# Performance of Time-series Momentum Strategy: US Evidence

Siyao Duan<sup>1,a,\*</sup>

<sup>1</sup>University of Glasgow, Glasgow G12 8QQ, UK a. 2803501d@student.gla.ac.uk \*corresponding author

Abstract: This paper examines the effectiveness of the time series momentum strategy in generating positive returns in the US stock market, with a focus on exploring its dynamics and performance using different moving average methods. The author conducted an empirical analysis of the time series momentum strategy using S&P500 data from 2000 to 2022. A regression model was applied to estimate the expected returns and volatility of each as-set, and then an evaluation of momentum trading strategy based on different moving average methods was developed. The author evaluates the performance of the strategy with and without transaction costs. The study contributes to the literature by providing empirical evidence on the effectiveness of the time series momentum strategy in the US stock market and by exploring the performance of different moving average methods on the strategy. The findings of this study can provide insights for investors and portfolio managers interested in implementing momentum strategies in their investment portfolios.

**Keywords:** momentum strategy, time-series momentum, quantitative portfolio

## 1. Introduction

Momentum investing is a popular quantitative investment strategy that has gained significant attention in the finance literature over the past few decades. Time series momentum strategy(TSMOS), a subcategory of momentum investing, uses past returns to predict future returns. In this paper, we conduct an empirical analysis of the time series momentum strategy using S&P500 data from 2000 to 2022.

Time series momentum strategy has gained significant attention in the past two decades in the finance literature. It is a quantitative investment strategy that uses past returns to predict future returns. In this paper, we conduct an empirical analysis of the time series momentum strategy using S&P500 data from 2000 to 2022. One of the most influential works in the field of momentum investing is Jegadeesh and Titman. The authors showed that past performance can predict future performance, especially for short-term investments [1]. Building on their work, Moskowitz et al. demonstrated that momentum strategies can generate positive returns in different asset classes, including equities, bonds, currencies, and commodities [2].

Further research has examined the effectiveness of different momentum strategies. Carhart proposed a four-factor model that includes momentum as one of the factors to explain the cross-section of stock returns [3]. Fama and French suggested that the momentum effect is not a distinct factor but rather a manifestation of other fac-tors such as market beta, size, and value [4].

<sup>© 2023</sup> The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

Recent literature has explored more applications of time series momentum strategies. Hutchinson and O'Brien delve into the relationship between time-series momentum and macroeconomic risk. Their study examines how macroeconomic factors impact the performance of momentum strategies and the findings contribute to understanding the underlying dynamics and risk factors associated with momentum investing [5]. Interest rate momentum is explored by Hartley across global yield curves. The study investigates the presence of momentum effects in interest rates in both developed and emerging market countries. The results highlight the pervasiveness of interest rate momentum and its potential implications for fixed income investors [6]. Intraday time-series momentum and investor trading behavior are explored by Onishchenko, Zhao, Kuruppuarachchi, and Roberts, they revealed how institutional and foreign investors' late-informed trading contributes to intraday momentum effects [7]. Examining the performance of time-series momentum strategies across different asset classes, Molyboga, Swedroe, and Qian analyzed data from 78 futures markets, their study highlights the profitability and practical applications of short-term trend-following strategies using daily returns [8]. The study by Molyboga, Qian, and He explores the practical applications of combining carry and time-series momentum strategies, their findings suggest that the combination of carry and time-series momentum can lead to improved risk-adjusted returns and enhance portfolio diversification [9]. Levy and Lopes introduce the concept of dynamic momentum learning and its application in trendfollowing strategies. The benefits of incorporating dynamic econometric models to adaptively adjust the importance of different look-back periods for individual assets are highlighted [10].

The previous studies, along with others not mentioned here, have significantly contributed to our understanding of time-series momentum in various financial markets. The literature on momentum investing provides valuable insights into the profitability, risk factors, and practical applications of momentum strategies. As momentum remains one of the widely accepted investment approaches, further research in this area will continue to enhance our understanding of its dynamics and implications for investors.

