Market Timing with Moving Averages: Anatomy and Performance of Trading Rules*

Valeriy Zakamulin†

This revision: March 25, 2015

Abstract

In this paper, we contribute to the literature in two important ways. The first contribution is to demonstrate the anatomy of market timing rules with moving averages. Our analysis offers a broad and clear perspective on the relationship between different rules and reveals that all technical trading indicators considered in this paper are computed in the same general manner. In particular, the computation of every technical trading indicator can be equivalently interpreted as the computation of the weighted moving average of price changes. The second contribution of this paper is to perform the longest out-of-sample testing of a set of trading rules. The trading rules in this set are selected to have clearly distinct weighting schemes. We report the detailed historical performance of the trading rules over the period from 1870 to 2010 and debunk several myths and common beliefs about market timing with moving averages.

Key words: technical analysis, market timing, momentum rule, price minus moving average rule, moving average change of direction rule, double crossover method, out-of-sample testing

JEL classification: G11, G17.

*This is the first draft. All comments are very welcome.

†a.k.a. Valeri Zakamouline, School of Business and Law, University of Agder, Service Box 422, 4604 Kristiansand, Norway, Tel.: (+47) 38 14 10 39, E-mail: Valeri.Zakamouline@uia.no

1 Introduction

Technical analysis represents a methodology of forecasting the future price movements through the study of past price data and uncovering some recurrent regularities, or patterns, in price dynamics. One of the fundamental principles of technical analysis is that prices move in trends. Analysts firmly believe that these trends can be identified in a timely manner to generate profits and limit losses. Market timing is an active trading strategy that implements this idea in practice. Specifically, this strategy is based on switching between the market and the cash depending on whether the prices trend upward or downward. A moving average is one of the oldest and most popular tools used in technical analysis for detecting a trend.

The great controversy about technical analysis is over whether it is scientific or non-scientific. One the one hand, technical analysis has been extensively used by traders for over a century and the majority of active traders strongly believe in market timing. On the other hand, academics had long been skeptical about the usefulness of technical analysis. Yet the academics' attitude towards the technical analysis is gradually changing. The findings in the papers on technical analysis of financial markets, published in prominent academic journals (examples are Brock, Lakonishok, and LeBaron (1992), Sullivan, Timmermann, and White (1999), Lo, Mamaysky, and Wang (2000), Okunev and White (2003), and Moskowitz, Ooi, and Pedersen (2012)), suggest that one should not dismiss the value of technical analysis. Recently we have witnessed a constantly increasing interest in technical analysis from both the practitioners and academics alike (see Park and Irwin (2007)). This interest developed because over the course of the last 15 years, especially over the decade of 2000s, many technical trading rules outperformed the market by a large margin.

However, despite a series of publications in academic journals, modern technical analysis is still largely based on superstitions and beliefs. Consequently, modern technical analysis remains art rather than science. The situation with market timing is as follows. There have been proposed many technical trading rules based on moving averages of prices calculated on a fixed size data window (called the "lookback" period). The main examples are: the momentum rule, the price-minus-moving-average rule, the change-of-direction rule, and the double-crossover method. In addition, there are several popular types of moving averages: simple (or equally-weighted) moving average, linearly-weighted moving average, exponentially-weighed moving average, etc. As a result, there exists a large number of potential combinations of trading rules with moving average weighting schemes. One of the controversies about market timing is over which trading rule in combination with which weighing scheme produces the best performance. The situation is further complicated because in order to compute a moving average one must define the length of the lookback period. Again, there is a big controversy about the length of the optimal lookback period. Nevertheless, one can easily note that technical traders do share a few common beliefs and myths. They are as follows. First, one can easily beat the market using some technical trading rules. Second, in the computation of a moving average one has to overweight the most recent prices because they contain more relevant information on the future direction of the price than earlier prices. Finally, in each trading rule there exits some specific, time-invariant, length of the lookback period that produces the best performance.

