Deriving momentum strategies in Chinese stock market:Using Gene Expression Programming

Yujie Zhu[[1]](#footnote-1) Tieqi Wang[[2]](#footnote-2)

**Abstract**

This paper presents how momentum strategies are generated using gene expression programming(GEP) in Chinese stock market. GEP, as a generating frame, can improve the efficiency of researches in the field of momentum strategy. In terms of empirical results, GEP generation mechanism is also outstanding. This study reveals that the GEP technique has important implications for both theory and practice.

**JEL classification:** C61; G11; G14

**Keywords:** GEP, Asset pricing, Momentum strategies, Weak method

1. **Introduction**

Dunis et al. reveal the fact that Artificial Intelligence (AI) techniques, including GEP, have been applied in many fields of finance [1]. Past papers have demonstrated the ability of using GEP to derive the trading rules. However, few, if any, papers focus on its application in asset pricing. In this research, momentum strategy and GEP are combined together to study the Chinese A-shares market.

Price momentum as documented by [2] is an asset pricing anomaly, which cannot be explained by Fama-French three-factor models [3, 4]. Since then, extensive studies have focused on two research branches, verifying the momentum effect internationally [5,6,7,8] and building momentum strategies with different sorting methods[9,10,11].

In this study, GEP is used to simulate the process of building momentum strategies and optimize the final expressions under the fundamental strategy frame. It could be treated as an abstract of the second momentum research branch and provide a unified solution to this type of problems.

The rest of the paper proceeds as follows. Section 2 gives the literature review about GEP and momentum strategy. In section 3, we illustrate the fundamental strategy frame and give GEP application. In Section 4, we describe the data set. Section 5 discusses the results. Finally, Section 6 concludes the paper.

1. **Literature review**

Momentum effect is one of the most attractive domains in asset pricing. Jegadeesh and Titman show that the momentum strategies which buy past winners and sell past losers realize significant abnormal returns over the 1965 to 1989 period [2]. This asset pricing anomaly cannot be explained by the Fama and French three-factor model. The paper uses parameter optimization method to verify the momentum effect under the self-financing strategy. The medium-term (3 to 12 months) abnormal returns are remarkable and this result is widely accepted in academic and practical fields. Thereafter, a series of studies aimed at validating whether the other countries' securities markets have the same momentum effect have been published. [5] documents international momentum effects in 12 European countries from 1980 to 1995. Iihara et al. use stocks listed on the Tokyo Stock Exchange from 1975 to 1997 to examine momentum effect. They find no momentum effect but return reversal over short period(1 month) instead [6]. Li et al. examine the momentum strategy on the S&P/ ASX200 for the sample period from January 2001 to December 2010 in Australia. They document that momentum strategy could generate significant abnormal returns [7]. In terms of Chinese stock market, Kang et al. show that short-horizon contrarian and intermediate-horizon momentum strategies both can generate statistically significant abnormal profits between 1993 and 2000 [12]. Wang shows that there’s a contrarian rather than momentum effect in Chinese A-share market during the period from 1994 to 2000 [13]. Naughton et al. demonstrate monthly momentum profits during the period from 1995 to 2005 for Shanghai Stock Exchange [14]. Pan et al. examine momentum profitability in Chinese A-share stock market by introducing return intervals concept. And this strategy could produce significantly positive profits on monthly and weekly returns [15].

Meanwhile, a branch of studies that focused on the selection of sorting indicator have also been published. Originally, Jegadeesh and Titman select stocks based on their past returns [2]. Wang and Chin use the “two-way-sorted” method which is based on past returns and past trading volume to build the momentum strategy portfolio [16]. Rachev et al. improve the strategy based on reward–risk stock selection criteria, say, standard Sharpe ratio with variance as a risk measure and alternative reward–risk ratios with the expected shortfall as a risk measure [11]. George and Hwang build the momentum strategy based on the ratio of current price to the highest price achieved within the past 12 months. As the emergence of the highest prices are stock specified, the traditional fixed window width is improved to the elastic window width, and each of the corresponding window width for target stocks would be different [10].

According to [1], GEP algorithm belongs to the wider category of evolutionary and genetic programming algorithms. “Survival of the fittest” is the principal behind this weak method. It can be applied in many areas of finance, optimizing asset allocation and searching trading rules for instance. Allen and Karjalainen use genetic programming to find and evaluate technical trading rules for daily S&P 500 index data. Although, when transaction costs considered, the rules do not earn consistent excess returns over buy-and-hold strategy in the out-of-sample test periods, its ability to identify the sign of daily returns and volatility level still stimulated a series of studies [17]. Based on Allen and Karjalainen’s working paper, Neely et al. study the trading rules in the foreign exchange market. Unlike the original paper, the trading rules’ excess returns, found by GEP, are all economically significant in out-of-sample period for each of six exchange rates. It means that GEP can detect patterns in data that are not captured by statistical models [18]. Fyfe et al. try to discuss the GEP method on Land Securities Plc for the period from 1980 to 1997. Due to the limited data sample, it really cannot derive any conclusion related to the market efficiency. But GEP, which can automatically generate the trading rules, is again confirmed to be useful in the context of investing[19].

