Time Series Momentum in the US Stock Market: Empirical Evidence and Theoretical Implications

Valeriy Zakamulin* and Javier Giner†

This revision: April 26, 2020

Abstract

We start this paper by presenting compelling evidence of short-term momentum in the excess returns on the S&P Composite stock price index. For the first time ever, we assume that the excess returns follow an autoregressive process of order p, $AR(p)$, and evaluate the parameters of this process. Armed with a fairly accurate knowledge of the momentum generating process, we continue this paper by providing a number of important theoretical implications. First, we present analytical results on the profitability of long-only and long-short time-series momentum (TSMOM) strategies. Our results suggest that the long-only TSMOM strategy is profitable, while the long-short one is not. We find that over multiple periods the risk profile of the long-only TSMOM strategy resembles the risk profile of a portfolio insurance strategy. We estimate the power of the statistical test for superiority of the TSMOM strategy and find that the power is much below the acceptable level. Consequently, any empirical study tends not to reject the null hypothesis of no profitability of TSMOM strategy. Finally, we evaluate the precision of identification of the optimal number of lags in the TSMOM rule using a standard back-testing methodology and find that this precision is extremely poor. However, we demonstrate that the performance of the TSMOM rule is robust to the choice of the number of lags.

Key words: technical trading rules, time series momentum, profitability, statistical power

JEL classification: G12, G14, G17

*School of Business and Law, University of Agder, Service Box 422, 4604 Kristiansand, Norway, E-mail: Valeri.Zakamouline@uia.no

†Department of Economics, Accounting and Finance, University of La Laguna, Camino La Hornera s/n, 38071, Santa Cruz de Tenerife, Spain, E-mail: jginer@ull.edu.es

1 Introduction

One of the key tenets of technical analysis is that prices tend to move in trends. The Dow Theory postulates that there are three types of trends in financial markets (see Brown, Goetzmann, and Kumar (1998) and references therein), the most important of which is the primary trend. Primary trends can be classified as bull and bear markets that tend to last for one year or more. These trends can be identified in a timely manner to generate profits and limit losses. However, whereas trend-following has been extensively employed by practitioners for almost a century, academics had long been skeptical of its usefulness. The study by Brock, Lakonishok, and LeBaron (1992) marked the onset of a change in the academics' attitudes toward existence of market trends because this study documented the profitability of trend-following strategies.

Since that time, the profitability of trend-following strategies has been documented in a large number of empirical studies.[1] The majority of these empirical studies find that these strategies are profitable in the long-run over periods ranging from 50 to 150 years. However, two issues of concern arise regarding the empirical performance of trend-following strategies. The first issue is that the researchers frequently report that, when they use the most recent 5 to 10 years in their sample of historical data, the trend-following strategies are not profitable (see, for example, Sullivan et al. (1999), Lee et al. (2001), Siegel (2002, Chapter 2), Okunev and White (2003), Olson (2004), Hutchinson and O'Brien (2014), and Zakamulin (2014)). This issue may imply that market efficiency improves over time and the trend-following strategies are no longer profitable. Another issue is the lack of scientific evidence on profitability of trend-following strategies. Specifically, even when the historical sample is rather long and the profitability of trend-following strategies is economically significant, quite often the researchers cannot reject the null hypothesis that the performance of a trend-following strategy is similar to the performance of the corresponding buy-and-hold strategy (see, among others, Kim, Tse, and Wald (2016), Zakamulin (2017), and Huang, Li, Wang, and Zhou (2020)). This issue may suggest that the profitability of trend-following strategies is an artifact of data-snooping.

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 asset

[1] Examples of such studies are: Sullivan, Timmermann, and White (1999), Lee, Pan, and Liu (2001), Siegel (2002), Okunev and White (2003), Olson (2004), Fifield, Power, and Sinclair (2005) Faber (2007), Gwilym, Clare, Seaton, and Thomas (2010), Kilgallen (2012), Moskowitz, Ooi, and Pedersen (2012), Clare, Seaton, Smith, and Thomas (2013), Neely, Rapach, Tu, and Zhou (2014), Hutchinson and O'Brien (2014), Marshall, Nguyen, and Visaltanachoti (2017), Zakamulin (2017), and Faber (2017).

classes, Moskowitz et al. (2012) demonstrate that the past 12-month returns predict the next month return; a trading strategy, which buys assets if their past 12-month returns are positive and sells them otherwise, earns significant risk-adjusted returns. However, recently the results reported by Moskowitz et al. (2012) have been heavily criticized. Specifically, Kim et al. (2016) find that the results by Moskowitz et al. (2012) are largely driven by volatility-scaling returns rather than by short-term momentum effect. 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. Even a pooled regression does not provide evidence of momentum. All in all, the papers by Kim et al. (2016) and Huang et al. (2020) cast serious doubts on the existence of short-term momentum in financial markets.

Everything considered, one can note that there is much controversy in the academic literature on trend-following. This controversy has different aspects and raises the following set of questions that have to be answered: Are there primary (short-term) trends in financial markets? If the answer is affirmative, what is the type of process that generates these trends? Are trend-following strategies profitable? An affirmative answer to this question leads to another question: Why the existing empirical evidence on profitability is often controversial?

The present paper restricts its attention to the study of time-series momentum in the US stock market. The objective of this paper is to suggest answers to all the questions raised in the preceding paragraph. 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 stock price index. For the first time ever, we assume that the excess returns follow an autoregressive process of order p, $AR(p)$, and estimate the parameters of this process. The difficulty is that over monthly horizons the autocorrelation in excess returns is very weak 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 produce the best fit to a theoretical model.

All previous results on the performance and profitability of trend-following rules have been obtained solely by means of 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 performance of the trend-following strategy 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, the statistical power of a test is directly related to the sample size; the power increases with an increase in sample size. Therefore, we suspect that in a great deal of studies the inability to reject the null hypothesis does not imply that the null hypothesis is true, but implies that the sample size is not long enough. Second, an empirical study is not integrated with a theoretical model that allows researchers justify and explain the observed results and provide deep insights into the nature of the trend process and the properties of trend-following strategies.

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

More specifically, we start the second half of this paper with providing analytical results on the one-period mean returns, variance of returns, and Sharpe ratios 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. Our evaluation results suggest that the long-short TSMOM strategy is not profitable, whereas the long-only TSMOM strategy has a clear edge over the buy-and-hold strategy. Specifically, we estimate that the Sharpe ratio of the long-only TSMOM strategy is about 30% higher than that of the buy-and-hold strategy.

Whereas the one-period properties of the TSMOM strategy can be examined using analytical results, the multi-period properties of the TSMOM strategy do not allow analytical solutions because their returns represent time series with rather complicated dependence. Therefore, we study the multi-period properties of the TSMOM strategy by means of carrying out extensive and careful computer simulations. We find that the shape of the probability density of multi-period returns to the long-only TSMOM strategy resembles the shape of density function of returns to a portfolio insurance strategy. Therefore, the risk profile of the long-only TSMOM strategy is totally different from that of the buy-and-hold strategy. For example, the results of our simulation analysis suggest that, as compared to the buy-and-hold strategy, over a 5-year horizon the long-only TSMOM strategy has twice as small the probability of loss and the expected loss if loss occurs. Consequently, the advantage of the long-only TSMOM strategy over the buy-and-hold strategy lies not only in the higher Sharpe ratio, but also in its superior downside protection.

Our results strongly suggest that over the long-run the (long-only) TSMOM strategy outperforms the buy-and-hold strategy. However, because of randomness, over a short-run the outperformance is not guaranteed. Using a simulation analysis, we find that over short- to medium-term horizons from 5 to 10 years, the probability that the TSMOM strategy outperforms the buy-and-hold strategy is less than 60%. Consequently, it is not surprising that in short samples the researchers often find that trend-following strategies are not profitable. By convention, the desired power of a statistical test is 80%. Our ballpark estimate is that, in order to reach the desired power level, the sample size must be about 250 years with monthly observations. The clear-cut implication from this result is that any empirical study tends not to reject the null hypothesis of no profitability because the power of the statistical test is much below the acceptable level.

Finally, by relying on a simulation study, we examine the precision of identification of the optimal number of return lags in the TSMOM rule using a standard back-testing methodology. We find that this precision is poor even when a sample of monthly observations covers several hundred years. However, we demonstrate that the performance of the TSMOM rule is robust to the choice of the number of lags. We further advocate that the performance of the TSMOM strategy depends mainly on the trend strength rather than on the choice of the number of lags.

The rest of the paper is organized as follows. Section 2 presents the TSMOM trading rules. Section 3 motivates for the choice of the autoregressive process for the excess returns 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 parameters of the autoregressive process for excess returns. Section 6 presents a number of analytical and simulation results and discusses the theoretical implications of time-series momentum. Finally, Section 7 concludes the paper.

2 Time Series Momentum Trading Rules

Denote by r_t and r_{ft} the month t log total return on the stock market and the log risk-free rate of return respectively. Denote by

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

the month t market excess return (that is, the market return in excess of the risk-free rate of return).

Generally, the value of the technical trading indicator of the TSMOM rule at the end of month $t-1$ is computed as the sum of the excess returns over the past n months:

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

It is worth emphasising that the technical indicator is computed at month-end $t-1$ and translated into a trading signal for the subsequent month t. If, for example, $MOM_{t-1}(n) > 0$, then the trading signal is Buy. This means that the trader buys the stocks at month $t-1$ closing price and holds them over the subsequent month t. If, on the other hand, $MOM_{t-1}(n) \leq 0$, the trading signal for the subsequent month t is Sell.

A Buy signal is always a signal to invest in the stocks (or stay invested in the stocks). When a Sell signal is generated, there are two alternative strategies. In the "long-only" strategy, a Sell signal is a signal to sell the stocks and invest the proceeds at the risk-free rate. In this case, in the absence of transaction costs, the return to the TSMOM strategy over month t is given by

$$R_t^{LO} = \begin{cases} r_t & \text{if signal is Buy,} \\ r_{ft} & \text{if signal is Sell.} \end{cases} \quad (1)$$

In the "long-short" strategy, a Sell signal is a signal to sell short the stocks and invest the proceeds at the risk-free rate. In this case the return to the TSMOM strategy over month t is given by

$$R_t^{LS} = \begin{cases} r_t & \text{if signal is Buy,} \\ 2r_{ft} - r_t & \text{if signal is Sell.} \end{cases} \quad (2)$$

Note that in this case it is assumed that, when a Sell signal is generated after a Buy signal, the trader sells all own shares of the stocks and, right after the selling own shares, the trader sells short the same number of shares of the stocks. The proceeds from the sale and the short sale are invested at the risk-free rate.

3 Process for Excess Returns

For the TSMOM strategy to be profitable, the future excess return must be predictable from the past excess returns. Specifically, the TSMOM requires persistence in excess returns. Put differently, the excess returns must exhibit positive autocorrelation. Therefore, to model the autocorrelation in excess returns, we assume that they follow the autoregressive process of order p. This $AR(p)$ model is defined as:

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

where c is a constant, p is the number of autoregressive terms, the coefficients $\{\phi_1, \phi_2, \ldots, \phi_p\}$ are the parameters of the model, X_{t-i} is the excess return observed at month $t-i$, and ε_t is the noise term which is i.i.d. random process with zero mean and variance of σ_ε^2. That is, $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$. We assume that the autoregressive coefficients ϕ_i satisfy the stationarity conditions.

In our paper, the persistence of the process is measured by the sum of the autoregressive coefficients, $\alpha = \sum_{i=1}^{p} \phi_i$. This measure was proposed by Andrews and Chen (1994) and subsequently by Marques (2005). Specifically, Marques (2005) starts with observing that every autoregressive process $AR(p)$ is, in fact, a mean-reverting process. The speed of mean reversion is inversely proportional to α. In particular, the larger the numerical value of α, the slower the reversion to the long-run mean and, hence, the stronger is the persistence. Consequently, the sum of the autoregressive coefficients can be used to measure the persistence. It is worth noting that if all coefficients $\{\phi_1, \phi_2, \ldots, \phi_p\}$ of the $AR(p)$ process are non-negative, then increasing the numerical value of some ϕ_i or increasing the order p increases the persistence of the $AR(p)$ process.

By multiplying equation (3) by X_{t-k}, taking expectation, and then dividing the resulting expression by the variance of X_t, we obtain the important recursive relationship for the autocorrelation coefficients of the $AR(p)$ process:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \ldots + \phi_p \rho_{k-p}, \tag{4}$$

where ρ_i denotes the autocorrelation between X_t and X_{t-i}. Putting $k = 1, 2, \ldots, p$ into equation (4) and using $\rho_0 = 1$ and $\rho_{-i} = \rho_i$, we obtain the set of Yule-Walker linear equations. Given numerical values for $\{\phi_1, \phi_2, \ldots, \phi_p\}$, these linear equations can be solved to obtain numerical values for $\{\rho_1, \rho_2, \ldots, \rho_p\}$. Equation (4) can then be recursively used to obtain numerical values for ρ_k for any $k > p$.