There is a simple explanation for the existing situation with market timing. First of all, technical traders do not really understand the response characteristics of the trading indicators they use. The selection of a trading rule is made based mainly on intuition rather than any deeper analysis of commonalities and differences between miscellaneous choices for trading rules and moving average weighting schemes. Second, there is usually no objective scientific evidence which supports the claim that some specific trading rule, coupled with some specific choices for the moving average weighting scheme and the length of the lookback period, produces the best performance. Often such a claim is supported by colorful narratives and anecdotal evidence. At best, such a claim is "supported" by finding the best performing combination in a back test using an arbitrary chosen period of historical prices. Yet this approach to selecting the best rule is commonly termed as "data-mining" (or "data-snooping") and has nothing to do with science (see Sullivan et al. (1999), White (2000), and Aronson (2006)).

In this paper, we contribute to the literature in two important ways. The first contribution is to demonstrate the anatomy of market timing rules with moving averages. Specifically, we begin the paper by presenting a methodology for examining how the value of a trading indicator is computed. Then using this methodology we study the computation of trading indicators in many market timing rules and analyze the commonalities and differences between the rules. We reveal that despite being computed seemingly different at the first sight, all technical trading indicators considered in this paper are computed in the same general manner. In particular, the computation of every technical trading indicator can be equivalently interpreted as the computation of the weighted moving average of price changes. The only real difference lies in the weighting scheme used to compute the moving average of price changes.

Our methodology of analyzing the computation of trading indicators for the timing rules based on moving averages offers a broad and clear perspective on the relationship between different rules. We show, for example, that every trading rule can also be presented as a weighted average of the momentum rules computed using different lookback periods. Thus, the momentum rule might be considered as an elementary trading rule on the basis of which one can construct more elaborate rules. In addition, we establish a one-to-one equivalence between a price-minus-moving-average rule and a corresponding moving-average-change-of-direction rule.

The second contribution of this paper is to perform an objective testing of the common beliefs and myths about the performance of market timing rules. These myths and common beliefs represent, in fact, meaningful claims that can be tested using historical data. Specifically, we perform out-of-sample testing of a few clearly distinct market timing rules based on moving averages in order to find out the following: Is it possible to beat the market by timing it? Does over- or under-weighting the recent prices allow one to improve the performance of market timing? Is there a single optimal lookback period in each trading rule?

We assemble a dataset of monthly returns on the Standard and Poor's Composite stock price index, as well as the risk-free rate of return, over the period from January 1860 to December 2009. We perform the longest out-of-sample test[1] of market timing rules over the period of 140 years, from January 1870 to December 2009. First and foremost, our results indicate that there is no single optimal lookback period in each trading rule. That is, contrary to the common belief, the length of the optimal lookback period is time-varying and depends probably on market conditions. Motivated by this finding, we perform the out-of-sample simulation of the returns to the market timing rules using not only the expanding-window estimation scheme to determine the length of the optimal lookback period, but also the rolling-window estimation scheme. This allows us to find out which estimation scheme produces the best performance of a market timing strategy. It is worth noting that the former estimation scheme is used when

[1] To the best knowledge of the author, there are only two papers to date in which the researchers implement an out-of-sample test of profitability for some trading rules in the stock market. Specifically, in the paper by Sullivan et al. (1999) the length of the out-of-sample period amounts to 10 years only. In the paper by Zakamulin (2014) the length of the out-of-sample period amounts to 80 years.

the parameter of estimation is supposed to be constant, whereas the latter estimation scheme is used when parameter instability is suspected.

The results of our out-of-sample testing suggest that the majority of market timing rules show a better risk-adjusted performance than that of the market. For a few market timing rules we find the evidence that the outperformance is statistically significant. There are indications that the use of the rolling-window estimation scheme produces better out-of-sample performance than the use of the expanding-window one. Contrary to the common belief, we find that neither over-weighting nor under-weighting the recent price changes improves the performance of a market timing strategy. Specifically, we find that the momentum rule, where the price changes are equally weighted, produces the best performance in out-of-sample tests.

