our risk mitigation strategies, we forecast a maximum drawdown of around 20%. This projection is grounded on the potential downside risk encountered by similar strategies in the past and our comprehensive risk management approach.

Considering the risk-free rate of 2.48800%, and our earlier projections for annualized return and standard deviation. Thus, we estimate a Sharpe Ratio of approximately 1.18364.

1.3. Literature Review

Effective stock selection and position management are pivotal elements for successful trading in the stock market. This comprehensive literature review examines the key strategies employed by traders, encompassing multi-factor stock selection strategies, market-neutral strategies, and those grounded in the risk parity model. These strategies offer essential theoretical underpinnings to guide investors in making informed decisions regarding stock selection and portfolio adjustments, with the ultimate goal of enhancing returns while prudently managing risk.

1.3.1. Multi-Factor Stock Selection Strategies

Multi-factor stock selection strategies revolve around the incorporation of various factors when crafting stock portfolios. Pioneering work by Fama and French introduced a three-factor model consisting of the market factor, market value factor, and book-to-market ratio factor. This model underscored the predictive capacity of these factors concerning stock returns [1]. Building upon this foundation, Carhart integrated a momentum factor, further amplifying the predictive potency of multi-factor stock selection strategies [2]. Asness, Moskowitz, and Pedersen's research offered compelling evidence that multi-factor stock selection strategies consistently yield substantial excess returns on a global scale, with certain factors exhibiting commonalities across diverse markets [3]. Additionally, Hou, Xue, and Zhang introduced the Q-factor model, which provided more robust explanations for stock returns and achieved significant excess returns across multiple markets [4]. Harvey, Liu, and Zhu contributed findings suggesting that multi-factor stock selection strategies can maintain relatively stable excess returns over the long term, impervious to shifting market conditions [5].

DOI: 10.54254/2754-1169/92/20231225

1.3.2. Market-Neutral Strategies

Market-neutral strategies are devised to maintain a neutral stance toward overall market movements by concurrently assuming long and short positions on stocks or other assets. Fung and Hsieh initiated early research into the application of market-neutral strategies within hedge funds, revealing that these strategies can yield relatively stable returns. Moreover, they exhibit a low correlation with the broader market trends and possess inherent risk control capabilities [6]. Avellaneda and Lee extended this research into the realm of high-frequency trading, introducing a market-neutral strategy grounded in market microstructure [7]. This approach exploits price differentials and liquidity in the market, ultimately delivering significant gains in high-frequency trading. Chen and Zhang contributed an exploration into risk management strategies tailored to market-neutral approaches [8]. They introduced a risk-parity-based method designed to equitably distribute risk through adjustments in the weights of different assets. The findings underscored the effectiveness of this strategy in reducing the overall risk exposure of portfolios.

1.3.3. The Arbitrage Pricing Theory (APT) Model

The APT model represents a significant departure from the Capital Asset Pricing Model (CAPM) and offers a more nuanced understanding of asset returns. Proposed by Ross, the APT model posits that an asset's expected return is a linear function of various systematic risk factors, encapsulating the inherent risks associated with the asset [9]. Diverging from the CAPM, which relies solely on the market risk factor (beta), the APT accommodates multiple factors, allowing for a more comprehensive assessment of an asset's expected return. While sharing some commonalities with the CAPM in its focus on the relationship between risk and return, the APT provides greater flexibility by considering various factors that may exert influence on asset pricing.

1.3.4. The Risk Parity Model

The risk parity model seeks to achieve a balanced allocation of risk among different assets. Early work by Maillard, Roncalli, and Teiletche introduced a portfolio construction method based on risk parity, demonstrating its capacity to equitably distribute risk among portfolio assets across diverse market conditions [10]. Ledoit and Wolf applied the risk parity model to stock portfolios, utilizing a covariance matrix-based approach to recalibrate stock weights and achieve balanced risk allocation. Their research illuminated the model's effectiveness in mitigating overall portfolio risk [11]. Roncalli further expanded the risk parity model by introducing a risk-contribution-based approach, which evaluates assets' contributions to the overall portfolio risk [12]. By adjusting asset weights accordingly, this model furnishes improved control over portfolio risk.