Set those drawbacks mentioned in [17] aside, three other issues should also be pointed out. Firstly, as a weak method, these GEP strategy applications are based on general standards which do not consider knowledge specific to any context or other former researches. The Genetic operators, say mutation and crossover, are tree pruning related, which cannot guarantee the effect of fine-tuning. Secondly, past studies only focus on limited targets. Under this circumstance, no solid conclusion could be derived due to the lack of empirical supports. Finally, there is no guarantee about its theoretical global minimum convergence, but only local minimum instead.

1. **Methodology**

The former studies [17,18,19] show that GEP actually has two types of functions in practice: real and Boolean. The fundamental generic operator would not be repeated in this paper, because they are basically the same as its original edition. We only focus on the differences.

Firstly, we like to point out that the information used in final trading rules(strategies) could be from the original set which is supplied by the environment (researchers specified) or the outcome of the GEP functions. This study uses accounting statement information, historical price data series and several indicators that are based on former two types, rather than price data alone, to find the proper strategies. The principal here is simple. Even if we provide mathematical operators that supports the famous Fourier transform, if you really need this information, it is better to provide that information to the algorithm rather than force it to find out all by itself.

Secondly, we modify the former GEP application to fit in momentum strategy frame in this paper. Technical trading rule is a controversial area in finance. Meanwhile, momentum effect is an asset pricing issue that even Fama and French have to admit its challenge [4]. The comprehensive momentum strategy frame is illustrated as follows.

Table 1 : Momentum strategy frame

|  |
| --- |
| Step 1: Set the strategy minimum interval. Whether to specify the periods of the actions(selecting and holding) or not, building strategy needs this to cope with the regulation constraints.  Step 2: Select some ranking criteria to construct portfolios. Fixed or elastic assessing window could be applied here to yield the final return series data. |
| Step 3: Evaluate the Returns of Relative Strength Portfolios. |

Past studies have shown that the ranking criteria could be an expression based on the information fed in the algorithm. For example, returns over past quarters in [2], sharp ratio in [11] and the ratio of current price to the highest price achieved within the past 12 months in [10] could be thought of as some expression based on price and risk measures information set. Meanwhile, [10] also shows how to apply the elastic window width when dealing with ranking stocks and building portfolios. Therefore, we replace the user identified criteria with GEP output.

Table 2: Improved GEP momentum frame

|  |
| --- |
| Step 1: Set the strategy minimum interval.  Step 2: Constrain GEP on real function only. Supply the information set whatever the numeric data that is supposed to be useful when building the strategies.  Step 3: Create random ranking criteria for population number of times, compute the fitness for each one of them. The fittest elite number rules would be retained for the next generation.  Step 4: apply the Genetic operators(selection, mutation and crossover) to reproduce the candidates in population.  Step 5: Check if the termination condition is met (generation numbers or performance improvement criteria), otherwise go to Step 4 |

From the Evolutionary computation point of view, fitness measure could be any indicator for performance and risk analysis of financial instruments or portfolios. For instance, the excess return over the buy-and-hold strategy is applied in [17]. In this paper, we have 4 different fitness measures to direct the GEP strategies, which are annualized return, sharpe ratio, max drawdown and complexity. When two or more measures are simultaneously selected, the process would keep all candidates that at least have one measure is better than others to make final output obey pareto optimality principal to the utmost extent.

We use a population size of 200 and set no limit to the genetic tree structure. Evolution continues for 30 generations. Considering the computational capability, for [17] has only one target and this paper has nearly 3000 targets, we set the only boundary here is that the product of population size and generation size is less than 60,000. We also set the elite number 20, crossover rate 0.75 and mutation rate 0.05.

Considering the setting of the momentum strategy backtest period in [2] and the principal that learning process of the GEP method should include at least one complete market cycle, we recommend use the period from 2004-01-01 to 2013-12-31 for training, and period from 2014-01-01 to 2017-06-30 for backtest.

1. **Data and variables description**

The data are collected from Shanghai Stock Exchange(SHSE) and Shenzhen Stock Exchange(SZSE) for all stocks listed. Although these excahnges are founded around 1990, we only use the period from 2004 to 2017. The reason behind this is the stability of trading rules. the “T+1” trading rule[[3]](#footnote-3) in 1995, “price limit”[[4]](#footnote-4) in 1996 and several years of government regulation on stock price manipulation after the securities law passed in 1999, there are obvious differences in statistical features and market operation mechanism between before-and-after data periods. There are 3031 stocks that meet the criteria[[5]](#footnote-5).