Despite the fact that over a very long horizon (which is beyond the investment horizon of most individual investors) an active timing strategy tends to outperform the market, the performance of an active strategy is highly uneven over time. Therefore, as argued by Zakamulin (2014), the traditional performance measurement, which consists in reporting a single number for performance, is very misleading for investors with medium-term horizons. To give a broader and clearer picture of market timing performance, we provide a detailed descriptive statistics of performance over 5- and 10-year horizons. Here the main new finding is that even for the best performing timing rules the probability of outperforming the market is barely above 50%. That is, there is absolutely no guarantee that a timing strategy beats the market over a medium run. Roughly, over medium-term horizons, the market timing strategy is equally likely to outperform as to underperform. Yet the average outperformance is greater than the average underperformance.

The rest of the paper is organized as follows. In the subsequent Section 2 we first present the moving averages and trading rules considered in the paper. Then we demonstrate the anatomy of trading rules with different moving averages. In Section 3 we perform an objective testing of the common beliefs and myths about the performance of market timing rules. We begin this section with a detailed description of our data, the set of tested rules, and our methodology. Then we perform the out-of-sample testing and present the results. Section 4 concludes the paper.

2 Anatomy of Market Timing with Moving Averages

2.1 Moving Averages

A moving average of prices is calculated using a fixed size data "window" that is rolled through time, each month adding the new price and taking off the oldest price. The length of this window of data, also called the "lookback" period or averaging period, is the time interval over which the moving average is computed. We follow the standard practice and use prices, not adjusted for dividends, in the computation of moving averages and all technical trading indicators. More formally, let $(P_1, P_2, \ldots, P_T)$ be the observations of the monthly closing prices of a stock price index. A moving average at time t is computed using the last closing price P_t and k lagged prices P_{t-j}, $j \in [1, k]$. It is worth noting that the time interval over which the moving average is computed amounts to k months and includes $k+1$ monthly observations. Generally, each price observation in the rolling window of data has its own weight in the computation of a moving average. More formally, a weighted Moving Average at month-end t with k lagged prices (denoted by $MA_t(k)$) is computed as

$$MA_t(k) = \frac{w_t P_t + w_{t-1} P_{t-1} + w_{t-2} P_{t-2} + \ldots + w_{t-k} P_{t-k}}{w_t + w_{t-1} + w_{t-2} + \ldots + w_{t-k}} = \frac{\sum_{j=0}^{k} w_{t-j} P_{t-j}}{\sum_{j=0}^{k} w_{t-j}},$$

where w_{t-j} is the weight of price P_{t-j} in the computation of the weighted moving average. It is worth observing that in order to compute a moving average one has to use at least one lagged price, this means that one should have $k \geq 1$. When the number of lagged prices is zero, a moving average becomes the last closing price

$$MA_t(0) = P_t.$$

The most commonly used type of moving average is the simple moving average. A Simple Moving Average (SMA) at month-end t is computed as

$$SMA_t(k) = \frac{P_t + P_{t-1} + P_{t-2} + \ldots + P_{t-k}}{k+1} = \frac{1}{k+1} \sum_{j=0}^{k} P_{t-j}.$$

A simple moving average is, in fact, an equally-weighted moving average where an equal weight is given to each price observation. Many analysts argue that the most recent stock prices contain more relevant information on the future direction of the stock price than earlier stock prices. Therefore, one should put more weight on the more recent price observations. For this purpose, analysts employ either the linearly weighted moving average or the exponentially weighted moving average.

A Linear (or linearly weighted) Moving Average (LMA) at month-end t is computed as

$$LMA_t(k) = \frac{(k+1)P_t + kP_{t-1} + (k-1)P_{t-2} \ldots + P_{t-k}}{(k+1) + k + (k-1) + \ldots + 1} = \frac{\sum_{j=0}^{k}(k-j+1)P_{t-j}}{\sum_{j=0}^{k}(k-j+1)}.$$

