The Effect of Momentum Strategy on the Stock Market - Based on the Linear Regression Model

Xiaonan Chen(✉)

Accounting Management, Scnu University, Guangzhou, Guangdong, China
chenxiaonan@fjmu.edu.cn

Abstract. With the development of the Internet and the increasing maturity of computer programming technology, the use of computer programming combined with mathematical models, which can assess the risk and return of financial markets, thus predicting the future market direction became a popular topic of interest. Quantitative finance is to fully capture the data changes in the financial market by computer and import the data into its designed mathematical model, so as to produce a series of accurate data results. Investors can make reasonable risk-return assessment based on the quantified data, making investment more rational. In this paper, people choose one of the strategies of quantitative finance, the momentum strategy, by selecting 100 stocks in the S&P 500, bringing them into the mathematical model for analysis, using parameters such as Sharpe ratio and max drawdown, and visualizing the data results to observe its overall direction with images. The results show that using the momentum strategy for the analysis, the overall trend of the stock is up and profitable, but after adding the transaction cost, the overall trend of the stock is down as the weight of the transaction cost increases, which means that it will face a large loss.

Keywords: Momentum Strategy · Linear Regression · Quantitative Finance

1 Introduction

With the development of the economy, the progress of the times and the renewal of the mindset, how to manage one's property and make reasonable investments to achieve the purpose of preserving the value and making money has become a very relevant subject for the people. In making investment decisions, people are most concerned about two issues, one is the risk and the other is the return. In the face of various investment products such as stocks, funds, bonds, futures, etc., whether people can fully evaluate the risks and returns for rational investment, and whether people can reduce the risk of investment as much as possible, has become a big problem for the investment community. Quantitative finance provides a pretty good solution to this problem [1].

Before introducing quantitative finance, this paper introduces "portfolio", any single investment product has its own limitations, for example, while investing stock can get a higher return you need to burden more risk. At the same time, though investing an insurance may not seem pretty profitable, people even can neglect it's risk. Through the

S. Tehseen et al. (Eds.): ICEDBC 2023, AEBMR 258, pp. 110–119, 2024.
https://doi.org/10.2991/978-94-6463-246-0_13

concept of portfolio, people can form a better combination about their assets. People can select 30% of their assets to invest in high-risk and high-yield products such as stocks, select 50% of their assets to invest in low-risk, and low-yield products such as bonds, and the remaining 20% is used for their cash flow for emergencies [2]. Such a portfolio model can quite effectively divide our risk, so that the risk as far as possible to keep within their own acceptable range, effectively enhance their asset flexibility, to achieve both the preservation of value, but also money to make money efficient investment and quantitative finance to build an effective portfolio, to achieve controlled risk and considerable benefits provide more possibilities [3].

Quantitative financial investment is different from traditional investment, traditional investment depends on the subjective judgment of people, relying on the decision maker's grasp of the market situation, this kind of investment method as far as possible to absorb objective data for subjective choice of risk is obvious, one of the objective data collected by the decision maker has limitations prone to inaccurate judgment, and the second decision maker is vulnerable to the influence of human factors error in judgment, resulting in a major collapse of the situation. Compared with traditional investment, quantitative financial investment constructs risk-return assessment models through mathematical modeling, and uses computer programs such as python to compile and execute the results, the process is completely by the computer's own judgment and decision-making, and its advantages are obvious [4]. Third, through quantification this paper can more accurately know the risk and return of investment projects, thus achieving better risk control.

The history of quantitative financial investment dates back to the 1950s, when Harry Markowitz was the first to propose a quantitative model of portfolio theory in 1952, to the establishment of the capital pricing model (taking risk-free rate of return into account), to Sharpe's alpha and beta, and then to Fama's multi-factor model, the organic combination of academic research and market practice has made quantitative According to the relevant statistics, quantitative investment has occupied 30% of foreign investment products [1].

Compared with the rapid development of quantitative research in foreign countries, China started late in the field of quantitative research, and because the domestic market is still in the stage of continuous reform and development, most of the quantitative financial products in China are still based on arbitrage strategies, and data acquisition has a common lack of characteristic data mining methods, and the brand positioning of quantitative finance, brand personalization construction, etc. are still in the initial development stage [5]. The development of quantitative investment in China can be said to be a long and difficult road.