2. Specification

2.1. E/I Factor

The E/I factor, representing the ratio of external trading volume to internal trading volume, is formulated to capture fluctuations in market demand and the trading activity of internal shareholders. The equation is expressed as:

$$E/I = \frac{External\ Trading\ Volume}{Internal\ Trading\ Volume}$$

Where,

DOI: 10.54254/2754-1169/92/20231225

* $External\ Trading\ Volume$ denotes the number of trades originating from non-insider shareholders.

*$Internal\ Trading\ Volume$pertains to the trading volume conducted by internal shareholders, including the company's management team, board members, or stakeholders owning over 10% of the company's shares.

The E/I factor offers insights into the confidence levels of external and internal traders. An increase in the E/I factor indicates that external trading volume surpasses that of the internal, possibly suggesting heightened market confidence in the stock. Conversely, a decline in the E/I ratio might imply a more vigorous trading activity by internal shareholders, potentially signaling apprehensions about the company's prospects.

In contrast to conventional factors like market capitalization, book-to-market ratio, or momentum, the E/I factor provides a unique perspective on the trading behavior of both internal and external shareholders of a company. This behavior may be a reaction to the company's fundamentals or market trends. Furthermore, compared to factors grounded on price or returns, the E/I factor might exhibit heightened sensitivity to short-term market dynamics, especially relevant in the volatile A-shares trading environment.

2.2. Qualitative analysis

The four-factor equal-weight scoring model offers a groundbreaking approach in the realm of SSE A-share investments. For investors, the strategy's comprehensive combination of the novel E/I factor and established metrics (PEG, RSI, beta) promises a deeper understanding of stock potential. Traders benefit immensely from the model's real-time responsiveness and daily rebalancing, ensuring swift actions in a volatile market. Meanwhile, institutional fund managers are presented with a novel blueprint that champions both potential returns and the essence of risk hedging, balancing reward with prudence.

The non-linear interplay between the incorporated factors and a stock's intrinsic value brings both depth and complexity. This model, deviating from traditional linear perspectives, has the potential for enhanced accuracy by embracing a factor rating approach. Initial results point towards a tug-of-war between risk management and returns. Yet, post-adjustment findings paint a promising picture, indicating the strategy's resilience and adaptability.

Regulatory bodies can glean valuable insights from the strategy's dynamics, especially its stance on short-selling, to inform future market guidelines. The two-phase analysis—initial and post-adjustment—reflects thoroughness and an unwavering commitment to excellence. In essence, this strategy offers stakeholders a fresh, innovative, and rigorously-tested tool in stock investment, emphasizing the bedrock principle of effective risk management.

2.3. Quantitative Analysis

Annualized Return---measures the average yearly return of the strategy, indicating its profitability over an annual period.

Standard Deviation of Return--- gauges the volatility or risk associated with the strategy, showing the extent of return fluctuations.

Sharpe Ratio---characterizes the investment strategy's ability to deliver returns above the risk-free rate for each unit of total risk undertaken.

Average Daily Return---provides insight into the strategy's daily performance, offering a granular view of its profitability.

Max Drawdown---measures the largest decline from a peak to a trough during a specific period, indicating the strategy's potential downside risk.

2.4. Data

2.4.1. Data Collection

In our research, we selected 2,224 stocks from the SSE A-Share to serve as our study sample, forming our stock pool. The choice of these stocks was guided by various data and attributes required by our investment strategy, including but not limited to, the daily fluctuations in beta, Price/Earnings to Growth ratio (PEG), Relative Strength Index (RSI), the ratio of External to Internal trading volumes, as well as opening and closing prices of each stock. Moreover, to set a benchmark in subsequent backtesting, we also collected data on the cumulative return rate of the CSI 300 Index.