In the linearly weighted moving average the weights decrease in arithmetic progression. In particular, in $LMA(k)$ the latest observation has weight $k+1$, the second latest k, etc. down to one. A disadvantage of the linearly weighted moving average is that the weighting scheme is too rigid. This problem can be addressed by using the exponentially weighted moving average instead of the linearly weighted moving average. An Exponential Moving Average (EMA) at month-end t is computed as

$$EMA_t(k) = \frac{P_t + \lambda P_{t-1} + \lambda^2 P_{t-2} + \ldots + \lambda^k P_{t-k}}{1 + \lambda + \lambda^2 + \ldots + \lambda^k} = \frac{\sum_{j=0}^{k} \lambda^j P_{t-j}}{\sum_{j=0}^{k} \lambda^j},$$

where $0 < \lambda \leq 1$ is a decay factor. When $\lambda < 1$, the exponentially weighted moving average assigns greater weights to the most recent prices. By varying the value of λ, one is able to adjust the weighting to give greater or lesser weight to the most recent price. The properties of the exponential moving average:

$$\lim_{\lambda \to 1} EMA_t(k) = SMA_t(k), \quad \lim_{\lambda \to 0} EMA_t(k) = P_t. \tag{1}$$

Probably the least commonly used type of moving average is the Reverse Exponential Moving Average (REMA) computed as

$$REMA_t(k) = \frac{\lambda^k P_t + \lambda^{k-1} P_{t-1} + \lambda^{k-2} P_{t-2} + \ldots + P_{t-k}}{\lambda^k + \lambda^{k-1} + \lambda^{k-2} + \ldots + 1} = \frac{\sum_{j=0}^{k} \lambda^{k-j} P_{t-j}}{\sum_{j=0}^{k} \lambda^{k-j}}.$$

Contrary to the regular exponential moving average that gives greater weights to the most recent prices, the reverse exponential moving average assigns greater weights to the most oldest prices and decreases the importance of the most recent prices. The use of the reverse exponential moving average can be justified if one assumes that earlier stock prices contain more relevant information on the future direction of the stock price than the most recent stock prices. The properties of the reverse exponential moving average:

$$\lim_{\lambda \to 1} REMA_t(k) = SMA_t(k), \quad \lim_{\lambda \to 0} REMA_t(k) = P_{t-k}. \tag{2}$$

Instead of the regular moving averages of prices considered above, traders sometimes use more elaborate moving averages that can be considered as "moving averages of moving averages". Specifically, instead of using a regular moving average to smooth the price series, some traders perform either double- or triple-smoothing of the price series. The main examples of this type of moving averages are: Triangular Moving Average, Double Exponential Moving Average, and Triple Exponential Moving Average (see, for example, Kirkpatrick and Dahlquist (2010)). To shorten and streamline the presentation, we will not consider these moving averages in our paper. Yet our methodology can be applied to the analysis of the trading indicators based on this type of moving averages in a straightforward manner.

2.2 Technical Trading Rules

Every market timing rule prescribes investing in the stocks (that is, the market) when a Buy signal is generated and moving to cash when a Sell signal is generated. In the absence of transaction costs, the time t return to a market timing strategy is given by

$$r_t = \delta_{t|t-1} r_{Mt} + \left(1 - \delta_{t|t-1}\right) r_{ft}, \tag{3}$$

where r_{Mt} and r_{ft} are the month t returns on the stock market (including dividends) and the risk-free asset respectively, and $\delta_{t|t-1} \in \{0, 1\}$ is a trading signal for month t (0 means Sell and 1 means Buy) generated at the end of month $t-1$.

In each market timing rule the generation of a trading signal is a two-step process. At the first step, one computes the value of a technical trading indicator using the last closing price and k lagged prices

$$\text{Indicator}_t^{TR(k)} = Eq(P_t, P_{t-1}, \ldots, P_{t-k}),$$

where TR denotes the timing rule and $Eq(\cdot)$ is the equation that specifies how the technical trading indicator is computed. At the second step, using a specific function one translates the value of the technical indicator into the trading signal. In all market timing rules considered in this paper the Buy signal is generated when the value of a technical trading indicator is positive. Otherwise, the Sell signal is generated. Thus, the generation of a trading signal can be interpreted as an application of the following (mathematical) *indicator function* to the value of the technical indicator