This paper focus on one of the strategies in quantitative financial investment - momentum strategy, first of all, a simple introduction to the principle of momentum (buy when the price is on an upward trend, sell when the price is on a downward trend), in a brief overview of several important parameters for judging risk-return. For example, sharp ratio: the higher the Sharpe ratio represents a large return at the same time, the risk is smaller Linear regression: mean regression equation through the attribution of past stock information to predict the future stock trend) then selected the S&P 500 index stocks inside the quantitative analysis of the momentum effect strategy, and then the transaction cost variable Then the transaction cost variable is added to the mathematical model to observe the effect of transaction cost on the momentum effect [6, 7]. The results are compared with the hypothetical results of the momentum effect strategy, they are analyzed in depth to see if there are problems such as look ahead bias in the process of back testing, or whether the number of stocks selected is too small to make the data convincing.

2 Method

2.1 Data Mining and Basic Data Processing

With the aim of performing this strategy in the real dataset, the first thing people need to do is the data selecting. In this paper, people select 100 stocks in the S&P 500 and set a looking period for 20 years, which from 2000 to 2020. This kinds of stocks tend a have a good performance as they have still existed for 20 years. For this whole dataset, people split the data into two parts, the first part is used for the training data set, people make use of this dataset to build a training linear regression model, and the second part is used for the testing dataset. In the testing data set, people can program this data and form a testing linear regression model. After this we can make a comparison between the two linear regression results to see whether the momentum strategy work well.

2.2 Linear Regression Model

In the momentum strategy, people use linear regression model to analyze [8, 9]. In this paper, they built two models, a basic model without transaction cost and the advanced costs that take transaction costs into consideration. For the linear regression model, there are some parameters in this formula should be sincerely introduced. The first parameter is called Ret + 1 which stands for expected return of the portfolio at time. The second parameter is called α. α stands for the intercept in the linear regression model. It also represents the expected return of the portfolio, when all the factors have the value of zero. The third parameter is called M. M is the number of factors included in the linear regression model. The fourth parameter is βi, it stands for the coefficients or weights allocated to each of the M factors, so it also represents the contribution of each factor to the expected return of the portfolio. The fifth parameter is called the moving average, moving average is a technical indicator that uses statistical analysis to average the prices of securities over a certain period of time and connect the average values over time to form a moving average. Moving average is used to observe the trend of the price movement of securities. It can be calculated by the second equation or the third equation. For those equation, they both have N. N stands for the look back period in the dataset. In the dataset, N is 20 years. The sixth parameter is called EMAT, this is an exponential moving average factor that is used in the momentum strategy. The EMAT calculates the average return of the portfolio over the past and periods with a decay factor. δ are applied to have more weights to recent returns. The seventh parameter is called ε_t, it stands for the unpredictable noise that can't be calculated in the linear regression model. Except the linear regression model, people will use the mean variance allocation model. People

Table 1. Parameter setting.

Parameters	Meanings
the sharp ratio	use the expected return to divide the standard deviation
Max-drawdown	Through this principle, people can see the volatility of the investment product. A larger max-drawdown means that the company needs a huge upward rebound in order to keep the company from losing money
Kurtosis	kurtosis is a statistic that describes the steepness or smoothness of the data distribution. By calculating kurtosis, people can determine whether the data distribution is steeper or smoother compared to the normal distribution.
skewness	Skewness, also known as skewness coefficient, is a measure that describes the direction and degree of skewness of the data distribution. It reflects the asymmetry of the data.

bring the dataset into the formula, they can find the exact values of alpha, beta, Sigma, R-square for each companies [3].

$$R_{t+1}^{e} = \alpha + \sum\nolimits_{i=1}^{M} \beta_i f_t^i + \varepsilon_{t+1} \tag{1}$$

$$MA_t^N = \frac{1}{N} \sum\nolimits_{s=t-N}^{t} R_s^e \tag{2}$$

$$EMA_t^N = \frac{1}{N} \sum\nolimits_{s=t-N}^{t} (1-\delta)^{t-s} R_s^e \tag{3}$$

2.3 Variable Selection and Description

After running the data in Python, people will come up with different kinds of results, all of which are useful tools to help us analyze the risk and return of stocks, but there are a variety of results that shown in Python. Next, in this paper, people will introduce how to analyze the result, it relates to several principles for judging the risk and return of investment products, as Table 1 shows.