As for variables supplied for the information set, there are two categories in fundamental data: financial statement information and historical transaction data. There are 106 pieces of information for each stock. Due to the limited hardware memory, only part of the variables could be loaded for each time. Therefore, this study would carry out 34 trails, roughly 20000 rounds of strategy searching, to cover the whole information set variables.

1. **Main Results**
   1. **Financial statement information**

The results of momentum strategies using financial statement information are listed below. We choose 4 criteria to analyze the strategy performance: Annualized rate of return(ROR), ActiveReturn(AR) ,Sharpe ratio(SR) and Max Drawdown(MD). The panel A shows that the RORs of FS-1, FS-2 and FS-4 for training are relatively lower than others, which uses information from income statement and cashflow statement. We use exact the same information set to run the GEP process again, their results(FS-1-1 and FS-2-1) are still lower than others, 4.6% and 5.9% respectively. Strategies with information from balance sheet involved can yield ROR at around 20% level.

Panel A

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Train** | | | | **Backtest** | | | |
| **TasK ID** | **ROR** | **AR** | **MD** | **SR** | **ROR** | **AR** | **MD** | **SR** |
| **FS-0** | **0.148** | **0.100** | **0.575** | **0.432** | **0.240** | **0.165** | **0.486** | **0.740** |
| **FS-1** | **0.051** | **0.003** | **0.520** | **0.109** | **0.225** | **0.149** | **0.420** | **0.792** |
| **FS-2** | **0.097** | **0.049** | **0.432** | **0.320** | **0.097** | **0.022** | **0.595** | **0.172** |
| **FS-3** | **0.224** | **0.176** | **0.570** | **0.823** | **0.341** | **0.266** | **0.454** | **1.096** |
| **FS-4** | **0.056** | **0.008** | **0.493** | **0.137** | **0.199** | **0.124** | **0.431** | **0.679** |
| **FS-5** | **0.202** | **0.154** | **0.572** | **0.723** | **0.247** | **0.172** | **0.434** | **0.865** |
| **FS-6** | **0.204** | **0.156** | **0.561** | **0.780** | **0.278** | **0.202** | **0.449** | **0.987** |
| **FS-7** | **0.195** | **0.147** | **0.592** | **0.727** | **0.265** | **0.190** | **0.443** | **0.968** |
| **FS-8** | **0.195** | **0.147** | **0.578** | **0.733** | **0.292** | **0.217** | **0.463** | **1.037** |
| **FS-9** | **0.194** | **0.146** | **0.579** | **0.751** | **0.263** | **0.188** | **0.476** | **0.925** |
| **FS-1-1** | **0.046** | **-0.001** | **0.515** | **0.083** | **0.238** | **0.161** | **0.420** | **0.893** |
| **FS-2-1** | **0.059** | **0.011** | **0.489** | **0.152** | **0.227** | **0.150** | **0.413** | **0.801** |

Compared with results from backtest, only FS-2 which uses cashflow statement information yields ROR at 9.7%. others’ RORs are all nearly at 20%, FS-3 even exceeds 30%. Considering the MD, the results of training and backtest are all in the range of 40%~60%. Nearly all samples except for FS-1-1 have positive AR, which means they beat the market. SRs show no consistency for most of the sample.

GEP can generate momentum strategies that performs better than market index based on financial statement information. But income statement and cashflow statement have limited effect.

* 1. **Transaction data of the day**

Panel B shows the results of momentum strategies based on daily transaction data. Although they use the exact the same info to yield strategies, the results are different as expected. For training part, their ROR are all above 20% level. They beat the market by 17% at least. The backtest results show similar pattern, with ROR above 20% level and beat the market by 16% at least. Considering the MD, the results of training are all around 60%. But backtest results are in the range of 30%~45%. SRs again show no consistency for most of the sample, four of which have higher backtest results.

Panel B

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Train** | | | | **Backtest** | | | |
| **ID** | **ROR** | **AR** | **MD** | **SR** | **ROR** | **AR** | **MD** | **SR** |
| **TS-0** | **0.225** | **0.177** | **0.689** | **0.795** | **0.288** | **0.213** | **0.326** | **1.043** |
| **TS-1** | **0.253** | **0.205** | **0.676** | **0.903** | **0.252** | **0.176** | **0.377** | **0.901** |
| **TS-2** | **0.227** | **0.179** | **0.675** | **0.798** | **0.309** | **0.234** | **0.329** | **1.097** |
| **TS-3** | **0.219** | **0.171** | **0.578** | **0.701** | **0.239** | **0.164** | **0.452** | **0.814** |
| **TS-4** | **0.236** | **0.188** | **0.668** | **0.836** | **0.293** | **0.218** | **0.321** | **1.058** |

GEP using data from transaction and financial statement can both generate momentum strategies that beat the market index at similar ROR level.