The mean and variance of X_t are given by (see, for example, Box, Jenkins, Reinsel, and Ljung (2016), Chapter 3)

$$\mu_x = \frac{c}{1 - \sum_{i=1}^{p} \phi_i}, \qquad \sigma_x^2 = \frac{\sigma_\varepsilon^2}{1 - \sum_{i=1}^{p} \phi_i \rho_i}. \tag{5}$$

One of the main goals of this paper is to evaluate the parameters of the $AR(p)$ process for excess returns using the long-run historical data for the US stock market. In particular, we are going to evaluate the values of ϕ_i and p. To make this task 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 the unknown parameter of the $AR(p)$ process for excess returns. Specifically, if Conjecture 1 is true, then all autoregressive coefficients in the $AR(p)$ process for returns are alike, $\phi_i = \phi$ for all $i \in [1, p]$, and the number of autoregressive terms equals the number of return lags in the TSMOM rule, $p = n$. This is because the $MOM(n)$ trading indicator has the highest possible correlation with the future return when $\phi_i = \phi$ and $n = p$, see Zakamulin and Giner (2020). In addition, 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. Hence, under conditions $\phi_i = \phi$ and $n = p$ the Sharpe ratio of the TSMOM strategy with n return lags represents the highest possible Sharpe ratio. Consequently, our task reduces to evaluating only two values: ϕ and p.

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

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

Therefore, equations (5) for the mean and variance of the process for excess returns reduce to

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

where the product $p\,\phi$ equals the persistence of the process α.

4 Data

The data used in our study are the monthly total returns on the Standard and Poor's Composite stock price index, as well as the risk-free rate of return proxied by the T-bill rate. Our sample period begins in January 1857 and ends in December 2018. The data on the S&P Composite index come from two sources. The index returns for the period January 1857 to December 1925 are provided by William Schwert.[2] The returns for the period January 1926 to December 2018 are computed from the closing monthly priced of the S&P Composite index and corresponding dividend data provided by Amit Goyal.[3] The T-bill rate for the period January 1920 to December 2018 is also provided by Amit Goyal. Because there was no risk-free short-term debt prior to the 1920s, we estimate it in the same manner as in Welch and Goyal (2008) using the monthly data for the Commercial Paper Rates for New York. These data are available for the period January 1857 to December 1971 from the National Bureau of Economic Research (NBER) Macrohistory database.[4]

We are primarily interested in the evaluation of autoregressive coefficients ϕ_i and the number of autoregressive terms p in the $AR(p)$ process for excess returns. 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 in between 6 and 12. The most straightforward approach to estimating the autore-

[2] `http://schwert.ssb.rochester.edu/data.htm`
[3] `http://www.hec.unil.ch/agoyal/`
[4] `http://research.stlouisfed.org/fred2/series/M13002US35620M156NNBR`

gressive coefficients is via using an OLS regression model. Table 1 reports the results of the estimation of the autoregressive coefficients ϕ_i and the empirical trend strength $\alpha = \sum_{i=1}^{p} \phi_i$ using the excess returns on the S&P Composite index. Specifically, using the excess return data for the total sample and the first and second halves of the sample, the table reports the estimated autoregressive coefficients and the sum of the coefficients. The number of lags $p = 12$ is chosen to capture the short-term momentum in the S&P Composite index.

	1858-2018		**1858-1937**		**1938-2018**	
Parameter	**Estimate**	**Std. error**	**Estimate**	**Std. error**	**Estimate**	**Std. error**
ϕ_1	0.067	0.042	**0.118**	0.057	0.017	0.044
ϕ_2	0.001	0.036	-0.011	0.054	-0.006	0.031
ϕ_3	-0.056	0.042	-0.094	0.060	0.034	0.034
ϕ_4	0.022	0.031	0.024	0.046	0.036	0.037
ϕ_5	**0.077**	0.027	**0.083**	0.040	**0.073**	0.034
ϕ_6	-0.042	0.036	-0.049	0.054	-0.065	0.041
ϕ_7	0.025	0.029	0.055	0.042	0.004	0.029
ϕ_8	0.043	0.032	**0.092**	0.041	-0.030	0.036
ϕ_9	0.016	0.038	0.026	0.046	-0.011	0.040
ϕ_{10}	0.013	0.030	0.047	0.040	-0.008	0.038
ϕ_{11}	0.004	0.034	-0.021	0.048	0.033	0.036
ϕ_{12}	-0.005	0.031	-0.023	0.044	0.048	0.032
α	0.167		0.246		0.127	

Table 1: Estimation of the autoregressive coefficients and the persistency in excess returns. ϕ_i denotes the autocorrelation coefficient at lag i. $\alpha = \sum_{i=1}^{12} \phi_i$ denotes the measure of persistency. The standard errors are computed using the Newey-West heteroskedasticity and autocorrelation-consistent estimator with 12 lags. Bold text highlights the values that are statistically significant at the 5% level.

In sum, the empirical results reported in Table 1 suggest the presence of relatively weak persistency ($\alpha = 0.246$) in excess returns in the first half of the sample and even weaker persistency ($\alpha = 0.127$) in the second half of the sample. Over the total sample and the second half of the sample, only ϕ_5 is positive and statistically significantly different from zero. In the first half of the sample, there are three positive and statistically significant coefficients: ϕ_1, ϕ_5, and ϕ_8. Consequently, only a small fraction of the estimated autoregressive coefficients are positive and statistically significantly different from zero. The value pattern of the estimated coefficients does not allow us either to evaluate the common value ϕ or the number of terms p.[5] These results seem to suggest that TSMOM is artifact. However, it is premature to jump to definite conclusions. In the subsequent section we present evidence that TSMOM does exist.

[5]We also tried to estimate the number of lags in the $AR(p)$ model using various model selection criteria. For example, the Bayesian information criterion (BIC) selects $p = 8$ for the total sample and the first half of the sample and $p = 12$ for the second half of the sample. However, the differences in the values of BICs for various lags $p \in [1, 12]$ are very marginal. Thus, the validity of this approach to select the lag length p is questionable.

5 Empirical Evidence of Persistence and Model Identification Procedure

In this section we document evidence of short-term persistence in the excess returns on the S&P Composite stock price index. The evidence is obtained by means of two different tests. The first test reveals the evidence of persistence and, under Conjecture 1, allows us to evaluate the value of the autoregressive coefficient ϕ given the order of the autoregression p. The second test reinforces the evidence of short-term persistence and allows us to jointly evaluate the parameters ϕ and p of the autoregressive process for excess returns. The results of the two tests agree with each other and present compelling evidence of the presence of time-series momentum in the S&P Composite index.

5.1 A Novel Test for Return Persistence

Is there persistence in the excess returns on the on the S&P Composite stock price index? The empirical results reported in Table 1 do not allow us to draw any definite conclusions. Table 1 reports the significance tests for a number of return lags and finds that, over the total sample and the second half of the sample, only one return lag is statistically significant. In this context, the well-known issue is that when one conducts testing many coefficients for statistical significance, one will inevitably find coefficients that are "significant." That is, statistical significance can be caused by luck or chance.

Recently, Zakamulin and Giner (2020) suggest a novel test for persistence in return series. This test is based on measuring the empirical correlation between the trading indicators of two TSMOM rules with n and m return lags and testing whether the empirical correlation is higher than the theoretical correlation under the random walk. Specifically, when the excess returns follow the $AR(p)$ process, the correlation coefficient between two trading indicators $MOM(n)$ and $MOM(m)$ is given by

$$Cor(MOM(n), MOM(m)) = \frac{\mathbf{1}_n' \boldsymbol{P}_{n,m} \mathbf{1}_m}{\sqrt{\mathbf{1}_n' \boldsymbol{P}_{n,n} \mathbf{1}_n} \sqrt{\mathbf{1}_m' \boldsymbol{P}_{m,m} \mathbf{1}_m}}, \tag{7}$$

where $\mathbf{1}_k$ is the $k \times 1$ vector of ones and $\boldsymbol{P}_{n,m}$ is the $n \times m$ matrix given by

$$\boldsymbol{P}_{n,m} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{m-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{m-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{m-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \dots & \rho_{|m-n|} \end{bmatrix}, \tag{8}$$

where ρ_i is the autocorrelation of order i of the $AR(p)$ process for returns. Under the random walk (where $\rho_i = 0$ for all $i > 0$), the formula for the correlation between two trading indicators $MOM(n)$ and $MOM(m)$ reduces to

$$Cor_{RW}(MOM(n), MOM(m)) = \frac{\min(n, m)}{\sqrt{n\,m}}. \tag{9}$$

Zakamulin and Giner (2020) prove that the correlation coefficient $Cor(MOM(n), MOM(m))$ increases with increasing persistence of the $AR(p)$ process for returns. Put differently, the correlation coefficient is lowest when the returns follow the random walk. The higher the return persistence α is, the larger the correlation coefficient between two trading indicators.

The formal description of the novel methodology to demonstrate the return persistence is as follows. First, we estimate the empirical correlation coefficient between the trading indicators of two TSMOM rules $Cor_{EMP}(MOM(n), MOM(m))$. Then we test whether the empirical correlation coefficient is statistically significantly higher than the correlation coefficient under the random walk $Cor_{RW}(MOM(n), MOM(m))$. For this purpose we formulate and conduct the test of the following null hypothesis

$$H_0 : Cor_{EMP}(MOM(n), MOM(m)) \leq Cor_{RW}(MOM(n), MOM(m)).$$

It is worth noting that under the null hypothesis the returns follow a random walk and the empirically estimated correlation coefficient is not greater than the true correlation under the random walk. Since under the null there is no dependence in return series, in order to conduct the test of the null hypothesis we employ the randomization method. In a nutshell, randomization consists of reshuffling the data to destroy any dependence and then recalculating the test statistics for each reshuffling in order to estimate its distribution under the null hypothesis of no dependence.

To be more specific, the estimation of the p-value of the test is conducted as follows. To learn the sampling distribution for $Cor_{RW}(MOM(n), MOM(m))$, we randomize the original excess return series. This is repeated 1,000 times, each time obtaining a new estimate for $Cor_{RW}(MOM(n), MOM(m))^*$.[6] In the end, to estimate the significance level, we count how many times the estimated value for $Cor_{RW}(MOM(n), MOM(m))^*$ after randomization falls above the value of the actual estimate for $Cor_{EMP}(MOM(n), MOM(m))$. In other words, under the null hypothesis we compute the probability of obtaining a more extreme value for the correlation coefficient than the actual estimate.

Once we reject the null and establish that $Cor_{EMP}(MOM(n), MOM(m))$ is greater than $Cor_{RW}(MOM(n), MOM(m))$, we can compute the implied return persistence. The notion of "implied return persistence" is motivated by the notion of implied volatility in option prices. In our context, the implied return persistence is the sum of the autoregressive coefficients which, when input in formula (7) for the correlation coefficient between two trading indicators, will return the empirically estimated value of the correlation coefficient. Specifically, when the returns follow a specific $AR(p)$ process, the correlation coefficient between two trading indicators is given by equation (7). The idea is to note that the correlation coefficient is the function of the return lags n and m and the autoregressive coefficients of the $AR(p)$ process

$$Cor(MOM(n), MOM(m)) = f(n, m, \phi_1, \ldots, \phi_p). \tag{10}$$

Even if we assume that all autoregressive coefficients are alike, the problem is that the implied alpha is not unique because it depends not only on the value of $Cor(MOM(n), MOM(m))$, but also on the order of the autoregressive process p. Since we expect that the number of autoregressive terms $p \in [6, 12]$, we compute the implied alpha under assumption that $p = 9$. The motivation is that $p = 9$ lies in the middle of the range for the expected value of p. It is generally not possible to invert formula (10) so that the implied α is expressed as a function of $Cor(MOM(n), MOM(m))$, n, m, and ϕ_i. However, the implied persistence can easily be computed using, for example, an iterative search procedure.

[6] The asterisk is used to indicate that each of these estimates is calculated on a randomized sample.

All correlations in our study are estimated using a highly robust covariance (and correlation) estimation method suggested by Rousseeuw (1984) and further developed by Rousseeuw (1985). The covariance is estimated using the minimum covariance determinant (MCD) method. The MCD method is highly resistant to outliers and can be thought of as estimating the covariance using the "good" part of the data. Specifically, in this method it is the volume of the Gaussian confidence ellipsoid, equivalently the determinant of the classical covariance matrix, that is minimized. The Mahalanobis distances of all the points from the initial location estimate for the covariance matrix are calculated, and those points within the 97.5% point under Gaussian assumptions are declared to be good. The final estimates are obtained by computing the covariance of the good points. The problem is that the exact MCD method is extremely time consuming. In our study, we rely on the FAST-MCD method developed by Rousseeuw and Driessen (1999).

In estimating the correlation coefficient between two trading indicators $MOM(n)$ and $MOM(m)$, we select n and m in order to maximize the potential difference between $Cor_{RW}(MOM(n), MOM(m))$ and $Cor_{EMP}(MOM(n), MOM(m))$. Roughly, this difference is maximized when m is about twice as small as n. However, the choice of n is somewhat arbitrary. Since our focus is on the short-term persistence in excess returns, the number of return lags n should not exceed 12. Therefore, we select three values for n, in particular, $n = \{8, 9, 10\}$.

Table 2 reports the estimated correlations $Cor_{EMP}(MOM(n), MOM(m))$ and the results of testing the null hypothesis for various choices for n and m using the total sample of data as well as the data for the first and second halves of the sample. In sum, the results reported in this table advocate that, regardless of the choice of the sample period and the values of n and m, the empirical correlation between the trading indicators $Cor_{EMP}(MOM(n), MOM(m))$ is statistically significantly higher than the correlation under the random walk $Cor_{RW}(MOM(n), MOM(m))$.