The motivation behind this study is to examine the effectiveness of the time series momentum strategy in generating positive returns in the US stock market, as well as to explore its dynamics and performance using different moving average methods. Specifically, this paper aims to investigate the relationship between time series momentum strategy and idiosyncratic volatility and to evaluate the performance of different moving average methods, such as simple moving average (MA) and exponential moving average (EMA), on the strategy. To achieve these goals, a regression model to estimate the expected returns and volatility of each asset was applied, and then an evaluation of momentum trading strategy based on different moving average was developed. The strategies were evaluated with and without transaction costs.

This paper contributes to the literature by providing empirical evidence on the effectiveness of time series momentum strategy in the US stock market and by exploring the performance of different moving average methods on the strategy. The findings of this study can provide insights for investors and portfolio managers who are interested in implementing momentum strategies in their investment portfolios.

# 2. Methodology

# 2.1. Data Description

The data set used in this study was from the S&P500 index, which includes daily closing prices of its constituent stocks. The S&P500 data set is representative of the US stock market and is widely used in academic and industry research, whose content from 2000 to 2022 provides a robust and comprehensive source of data for analyzing the performance of the time series momentum strategy.

The data is organized into time series, with each observation corresponding to a trading day from January 1, 2000, to December 31, 2022.

The data cleaning and preprocessing procedures were conducted to ensure data quality, including identifying and correcting missing data, handling outliers, and checking for data consistency.

# 2.2. The Regression Model

Suppose there is a time series of returns  $R_t$ , which will be predict using a linear model with past returns at various horizons, or computed with different smoothing methods:

$$R_{t+1}^{e} = \alpha + \sum_{i=1}^{M} \beta_{i} f_{t}^{i} + \varepsilon_{t+1}$$
 (1)

Here,  $\epsilon_{t+1}$  is noise and each of the M factors  $f_t^i$  can take the form:

$$MA_t^N = \frac{1}{N} \sum_{s=t-N}^t R_s^e$$
 (2)

Or

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

where N is the lookback period,  $R_{t+1}^e$  is the expected return of the portfolio at time t+1,  $\alpha$  is the intercept or constant term in the linear regression model, which represents the expected return of the portfolio when all the factors have a value of zero. M is the number of factors included in the linear regression model;  $\beta_i$  is the coefficient or weights assigned to each of the M factors, which represents the contribution of each factor to the expected return of the portfolio, where the coefficients are estimated using historical data and are optimized to maximize the predictive power of the linear regression model;  $f_t^i$  is the values of the M factors at time t, which can be represented by the second equation or third equation;  $\varepsilon_{t+1}$  is the error term or noise in the linear regression model, represents the part of the expected return that is not explained by the factors included in the model.  $MA_t^N$  is the moving average factor that used in the Momentum Strategy, which calculates the average return of the portfolio over the past N periods.  $EMA_t^N$  is the exponential moving average factor that is used in the Momentum Strategy, which calculates the average return of the portfolio over the past N periods with a decay factor  $\varepsilon$  applied to give more weight to recent returns.

The parameters of the model, including the variance  $\sigma_{\epsilon}^2$  of the noise) can be estimated using linear regression (or more sophisticated technique like generalized least squares). Once that is done, the model can be used to make predictions about the next period returns. N = 24 was set for this model. The training set is set before January 1, 2010, and the test set afterwards.

#### 2.3. Portfolio Evaluation

After fitting the regression model, a momentum trading strategy that generates portfolio weights based on the simple moving average (MA) or exponential moving average (EMA) of asset returns was developed. The MA and EMA of each stock's returns using different window lengths or half-lives were calculated firstly. Based on these averages, the regression coefficients ( $\alpha$  and  $\beta$ ) for each stock's returns as a function of its MA or EMA, and the portfolio's expected volatility ( $\sigma$ ) were estimated. The portfolio weights were then calculated as:

$$portfolio\ weights = \frac{(moving\ average*\beta+\alpha)}{(\sigma^2)} \tag{4}$$

The returns for each portfolio using the weights computed were worked out and then evaluated for different transaction cost coefficients. The performances of equal-weighted portfolios were also calculated for comparison.