* 1. **Data from financial statement and transaction**

When we combine the information from financial statement with transaction dataset, the ROR results are all improved. The best performance is FTS-0 which uses the information from balance sheet and transaction dataset. Its ROR is nearly 50%, which is two times bigger than its individual tasks: FS-3 and TS series. The tasks that use info from income statement, cashflow statement and transaction data, also achieve better ROR level.

Panel C

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Train** | | | | **Backtest** | | | |
| **ID** | **ROR** | **AR** | **MD** | **SR** | **ROR** | **AR** | **MD** | **SR** |
| **FTS-0** | **0.518** | **0.470** | **0.568** | **1.589** | **0.316** | **0.241** | **0.403** | **1.099** |
| **FTS-1** | **0.314** | **0.266** | **0.649** | **1.126** | **0.660** | **0.586** | **0.385** | **2.283** |
| **FTS-2** | **0.387** | **0.339** | **0.694** | **1.034** | **0.276** | **0.201** | **0.462** | **0.845** |
| **FTS-3** | **0.386** | **0.338** | **0.659** | **1.078** | **0.300** | **0.225** | **0.426** | **0.982** |
| **FTS-A-0** | **0.376** | **0.328** | **0.672** | **1.080** | **0.377** | **0.299** | **0.414** | **1.365** |
| **FTS-A-1** | **0.383** | **0.335** | **0.653** | **1.202** | **0.266** | **0.188** | **0.473** | **0.912** |
| **FTS-0-1** | **0.319** | **0.271** | **0.673** | **0.993** | **0.219** | **0.142** | **0.445** | **0.741** |
| **FTS-0-2** | **0.498** | **0.450** | **0.610** | **1.265** | **0.326** | **0.249** | **0.398** | **1.160** |

There are two points that should be mentioned. First, when dealing with all information from financial statement and transaction data, the strategies FTS-A-0 and FTS-A-1 have similar RORs but lower than FTS-0. Second, we use exact the same info and fitness functions to yield strategy FTS-0-1, whose ROR is obviously lower than FTS-0. The implication behind this is that GEP frame parameters are probably not enough to cope with the complexity of this trial. Therefore, we constrain the fitness function to only one criteria: ROR. FTS-0-2 then yields ROR at nearly 50% again.

Considering the MD, the results of FTS show similar pattern like FS and TS series, the range from 55% to 70% for training and 30% to 50% for backtest. SRs also show no consistency for most of the sample.

* 1. **Technical indicators with classic parameters**

Previous studies also point out the fact that technical analysis could be useful when dealing with building strategies. There is no way to enumerate all technical indicators available. Therefore, we only put dozens of technical indicators with their classic parameters in information set. The purpose of this section is to find out whether there is any indicator can yield better performance than data from financial statement or transaction. Once the fact is confirmed, we will stop the searching process.

Panel D shows that most of the technical indicators have the similar effect on building momentum strategies. Their RORs are about 20% level which is the same as section 5.1 and 5.2. However, TID-2 and TID-3 show opposite results for training and backtest. TID-2’s training ROR is twice of TID-3 in amount. But TID-2’s backtest ROR is only one half of TID-3’s. This comparison indicates that the consistency of any strategy cannot be guaranteed. Even if the pattern is still functional or we lack the evidence to prove its invalidity, we still cannot deny the fact that the patterns do shift. TID-6’s ROR reaches 40% level, which is higher than any sample from strategies of financial statement or transaction data.

Panel D

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Train** | | | | **Backtest** | | | |
| **ID** | **ROR** | **AR** | **MD** | **SR** | **ROR** | **AR** | **MD** | **SR** |
| **TID-0** | **0.237** | **0.189** | **0.625** | **0.849** | **0.322** | **0.248** | **0.355** | **1.118** |
| **TID-1** | **0.222** | **0.174** | **0.560** | **0.750** | **0.231** | **0.156** | **0.458** | **0.792** |
| **TID-2** | **0.275** | **0.227** | **0.631** | **0.758** | **0.098** | **0.023** | **0.402** | **0.263** |
| **TID-3** | **0.131** | **0.083** | **0.664** | **0.355** | **0.205** | **0.130** | **0.425** | **0.663** |
| **TID-4** | **0.236** | **0.188** | **0.621** | **0.856** | **0.311** | **0.237** | **0.362** | **1.083** |
| **TID-5** | **0.239** | **0.191** | **0.616** | **0.858** | **0.318** | **0.243** | **0.356** | **1.099** |
| **TID-6** | **0.400** | **0.352** | **0.654** | **1.124** | **0.281** | **0.206** | **0.452** | **0.942** |

MDs lie within exact the same range in section 5.3. SRs show no consistency again. No matter what information set is chosen, MD and SR show similar results. It is the momentum strategy structure rather than expressions governed by GEP frame that plays a more important role here.

