Evaluating Technical Trading Strategies in US Stocks: Insights From Data-Snooping Test

Yifei Wang
Hong Kong Baptist University

Yidong Chen
Tsinghua University

Hongyu Tian
Beijing Normal University
Hong Kong Baptist University
United International College

Zane Wayne
Independent

Our comprehensive evaluation examines the in-sample and out-of-sample effectiveness of technical trading strategies for Apple Computer Inc (AAPL), Microsoft Corp (MSFT), and Nvidia Corp (NVDA) over the period from January 2000 to December 2022. Utilizing advanced methods such as reality checks and stepwise tests, we address the potential data-snooping bias—a scenario where apparently profitable strategies might arise by chance rather than through genuine predictive accuracy. Despite rigorous analytical techniques, our findings indicate an inability to identify any technical trading strategies that yield consistent profits in both the in-sample and out-of-sample periods. Further analysis, with a specific cutoff established in February 2016 and incorporating corrections for data snooping, consistently demonstrates the unprofitability of these strategies. This highlights a significant challenge in financial markets: the intrinsic difficulty in identifying technical trading strategies that can consistently produce profitable outcomes over time. Our conclusions underscore the complexities inherent in technical analysis and the substantial obstacles in deriving actionable insights for stock market trading based on technical trading frameworks.

Keywords: technical trading strategies, data snooping bias, US stocks

INTRODUCTION

Technical analysis (TA) involves studying historical price data, often via charts or mathematical models, to generate trading signals without relying on asset fundamentals. Despite this, TA is widely used and has shown profitability in various markets. For instance, Kwon and Kish (2002) demonstrated significant returns on the NYSE index (1962–1996), while Ratner and Leal (1999) found profitable variable-length Moving Average rules in emerging markets, and Tam and Cuong (2018) corroborated the effectiveness of key indicators in Vietnam. Studies on equities also highlight TA's importance. Taylor (2014) identified momentum-based gains on the DJIA (1928–2012), contingent on short-covering, while Brock et al. (1992) reported superior returns from moving average and trading range break strategies on the Dow Jones (1897–1986), and Ratner and Leal (1999) noted success in Latin American and Asian markets.

Here, we examine real-world TA profitability from 2000 to 2022 using both in- and out-of-sample tests, reality checks, and stepwise analyses to address data-snooping concerns. We focus on AAPL, MSFT, and NVDA, applying up to 90 TA strategies and measuring average annualized returns and Sharpe ratios. We assess whether in-sample profitability carries over out-of-sample, with thorough robustness checks inspired by Neely et al. (2014).

Our findings indicate that TA strategies generally fail to produce significant profits in- or out-of-sample, even under varying performance measures and thresholds. This conclusion remains robust across diverse methodologies, implying that TA-based returns are often tied to parameter optimization rather than true market inefficiencies. Consequently, strategies appearing profitable ex post face substantial hurdles ex ante.

These insights highlight the need for caution when relying solely on TA. A diversified approach—combining TA with fundamentals, timing, and advanced techniques—may better enhance returns and manage risks. By addressing data-snooping and parameter-sensitivity issues, this research offers valuable guidance for investors and algorithmic traders aiming for consistent profitability. The rest of this paper discusses TA's theoretical underpinnings, data, strategy design, performance metrics, and detailed empirical results, culminating in concluding remarks on the intricate interplay of economic indicators and advanced quantitative methods in finance.

Technical Analysis

Quantitative investment integrates computer engineering, statistical mathematics, finance, and related fields to build models that mine large datasets for high-probability trading opportunities. By leveraging computational power and historical databases, it aims to produce measurable, repeatable, and predictable results while reducing emotional bias. Strategies typically fall into two categories: market timing and equity selection. Market timing uses algorithms to forecast market movements and trade accordingly, whereas equity selection focuses on evaluating stocks through criteria like financial metrics and market conditions.

A key component is stock timing: buying or holding when an uptrend is expected, selling or liquidating when a downtrend looms, and adjusting tactics when trends are uncertain. However, forecasting market changes is complex, requiring sophisticated quantitative tools to analyze macroeconomic and microeconomic factors. Technical analysis—one of the most common methods—relies on price and volume patterns, grounded in assumptions that market prices reflect all information, follow discernible trends, and often repeat historical behaviors.

Effective Market Hypothesis

The Efficient Market Hypothesis (EMH), introduced by Eugene Fama, posits that markets integrate all available information into security prices, limiting investors' ability to consistently earn above-average risk-adjusted returns. EMH is categorized into weak, semi-strong, and strong forms, each reflecting the extent to which historical, public, and private information is priced in. However, the still-developing equities market and the surge in quantitative investing suggest inefficiencies remain, creating opportunities for excess returns. Behavioral finance research underscores how biases like overconfidence, representativeness, anchoring, and loss aversion distort investor decisions and challenge pure market rationality. Recognizing these biases informs more accurate financial models and strategies, ultimately helping investors make better decisions and enhancing overall market stability.

DATA

We compiled daily transaction data (open, high, low, close, volume) for Apple, Microsoft, and Nvidia from January 1, 2000, to December 31, 2022. Using log returns, we assessed average daily returns and split the sample into in-sample (January 1, 2000–February 6, 2016) and out-of-sample (February 6, 2016–December 31, 2022) periods to evaluate investment performance.

Meanwhile, liquidity risk has emerged as a pivotal factor in asset pricing, underscored by severe market dislocations during the 2007–2009 crisis. Research increasingly integrates liquidity risk into pricing models, highlighting its role in systemic stability. Regulatory frameworks, including Basel III, also reflect this focus, emphasizing liquidity requirements to fortify financial markets against future shocks.

TECHNICAL TRADING STRATEGIES

TEMA Strategy

Moving average (MA) trading strategies identify trends and signal reversals. The simplest method uses price crossing an MA to trigger trades. Another approach compares two MAs: a short-term one crossing above or below a longer-term one indicates a reversal. MAs under 20 days are short-term, 20–60 days are medium, and over 60 days are long-term.

Weighted Moving Average (WMA) emphasizes recent data for more precise insights. Mathematically, WMA is a convolution of price and a weighting function, enhancing analysis and forecasts.

$$weights = range(1, n + 1) \tag{1}$$

$$\mathrm{WMA}_t(\mathrm{n}) = \frac{\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{weights_i} \times \mathrm{P_{t-i+1}}}{\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{weights_i}} \tag{2}$$

where n signifies the temporal parameter. The Exponential Moving Average (EMA), a type of Weighted Moving Average (WMA), assigns greater weight to the most recent price information. This weighting method ensures that the latest observations have a more significant impact on the average, rendering the EMA a highly responsive and adaptable tool for assessing financial market trends. By emphasizing recent price movements, the EMA provides a more fluid and immediate depiction of market conditions compared to traditional moving averages, which equally weigh all data points. This characteristic makes the EMA exceptionally useful for investors seeking to identify recent market changes and execute timely trading decisions.

$$EMA_t(P, n) = EMA_{t-1}(P, n) \times \left(1 - \frac{2}{1+n}\right) + P_t \times \frac{2}{1+n} \tag{3}$$

where n represents the chronological factor.

The Triple Exponential Moving Average (TEMA) was introduced in 1994 through Patrick G. Mulloy’s manuscript, titled ”Smoothing Data with Faster Moving Averages,” published in the journal Technical Analysis of Stocks Commodities. This technique seeks to mitigate the inherent lag associated with Moving Averages.

$$EMA_t^{(2)}(n) = EMA_t(EMA(P, n), n) \tag{4}$$

$$EMA_t^{(3)}(n) = EMA_t\left(EMA^{(2)}(n), n\right) \tag{5}$$

$$TEMA_t(n) = 3 \times EMA_t(P, n) - 3 \times EMA_t^{(2)}(n) + EMA_t^{(3)}(n) \tag{6}$$

Period parameter n is crucial for time series analysis, segmenting data into intervals that reveal trends, seasonality, and cyclical behaviors. This segmentation sharpens forecasting, model accuracy, and financial research, informing decisions and policy. We then apply the TEMA strategy, evaluating performance via backtesting. A grid search tests fast periods from 5 to 25 (increments of 2) and slow periods from 30 to 40 (increments of 2).

Zero-Lag Indicator Strategy

The filtration criterion is a core technical analysis tool, buying when price rises by a set percentage from its latest low and selling when it falls by a set percentage from its latest high. The Zero-Lag Indicator (ZLIndicator) refines the EMA by reducing the lag between true price and its correction, offering more timely, precise signals. Replacing the EMA with an Error Correcting (EC) filter and modulating a gain factor within set bounds minimizes error, boosting precision and adaptability. This synergy delivers more accurate forecasts and a robust, responsive model for financial analytics.

$$EC_t(n,l) = \frac{2}{n+1} \times \left(P_t + g_t \times (P_t - EC_{t-1})\right) + \left(1 - \frac{2}{n+1}\right) \times EC_{t-1} \tag{7}$$

$$g_t = argmin_{-l<g_t<l} \mid P_t - EC_t \tag{8}$$

Gain term (g), period (n), and gain limit ($l > 0$) are crucial parameters in this framework. G measures profit, n defines the time frame for assessing gains (e.g., quarterly, annually), and l caps the maximum gain to ensure realistic financial analysis. Together, they provide a structured approach for evaluating financial data. The Zero-Lag Indicator's intersections enable an automated crossover trading system using Exponential Moving Average (EMA) and Error-Correcting (EC) lines to signal trade entries and exits, enhancing trading precision by reducing delays. Additionally, comprehensive backtesting with grid search on period values (20–40) and gain thresholds (50–70) optimizes parameters, improving the strategy's reliability and predictive power across diverse market conditions.