## 3. Results and Discussion

## 3.1. Model Without Transaction Cost

The numerical results of this model are shown in Table 1. The result indicates a relatively high Sharpe ratio on the training dataset and positive on the testing dataset, which suggests that the strategy is capable of generating significant returns with low risk. However, the maximum drawdown on the test data set is much lower than on the training data, which suggests that the strategy may be overfitting to the training dataset and may not generalise well to unseen data. The skewness and kurtosis values in the result indicate positive for both datasets, means that the returns distribution is positively skewed and leptokurtic. From the below four parameters, the results suggest that the time series momentum trading strategy performs well with the current model, generating significant returns with low risk.

	Training Set	Test Set
Sharpe Ratio	2.403	0.439
Max Drawdown	14.403	3.304
Skewness	-0.355	-0.355
Kurtosis	23.544	36.877

Table 1: Numerical result of the model without transaction cost.

The percentage change of the investment value on both the training and test data set were computed and plotted, which are shown in Figure 1. Based on the trends observed in the two plots, it can be said that the model performs well on both the training and test data sets as the cumulative returns remained stable throughout most of the period. However, there were a few special periods, where the model shows a downward trend, such as during the global financial crisis in 2009, the European sovereign debt crisis in 2012, the US presidential election in 2016, and the COVID-19 pandemic in 2021. These periods of high market volatility had a significant impact on the model's performance. In contrast, during the recovery period of the US stock market in 2008 and 2010, where there were strong upward trends, the momentum strategy performs better as it was designed to capture such trends. Therefore, it can be concluded that the stronger the trend, the better the model's performance, as figure 1 shows.

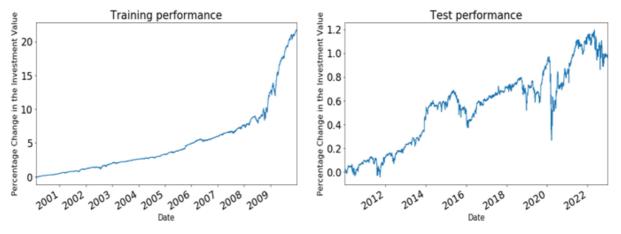


Figure 1: The percentage change of investment value on both training and test data set.

However, it should be noted that the model's performance is better in the training data set than in the test data set. This difference in performance could be attributed to the competitive market environment as this model was more widely known and utilized by investors and the US stock market's relatively low volatility and gradual upward trend, making it harder to capture strong trends. As a result, the returns generated by the model gradually decreased over time.

#### 3.2. Model with Transaction Cost

In this model, three different transaction cost rates were considered, 0.001, 0.005, and 0.01. The Results are shown in Table 2 and 3, and Figure 2, 3 and 4.

Transaction Cost Rate	Sharpe Ratio	Max Drawdown	Skewness	Kurtosis
0.01	1.141	2049.468	1.702	22.856
0.005	1.770	0.768	1.695	23.327
0.001	2.276	23.810	1.696	23.551

Table 2: Numerical result of the model with transaction cost on training data.

Table 3: Numerical result of the model with transaction cost on test data.