For the first half of the sample, the results in Table 2 on estimates of the implied return persistence based on using the empirical correlation between two trading indicators agree with the results in Table 1 on the estimates of the return persistence based on using the sum of the estimated autoregressive coefficients. In contrast, the results for the total sample and the second half of the sample differ remarkably. In particular, the results in Table 1 suggest

Parameter	1857-2018 Estimate	1857-2018 P-value	1857-1937 Estimate	1857-1937 P-value	1938-2018 Estimate	1938-2018 P-value
Panel A: Results for $n = 8$ and $m = 4$.						
Cor_{RW}	0.707		0.707		0.707	
Cor_{EMP}	**0.767**	0.000	**0.773**	0.005	**0.764**	0.024
Implied α	0.327		0.354		0.310	
Panel B: Results for $n = 9$ and $m = 4$.						
Cor_{RW}	0.667		0.667		0.667	
Cor_{EMP}	**0.744**	0.000	**0.716**	0.032	**0.745**	0.003
Implied α	0.354		0.239		0.358	
Panel C: Results for $n = 10$ and $m = 5$.						
Cor_{RW}	0.707		0.707		0.707	
Cor_{EMP}	**0.772**	0.001	**0.759**	0.039	**0.775**	0.010
Implied α	0.299		0.244		0.312	

Table 2: Detection of persistence in excess returns and measuring the persistence using the empirical correlation between the trading indicators of the $MOM(n)$ and $MOM(m)$ rules. Cor_{RW} denotes the theoretical correlation between the trading indicators under the random walk. Cor_{EMP} denotes the empirical estimate of the correlation between the trading indicators. Implied α denotes the implied persistence of the $AR(p)$ process for excess returns under assumption that $\phi_i = \phi$ and $p = 9$. Bold text highlights the values that are statistically significant at the 5% level.

the presence of a weak return persistence in the first half of the sample and a substantially weaker return persistence in the total sample and the second half of the sample. In contrast, the results in Table 2 argue that over the total sample and the second half of the sample the return persistence was actually a bit stronger than that in the first half of the sample.

It is important to emphasize that the results reported in Table 2 reveal the evidence of persistence in the excess returns on the S&P Composite stock price index. Specifically, if the excess returns follow the $AR(p)$ process, we have evidence that $\alpha = \sum_{i=1}^{p} \phi_i > 0$. Unfortunately, the test for return persistence does not allow us to estimate the order of autoregression p and the values of the autoregressive coefficients ϕ_i. Anyway, the results of this novel test for persistence are highly interesting because they allow us to evaluate the value of ϕ given the order p.[7]

Table 3 reports the value of the autoregressive coefficients ϕ and the implied $\alpha = \phi p$ for each $p \in [1, 12]$. Specifically, assuming that $Cor(MOM(10), MOM(5)) = 0.772$ (the estimated correlation using the data for the total sample) and using an iterative search procedure, we

[7]We remind the reader that Conjecture 1 allows us to narrow the number of the unknown parameter down to two: p and ϕ.

Order p	Coefficient ϕ	Implied α
1	0.5348	0.5348
2	0.2083	0.4167
3	0.1159	0.3477
4	0.0767	0.3068
5	0.0553	0.2765
6	0.0448	0.2686
7	0.0388	0.2719
8	0.0353	0.2827
9	0.0333	0.2995
10	0.0322	0.3221
11	0.0312	0.3433
12	0.0303	0.3631

Table 3: Assuming that $Cor(MOM(10), MOM(5)) = 0.772$, this table reports the value of the autoregressive coefficients ϕ and the implied α for each order $p \in [1, 12]$.

compute numerically the value of ϕ for each given p. We remind the reader that it is reasonable to assume that the number of autoregressive terms p lies somewhere in between 6 and 12. Even though we are not able to evaluate the order p exactly, the results reported in Table 3 suggest that $\phi \in [0.0303, 0.0448]$. Roughly, this means that we expect the value of ϕ to be about of the same largeness as that of the standard error of estimation of ϕ using the OLS regression model, see Table 1. This conclusion explains the reason for why we do not see statistically significant ϕ_i coefficients when we rely on the OLS methodology. This is because in order to detect statistical significance at the 5% level, a coefficient in an OLS model must be more than about twice as large as the standard error of estimation of this coefficient.

5.2 A Methodology for Joint Evaluation of Model Parameters

The results reported in the previous section reveal the evidence of persistence in the excess returns on the S&P Composite stock price index. Additionally, these results allowed us to evaluate the magnitude of the autoregressive coefficients in the $AR(p)$ model for excess returns. It turns out that over monthly horizons the autocorrelation in excess returns is very weak and hence escapes detection. Therefore, a reliable detection of short-term persistence in excess returns is only possible using excess returns aggregated over multiple months. A similar idea was put forward already by Fama and French (1988) who, in order to detect the presence of weak mean-reversion in stock prices, suggested using the first-order autocorrelation of returns aggregated over multiple periods. In the context of our study, a relevant statistics of interest is the first-order autocorrelation of k-month excess returns:

$$Cor(X_{t+k,t+1}, X_{t,t-k+1}) \tag{11}$$

where

$$X_{t+k,t+1} = \sum_{i=1}^{k} X_{t+i}, \quad X_{t,t-k+1} = \sum_{i=1}^{k} X_{t-k+i}.$$

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

$$Cor(X_{t+k,t+1}, X_{t,t-k+1}) = \frac{\mathbf{1}_k' \boldsymbol{Q}_{k,k} \mathbf{1}_k}{\mathbf{1}_k' \boldsymbol{P}_{k,k} \mathbf{1}_k}, \tag{12}$$

where $\mathbf{1}_k$ is the $k \times 1$ vector of ones, $\boldsymbol{P}_{k,k}$ is the $k \times k$ matrix given by equation (8) and $\boldsymbol{Q}_{k,k}$ is the $k \times k$ matrix given by

$$\boldsymbol{Q}_{k,k} = \begin{bmatrix} \rho_k & \rho_{k+1} & \rho_{k+2} & \cdots & \rho_{2k-1} \\ \rho_{k-1} & \rho_k & \rho_{k+1} & \cdots & \rho_{2k-2} \\ \rho_{k-2} & \rho_{k-1} & \rho_k & \cdots & \rho_{2k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_1 & \rho_2 & \rho_3 & \cdots & \rho_k \end{bmatrix}, \tag{13}$$

where ρ_i is the autocorrelation of order i of the process for X_t.

The proof is given in the Appendix.

Figure 1 plots the shape of the first-order autocorrelation function of k-month excess returns under the assumption that the monthly excess returns follow the $AR(p)$ process where $\phi_i = \phi$ for all $i \in [1, p]$. For each order p, the value of the autoregressive coefficient ϕ is taken from Table 3. This means that, for each order $p \in [6, 12]$, Figure 1 plots the theoretical shape of the first-order autocorrelation of k-month excess returns that follow the $AR(p)$ process which induces the correlation between the MOM(10) and MOM(5) trading indicators that corresponds to the empirically estimated correlation of 0.772.

Our first observation is that each curve in Figure 1 is a positively skewed bell-shaped curve with a clearly evident top. The location of the top is determined by the order of the $AR(p)$ process. 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 $k < p$. Our second observation is

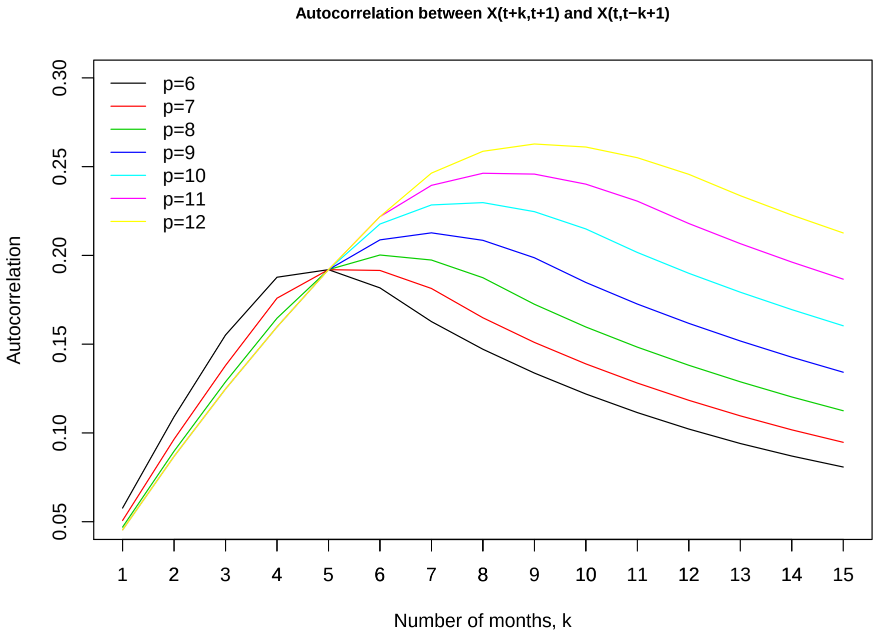

Figure 1: The shapes of the $Cor(X_{t+k,t+1}, X_{t,t-k+1})$ when X_t follows the $AR(p)$ process where $\phi_i = \phi$ for all $i \in [1, p]$ and $p \in [6, 12]$. For each order p, the value of the autoregressive coefficient ϕ is taken from Table 3.

that the value of the $Cor(X_{t+k,t+1}, X_{t,t-k+1})$ at the top monotonically increases as the order p of the autoregressive process increases. In addition, our numerical experiments indicate that the value of the $Cor(X_{t+k,t+1}, X_{t,t-k+1})$ at the top monotonically increases as the value of the autoregressive coefficient ϕ increases. In sum, the parameters ϕ and p of the autoregressive process for excess returns define the shape of the first-order autocorrelation function of k-month excess returns.

The observations presented in the preceding paragraph suggest an indirect approach to the joint evaluation of the parameters ϕ and p of the $AR(p)$ process for excess returns. Specifically, the idea is to evaluate ϕ and p by fitting the theoretical shape of the correlation $Cor(X_{t+1,t+k}, X_{t,t-k+1})$ to the empirically estimated shape. For this purpose we solve the following problem:

$$\min_{p,\phi} \sum_{k=k_{min}}^{k=k_{max}} \Big| Cor(X_{t+1,t+k}, X_{t,t-k+1}, p, \phi) - Cor_{EMP}(X_{t+1,t+k}, X_{t,t-k+1}) \Big|, \tag{14}$$

where $Cor_{EMP}(X_{t+1,t+k}, X_{t,t-k+1})$ is the empirically estimated autocorrelation function of k-month excess returns and $Cor(X_{t+1,t+k}, X_{t,t-k+1}, p, \phi)$ is the theoretical correlation function of k-month excess returns given some specific values for p and ϕ. That is, the parameters ϕ and p 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.

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 length 14 months. Therefore, as in Fama and French (1988), in order 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 model, we estimate the autocorrelations using a highly robust to outliers correlation estimation method developed by Rousseeuw and Driessen (1999).

After the estimation of $Cor(X_{t+1,t+k}, X_{t,t-k+1})$ and before attacking the problem of the joint evaluation of the parameters ϕ and p of the $AR(p)$ process for excess returns, we want to make sure that the estimated first-order autocorrelations are statistically significantly positive. For this purpose we formulate and test the following null hypothesis:

$$H_0 : Cor(X_{t+1,t+k}, X_{t,t-k+1}) \leq 0.$$

This test is motivated by the notion that if the excess returns are independent and identically distributed, then the first-order autocorrelation function is zero irrespective of the number of months, k. In other words, absent persistence in the excess returns, there is no (positive) correlation between two successive non-overlapping k-month excess returns. The null hypothesis of a random walk is rejected in favor of persistence in the excess returns if the first-order autocorrelation is significantly above zero. Note that this test complements the novel test presented in the preceding section and reinforces the evidence of persistence in the excess returns on the S&P Composite stock price index. Besides, by comparing and contrasting the results from two different tests, we can check for consistency in the evaluated parameters of the $AR(p)$ process for excess returns.

As in the preceding section, in order to conduct the test of the null hypothesis we employ the randomization method. In particular, to estimate the sampling distribution for

$Cor(X_{t+1,t+k}, X_{t,t-k+1})$ under the null, we randomize the original excess return series. This is repeated 1,000 times, each time obtaining a new estimate for $Cor(X_{t+1,t+k}, X_{t,t-k+1})^*$. In the end, to estimate the significance level, we count how many times the estimated value for $Cor(X_{t+1,t+k}, X_{t,t-k+1})^*$ after randomization falls above the value of the actual estimate for $Cor(X_{t+1,t+k}, X_{t,t-k+1})$. That is, under the null hypothesis of random walk, we compute the probability of obtaining a more extreme value for the autocorrelation coefficient than the actual estimate.

The use of overlapping returns leads to some potentially serious econometric issues which are commonly termed as "small-sample bias". In particular, it is known that an estimate of autocorrelation obtained using overlapping blocks of data is downward biased. The randomization method allows us not only to conduct the test of the null hypothesis, but also to estimate the bias and conduct the bias correction.[8] The idea behind the bias correction is as follows. Since the true value of the autocorrelation is zero under the null hypothesis, the bias correction is done by subtracting the mean value of $Cor(X_{t+1,t+k}, X_{t,t-k+1})^*$ from the empirical estimate for $Cor(X_{t+1,t+k}, X_{t,t-k+1})$. That is, the bias adjusted values of the first-order autocorrelation of k-month excess returns are computed as $Cor(X_{t+1,t+k}, X_{t,t-k+1}) - E[Cor(X_{t+1,t+k}, X_{t,t-k+1})^*]$.

For the total sample as well as the first and second half of the sample, Figure 2 plots the estimated first-order autocorrelation function of k-month excess returns on the S&P Composite stock price index. As expected, the first-order autocorrelation of k-month excess returns first increases, attains a maximum, and then decreases. The main discrepancy between the model-implied autocorrelation function depicted in Figure 1 and the sample autocorrelation function plotted in Figure 2 concerns the behavior of the first-order autocorrelation function for the excess returns aggregated over periods longer than about 12 months. Specifically, whereas the model-implied autocorrelation function decreases gradually toward zero as the aggregation period k increases, the sample autocorrelation function becomes eventually negative when k increases. This behavior of the sample 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.