RETURNS AND PERFORMANCE METRICS

Logarithmic Returns

Logarithmic returns effectively measure exponential growth in financial markets by capturing the continuous nature of asset price movements. Unlike traditional percentage changes over discrete periods, they account for compounding, providing a more accurate reflection of asset performance and growth patterns over time.

Calculating total compound return involves aggregating individual returns over a period, considering reinvested earnings and both positive and negative fluctuations. This comprehensive metric offers a precise view of long-term portfolio performance, aiding analysts and investors in evaluating investment strategies, asset allocation, and risk management. Compound return is essential for assessing investment sustainability and effectiveness under varying market conditions.

The arithmetic mean of logarithmic returns offers a more accurate representation of compounded returns than the simple arithmetic mean of raw returns. It mitigates volatility effects, aligns with continuous compounding, and supports advanced financial modeling and decision-making. Logarithmic returns also enhance risk management by enabling easier data analysis and helping develop resilient investment strategies in volatile markets.

Sharpe Ratio

The Sharpe Ratio is widely acknowledged as a traditional performance metric in the financial domain, evaluating the average excess return per unit of risk, where risk is quantified by the standard deviation of excess returns. A higher Sharpe Ratio indicates superior risk adjusted profitability of investments. In this study, we use the ex-post Sharpe Ratio (SR).

Max Drawdown

The term ”maximum drawdown” denotes the most significant proportional decline in a portfolio’s value from its peak to a subsequent trough. This metric gauges the downside risk inherent in an investment strategy, representing the largest potential loss an investor could have experienced when buying at the peak and selling at the nadir.

REALITY CHECK AND STEPWISE TEST

The importance of testing for data snooping is paramount, especially given the inherent flexibility in developing technical trading strategies, which allows for a wide range of parameters to be selected. This flexibility results in the potential consideration of numerous alternative hypotheses for statistical inference. Consequently, it is essential to determine whether these ostensibly profitable trading strategies genuinely provide predictive superiority over a benchmark model. However, this determination is challenging within traditional statistical frameworks that reject the null hypothesis based on the improbability of the observed data under the null. Evaluating various trading strategies increases the number of hypotheses tested, as weaker models or rules are routinely filtered out. This practice introduces the multiplicity problem: as the number of hypotheses tested increases, so does the probability of encountering a rare event, which in turn heightens the risk of falsely rejecting the null hypothesis, or committing a Type I error. Thus, the impressive performance observed in individual models through the rejection of separate null hypotheses might not truly reflect predictive superiority over a benchmark; rather, they could be the outcome of exhaustive specification searches. Particularly in our research, where up to 90 variants of technical trading strategies are typically examined, it is plausible for a skeptic to argue that discovering any well-performing strategies becomes almost inevitable given the extensive number of variants evaluated.

Scholars engaged in application-oriented research will recognize the dilemma of data mining—often referred to as data snooping—involving the overly specific tailoring of information. This issue has been exacerbated by the increased use of “big data,” and remains a persistent challenge in practical economics and finance. The prevalence of this problem is well-documented in foundational literature (e.g., Leamer (1978) and the references therein), sparking a considerable wave of innovative advancements. Such progress underscores the critical need to address the methodological pitfalls posed by data mining to ensure the integrity and reliability of econometric and financial analyses.

More rigorously, consider $fG = \{G_1, G_2, \cdots, G_K\}$ as $1 \times K$ matrices where the k_{th} element, G_k, represents the expected return or Sharpe ratio of the k_{th} strategy. Here, K denotes the total number of technical trading strategies evaluated in each experiment. Data snooping becomes problematic when an analyst identifies the top-performing element in the performance array G denoted as $G_j = \max(G)$. The analyst then conducts hypothesis testing under the null hypothesis that the specified strategy yields zero returns. This practice can inadvertently inflate the perceived performance of the strategy due to selection bias based on prior data. By isolating G_j and ignoring the wide array of examined strategies, the probability of a Type I error increases, leading analysts to erroneously reject the null hypothesis more frequently than warranted. Consequently, this can result in an overestimation of the effectiveness of the selected trading strategy, undermining the robustness of empirical findings in financial research.

$$H_0: G_j = 0. \tag{9}$$

"An evaluation of the null hypothesis, as stated in Equation (9), is considered a 'solitary examination'. The nominal t-statistic for this solitary examination is calculated as follows:"

$$t_{G_j} = \frac{G_j}{Std(G_j)\sqrt{n}}, \tag{10}$$

The nominal p-value, an essential element in statistical analysis, is derived by utilizing the cumulative distribution function (CDF) after determining $\mathrm{Std}(\mathrm{G_j})$, which represents the standard deviation of the variable $\mathrm{G_j}$. To obtain $\mathrm{Std}(\mathrm{G_j})$, one must analyze the relevant dataset, where n denotes the sample size that forms the basis of these statistical computations. The derived p-value quantifies the likelihood of encountering a result at least as extreme as the one observed, assuming the null hypothesis is correct. This calculation is based on the properties of the CDF, which converts the observed statistical value into a corresponding significance level. This conversion process is crucial for drawing robust inferential conclusions within the domain of financial econometrics and broader contexts. By integrating the standard deviation $\mathrm{Std}(\mathrm{G_j})$ with the sample size n, scholars can precisely assess the probability distribution of various outcomes, thereby ensuring the accuracy and reliability of the p-value in hypothesis testing frameworks. This methodological rigor underpins the reliability of conclusions drawn from statistical analyses in financial research.

As highlighted in White (2000), an individual appraisal of this nature fails to consider that G_j might represent the highest performance among K strategies when researchers aim to emphasize notable findings. Consequently, it is not dependent on the exact distribution of figures. Specifically, when K is exceedingly large, the nominal significance proposed by the analysis based on isolated testing might understate the true probability of a Type I error regarding the profitability of technical trading strategies, because the methodology under investigation has already been identified as the best. Therefore, a single assessment often tends to overly reject the null hypothesis due to data mining bias, thereby overstating the statistical significance of the profitability of technical trading.

To address this data-mining issue, White (2000) introduces a “reality check” evaluation, utilizing bootstrap techniques to construct the empirical distribution for G. This method examines a composite null hypothesis based on the combined distribution of each component of G.

$$H_0: \max_{k=1,\cdots,K} G_k \leq 0, \tag{11}$$

The variable denoted G_k represents the average yield or Sharpe ratio of the k_{th} strategy. To appropriately assess the specified composite null hypothesis, it is essential to adopt a multiple-testing framework that establishes suitable significance thresholds for the returns of various technical trading methodologies. Consistent with the bootstrap reality check technique proposed by by White (2000), we compute the reality check p-value, exemplified by utilizing daily time series data. This methodological approach ensures our evaluation of each strategy’s outcomes considers the multitude of tests, thereby providing a more robust and credible inference of their statistical significance.

The construction of the daily return matrix, denoted as Z, constitutes the initial step. Each element Z_{kt} captures the daily return of the kth strategy on the t indexed from 1 to K and t from 1 to T. Subsequently, the matrix Z undergoes the stationary bootstrap method, as formulated by Politis and Romano (1994), incorporating a predefined parameter set E. This bootstrap resampling procedure is replicated B times, resulting in resampled matrices designated as Z_b, with b spanning from 1 to B.

For each strategy k, performance metrics like the mean return or the Sharpe ratio, denoted by G_k, are computed using the original matrix Z. Correspondingly, for each resampled matrix Z_b, the metric $\mathrm{G_{kb}}$ is calculated.

Advancing in the process, let $\overline{\overline{V_1}}$ be defined as $T^{1/2}G_1$, and compute $\overline{\overline{V_{1b}^*}}$ as $T^{1/2}(G_{1b}-G_1)$. For indices $k > 1$, define $\overline{\overline{V_k}}$ as $\max\{T^{1/2}G_k, \overline{\overline{V_{k-1}}}\}$ and $\overline{\overline{\mathrm{V_{kb}^*}}}$ as $\max\{\mathrm{T}^{1/2}(\mathrm{G_{kb}}-\mathrm{G_k}), \overline{\overline{\mathrm{V_{k-1,b}^*}}}\}$. The next step involves sorting the series $\overline{\overline{\mathrm{V_{kb}^*}}}$ to obtain an ordered sequence $\overline{\overline{\mathrm{V_{k(1)}^*}}}, \dots, \overline{\overline{\mathrm{V_{k(B)}^*}}}$. Subsequently, determine the integer M such that $\overline{\overline{\mathrm{V_{k(M)}^*}}} \leq \overline{\overline{\mathrm{V_k}}} \leq \overline{\overline{\mathrm{V_{k(M+1)}^*}}}$. The ultimate bootstrap reality check p-value, denoted as $\mathrm{p_{rc}}$, is then calculated using the formula $\mathrm{p_{rc}} = 1 - \mathrm{M/B}$.