Transaction Cost Rate	Sharpe Ratio	Max Drawdown	Skewness	Kurtosis
0.01	-0.250	7511.158	-54.090	3043.681
0.005	0.176	306.724	50.238	2792.062
0.001	0.254	4.392	-0.113	48.099

We can observe that the model performs well on both datasets in the absence of transaction costs in 3.1, however, the performance is impacted as transaction costs increased. On the training set, considering the relatively low maximum drawdown as well as the positive Sharpe Ratio at all transaction cost rates, it shows the model can generates excess returns with relatively low risk. Moreover, the Sharpe Ratio increases as the transaction cost rate decreases, indicating better performance with lower transaction costs. The skewness is shown relatively high and stable at all transaction cost rates, indicating a right-skewed distribution of returns. The high kurtosis for all transaction cost rates also indicates a sharper distribution than the normal distribution, which however suggests that opportunities to generate large excess returns in this time period were relatively low in the market. On the test set, the impact of transaction cost rate was indicated more obviously. The negative Sharpe Ratio at the 0.01 transaction cost rate shows that it is too high to generate positive returns straightforward. The maximum drawdown also turned to over 7000 when the rate was 0.01 while it was relatively low with lower rates. The low skewness on the 0.01 transaction cost rate indicates a relatively balanced distribution of returns, however the large positive and negative value of skewness on the other two transaction cost rates indicates very right and left skewed distributions of returns. The positively skewed and peaked return distribution suggests that the model is probably taking advantage of trends in the market and profiting from positive momentum. However, considering with the high kurtosis, it indicates again there are not many opportunities to generate large excess returns in the market with high transaction costs.

By checking the graph results - the change of investment value, it is not hard to see that when transaction cost rate is at 0.001, there is little difference in returns before and after costs are added. However, when transaction cost rate increases to 0.005 to 0.01, the difference in returns becomes significant. This indicates that the original model that with-out transaction costs is too idealistic, and transaction costs play a critical role in determining profitability. Furthermore, when the test data was applied to this model, the cumulative returns steadily de-creased. This result is likely because the

portfolios constructed by this model were not profitable in the market environment over the last decade, even without transaction costs. Therefore, it highlights again that transaction costs cannot be ignored as they have a significant impact on profitability in this model, and if profits are not high enough to cover the transaction costs, significant losses could occur, as figure 2, 3 and 4 show.

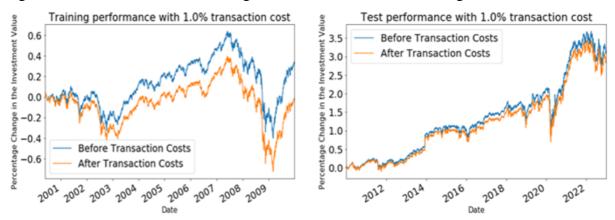


Figure 2: The percentage change of investment value when transaction cost rate = 0.01.

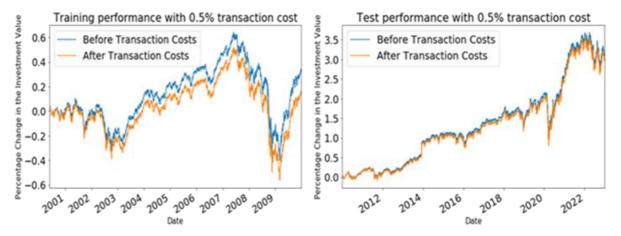


Figure 3: The percentage change of investment value when transaction cost rate = 0.005.

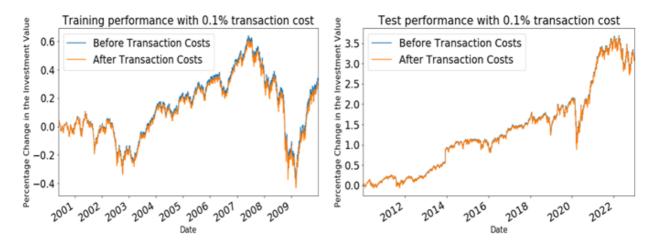


Figure 4: The percentage change of investment value when transaction cost rate = 0.001.

In summary, the model performs well on the training data, generating excess returns with relatively low risk at all transaction cost rates. However, on the test data, varying transaction costs can impact its performance significantly. The better performance at lower transaction cost rates indicates the importance of considering transaction costs in investment strategies. The skewness and kurtosis of the return distribution also vary in different market conditions, therefore, the market conditions, transaction costs and return risks should be considered when implementing the model in actual investments.