* 1. **Historical transaction information**

Some previous studies only use historical price series to build trading rules. Although their trading rules focus on market timing which is different from the momentum strategy, we could also apply the GEP to similar information set.

HSD-0, HSD-1 and HSD-2 use exact the same historical transaction information to yield momentum strategies. HSD-3 adds some other technical indicators. The training RORs of these tasks are quite similar, all around 23%. Their backtest RORs are also quite the same at 32% level except for HSD-3 whose value is 29%. Panel E shows the fact that momentum strategies use historical information set yield similar ROR. When adding other information set whose performance is dominated by historical transaction information, the ROR shows no improvement.

Panel E

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Train** | | | | **Backtest** | | | |
| **ID** | **ROR** | **AR** | **MD** | **SR** | **ROR** | **AR** | **MD** | **SR** |
| **HSD-0** | **0.239** | **0.191** | **0.625** | **0.856** | **0.322** | **0.244** | **0.355** | **1.120** |
| **HSD-1** | **0.237** | **0.189** | **0.626** | **0.847** | **0.322** | **0.244** | **0.351** | **1.115** |
| **HSD-2** | **0.237** | **0.189** | **0.628** | **0.841** | **0.319** | **0.241** | **0.356** | **1.096** |
| **HSD-3** | **0.236** | **0.188** | **0.627** | **0.844** | **0.292** | **0.215** | **0.389** | **1.023** |

Considering the MD, the results of HSD show more concentrated effect, nearly all around 62% for training and 35% to 39% for backtest. SRs also show no consistency for most of the sample.

Although they also beat the market, the concentration of the results is quite different from former sections. The implication behind this is probably that the solution space built on this specific information set is more flat than former sections’ examples. Their local extremums are pretty at the same level could be one explanation. This could also mean that complex expressions, which are frequently caused by overfitting, may not contribute to the strategy performance at the training phase at all under some circumstances.

* 1. **Momentum effect V.S Fama-French 3-factor**

[4] demonstrates the methodology to test whether anomalies could be explained by Fama-French 3-factor model(FF3). We use the same frame to test GEP momentum strategies. Panel F shows FTS-0’s results of 3-factor model. For Tradable Market Value weighted FF3 model, the t value is 4.625 and P value is 3.89e-06. Meanwhile, the t value is 4.652 and P value is 3.43e-06 for Market Capitalization weighted FF3 model. The null hypothesis that intercept is equal to zero is rejected. As found in [4], the intercepts are also strongly positive. It means FF3 model fails to explain the FTS-0 return series.

Panel F: **FTS-0 3 factors analysis results**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Tradable Market Value Weighted** | | | **Market Capitalization Weighted** | | |
| **Estimate** | **t value** | **Pr(>|t|)** | **Estimate** | **t value** | **Pr(>|t|)** |
| **(10-3)** | 0.919 | 4.625 | 3.89e-06 \*\*\* | 0.926 | 4.652 | 3.43e-06 \*\*\* |
| **Rmrf** | 0.657 | 59.842 | < 2e-16 \*\*\* | 0.661 | 59.177 | < 2e-16 \*\*\* |
| **Smb** | 0.599 | 20.066 | < 2e-16 \*\*\* | 0.670 | 23.803 | < 2e-16 \*\*\* |
| **Hml** | 0.129 | 3.332 | 0.000873 \*\*\* | 0.160 | 4.910 | 9.58e-07 \*\*\* |

**Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1**

Compared with previous studies of momentum effect in Chinese A-share market [12,13,14,15], GEP momentum frame can yield more significant results. Panel G shows all tasks’ FF3 intercept statistics in this paper. More than 85% of the GEP instances can reject the null hypothesis at 0.05 significant level. Several facts should be mentioned here. First, some of them are in FS section which means FF3 can even be improved only within financial statement information set. Second, the significant level of Hml factor for some instances are relatively lower than other two factors. It cannot even reject the null hypothesis in task TID-0 at any reasonable level. As pointed out in [4], FF3 is just a model. Future work should look for a richer model with more risk factors. GEP frame could be applied under this circumstance.