[8] For a similar approach to the bias correction, see Fama and French (1988), Kim, Nelson, and Startz (1991), and Nelson and Kim (1993).

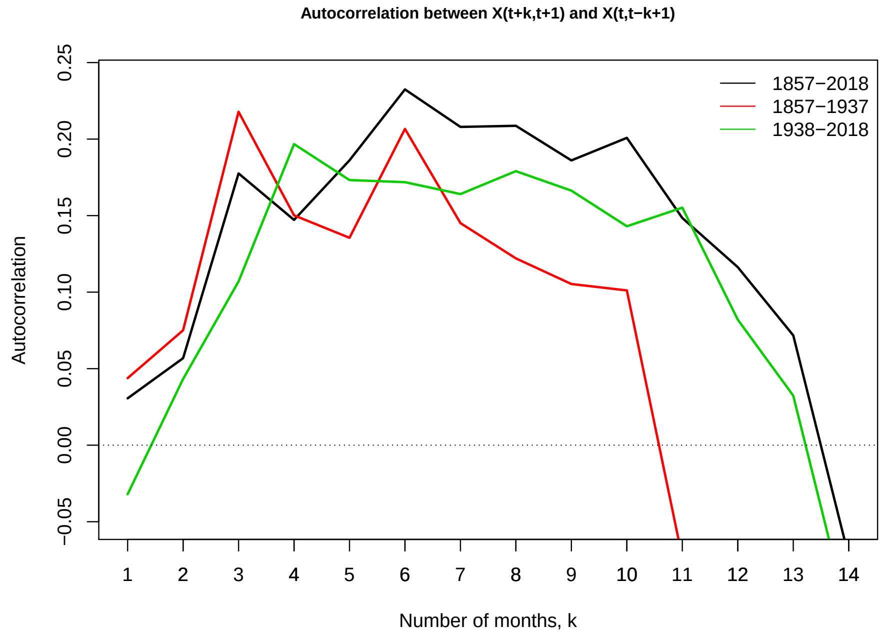

Figure 2: For the total sample as well as the first and second half of the sample, this figure plots the shape of the empirically estimated first-order autocorrelation function of k-month excess returns, $Cor(X_{t+1,t+k}, X_{t,t-k+1})$, on the S&P Composite stock price index.

Table 4 reports the estimated first-order autocorrelations and the results of testing the null hypothesis using the total sample of data as well as the data for the first and second halves of the sample. The results reported in this table exhibit clear evidence against the random walk in favor of a short-term persistence in the excess returns. Specifically, over the total sample the first-order autocorrelation values are statistically significantly greater than zero at the 5% level over periods of $k \in [2, 12]$ months. For the first (second) half of the sample, the first-order autocorrelation values are statistically significantly positive at the 5% level for periods of $k \in [2, 7]$ ($k \in [3, 11]$) months.

Provided the evidence that the excess returns do not follow a random walk but rather an $AR(p)$ process, we jointly evaluate the parameters ϕ and p of the $AR(p)$ process for excess returns by fitting the theoretical shape of the correlation $Cor(X_{t+1,t+k}, X_{t,t-k+1})$ to the empirically estimated shape according to the optimization procedure given by equation (14). The optimization procedure finds the parameters ϕ and p that produce the best fit over the range $k_{min} = 1$ month to $k_{max} = 13$ moths. The motivation for limiting the maximum value for k to 13 months is to confine our attention solely to the short-term persistence effect and overlook the effect of the subsequent medium-term reversion to the mean.

	1857-2018		1857-1937		1938-2018	
Period k	**Estimate**	**P-value**	**Estimate**	**P-value**	**Estimate**	**P-value**
1	0.031	0.241	0.044	0.208	-0.032	0.754
2	**0.057**	0.085	**0.075**	0.078	0.043	0.204
3	**0.178**	0.000	**0.218**	0.000	**0.107**	0.052
4	**0.147**	0.001	**0.150**	0.018	**0.197**	0.004
5	**0.186**	0.000	**0.136**	0.040	**0.173**	0.007
6	**0.232**	0.000	**0.207**	0.007	**0.172**	0.020
7	**0.208**	0.000	**0.145**	0.046	**0.164**	0.032
8	**0.209**	0.000	0.122	0.119	**0.179**	0.042
9	**0.186**	0.003	0.105	0.142	**0.166**	0.055
10	**0.201**	0.003	0.101	0.178	**0.143**	0.095
11	**0.148**	0.033	-0.073	0.744	**0.155**	0.081
12	**0.116**	0.071	-0.100	0.794	0.082	0.234
13	0.072	0.185	-0.126	0.856	0.032	0.394
14	-0.073	0.811	-0.134	0.876	-0.113	0.802

Table 4: 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 text indicates values that are statistically significant at the 5% level. The p-values of the hypothesis test $H_0 : Cor(X_{t+1,t+k}, X_{t,t-k+1}) \leq 0$ are computed using the randomization method.

Table 5 documents the estimated best-fit parameters p and ϕ of the $AR(p)$ process for the excess returns on the S&P Composite stock price index for the total sample as well as the first and second halves of the sample. The results reported in this table suggest that over the fist half of the sample the excess returns followed the $AR(p)$ process with $p = 6$ and $\phi = 0.0368$, whereas over the second half of the sample the parameters of the process have been $p = 8$ and $\phi = 0.0322$. Over the total sample, the estimated parameters of the autoregressive process for excess returns are $p = 9$ and $\phi = 0.0324$. Note that in this case the estimated value of ϕ closely corresponds to the implied value of ϕ when $p = 9$ in Table 3. Consequently, our method of the joint evaluation of the parameters ϕ and p of the $AR(p)$ process for excess returns produces the estimates that are consistent with the results reported in the previous section. For the sake of illustration, using the data for the whole sample Figure 3 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 $\phi = 0.0324$ that produce the best fit to the empirical data.

Note that our methodology for the joint evaluation of the model parameters does not allow us to estimate the confidence intervals for each parameter. Later in the paper we will argue that the main parameter of interest is the trend strength α. This is because the trend strength

Period	Order p	Coefficient ϕ	Persistence α
1857-2018	9	0.0324	0.2916
1857-1937	6	0.0368	0.2208
1938-2018	8	0.0322	0.2576

Table 5: Estimated best-fit parameters p and ϕ of the $AR(p)$ process for the excess returns on the S&P Composite stock price index. α denotes the measure of persistence of the process for excess returns, $\alpha = p\,\phi$.

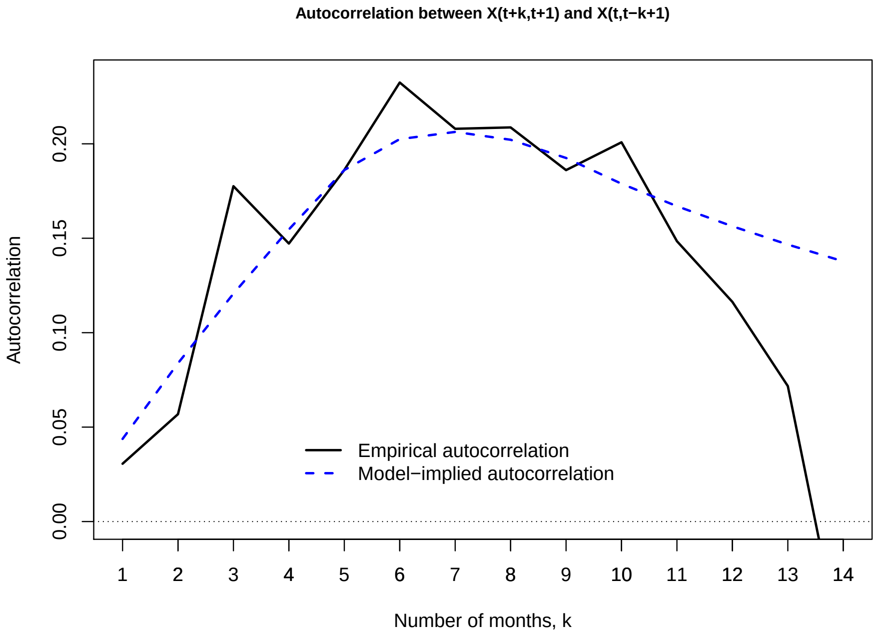

Figure 3: Using the data for the whole sample, this figure plots the shape of the empirically estimated first-order autocorrelation function of k-month excess returns and the shape of the theoretical first-order autocorrelation function with the parameters $p = 9$ and $\phi = 0.0324$ that produce the best fit to the empirical data.

is the main driving factor that determines the success of the TSMOM strategy. Therefore, it is important to know the error bounds for the point estimate of α.

To learn the sampling distribution of the estimated α, we rely on a parametric bootstrap method. For this purpose, we resample with replacement the original excess returns series to destroy the serial dependence in data. Then, using the estimated parameters p and ϕ, the resampled series are employed to simulate a sample from the $AR(p)$ process for the excess returns with known trend strength α. Finally, the artificially simulated series of excess returns are used as input in the procedure for the joint evaluation of the parameters ϕ^* and p^* of the $AR(p)$ process. This sequence of steps is repeated 1,000 times, each time obtaining a new estimate for $\alpha^* = \phi^* p^*$.

Using the data for the whole sample that covers the period from 1857 to 2018, Figure 4 plots the estimated density of the sampling distribution of α using a Gaussian kernel smoother. The vertical black dashed lines show the location of the 95% confidence interval for α. These boundaries are calculated by taking the 0.025th and 0.975th empirical quantiles of the sampling distribution of α. The estimated lower and upper boundaries are 0.128 and 0.437 respectively; our point estimate is $\alpha = 0.2916$. These numbers (together with the visual observation of the graph plotted in Figure 4) suggest that the sampling distribution of α is non-normal: the left tail is a bit longer than the right tail. Therefore, the boundaries of the 95% confidence interval can be written as:

$$\text{Lower bound} = \text{point estimate} - 0.163,$$

$$\text{Upper bound} = \text{point estimate} + 0.145,$$

where 0.163 and 0.145 are the lower and upper error bounds respectively.

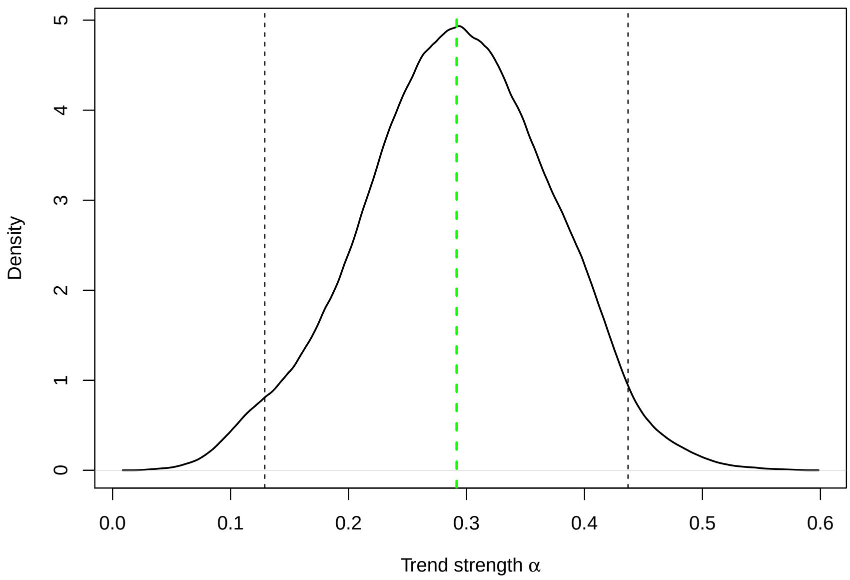

Figure 4: Kernel-smoothed density of the sampling distribution of α. The vertical green dashed line shows the location of the estimated $\alpha = 0.2916$. The vertical black dashed lines show the location of the 95% confidence interval for α.

It is worth emphasizing that our bootstrap method answers the following question: Given that the true trend strength $\alpha = 0.2916$, what are the error bounds on the estimated trend strength? In contrast, we would like the answers to the following two questions. The first question is: what is the value of α such that the probability of observing $\alpha > 0.2916$ equals 2.5%? The second question is: what is the value of α such that the probability of observing $\alpha < 0.2916$ equals 2.5%? Assuming that the lower and upper error bounds on α do not depend on the value of α,[9] these bounds are 0.146 and 0.454 respectively. Finally, assuming that $p = 9$, the implied lower and upper bounds on ϕ are 0.0163 and 0.0505 respectively.

	Persistence α		**Coefficient ϕ**	
Period	**Lower bound**	**Upper bound**	**Lower bound**	**Upper bound**
1857-2018	0.146	0.454	0.0163	0.0505
1857-1937	0.029	0.413	0.0048	0.0687
1938-2018	0.051	0.479	0.0064	0.0598

Table 6: Estimated lower and upper boundaries of the 95% confidence interval for the trend strength α and the value of the autoregressive coefficient ϕ. The estimations are obtained using a parametric bootstrap method. The lower and upper bounds on ϕ are estimated under the assumption that the number of autoregressive terms p in the $AR(p)$ process for the excess returns equals the number of estimated terms.

In the similar manner, using the data for the first and second halves of the sample, we compute the lower and upper bounds for the trend strength α and the value of the autoregressive coefficient ϕ. Table 6 reports the estimated lower and upper boundaries of the 95% confidence interval for α and ϕ. Two observations deserve mentioning. First, the 95% confidence interval is narrower when the data for the whole sample are used, indicating less error and greater precision of estimation. Second, there are no any sings that, for example, the trend strength α has increased or decreased over time. Therefore, we cannot reject the hypothesis that the trend strength has been stable over the whole sample.