## 3.3. Portfolio Evaluation

The results of portfolio evaluation are shown in Table 4, 5, 6 and 7, where MA and EMA mean simple moving average and exponential moving average, and the followed numbers mean the length of halved life. SR, MD and K mean Sharp Ratio, Maximum Drawdown and Kurtosis respectively; (tr) means on training data and (te) means on test data. The strategies are ranked by the Sharpe Ratio computed on test data.

Strategy	SR(tr)	MD(tr)	K(tr)	SR(te)	MD(te)	K(te)
MA 1	3.148	0.482	151.827	0.794	5.103	29.066
MA 5	2.469	5.793	29.832	0.654	0.926	32.815
Equal Weight	0.244	3.978	10.197	0.652	4.027	35.601
EMA 1	3.048	0.22	50.37	0.602	4.781	30.657
EMA 2	2.924	0.199	28.226	0.57	8.781	31.479
EMA 5	2.763	0.214	26.781	0.567	1.06	38.038
EMA 40	2.111	0.33	24.009	0.551	2.871	45.3
EMA 20	2.409	0.387	25.057	0.549	4.977	41.934
MA 100	2.328	0.28	20.138	0.533	2.739	48.184
EMA 10	2.65	0.27	25.605	0.533	38.41	40.058
EMA 60	1.957	0.258	23.587	0.532	6.018	48.521
MA 10	2.609	3.655	27.508	0.518	1.234	34.187
EMA 80	1.864	0.264	23.507	0.516	24.644	51.469
EMA_100	1.804	0.284	23.476	0.511	2.155	54.303
MA_80	2.056	0.289	22.183	0.509	1.581	42.745
MA 60	1.984	0.283	23.16	0.507	1.24	42.72
MA_20	2.515	13.612	22.706	0.487	3.467	36.5
MA 40	2.034	3.435	21.845	0.389	4.455	41.673
MA 2	2,882	0.65	37.182	0.222	3.546	26.69

Table 4: Portfolio evaluation – ranking of strategies when transaction cost rate = 0.

Based on the results, it indicates again that the transaction cost rate has a significant impact on the performance of the trading strategies. In general, the higher the transaction cost rate, the lower the performance of the strategies. This makes sense since transaction costs can eat into the profits of the strategies. For example, when the transaction cost rate is zero, the "MA 1" and "MA 5" strategy has the highest Sharpe ratios in both the training and testing sets, however, when the transaction cost rate is increased to 0.001, the performance of both these strategies deteriorates, and they are outperformed by several other strategies, such as the "Equal Weight" strategy and the "MA 100" strategy, as Table 4, 5, 6 and 7 show.

It seems that the strategies that perform well with low transaction cost rates are also the ones that have relatively high turnover rates, such as the "MA 1" and "MA 5" strategies. On the other hand, the strategies that perform well with high transaction cost rates tend to have lower turnover rates and are more diversified, such as the "Equal Weight" strategy and the "MA 100" strategy.

It is also interesting to note that the performance of some strategies, such as the "MA 2" strategy, deteriorates significantly when the transaction cost rate is increased. This may be because these strategies rely heavily on frequent trading, and the transaction costs can add up quickly and erode their profitability.

Table 5: Portfolio evaluation – ranking of strategies when transaction cost rate = 0.001.