For the division of the data sample, we designated the data from August 30, 2019 to July 21, 2022, spanning about three years, as our in-sample data, which was utilized to develop and optimize our investment model. Meanwhile, the data from July 22, 2022 to July 2023 was chosen as our out-of-sample data, purposed for verifying the performance of our model on unseen data.

All these stock data were sourced from the Choice data platform by East Money, a platform widely recognized as an authoritative source for Chinese financial market data. Using the data tools and services provided by this platform, we conducted data acquisition and preliminary processing to provide data support for our investment strategy.

2.4.2. Data Processing

Handling of Newly Listed & Delisted Stocks

In our initial data cleansing phase, we eliminated all stocks that were delisted during the period from August 30, 2019 to July 22, 2023. Secondly, given that the initial data values for newly listed stocks are zero, we chose to exclude these stocks' data and company names before their listing dates. In other words, they do not participate in the ranking of stock scores before they are officially listed. In this operation, we utilized the E/I ratio of zero as the criterion to exclude unlisted stocks, as there cannot be any external or internal trading volume before a stock is listed. This measure primarily serves to facilitate the selection of stocks for short selling, avoiding the shorting of those stocks that rank at the bottom due to a score of zero but have not yet been listed.

Data Standardization

We selected 2,224 stocks from the SSE A-Share to form our stock pool and proceeded to standardize their data. The standardization formula is as follows:

$$Y_{Kij} = \frac{X_{Kij} - X_{KMini}}{X_{KMaxi} - X_{KMini}}$$

Where, X_{Kij} represents the Kth factor data of the Jth stock on day i, X_{KMini} represents the minimum value of the Kth factor data of all stocks on day i, X_{KMaxi} represents the maximum value of the Kth factor data of all stocks on day i, and Y_{Kij} represents the standardized data of the Kth factor data of the Jth stock on day i.

2.4.3. Dataset

Below “features” of universe stocks will be used to build our strategy database.

- Stock-name(type:char)
- Stock-RSI(type:float e.g, 0.676810700834051)
- Stock-EI(type:float e.g, 0.960034638456848)

- Stock-PEG(type:float e.g, 0.402043701473152)
- Stock-BETA(type:float e.g, 0.0606451624271386)
- Stock-opening price(type:float)
- Stock-ending price(type:float)
- Risk-free rate(type:float)
- Time(type:datetime)
- Standard deviation of return(type:float)
- Stock-return rate(type:float)
- Stock-portfolio(type:char)
- Factors restriction region(type:float e.g, 0.5)
- Stock-score(type:float)
- Stock-industry(type:char)
- Stock-max drawdown(type:float e.g, -0.052345623483127)

2.5. Methodology

2.5.1. Model Hypothesis

According to the trading rules of China's securities market and the trading restrictions of the backtest model, this paper makes the following assumptions:

(1) Assume that the data obtained are true and reliable.

(2) Assume that the A-share market can be short-sold

(3) Assume that the policy can be run in real-time and updated with each transfer day

(4) Assume that all trades take place under conditions of high market liquidity, and the trading order for all target traded stocks is market order, and transactions are made for one lot

2.5.2. Strategy Implementation

First of all, we construct a comprehensive stock scoring and ranking model. For each stock i in the SSE A-Share , we define its score as the following formula:

$$Score(i) = -\frac{1}{4}X_{PEG_i} + \frac{1}{4}X_{RSI_i} + \frac{1}{4}X_{E/I_i} + \frac{1}{4}X_{beta_i} \qquad \text{-----(1)}$$

Here, PEG_i , RSI_i , E/I_i , $beta_i$ represent the stock i's Beta, relative strength index, external/internal trading volume, and price/earnings growth ratio, respectively.