3 Results and Discussion

3.1 Linear Regression Results

Looking at Fig. 1 and Fig. 2, people can see that the cumulative rates for both training performance and test performance show an overall upward trend, which means that both training performance (Fig. 2) and test performance (Fig. 1) are profitable. This is a deep reflection of the effectiveness of the momentum strategy. However, from the details, people can find that both training performance and test performance show downward trend at some time points, which mean there is still some volatility in the overall upward trend, and in comparison, the volatility and downward trend of test performance is greater than that of training performance. This is basically cope with the conclusion we draw from the above table. By carefully exploring the time nodes that are in a downward trend, people will find that the reasons that lead to the downward trend of stocks are often closely related to the general international economic environment. In training performance, from 2008 to 2010, the volatility of stocks is extremely strong, and after 2010, the running trend area of stocks is smooth and shows a steady rise, combined with the world economic context, the global financial crisis started to spread in 2008, and the stock market gradually recovered from the crisis in 2010. In test performance, 2012 shows a downward trend, which is most likely related to the European sovereign debt crisis in 2012. 2019 to 2020, test performance shows an extremely exaggerated downward trend with extremely strong volatility, which is highly correlated with the spread of Covid-19 around the world. The external events of linear regression model are concentrated in the formula of linear human regression as δ, unpredictable noise, these external events cannot be expected, but are closely related to the risk and return of the stock, which also reflects the risk of momentum strategy.

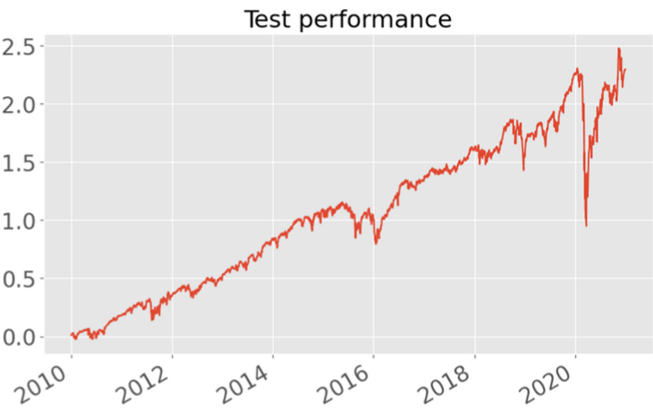

Fig. 1. The test performance model.

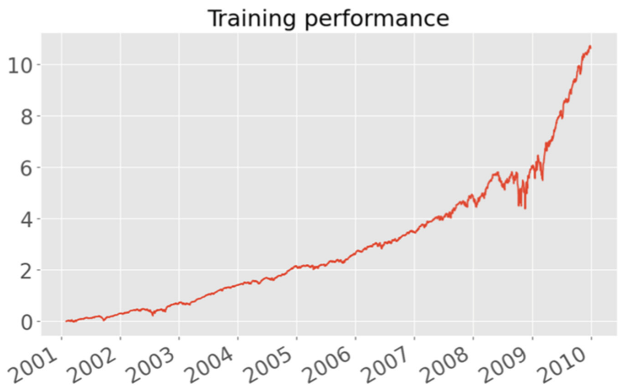

Fig. 2. The training performance model.

After analyzing the test performance and training performance model, let's take transaction cost into consideration to see how the transaction cost affect the momentum strategy. For the training performance model. At first, people put 0.1 point of the transaction cost into the training performance model (Fig. 3). They can find that the transaction cost has negligible influence on the whole model. However, when the transaction cost tends to point 0.5% (Fig. 4) and point 1% (Fig. 5), people gradually find the transaction cost were affecting the model in a potent way. As the point came to 0.5%, the whole model was still showed an upward trend, but the trend is quit flat comparing to the basic model that without transaction cost. When the transaction cost up to 1% point, the upward trend become flatter.