Hansen (2005) underscores the susceptibility of the reality check test to the inclusion of irrelevant and inconsequential options. The presence of these non-essential options can markedly undermine the test's capacity to refute the null hypothesis when it is indeed false. To mitigate this issue, it is recommended to studentize the test statistic and integrate a sample-dependent null distribution, thereby improving the identification of pertinent alter- natives. Expanding on this methodology, we propose a stepwise testing procedure rooted in various techniques derived from White's reality check test, as discussed by Hansen (2005), Romano and Wolf (2005), and Hsu, Hsu, and Kuan (2010). We commence by explicitly defining alternative hypotheses corresponding to the null hypothesis, adhering to the formulation provided in Equation 16.

$$H_A^k : G_k \geq 0, \text{for } k = 1, \cdots, K. \tag{12}$$

Rejecting the k_{th} individual null hypothesis suggests that the k_{th} technical strategy exhibits significant profitability after a comprehensive evaluation of all plausible alternative hypotheses, thereby effectively eliminating any potential data snooping bias. We detail the stepwise testing procedure with a Type I error rate α_0 for a specified temporal span $(t = 1, \cdots, T)$ as elucidated:

1-3. The preliminary set of actions aligns with those utilized in determining reality check p-values

4. We generate an empirical null distribution for the test statistics as illustrated here:
   (a) For each b, compute

$$Y_{bi} = T^{1/2} \max_{k=1,\cdots,K} \{G_{kb} - G_k + G_k \backslash mathbbm1\big(T^{1/2} G_k \leq -\sigma_k [2 \log(\log(T))]^{1/2}\big)\}, \tag{13}$$

where $\mathbb{1}(U)$ represents the characteristic function of the event U, and σ_k denotes the standard deviation of the daily return sequence for the k_{th} method.

The threshold

$$\mathbb{1}\big(T^{1/2} G_k \leq -\sigma_k [2 \log(\log(T))]^{1/2}\big) \tag{14}$$

utilized by Hansen (2005), is employed to adjust the distribution of G to mitigate the distortion caused by numerous "ineffective" strategies.

   (b) Aggregate all $\{Y_{bi}\}_{b=1,\ldots,B}$, sort them in decreasing order, and determine their $(1 - \alpha_0)_{th}$ quantile as $q_i(\alpha_0)$.

5. Compare each strategy's $T^{1/2}G_k$ to $q_i(\alpha_0)$. The k_{th} null hypothesis is rejected at the i_{th} stage if $T^{1/2}G_k > q_i(\alpha_0)$, as noted by Romano and Wolf (2005). Record all characteristics of these rejected strategies and mark them as dismissed at the i_{th} stage. Subsequently, return to Step 5, set $G_k = 0$ and $G_{kb} = 0$ for each rejected hypothesis k, and increment the loop counter from i to $i + 1$. However, if no strategy is rejected given $q_i(\alpha_0)$, i.e., $T^{1/2}G_k \leq q_j(\alpha_0)$ for the remaining j, cease and proceed to Step 7.
6. Restore the initial G_k from Z and compute each technical rule's marginal p-value, p_k, as the percentile of $T^{1/2}G_k$ in the final $\{Y_{bi}\}_{b=1,\ldots,B}$, serving as an empirical null reference.
7. Compare each technical rule's p_k to α_0. If $p_k < \alpha_0$, we conclude that the k_{th} strategy is advantageous over the sample period at the significance level of α_0. If at least one profitable strategy is identified within the sample period, we assert that technical trading is advantageous at the significance level of α_0, and the stepwise test p-value is $1 - \alpha_0$.

In our empirical evaluations, we begin by setting $\alpha_0 = 0.05$, indicating that our statistical significance is assessed at a 5% threshold. We also use $Q = 0.9$ and $B = 1000$ following established scholarly conventions. This framework ensures the robustness and consistency of our statistical analysis. Specifically, if an investment strategy shows positive returns but fails to pass our rigorous data-snooping assessments, there is an increased likelihood that its success may be due to random variance rather than genuinely strong performance. Thus, our methodology carefully examines the validity of these strategies' profitability by mitigating the impact of stochastic discrepancies, thereby enhancing the reliability of our empirical findings.

THE EMPIRICAL PERFORMANCE OF TRADING STRATEGY

In this section, we thoroughly analyze returns from various technical trading strategies using both in-sample and out-of-sample data. Starting with an initial capital of $100,000, we limit each buy and sell action to a maximum of one unit to effectively manage risk and maintain parity. Fractional trading is prohibited to ensure consistency across simulations. The model accounts for a slippage rate of 0.001 and a commission rate of 0.0003 per trade to cover transaction costs and brokerage fees. Our strategy is based on the previous day's closing prices, incorporating the latest market information into decisions. Additionally, the framework restricts trading to long positions only, avoiding short selling to comply with regulatory requirements and investor preferences.

TABLE 1
SUMMARY STATISTICS OF TRADING DATA

	1d			
	mean	std	min	max
AAPL				
open	173.638	153.107	12.990	702.410
high	175.498	154.362	13.190	705.070
low	171.511	151.628	12.720	699.570
close	173.560	153.031	13.120	702.100
volume	31444683.351	30048361.935	719436.000	331572938.000
returns	0.001	0.030	-0.855	0.139
MSFT				
open	72.662	73.862	15.200	344.620
high	73.459	74.624	15.620	349.670
low	71.841	73.012	14.870	342.200
close	72.669	73.859	15.150	343.110
volume	47531379.338	28054393.567	5908046.000	592924962.000
returns	0.000	0.020	-0.483	0.196
NVDA				
open	89.251	129.445	6.000	834.140
high	90.966	131.512	6.380	835.000
low	87.456	127.168	5.750	814.010
close	89.264	129.459	5.900	827.940
volume	13721826.608	12824781.620	149328.000	117779270.000
returns	0.001	0.041	-0.752	0.424

FIGURE 1
THE TREND OF CLOSING PRICE FOR AAPL

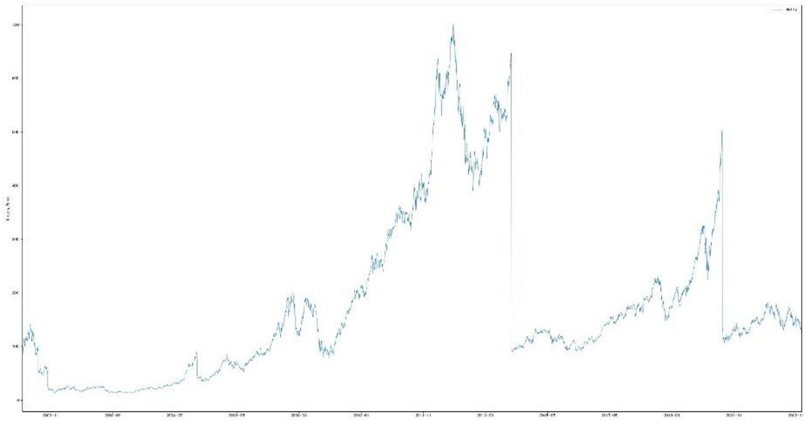


FIGURE 2
THE TREND OF CLOSING PRICE FOR MSFT

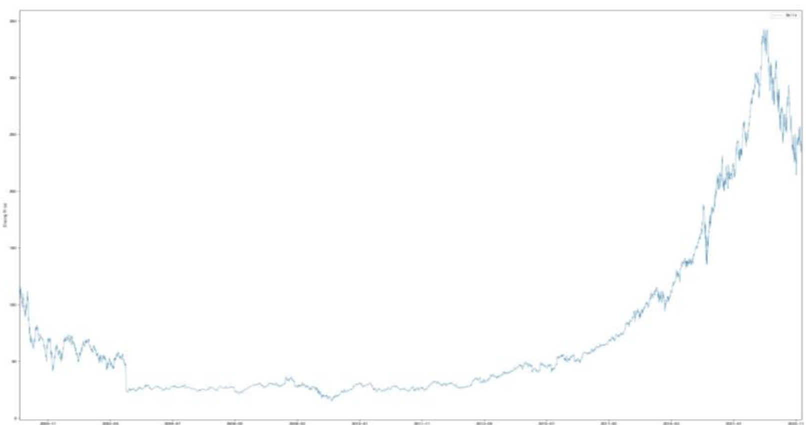
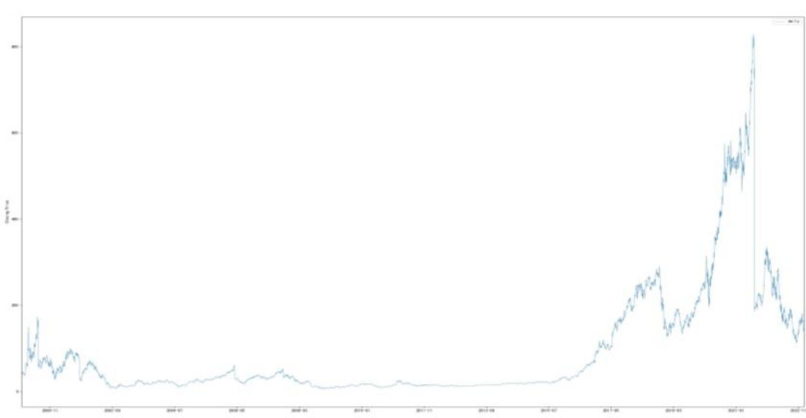