Our trading signal is based on the stock score calculated by the above scoring model. For each day t, we rank the stocks based on their scores:

$$\text{Rank}(i, t) = \text{Rank}(\text{Score}(i, t)) \qquad \text{------ (2)}$$

The Rank function returns the score rank of stock i on day t. The higher the ranking, the higher the score, which means that our expected return on the stock is higher.

Next, we build a dynamically adjusted portfolio. We select the top 100 stocks on T-day for long investment, and select the bottom 100 stocks for short investment:

$$\text{Long}(t) = \{i : \text{Rank}(i, t) <= 100\} \qquad \text{------ (3)}$$

$$\text{Short}(t) = \{i : \text{Rank}(i, t) > 1900\} \qquad \text{------ (4)}$$

DOI: 10.54254/2754-1169/92/20231225

Before making an investment, we first need to determine the investment weight of each stock. For long position stocks, we use a total risk contribution (TRC) based approach to allocate capital:

$$W_{Long(i,t)} = \frac{\text{TRC(i,Cov(Long(t),t))}}{\sum_j \text{TRC(j,Cov(Long(t),t))}} \quad \text{------ (5)}$$

Where Cov(Long(t), t) represents the covariance matrix of the stock long invested on day t, and the TRC function returns the total risk contribution of stock i. The total risk contribution can be calculated by first calculating the marginal risk contribution (MRC) and then calculating it based on the MRC and weight of each stock.

For stocks that are short position, we choose to allocate capital in the form of equal weight:

$$W_{\text{Short(i,t)}} = \frac{1}{|\text{Short(t)}|} \quad \text{------ (6)}$$

Where |Short(t)| indicates the number of stocks with short interest on day t.

At the end of each trading day, we update the four-factor score of each stock using the latest stock price data and then re-calculate the ranking and investment weighting based on the score. When the market opens the next day, we will perform buy and sell operations based on the new investment weights.

2.5.3. Trading Cost

Investors maintain "long" security positions in the expectation that the stock will rise in value in the future [13]. Buyers long the stock at a low price and sell it at a higher price, and the spread could allow them to make a profit. For the transaction fee, from April 29, 2022, the overall transfer fee for stock trading will be reduced by 50%, which means the transfer fee for stock trading is 0.01‰ of the transaction amount [14].

A short sale generally involves the sale of a stock that the investors do not own and need to borrow for delivery. Aiming to make a profit or hedge the risk based on the forecast that the stock price will fall, short sellers borrow shares initially, sell them at a high price, and buy shares to repay them after the price falls. The typical fee for a stock loan is 0.30% per annum. In case of short supply, when many investors are going short on a stock, the fee may go up to 20-30% per annum [15]. In this paper, we select the minimum fee to be the actual rate.

3. Result (In-sample performance)

3.1. P&L graph & Summary statistic

Our investment strategy was deployed on an in-sample data set ranging from August 30, 2019, to July 21, 2022. We executed a long-short strategy, taking long positions in the top 100 stocks according to our composite scoring model and short positions in the bottom 100 stocks. The Profit and Loss (P&L) graph derived from this strategy is presented below:

Figure 1: In sample cumulative return [16]

As shown in Figure 1, The strategy moves almost in line with the benchmark trend until July 2020, but the yield is slightly lower overall and only higher than the benchmark yield in some periods. Since July 2020, there has been a large gap between the benchmark and its return rate. With the summary statistics shown in Table 1, during 700 trading days, the cumulative return is 1.0532 times than the original asset. Annualized return is about 1.893571% while Sharpe ratio is nearly -0.015540.

Table 1: In-sample data [16]

Category	Value
Annualized return	1.893571%
The standard deviation of return	3.213546%
Sharpe ratio	-0.015540
Max Drawdown	33.934439%
Average Daily return	0.007574%
Risk-free rate	2.4888%
Times of Trading	140000
Total Trading Days	700

3.2. Abnormal Analysis

The Sharpe ratio derived from this strategy is negative, indicating that the return of our strategy is below the risk-free rate. Concurrently, the strategy's return experienced a precipitous drop in January 2020 and May 2022. It exhibited an inverse relationship with the benchmark yield curve from May 2020 to January 2021, a deviation we attribute to multiple contributing factors:

1) Our equity universe comprises 1,417 stocks predominantly from the manufacturing sector, leading to a sectoral concentration in our portfolio. Consequently, our performance is highly correlated with the market dynamics of the manufacturing industry.