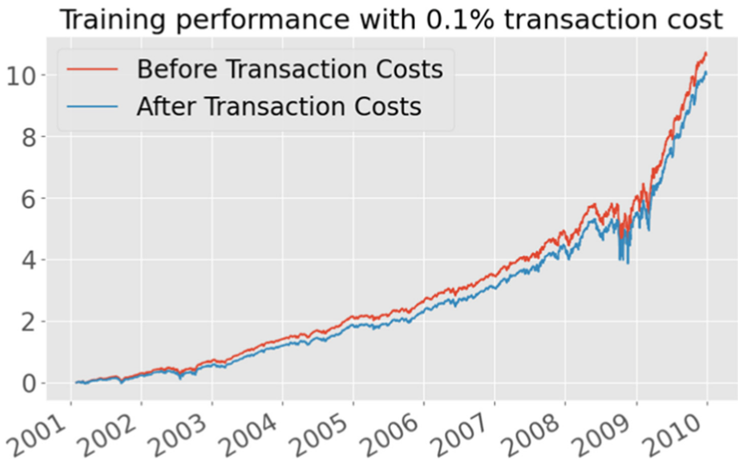

Fig. 3. The training performance model with 0.1% transaction cost.

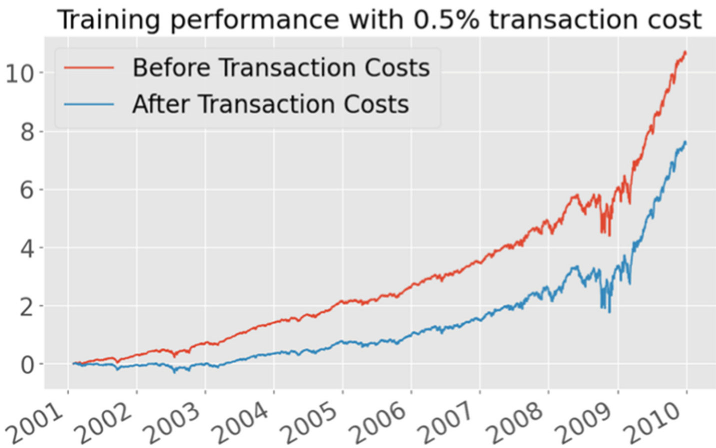

Fig. 4. The training performance model with 0.5% transaction cost.

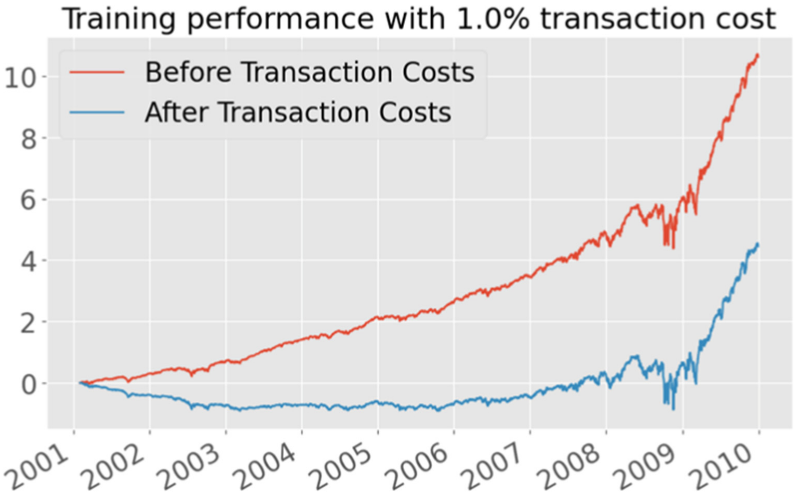

Fig. 5. The training performance model with 1.0% transaction cost.

For the test performance, when people put 0.1 point of the transaction cost into the test performance model (Fig. 6), they can find the image still showed an upward trend, but the trend signal is not as strong as the basic model that without transaction cost. However, when the transaction cost up to point 0.5% (Fig. 7) and 1% (Fig. 8), the image shows a downward trend which means if the transaction cost up to the level of 0.5% and 1%, the stocks are keep losing money and the losses it brings is tremendous. Through this these images, people can conclude the truth that the transaction cost plays a significant role in the stock market. It can affect the profitability and lead to huge losses.