Panel G:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **TMV**  **Intercept** | | **MC**  **Intercept** | |  | **TMV**  **Intercept** | | **MC**  **Intercept** | |
| **ID** | **t** | **Pr(>|t|)** | **t** | **Pr(>|t|)** | **ID** | **t** | **Pr(>|t|)** | **t** | **Pr(>|t|)** |
| FS-0 | -0.222 | 0.824 | 0.573 | 0.567 | FTS-0 | 4.625 | 3.89e-06 \*\*\* | 4.652 | 3.43e-06 \*\*\* |
| FS-1 | -0.073 | 0.942 | 0.275 | 0.783 | FTS-1 | 6.327 | 2.85e-10 \*\*\* | 6.518 | 8.27e-11 \*\*\* |
| FS-2 | 0.023 | 0.982 | 0.303 | 0.762 | FTS-2 | 2.947 | 0.003\*\* | 2.981 | 0.002 \*\* |
| FS-3 | 2.962 | 0.003 \*\* | 3.480 | 0.000 \*\*\* | FTS-3 | 3.432 | 0.000 \*\*\* | 3.436 | 0.000 \*\*\* |
| FS-4 | -0.157 | 0.875 | 0.205 | 0.838 | FTS-4 | 3.478 | 0.000 \*\*\* | 3.556 | 0.000 \*\*\* |
| FS-5 | 2.012 | 0.044 \* | 2.283 | 0.022 \* | FTS-5 | 3.028 | 0.002 \*\* | 3.171 | 0.001 \*\* |
| FS-6 | 2.429 | 0.015 \* | 2.715 | 0.006 \*\* | FTS-6 | 2.527 | 0.011 \* | 2.684 | 0.007 \*\* |
| FS-7 | 1.975 | 0.048 \* | 2.191 | 0.028 \* | FTS-7 | 3.342 | 0.000 \*\*\* | 3.352 | 0.000 \*\*\* |
| FS-8 | 2.484 | 0.013 \* | 2.825 | 0.004 \*\* | TID-0 | 3.414 | 0.000 \*\*\* | 3.418 | 0.000 \*\*\* |
| FS-9 | 1.997 | 0.045 \* | 2.203 | 0.027 \* | TID-1 | 2.950 | 0.003 \*\* | 2.882 | 0.003 \*\* |
| TS-0 | 2.363 | 0.018 \* | 2.342 | 0.019 \* | TID-2 | 2.364 | 0.018 \* | 2.362 | 0.018 \* |
| TS-1 | 2.496 | 0.012 \* | 2.485 | 0.013 \* | TID-3 | 1.792 | 0.073 | 1.893 | 0.058 |
| TS-2 | 2.536 | 0.011 \* | 2.501 | 0.012 \* | TID-4 | 2.885 | 0.003 \*\* | 2.818 | 0.004 \*\* |
| TS-3 | 2.065 | 0.039 \* | 2.006 | 0.045 \* | TID-5 | -1.170 | 0.242 | -0.492 | 0.623 |
| TS-4 | 2.596 | 0.009 \*\* | 2.564 | 0.010 \* | TID-6 | 2.948 | 0.003 \*\* | 2.880 | 0.004 \*\* |
| HSD-0 | 2.981 | 0.002 \*\* | 2.913 | 0.003 \*\* | HSD-2 | 2.935 | 0.003 \*\* | 2.868 | 0.004 \*\* |
| HSD-1 | 2.713 | 0.006 \*\* | 2.649 | 0.008 \*\* | HSD-3 | 2.876 | 0.004 \*\* | 2.807 | 0.005 \*\* |

**Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1**

1. **Conclusion**

As an empirical study, this paper shows that the Chinese A-share market does have momentum effect, when we consider the momentum definition as “buying winners and selling losers”. The sort and buy frame can incorporate all kinds of information set available. Each set or their combinations can yield better performance series data than market index which cannot be explained by FF3 model.

The more important issue we want to point out is that Evolutionary Computation, which is one branch of AI techniques could be applied in strategy building process when dealing with asset pricing problem. This combination can level up the efficiency of former related area of studies, provides more prudent evidence to accept or reject any given hypothetical conclusion. Specifically, whether there is a momentum effect in Chinese A-share market in this paper or follow the implication of improving FF3 model mentioned in [4] could be the specific application scenarios for GEP technique.

**References**

[1] C. L. Dunis, P. W. Middleton, A. Karathanasopolous, and K. A. Theofilatos (Eds.), “Artificial Intelligence in Financial Markets: Cutting Edge Applications for Risk Management, Portfolio Optimization and Economics”. Springer. 2017

[2] N. Jegadeesh, and S. Titman, “Returns to buying winners and selling losers: Implications for stock market efficiency”. The Journal of finance,48(1), 1993, 65-91.

[3] E. F. Fama, and K. R. French, “Common risk factors in the returns on stocks and bonds”. *Journal of financial economics*, *33*(1), 1993, 3-56.

[4] E. F. Fama, and K. R. French, “Multifactor explanations of asset pricing anomalies”. The journal of finance, 51(1), 1996, 55-84.