2) The onset of COVID-19 around January 2020 severely impacted the manufacturing sector, with widespread shutdowns and significant operating losses across firms. This external shock precipitated a sharp decline in our returns during that period, causing a deviation from the benchmark yield curve for a subsequent duration.

4. Refinement

To bolster our multi-factor scoring model, we've integrated a factor rating method. The four factors—Price/Earnings to Growth (PEG), Relative Strength Index (RSI), beta, and External/Internal trading volume (E/I)—are assigned ratings within five categories (Rating 1 - Rating 5). Positive coefficient

DOI: 10.54254/2754-1169/92/20231225

factors use a 5-1 rating scale, while factors with negative coefficients, like PEG, use a 1-5 scale. This approach minimizes systematic risk and enhances stock selection accuracy by mitigating undue influence from any single factor.

New data processing method:

First, we add the initial standardized data to the level assigned by the factor to obtain a new data.

Step1: Z_{Kij}=Y_{Kij}+A_{Kij},

Where A_{Kij} represents the value of the Kth factor data of the Jth stock on the i day according to the grade, and Z_{Kij} represents the basic value of the Kth factor data of the Jth stock on the i day.

Second, we standardize Z_{Kij} to let it fall between 0 and 1.

Step2: $U_{Kij}=\frac{Z_{Kij}-Z_{KMini}}{Z_{KMaxi}-Z_{KMini}}$

Where, Z_{KMini} represents the minimum value of the Kth factor data of all stocks on day i, Z_{KMaxi} represents the maximum value of the Kth factor data of all stocks on day i, and U_{ij} represents the final standardized value of the Kth factor data of the Jth stock on day i.

According to the range of factor data we obtained, the following are the rating divisions and reasons for the 4 factors:

Table 2: PEG Factor

PEG Ratio	**Rating**
PEG < 0	5
0 < PEG < 0.5	1
0.5 ≤ PEG < 1	2
1 ≤ PEG < 2	3
2 ≤ PEG < 3	4
PEG ≥ 3	5

The lower the PEG factor, the lower the price of the stock relative to the expectation of future earnings, and therefore the likely value of the investment is higher. Conversely, the higher the PEG factor, the higher the price of the stock relative to the expectation of future earnings, and thus the likely lower the investment value.

Table 3: Beta Factor

Beta Coefficient	**Rating**
Beta < 0 & 0 < Beta < 0.8	1
0.8 ≤ Beta < 1.0	2
1.0 ≤ Beta < 1.2	3
1.2 ≤ Beta < 1.5	4
Beta ≥ 1.5	5

The beta factor is a measure of the systemic risk of a stock or portfolio relative to the market. If the beta is equal to 1, then the price of the stock moves exactly with the market; If the beta is greater than 1, then the stock price will be more volatile than the market, and vice versa. According to our investment philosophy, we choose stocks with a beta close to 1, because such stocks are not too sensitive to market movements, but can also make some gains when the market rises.

DOI: 10.54254/2754-1169/92/20231225

Table 4: RSI Factor

RSI Value	Rating
$70 < RSI \leq 100$	1
$50 < RSI \leq 70$	2
$30 < RSI \leq 50$	3
$10 < RSI \leq 30$	4
$0 \leq RSI \leq 10$	5

RSI (Relative Strength Index) is a momentum indicator used to determine if a stock has been overbought or sold. When the RSI value is above 70, the stock may be overbought and there may be a risk of correction; When the RSI value is below 30, the stock may be over-sold and there may be opportunities for a rebound.