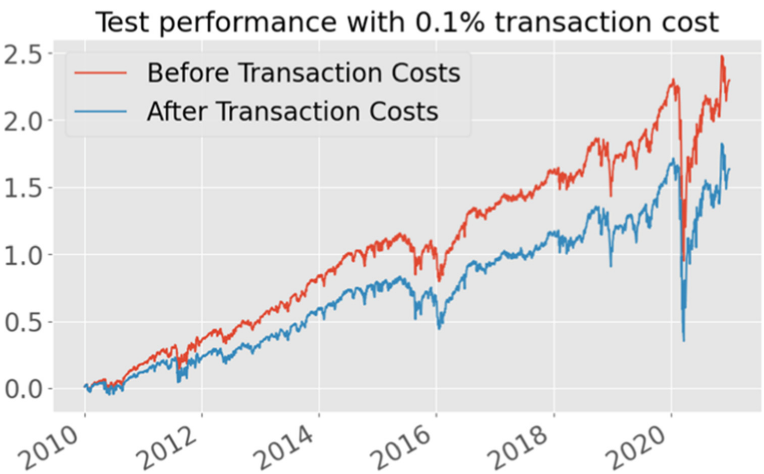

Fig. 6. The training performance model with 0.1% transaction cost.

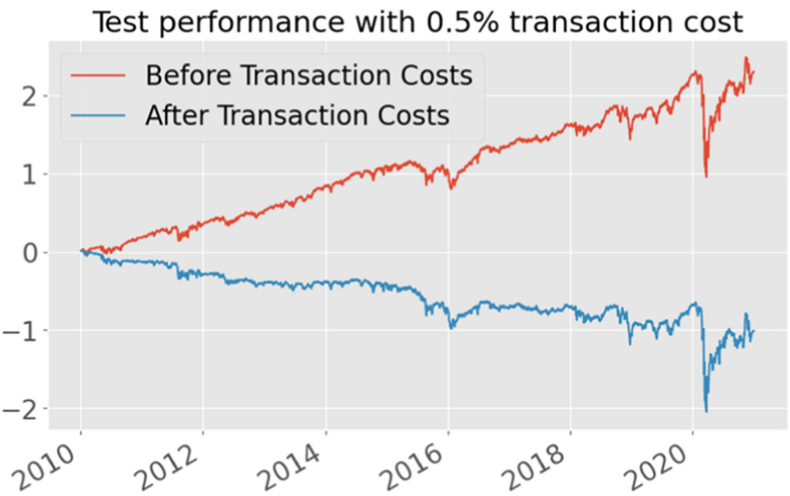

Fig. 7. The training performance model with 0.5% transaction cost.

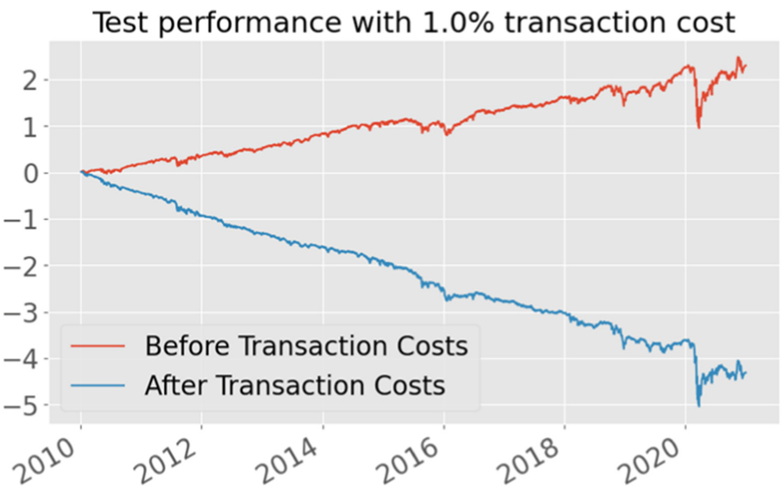

Fig. 8. The training performance model with 1.0% transaction cost.

3.2 Evaluating Results

Through the graph, people can find that the transaction cost are made a bad influence on every parameter and the overall performance of the training dataset is more outstanding than that of the testing dataset. With the aim of making a better comparison between the training dataset and testing dataset, this essay will emphasize in analyzing the line that without transaction cost. In the Sharpe ratio, the Sharpe ratio of training-dataset is 1.737 and the Sharpe ratio of testing-dataset is 0.730. As we mentioned earlier, the

Table 2. Evaluating parameters.