FIGURE 3
THE TREND OF CLOSING PRICE FOR NVDA



AAPL

Profitability Over the Whole Sample

Table 2 presents the results for the entire sample period, evaluated by annualized return and Sharpe ratio in the left and right panels, respectively. Each panel includes column for various technical trading strategy classifications based on daily data. Our analysis focuses on two metrics from the data snooping test: 1) performance indicators and p-values of the top strategy, and 2) the number of profitable strategies with significantly positive metrics, using a 5% significance level. The "Description" row specifies the optimal daily strategy, followed by the nominal p-value from the individual test. The next two rows show p-values from the reality check and stepwise tests. For example, column (1) identifies the top strategy as TEMA(5,30), with an annualized return of 13.256%, a Sharpe ratio of 0.446, and a maximum drawdown of 66.884%. However, its average return is not statistically significant (p-values: 0.016, 0.307, and 0.024). Figure 4 illustrates the cumulative log returns of the best-performing daily technical trading strategies.

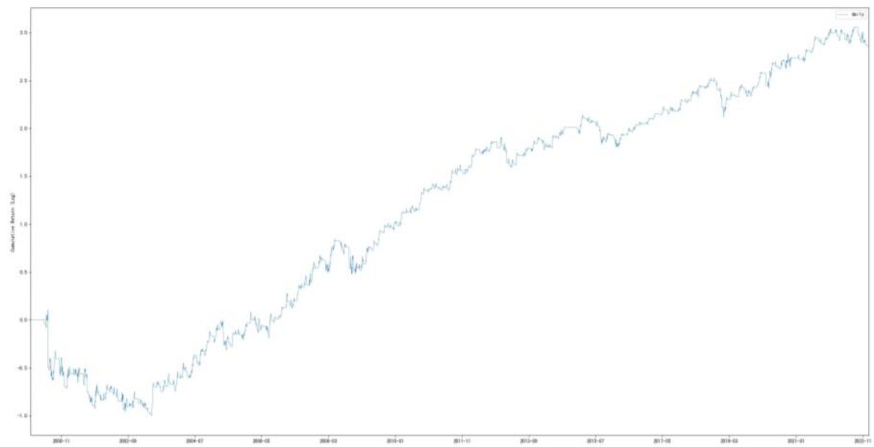
TABLE 2
THE PERFORMANCE OF TECHNICAL TRADE STRATEGIES IN THE WHOLE SAMPLE – AAPL

Performance Metric	Annualized Return	Sharpe Ratio
	(1)	(2)
Time Resolution	1d	1d
Number of Strategies	90	90
Best Strategy		
Description	tema(5, 30)	tema(5, 30)
Annualized Return	13.256%	13.256%
Sharpe Ratio	0.446	0.446
Max Drawdown	66.884%	66.884%
P-Value (Nomial)	0.016	0.000
P-Value (Reality Check)	0.307	0.021
P-Value (Stepwise Check)	0.024	0.019
All Profitable Strategies (500 Tests)		
Minimum Number	2	2
Maximum Number	4	4
Average Number	3.656	2.006
Average Number / Number of Strategies	2.611%	1.433%

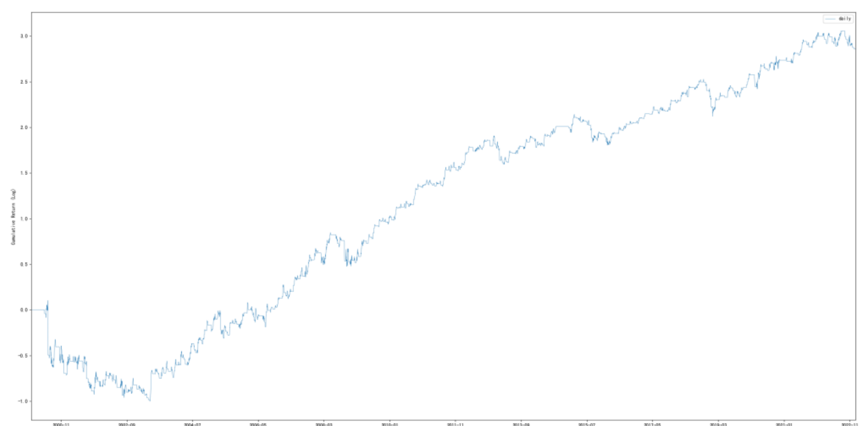
This table presents the profitability of technical trading strategies in the whole sample period. The left and right panels are based on the annualized return criterion and Sharpe ratio criterion, respectively. Within each panel, we have 1 column for different groups of technical trading strategies based on daily trading data we considered. In the panel titled "Best Strategy", we list the description, annualized return, Sharpe ratio, maximum drawdown, nominal p-value, reality check p-value, and stepwise test p-value of the best-performing strategy. In the panel titled "All Profitable Strategies", we list the average, minimum, and maximum number of profitable technical trading strategies from 500 stepwise tests. In the bottom row, we provide the ratio of the average number of profitable technical trading strategies to the total number of technical trading strategies considered. We use 5% significance level in our tests.

FIGURE 4
PERFORMANCE OF THE BEST TECHNICAL TRADE STRATEGY IN THE WHOLE SAMPLE – AAPL

Panel A. Annualized Return Criterion



Panel B. Sharpe Ratio Criterion



This figure plots the performances of the best technical trade strategies of daily trading data from Jan 01 2000 to Dec 31, 2022. Panel A and B plot the results based on the annualized return criterion and Sharpe ratio criterion, and we use Feb 06, 2016, as the cutoff of the in-/out-of-sample period.

Profitability in In-Sample and Out-of-Sample Periods

Test results for the Filter approach are detailed in Table 3, showing in-sample and out-of-sample metrics like annualized return and Sharpe ratio. The "Best Strategy (In-Sample)" and "Best Strategy (Out-of-Sample)" sections include key statistics such as maximum drawdown and p-values from various tests. To address data snooping, 500 tests were conducted, revealing the average number and proportion of profitable out-of-sample strategies. The top strategy, TEMA with a fast period of 5 and slow period of 30, achieved an 11.936% annualized return and a Sharpe ratio of 0.364 in-sample but showed no significant profitability out-of-sample, indicating the need for more robust methods. Figure 5 visualizes the cumulative log returns of top trading strategies.

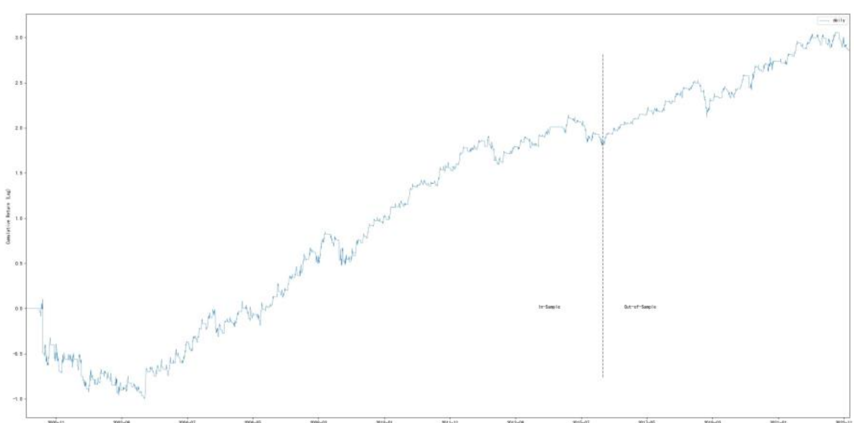
TABLE 3
THE IN-/OUT-OF-SAMPLE PERIOD PERFORMANCE OF TECHNICAL TRADE STRATEGIES WITH TEMA STRATEGIES – AAPL

