International Review of Financial Analysis 82 (2022) 102173

Contents lists available at ScienceDirect

International Review of Financial Analysis

journal homepage: www.elsevier.com/locate/irfa

# Time series momentum in the US stock market: Empirical evidence and theoretical analysis

Valeriy Zakamulin [a,*], Javier Giner [b]

[a] School of Business and Law, University of Agder, Service Box 422, 4604 Kristiansand, Norway
[b] Department of Economics, Accounting and Finance, University of La Laguna, Camino La Hornera s/n, 38071, Santa Cruz de Tenerife, Spain

ARTICLE INFO

*JEL classification:*
G12
G14
G17

*Keywords:*
Time series momentum
Trend-following
Profitability
Statistical power

ABSTRACT

There is much controversy in the academic literature on the presence of short-term trends in financial markets and the trend-following strategy's profitability. We restrict our attention to studying the time series momentum in the S&P Composite stock price index. Our contributions are both empirical and theoretical. On the empirical side, we present compelling evidence of the presence of short-term momentum. For the first time, we suppose that the returns follow a $p$-order autoregressive process and evaluate this process's parameters. On the theoretical side, we develop a tractable theoretical model that contributes to our fundamental understanding of the trend-following strategy's risk, return, and performance. Using our model, we also estimate the power of statistical tests on the trend-following strategy's profitability and find that these tests suffer from the low power problem.

## 1. Introduction

Investment professionals have always believed that asset prices tend to move in trends. Trend-following strategies that try to identify and profit from market trends have existed for a century or more. In contrast, academics had long been skeptical about the benefits of trend-following. In their landmark study, Brock et al. (1992) demonstrate for the first time the benefits of trend-following strategies. This study dramatically changed the academics' attitude toward the existence of market trends. Since then, many empirical studies have documented the profitability of trend-following strategies.[1]

The time series momentum (TSMOM) strategy, presented by Moskowitz et al. (2012), is an example of a trend-following strategy. Using a comprehensive dataset of different US asset classes, Moskowitz et al. (2012) show that the past 12-month returns predict the next month's return; a trading strategy, which buys assets if their past 12-month returns are positive and sells them otherwise, earns significant risk-adjusted returns. In essence, a TSMOM strategy is a long–short portfolio strategy where each return is scaled by its volatility. The profitability of the TSMOM strategy in individual asset classes and international markets is further documented among others by Georgopoulou and Wang (2016), Lim et al. (2018), and Ham et al. (2019).

However, recently the profitability of the TSMOM strategy has been seriously questioned on two grounds. First, Kim et al. (2016) find that the superiority of the TSMOM strategy is primarily driven by volatility-scaling returns rather than by the short-term momentum effect. Effectively, there is no scientific evidence of the TSMOM strategy's profitability without volatility-scaling. Second, Huang et al. (2020) show, among other things, that asset-by-asset time series regressions reveal almost no evidence of short-term momentum. For example, they find no evidence of short-term momentum in the S&P 500 index. All in all, the papers by Kim et al. (2016) and Huang et al. (2020) cast severe doubts on the existence and profitability of short-term momentum in financial markets.

As a matter of fact, the scientific evidence of trend-following strategies' profitability has always been a problematic issue in the literature. Many studies find that trend-following strategies are profitable in the long run over periods ranging from 50 to 150 years. However, researchers frequently report that, when they use the most recent 5 to 10 years in their historical data sample, the trend-following strategies are not profitable (see, for example, Hutchinson & O'Brien, 2014; Lee et al., 2001; Okunev & White, 2003; Olson, 2004, Siegel, 2002, Chapter 17, Sullivan et al. (1999), and Zakamulin, 2014). In some other studies, the researchers fail to establish the statistical significance of

* Corresponding author.
*E-mail addresses:* Valeri.Zakamouline@uia.no (V. Zakamulin), jginer@ull.edu.es (J. Giner).

[1] Examples of such studies are: Clare et al. (2013), Faber (2007, 2017), Fifield et al. (2005), Georgopoulou and Wang (2016), Gwilym et al. (2010), Ham et al. (2019), Hutchinson and O'Brien (2014), Kilgallen (2012), Lee et al. (2001), Lim et al. (2018), Marshall et al. (2017), Moskowitz et al. (2012), Neely et al. (2014), Okunev and White (2003), Olson (2004), Siegel (2002), Sullivan et al. (1999), and Zakamulin (2017).

https://doi.org/10.1016/j.irfa.2022.102173
Received 22 December 2021; Received in revised form 14 March 2022; Accepted 19 April 2022
Available online 22 April 2022

profitability even when the historical sample is rather long, and the profitability of trend-following strategies is economically significant (see, among others, Kim et al., 2016; Zakamulin, 2017, and Huang et al., 2020).

There are two alternative explanations for the lack of scientific evidence on the profitability of trend-following strategies. The first explanation is that the profitability of trend-following strategies is an artifact of data-snooping. A variant of this explanation is that market efficiency improves over time. As a result, the trend-following rules that were profitable in the past have lost their effectiveness over time. The second alternative explanation is that the trend-following strategies are profitable, but the statistical tests for profitability have low power. It is known that a test's statistical power is directly related to the sample size; the power increases with increased sample size. Therefore, one can argue that the inability to reject the null hypothesis of no superiority does not imply that the null hypothesis is true but means that the sample size is not long enough.

In this paper, we restrict our attention to studying the time series momentum in the S&P Composite stock price index. Our contributions are both empirical and theoretical. On the empirical side, we aim to resolve the existing controversy regarding the presence of short-term trends in the market. On the theoretical side, we develop a tractable theoretical model that can significantly contribute to our fundamental understanding of the TSMOM strategy and help explain the existing controversy on its profitability.

In particular, in the first half of this paper, we conduct a novel empirical study that presents compelling evidence of short-term momentum in the excess returns on the S&P Composite index. For the first time, we assume that the excess returns follow a $p$-order autoregressive process, $AR(p)$, with $p > 1$ and evaluate the parameters of this process. The difficulty is that the autocorrelation in excess returns is very weak over monthly horizons and escapes detection when traditional estimation methods are used. We suggest a methodology that uses excess returns aggregated over multiple months and finds the parameters of the $AR(p)$ process that produces the best fit to a theoretical model.

All previous results on the performance and profitability of trend-following rules have been obtained solely through empirical research. Typically, an empirical study uses historical data to simulate the returns to a trend-following strategy and subsequently tests the null hypothesis that the trend-following strategy's performance is equal to that of the buy-and-hold strategy. The major limitation of such empirical studies is that they are conducted using a single and relatively short historical realization of a random process. In this regard, two obstacles deserve mentioning. First, one has no idea whether the number of observations is sufficient to detect the profitability of a trend-following strategy. Second, an empirical study is not integrated with a theoretical model that allows researchers to justify and explain the observed results and provide deep insights into the nature of the trend process and the properties of trend-following rules.

In contrast to previous studies, the outcome of our empirical analysis is not only the evidence of short-term momentum but also a tractable and well-understood statistical model that describes the dynamics of trends. Fairly accurate knowledge of this model allows us to provide several important theoretical implications for the profitability of the TSMOM strategy and its other properties. The model also allows us to evaluate the power of statistical tests and the required sample size to achieve the desired statistical power.

More specifically, we start the second half of this paper by providing analytical results on the mean return, variance of returns, Sharpe ratio, and CAPM beta and alpha of the long-only and long–short TSMOM strategies. The estimated parameters of the model for the excess returns are utilized to evaluate the profitability of these strategies analytically. We demonstrate that our analytical results agree well with the empirical results obtained by simulating the historical returns to the TSMOM strategies. Subsequently, we conduct a comparative static analysis of our theoretical model to examine how the profitability of the TSMOM strategies depends on the model's parameters. Finally, by relying on a simulation analysis, we explore how the evidence of the TSMOM strategy's superior performance depends on the investment horizon and evaluate the power of statistical tests.

The Sharpe ratio and CAPM alpha are the two most common performance measures in modern finance. The results of our analysis reveal that the choice of performance measure plays a crucial role in determining whether a TSMOM strategy is superior to its passive counterpart and the outcome of comparing the performance of the long–short TSMOM strategy with that of the long-only TSMOM strategy. In particular, our analysis suggests that, when the Sharpe ratio measures the performance, the long-only TSMOM strategy outperforms the buy-and-hold strategy, while the superiority of the long–short TSMOM strategy is questionable. By contrast, when the CAPM alpha measures the performance, the long–short TSMOM strategy is superior to the long-only TSMOM strategy, which, in turn, is superior to the buy-and-hold strategy.

By convention, the desired power of a statistical test is 80%. Our ballpark estimate is that the sample size must be about 250 years with monthly observations to reach the desired power level when the Sharpe ratio measures the performance. When the performance is measured by alpha, the sample size must be about 90 years to ensure the recommended statistical power. Even though the alpha test has notably larger statistical power than the Sharpe ratio test, it still requires a sample size that spans almost a century. Consequently, we conclude that in virtually all empirical studies that evaluate a trend-following strategy's profitability, the power of a statistical test is much below the acceptable level. This finding explains the lack of scientific evidence on the TSMOM strategy's profitability.

The rest of the paper is organized as follows. Section 2 presents the TSMOM trading rules. Section 3 motivates the choice of the $AR(p)$ process to model the price trends, while Section 4 describes the data. The empirical results are presented in Section 5. In particular, this section documents the evidence of short-term momentum and evaluates the autoregressive process parameters for excess returns. Section 6 presents a number of analytical and simulation results on the profitability of the TSMOM strategies and the statistical power of the tests on profitability. Finally, Section 7 concludes the paper.

## 2. Time series momentum trading strategies

Denote by $r_t$ and $r_{ft}$ the month $t$ total return on the stock market and the risk-free rate of return, respectively. Further, denote by $X_t$ the month $t$ market excess return

$$X_t = r_t - r_{ft}.$$

The TSMOM trading signal for month $t$ is generated in a two-step process. First, the TSMOM trading indicator's value is computed at month-end $t-1$ by summing the excess returns over the past $n$ months (including the last one):

$$MOM_{t-1}(n) = \sum_{i=1}^{n} X_{t-i}.$$

Subsequently, this value is translated into a trading signal for month $t$ as follows. If $MOM_{t-1}(n) > 0$, then the trading signal is Buy. Otherwise, the trading signal is Sell.

This paper considers both the "long-only" and "long–short" trading strategies. In both strategies, a Buy signal commands buying the stocks. The long-only TSMOM strategy seeks to generate profits and limit losses by investing in the stocks only when prices trend upwards. In this strategy, if a Sell signal is generated after a Buy signal, the trader sells the stocks and invests the proceeds at the risk-free rate. Therefore, the return to the long-only TSMOM strategy over month $t$ is given by

$$R_t^{LO} = \begin{cases} r_t & \text{if } MOM_{t-1}(n) > 0, \\ r_{ft} & \text{otherwise.} \end{cases} \quad (1)$$

The long–short TSMOM strategy tries to take advantage of profit opportunities in both upward- and downward-trending markets. In this case, when a Sell signal is generated after a Buy signal, the trader sells all own shares of the stocks and, subsequently, sells short the same number of shares of the stocks. The proceeds from both sales are invested at the risk-free rate. Thus, the return to the long–short TSMOM strategy over month $t$ is given by

$$R_t^{LS} = \begin{cases} r_t & \text{if } MOM_{t-1}(n) > 0, \\ 2r_{ft} - r_t & \text{otherwise.} \end{cases} \tag{2}$$

The strategy introduced by Moskowitz et al. (2012) is an extension of the long–short TSMOM strategy that uses the window size of $n = 12$ months and involves several risky assets weighted in inverse proportion to the volatility of each asset.

## 3. Process for excess returns

For the TSMOM strategy to be profitable, excess returns must exhibit short-term persistence. In other words, excess returns must be positively autocorrelated. To this end, we assume that the excess returns follow a $p$-order autoregressive process, $AR(p)$. The following equation defines this process:

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t = c + \sum_{i=1}^{p} \phi_i X_{t-i} + \varepsilon_t, \tag{3}$$

where $X_{t-i}$ is the excess return observed at month $t-i$, $c$ is a constant, $\{\phi_1, \phi_2, \dots, \phi_p\}$ are the parameters of the model, and $\varepsilon_t$ are i.i.d. random variables with zero mean and variance of $\sigma_\varepsilon^2$. We assume that the $AR(p)$ process satisfies the stationarity conditions.

In our paper, the process's persistence is measured by the sum of the autoregressive coefficients, $\kappa = \sum_{i=1}^{p} \phi_i$. This measure of persistence was originally introduced by Andrews and Chen (1994). The rationale for this measure is that every $AR(p)$ process exhibits mean-reverting behavior, and the mean reversion's speed is inversely proportional to $\kappa$. Specifically, a large value of $\kappa$ implies a slow reversion to the long-run mean and, hence, a strong persistence of the $AR(p)$ process. For more details on this measure of persistence, the interested reader is referred to Marques (2005).

The mean and variance of $X_t$ are given by (see, for example, Box et al. (2016), Chapter 3)

$$\mu_x = \frac{c}{1 - \sum_{i=1}^{p} \phi_i}, \qquad \sigma_x^2 = \frac{\sigma_\varepsilon^2}{1 - \sum_{i=1}^{p} \phi_i \rho_i}, \tag{4}$$

where $\rho_i$ denotes the autocorrelation between $X_t$ and $X_{t-i}$.