Parameter index	Value
Train sharp ratio	1.737
Train max drawdown	0.211
Train skewness	0.688
Train kurtosis	18.613
Test sharp ratio	0.730
Test max drawdown	0.409
Test skewness	−0.895
Test kurtosis	35.685

higher the Sharpe ratio, the higher the return while proving the lower the risk. In the max-drawdown, the max-drawdown of the training-dataset is 0.211 and the max-drawdown of the testing dataset is 0.409, as Table 2 shows.

By comparing, people can find that the max-drawdown of the training dataset is smaller than that of the testing dataset, which indicates that the overall performance of the training dataset is more stable and the volatility is not particularly strong, which also represents that the loss of the stock is controllable. In skewness, the value of skewness of training dataset is 0.688 and the skewness of testing dataset is - −0.895, which shows that the absolute value of skewness of training dataset is smaller than that of The absolute value of testing dataset is smaller than the absolute value of testing dataset, which means that the symmetry of training dataset is better than testing dataset, (in normal distribution skewness = 0), in kurtosis, the peak of training dataset is 18.612 and the peak value of testing dataset is 35.684, through comparison one can see that the peak value of training dataset is lower than that of testing dataset, in kurtosis evaluation, the larger the kurtosis coefficient represents the more extreme values in the data series and its stability is less. In summary, by comparing the data of training dataset and testing dataset one can conclude that training dataset is more profitable and stable compared to testing dataset.

4 Conclusion

Based on the analysis, people can summarize the momentum strategy as a whole. First of all, it is very good in terms of operability. And its principle is very easy to understand, buying when the price rises and selling when the price falls. Secondly, when the model shows a relatively strong uptrend, the momentum strategy will give investors very significant returns. However, the drawbacks of the momentum strategy are also obvious. First, because the principle of the momentum strategy is extremely simple and well known, the number of people using it is very large, which means that the competition among the investment community is more intense, and the income that can be obtained will be significantly reduced due to the intensity of competition. Thirdly, in the momentum strategy, traders need to closely follow the trend of the stock and buy and sell frequently, which can lead to very large transaction costs, which can lead to huge losses.

In the practice of this momentum strategy, there are still some areas that can be optimized. In terms of stock selection, the number of stocks selected in this strategy is only 100, which is limited, and it is difficult to reflect the impact of the momentum strategy globally, and the selected stocks are mostly high quality stocks with outstanding performance in the past, so the results obtained have certain limitations. In terms of time horizon, the strategy is selected over a 20-year time horizon, which is too long to capture the dynamic changes of the upward and downward movements of the momentum strategy in a refined manner. In the comparison of stock results, this strategy adopts Sharpe ratio, max-drawdown, kurtosis, skewness as the measures, in the future, one can use more measures such as SORTINO RATIO, TREYNOR RATIO, INFORMATION RATIO, to analyze our risk and return more comprehensively. This will give us a more comprehensive and objective analysis of the momentum effect.

References

1. Liu, S. M.: The market calls for quantitative financial brands. China Finance, (2020).
2. Yang, X. H.: Research on investment strategy of FOF funds based on momentum effect and machine learning scoring model. Anhui University of Finance and Economics, (2022).
3. Liao, F. H.: Risk and return analysis of securities investment portfolio. Modern Business Industry, (2019).
4. Zhou, J. H., et al.: Research on stock selection strategy based on machine learning multi-factor quantitative model. Science and Technology Innovation, (2022).
5. Sun, X. Q., et al.: Decoding Chinese style quantitative investment. Tsinghua Financial Review, (2016).
6. Jia, L. F.: Research on quantitative trading strategy based on multivariate hidden Markov model. Southwest University of Finance and Economics, (2020).
7. Zhang, X. T., Chen, L.: Design and implementation of a linear regression-based popular stock analysis and recommendation system. Modern Information Technology, (2022).
8. Yuan, Z. B.: Market model R~2: A measure of private information arbitrage or a measure of market efficiency. Southwest University of Finance and Economics, (2009).
9. Zhang, B.: Evaluation analysis of the macroeconomic impact of the new crown pneumonia epidemic and the hedging effect of fiscal policy. Industrial Technology Economics, (2020).