Performance Metric	Annualized Return	Sharpe Ratio
	(1)	(2)
Time Resolution	1d	1d
Number of Strategies	50	50
Best Strategy (In-Sample)		
Description	tema(5, 30)	tema(5, 30)
Annualized Return	11.936%	11.936%
Sharpe Ratio	0.364	0.364
Max Drawdown	66.884%	66.884%
P-Value (Nomial)	0.072	0.001
P-Value (Reality Check)	0.412	0.000
P-Value (Stepwise Check)	0.152	0.000
Performance of the Best Strategy (Out-of-Sample)		
Annualized Return	20.735%	20.735%
Sharpe Ratio	1.021	1.021
Max Drawdown	34.625%	34.625%
P-Value (Nomial)	0.035	0.000
P-Value (Reality Check)	0.358	0.000
P-Value (Stepwise Check)	0.034	0.000
All Profitable Strategies (In-Sample, 500 Tests)		
Average Number	0.000	50.000
Average Number / Number of Strategies	0.0%	100.0%
Performance of Profitable Strategies (Out-of-Sample, 500 Tests)		
Average Number	10.000	50.000
Average Number / Number of Strategies (In-Sample)	20.0%	100.0%

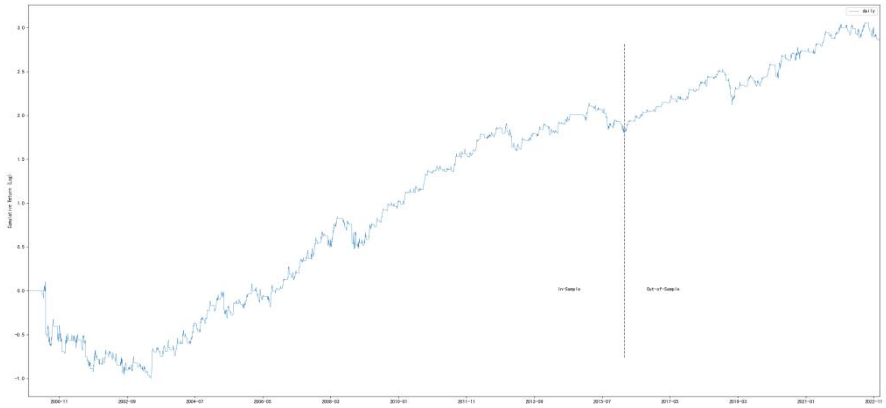
This table's settings are similar to Table 2.

FIGURE 5
PERFORMANCE OF THE BEST TECHNICAL TRADE TEMA STRATEGY – AAPL

Panel A. Annualized Return Criterion



Panel B. Sharpe Ratio Criterion



This figure's settings are similar to figure 4.

Table 4 presents the Zero-lag indicator method, identifying the best configuration with a duration of 38 and gain limit of 50, which yielded a 3.866% annualized return and a Sharpe ratio of 0.065 in-sample. However, similar to the Filter approach, it did not show significant profitability out-of-sample, with a 9.017% annualized return. This consistent underperformance highlights the necessity for more sustainable trading methodologies. Figure 6 illustrates the cumulative log returns of the leading technical trading methods.

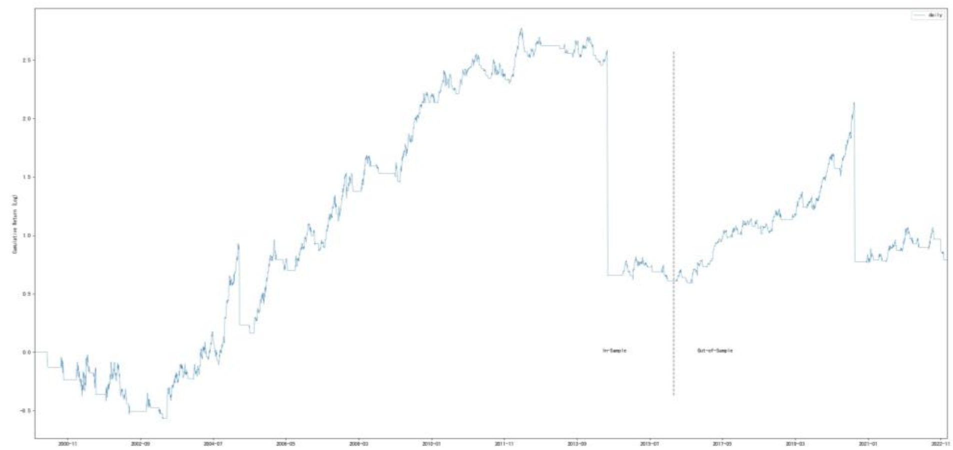
TABLE 4
THE IN-/OUT-OF-SAMPLE PERIOD PERFORMANCE OF TECHNICAL TRADE STRATEGIES WITH ZERO-LAG INDICATOR STRATEGIES – AAPL

Performance Metric	**Annualized Return**	**Sharpe Ratio**
	(1)	(2)
Time Resolution	1d	1d
Number of Strategies	40	40
Best Strategy (In-Sample)		
Description	zlind(38, 50)	zlind(38, 50)
Annualized Return	3.866%	3.866%
Sharpe Ratio	0.065	0.065
Max Drawdown	88.521%	88.521%
P-Value (Nomial)	0.397	0.000
P-Value (Reality Check)	0.543	0.000
P-Value (Stepwise Check)	0.542	0.000
Performance of the Best Strategy (Out-of-Sample)		
Annualized Return	9.017%	9.017%
Sharpe Ratio	0.158	0.158
Max Drawdown	74.651%	74.651%
P-Value (Nomial)	0.455	0.000
P-Value (Reality Check)	0.534	0.000
P-Value (Stepwise Check)	0.533	0.000
All Profitable Strategies (In-Sample, 500 Tests)		
Average Number	0.000	40.000
Average Number / Number of Strategies	0.0%	100.0%
Performance of Profitable Strategies (Out-of-Sample, 500 Tests)		
Average Number	0.000	40.000
Average Number / Number of Strategies (In-Sample)	0.0%	100.0%

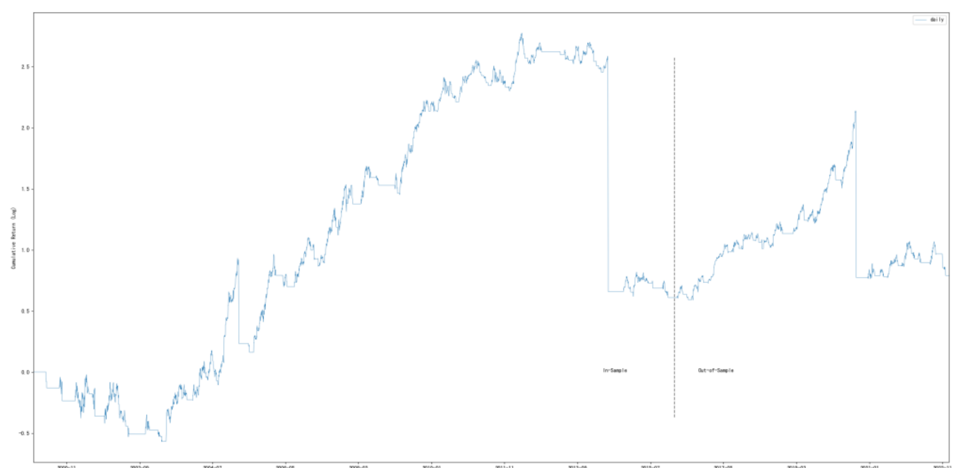
This table's settings are similar to Table 2.

FIGURE 6
PERFORMANCE OF THE BEST TECHNICAL TRADE ZERO-LAG INDICATOR STRATEGY – AAPL

Panel A. Annualized Return Criterion



Panel B. Sharpe Ratio Criterion



This figure’s settings are similar to Figure 4.

MSFT

Profitability Over the Whole Sample

Table 5 summarizes results for the entire sample period, evaluating technical trading strategies using annualized returns and Sharpe ratios in separate daily data panels. The analysis focuses on data snooping tests, highlighting the top strategy’s performance metrics and p-values, and counting strategies with significantly positive metrics at a 5% significance level. The "Description" row identifies the best daily strategy, presenting nominal p-values from individual tests alongside reality check and stepwise test p-values. For example, the TEMA strategy (column 1) has a 6.331% return and 0.325 Sharpe ratio but fails significance tests (p-values = 0.06, 0.263, 0.169). Additionally, Figure 7 shows cumulative log returns of top strategies.