One of this paper's main goals is to evaluate the parameters of the $AR(p)$ process for excess returns using long-term data for the S&P Composite index. In particular, our goal is to evaluate the values of $p$ and $\phi_i$ for all $i \in [1, p]$. The main obstacle is that we have no clue about the process that governs the return persistence. As a result, there is a wide variety in the number of parameters and their functional form. Therefore, to make this goal feasible, we propose the following conjecture:

**Conjecture 1.** *The TSMOM rule with $n$ return lags represents the most optimal trading rule among all feasible trend-following rules. In particular, the Sharpe ratio of the strategy based on using the $MOM(n)$ trading indicator represents the highest possible Sharpe ratio.*

Conjecture 1 allows us to narrow down the number of unknown parameters of the $AR(p)$ process dramatically. The logic is as follows. Acar (2003) proves that the Sharpe ratio of a trend following strategy increases as the correlation between the trading indicator and the future return increases. Therefore, for the TSMOM rule to have the highest possible Sharpe ratio among all feasible trend-following strategies, the $MOM(n)$ trading indicator must have the highest possible correlation with the future return. When returns follow the $AR(p)$ process, Zakamulin and Giner (2020) show that the $MOM(n)$ trading indicator has the highest possible correlation with the future return when all autoregressive coefficients are alike, $\phi_i = \phi$, and the number of autoregressive terms equals the number of return lags in the TSMOM rule, $p = n$. Consequently, our task reduces to evaluating only two values: $\phi$ and $p$.

The Yule–Walker equations imply that, in the special case where $\phi_i = \phi$ for all $i \in [1, p]$, the first $p$ autocorrelation coefficients are alike, $\rho_i = \rho$, and equal to

$$\rho = \frac{\phi}{1 - (p-1)\phi}. \tag{5}$$

Therefore, Eqs. (4) for the mean and variance of the process for excess returns reduce to

$$\mu_x = \frac{c}{1-\kappa}, \qquad \sigma_x^2 = \frac{\sigma_\varepsilon^2}{1-\rho\kappa}, \qquad \kappa = p\phi. \tag{6}$$

## 4. Data

The S&P Composite index is a proxy for the U.S. stock market over the very long run. The risk-free rate of return is proxied by the T-Bill rate. All data come at the monthly frequency. The sample period used in our study starts in January 1857 and ends in December 2018 (162 full years).

William Schwert[2] provides the data for the U.S. market returns from January 1857 to December 1925. The market returns for this period are constructed using a collection of early stock market indices for the U.S. The methodology of construction is described in all detail in Schwert (1990). From January 1926 to February 1957, the market returns are the returns on the S&P 90 stock market index. Starting from March 1957, the market returns are the returns on the S&P 500 stock market index. Amit Goyal[3] provides the returns for the period from January 1926 to December 2018.

Amit Goyal also provides the data for the T-Bill rate from January 1920 to December 2018. There are no data for the short-term risk-free debt before January 1920. To estimate the risk-free rate over the period from January 1857 to December 1919, we apply the methodology suggested by Welch and Goyal (2008). This methodology constructs the instrumented risk-free rate from the Commercial Paper Rates for New York.[4]

We are primarily interested in determining the order $p$ of the autoregressive process for excess returns and evaluating coefficients $\phi_i$. Moskowitz et al. (2012) report that the TSMOM strategy produces a rather stable performance when the number of return lags $n \in [6, 12]$. Therefore, we expect that the number of autoregressive terms $p$ lies somewhere between 6 and 12. The most straightforward approach to estimating the autoregressive coefficients is to use an OLS regression model.

Table 1 reports the estimated autoregressive coefficients $\phi_i$ of the $AR(p)$ process for excess returns. Additionally, this table shows the estimated trend strength $\kappa = \sum_{i=1}^{12} \phi_i$. In brief, our results suggest evidence in favor of weak persistence ($\kappa = 0.246$) in excess returns in the first half of the sample. In this half, there are three positive and statistically significant coefficients: $\phi_1$, $\phi_5$, and $\phi_8$. In the whole sample period and the second half, the persistence ($\kappa = 0.167$ and $\kappa = 0.127$ respectively) is much weaker than in the first half, and the only positive and statistically significant coefficient is $\phi_5$. In this context, the well-known issue is that when one conducts testing of many coefficients for statistical significance, one will inevitably find coefficients that are "significant". That is, statistical significance can be caused by luck or

[2] http://schwert.ssb.rochester.edu/data.htm

[3] http://www.hec.unil.ch/agoyal/

[4] The data for the Commercial Paper Rates for New York from January 1857 to December 1971 are available from the NBER Macrohistory database. http://research.stlouisfed.org/fred2/series/M13002US35620M156NNBR.

**Table 1**
Estimated autoregressive coefficients $\phi_i$ of the $AR(p)$ process for the excess returns and the empirical trend strength $\kappa = \sum_{i=1}^{12} \phi_i$. The reported standard errors are Newey–West heteroscedasticity and autocorrelation consistent standard errors computed using 12 lags. Bold values indicate statistical significance at the 5% level.

| Parameter | 1858–2018 | | 1858–1937 | | 1938–2018 | |
|---|---|---|---|---|---|---|
| | Estimate | Std. error | Estimate | Std. error | Estimate | Std. error |
| $\phi_1$ | 0.067 | 0.042 | **0.118** | 0.057 | 0.017 | 0.044 |
| $\phi_2$ | 0.001 | 0.036 | −0.011 | 0.054 | −0.006 | 0.031 |
| $\phi_3$ | −0.056 | 0.042 | −0.094 | 0.060 | 0.034 | 0.034 |
| $\phi_4$ | 0.022 | 0.031 | 0.024 | 0.046 | 0.036 | 0.037 |
| $\phi_5$ | **0.077** | 0.027 | **0.083** | 0.040 | **0.073** | 0.034 |
| $\phi_6$ | −0.042 | 0.036 | −0.049 | 0.054 | −0.065 | 0.041 |
| $\phi_7$ | 0.025 | 0.029 | 0.055 | 0.042 | 0.004 | 0.029 |
| $\phi_8$ | 0.043 | 0.032 | **0.092** | 0.041 | −0.030 | 0.036 |
| $\phi_9$ | 0.016 | 0.038 | 0.026 | 0.046 | −0.011 | 0.040 |
| $\phi_{10}$ | 0.013 | 0.030 | 0.047 | 0.040 | −0.008 | 0.038 |
| $\phi_{11}$ | 0.004 | 0.034 | −0.021 | 0.048 | 0.033 | 0.036 |
| $\phi_{12}$ | −0.005 | 0.031 | −0.023 | 0.044 | 0.048 | 0.032 |
| $\kappa$ | 0.167 | | 0.246 | | 0.127 | |

chance. Therefore, the results for the second half of the sample seem to suggest that TSMOM is an artifact. However, it is premature to jump to definite conclusions. In the subsequent section, we present strong evidence that TSMOM does exist.

## 5. Empirical evidence of persistence and model identification procedure

This section starts with presenting strong evidence of short-term persistence in the excess returns on the S&P Composite stock price index. The section continues with evaluating the parameters $p$ and $\phi$ of the autoregressive process for excess returns.

### 5.1. A test of return persistence

Moskowitz et al. (2012) show the time series predictability of the excess returns by regressing the month $t$ excess return on the excess return lagged $h$ months:

$$X_t = \delta_h + \gamma_h X_{t-h} + \varepsilon_t. \tag{7}$$

To increase the test power, Moskowitz et al. (2012) compute the $t$-statistic of $\gamma_h$ using the pooled regression across many asset classes. They demonstrate that the $t$-statistic is positive for the first 12 lags; for the most lags, the value of the $t$-statistic implies that $\gamma_h$ is statistically significantly positive.

Since the TSMOM strategy is profitable, it is natural to assume that the past 12-month returns predict the next month's return. However, when Huang et al. (2020) run asset-by-asset time series predictive regressions

$$X_t = \delta + \gamma \sum_{i=1}^{12} X_{t-i} + \varepsilon_t, \tag{8}$$

these regressions reveal almost no evidence of predictability. In particular, Huang et al. (2020) find no evidence that the past 12-month returns on the S&P 500 index predict the next month return.

We maintain that the autocorrelation in excess returns at any lag is very weak over monthly horizons and escapes detection when traditional estimation methods (e.g., OLS) are used. To provide reliable evidence of short-term persistence in excess returns, we suggest a new methodology that is based on two key elements. The first key element is to use excess returns aggregated over multiple months. A similar idea was put forward already by Fama and French (1988) who, to detect the presence of weak mean-reversion in stock prices, advocate for using the first-order autocorrelation of returns aggregated over multiple periods. Formally, the methodology suggested by Fama and French (1988) consists in running the following regression for various aggregation periods of $k$ months:

$$\sum_{i=1}^{k} X_{t+i} = \delta_k + \gamma_k \sum_{i=1}^{k} X_{t-k+i} + \varepsilon_t. \tag{9}$$

The slope coefficient $\gamma_k$, which we further use as a test statistic and denote by $AC1(k)$, is the first-order autocorrelation of $k$-month excess returns:

$$\begin{aligned}\gamma_k = AC1(k) &= Cor\left(\sum_{i=1}^{k} X_{t+i}, \sum_{i=1}^{k} X_{t-k+i}\right)\\ &= Cor(MOM_{t+k}(k), MOM_t(k)),\end{aligned} \tag{10}$$

where $Cor(\cdot,\cdot)$ denotes the correlation coefficient. The test statistic $AC1(k)$ is nothing else than the correlation between $MOM_t(k)$ and $MOM_{t+k}(k)$ trading indicators.

We wish to estimate the first-order autocorrelation over periods $k \in [1, 14]$ months. The fundamental problem with these estimations is that we have only a relatively small number of non-overlapping intervals of 14 months. Therefore, as in Fama and French (1988), to increase the number of observations of $k$-month excess returns, we employ overlapping intervals of $k$ months. However, in contrast to Fama and French (1988) who estimate the first-order autocorrelations using a standard OLS regression, the second key element of our methodology is to estimate the autocorrelations using a highly robust covariance (and correlation) estimation method suggested by Rousseeuw (1984) and further developed by Rousseeuw (1985).

The justification of our approach is as follows. It is well-known the stock return distribution is leptokurtic and skewed to the left. Put differently, as compared to the normal distribution, the stock returns contain lots of "outliers", and the negative outliers are larger (in absolute value) than the positive outliers. Since we expect that the slope coefficient $\gamma_k$ is rather small, its estimation and statistical significance are heavily influenced by outliers, especially negative ones. To decrease the role of outliers, one possibility is to run a robust linear regression. Typically, in robust regressions, one reduces the weight of outliers. The problem is that the outliers are still present in the estimation and, thus, they can nevertheless distort the estimation and decrease the test power. In contrast, the estimation method suggested by Rousseeuw (1984) uses the minimum covariance determinant (MCD) method, which is highly resistant to outliers. In brief, this method only uses observations that lie within 97.5% volume of a Gaussian confidence ellipsoid. The problem is that the exact MCD method is extremely time-consuming. Our study relies on the FAST-MCD method developed by Rousseeuw and Driessen (1999).

After estimating $AC1(k)$, we want to ensure that the estimated first-order autocorrelations are statistically significantly positive. To this end, we test the following null hypothesis:

$$H_0 : AC1(k) = 0 \text{ versus } H_A : AC1(k) > 0.$$

The motivation for this test is as follows. If the excess returns are independent and identically distributed, then $AC1(k) = 0$ for all $k$. In other words, absent persistence in the excess returns, there is no correlation between two consecutive and non-overlapping $k$-month excess returns. The null hypothesis of independence is rejected in favor of persistence in the excess returns if $AC1(k)$ is significantly above zero. Testing this null hypothesis can be carried out using the randomization method.

To be more specific, we randomize the original excess return series $N = 1{,}000$ times, each time obtaining a new estimate for $AC1_j^*(k)$, where $j$ is an index for the randomization round (so $j = 1$ for the round 1).[5] In the end, the collection of all estimates for $AC1_j^*(k)$ constitutes

[5] Asterisk is used to indicate that each of these estimates is calculated on a randomized sample.

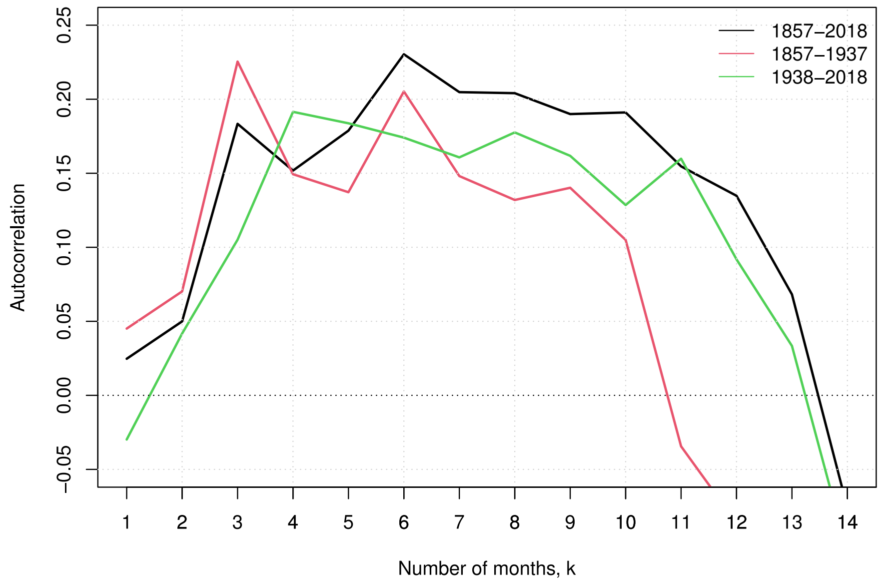


**Fig. 1.** For the whole sample and both halves of the sample, this figure plots the shape of the empirically estimated first-order autocorrelation function of $k$-month excess returns, $AC1(k)$.

the probability distribution of $AC1(k)$ under the null hypothesis. The test's $p$-value is estimated as the proportion of test statistics $AC1_j^*(k)$ that are at least as extreme as the observed statistic $AC1(k)$. Formally, the $p$-value is computed as:

$$p(k) = \frac{\sum_{j=1}^{N} \mathbf{1}_{AC1_j^*(k) \geq AC1(k)}}{N},$$

where $\mathbf{1}_{AC1(k)_j^* \geq AC1(k)}$ denotes the indicator function that takes one if $AC1(k)_j^* \geq AC1(k)$ and zero otherwise.