6 Theoretical Implications of Time-Series Momentum

Using the estimated parameters of the $AR(p)$ process for the excess returns over the total historical sample of data,[10] in this section we present and discuss a number of theoretical implications of time-series momentum. In particular, in the subsequent section we present analytical results on the one-period performance of the TSMOM strategy. Subsequently, we examine the one- and multi-period properties of the TSMOM strategy by means of a simulation study. After that, by relying again on a simulation analysis, we explore how the evidence of

[9] This assumption is supported by our simulation experiments.

[10] We use the data for the whole sample because they provide a greater precision of estimation of the model parameters.

superior performance of the TSMOM strategy depends on the investment horizon and evaluate the power of statistical tests. Finally, we study the precision of identification of the optimal number of lags in the TSMOM rule using a standard back-testing methodology.

6.1 Mean, Variance, and Sharpe ratio of TSMOM Strategy

In this section, we present analytical solutions for the one-period mean and variance of returns of the TSMOM strategy, as well as for the Sharpe ratio of this 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

$$\begin{bmatrix} r_t \\ MOM_{t-1}(n) \end{bmatrix} = \mathcal{N}\left(\begin{bmatrix} \mu \\ m \end{bmatrix}, \begin{bmatrix} \sigma^2 & \rho_m \sigma v \\ \rho_m \sigma v & v^2 \end{bmatrix}\right), \tag{15}$$

where μ and σ^2 are the mean and variance of r_t, m and v^2 are the mean and variance of $MOM_{t-1}(n)$, and ρ_m is the correlation coefficient between r_t and $MOM_{t-1}(n)$. The mean and variance of $MOM_{t-1}(n)$ are given by

$$m = n\mu_x, \quad v^2 = \mathbf{1}_n' \boldsymbol{P}_{n,n} \mathbf{1}_n \sigma_x^2, \tag{16}$$

where μ_x and σ_x^2 are given by (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 (8). The correlation coefficient is computed as (see Zakamulin and Giner (2020))

$$\rho_m = \frac{\mathbf{1}_n' \boldsymbol{P}_{n,p}\, \boldsymbol{\phi}_p}{\sqrt{\mathbf{1}_n' \boldsymbol{P}_{n,n} \mathbf{1}_n}}, \tag{17}$$

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 (8).

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

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

$$Var[R_t^{LO}] = (\mu^2 + r_f^2)\Phi(-d) + g(2\mu + \sigma\rho_m d) + r_f^2 \Phi(d) - \left[(\mu - r_f)\Phi(-d) + r_f + g\right]^2, \tag{19}$$

where

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

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 and variance of returns of the long-short TSMOM strategy are given by

$$E[R_t^{LS}] = (2\Phi(-d) - 1)\mu + 2(g + \Phi(d) r_f), \tag{20}$$

$$Var[R_t^{LS}] = (\mu^2 + \sigma^2) + 4r_f(g - (\mu - r_f)\Phi(d)) - \left[(2\Phi(-d) - 1)\mu + 2(g + \Phi(d) r_f)\right]^2. \tag{21}$$

The proof is given in the Appendix.

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

Remark 2. Given the expressions for the mean and variance of returns to the long-only and long-short TSMOM strategies, one can easily compute the Sharpe ratio of each strategy

$$SR_{LO} = \frac{E[R_t^{LO}] - r_f}{\sqrt{Var\left[R_t^{LO}\right]}}, \quad SR_{LS} = \frac{E[R_t^{LS}] - r_f}{\sqrt{Var\left[R_t^{LS}\right]}}, \tag{22}$$

where SR_{LO} and SR_{LS} denote the Sharpe ratios of the long-only and long-short TSMOM strategy respectively.

For the sake of illustration, using the data for the whole sample that covers the period from 1857 to 2018, we estimate the monthly parameters[11] μ, σ, and r_f and compute the theoretical mean, standard deviation, and Sharpe ratios of the buy-and-hold strategy, the long-only $MOM(n)$ strategy, and the long-short $MOM(n)$ strategy. We assume that the excess market returns $r_t - r_f$ follow the $AR(p)$ process with $p = 9$ and $\phi = 0.0324$; these are the estimated best-fit parameters reported in the previous section. We further assume that the window size in the TSMOM rule equals $n = p$; this is the optimal window size that produces the highest correlation ρ_m. Note that under our assumptions $\mu_x = \mu - r_f$ and $\sigma_x = \sigma$.

[11] The risk-free rate of return is estimated as the mean risk-free rate of return over the whole sample.

Table 7 reports the descriptive statistics of the buy-and-hold strategy, the long-only $MOM(9)$ strategy, and the long-short $MOM(9)$ strategy. In particular, this table reports theoretical monthly means, standard deviations, and Sharpe ratios of each strategy under investigation. The descriptive statistics of the long-short TSMOM strategy are virtually identical to those of the buy-and-hold strategy. In contrast, the long-only TSMOM strategy has about the same mean monthly return as that of the buy-and-hold strategy, but with a substantially lower standard deviation. As a consequence, the monthly Sharpe ratio of the long-only TSMOM strategy is considerably above that of the buy-and-hold strategy. Namely, the Sharpe ratio of the long-only strategy is about 30% higher than the Sharpe ratio of the buy-and-hold strategy. The results reported in this table agree very well with the published results. Specifically, numerous papers report that the long-only TSMOM strategy outperforms the buy-and-hold strategy in both in- and out-of-sample tests. Using virtually the same historical data set, Zakamulin (2017, Chapter 9) finds that, after accounting for realistic transaction costs, the long-short TSMOM strategy does not outperform the buy-and-hold strategy even in in-sample tests.

	Strategy		
	Buy and hold	**Long-only**	**Long-short**
Mean return, %	0.856	0.864	0.872
Std. deviation, %	5.024	3.930	5.022
Sharpe ratio	0.107	0.139	0.110

Table 7: The theoretical monthly mean return, standard deviation, and the Sharpe ratio of the buy-and-hold strategy, the long-only $MOM(9)$ strategy, and the long-short $MOM(9)$ strategy. The model parameters μ, σ, r_f, p, and ϕ are estimated using the data on the S&P Composite index and the risk-free rate of return over the whole sample 1857-2018.

The mean, variance, and Sharpe ratio of a TSMOM strategy depends on the correlation coefficient, ρ_m, between the trading indicator $MOM_{t-1}(n)$ and the next period return r_t. The correlation coefficient, in its turn, depends on the value of the autoregressive coefficient ϕ. Since there is uncertainty about the true value of ϕ, Figure 5 plots how the Sharpe ratio of the long-only and long-short MOM(9) strategy depends on the value of ϕ in the $AR(9)$ process for excess returns.

On the basis of the results reported in Figure 5, the following observations can be made. First, when the excess market returns follow a random walk, $\phi = 0$, the correlation coefficient $\rho_m = 0$ and the TSMOM strategy underperforms the buy-and-hold strategy. For example, in

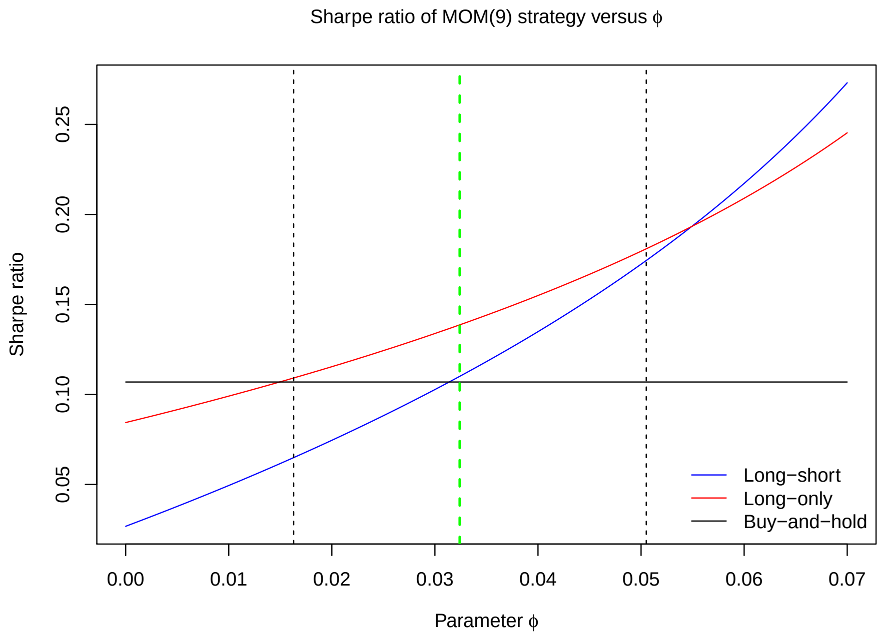

Figure 5: The theoretical Sharpe ratio of the buy-and-hold strategy, the long-only $MOM(9)$ strategy, and the long-short $MOM(9)$ strategy versus the value of ϕ in the $AR(p)$ process for the excess returns. The parameters μ, σ, r_f, and p are estimated using the data for the whole sample. The vertical green dashed line show the location of the estimated value of $\phi = 0.0324$. The vertical black dashed lines show the location of the 95% confidence interval for ϕ under assumption that $p = 9$.

this case the mean return to the long-only strategy is given by $E[R_t^{LO}] = (\mu - r_f)\Phi(-d) + r_f$. Since $\Phi(-d) < 1$, the mean return to the long-only strategy is smaller than the mean market return. In this case the variance of returns to the long-only strategy is also smaller than the variance of the market returns. The combined impact of these two effects is to decrease the Sharpe ratio of the long-only strategy to a level below the Sharpe ratio of the buy-and-hold strategy. Only when the market exhibits very strong upward drift, the value of $\Phi(-d)$ approaches unity and, as a consequence, the Sharpe ratio of the long-only strategy approaches the Sharpe ratio of the buy-and-hold strategy.

Second, when the autoregressive coefficient ϕ increases, both the trend strength $\alpha = p\phi$ and the correlation ρ_m increases. As a result, the Sharpe ratio of the TSMOM strategy increases. For the long-only strategy, the break-even value of ϕ, at which the Sharpe ratio of the long-only strategy equals that of the buy-and-hold strategy, amounts to $\phi = 0.0149$. The break-even value is about twice as small as the estimated value of ϕ (0.0324) and just a bit below than the lower boundary of the 95% confidence interval for ϕ (0.0163).[12] For the long-short strategy, the

[12]We remind the reader that the lower and upper bounds on ϕ are computed under the assumption that

break-even value of ϕ amounts to $\phi = 0.0314$ which is roughly equal to the estimated value of ϕ. Third, in order the long-short TSMOM strategy outperforms the long-only TSMOM strategy, the trend must be rather strong. In our case, the value of the autoregressive parameter ϕ must exceed 0.0549 to make the long-short TSMOM strategy worthwhile. Roughly, the market trend strength must be twice as strong as the estimated trend strength to make the long-short TSMOM strategy superior to the long-only TSMOM strategy. However, the value 0.0549 is above the upper boundary of the 95% confidence interval for ϕ (0.0505). Thus, we can say with a very high degree of confidence that the long-short TSMOM strategy is inferior to the long-only TSMOM strategy.

6.2 Probability Distribution of Returns to TSMOM Strategy

6.2.1 Probability Distribution of One-Period Returns

In the preceding section we presented analytical results for the one-period mean and variance of returns to the TSMOM strategy. The mean return is the average of all possible values of the probability distribution of returns. The variance of returns is the average value of squared difference between each value and the mean of the distribution. In essence, the mean and variance describe the first two moments of a probability distribution: the center and the spread. Yet, the first two moments are not enough to characterize the probability distribution function of returns to the TSMOM strategy because its shape deviates significantly from the shape of a normal distribution.

To learn the shape of the probability distribution function of one-period returns to the realistic TSMOM strategy, we rely on a simulation method. In particular, using the data for the whole sample, we estimate the monthly parameters μ and σ 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 $\phi = 0.0324$. The parameters c and σ_ε of the $AR(p)$ process for the excess returns are computed using formulas given by (6).

First, we simulate a time-series of monthly returns to the buy-and-hold strategy of length $N = 1,000,000$. Then, using the returns to the buy-and-hold strategy and the risk-free rate of return, we simulate the trading signal to the MOM(9) rule and subsequently the returns

$p = 9$, see the preceding section.

to the long-only and long-short TSMOM strategies. Specifically, the returns to the long-only strategy are given by (1), whereas the returns to the long-short strategy are given by (2). Figure 6 plots the estimated densities of returns to the buy-and-hold strategy and the long-only TSMOM strategy using a Gaussian kernel smoother.

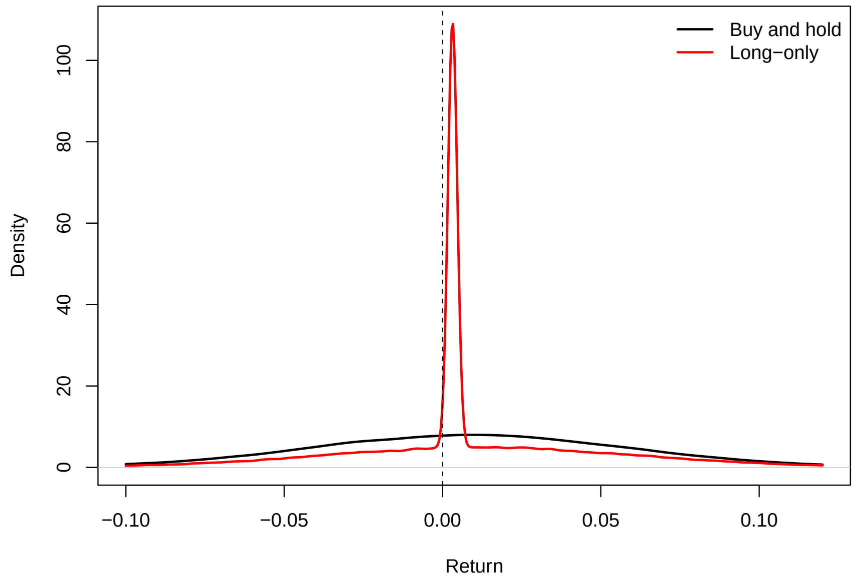

Figure 6: Kernel-smoothed densities of one-month returns to the buy-and-hold strategy and the long-only $MOM(9)$ strategy.