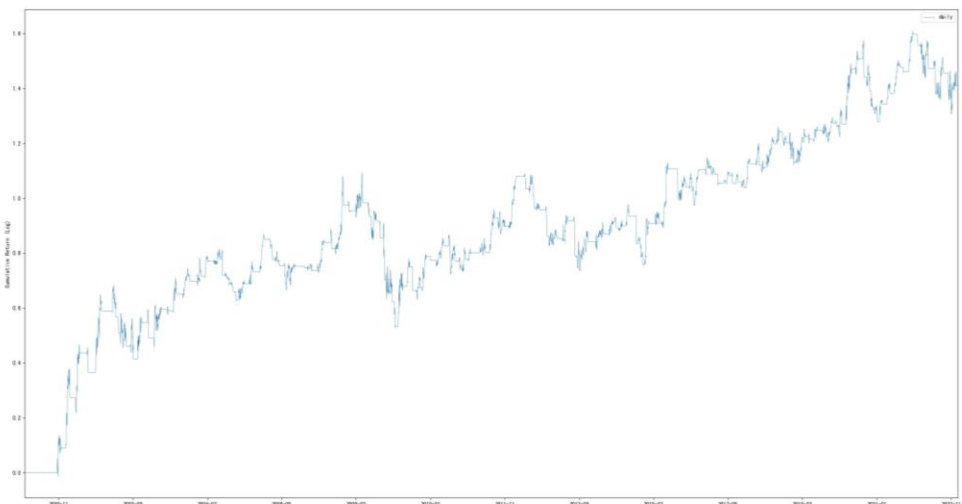
TABLE 5
THE PERFORMANCE OF TECHNICAL TRADE STRATEGIES IN THE WHOLE SAMPLE – MSFT

Performance Metric	Annualized Return	Sharpe Ratio
	(1)	(2)
Time Resolution	1d	1d
Number of Strategies	90	90
Best Strategy		
Description	tema(21, 32)	tema(21, 32)
Annualized Return	6.331%	6.331%
Sharpe Ratio	0.325	0.325
Max Drawdown	43.107%	43.107%
P-Value (Nomial)	0.060	0.000
P-Value (Reality Check)	0.263	0.000
P-Value (Stepwise Check)	0.169	0.000
All Profitable Strategies (500 Tests)		
Minimum Number	0	15
Maximum Number	0	16
Average Number	0.000	15.018
Average Number / Number of Strategies	0.0%	10.727%

This table's settings are similar to Table 2.

FIGURE 7
PERFORMANCE OF THE BEST TECHNICAL TRADE STRATEGY IN THE WHOLE SAMPLE – MSFT

Panel A. Annualized Return Criterion



Panel B. Sharpe Ratio Criterion

This figure's settings are similar to Figure 4.

Profitability in In-Sample and Out-of-Sample Periods

The validation of the TEMA methodology in Table 6 uses annualized return and Sharpe ratio as key metrics. The "Best Strategy (In-Sample)" section details the top in-sample strategy's description, annualized return (7.106%), Sharpe ratio (0.35), maximum drawdown (43.107%), and p-values from various tests. The "Best Strategy (Out-of-Sample)" section presents similar metrics without maximum drawdown for the leading out-of-sample strategy. Additionally, 500 data snooping tests address sampling bias, and the out-of-sample performance averages the profitability of these strategies. For example, the optimal TEMA strategy underperforms both in-sample and out-of-sample, failing to reject the null hypothesis and indicating the need for a more robust method. Figure 8 shows the cumulative log returns of top technical trading strategies.

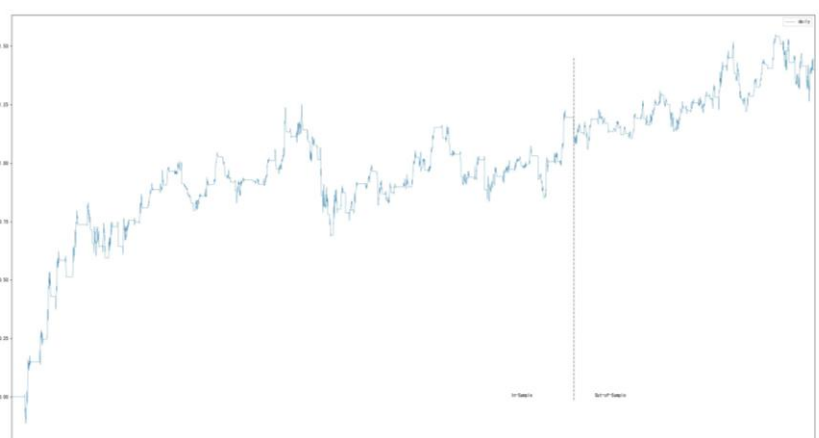
TABLE 6
THE IN-/OUT-OF-SAMPLE PERIOD PERFORMANCE OF TECHNICAL TRADE STRATEGIES WITH TEMA STRATEGIES – MSFT

Performance Metric	Annualized Return	Sharpe Ratio
	(1)	(2)
Time Resolution	1d	1d
Number of Strategies	50	50
Best Strategy (In-Sample)		
Description	tema(23, 30)	tema(23, 30)
Annualized Return	7.106%	7.106%
Sharpe Ratio	0.350	0.350
Max Drawdown	43.107%	43.107%
P-Value (Nomial)	0.081	0.000
P-Value (Reality Check)	0.302	0.000
P-Value (Stepwise Check)	0.188	0.000
Performance of the Best Strategy (Out-of-Sample)		
Annualized Return	8.812%	8.812%
Sharpe Ratio	0.489	0.489
Max Drawdown	30.524%	30.524%
P-Value (Nomial)	0.231	0.000
P-Value (Reality Check)	0.460	0.000
P-Value (Stepwise Check)	0.434	0.000
All Profitable Strategies (In-Sample, 500 Tests)		
Average Number	0.000	37.964
Average Number / Number of Strategies	0.0%	75.928%
Performance of Profitable Strategies (Out-of-Sample, 500 Tests)		
Average Number	0.000	37.964
Average Number / Number of Strategies (In-Sample)	0.0%	75.928%

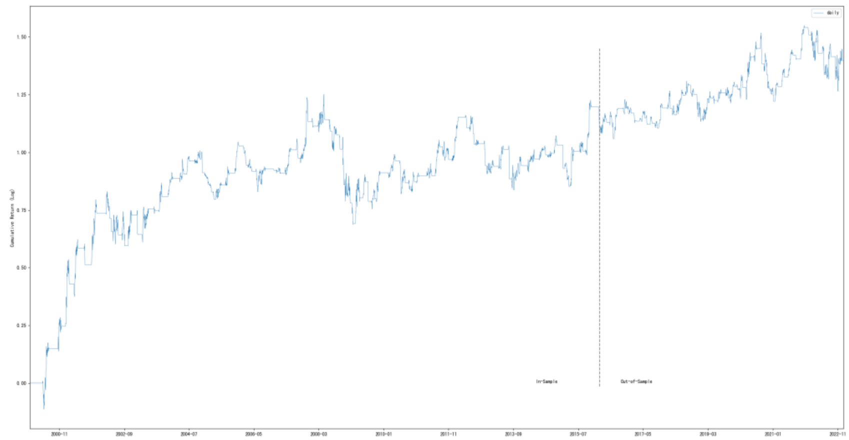
This table's settings are similar to Table 2.

FIGURE 8
PERFORMANCE OF THE BEST TECHNICAL TRADE TEMA STRATEGY – MSFT

Panel A. Annualized Return Criterion



Panel B. Sharpe Ratio Criterion



This figure's settings are similar to Figure 4.

Similarly, Table 7 evaluates the Zero-lag indicator approach. The optimal strategy, with a period of 26 and gain threshold of 50, records an annualized return of –0.825%, a Sharpe ratio of –0.042, and a maximum drawdown of 55.323%. This strategy performs poorly in both in-sample and out-of-sample tests, failing to reject the null hypothesis and demonstrating ineffectiveness. Figure 9 illustrates the cumulative logarithmic returns of these top trading strategies based on daily data.

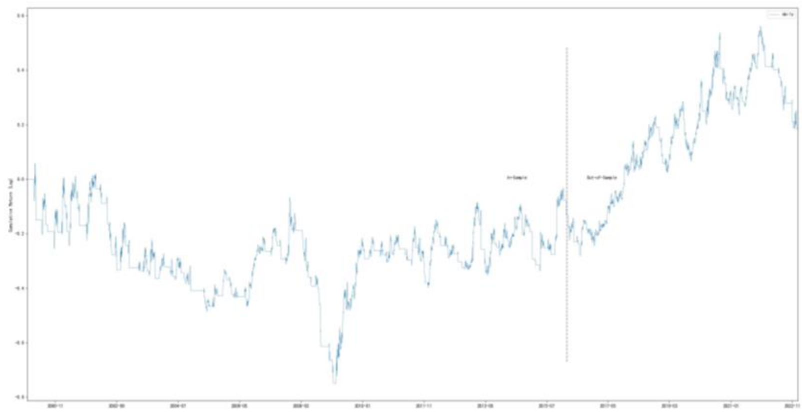
TABLE 7
THE IN-/OUT-OF-SAMPLE PERIOD PERFORMANCE OF TECHNICAL TRADE STRATEGIES WITH ZERO-LAG INDICATOR STRATEGIES – MSFT

Performance Metric	Annualized Return	Sharpe Ratio
	(1)	(2)
Time Resolution	1d	1d
Number of Strategies	40	40
Best Strategy (In-Sample)		
Description	zlind(26, 50)	zlind(26, 50)
Annualized Return	-0.825%	-0.825%
Sharpe Ratio	-0.042	-0.042
Max Drawdown	55.323%	55.323%
P-Value (Nomial)	0.434	0.000
P-Value (Reality Check)	0.759	0.000
P-Value (Stepwise Check)	0.705	0.000
Performance of the Best Strategy (Out-of-Sample)		
Annualized Return	11.27%	11.27%
Sharpe Ratio	0.574	0.574
Max Drawdown	26.056%	26.056%
P-Value (Nomial)	0.188	0.000
P-Value (Reality Check)	0.331	0.000
P-Value (Stepwise Check)	0.330	0.000
All Profitable Strategies (In-Sample, 500 Tests)		
Average Number	0.000	40.000
Average Number / Number of Strategies	0.0%	100.0%
Performance of Profitable Strategies (Out-of-Sample, 500 Tests)		
Average Number	0.000	40.000
Average Number / Number of Strategies (In-Sample)	0.0%	100.0%