It is known that estimates obtained using overlapping blocks of data are biased (see Fama & French, 1988; Kim et al., 1991, and Nelson & Kim, 1993 among others). Therefore, we estimate the bias in $AC1(k)$ and conduct the bias correction. The bias is estimated as the mean value of the sampling distribution for $AC1(k)$ obtained by assuming the null hypothesis is true. The bias is corrected by subtracting the estimated mean from the observed $AC1(k)$.

For the whole sample and both halves of the sample, Fig. 1 plots the estimated first-order autocorrelation function of $k$-month excess returns on the S&P Composite stock price index. Table 2 reports the estimated first-order autocorrelations and the results of testing the null hypothesis using the whole sample of data and both halves of the sample. The results reported in this table exhibit clear evidence against independence in favor of a short-term persistence in the excess returns. Specifically, in the whole sample, the first-order autocorrelation is statistically significantly above zero at the 5% level for periods of $k \in [3, 11]$ months. In the first (second) half of the sample, the first-order autocorrelation is statistically significantly above zero at the 5% level for periods of $k \in [3, 7]$ ($k \in [4, 8]$) months. For the sake of illustration, Fig. 2 plots the estimated first-order autocorrelation function of $k$-month excess returns, $AC1(k)$, using the data for the whole sample period. The shaded area indicates the 90% confidence interval for the estimated autocorrelation under the null hypothesis that the excess returns are independent and identically distributed. With this choice, if the estimated value of $AC1(k)$ lies above the 90% confidence interval, this value is statistically significantly positive at the 5% level in a one-tailed test.

### 5.2. Evaluation of model parameters

The $AR(p)$ process for the excess returns is tractable and allows us to compute the theoretical shape of the test statistic $AC1(k)$ given a pair of parameters $p$ and $\phi$.

**Table 2**
First-order autocorrelation of $k$-month excess returns on the S&P Composite stock price index. The estimates are corrected for bias under the null hypothesis. Bold values indicate statistical significance at the 5% level.

| Period $k$ | 1857–2018 | | 1857–1937 | | 1938–2018 | |
|---|---|---|---|---|---|---|
| | Estimate | $P$-value | Estimate | $P$-value | Estimate | $P$-value |
| 1 | 0.025 | 0.279 | 0.045 | 0.199 | −0.030 | 0.710 |
| 2 | 0.050 | 0.113 | 0.070 | 0.105 | 0.042 | 0.231 |
| 3 | **0.183** | 0.000 | **0.225** | 0.000 | 0.105 | 0.052 |
| 4 | **0.152** | 0.002 | **0.149** | 0.022 | **0.191** | 0.004 |
| 5 | **0.179** | 0.000 | **0.137** | 0.037 | **0.184** | 0.009 |
| 6 | **0.230** | 0.000 | **0.205** | 0.006 | **0.174** | 0.018 |
| 7 | **0.205** | 0.000 | **0.148** | 0.042 | **0.161** | 0.037 |
| 8 | **0.204** | 0.002 | 0.132 | 0.086 | **0.178** | 0.032 |
| 9 | **0.190** | 0.004 | 0.140 | 0.087 | 0.162 | 0.056 |
| 10 | **0.191** | 0.008 | 0.105 | 0.170 | 0.129 | 0.122 |
| 11 | **0.155** | 0.024 | −0.034 | 0.614 | 0.160 | 0.082 |
| 12 | 0.135 | 0.055 | −0.086 | 0.767 | 0.092 | 0.203 |
| 13 | 0.068 | 0.213 | −0.128 | 0.852 | 0.033 | 0.396 |
| 14 | −0.076 | 0.808 | −0.138 | 0.866 | −0.106 | 0.799 |

**Proposition 1.** *If $X_t$ is a wide-sense stationary stochastic process, then*

$$AC1(k) = \frac{\mathbf{1}_k' \boldsymbol{Q}_{k,k} \mathbf{1}_k}{\mathbf{1}_k' \boldsymbol{P}_{k,k} \mathbf{1}_k}, \tag{11}$$

*where $\mathbf{1}_k$ is the $k \times 1$ vector of ones, $\boldsymbol{P}_{k,k}$ and $\boldsymbol{Q}_{k,k}$ are the $k \times k$ matrices given by*

$$\boldsymbol{P}_{k,k} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix},$$
$$\boldsymbol{Q}_{k,k} = \begin{bmatrix} \rho_k & \rho_{k+1} & \rho_{k+2} & \cdots & \rho_{2k-1} \\ \rho_{k-1} & \rho_k & \rho_{k+1} & \cdots & \rho_{2k-2} \\ \rho_{k-2} & \rho_{k-1} & \rho_k & \cdots & \rho_{2k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_1 & \rho_2 & \rho_3 & \cdots & \rho_k \end{bmatrix}, \tag{12}$$

*where $\rho_i$ is the autocorrelation of order $i$ of the process for $X_t$.*

The proof is given in Appendix.

Fig. 3 plots the theoretical shapes of the $AC1(k)$ where each model parameter can take only two values: $p \in \{6, 8\}$ and $\phi \in \{0.03, 0.04\}$.

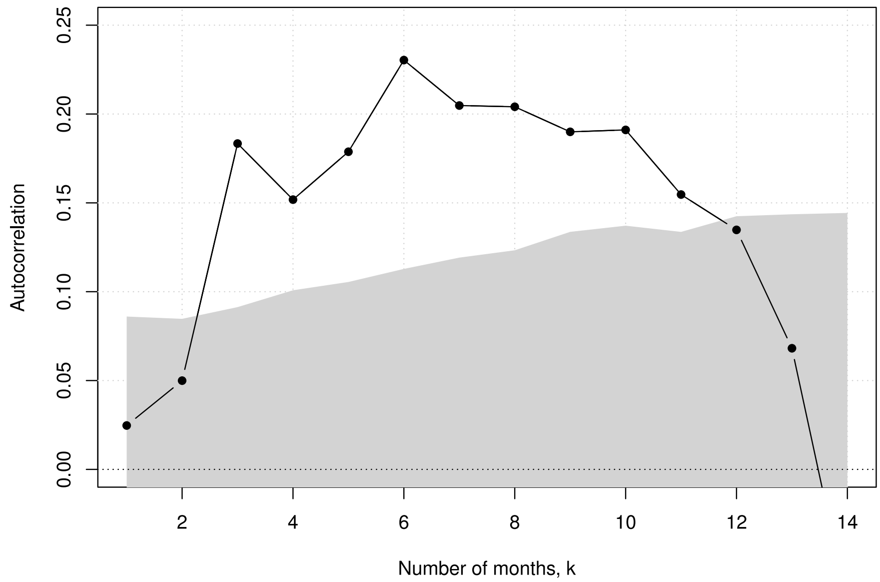


**Fig. 2.** Using the data for the whole sample, this figure plots the shape of the empirically estimated first-order autocorrelation function of $k$-month excess returns, $AC1(k)$, on the S&P Composite stock price index. The shaded area indicates the 90% confidence interval for the estimated autocorrelation under the null hypothesis that the excess returns are independent and identically distributed.

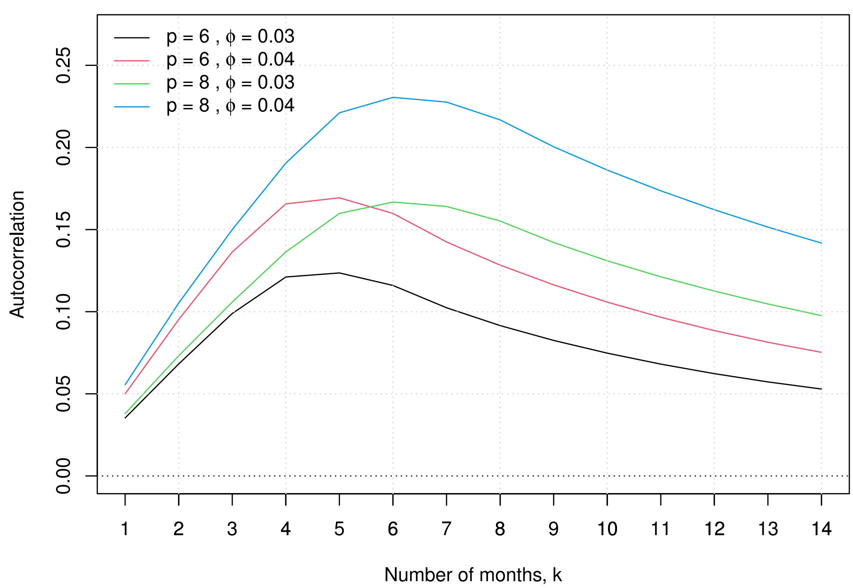


**Fig. 3.** The theoretical shapes of the $AC1(k)$ where each model parameter can take only two values: $p \in \{6, 8\}$ and $\phi \in \{0.03, 0.04\}$.

Our first observation is that each curve in Fig. 3 is a positively skewed bell-shaped curve with an evident top. Our second observation is that the top's location is governed by the order of the $AR(p)$ process, that is, the parameter $p$.[6] Our third and final observation is that, for a fixed value of $p$, the value of $AC1(k)$ for any $k$ increases as the parameter $\phi$ increases. Altogether, the curves in the figure imply that the shape of $AC1(k)$ is uniquely determined by the parameters $p$ and $\phi$.[7]

The observations presented in the preceding paragraph suggest an indirect approach to the joint evaluation of the parameters $p$ and $\phi$ of the $AR(p)$ process for excess returns. Specifically, the idea is to evaluate $p$ and $\phi$ by fitting the theoretical shape of $AC1(k)$ to the empirically estimated shape. For this purpose, we numerically solve the following problem:

$$\min_{p,\phi} \sum_{k=k_{min}}^{k=k_{max}} \left| AC1(k, p, \phi) - AC1_{EMP}(k) \right|, \tag{13}$$

where $AC1_{EMP}(k)$ is the empirically estimated first-order autocorrelation function of $k$-month excess returns and $AC1(k, p, \phi)$ is the theoretical autocorrelation function of $k$-month excess returns for some specific values of $p$ and $\phi$. That is, the parameters $p$ and $\phi$ are evaluated by a numerical procedure that finds the pair $\{p, \phi\}$ which minimizes the sum of the absolute deviations between the empirically observed and the model-implied values of the first-order autocorrelation function of $k$-month excess returns.

The main discrepancy between the model-implied autocorrelation function depicted in Fig. 3 and the sample autocorrelation function

[6] Our numerical experiments suggest that for a small $p \in [1, 5]$ the top is located at $k = p$. When the order $p$ exceeds 5, the top is located at $k < p$.

[7] Using an approximate solution for the value of $AC1(k)$, we prove that $\partial AC1(k)/\partial p > 0$ and $\partial AC1(k)/\partial \phi > 0$. These proofs are available from the authors upon request.

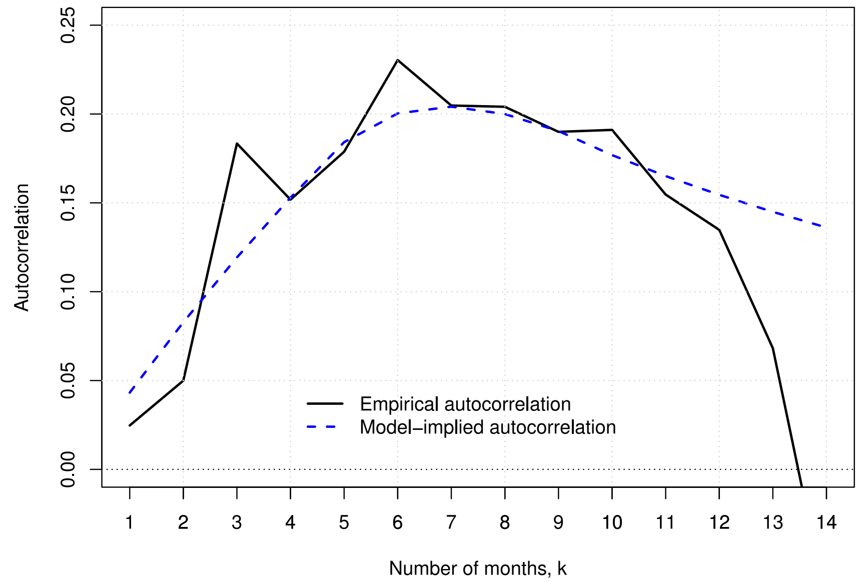


**Fig. 4.** Using the data for the whole sample, this figure plots the shape of the empirically estimated first-order autocorrelation function of $k$-month excess returns and the shape of the theoretical first-order autocorrelation function with the parameters $p = 9$ and $\phi = 0.0321$ that produce the best fit to the empirical data.