[5] K. G. Rouwenhorst, “International momentum strategies”. *The Journal of Finance*, *53*(1), 1998, 267-284.

[6] Y. Iihara, H. K. Kato, and T. Tokunaga, “The winner–loser effect in Japanese stock returns”. *Japan and the World Economy*, *16*(4), 2004, 471-485.

[7] B. Li, T. Stork, D. Chai, M. S. Ee, and H. N. Ang, “Momentum effect in Australian equities: Revisit, armed with short-selling ban and risk factors”.*Pacific-Basin Finance Journal*, *27*, 2014,19-31.

[8] C. H. D. Hung and A. N. Banerjee, “How do momentum strategies ‘score’ against individual investors in Taiwan, Hong Kong and Korea?”. Emerging Markets Review, 21, 2014, 67-81.

[9] C. Wang and S. Chin, “Profitability of return and volume-based investment strategies in China's stock market”. *Pacific-Basin Finance Journal*, *12*(5), 2004, 541-564.

[10] T. J. George and C. Y. Hwang, “The 52‐week high and momentum investing”. *The Journal of Finance*, *59*(5), 2004, 2145-2176.

[11] S. Rachev, T. Jašić, S. Stoyanov, F. J. Fabozzi, “Momentum strategies based on reward–risk stock selection criteria”. *Journal of Banking & Finance*, *31*(8), 2007, 2325-2346.

[12] J. Kang, M. H. Liu, and S. X. Ni, “Contrarian and momentum strategies in the China stock market: 1993–2000”. *Pacific-Basin Finance Journal*, *10*(3), 2002, 243-265.

[13] C. Wang, “Relative strength strategies in China's stock market: 1994–2000”. *Pacific-Basin Finance Journal*, *12*(2), 2004, 159-177.

[14] T. Naughton, C. Truong, and M. Veeraraghavan, “Momentum strategies and stock returns: Chinese evidence”. *Pacific-Basin Finance Journal*, *16*(4), 2008, 476-492.

[15] L. Pan, Y. Tang, and J. Xu, “Weekly momentum by return interval ranking”. *Pacific-Basin Finance Journal*, *21*(1), 2013, 1191-1208.

[16] C. Wang, and S. Chin, “Profitability of return and volume-based investment strategies in China's stock market”. *Pacific-Basin Finance Journal*, *12*(5), 2004, 541-564.

[17] F. Allen, and R. Karjalainen, “Using genetic algorithms to find technical trading rules”. *Journal of financial Economics*, *51*(2), 1999, 245-271.

[18] C. Neely, P. Weller, and R. Dittmar, “Is technical analysis in the foreign exchange market profitable? A genetic programming approach”. Journal of financial and Quantitative Analysis, 32(04), 1997, 405-426.

[19] C. Fyfe, J. P. Marney, and H. F. Tarbert, “Technical analysis versus market efficiency-a genetic programming approach”. Applied Financial Economics, 9(2), 1999, 183-191.

**Appendix**

|  |  |  |
| --- | --- | --- |
| **ID** | **Task\_ID\_on\_GPXServer** | **Information set** |
| FS-0 | 59508b58f8975e0079bd06c6 | Income statement & NO. of shares |
| FS-1 | 594cdc7cf8975e000c2b26ec | Income statement |
| FS-2 | 5949fec2888da30277245056 | Cashflow statement |
| FS-3 | 59509f27f8975e00596582b4 | Balance sheet |
| FS-4 | 5950a0c8f8975e004afbd76d | Cashflow statement & Income statement |
| FS-5 | 5951c19d1ce636000e9451b6 | Balance sheet & Income statement |
| FS-6 | 5951c4f3a8959500166575d9 | Balance sheet & Cashflow statement |
| FS-7 | 5951c88da895950011b422c5 | Balance sheet & Income statement &Cashflow statement |
| FS-8 | 59520dfca895950069e6b55a | Above |
| FS-9 | 59520de5a8959500893d84e7 | Above |
| TS-0 | 5950a286f8975e0099c45e0d | Transaction data |
| TS-1 | 59521c04a89595008c3c7093 | Above |
| TS-2 | 59521bf4a8959500d3b2a530 | Above |
| TS-3 | 59521f0ba89595008c3c709f | Above |
| TS-4 | 59521f19a895950102aaa063 | Above |
| FTS-0 | 5949e66a888da3021ac10890 | Balance sheet & Transaction data |
| FTS-1 | 594a2bfb888da302b3581587 | NO. of shares & Transaction data |
| FTS-2 | 5949feea888da30286775561 | Cashflow statement & Transaction data |
| FTS-3 | 5949f8a0888da302687b6a18 | Income statement & Transaction data |
| FTS-4 | 5959adf61b4a060099c5e1bc | Balance sheet & Transaction data  Cashflow statement & Income statement |
| FTS-5 | 5959b1431b4a0600a8fa6d0a | Above |
| FTS-6 | 5959b1971b4a060099c5e1c4 | Balance sheet & Transaction data |
| FTS-7 | 5959bc691b4a060160f19a41 | Balance sheet & Transaction data(ROR) |
| TID-0 | 5949e152888da3021ac1088c | TID SET 00 |
| TID-1 | 5949e5e2888da3025941b930 | TID SET 01 |
| TID-2 | 594a3b6e888da302d2d74380 | TID SET 02 |
| TID-3 | 594a3b56888da302e16f0ef6 | TID SET 03 |
| TID-4 | 594a3240888da302d2d7437c | TID SET 04 |
| TID-5 | 594a321b888da302e16f0ef2 | TID SET 05 |
| TID-6 | 594a2c33888da302c2745887 | TID SET 06 |
| HSD-0 | 595b5c601b4a0601fe73e30a | HSD SET 00 |
| HSD-1 | 595b20cf1b4a0601ef295bb1 | Above |
| HSD-2 | 5959ed261b4a06019c928bca | Above |
| HSD-3 | 5959ec761b4a0601af7d2ca4 |  |