Figure 6 does not plot the estimated density of returns to the long-short strategy because it is almost identical to the estimated density of returns to the buy-and-hold strategy. This similarity is not surprising given the fact that the descriptive statistics of the long-short TSMOM strategy are virtually identical to those of the buy-and-hold strategy, see Table 7. Whereas the probability distribution of one-period returns to the buy-and-hold strategy has the familiar shape of the normal distribution function, the probability distribution function of returns to the long-only TSMOM strategy has a spike at r_f. The reason for this spike is evident: when the trading signal is Buy, the returns equal the market returns, but when the trading signal is Sell, the returns equal the risk-free rate of return.

6.2.2 Probability Distribution of Multi-Period Returns

It is a long tradition in empirical finance to measure the portfolio performance using returns sampled at the monthly frequency. The main reason for using monthly returns rather than, say, annual returns is to increase the sample size. It is known from statistics that the larger the sample size, the higher the precision of estimation and the power of statistical tests. However, already Levy (1972) argued that, in order to correctly make asset allocation decisions, the sampling interval for returns must equal the length of the investment horizons. Specifically, Levy (1972) was the first to show that a Sharpe ratio computed using short return intervals to evaluate portfolios or make asset allocation decisions is biased for long-term investors and may lead to suboptimal results. Thereafter, several studies further developed this idea and identified the investment horizon as an important factor affecting the performance measurement using the Sharpe ratio (see, for example, Chen and Lee (1981), Gunthorpe and Levy (1994), and Hodges, Taylor, and Yoder (1997)).

When the returns are independent and identically distributed, Levy (1972) demonstrated how to compute the multi-period Sharpe ratio. In the context of our study, the major obstacle is that the returns to both the buy-and-hold strategy and the TSMOM strategy are serially dependent. Hence, the analytical formulas derived by Levy (1972) cannot be used. Therefore, we rely on a simulation method. The simulations are conducted in the same manner as in the preceding section. But this time we simulate time-series of monthly returns to the buy-and-hold strategy, the long-only $MOM(9)$ strategy, and the long-short $MOM(9)$ strategy over the period of 5-years (60 monthly observations). The simulations are repeated $N = 1,000,000$ times. Figure 7 plots the estimated densities of 5-year returns to the buy-and-hold strategy, the long-only TSMOM strategy, and the long-short TSMOM strategy.

Figure 7 advocates that the shape of the probability density function of multi-period returns to the buy-and-hold strategy resembles the shape of the log-normal probability density function (that is displaced towards left). In contrast, the shapes of the probability density of multi-period returns to the TSMOM strategy differ from that of the buy-and-hold strategy. The most striking difference is observed between the shapes of estimated densities to the buy-and-hold strategy and the long-only TSMOM strategy. In fact, the shape of the estimated density function of multi-period returns to the long-only TSMOM strategy resembles the shape of density function of returns to a portfolio insurance strategy.[13]

[13]Portfolio insurance is a hedging strategy designed to limit the potential losses an investor might face from a declining stock index price. A variety of portfolio insurance strategies have been suggested in the literature. Examples are a protective put strategy, a constant proportion portfolio insurance (CPPI) strategy, a stop-loss strategy, etc. For a review of portfolio insurance strategies, see Dichtl, Drobetz, and Wambach (2017) and references therein.

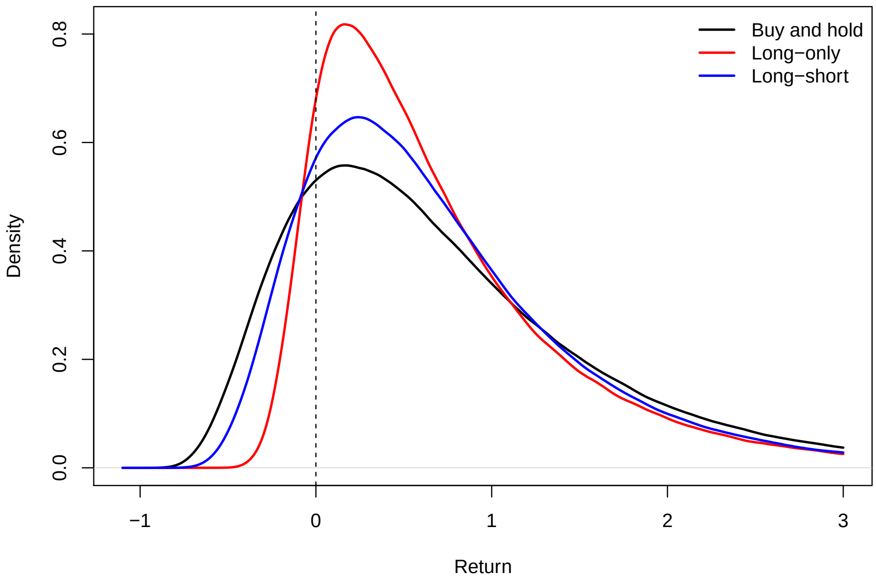

Figure 7: Kernel-smoothed densities of 5-year returns to the buy-and-hold strategy, the long-only $MOM(9)$ strategy, and long-short $MOM(9)$ strategy.

It is important to emphasize that, as compared to the empirical density of returns to the buy-and-hold strategy, the empirical density of returns to the long-only TSMOM strategy has a notably shorter left tail. That is, the two return distributions are significantly different in the domain of losses where the returns are negative. This observation suggests that a correct comparison of riskiness of the alternative strategies requires taking into account the differences between the shapes of the probability density functions. Therefore, to provide a deeper insight into the comparative riskiness of the alternative strategies, in addition to the mean and standard deviation of returns we will also compute the probability of loss. Formally, the probability of loss is defined by

$$\text{Probability of loss} = Prob(r < 0),$$

where r denotes the return and $Prob(\cdot)$ denotes the probability. The problem in using the probability of loss as a risk measure is the fact that this measure tells nothing about the magnitude of potential loss if loss occurs. That is, in principle, one financial asset may have a higher probability of loss than the other asset, but the losses on the latter asset might be much more severe than the losses on the former asset. To complete the picture of losses, we will also compute the expected return if loss occurs. This risk measure represents a specific realization of the popular risk measure that is known under different aliases: the Conditional Value-at-Risk (CVaR), the Expected Shortfall (ES), and the Expected Tail Loss (ETL). Formally, the expected return if loss occurs is computed as

$$\text{Expected return if loss occurs} = E[r|r < 0],$$

where $E[r|r < 0]$ denotes the expected return conditional on the outcome $r < 0$.

	Buy and Hold	**Long-only**	**Long-short**
Mean return, %	77.58	74.81	73.89
Std. deviation, %	99.76	84.18	88.18
Sharpe ratio	0.57	0.64	0.60
Probability of loss	0.20	0.10	0.16
Expected return if loss occurs, %	-23.76	-10.55	-18.93

Table 8: The descriptive statistics of 5-year returns to the buy-and-hold strategy, the long-only $MOM(9)$ strategy, and the long-short $MOM(9)$ strategy.

Table 8 reports the descriptive statistics of 5-year returns the buy-and-hold strategy, the long-only $MOM(9)$ strategy, and the long-short $MOM(9)$ strategy. Judging by the Sharpe ratio, over a 5-year horizon the advantage of the long-only TSMOM strategy over the buy-and-hold strategy becomes smaller as compared to that over a 1-month horizon. Namely, over a 5-year (1-month) horizon the Sharpe ratio of the long-only strategy is about 12% (30%) higher than the Sharpe ratio of the buy-and-hold strategy. However, the risk profile of the long-only TSMOM strategy is totally different from that of the buy-and-hold strategy. In particular, the long-only TSMOM strategy not only has notably lower standard deviation of returns, but twice as small the probability of loss and the expected loss (return) if loss occurs. Consequently, the advantage of the long-only TSMOM strategy over the buy-and-hold strategy lies not only in the higher Sharpe ratio, but also in its superior downside protection.

Finally, when it comes to the long-short TSMOM strategy, it is marginally better than the buy-and-hold strategy.[14] Specifically, as compared to the buy-and-hold strategy, the long-short TSMOM strategy has a bit higher Sharpe ratio and somewhat lower probability of loss and the expected loss if loss occurs. However, the long-short TSMOM strategy is definitely inferior to the long-only TSMOM strategy according to all descriptive statistics presented in Table 8.

[14]Under the assumption that the value of the autoregressive coefficient ϕ equals its point estimate.

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, Sullivan et al. (1999), Lee et al. (2001), Siegel (2002, Chapter 2), Okunev and White (2003), Olson (2004), Hutchinson and O'Brien (2014), and Zakamulin (2014)). Typically, this result is attributed to increased market efficiency over time. Another issue in the context of these studies is the lack of scientific evidence on profitability of trend-following strategies. Specifically, quite often the researchers cannot reject the null hypothesis that the performance of a trend-following strategy is similar to the performance of the corresponding buy-and-hold strategy (see, among others, Kim et al. (2016), Zakamulin (2017), and Huang et al. (2020)).

The results reported in the previous sections strongly suggest that the long-only TSMOM strategy has an edge over the buy-and-hold strategy. Put differently, the performance of the long-only TSMOM strategy is superior to that of the buy-and-hold strategy. Strictly speaking, this says that over the long-run the long-only TSMOM strategy tends to outperform the buy-and-hold strategy. However, because of randomness, over a short-run there is no guarantee that the long-only TSMOM strategy outperforms the buy-and-hold strategy. In this section we examine how the evidence of superior performance of the long-only TSMOM strategy depends on the length of investment horizon. The goal is to understand and explain why trend-following strategies very often demonstrate inferior performance over a short-run and lack of scientific evidence of superior performance even over a long-run.

In this section we again rely on a simulation method that is described in the preceding section. Specifically, for a fixed horizon of Y years, we simulate monthly returns to the buy-and-hold strategy and the long-only $MOM(9)$ strategy $N = 100,000$ times. After each simulation round, we compute the (monthly) Sharpe ratios of the buy-and-hold strategy, SR_{BH}, and the long-only TSMOM strategy, SR_{LO}. Subsequently, we conduct the test of the following null hypothesis

$$H_0 : SR_{LO} \leq SR_{BH}.$$

In words, the null hypothesis says that the Sharpe ratio of the long-only TSMOM strategy is not greater than the Sharpe ratio of the buy-and-hold strategy. The alternative hypothesis is, therefore, that the Sharpe ratio of the long-only TSMOM strategy is greater than the Sharpe ratio of the buy-and-hold strategy.

To conduct the test of the null hypothesis, we apply the Jobson and Korkie (1981) test with the Memmel (2003) correction. Specifically, given SR_{LO}, SR_{BH}, and ρ_b as the estimated Sharpe ratios and correlation coefficient over a sample of size $T = Y \times 12$ months, the test of the null hypothesis is obtained via the test statistic

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

which is asymptotically distributed as a standard normal. Finally, after carrying out all simulations for a specific horizon Y, we compute two probabilities. The first probability is the probability that over horizon of Y years the Sharpe ratio of the long-only TSMOM strategy is greater than the Sharpe ratio of the buy-and-hold strategy, $Prob(SR_{LO} > SR_{BH}) = Prob(z > 0)$. To compute this probability, we count how many times the Sharpe ratio of the long-only TSMOM strategy is greater than the Sharpe ratio 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 horizon of Y years the p-value of the Sharpe ratio test is lower than 5%. This probability is the probability of rejecting the null hypothesis at a significance level of 5%. 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 9 reports the results of this simulations study. The information in the table is very insightful and suggest the following conclusions. 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. However, even over the horizon of 50 years the probability of outperformance is about 80%. That is, even over a very long horizon (which is beyond the investment horizon of most individual investors) there is no guarantee that the trend following strategy outperforms the buy-and-hold strategy.

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

Table 9: The results of the simulation study. For each horizon, the table reports two probabilities. Probability $Prob(z > 0)$ denotes the probability that the Sharpe ratio of the long-only TSMOM strategy is greater than the Sharpe ratio of the buy-and-hold strategy. Probability $Prob(z > 1.64)$ denotes the probability of rejecting the null hypothesis at a significance level of 5%. The rejection of the null hypothesis means that the Sharpe ratio of the long-only TSMOM strategy is statistically significantly greater than the Sharpe ratio of the buy-and-hold strategy.

The probability that the long-only TSMOM strategy statistically significantly outperforms the buy-and-hold strategy also increases as the investment horizon lengthens. Over medium- and long-term horizons ranging from 5 to 50 years, the probability of observing statistically significant outperformance does not exceed 31%. Even over 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 horizon of 500 years. Yet, even in this case there is 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 of a test 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.