Table 5: E/I Factor

E/I Ratio	Rating
$0 < E/I < 0.75$	1
$0.75 \leq E/I < 1$	2
$1 \leq E/I < 1.5$	5
$1.5 \leq E/I < 2$	4
$2 \leq E/I < 2.5$	3
$E/I \geq 2.5$	1

The E/I factor value represents the ratio of a stock's external volume to its internal volume. A low E/I value means that the internal volume is greater than the external volume, which may cause the stock to underperform; A high E/I value means that external trading volume is greater than internal trading volume, which may result in superior stock performance.

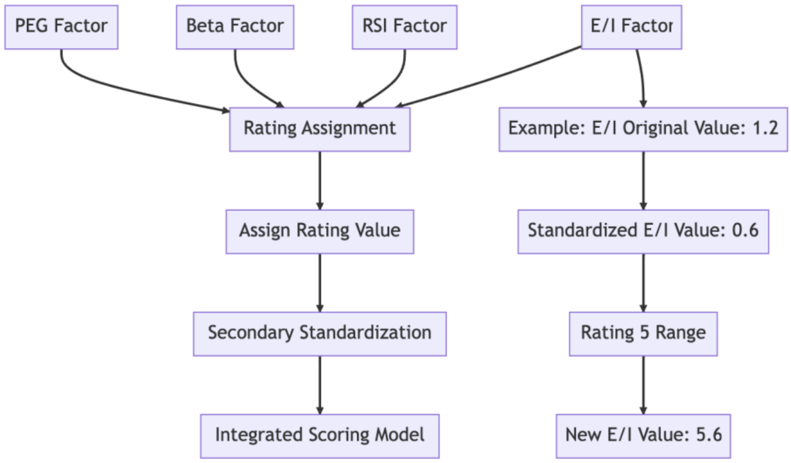

Figure 2: Factor rating process of E/I

This procedure adds selectivity to the model, enabling the score to better reflect the differences between each stock on different factors. Afterward, the numerical values of each factor of the stock are inserted as independent variables into the integrated scoring model equation.

Upon implementing the factor rating refinement, we observed a substantial alteration in our stock selection. Specifically, our daily stock pool exhibited an average change of 94.5898491%. Such a pronounced variation implies a significant reshuffling in stock selection, underscoring the weight and impact of the factor rating method.

We assigned rankings to the four factors and re-scored and re-ordered the 2,224 equities within the SSE A-share market. Leveraging the time horizon mentioned in Section 3, we recalibrated our long and short positions, subsequently recalculating the cumulative returns as illustrated in the following figure:

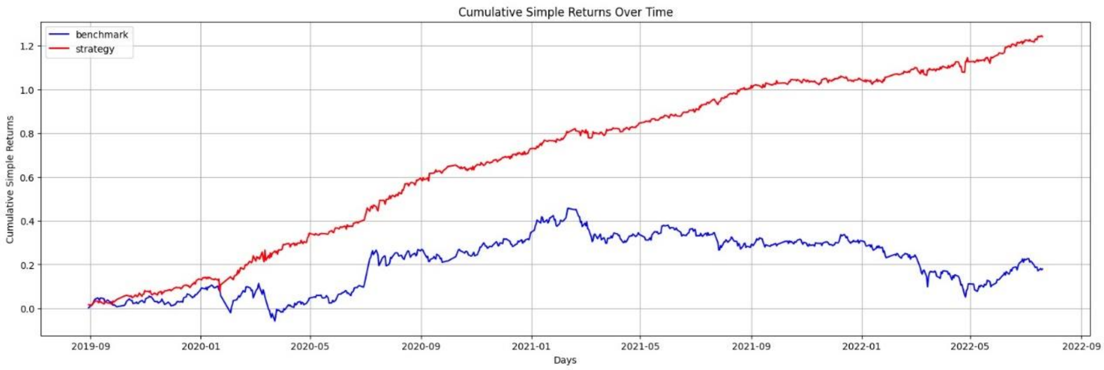

Figure 3: In sample cumulative return after refinement [16]

As shown in Figure 2, the strategy never experienced losses from April 30, 2019, to July 21, 2022. With the summary statistics shown in Table 2, during 700 trading days, the cumulative return is 2.23711 times than the original asset. Annualized return is about 44.1825% while Sharpe ratio is nearly 0.275119.