**Table 3**
Estimated best-fit parameters $p$ and $\phi$ of the $AR(p)$ process for the excess returns on the S&P Composite stock price index. $\kappa = p\phi$ measures the persistence of the process for excess returns.

| Period | Order $p$ | Coefficient $\phi$ | Persistence $\kappa$ |
|---|---|---|---|
| 1857–2018 | 9 | 0.0321 | 0.2889 |
| 1857–1937 | 7 | 0.0325 | 0.2275 |
| 1938–2018 | 8 | 0.0312 | 0.2496 |

plotted in Fig. 1 concerns the behavior for $k$ greater than about 12 months. Specifically, whereas the model-implied autocorrelation function decreases gradually toward zero as $k$ increases, the sample autocorrelation function eventually becomes negative. This behavior of the sample first-order autocorrelation function is a direct consequence of the well-known empirical fact that "the time series momentum or 'trend' effect persists for about a year and then partially reverses over longer horizons" (Moskowitz et al., 2012, page 228). That is, there is both a short-term momentum and a subsequent medium-term mean reversion in excess returns. Motivated by this observation, our model identification procedure finds the parameters $p$ and $\phi$ that produce the best fit over the range $k_{min} = 1$ month to $k_{max} = 12$ months. The motivation for limiting the maximum value for $k$ to 12 months is to confine our attention solely to the short-term persistence effect and overlook the impact of the subsequent medium-term reversion to the mean.

Table 3 documents the estimated best-fit parameters $p$ and $\phi$ of the $AR(p)$ process for the excess returns on the S&P Composite stock price index for the whole sample and both halves of the sample. The results reported in this table suggest that over the first (second) half of the sample, the excess returns followed the $AR(p)$ process with $p = 7$ and $\phi = 0.0325$ ($p = 8$ and $\phi = 0.0312$). Over the whole sample, the estimated parameters are $p = 9$ and $\phi = 0.0321$. For the sake of illustration, using the data for the whole sample, Fig. 4 plots the shape of the empirically estimated first-order autocorrelation function of $k$-month excess returns and the shape of the theoretical first-order autocorrelation function with the parameters $p = 9$ and $\phi = 0.0321$ that produce the best fit to the empirical data.

It is important to emphasize that the results of our estimations imply that, when the whole sample data are used, the value of $\phi$ is about the same largeness as that of the standard error of estimation of $\phi$ using the OLS regression model, see Table 1. This observation explains why we generally do not see statistically significant $\phi_i$ coefficients when we rely on the OLS methodology. This is because, to detect statistical significance at the 5% level, a coefficient in an OLS model must be more than twice as large as the standard error of estimation of this coefficient. To reduce the standard error by half, one needs to quadruple the sample size. Therefore, our ballpark estimate is that, to detect statistically significant autoregressive coefficient using the OLS methodology, the sample size must cover more than $162 \times 4 \approx 650$ years.[8] Thus, the required sample size is way beyond the currently available sample sizes.

It is also instructive to compare the results on the estimated return persistence calculated as the sum of the autoregressive coefficients from the OLS regression and the return persistence estimated by our model identification procedure. Whereas the two results agree well when the data for the first half of the sample is used, they differ remarkably when the data for the whole sample or the second half are utilized. Specifically, the results in Table 1 indicate a weak return persistence in the first half of the sample and a substantially weaker return persistence in the whole sample and the second half of it. By contrast, the results in Table 3 argue that in the entire sample period and the second half, the return persistence was a bit stronger than that in the first half of the sample.

## 6. Theoretical analysis of time series momentum strategy

Armed with a reasonably accurate knowledge of the momentum generating process, we continue this paper with a theoretical analysis of the TSMOM strategy. In particular, in the subsequent section, we offer analytical results on the TSMOM strategy's risk, return, and performance. We demonstrate that our analytical results agree well with the empirical data. Subsequently, we provide a comparative static analysis of the profitability of the TSMOM strategy. Finally, by relying on a simulation analysis, we explore how the evidence of the TSMOM strategy's superior performance depends on the investment horizon and evaluate the power of statistical tests.

[8] This number is verified using a parametric bootstrap simulation method.

### 6.1. Risk, return, and performance of TSMOM strategy

To start with, this section presents analytical solutions for the mean returns, the variance of returns, the Sharpe ratio, and the CAPM alpha and beta of the TSMOM strategy. For simplicity, we assume that the risk-free interest rate, $r_f$, is constant and the joint distribution of the market returns $r_t$ and the $MOM_{t-1}(n)$ trading indicator follows a bivariate normal distribution

$$\begin{bmatrix} r_t \\ MOM_{t-1}(n) \end{bmatrix} = \mathcal{N}\left( \begin{bmatrix} \mu \\ m \end{bmatrix}, \begin{bmatrix} \sigma^2 & \varrho_n \sigma \upsilon \\ \varrho_n \sigma \upsilon & \upsilon^2 \end{bmatrix} \right), \tag{14}$$

where $\mu$ and $\sigma^2$ are the mean and variance of $r_t$, $m$ and $\upsilon^2$ are the mean and variance of $MOM_{t-1}(n)$, and $\varrho_n$ is the correlation coefficient between $r_t$ and $MOM_{t-1}(n)$. The mean and variance of $MOM_{t-1}(n)$ are given by

$$m = n\mu_x, \quad \upsilon^2 = \mathbf{1}_n' \boldsymbol{P}_{n,n} \mathbf{1}_n \sigma_x^2, \tag{15}$$

where $\mu_x$ and $\sigma_x^2$ are given by Eq. (6), $\mathbf{1}_n$ is the $n \times 1$ vector of ones, and matrix $\boldsymbol{P}_{n,n}$ is the $n \times n$ matrix given by (12). The correlation coefficient is computed as (see Zakamulin & Giner, 2020)

$$\varrho_n = \frac{\mathbf{1}_n' \boldsymbol{P}_{n,p} \boldsymbol{\phi}_p}{\sqrt{\mathbf{1}_n' \boldsymbol{P}_{n,n} \mathbf{1}_n}}, \tag{16}$$

where $\boldsymbol{\phi}_p' = [\phi, \phi, \ldots, \phi]$ is the $p \times 1$ vector of autoregressive coefficients of $X_t$ and $\boldsymbol{P}_{n,p}$ is the $n \times p$ matrix given by

$$\boldsymbol{P}_{n,p} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{p-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \cdots & \rho_{|p-n|} \end{bmatrix}.$$

**Proposition 2.** *The mean return, variance of returns, beta, and alpha of the long-only TSMOM strategy are given by*

$$E[R_t^{LO}] = (\mu - r_f)\Phi(-d) + r_f + g, \tag{17}$$

$$Var[R_t^{LO}] = (\mu^2 + \sigma^2)\Phi(-d) + g(2\mu + \sigma\varrho_n d) + r_f^2 \Phi(d) - E[R_t^{LO}]^2, \tag{18}$$

$$\beta^{LO} = \Phi(-d) + \frac{g(\mu - r_f + \sigma\varrho_n d)}{\sigma^2}, \tag{19}$$

$$\alpha^{LO} = g\left(1 - \frac{(\mu - r_f)(\mu - r_f + \sigma\varrho_n d)}{\sigma^2}\right), \tag{20}$$

*where*

$$d = -\frac{m}{\upsilon}, \quad g = \sigma\varrho_n \varphi(d),$$

*and $\varphi(.)$ and $\Phi(.)$ denote the probability density and the cumulative probability distribution function, respectively, of the standard normal random variable*

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad \Phi(d) = \int_{-\infty}^{d} \varphi(z) dz.$$

*The mean return, variance of returns, beta, and alpha of the long–short TSMOM strategy are given by*

$$E[R_t^{LS}] = (2\Phi(-d) - 1)\mu + 2(g + \Phi(d) r_f), \tag{21}$$

$$Var[R_t^{LS}] = (\mu^2 + \sigma^2) + 4r_f(g - (\mu - r_f)\Phi(d)) - E[R_t^{LS}]^2, \tag{22}$$

$$\beta^{LS} = \Phi(-d) - \Phi(d) + \frac{2g(\mu - r_f + \sigma\varrho_n d)}{\sigma^2}, \tag{23}$$

$$\alpha^{LS} = 2\, g\left(1 - \frac{(\mu - r_f)(\mu - r_f + \sigma\varrho_n d)}{\sigma^2}\right). \tag{24}$$

The proof is given in Appendix.

**Remark 1.** We remind the reader that the returns to the long-only strategy, $R_t^{LO}$, are given by Eq. (1), whereas the returns to the long–short strategy, $R_t^{LS}$, are specified by Eq. (2).

**Remark 2.** The Sharpe ratio of the buy-and-hold strategy, the long-only TSMOM strategy, and long–short TSMOM strategy are given by

$$SR_{BH} = \frac{\mu - r_f}{\sigma}, \quad SR_{LO} = \frac{E[R_t^{LO}] - r_f}{\sqrt{Var\left[R_t^{LO}\right]}}, \quad SR_{LS} = \frac{E[R_t^{LS}] - r_f}{\sqrt{Var\left[R_t^{LS}\right]}}, \tag{25}$$

where $SR_{BH}$, $SR_{LO}$, and $SR_{LS}$ denote the Sharpe ratio of the buy-and-hold strategy, the long-only TSMOM strategy, and the long–short TSMOM strategy, respectively.

**Remark 3.** Note that $\alpha^{LS} = 2\alpha^{LO}$. In words, the alpha of the long–short strategy is twice as large as the alpha of the long-only strategy.

First of all, we check how well the results of our theoretical modeling agree with the empirical results. The empirical results are obtained by simulating the historical returns to the TSMOM strategies over the whole sample and each half of the sample. We assume that the optimal window size in the $MOM(n)$ trading indicator is known in these simulations. We remind the reader that the optimal window size equals the number of lags in the $AR(p)$ process for the excess returns. Thus, when the whole sample data are used, the optimal window size is $n = 9$. The optimal window sizes for the first and second half of the sample are 7 and 8, respectively.

The theoretical results are computed as follows. First, using a particular sample period, we estimate the monthly parameters $\mu$, $\sigma$, and $r_f$. Then, using the analytical solutions, we compute the mean return, standard deviation, Sharpe ratio, and CAPM alpha and beta of the buy-and-hold strategy, the long-only TSMOM strategy, and the long–short TSMOM strategy.[9] We assume that the excess market returns $r_t - r_f$ follow the $AR(p)$ process with the estimated best-fit parameters reported in Table 3. As in the historical simulations, we assume that the optimal window size in the TSMOM rule is known, that is, $n = p$. In this case, the expression for the correlation coefficient specified by Eq. (16) reduces to

$$\varrho_p = \frac{(1 + (p-1)\rho)p\,\phi}{\sqrt{p(1 + (p-1)\rho)}}, \tag{26}$$

where $\rho$ is given by Eq. (5). Assuming that $p - 1 \approx p$, a useful approximate solution for this correlation coefficient is given by

$$\varrho_p \approx \frac{\kappa}{\sqrt{p(1-\kappa)}}. \tag{27}$$

Note that this correlation coefficient is the main driving factor determining the TSMOM strategy's success. This is because the performance of any trend-following strategy increases as the correlation between the trading indicator and the future return increases.

Both the Sharpe ratio and CAPM alpha measure the performance of an investment strategy. To check whether the TSMOM strategy outperforms the buy-and-hold strategy, we conduct two statistical tests. The first test uses the Sharpe ratio as a performance measure. The null hypothesis in this test is

$$H_0 : SR_{MOM} = SR_{BH} \text{ versus } H_A : SR_{MOM} > SR_{BH}, \tag{28}$$

where $SR_{MOM}$ and $SR_{BH}$ are the Sharpe ratios of the TSMOM and buy-and-hold strategy, respectively. In words, we test for the difference between the Sharpe ratios of two alternative strategies. To this end, we

[9] The risk-free rate of return is estimated as the mean risk-free rate of return over the sample. Further note that under our assumptions $\mu_x = \mu - r_f$ and $\sigma_x = \sigma$.