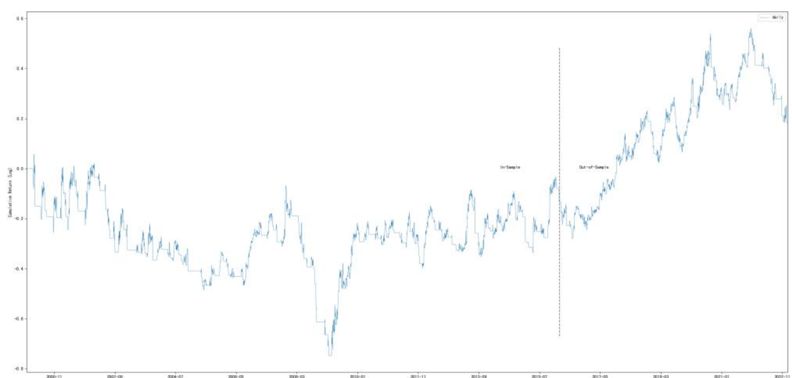
This table's settings are similar to Table 2.

FIGURE 9
PERFORMANCE OF THE BEST TECHNICAL TRADE ZERO-LAG INDICATOR STRATEGY – MSFT

Panel A. Annualized Return Criterion



Panel B. Sharpe Ratio Criterion



This figure's settings are similar to Figure 4.

NVDA

Profitability Over the Whole Sample

Table 8 presents the test results for the entire sample period, evaluating annualized return (left panel) and Sharpe ratio (right panel) for various technical trading strategies based on daily data. We focus on two main indicators from the data snooping test: the optimal strategy's performance and p-value, and the number of significantly profitable strategies, using a 5% significance level. The "Description" row identifies the top daily strategy, followed by its nominal p-value for comparison with previous studies. Subsequent rows show p-values from the reality check and stepwise tests. For example, column (1) highlights the tema(5,34) strategy with a 6.433% annualized return and a Sharpe ratio of 0.132, but its performance is not statistically significant (p-values: 0.264, 0.601, 0.598). Figure 10 illustrates the cumulative log returns of the top daily trading strategies.

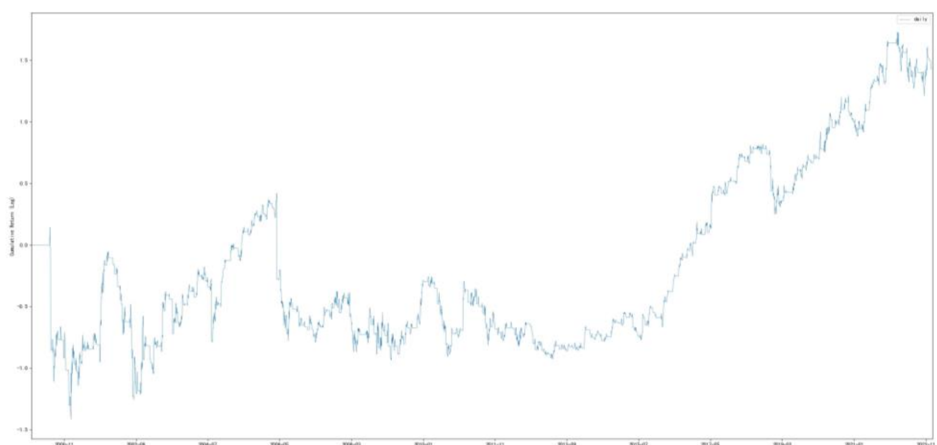
TABLE 8
THE PERFORMANCE OF TECHNICAL TRADE STRATEGIES IN THE WHOLE SAMPLE – NVDA

Performance Metric	Annualized Return	Sharpe Ratio
	(1)	(2)
Time Resolution	1d	1d
Number of Strategies	90	90
Best Strategy		
Description	tema(5, 34)	tema(5, 34)
Annualized Return	6.433%	6.433%
Sharpe Ratio	0.132	0.132
Max Drawdown	78.984%	78.984%
P-Value (Nomial)	0.264	0.000
P-Value (Reality Check)	0.601	1.000
P-Value (Stepwise Check)	0.598	0.777
All Profitable Strategies (500 Tests)		
Minimum Number	0	26
Maximum Number	1	27
Average Number	0.002	26.814
Average Number / Number of Strategies	0.001%	19.153%

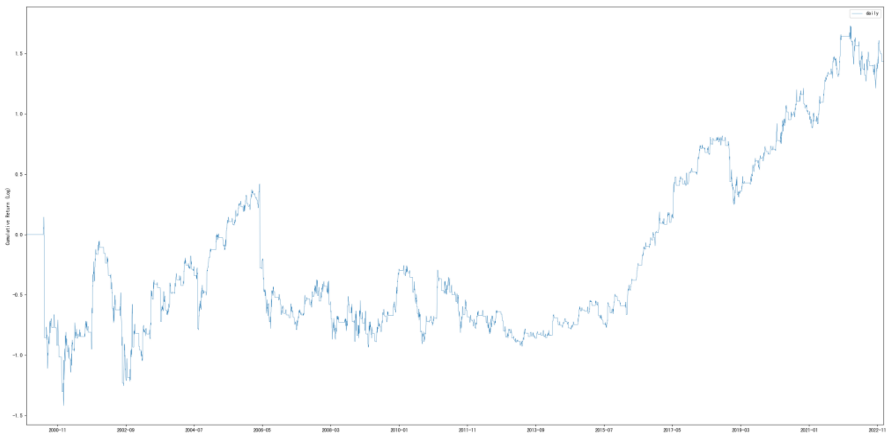
This table's settings are similar to Table 2.

FIGURE 10
PERFORMANCE OF THE BEST TECHNICAL TRADE STRATEGY IN THE WHOLE SAMPLE – NVDA

Panel A. Annualized Return Criterion



Panel B. Sharpe Ratio Criterion



This figure's settings are similar to Figure 4.

Profitability in In-Sample and Out-of-Sample Periods

Table 9 presents the empirical results for the TEMA strategy, evaluated both in-sample and out-of-sample using annualized return and Sharpe ratio. The "Best Strategy (In-Sample)" section details the top in-sample strategy's metrics, including p-values from nominal, reality check, and stepwise tests. Similarly, the "Best Strategy (Out-of-Sample)" section shows the best out-of-sample performance with corresponding p-values. Additionally, 500 data snooping tests are reported to minimize sampling bias, and the profitability of strategies out-of-sample is summarized. For example, the TEMA strategy with fast=5 and slow=32 periods shows poor performance, with negative returns and ineffective p-values in both samples, highlighting its ineffectiveness. Figure 11 illustrates the cumulative log returns of top technical trading strategies.

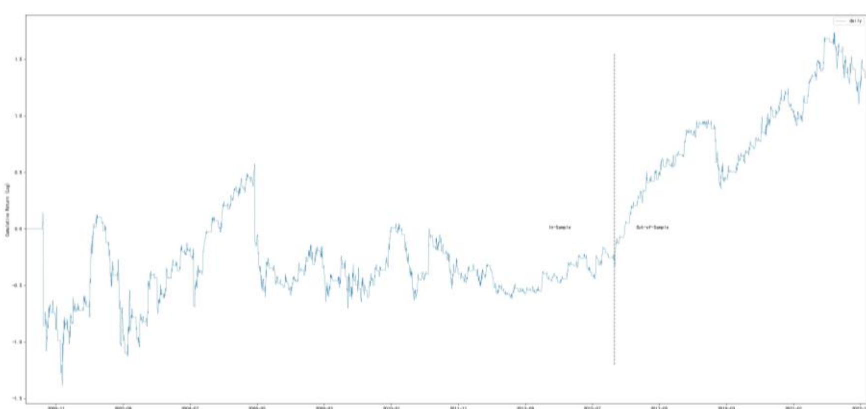
TABLE 9
THE IN-/OUT-OF-SAMPLE PERIOD PERFORMANCE OF TECHNICAL TRADE STRATEGIES WITH TEMA STRATEGIES – NVDA