Table 6: In sample data after refinement [16]

Category	Value
Annualized return	44.182500%
The standard deviation of return	12.73%
Sharpe ratio	0.275119
Max Drawdown	54.940537%
Average Daily return	0.176730%
Risk-free rate	2.488800%
Times of Trading	14000
Total Trading Days	700

5. Conclusion

5.1. Out-of-sample Performance

5.1.1. P&L graph & Summary statistic

According to the method mentioned in refinement, we improve the overall strategy and output a new P&L from July 22, 2022 to July 21, 2023 as the out-of-sample, and compare it with the daily rate of return of the CSI 300 stock pool again. The following figure is obtained:

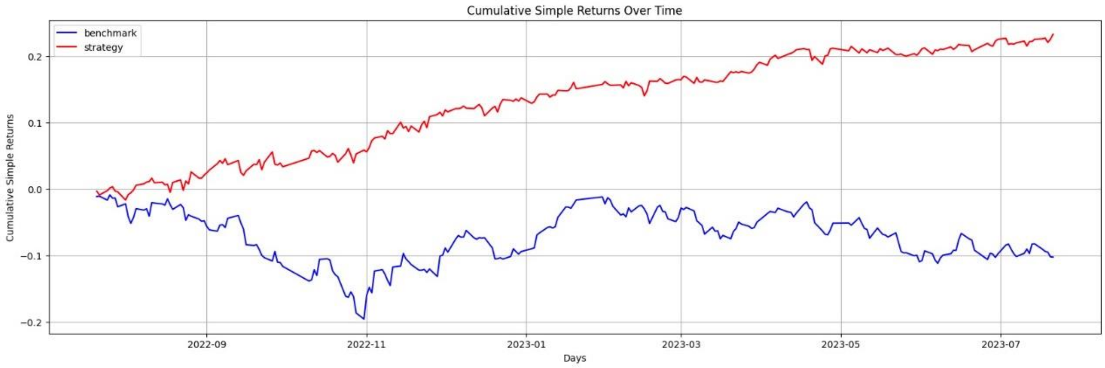

Figure 4: Out-of-sample cumulative return [16]

As shown in Figure 3, the strategy only experienced losses from July 22, 2021, to July 29, 2022, with profits for the rest of the period. With the summary statistics shown in Table 2, during 243 trading days, the cumulative return is 1.23296 times than the original asset. Annualized return is about 23.967% while Sharpe ratio is nearly 1.146165.

Table 7: Out-of-sample data [16]

Category	**Value**
Annualized return	23.967078%
The standard deviation of return	1.574092%
Sharpe ratio	1.146165
Max Drawdown	20.659890%
Average Daily return	0.095868%
Risk-free rate	2.488800%
Times of Trading	48600
Total Trading Days	243

5.1.2. Abnormal analysis

Benchmark's cumulative return is below zero, but our strategy is always growing upward. Even in certain periods, when the market has a downward trend, our strategy yield has also declined to a certain extent, but it still means that the yield trend of our strategy is opposite to the market. We believe there are several reasons for this result:

1) There are many retail investors in the A-share market, and the buying and selling of stocks is highly subjective. However, our stock selection model adopts 4 factors that cannot be categorized, and the stocks traded are more reasonable and objective.

2) Retail investors in the A-share market do not have high trading frequency, which is likely to lead to long-term holding of stocks with high fluctuations, resulting in large losses. However, our strategy takes the daily trading rate as the trading rate, which can indeed carry out arbitrage from stocks with high fluctuations while bearing high fees.