**Table 4**
The descriptive statistics of the buy-and-hold strategy, the long-only TSMOM strategy, and the long–short TSMOM strategy. The mean return, standard deviation, and alpha are annualized and reported in percentages. The Sharpe ratios are also annualized. For each TSMOM strategy, the table reports the empirically estimated descriptive statistics and the theoretically computed (model-implied) descriptive statistics. The $p$-values are reported in round brackets. The $p$-value of alpha is computed using Newey–West errors with 12 lags. Bold indicates statistical significance at the 5% level.

| | Buy and hold | Long-only | | Long–short | |
|---|---|---|---|---|---|
| | | Empirical | Theoretical | Empirical | Theoretical |
| **Full sample: 1857–2018** | | | | | |
| Mean return | 10.27 | 10.06 | 10.34 | 9.64 | 10.40 |
| Std. deviation | 17.40 | 12.88 | 13.61 | 17.46 | 17.40 |
| Sharpe ratio | 0.37 | **0.48** | 0.48 | 0.33 | 0.38 |
| | | (0.04) | | (0.68) | |
| Beta | 1.00 | 0.55 | 0.61 | 0.10 | 0.22 |
| Alpha | 0.00 | **2.59** | 2.56 | **5.18** | 5.13 |
| | | (0.00) | | (0.00) | |
| **First half: 1857–1937** | | | | | |
| Mean return | 8.99 | 9.19 | 9.36 | 9.12 | 9.73 |
| Std. deviation | 19.25 | 12.87 | 14.56 | 19.34 | 19.23 |
| Sharpe ratio | 0.26 | 0.40 | 0.37 | 0.27 | 0.30 |
| | | (0.08) | | (0.52) | |
| Beta | 1.00 | 0.45 | 0.57 | −0.11 | 0.14 |
| Alpha | 0.00 | **2.85** | 2.50 | **5.69** | 5.00 |
| | | (0.01) | | (0.01) | |
| **Second half: 1938–2018** | | | | | |
| Mean return | 11.56 | 10.37 | 10.72 | 9.41 | 9.89 |
| Std. deviation | 15.34 | 11.64 | 12.38 | 14.91 | 15.40 |
| Sharpe ratio | 0.52 | 0.57 | 0.57 | 0.39 | 0.41 |
| | | (0.22) | | (0.83) | |
| Beta | 1.00 | 0.62 | 0.65 | 0.23 | 0.30 |
| Alpha | 0.00 | **1.97** | 1.94 | **3.95** | 3.88 |
| | | (0.02) | | (0.02) | |

apply the Jobson and Korkie (1981) test corrected by Memmel (2003). The test statistic in the Jobson and Korkie (1981) test is

$$z = \frac{SR_{MOM} - SR_{BH}}{\sqrt{\left[2(1-\rho_b) + \frac{1}{2}(SR_{MOM}^2 + SR_{BH}^2 - 2\rho_b^2 SR_{MOM} SR_{BH})\right]/T}}, \quad (29)$$

where $SR_{MOM}$, $SR_{BH}$, and $\rho_b$ are the estimated Sharpe ratios and correlation coefficient between the returns of the two strategies over a sample size of $T$ months. Under the null hypothesis, $z$ is asymptotically standard normal.

The second test uses the CAPM alpha as a performance measure. The null hypothesis in this test is

$$H_0 : \alpha_{MOM} = 0 \text{ versus } H_A : \alpha_{MOM} > 0, \quad (30)$$

where $\alpha_{MOM}$ is the alpha of the TSMOM strategy. The test statistic in this test

$$z = \frac{\alpha_{MOM}}{se\left(\alpha_{MOM}\right)}, \quad (31)$$

where $se\left(\alpha_{MOM}\right)$ is the standard error of estimation of $\alpha_{MOM}$. When the number of observations is sufficiently large, $z$ is asymptotically standard normal. The alpha is estimated using the CAPM regression. Because the regression residuals are autocorrelated, the standard error of estimation of alpha is computed using the Newey–West estimator with 12 lags. Finally, note that both our tests are one-sided.

Table 4 reports the descriptive statistics of the buy-and-hold strategy using the whole sample of data and both halves of the sample. The descriptive statistics include the mean return, standard deviation, Sharpe ratio, and CAPM alpha and beta. Besides, for both the long-only and long–short TSMOM strategies, this table reports the empirically estimated descriptive statistics and the theoretically computed (model-implied) descriptive statistics. The empirical performance of each TSMOM strategy is tested against the performance of the corresponding buy-and-hold strategy. The p-values are reported in round brackets.

Our first observation concerns the agreement between the empirically estimated and model-implied descriptive statistics of the TSMOM strategies. As the numbers in Table 4 suggest, most often, the analytically calculated descriptive statistics slightly overestimate the empirically estimated descriptive statistics. Yet, the general picture is that the model-implied results agree well with the empirical data.

Our second observation is related to the profitability of the TSMOM strategies. Judging by the Sharpe ratio, the long–short TSMOM strategy is not superior to the buy-and-hold strategy, while the long-only TSMOM strategy outperforms the buy-and-hold strategy. In particular, the evaluated Sharpe ratios of the long-only TSMOM strategy are economically significantly larger than those of the buy-and-hold strategy over the whole sample and both halves of the sample. Yet, the statistical evidence of outperformance is obtained only when the data for the whole sample are used. Judging by the alpha, both TSMOM strategies are statistically significantly superior to the buy-and-hold strategy. As our theoretical model correctly predicts, the empirically estimated alpha of the long–short TSMOM strategy tends to be double as large as the alpha of the long-only TSMOM strategy.

It is worth emphasizing that the choice of performance measure plays a crucial role in determining which TSMOM strategy is best. When the Sharpe ratio measures the performance, the long-only TSMOM strategy is best. By contrast, the long–short TSMOM strategy is best according to the CAPM alpha. The difference in conclusions may seem surprising but can be explained by how each performance criterion measures the risk. While the Sharpe ratio is a performance measure that considers the total risk (total standard deviation), the CAPM alpha is a performance measure that accounts for the systematic risk (beta) only. As the numbers in Table 4 reveal, the standard deviation of returns of the long–short TSMOM strategy is comparable to that of the buy-and-hold strategy and larger than the standard deviation of the long-only TSMOM strategy. Conversely, the beta of the long–short TSMOM strategy is substantially below the beta of the long-only TSMOM strategy, which, in turn, has a beta that is notably below the beta of the buy-and-hold strategy. Therefore, while the long-only TSMOM strategy has the lowest total risk, the long–short TSMOM strategy has the lowest systematic risk.

### 6.2. Comparative static analysis of TSMOM strategy's performance

The results presented in the preceding section advocate that our theoretical model is in good agreement with empirical data. Consequently, the comparative static analysis of our theoretical model can significantly contribute to our fundamental understanding of the TSMOM strategy. Given the space limitation, it is impossible to present a full-blown analysis. With this in mind, we focus only on studying how the TSMOM strategy's profitability depends on the parameters of the $AR(p)$ process for the excess returns and the number of lags in the $MOM(n)$ trading indicator.

We start our analysis by investigating how the performance of the long-only and long–short TSMOM strategy depends on the value of $\phi$ in the $AR(p)$ process for the excess returns. The model parameters $\mu$, $\sigma$, and $r_f$ are evaluated using the data for the whole sample. We hold the value of $p$ fixed at 9 (the estimated value) and vary the value of $\phi \in [0, 0.07]$. Panel A (B) in Fig. 5 plots the theoretical Sharpe ratio (alpha) of the buy-and-hold strategy, the long-only TSMOM strategy, and the long–short TSMOM strategy versus the value of $\phi$ in the $AR(9)$ process for the excess returns. In both strategies, the trading signal is generated using the $MOM(9)$ trading indicator.

To begin with, we analyze how the superiority of the TSMOM strategy depends on $\phi$ when the performance is measured by the Sharpe ratio. Based on the visual observations of the curves in Fig. 5 Panel A, we can make the following observations. First, when the excess market

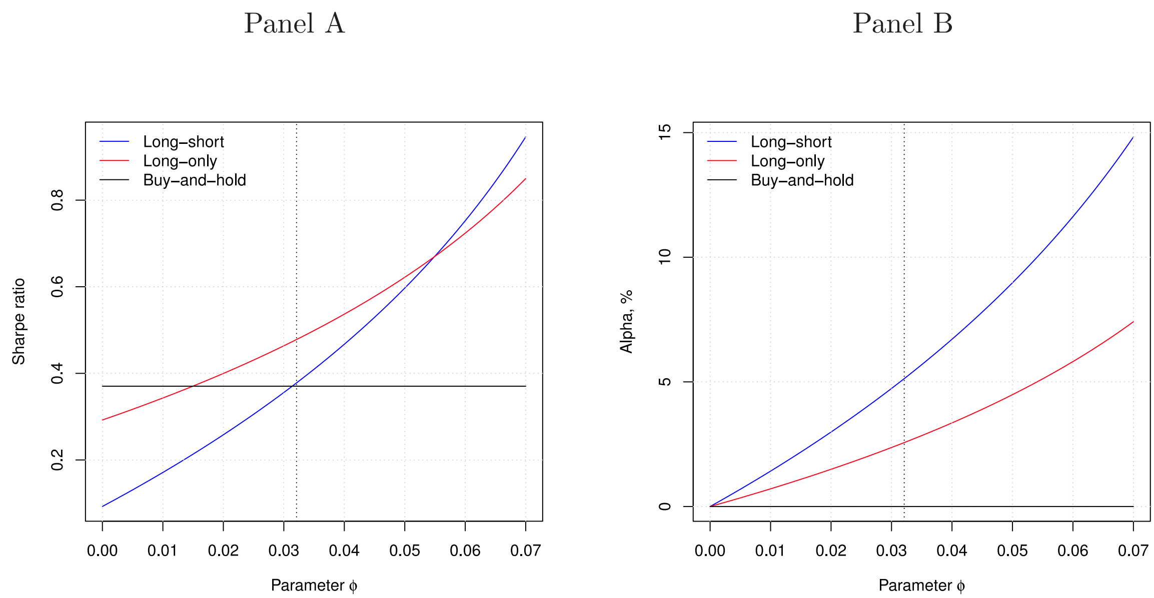


**Fig. 5.** Panel A (B) plots the theoretical Sharpe ratio (alpha) of the buy-and-hold strategy, the long-only TSMOM strategy, and the long–short TSMOM strategy versus the value of $\phi$ in the $AR(9)$ process for the excess returns. In both strategies, the trading signal is generated using the $MOM(9)$ trading indicator. The parameters $\mu$, $\sigma$, and $r_f$ are evaluated using the data for the whole sample. The vertical dotted line shows the location of the estimated value of $\phi = 0.0321$.

returns follow a random walk, $\phi = 0$, the correlation coefficient $\varrho_p = 0$, and, consequently, the trading signals do not have any predictive ability. In this case, both TSMOM strategies underperform the buy-and-hold strategy. It is also worth mentioning that, in this case, the Sharpe ratio of the long–short TSMOM strategy is notably below that of the long-only TSMOM strategy.

Second, when the autoregressive coefficient $\phi$ increases, the return persistence $\kappa = p\phi$ and the correlation $\varrho_p$ increase (see Eq. (27)). As a result, the Sharpe ratio of both TSMOM strategies increases. Third, the long-only TSMOM strategy starts outperforming the buy-and-hold strategy when the value of the autoregressive parameter $\phi$ exceeds 0.0149. This value is twice as low as the estimated value. Fourth, the value of $\phi$ must surpass 0.0549 to make the long–short TSMOM strategy superior to the long-only TSMOM strategy. That is, the value of $\phi$ must be more than twice as large as the estimated value to make the long–short TSMOM strategy worthwhile according to the Sharpe ratio criterion. Finally, when the performance is measured by alpha and $\phi > 0$, the long–short TSMOM strategy always outperforms the long-only TSMOM strategy, which, in turn, outperforms the buy-and-hold strategy.

Our next analysis is motivated by the following observation concerning the superiority of the long-only TSMOM strategy. Both the empirical and model-implied Sharpe ratios of the long-only TSMOM strategy are higher than the Sharpe ratios of the buy-and-hold strategy. Note, however, that in the first half of the sample, the Sharpe ratio of the long-only TSMOM strategy is about 40% higher than that of the buy-and-hold strategy. In contrast, in the second half of the sample, the Sharpe ratio of the long-only TSMOM strategy is only about 12% higher than that of the buy-and-hold strategy. This difference seems surprising given that the estimated persistence of the return process is almost the same in both halves of the sample. A similar observation can be made regarding the alpha of a TSMOM strategy. In particular, the alpha is notably greater in the first half of the sample than in the second half.

Our theoretical model, coupled with the estimated parameters of the $AR(p)$ process for the excess returns, can explain the observed discrepancy between the relative performance of the TSMOM strategy versus the buy-and-hold strategy in two halves of the sample. Note that the correlation coefficient $\varrho_p$ increases as the return persistence $\kappa = p\phi$ increases. In contrast, for a fixed $\kappa$, the correlation coefficient increases when $p$ decreases (see Eq. (27)). As a result, two different pairs $\{p, \phi\}$ can yield notably different performances even though the product of each pair is the same.

For the sake of illustration, Panel A (B) in Fig. 6 plots the theoretical Sharpe ratio (alpha) of the long-only and long–short $MOM(p)$ strategy versus $p$ for a constant $\kappa = p\phi = 0.2889$.[10] The parameters $\mu$, $\sigma$, and $r_f$ are estimated using the data for the whole sample. This figure demonstrates that, for a fixed value of $\kappa = p\phi$, the Sharpe ratio and alpha of the TSMOM strategy increase exponentially when $p$ decreases. Note that the estimated value of $p$ for the first half of the sample equals 7, whereas the estimated value equals 8 for the second half of the sample. At the same time, the estimated return persistence is about the same in both halves of the sample. Consequently, our model correctly predicts that a TSMOM strategy outperforms the buy-and-hold strategy to a greater extent in the first half of the sample than in the second half.

The results reported above are obtained under the implicit assumption that we know the optimal number of lags in the TSMOM rule. In reality, the optimal number of lags is never known for sure. The final analysis in this section investigates the TSMOM strategy's robustness to the number of lags in the rule. Panel A (B) in Fig. 7 plots the theoretical Sharpe ratio (alpha) of the buy-and-hold strategy, the long-only TSMOM strategy, and the long–short TSMOM strategy versus the number of lags $n$. In both strategies, the trading signal is generated using the $MOM(n)$ trading indicator. As before, we assume that the excess returns follow the $AR(9)$ process and the parameters $\mu$, $\sigma$, and $r_f$ are estimated using the whole sample data.