Figure 8 plots the estimated power of the statistical test of the null hypothesis $H_0 : SR_{LO} \leq SR_{BH}$ versus the number of years in a sample. Specifically, this figure plots the estimated sta-

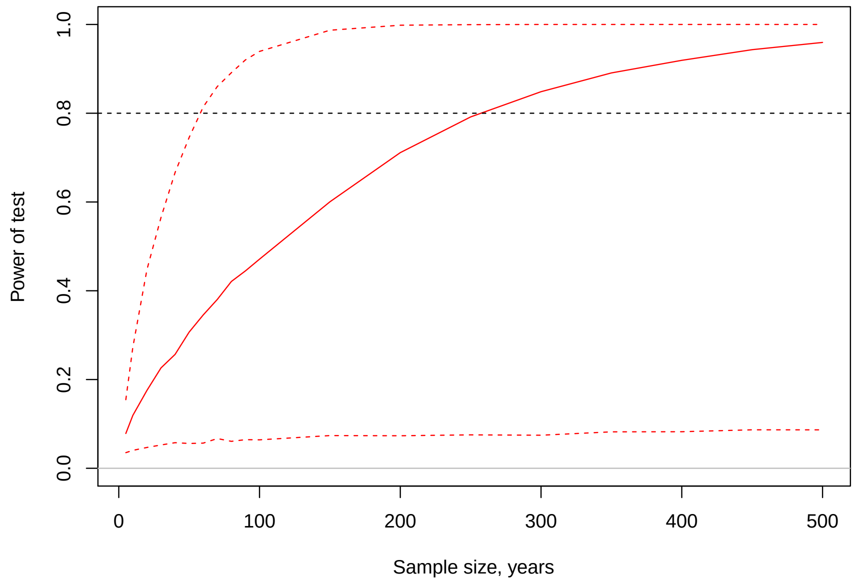

Figure 8: Estimated power of the statistical test of the null hypothesis $H_0 : SR_{LO} \leq SR_{BH}$ versus the number of years in a sample. The market excess returns follows an $AR(9)$ process. The solid red line in the figure plots the estimated statistical power when the value of ϕ equals its point estimate of 0.0324. The dashed red lines plot the estimated statistical power using the boundaries of the 95% confidence interval for ϕ.

tistical power when the value of ϕ equals its point estimate and when ϕ attains the boundaries of the 95% confidence interval for ϕ. The curves in Figure 8 suggest that, even in the best-case scenario when the value of ϕ equals its upper confidence bound, in order to reach the desired power level of 80% the sample size must be approximately 60 years. When ϕ equals its point estimate, the sample size must be about 250 years. Yet, such sample size is beyond the currently available historical sample sizes. Generalizing this result we conclude that in virtually all empirical studies that evaluate the profitability of a trend-following strategy the power of the statistical test is much below the acceptable level.

To sum up, the results of our simulation study agree very well with the findings reported in the numerous papers. Despite the fact that in our simulation study the true monthly Sharpe ratio of the long-only TSMOM strategy is about 30% higher than that of the buy-and-hold strategy, over short-term horizons the outperformance is not guaranteed. And this result has nothing to do with increased efficiency of financial markets. It is simply the result of randomness. In addition, because of the low power of the statistical test, the statistical significant outperformance is not guaranteed even over very long-term horizons of several hundred years.

6.4 Precision of Estimation of Optimal Number of Lags in TSMOM Rule

The results reported in the preceding sections 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 in the TSMOM rule is never known for certain. Typically, the optimal number of lags in any trading rule is found using the back-testing methodology. In the context of our study, the selection of the optimal number of lags in the $MOM(n)$ rule is conducted as follows. Using relevant historical data, one evaluates the performance of the $MOM(n)$ rule for various number of lags n and selects the value of n which maximizes the performance. To be more specific, given the number of lags n in the $MOM(n)$ rule, one simulates the returns to the long-only TSMOM strategy over a given historical sample $\left(R^{LO}_{1,n}, R^{LO}_{2,n}, \ldots, R^{LO}_{T,n}\right)$. The optimal number of lags n^* is found by maximizing the performance of the $MOM(n)$ strategy. Formally,

$$n^* = \arg \max_{n \in [n^{\min}, n^{\max}]} SR\left(R^{LO}_{1,n}, R^{LO}_{2,n}, \ldots, R^{LO}_{T,n}\right),$$

where T denotes the length of the historical sample, $n^{\min}$ and $n^{\max}$ are the minimum and maximum values of n, respectively, and $SR(\cdot)$ denotes the Sharpe ratio.

In principle, if Conjecture 1 is true and the historical sample is very long, this back-testing procedure must correctly reveal that $n^* = p$. However, because of randomness, in a short sample there is no guarantee that the back-testing procedure finds the true value of n^*. The goal of this section is to evaluate the precision of estimation of the optimal number of lags in the TSMOM rule versus the length of the sample. Again we rely on a simulation method that is described in the preceding sections. Specifically, for a fixed horizon of Y years, we simulate monthly returns to the buy-and-hold strategy. Then by varying $n \in [1, 24]$ we simulate the returns to a set of long-only $MOM(n)$ strategies and select the value of n that maximizes the Sharpe ratio of the TSMOM strategy. For each horizon, we repeat this procedure $N = 100,000$ times.

The number of autoregressive terms in the $AR(p)$ process for the excess returns to the buy-and-hold strategy equals 9. To evaluate the precision of estimation of the optimal number of lags in the TSMOM rule, for each horizon of Y years we compute the probability of n^* being 9, 9 ± 1, 9 ± 2, 9 ± 3, and 9 ± 4. Table 10 reports the results of this simulations study. The information in the table reveals that the precision of estimation of the optimal number of lags in the TSMOM rule is extremely poor. For example, for the US stock market we have historical data on the monthly excess returns that cover a period of about 150 years. However, even over such a long-term period the probability of correctly identifying that $n = 9$ is below 20%. There is about 25% probability that the estimated value of n lies outside of the interval 9 ± 4. Even if we had data that cover a period of 500 years, the probability of correctly identifying that $n = 9$ is only 34%.

Horizon,	**Probability of n^* being**				
years	9	9 ± 1	9 ± 2	9 ± 3	9 ± 4
5	0.05	0.16	0.25	0.35	0.44
10	0.06	0.17	0.27	0.37	0.46
20	0.07	0.20	0.31	0.41	0.51
30	0.09	0.23	0.34	0.45	0.55
40	0.10	0.25	0.37	0.48	0.58
50	0.11	0.27	0.40	0.51	0.61
75	0.13	0.31	0.45	0.57	0.66
100	0.15	0.35	0.50	0.61	0.71
200	0.21	0.46	0.62	0.73	0.81
300	0.26	0.53	0.70	0.80	0.86
500	0.34	0.64	0.80	0.88	0.93

Table 10: The results of the simulation study. The market excess returns follows an $AR(9)$ process with $\phi = 0.0324$. Consequently, the optimal number of lags in the TSMOM rule equals $n = 9$. The value of n^* denotes the optimal number of lags in the TSMOM rule found using the back-testing methodology. For each horizon, the table reports the probability of n^* being 9, 9 ± 1, 9 ± 2, 9 ± 3, and 9 ± 4.

The majority of traders believe that the market's dynamics is changing over time. Therefore, they insist that in a back-test one has to use the most recent data that cover a period from 5 to 10 years. In this case, the probability of correctly identifying that $n = 9$ is about 5% only; in more than 50% of cases the estimated value of n will lie outside of the interval 9 ± 4.

So the bad news is that in real situations the precision of estimation of the optimal number of lags in the TSMOM rule is very poor. However, there is also good news: the performance of the TSMOM rule is robust to the choice of the number of lags n. Recently, Zakamulin and Giner (2020) provide, among other things, a number of theoretical and empirical results on the similarity between two TSMOM rules. Specifically, one of these rules uses the number of lags n and the other uses the number of lags $m \neq n$. Zakamulin and Giner (2020) document that the similarity between the rules is rather high even when the market returns follow a random walk; the similarity increases when the trend strength increases.

For the sake of illustration, Figure 9 plots the theoretical monthly Sharpe ratio of the buy-and-hold strategy and the long-only $MOM(n)$ strategy versus the number of lags n. As before, we assume that the excess market returns $\mu - r_f$ follow the $AR(p)$ process with $p = 9$. The monthly parameters μ, σ, and r_f are estimated using the data for the whole sample that covers the period from 1857 to 2018. The solid red line in the figure plots the Sharpe ratio of the long-only TSMOM strategy when the value of ϕ equals its point estimate of 0.0324. The dashed red lines plot the Sharpe ratio of the TSMOM strategy using the boundaries of the 95% confidence interval for ϕ.

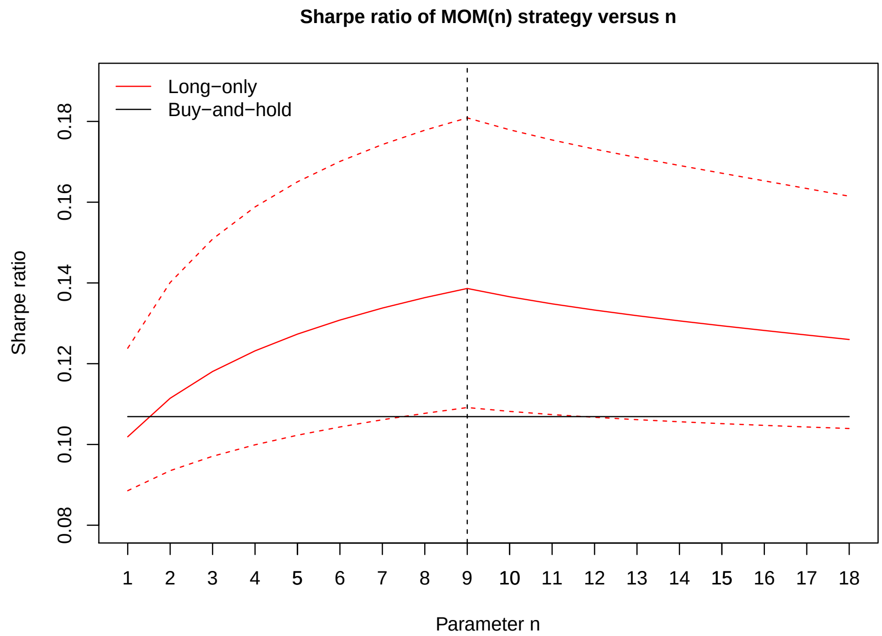

Figure 9: The theoretical Sharpe ratio of the buy-and-hold strategy and the long-only $MOM(n)$ strategy versus the number of lags n. The market excess returns follows an $AR(9)$ process. The solid red line in the figure plots the Sharpe ratio of the long-only TSMOM strategy when the value of ϕ equals its point estimate of 0.0324. The dashed red lines plot the Sharpe ratio of the TSMOM strategy using the boundaries of the 95% confidence interval for ϕ. The vertical dashed line show the location of $n = 9$.

The curves in this figure advocate that the performance of the TSMOM strategy is rather stable with respect to the number of lags n (at least when n is not very much different from p). For example, when ϕ equals its point estimate, the Sharpe ratio of either $MOM(6)$ or $MOM(14)$ strategy is only about 6% below the Sharpe ratio of the $MOM(9)$ strategy and still about 23% above the Sharpe ratio of the buy-and-hold strategy. In the worst-case scenario where ϕ equals the lower 95% confidence bound, the performance of the long-only TSMOM strategy roughly equals the performance of the buy-and-hold strategy when n is not far away from p. This theoretical result agrees very well with the results of empirical studies. In particular, Moskowitz et al. (2012) document that the TSMOM strategy delivers a rather stable performance when the number of return lags $n \in [6, 12]$. A similar finding is documented in Zakamulin and Giner (2020). All this suggests that the performance of the TSMOM strategy depends mainly on the strength of return persistence $\alpha = \phi\, p$ rather than on the choice of the number of lags n.

7 Conclusions

There is much controversy in the academic literature on the presence of short-term trends in financial markets and the profitability of trend-following strategies. This controversy has different aspects and raises many questions that need to be answered. This paper restricts its attention to the study of time-series momentum in the US stock market. The objective of this paper is to suggest answers to several major questions regarding time series momentum and to explain the existing controversy.

Our answer to the first question, whether short-term trends exist, is strongly affirmative. We present compelling evidence of short-term momentum in the excess returns on the S&P Composite stock price index. The affirmative answer to the first question leads to the second question: What is the type of process that generates these trends? We assume that the excess returns follow an autoregressive process of order p and, using a novel methodology, estimate the parameters of this process. Our estimation results reveal that over monthly horizons the autocorrelation in excess returns is very weak and, hence, escapes detection when traditional estimation methods are used.

Our answer to the third question, whether the TSMOM strategies are profitable, is affirmative for the long-only TSMOM strategy and negative for the long-short TSMOM strategy. Our results uncover that the shape of the probability density of multi-period returns to the long-only TSMOM strategy resembles the shape of density function of returns to a portfolio insurance strategy. Consequently, the advantage of the long-only TSMOM strategy over the buy-and-hold strategy lies not only in better performance, but also in its superior downside protection.

Given clear indications that the TSMOM strategy is superior to the buy-and-hold strategy, the fourth question arises: Why the existing empirical evidence on profitability is often controversial? The problem is that, because of randomness and the fact that the short-term trends are not strong enough, there is absolutely no guarantee that the TSMOM strategy will be profitable over a short-run. Our extensive simulation results show that over short- to medium-term horizons the probability that the TSMOM strategy outperforms the buy-and-hold strategy is less than 60%. Our ballpark estimate is that, in order to reach the desired power level of the statistical test on the profitability of the TSMOM strategy, the sample size must be about 250 years with monthly observations. Consequently, any empirical study tends not to reject the null hypothesis of no profitability because the power of the statistical test is much below the acceptable level. A related issue is a very poor precision of identification of the optimal number of return lags in the TSMOM rule using a standard back-testing methodology. Luckily, the performance of the TSMOM rule is robust to the choice of the number of lags.