Strategy	SR(tr)	MD(tr)	K(tr)	SR(te)	MD(te)	K(te)
Equal Weight	0.232	4.246	10.563	0.647	4.046	35.660
MA 100	2.247	0.301	20.200	0.440	3.021	50.304
EMA 40	2.041	0.352	23.982	0.436	3.379	51.474
EMA 60	1.898	0.274	23.603	0.435	7.166	53.760
EMA 100	1.756	0.291	23.518	0.430	2.531	58.979
EMA 80	1.811	0.271	23.540	0.428	44.061	56.331
MA 2	2.631	0.990	39.339	0.423	27.602	616.584
MA 80	1.978	0.312	22.221	0.402	1.898	48.771
EMA 20	2.316	0.414	24.874	0.394	7.060	51.551
MA 60	1.911	0.301	23.276	0.391	1.758	50.607
EMA 10	2.531	0.301	25.136	0.318	130.217	54.768
MA 5	2.285	15.251	28.460	0.304	2.260	60.183
MA 20	2.382	15.710	22.671	0.288	4.723	48.439
EMA 5	2.611	0.261	25.941	0.280	2.376	60.687
MA 1	2.801	1.725	180.229	0.259	14.976	3211.455
MA 40	1.957	3.523	21.865	0.250	6.317	50.289
MA 10	2.484	2.216	27.582	0.237	1.909	50.498
EMA 1	2.782	0.414	55.704	0.198	30.236	174.621
EMA 2	2.710	0.361	27.704	0.183	2.134	68.944

Overall, it is important to take transaction costs into account when evaluating the performance of trading strategies, as they can have a significant impact on the profitability of the strategies. The best strategy will depend on the specific trading scenario, including the expected transaction cost rate, the expected turnover rate, and the expected market conditions.

Table 6: Portfolio evaluation – ranking of strategies when transaction cost rate = 0.005.

Strategy	SR(tr)	MD(tr)	K(tr)	SR(te)	MD(te)	K(te)
Equal Weight	0.189	5.673	12.488	0.626	4.122	35.909
EMA 2	1.672	18.005	32.927	0.323	0.000	2370.539
MA 1	0.520	0.476	1126.713	0.304	0.000	3218.314
EMA 1	1.395	71.981	93.986	0.278	0.000	3273.889
EMA 40	1.765	0.448	23.759	0.222	13.940	482.907
EMA 10	2.042	0.797	22.547	0.194	0.000	447.043
EMA 20	1.947	0.528	23.808	0.185	123.087	279.989
MA 80	1.672	0.444	22.327	0.178	19.518	324.668
EMA 100	1.568	0.320	23.665	0.174	5.295	108.344
EMA 60	1.666	0.356	23.613	0.168	18.539	147.249
EMA 80	1.605	0.334	23.643	0.165	5.695	117.401
MA 100	1.926	0.577	20.423	0.140	9.235	76.219
MA 2	1.291	0.837	79.597	0.122	0.000	1176.659
MA 40	1.654	3.904	21.979	-0.203	98.300	1445.151
MA 10	1.971	16.321	28.656	-0.225	0.000	696.500
MA 20	1.854	27.485	22.270	-0.257	160.969	3215.537
MA 5	1.457	193.542	18.244	-0.258	0.000	458.657
MA 60	1.627	0.397	23.733	-0.308	23.034	3078.939
EMA 5	1.960	2.271	21.298	-0.564	0.000	1154.586

Table 7: Portfolio evaluation – ranking of strategies when transaction cost rate = 0.01.

Strategy	SR(tr)	MD(tr)	K(tr)	SR(te)	MD(te)	K(te)
Equal Weight	0.155	9.026	16.972	0.601	4.219	36.248
MA 60	1.289	0.631	24.294	0.427	571.921	821.039
EMA 2	0.432	355.531	1200.537	0.403	0.000	2271.514
EMA 10	1.377	21.645	20.201	0.354	0.000	1419.524
MA 100	1.542	2.770	20.644	0.342	175.640	694.735
MA 2	0.250	4.182	987.252	0.280	0.000	3273.526
MA 1	0.067	5.125	392.233	0.265	0.000	1230.082
MA 10	1.276	2.823	40.205	0.160	0.000	2131.793
EMA 100	1.341	0.718	23.805	0.146	320.605	712.969
MA 80	1.310	1.271	22.357	0.036	281.725	1009.105
EMA 80	1.358	0.793	23.691	-0.062	0.000	370.245
EMA 5	0.820	76.515	387.912	-0.153	0.000	2768.759
MA 20	1.199	185.184	22.220	-0.179	15830.950	527.220
MA 5	-0.241	4.603	2229.054	-0.187	0.000	2194.113
MA 40	1.290	4.454	22.151	-0.217	8030.596	3184.757
EMA 60	1.386	0.916	23.470	-0.307	11953.719	3240.147
EMA 40	1.430	1.211	23.146	-0.377	2377.669	2106.976
EMA 20	1.484	3.354	21.715	-0.386	0.000	1163.415
EMA 1	-0.151	1137.854	1151.489	-0.454	0.000	1167.909