Performance Metric	Annualized Return	Sharpe Ratio
	(1)	(2)
Time Resolution	1d	1d
Number of Strategies	50	50
Best Strategy (In-Sample)		
Description	tema(5, 32)	tema(5, 32)
Annualized Return	-2.029%	-2.029%
Sharpe Ratio	-0.039	-0.039
Max Drawdown	78.382%	78.382%
P-Value (Nomial)	0.437	0.000
P-Value (Reality Check)	0.800	0.000
P-Value (Stepwise Check)	0.749	0.000
Performance of the Best Strategy (Out-of-Sample)		
Annualized Return	29.192%	29.192%
Sharpe Ratio	0.745	0.745
Max Drawdown	50.869%	50.869%
P-Value (Nomial)	0.065	0.000
P-Value (Reality Check)	0.157	0.000
P-Value (Stepwise Check)	0.134	0.000
All Profitable Strategies (In-Sample, 500 Tests)		
Average Number	0.000	50.000
Average Number / Number of Strategies	0.0%	100.0%
Performance of Profitable Strategies (Out-of-Sample, 500 Tests)		
Average Number	2.116	50.000
Average Number / Number of Strategies (In-Sample)	4.232%	100.0%

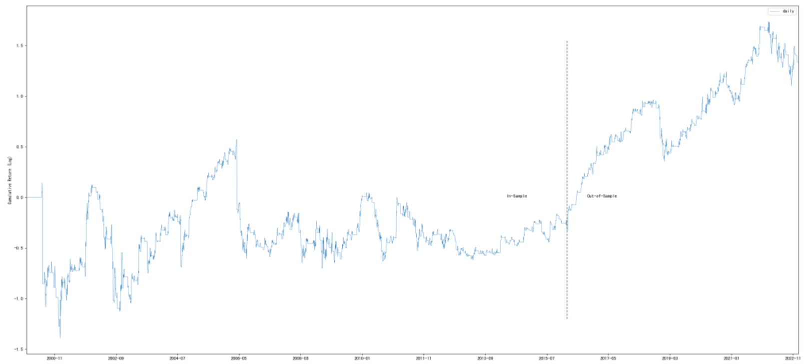
This table's settings are similar to Table 2.

FIGURE 11
PERFORMANCE OF THE BEST TECHNICAL TRADE TEMA STRATEGY – NVDA

Panel A. Annualized Return Criterion



Panel B. Sharpe Ratio Criterion



This figure's settings are similar to Figure 4.

Table 10 details the Filter approach's performance for both in-sample and out-of-sample periods. The best Filter strategy, with high=28 and low=18 intervals, shows zero annualized return, Sharpe ratio, and maximum drawdown, with p-values failing to reject the null hypothesis. This indicates the strategy is ineffective and unprofitable in both intervals. Figure 12 further depicts the cumulative log returns of these top technical trading strategies based on daily trading data.

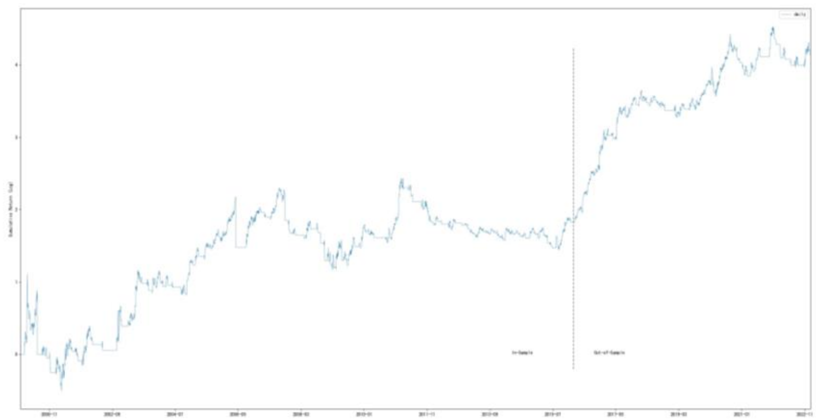
TABLE 10
THE IN-/OUT-OF-SAMPLE PERIOD PERFORMANCE OF TECHNICAL TRADE STRATEGIES WITH ZERO-LAG INDICATOR STRATEGIES – NVDA

Performance Metric	Annualized Return	Sharpe Ratio
	(1)	(2)
Time Resolution	1d	1d
Number of Strategies	40	40
Best Strategy (In-Sample)		
Description	zlind(24, 50)	zlind(24, 50)
Annualized Return	12.029%	12.029%
Sharpe Ratio	0.225	0.225
Max Drawdown	80.031%	80.031%
P-Value (Nomial)	0.184	0.000
P-Value (Reality Check)	0.259	0.000
P-Value (Stepwise Check)	0.258	0.000
Performance of the Best Strategy (Out-of-Sample)		
Annualized Return	39.376%	39.376%
Sharpe Ratio	0.913	0.913
Max Drawdown	40.201%	40.201%
P-Value (Nomial)	0.010	0.000
P-Value (Reality Check)	0.123	0.000
P-Value (Stepwise Check)	0.029	0.000
All Profitable Strategies (In-Sample, 500 Tests)		
Average Number	0.000	40.000
Average Number / Number of Strategies	0.0%	100.0%
Performance of Profitable Strategies (Out-of-Sample, 500 Tests)		
Average Number	14.684	40.000
Average Number / Number of Strategies (In-Sample)	36.71%	100.0%

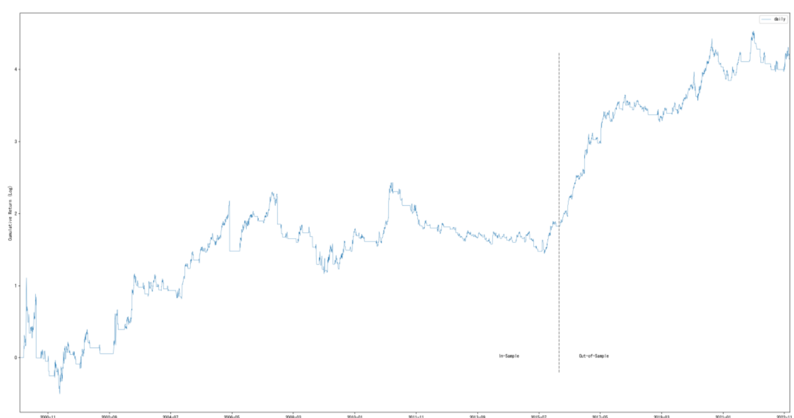
This table's settings are similar to Table 2.

FIGURE 12
PERFORMANCE OF THE BEST TECHNICAL TRADE ZERO-LAG INDICATOR STRATEGY – NVDA

Panel A. Annualized Return Criterion



Panel B. Sharpe Ratio Criterion



This figure's settings are similar to Figure 4.

CONCLUSION

In this paper, we analyze a comprehensive dataset of AAPL, MSFT, and NVDA stocks from January 2000 to December 2022 to evaluate the profitability of various technical trading strategies both in-sample and out-of-sample, using February 2016 as the primary cutoff and May 2018 as an alternative. We construct strategies based on multiple indicators and timeframes, conducting thorough statistical analyses to ensure robustness against data-snooping bias. Our results consistently demonstrate that apparent profitability often stems from parameter selection rather than true market inefficiencies, supporting the efficient market hypothesis. This highlights the difficulty in predicting profitable strategies ahead of time, emphasizing the unpredictable nature of achieving sustained trading success.

Additionally, we investigate the impact of corporate governance mechanisms on firm performance using an extensive dataset. Our findings reveal that strong governance structures, such as board independence and the separation of CEO and chair roles, significantly enhance financial outcomes. Shareholder activism also plays a crucial role in improving governance and increasing firm value. This research provides robust empirical evidence that effective governance mechanisms are pivotal in driving superior financial performance, contributing valuable insights to the existing literature.

REFERENCES

Alexander, S.S., (1961). Price movements in speculative markets: Trends or random walks. *Industrial Management Review (pre-1986)*, *2*, 7.

Brock, W., Lakonishok, J., & LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *The Journal of finance*, *47*, 1731–1764.

Fama, E.F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*, 383–417.

Hansen, P.R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics*, *23*, 365–380.

Hsu, P.-H., Hsu, Y.-C., & Kuan, C.-M. (2010). Testing the predictive ability of technical analysis using a new stepwise test without data snooping bias. *Journal of Empirical Finance*, *17*, 471–484.

Kwon, K.-Y., & Kish, R.J. (2002). Technical trading strategies and return predictability: Nyse. *Applied Financial Economics*, *12*, 639–653.

Leamer, E.E. (1978). Regression selection strategies and revealed priors. *Journal of the American Statistical Association*, *73*, 580–587.

Politis, D.N., & Romano, J.P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, *89*, 1303–1313.

Ratner, M., & Leal, R.P.C. (1999). Tests of technical trading strategies in the emerging equity markets of Latin America and Asia. *Journal of Banking Finance*, *23*, 1887–1905.

Romano, J.P., & Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, *73*, 1237–1282.

Tam, P.H., & Cuong, N.T. (2018). Effectiveness of investment strategies based on technical indicators: Evidence from Vietnamese stock markets. *Journal of Insurance and Financial Management*, *3*, 55–68.

Taylor, N. (2014). The rise and fall of technical trading rule success. *Journal of Banking & Finance*, *40*, 286–302.

White, H. (2000). A reality check for data snooping. *Econometrica*, *68*, 1097–1126.