References

Acar, E. (2003). "Chapter 8 - Modelling Directional Hedge Funds-Mean, Variance and Correlation with Tracker Funds", In Satchell, S. and Scowcroft, A. (Eds.), *Advances in Portfolio Construction and Implementation*, pp. 193 – 214. Butterworth-Heinemann, Oxford.

Andrews, D. W. K. and Chen, H.-Y. (1994). "Approximately Median-Unbiased Estimation of Autoregressive Models", *Journal of Business & Economic Statistics*, *12*(2), 187–204.

Ang, A., Goetzmann, W. N., and Schaefer, S. M. (2011). "The Efficient Market Theory and Evidence: Implications for Active Investment Management", *Foundations and Trends in Finance*, *5*(3), 157–242.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2016). *Time Series Analysis: Forecasting and Control* (5 edition). John Wiley & Sons Inc., New Jersey.

Brock, W., Lakonishok, J., and LeBaron, B. (1992). "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns", *Journal of Finance*, *47*(5), 1731–1764.

Brown, S. J., Goetzmann, W. N., and Kumar, A. (1998). "The Dow Theory: William Peter Hamilton's Track Record Reconsidered", *Journal of Finance*, *53*(4), 1311–1333.

Chen, S.-N. and Lee, C. F. (1981). "The Sampling Relationship between Sharpe's Performance Measure and Its Risk Proxy: Sample Size, Investment Horizon and Market Conditions", *Management Science*, *27*(6), 607–618.

Clare, A., Seaton, J., Smith, P. N., and Thomas, S. (2013). "Breaking Into the Blackbox: Trend Following, Stop losses and the Frequency of Trading - The Case of the S&P500", *Journal of Asset Management*, *14*(3), 182–194.

Dichtl, H., Drobetz, W., and Wambach, M. (2017). "A Bootstrap-Based Comparison of Portfolio Insurance Strategies", *European Journal of Finance*, *23*(1), 31–59.

Faber, M. (2017). "A Quantitative Approach to Tactical Asset Allocation Revisited 10 Years Later", *Journal of Portfolio Management*, *44*(2), 156–167.

Faber, M. T. (2007). "A Quantitative Approach to Tactical Asset Allocation", *Journal of Wealth Management*, *9*(4), 69–79.

Fama, E. F. (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work", *Journal of Finance*, *25*(2), 383–417.

Fama, E. F. and 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., and Sinclair, C. D. (2005). "An Analysis of Trading Strategies in Eleven European Stock Markets", *European Journal of Finance*, *11*(6), 531–548.

Gunthorpe, D. and Levy, H. (1994). "Portfolio Composition and the Investment Horizon", *Financial Analyst Journal*, *50*(1), 51–56.

Gwilym, O., Clare, A., Seaton, J., and Thomas, S. (2010). "Price and Momentum as Robust Tactical Approaches to Global Equity Investing", *Journal of Investing*, *19*(3), 80–91.

Hodges, C. W., Taylor, W. R., and Yoder, J. A. (1997). "Stock, Bonds, the Sharpe Ratio, and the Investment Horizon", *Financial Analyst Journal*, *53*(6), 74–80.

Huang, D., Li, J., Wang, L., and Zhou, G. (2020). "Time Series Momentum: Is It There?", *Journal of Financial Economics*, *135*(3), 774 – 794.

Hutchinson, M. C. and O'Brien, J. (2014). "Is This Time Different? Trend-Following and Financial Crises", *Journal of Alternative Investments*, *17*(2), 82–102.

Jobson, J. D. and Korkie, B. M. (1981). "Performance Hypothesis Testing with the Sharpe and Treynor Measures", *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, A. Y., Tse, Y., and Wald, J. K. (2016). "Time Series Momentum and Volatility Scaling", *Journal of Financial Markets*, *30*, 103 – 124.

Kim, M. J., Nelson, C. R., and Startz, R. (1991). "Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence", *Review of Economic Studies*, *58*(3), 515–528.

Kotz, S., Balakrishnan, N., and Johnson, N. L. (2000). *Continuous Multivariate Distributions, Volume 1: Models and Applications* (2 edition). John Wiley & Sons, Inc., New York.

Lee, C. I., Pan, M.-S., and 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*, *11*(2), 199 – 214.

Levy, H. (1972). "Portfolio Performance and the Investment Horizon", *Management Science*, *18*(12), 645–653.

Malkiel, B. G. (2003). "The Efficient Market Hypothesis and Its Critics", *Journal of Economic Perspectives*, *17*(1), 59–82.

Marques, C. R. (2005). "Inflation Persistence: Facts or Artifacts?", *Banco de Portugal Economic bulletin*, *11*(2), 69–79.

Marshall, B. R., Nguyen, N. H., and 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., and Pedersen, L. H. (2012). "Time Series Momentum", *Journal of Financial Economics*, *104*(2), 228–250.

Neely, C. J., Rapach, D. E., Tu, J., and Zhou, G. (2014). "Forecasting the Equity Risk Premium: The Role of Technical Indicators", *Management Science*, *60*(7), 1772–1791.

Nelson, C. R. and Kim, M. J. (1993). "Predictable Stock Returns: The Role of Small Sample Bias", *Journal of Finance*, *48*(2), 641–661.

Okunev, J. and 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 Grossmann, W., Pflug, G., Vincze, I., and Wertz, W. (Eds.), *Mathematical Statistics and Applications*, Vol. B, pp. 283–297. Reidel Publishing Company, Dordrecht, Nederland.

Rousseeuw, P. J. and Driessen, K. V. (1999). "A Fast Algorithm for the Minimum Covariance Determinant Estimator", *Technometrics*, *41*(3), 212–223.

Siegel, J. (2002). *Stocks for the Long Run.* McGraw-Hill Companies.

Sullivan, R., Timmermann, A., and White, H. (1999). "Data-Snooping, Technical Trading Rule Performance, and the Bootstrap", *Journal of Finance*, *54*(5), 1647–1691.

Welch, I. and 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. and Giner, J. (2020). "Trend Following with Momentum Versus Moving Average: A Tale of Differences", *Quantitative Finance*, *Forthcoming.*

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,

$$Cor(X_{t+k,t+1}, X_{t,t-k+1}) = \frac{Cov(X_{t+k,t+1}, X_{t,t-k+1})}{Var(X_{t,t-k+1})}, \tag{23}$$

where $Cov(X_{t+k,t+1}, X_{t,t-k+1})$ is the covariance between $X_{t+k,t+1}$ and $X_{t,t-k+1}$. Note that, because of the stationary assumption, the variance of $X_{t+k,t+1}$ equals that of $X_{t,t-k+1}$.

The variance of $X_{t+k,t+1}$ is given by

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

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

$$Var(X_{t,t-k+1}) \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} \rho_{|i-j|}\sigma_x^2.$$

In matrix notation, the expression for the variance becomes $Var(X_{t,t-k+1}) = \mathbf{1}_k' \boldsymbol{P}_{k,k} \mathbf{1}_k \sigma_x^2$, where $\mathbf{1}_k$ is the $k \times 1$ vector of ones and matrix $\boldsymbol{P}_{k,k}$ is the $k \times k$ matrix given by equation (8).

By similar reasoning, the covariance between $X_{t+k,t+1}$ and $X_{t,t-k+1}$ is given by

$$\begin{aligned} Cov(X_{t+k,t+1}, X_{t,t-k+1}) &= Cov\left(\sum_{i=1}^{k} X_{t+i}, \sum_{j=1}^{k} X_{t-k+j}\right) = \sum_{i=1}^{k}\sum_{j=1}^{k} Cov(X_{t+i}, X_{t-k+j}) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{k} \rho_{|k-j+i|}\sigma_x^2. \end{aligned}$$

In matrix notation, the expression for the covariance becomes $Cov(X_{t+k,t+1}, X_{t,t-k+1}) = \mathbf{1}_k' \boldsymbol{Q}_{k,k} \mathbf{1}_k \sigma_x^2$ where $\boldsymbol{Q}_{k,k}$ is the $k \times k$ matrix given by equation (13).

Inserting the expressions for $Cov(X_{t+k,t+1}, X_{t,t-k+1})$ and $Var(X_{t,t-k+1})$ into equation (23) completes the proof.

Proof of Proposition 2

Consider the returns to the following generalized trading strategy:

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

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

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

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 Useful 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 (15), we introduce variables z_t and y_{t-1} that are standardized variables 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.

The following results are straightforward to derive:

$$E[B_t] = a\Phi(-d) + b\Phi(d), \tag{24}$$

$$E[C_t] = c_a\Phi(-d) + c_b\Phi(d), \tag{25}$$

$$E[B_t^2] = a^2\Phi(-d) + b^2\Phi(d), \tag{26}$$

$$E[C_t^2] = c_a^2\Phi(-d) + c_b^2\Phi(d), \tag{27}$$

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

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, Balakrishnan, and Johnson (2000, pages 311-315) on the first and second moments of truncated bivariate distributions, we obtain

$$E[z_t|y_{t-1} > d] = \frac{\rho_m\varphi(d)}{\Phi(-d)}, \tag{28}$$

$$E[z_t|y_{t-1} \leq d] = -\frac{\rho_m\varphi(d)}{\Phi(d)}, \tag{29}$$

$$E[z_t^2|y_{t-1} > d] = \frac{(1 + \rho_m^2 d\varphi(d))}{\Phi(-d)}, \tag{30}$$

$$E[z_t^2|y_{t-1} \leq d] = \frac{(1 - \rho_m^2 d\varphi(d))}{\Phi(d)}. \tag{31}$$

As an example, the expectation $E[B_t z_t]$ can be obtained by combining equations (28) and (29)

$$E[B_t z_t] = a\Phi(-d)E[z_t|y_{t-1} > d] + b\Phi(d)E[z_t|y_{t-1} \leq d] = (a - b)\rho_m\varphi(d). \tag{32}$$

As an another example, the expectation $E[B_t^2 z_t^2]$ can be obtained by combining equations (30) and (31)

$$\begin{aligned} E[B_t^2 z_t^2] &= a^2\Phi(-d)E[z_t^2|y_{t-1} > d] + b^2\Phi(d)E[z_t^2|y_{t-1} \leq d] \\ &= a^2\Phi(-d) + b^2\Phi(d) + (a^2 - b^2)\rho_m^2 d\varphi(d). \end{aligned} \tag{33}$$

Both of these results are needed later in the derivation.

Mean Returns of Generalized Trading Strategy

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

$$E[R_t] = E[B_t r_t + C_t] = E[B_t r_t] + E[C_t]. \tag{34}$$

The term $E[C_t]$ is given by equation (25). The term $E[B_t r_t]$ can be analyzed 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 equation (24), whereas the term $E[B_t z_t]$ is given by equation (32). Putting everything together into equation (34), we obtain

$$E[R_t] = \mu \left(a\Phi(-d) + b\Phi(d)\right) + (a-b)g + c_a\Phi(-d) + c_b\Phi(d), \tag{35}$$

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

Variance of Returns of Generalized Trading Strategy

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

$$Var[R_t] = E[R_t^2] - E[R_t]^2. \tag{36}$$

The term $E[R_t]$ is given by equation (35). Consider the term $E[R_t^2]$ which is the mean squared return of the generalized trading strategy:

$$E[R_t^2] = E\left[(B_t r_t + C_t)^2\right] = E\left[B_t^2 r_t^2 + 2B_t C_t r_t + C_t^2\right] = E\left[B_t^2 r_t^2\right] + 2E\left[B_t C_t r_t\right] + E\left[C_t^2\right]. \tag{37}$$

The term $E[C_t^2]$ is given by equation (27). Consider the term $E[B_t C_t r_t]$

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

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

$$E[B_t C_t] = a c_a \Phi(-d) + b c_b \Phi(d).$$

The expression for $E[B_t C_t z_t]$ can be obtained in a similar manner to that of the expression for $E[B_t z_t]$ (see equation (32))

$$E[B_t C_t z_t] = (a c_a - b c_b)\rho_m \varphi(d).$$

Therefore,

$$E[B_tC_tr_t] = \mu \left(ac_a\Phi(-d) + bc_b\Phi(d)\right) + (ac_a - bc_b)g. \tag{38}$$

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

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

The terms with $E[B_t^2]$ and $E[B_t^2 z_t^2]$ are given by equations (26) and (33) respectively. The expression for the term $E[B_t^2 z_t]$ is obtained in a similar manner to that of $E[B_t z_t]$ (see equation (32))

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

Consequently, we obtain

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

In the end, substituting $E[C_t^2]$, $E[B_tC_tr_t]$ and $E[B_t^2 r_t^2]$ into equation (37) for $E[R_t^2]$, the final expression for the variance of returns of the generalized trading strategy becomes:

$$\begin{aligned} Var[R_t] = &\left(\mu^2 + \sigma^2\right)\left(a^2\Phi(-d) + b^2\Phi(d)\right) + \left(a^2 - b^2\right) g(2\mu + \sigma\rho_m d) \\ &+ 2\mu\left(ac_a\Phi(-d) + bc_b\Phi(d)\right) + 2(ac_a - bc_b)g + c_a^2\Phi(-d) + c_b^2\Phi(d) - E[R_t]^2. \end{aligned} \tag{39}$$

Results for the Long-Only and Long-Short Trading Strategy

The particular results for the long-only trading strategy are obtained through equations (35) and (39) 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 equations (35) and (39) by using $a = 1$, $b = -1$, $c_a = 0$ and $c_b = 2\,r_f$.