#### 4. Results and Discussion

The results of the study suggest that the time series momentum strategy can be an effective investment strategy for generating positive returns with low risk in the US stock market. The regression model developed in this study shows a high Sharpe ratio on the training dataset and a moderate Sharpe ratio on the testing dataset, indicating that the strategy has the potential to outperform the market. The results also suggest that the momentum effect is present in the US stock market, as the time series momentum strategy generates significant excess returns after controlling for risk factors. The strategy's performance is consistent with previous studies that have shown the effectiveness of momentum strategies in different asset classes and regions. The importance of considering transaction cost and market environment when implementing this model is also emphasized by the results. However, the results also suggest that the strategy was probably overfitting to the training data in this model, as indicated by the high maximum drawdown on the training dataset. Therefore, future studies could explore methods to reduce overfitting, such as using more robust machine learning algorithms or incorporating additional factors into the regression model.

In conclusion, this study provides empirical evidence of the effective-ness of the time series momentum strategy in the US stock market using S&P500 data from 2000 to 2022. The strategy generates positive excess returns with low risk in particular situations, indicating its potential as a profitable investment strategy. The regression model developed in this study provides a framework for implementing the strategy, which involves calculating simple or exponential moving averages of asset returns and estimating the regression coefficients and portfolio weights based on these averages. The strategy can be further optimized by adjusting the window length or half-life of the moving averages and considering transaction costs.

# References

- [1] Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. The Journal of Finance, 48(1), 65-91.
- [2] Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H. (2012). Time series momentum. Journal of Financial Economics, 104(2), 228-250.
- [3] Carhart, M. M. (1997). On persistence in mutual fund performance. The Journal of Finance, 52(1), 57-82.
- [4] Fama, E. F., & French, K. R. (2012). Size, value, and momentum in international stock returns. Journal of Financial Economics, 105(3), 457-472.
- [5] Hutchinson, M. C., & O'Brien, J. (2020). Time series momentum and macroeconomic risk. International Review of Financial Analysis, 69, 101469. https://doi.org/10.1016/j.irfa.2020.101469
- [6] Hartley, J. (2020). Interest rate momentum everywhere across global yield curves. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3542813
- [7] Molyboga, M., Swedroe, L., & Qian, J. (2020). Short-Term Trend: A Jewel Hidden in Daily Returns. The Journal of Portfolio Management, 46(3), 80-93. https://doi.org/10.3905/jpm.2020.1.186
- [8] Onishchenko, O., Zhao, J., Kuruppuarachchi, D., & Roberts, H. (2021). Intraday time-series momentum and investor trading behavior. Journal of Behavioral and Experimental Finance, 31, 100557. https://doi.org/10.1016/j.jbef.2021.100557
- [9] Molyboga, M., Qian, J., & He, C. (2021). Practical applications of carry and time-series momentum: A match made in heaven. The Journal of Alternative Investments. https://doi.org/10.4324/9781315100944-5
- [10] Levy, B. P. C., & Lopes, H. F. (2021). Trend-following strategies via dynamic momentum learning. https://doi.org/10.48550/arXiv.2106.08420