The curves in Fig. 7 advocate that the TSMOM strategy's performance is rather stable with respect to changes in $n$ (at least when $n$ is not very much different from the optimal $n = p$). For example, the Sharpe ratio of the long-only TSMOM strategy with either $n = 6$ or $n = 14$ is only about 6% below the Sharpe ratio of the optimal strategy with $n = 9$ and still about 23% above the estimated Sharpe ratio of the buy-and-hold strategy. A similar conclusion can be drawn regarding the TSMOM strategy's robustness when the performance is measured by alpha. In this case, the alpha of both TSMOM strategies is only a fraction below its maximum value when $n$ varies from 6 to 14. Even when $n = 1$, both TSMOM strategies have an edge over the buy-and-hold strategy.

[10] For each value of $p$, the value of $\phi$ is computed as $\phi = \kappa/p$.

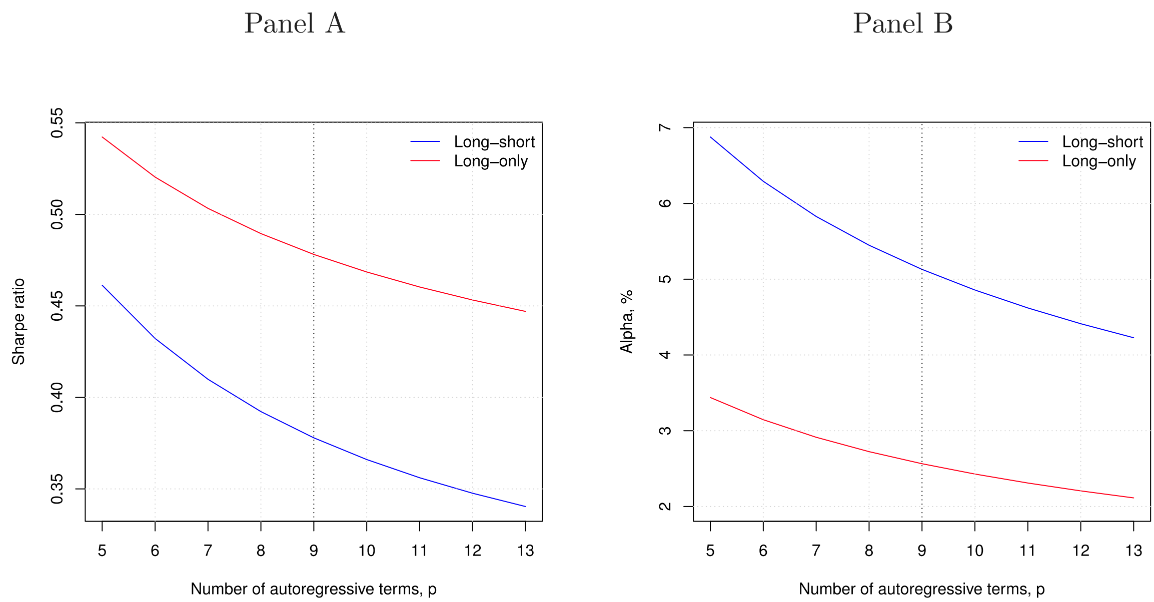


**Fig. 6.** Panel A (B) plots the theoretical Sharpe ratio (alpha) of the long-only and long–short $MOM(p)$ strategy versus the value of $p$ in the $AR(p)$ process for the excess returns. It is assumed that the return persistence $\kappa = p\phi = 0.2889$ is constant for all $p$. The parameters $\mu$, $\sigma$, and $r_f$ are evaluated using the data for the whole sample. The vertical dotted line shows the location of the estimated value of $p = 9$.

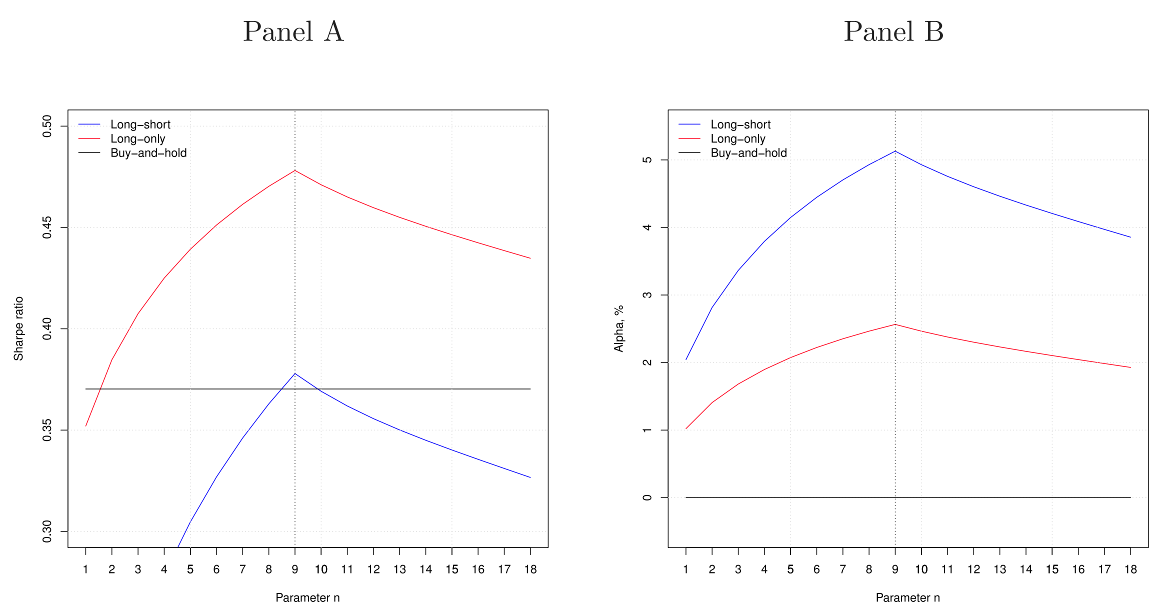


**Fig. 7.** Panel A (B) plots the theoretical Sharpe ratio (alpha) of the buy-and-hold strategy, the long-only TSMOM strategy, and the long–short TSMOM strategy versus the number of lags $n$. In both strategies, the trading signal is generated using the $MOM(n)$ trading indicator. It is assumed that the excess returns follow the $AR(9)$ process. The parameters $\mu$, $\sigma$, $r_f$, and $\phi$ are estimated using the data for the whole sample. The vertical dotted line shows the location of the optimal $n = 9$.

### 6.3. Investment Horizon and evidence of superior performance

The majority of the empirical studies on the profitability of trend-following strategies find that these strategies are profitable in the long run over periods ranging from 50 to 150 years. However, when the researchers use the most recent historical period (from 5 to 10 last years in the sample of data used in a study), they frequently report that the trend-following strategies are not profitable (see, for example, Hutchinson & O'Brien, 2014; Lee et al., 2001; Okunev & White, 2003; Olson, 2004, Siegel, 2002, Chapter 2, Sullivan et al., 1999, and Zakamulin, 2014). Typically, this result is attributed to increased market efficiency over time. Another issue in these studies is the lack of scientific evidence on the profitability of trend-following strategies. Specifically, even when the performance of a trend-following strategy is economically significantly higher than that of its passive counterpart, the researchers often fail to reject the null hypothesis that both strategies have equal performance (see, among others, Huang et al., 2020; Kim et al., 2016, and Zakamulin, 2017).

Our theoretical and empirical results reported above suggest that the performance of the TSMOM strategy is superior to that of the buy-and-hold strategy.[11] Strictly speaking, this says that the TSMOM strategy tends to outperform the buy-and-hold strategy over the long run. However, because of randomness, there is no guarantee that the TSMOM strategy outperforms the buy-and-hold strategy over the short

[11] Except the case of the long–short TSMOM strategy when the performance is measured by the Sharpe ratio.

run. In this section, we examine how the evidence of the TSMOM strategy's superior performance depends on the investment horizon length. The goal is to understand and explain why trend-following strategies often demonstrate inferior performance over the short run and lack of scientific evidence of superior performance even over the long run.

To achieve this goal, we rely on a simulation method. In particular, using the data for the whole sample, we estimate the monthly parameters $\mu$ and $\sigma$ of the returns on the S&P Composite index and the risk-free rate of returns $r_f$. The mean and variance of the excess market returns are computed as $\mu_x = \mu - r_f$ and $\sigma_x = \sigma$. We assume that the excess market returns follow the $AR(p)$ process with $p = 9$ and $\phi = 0.0321$. The parameters $c$ and $\sigma_\varepsilon$ of the $AR(p)$ process for the excess returns are computed using Eqs. (6). Subsequently, for a fixed horizon of $Y$ years, we simulate monthly returns to the buy-and-hold strategy and the TSMOM strategy $N = 100{,}000$ times. The trading signal in the TSMOM strategy is generated using the $MOM(9)$ rule. After each simulation round, we compute the Sharpe ratio of the long-only TSMOM strategy[12] and the buy-and-hold strategy, and the CAPM alpha of both the long-only and long–short TSMOM strategies. Subsequently, we conduct the tests of the null hypothesis of equal performance. The null hypothesis in the Sharpe ratio test is given by Eq. (28), and its test statistic is given by Eq. (29). In the alpha test, the null hypothesis and its test statistic are given by Eqs. (30) and (31), respectively. The standard error of estimation of alpha is computed using the Newey–West estimator with 12 lags.

After carrying out all simulations for a specific horizon $Y$, we compute two probabilities. The first probability is the probability that over the horizon of $Y$ years, the performance of the TSMOM strategy is better than that of the buy-and-hold strategy, $Prob(z > 0)$. To compute this probability, we count how many times the performance (Sharpe ratio or alpha) of the TSMOM strategy is greater than the performance of the buy-and-hold strategy. Denoting this value by $q_1$, the probability is computed as $Prob(z > 0) = q_1/N$. The second probability is the probability that over the horizon of $Y$ years, the $p$-value of the null hypothesis is lower than 5%. This probability is the probability of rejecting the null hypothesis at the 5% level. To compute this probability, we count how many times the value of the $z$-statistic exceeds the value of 1.64 (which is the critical value in a one-tailed test). Denoting this value by $q_2$, the probability is computed as $Prob(z > 1.64) = q_2/N$.

Table 5 reports the results of this simulation study. The information in the table is very insightful and allows us to draw important conclusions. We start with discussing the conclusions that can be reached when the performance is measured by the Sharpe ratio criterion. Over short- to medium-term horizons of up to 5 years, the probability that the long-only TSMOM strategy outperforms the buy-and-hold strategy is around 50%. That is, over these horizons, the long-only TSMOM strategy is equally like to underperform the buy-and-hold strategy as to outperform. The probability of outperformance increases as the investment horizon lengthens. For example, over horizons ranging from 5 to 10 years, the TSMOM strategy's probability of outperforming the buy-and-hold strategy is about 60%. However, even over 50 years, the probability of outperformance is about 80% only. Consequently, there is no guarantee that the trend-following strategy outperforms the buy-and-hold strategy even over a very long horizon.

The probability that the long-only TSMOM strategy statistically significantly outperforms the buy-and-hold strategy also increases as the investment horizon lengthens. Over the medium- and long-term horizons ranging from 5 to 50 years, the probability of observing statistically significant outperformance does not exceed 31%. Even over

**Table 5**
The results of the simulation study. For each horizon, the table reports two probabilities. Probability $Prob(z > 0)$ denotes the probability that the performance of a TSMOM strategy is greater than the performance of the buy-and-hold strategy. Probability $Prob(z > 1.64)$ denotes the probability of rejecting the null hypothesis of equal performances at the 5% level. The rejection of the null hypothesis means that the performance of a TSMOM strategy is statistically significantly higher than the performance of the buy-and-hold strategy.

Panel A: Simulation study results for the Sharpe ratio

| Probability | Horizon, years | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 5 | 10 | 20 | 30 | 40 | 50 | 75 | 100 | 200 | 300 | 500 |
| $z > 0$ | 0.50 | 0.59 | 0.69 | 0.75 | 0.79 | 0.82 | 0.88 | 0.92 | 0.98 | 0.99 | 1.00 |
| $z > 1.64$ | 0.08 | 0.12 | 0.18 | 0.22 | 0.26 | 0.31 | 0.39 | 0.48 | 0.71 | 0.85 | 0.96 |

Panel B: Simulation study results for the CAPM alpha

| Probability | Horizon, years | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 10 | 25 | 40 | 50 | 60 | 70 | 80 | 90 | 100 | 110 | 120 |
| $z > 0$ | 0.70 | 0.88 | 0.94 | 0.96 | 0.98 | 0.98 | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 |
| $z > 1.64$ | 0.18 | 0.34 | 0.48 | 0.56 | 0.64 | 0.70 | 0.76 | 0.81 | 0.83 | 0.87 | 0.90 |

the horizon of 100 years in 52% of cases, the long-only TSMOM strategy does not statistically significantly outperform the buy-and-hold strategy. The Sharpe ratio of the long-only TSMOM strategy almost surely exceeds the Sharpe ratio of the buy-and-hold strategy over the horizon of 500 years. Yet, even in this case, there is a 4% probability that the Sharpe ratio of the long-only TSMOM strategy is not statistically significantly higher than the Sharpe ratio of the buy-and-hold strategy.

It is worth noting that the probability $Prob(z > 1.64)$ is the probability of correctly rejecting the false null hypothesis at the 5% level. This probability is commonly known as the "statistical power" of the test. From elementary statistics, it is known that when the statistical power is low, there is a high probability of a Type II error or concluding there is no effect when a true effect exists. By convention, 80% is an acceptable level of power in a statistical test. The left panel in Fig. 8 plots the estimated power of the Sharpe ratio test versus the number of years in a sample. The plot in this panel advocates that, to reach the desired power level of 80%, the sample size must be approximately 250 years. Consequently, our simulation results reveal that the Sharpe ratio test suffers from extremely low statistical power.