|  |  |
| --- | --- |
| **Information set** | **Indicators** |
| Income statement | profit\_TotalRevenue; profit\_Revenue ;profit\_TotalCOGS;  profit\_COGS; profit\_operatingProfit; profit\_NonoperatingExpen; profit\_NonoperatingRevenue;  profit\_NetIncome; profit\_NetIncomeAttrP |
| Cashflow statement | CF\_CFOIn;CF\_CFOOut;CF\_CFO; CF\_CFIIn;CF\_CFIOut;  CF\_CFI;CF\_CFFIn;CF\_CFFOut;CF\_CFF;CF\_C&CE(NI) |
| Balance sheet | BS\_C&CE;BS\_ Receivables;BS\_Inventory;BS\_CurrentAssets;  BS\_FixedAsset;BS\_CIP;BS\_TotalAssets;BS\_ShortLoan;  BS\_Payables;BS\_CurrentDebt;BS\_LongTermLoan;  BS\_UndistributedProfit;BS\_TotalEquityAttriParentCompany |
| NO. of shares | Share\_totalShares; Share\_tradableA |
| Transaction data | High; Open; Low; Close; Volume |
| TID SET 00 | High; Open; Low; Close; Volume;  PER;PBR;PSR;PCFR;OPR;CRatio;ROE;TAT |
| TID SET 01 | MA;EMA;LLV;HHV;MACD;KDJ\_RSV;KDJ\_K  KDJ\_D;KDJ\_J;AVEDEV;STD;SUM;DMA;SMA |
| TID SET 02 | RWI\_H;RWI\_L;ATR;VC;RVI;VHF |
| TID SET 03 | OBV;ADL;ASI;NVI;TVI;CR;CRMA;KVO;CMF;  ARBR\_AR;ARBR\_BR |
| TID SET 04 | BOLL\_MID;BOLL\_UPPER;BOLL\_LOWER;EOM;PCU;PCD;  AROON\_UP ; AROON\_DOWN;MFI |
| TID SET 05 | DMI\_PDI;DMI\_MDI;DMI\_ADX;DMI\_ADXR;MTM;RMI;CCI;  IMI;WAD;PSY |
| TID SET 06 | WC;QSTICK;TRIX;WMA;SMI;WR;RSI;BIAS;SO |
| HSD SET 00 | MA(Close,2); MA(Close,3); MA(Close,4); MA(Close,6);  MA(Close,8); MA(Close,10);High;Open;Low;Close;Volume |
| HSD SET 03 | MA(Close,2); MA(Close,3); MA(Close,4); MA(Close,6);  MA(Close,8); MA(Close,10);High;Open;Low;Close;Volume  WC; QSTICK; TRIX; WMA; SMI |

1. 朱玉杰(in Chinese). School of Economics and Management, Tsinghua University. e-mail: zhuyj@sem.tsinghua.edu.cn [↑](#footnote-ref-1)
2. 王铁琪(in Chinese). School of Economics and Management, Tsinghua University. e-mail: tieqiw@hotmail.com [↑](#footnote-ref-2)
3. Securities purchased by investors shall not be resold before settlement. [↑](#footnote-ref-3)
4. The Exchange imposes the daily price limit on trading of stocks and mutual funds, with a daily price up/down limit of 10% for stocks and mutual funds and a daily price up/down limit of 5% for stocks under special treatment (ST shares or \*ST shares). [↑](#footnote-ref-4)
5. Thanks gpxtrade.com for the support of professional test environment. [↑](#footnote-ref-5)