3) The A-share market is difficult to short, but our strategy needs to short 100 stocks per day, which is a good risk hedge, to obtain a large return.

5.2. Trading Recommendation

Our comprehensive stock scoring model combines four different kinds of factors and assigns them to grades, which can select stocks with high investment value to a certain extent. We trade daily and are

DOI: 10.54254/2754-1169/92/20231225

well-positioned to arbitrage some of the most volatile stocks. At the same time, our strategy selected the top 100 stocks in the stock pool to go long, and the bottom 100 stocks to go short, which achieved good risk hedging and finally the annualized return is 23.967078%. To sum up, the investment strategy described in this paper is a strategy with certain feasibility and stable profits.

5.3. Suggestions

In our scoring model, we introduce the E/I factor, and we preliminarily believe that when the external trading volume is relatively high, the stock will have greater investment value. However, in practice, such data can be easily influenced and controlled by investment companies, leading to a lack of credibility in some cases. Therefore, we recommend that investors make some mathematical changes to this variable to make it a more stable factor.

In our scoring model, we calculate the score with equal weight of each factor, but in fact, the influence of each factor is different, and equal weight calculation is a subjective method. Therefore, we suggest that traders should judge the influence weight of each factor on the potential value of stocks when adopting our trading strategy, and rematch the coefficient of the scoring model to make the whole model more objective.

References

[1] *Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. The Journal of Finance, 47(2), 427-465.*

[2] *Carhart, M. M. (1997). On persistence in mutual fund performance. The Journal of Finance, 52(1), 57-82.*

[3] *Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. The Journal of Finance, 68(3), 929-985.*

[4] *Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. Review of Financial Studies, 28(3), 650-705.*

[5] *Harvey, C. R., Liu, Y., & Zhu, H. (2016). ... And the cross-section of expected returns. Review of Financial Studies, 29(1), 5-68.*

[6] *Fung, W., & Hsieh, D. A. (2004). Hedge-fund benchmarks: A risk-based approach. Financial Analysts Journal, 60(5), 65-80.*

[7] *Avellaneda, M., & Lee, J. (2010). Statistical arbitrage in the US equities market. Quantitative Finance, 10(7), 761-782.*

[8] *Chen, L., & Zhang, H. (2020). Risk management of market-neutral strategies: A risk-parity-based approach. The Journal of Portfolio Management, 46(1), 103-113.*

[9] *Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. Journal of Economic Theory, 13(3), 341-360. doi:10.1016/0022-0531(76)90046-6.*

[10] *Maillard, S., Roncalli, T., & Teiletche, J. (2010). The properties of equally weighted risk contribution portfolios. Journal of Portfolio Management, 36(4), 60-70.*

[11] *Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. Journal of Empirical Finance, 15(5), 850-859.*

[12] *Roncalli, T. (2013). Introduction to risk parity and budgeting. CRC Press.*

[13] *Stock Purchases and Sales: Long and Short. (n.d.). Retrieved from https://www.investor.gov/introduction-investing/investing-basics/how-stock-markets-work/stock-purchases-and-sales-long-and#:~:text=Having%20a%20%E2%80%9Clong%E2%80%9D%20position%20in,stock%20you%20do%20not%20own.*

[14] *CSDC.(2022). Notice of China Clearing on lowering the fee standard of stock transaction transfer fee. Retrieved from http://www.chinaclear.cn/zdjs/gszb/202204/837e3c5031104aa099d6597ba381342 a.shtml.*

[15] *Short Selling and Stock Borrowing Costs. (n.d.). Retrieved from https://financetrain.com/short-selling-and-stock-borrowing-costs.*

[16] *Lin, Y., Ren, Y., Chen, W., Guo, L., & Li, J. (2023). Construction of a Scoring and Ranking System for Multi-Factor Quantitative Stock Selection Strategy with Daily Rebalancing: An Analysis of SSE A-shares Component Stocks. Retrieved from https://github.com/wwang0731/Multi-factor_scoring_model.git.*