Next, we briefly discuss the conclusions that can be drawn when the CAPM alpha measures the performance. First, our simulation study results suggest that the alpha test's statistical power is similar for both the long-only and long–short TSMOM strategies. This result seems surprising at first glance because the theoretical alpha of the long–short TSMOM strategy is twice as high as that of the long-only TSMOM strategy. However, the long-only TSMOM strategy has a lower standard deviation of returns and higher beta than the long–short TSMOM strategy. Therefore, the long-only TSMOM strategy has a substantially lower residual variance (in the CAPM regression) and, hence, a lower standard error of estimation of alpha than the long–short TSMOM strategy.[13] Our second conclusion is that the alpha test has a notably higher statistical power than the Sharpe ratio test. Our ballpark estimate is that the sample size must be about 90 years to ensure the recommended statistical power in the alpha test. This sample size is approximately three times smaller than the necessary sample size in the Sharpe ratio test. Still, to reach the desired power level in the alpha test, the sample size must span almost a century.

To sum up, we find that both the Sharpe ratio and CAPM alpha tests suffer from the low statistical power problem. Consequently, we conclude that in virtually all empirical studies that evaluate a trend-following strategy's profitability, the power of the statistical test is

[12] In the Sharpe ratio test, we focus exclusively on the long-only TSMOM strategy because our results suggest that the long–short TSMOM strategy is not superior to the buy-and-hold strategy according to the Sharpe ratio criterion.

[13] Consequently, our results advocate that the standard error of estimation of alpha of the long-only TSMOM strategy is twice as small as that of the long–short strategy.

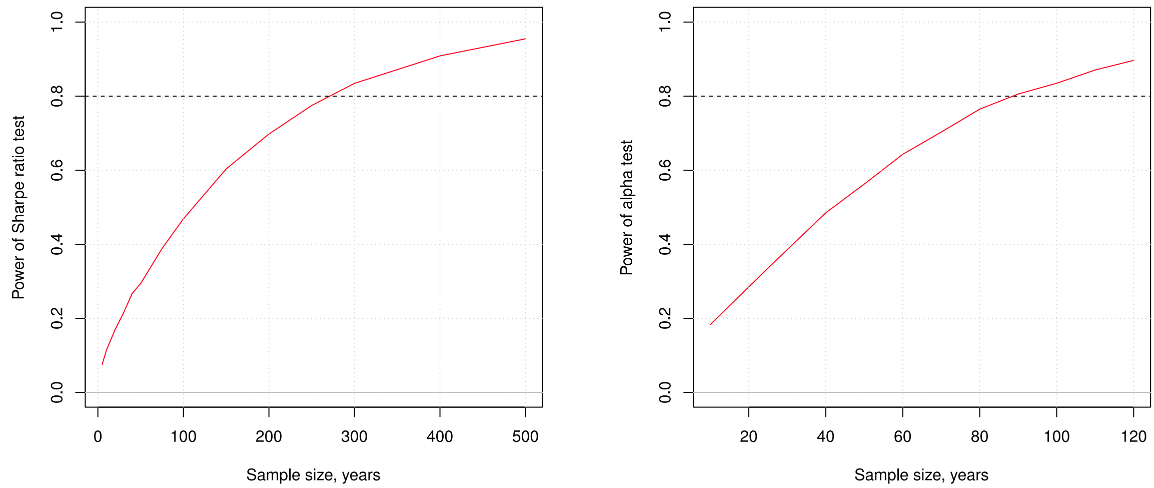


**Fig. 8.** Estimated statistical power of the test of equal performance of the TSMOM and buy-and-hold strategies versus the number of years in a sample. The left (right) panel plots the statistical power of the Sharpe ratio (alpha) test. The statistical power is evaluated using the simulation method. The dashed horizontal line shows the recommended power level in a statistical test.

much below the acceptable level. Finally, it is worth noting that our simulation study results agree very well with the findings reported in the numerous papers. Although the TSMOM strategy's performance in our simulation study is notably higher than that of the buy-and-hold strategy, the outperformance is not guaranteed over short-term horizons. And this result has nothing to do with the increased efficiency of financial markets. It is merely the result of randomness. Besides, the statistically significant outperformance is not guaranteed even over very long-term horizons.

## 7. Conclusions

There is much controversy in the academic literature on the presence of short-term trends in financial markets and the trend-following strategy's profitability. This controversy has different aspects and raises many questions that need to be answered. In this paper, we restrict our attention to studying time series momentum in the S&P Composite stock price index. We aim to answer several important questions regarding time series momentum and explain the existing controversy.

Our contributions are both empirical and theoretical. On the empirical side, we present compelling evidence of short-term momentum in the excess returns on the S&P Composite stock price index. For the first time, we assume that the excess returns follow an autoregressive process of order $p$ and, using a novel methodology, estimate this process's parameters. Our estimation results reveal that the autocorrelation in excess returns is very weak over monthly horizons and, hence, escapes detection when traditional estimation methods are used.

On the theoretical side, armed with a fairly accurate knowledge of the momentum generating process, we provide analytical results on the risk, return, and performance of the long-only and long–short TSMOM strategies. We demonstrate that our analytical results agree well with the empirical results obtained by simulating the historical returns to the TSMOM strategies. We conduct a comparative static analysis of our analytical model to examine how the performance of the TSMOM strategies depends on the model's parameters. This analysis provides new and valuable insights into the properties and profitability of the TSMOM strategies.

The Sharpe ratio and CAPM alpha are the two most common performance measures in modern finance. We find that the choice of performance measure plays a crucial role in determining whether a TSMOM strategy is superior to its passive counterpart. The explanation for this finding is that these performance criteria use totally different risk measures. By relying on a simulation study, we explore how the evidence of the TSMOM strategy's profitability depends on the investment horizon and evaluate the power of statistical tests. This study reveals that both the Sharpe ratio and alpha tests suffer from the low statistical power problem.

By convention, the recommended power of a statistical test is 80%. Our ballpark estimate is that the sample size must be about 250 years with monthly observations to reach the desired power level when the performance is measured by the Sharpe ratio. When the performance is measured by alpha, the sample size must be about 90 years to ensure the recommended statistical power. Even though the alpha test has notably larger statistical power than the Sharpe ratio test, it still requires a sample size that spans almost a century. Consequently, we conclude that in virtually all empirical studies that evaluate a trend-following strategy's profitability, the power of the statistical test is much below the acceptable level.

## CRediT authorship contribution statement

**Valeriy Zakamulin:** Conceptualization, Methodology (empirical), Software, Data curation, Formal analysis, Investigation, Validation, Visualization, Writing – original draft, Writing – review & editing, Supervision. **Javier Giner:** Methodology (theoretical), Formal analysis, Investigation, Writing – original draft.

## Data availability

Data will be made available on request.

## Appendix

*Proof of Proposition 1*

We suppose that $X_t$ is wide-sense stationary process, that is, the process whose mean and autocovariance do not vary with respect to time: $E[X_t] = \mu_x$ and $E[(X_t - \mu_x)(X_{t-k} - \mu_x)] = Cov(X_t, X_{t-k}) = \gamma_k$ for any $t$ and $k$.

By definition,

$$Cor(MOM_{t+k}(k), MOM_t(k)) = \frac{Cov(MOM_{t+k}(k), MOM_t(k))}{Var(MOM_t(k))}, \tag{32}$$

where $Cov(MOM_{t+k}(k), MOM_t(k))$ is the covariance between $MOM_{t+k}(k)$ and $MOM_t(k)$. Note that, because of the stationarity assumption, the variance of $MOM_{t+k}(k)$ equals that of $MOM_t(k)$.

The variance of $MOM_t(k)$ is given by

$$Var(MOM_t(k)) = Var\left(\sum_{i=1}^{k} X_{t-k+i}\right) = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} Cov(X_{t-i}, X_{t-j}).$$

By definition, $Cov(X_{t-i}, X_{t-j}) = \rho_{|i-j|}\gamma_0^2 = \rho_{|i-j|}\sigma_x^2$, where $\rho_m$ denotes the autocorrelation of order $m$ of $X_t$ (with $\rho_0 = 1$) and $\sigma_x^2$ denotes the variance of $X_t$. Consequently, the expression for the variance can be written as

$$Var(MOM_t(k)) = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} \rho_{|i-j|}\sigma_x^2 = \mathbf{1}_k' \boldsymbol{P}_{k,k}\mathbf{1}_k\sigma_x^2.$$

By similar reasoning, the covariance between $MOM_{t+k}(k)$ and $MOM_t(k)$ is given by

$$\begin{aligned} Cov(MOM_{t+k}(k), MOM_t(k)) &= Cov\left(\sum_{i=1}^{k} X_{t+i}, \sum_{j=1}^{k} X_{t-k+j}\right) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{k} Cov(X_{t+i}, X_{t-k+j}) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{k} \rho_{|k-j+i|}\sigma_x^2 = \mathbf{1}_k' \boldsymbol{Q}_{k,k}\mathbf{1}_k\sigma_x^2. \end{aligned}$$

Inserting the expressions for $Cov(MOM_{t+k}(k), MOM_t(k))$ and $Var(MOM_t(k))$ into Eq. (32) completes the proof.

*Proof of Proposition 2*

Consider the returns to the following generalized trading strategy:

$$R_t = \begin{cases} ar_t + c_a & \text{if } MOM_{t-1}(n) > 0, \\ br_t + c_b & \text{if } MOM_{t-1}(n) \le 0. \end{cases}$$

This trading strategy can be seen as an investment directed by binary operators $B_t$ and $C_t$ defined in the following way:

$$B_t = \begin{cases} a & \text{if } MOM_{t-1}(n) > 0, \\ b & \text{if } MOM_{t-1}(n) \le 0, \end{cases} \quad \text{and} \quad C_t = \begin{cases} c_a & \text{if } MOM_{t-1}(n) > 0, \\ c_b & \text{if } MOM_{t-1}(n) \le 0, \end{cases}$$

so that the returns to the generalized trading strategy can be expressed as $R_t = B_t r_t + C_t$. The derivation of the formulas for the mean and variance of $R_t$ follows along the lines of the derivation in Acar (2003) who considered the simplified case $R_t = B_t r_t$.

*Some necessary intermediate results*

Taking into account that the joint distribution of $r_t$ and $MOM_{t-1}(n)$ can be represented by the bivariate normal distribution given by (14), we introduce variables $z_t$ and $y_{t-1}$ that are standardized versions of $r_t$ and $MOM_{t-1}(n)$ respectively. That is, $z_t = \frac{r_t-\mu}{\sigma}$ and $y_{t-1} = \frac{MOM_{t-1}(n)-m}{v}$. Then the pair of variables $(z_t, y_{t-1})$ follows a bivariate standard normal distribution with correlation coefficient $\varrho_n$.

We introduce the following variable:

$$d = -\frac{m}{v}.$$

With this notation, the probability of $MOM_{t-1}(n)$ being less or greater than 0 equals

$$Prob(MOM_{t-1}(n) \le 0) = Prob(y_{t-1} \le d) = \Phi(d),$$

$$Prob(MOM_{t-1}(n) > 0) = Prob(y_{t-1} > d) = 1 - \Phi(d) = \Phi(-d).$$

Note that in the $y_{t-1}z_t$ plane the line $y_{t-1} = d$ divides the whole plane into two half-planes: in the half-plane where $y_{t-1} \le d$ the trading signal is Sell, whereas in the other half-plane where $y_{t-1} > d$ the trading signal is Buy.

The following results are straightforward to derive:

$$E[B_t] = a\Phi(-d) + b\Phi(d), \tag{33}$$

$$E[C_t] = c_a\Phi(-d) + c_b\Phi(d), \tag{34}$$

$$E[B_t^2] = a^2\Phi(-d) + b^2\Phi(d), \tag{35}$$

$$E[C_t^2] = c_a^2\Phi(-d) + c_b^2\Phi(d). \tag{36}$$

Consider the first and second moments of $z_t$ conditioned on $y_{t-1}$ being less or greater than the value of $d$. Using the results of Kotz et al. (2000, pages 311-315) on the first and second moments of truncated bivariate distributions, we obtain

$$E[z_t|y_{t-1} > d] = \frac{\varrho_n\varphi(d)}{\Phi(-d)}, \tag{37}$$

$$E[z_t|y_{t-1} \le d] = -\frac{\varrho_n\varphi(d)}{\Phi(d)}, \tag{38}$$

$$E[z_t^2|y_{t-1} > d] = 1 + \frac{\varrho_n^2 d\varphi(d)}{\Phi(-d)}, \tag{39}$$

$$E[z_t^2|y_{t-1} \le d] = 1 - \frac{\varrho_n^2 d\varphi(d)}{\Phi(d)}. \tag{40}$$

As an example, the expectation $E[B_t z_t]$ can be obtained by combining Eqs. (37) and (38)

$$E[B_t z_t] = a\Phi(-d)E[z_t|y_{t-1} > d] + b\Phi(d)E[z_t|y_{t-1} \le d] = (a-b)\varrho_n\varphi(d). \tag{41}$$

As another example, the expectation $E[B_t^2 z_t^2]$ can be obtained by combining Eqs. (39) and (40)

$$\begin{aligned} E[B_t^2 z_t^2] &= a^2\Phi(-d)E[z_t^2|y_{t-1} > d] + b^2\Phi(d)E[z_t^2|y_{t-1} \le d] \\ &= a^2\Phi(-d) + b^2\Phi(d) + (a^2 - b^2)\varrho_n^2 d\varphi(d). \end{aligned} \tag{42}$$

As a final example, the solution for $E[B_t z_t^2]$ is given by

$$\begin{aligned} E[B_t z_t^2] &= a\Phi(-d)E[z_t^2|y_{t-1} > d] + b\Phi(d)E[z_t^2|y_{t-1} \le d] \\ &= a\Phi(-d) + b\Phi(d) + (a-b)\varrho_n^2 d\varphi(d). \end{aligned} \tag{43}$$

All these three results are needed later in the derivations.

*Mean returns of generalized trading strategy*

Consider the expression for the mean returns of the generalized trading strategy:

$$E[R_t] = E[B_t r_t + C_t] = E[B_t r_t] + E[C_t]. \tag{44}$$

The term $E[C_t]$ is given by Eq. (34). The expression for the term $E[B_t r_t]$ can be represented in the following manner:

$$E[B_t r_t] = E[B_t(\mu + \sigma z_t)] = \mu E[B_t] + \sigma E[B_t z_t].$$

The term $E[B_t]$ is given by Eq. (33), whereas the term $E[B_t z_t]$ is given by Eq. (41). Putting everything together into Eq. (44), we obtain

$$E[R_t] = \mu\left(a\Phi(-d) + b\Phi(d)\right) + (a-b)g + c_a\Phi(-d) + c_b\Phi(d), \tag{45}$$

where, for the sake of brevity, we denote the product $\sigma\varrho_n\varphi(d)$ by $g$.

*Variance of returns of generalized trading strategy*

The variance of returns of the generalized trading strategy can be computed as:

$$Var[R_t] = E[R_t^2] - E[R_t]^2. \tag{46}$$

The term $E[R_t]$ is given by Eq. (45). Consider the term $E[R_t^2]$ which is the mean squared return of the generalized trading strategy:

$$\begin{aligned} E[R_t^2] &= E\left[(B_t r_t + C_t)^2\right] = E\left[B_t^2 r_t^2 + 2B_t C_t r_t + C_t^2\right] \\ &= E\left[B_t^2 r_t^2\right] + 2E\left[B_t C_t r_t\right] + E\left[C_t^2\right]. \end{aligned} \tag{47}$$

The term $E[C_t^2]$ is given by Eq. (36). Consider the term $E[B_tC_tr_t]$

$$E[B_tC_tr_t] = E[B_tC_t(\mu + \sigma z_t)] = \mu E[B_tC_t] + \sigma E[B_tC_tz_t].$$

The expression for $E[B_tC_t]$ is straightforward to derive:

$$E[B_tC_t] = ac_a\Phi(-d) + bc_b\Phi(d).$$

The expression for $E[B_tC_tz_t]$ can be obtained in a similar manner to that of the expression for $E[B_tz_t]$ (see Eq. (41))

$$E[B_tC_tz_t] = (ac_a - bc_b)\varrho_n\varphi(d).$$

Therefore,

$$E[B_tC_tr_t] = \mu\left(ac_a\Phi(-d) + bc_b\Phi(d)\right) + (ac_a - bc_b)g. \tag{48}$$

Now consider the expression for $E[B_t^2r_t^2]$

$$E[B_t^2r_t^2] = E[B_t^2(\mu + \sigma z_t)^2] = \mu^2E[B_t^2] + 2\mu\sigma E[B_t^2z_t] + \sigma^2E[B_t^2z_t^2].$$

The terms with $E[B_t^2]$ and $E[B_t^2z_t^2]$ are given by Eqs. (35) and (42) respectively. The expression for the term $E[B_t^2z_t]$ is obtained in a similar manner to that of $E[B_tz_t]$ (see Eq. (41))

$$E[B_t^2z_t] = \left(a^2 - b^2\right)\varrho_n\varphi(d).$$

Consequently, we obtain

$$E[B_t^2r_t^2] = \left(\mu^2 + \sigma^2\right)\left(a^2\Phi(-d) + b^2\Phi(d)\right) + \left(a^2 - b^2\right)g(2\mu + \sigma\varrho_nd).$$

Finally, substituting $E[C_t^2]$, $E[B_tC_tr_t]$, and $E[B_t^2r_t^2]$ into Eq. (47) for $E[R_t^2]$, the expression for the variance of returns of the generalized trading strategy becomes:

$$\begin{aligned}Var[R_t] = {} & \left(\mu^2 + \sigma^2\right)\left(a^2\Phi(-d) + b^2\Phi(d)\right) + \left(a^2 - b^2\right)g(2\mu + \sigma\varrho_nd)\\ & + 2\mu\left(ac_a\Phi(-d) + bc_b\Phi(d)\right) + 2(ac_a - bc_b)g\\ & + c_a^2\Phi(-d) + c_b^2\Phi(d) - E[R_t]^2.\end{aligned} \tag{49}$$

*Alpha and beta of generalized trading strategy*

The alpha in the CAPM is given by:

$$\alpha = (E[R_t] - r_f) - \beta(\mu - r_f), \tag{50}$$

where $E[R_t]$ is given by Eq. (45), while the beta is defined by:

$$\beta = \frac{Cov(R_t, r_t)}{Var(r_t)} = \frac{E[R_tr_t] - E[R_t]\mu}{\sigma^2}. \tag{51}$$

The only new term in Eq. (51) is:

$$E[R_tr_t] = E[(B_tr_t + C_t)r_t] = E[B_tr_t^2] + E[C_tr_t].$$

To obtain the analytical result for $E[R_tr_t]$, first, we derive the solution for $E[B_tr_t^2]$;

$$E[B_tr_t^2] = E[B_t(\mu + \sigma z_t)^2] = \mu^2E[B_t] + 2\mu\sigma E[B_tz_t] + \sigma^2E[B_tz_t^2].$$

The solution for $E[B_t]$ is given by Eq. (33), the solution for $E[B_tz_t]$ is given by Eq. (41), while the solution for $E[B_tz_t^2]$ is given by Eq. (43). Putting it all together we get after simplification

$$E[B_tr_t^2] = \left(\mu^2 + \sigma^2\right)(a\Phi(-d) + b\Phi(d)) + (a - b)g(2\mu + \sigma\varrho_nd). \tag{52}$$

Second, we derive the solution for $E[C_tr_t]$:

$$E[C_tr_t] = E[C_t(\mu + \sigma z_t)] = \mu E[C_t] + \sigma E[C_tz_t].$$

The solution for $E[C_t]$ is given by Eq. (34). The solution for $E[C_tz_t]$ is obtained along the same lines as the solution for $E[B_tz_t]$ (see Eq. (41)):

$$E[C_tz_t] = (c_a - c_b)\varrho_n\varphi(d).$$

Putting it all together yields:

$$E[C_tr_t] = \mu\left(c_a\Phi(-d) + c_b\Phi(d)\right) + (c_a - c_b)\varrho_n\varphi(d). \tag{53}$$

Finally, inserting the solutions for $E[R_tr_t]$ and $E[R_t]$ into Eq. (51) we get after simplification

$$\beta = a\Phi(-d) + b\Phi(d) + \frac{(a - b)g(\mu + \sigma\varrho_nd) + g(c_a - c_b)}{\sigma^2}. \tag{54}$$

*Results for the long-only and long–short trading strategy*

The particular results for the long-only trading strategy are obtained through Eqs. (45), (49), (54), and (50) by using $a = 1$, $b = 0$, $c_a = 0$ and $c_b = r_f$.

The particular results for the long–short trading strategy are obtained through Eqs. (45), (49), (54), and (50) by using $a = 1$, $b = -1$, $c_a = 0$ and $c_b = 2r_f$.

## References


Acar, E. (2003). Chapter 8 - modelling directional hedge funds-mean, variance and correlation with tracker funds. In S. Satchell, & A. Scowcroft (Eds.), *Advances in portfolio construction and implementation* (pp. 193–214). Oxford: Butterworth-Heinemann.

Andrews, D. W. K., & Chen, H.-Y. (1994). Approximately median-unbiased estimation of autoregressive models. *Journal of Business & Economic Statistics*, *12*(2), 187–204.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time series analysis: Forecasting and control* (5). John Wiley & Sons Inc., New Jersey.

Brock, W., Lakonishok, J., & LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance*, *47*(5), 1731–1764.

Clare, A., Seaton, J., Smith, P. N., & Thomas, S. (2013). Breaking into the blackbox: Trend following, stop losses and the frequency of trading - the case of the S&P 500. *Journal of Asset Management*, *14*(3), 182–194.

Faber, M. T. (2007). A quantitative approach to tactical asset allocation. *Journal of Wealth Management*, *9*(4), 69–79.

Faber, M. (2017). A quantitative approach to tactical asset allocation revisited 10 years later. *Journal of Portfolio Management*, *44*(2), 156–167.

Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy*, *96*(2), 246–273.

Fifield, S. G. M., Power, D. M., & Sinclair, C. D. (2005). An analysis of trading strategies in eleven European stock markets. *European Journal of Finance*, *11*(6), 531–548.

Georgopoulou, A., & Wang, J. G. (2016). The trend is your friend: Time-series momentum strategies across equity and commodity markets. *Review of Finance*, *21*(4), 1557–1592.

Gwilym, O., Clare, A., Seaton, J., & Thomas, S. (2010). Price and momentum as robust tactical approaches to global equity investing. *Journal of Investing*, *19*(3), 80–91.

Ham, H., Cho, H., Kim, H., & Ryu, D. (2019). Time-series momentum in China's commodity futures market. *Journal of Futures Markets*, *39*(12), 1515–1528.

Huang, D., Li, J., Wang, L., & Zhou, G. (2020). Time series momentum: Is it there? *Journal of Financial Economics*, *135*(3), 774–794.

Hutchinson, M. C., & O'Brien, J. (2014). Is this time different? Trend-following and financial crises. *Journal of Alternative Investments*, *17*(2), 82–102.

Jobson, J. D., & Korkie, B. M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *The Journal of Finance*, *36*(4), 889–908.

Kilgallen, T. (2012). Testing the simple moving average across commodities, global stock indices, and currencies. *Journal of Wealth Management*, *15*(1), 82–100.

Kim, M. J., Nelson, C. R., & Startz, R. (1991). Mean reversion in stock prices? A reappraisal of the empirical evidence. *Review of Economic Studies*, *58*(3), 515–528.

Kim, A. Y., Tse, Y., & Wald, J. K. (2016). Time series momentum and volatility scaling. *Journal of Financial Markets*, [ISSN: 1386-4181] *30*, 103–124.

Kotz, S., Balakrishnan, N., & Johnson, N. L. (2000). *Continuous multivariate distributions, Vol. 1: Models and applications* (2). New York: John Wiley & Sons, Inc..

Lee, C. I., Pan, M.-S., & Liu, Y. (2001). On market efficiency of Asian foreign exchange rates: Evidence from a joint variance ratio test and technical trading rules. *Journal of International Financial Markets, Institutions and Money*, [ISSN: 1042-4431] *11*(2), 199–214.

Lim, B. Y., Wang, J. G., & Yao, Y. (2018). Time-series momentum in nearly 100 years of stock returns. *Journal of Banking & Finance*, *97*, 283–296.

Marques, C. R. (2005). Inflation persistence: Facts or artifacts? *Banco de Portugal Economic Bulletin*, *11*(2), 69–79.

Marshall, B. R., Nguyen, N. H., & Visaltanachoti, N. (2017). Time series momentum and moving average trading rules. *Quantitative Finance*, *17*(3), 405–421.

Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Letters*, *1*, 21–23.

Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics*, *104*(2), 228–250.

Neely, C. J., Rapach, D. E., Tu, J., & Zhou, G. (2014). Forecasting the equity risk premium: The role of technical indicators. *Management Science*, *60*(7), 1772–1791.

Nelson, C. R., & Kim, M. J. (1993). Predictable stock returns: The role of small sample bias. *The Journal of Finance*, *48*(2), 641–661.

Okunev, J., & White, D. (2003). Do momentum-based strategies still work in foreign currency markets? *Journal of Financial and Quantitative Analysis*, *38*(2), 425–447.

Olson, D. (2004). Have trading rule profits in the currency markets declined over time? *Journal of Banking & Finance*, *28*(1), 85–105.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American Statistical Association*, *79*(388), 871–880.

Rousseeuw, P. J. (1985). Multivariate estimation with high breakdown point. In W. Grossmann, G. Pflug, I. Vincze, & W. Wertz (Eds.), *Mathematical statistics and applications, Vol. B* (pp. 283–297). Dordrecht, Nederland: Reidel Publishing Company.

Rousseeuw, P. J., & Driessen, K. V. (1999). A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, *41*(3), 212–223.

Schwert, G. W. (1990). Indexes of U.S. stock prices from 1802 to 1987. *Journal of Business*, *63*(3), 399–426.

Siegel, J. (2002). *Stocks for the long run*. McGraw-Hill Companies.

Sullivan, R., Timmermann, A., & White, H. (1999). Data-snooping, technical trading rule performance, and the bootstrap. *The Journal of Finance*, *54*(5), 1647–1691.

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, *21*(4), 1455–1508.

Zakamulin, V. (2014). The real-life performance of market timing with moving average and time-series momentum rules. *Journal of Asset Management*, *15*(4), 261–278.

Zakamulin, V. (2017). *Market timing with moving averages: The anatomy and performance of trading rules*. Palgrave Macmillan.

Zakamulin, V., & Giner, J. (2020). Trend following with momentum versus moving average: A tale of differences. *Quantitative Finance*, *20*(6), 985–1007.