$$\delta_{t+1|t} = \mathbf{1}_{+}\left(\text{Indicator}_t^{TR(k)}\right),$$

where the indicator function $\mathbf{1}_{+}(\cdot)$ is defined by

$$\mathbf{1}_{+}(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

We start with the Momentum rule which is the simplest and most basic market timing rule. In the Momentum rule one compares the last closing price, P_t, with the closing price k months ago, P_{t-k}. In this rule a Buy signal is generated when the last closing price is greater than the closing price k months ago. Formally, the technical trading indicator for the Momentum rule is computed as

$$\text{Indicator}_t^{\text{MOM}(k)} = MOM_t(k) = P_t - P_{t-k}.$$

Then the trading signal is generated by

$$\delta_{t+1|t}^{\text{MOM}(k)} = \mathbf{1}_{+}\left(MOM_t(k)\right).$$

Most often, in order to generate a trading signal, the analysts compare the last closing price with the value of a k-month moving average. In this case a Buy signal is generated when the last closing price is above a k-month moving average. Otherwise, if the last closing price is below a k-month moving average, a Sell signal is generated. Formally, the technical trading indicator for the Price-Minus-Moving-Average rule is computed as

$$\text{Indicator}_t^{\text{P-MA}(k)} = P_t - MA_t(k).$$

The trading signals for each considered type of moving average are generated by

$$\begin{aligned}
\delta_{t+1|t}^{\text{P-SMA}(k)} &= \mathbf{1}_{+}\left(P_t - SMA_t(k)\right), \\
\delta_{t+1|t}^{\text{P-LMA}(k)} &= \mathbf{1}_{+}\left(P_t - LMA_t(k)\right), \\
\delta_{t+1|t}^{\text{P-EMA}(k)} &= \mathbf{1}_{+}\left(P_t - EMA_t(k)\right), \\
\delta_{t+1|t}^{\text{P-REMA}(k)} &= \mathbf{1}_{+}\left(P_t - REMA_t(k)\right).
\end{aligned}$$

Some analysts argue that the price is noisy and the Price-Minus-Moving-Average rule produces many false signals (whipsaws). They suggest to address this problem by employing two moving averages in the generation of a trading signal: one shorter average with lookback period s and one longer average with lookback period $k > s$. This technique is called the Double Crossover Method (see, for example, Murphy (1999), Chapter 9). In this case the technical trading indicator is computed as

$$\text{Indicator}_t^{\text{DCM}(s,k)} = MA_t(s) - MA_t(k).$$

It is worth noting the obvious relationship

$$\text{Indicator}_t^{\text{DCM}(0,k)} = \text{Indicator}_t^{\text{P-MA}(k)}.$$

Less often, in order to generate a trading signal, the analysts compare the most recent value of a k-month moving average with the value of a k-month moving average in the preceding month. Intuitively, when the stock prices are trending upward (downward) the moving average is increasing (decreasing). Consequently, in this case a Buy signal is generated when the value of a k-month moving average has increased over a month. Otherwise, a Sell signal is generated. Formally, the technical trading indicator for the Moving-Average-Change-of-Direction rule is computed as

$$\text{Indicator}_t^{\Delta\text{MA}(k)} = MA_t(k) - MA_{t-1}(k).$$

The trading signals for each considered type of moving average are generated by

$$\begin{aligned}
\delta_{t+1|t}^{\Delta\mathrm{SMA}(k)} &= \mathbf{1}_+\left(SMA_t(k) - SMA_{t-1}(k)\right),\\
\delta_{t+1|t}^{\Delta\mathrm{LMA}(k)} &= \mathbf{1}_+\left(LMA_t(k) - LMA_{t-1}(k)\right),\\
\delta_{t+1|t}^{\Delta\mathrm{EMA}(k)} &= \mathbf{1}_+\left(EMA_t(k) - EMA_{t-1}(k)\right),\\
\delta_{t+1|t}^{\Delta\mathrm{REMA}(k)} &= \mathbf{1}_+\left(REMA_t(k) - REMA_{t-1}(k)\right).
\end{aligned}$$

2.3 Anatomy of Trading Rules

2.3.1 Preliminaries

It has been known for years that there is a relationship between the Momentum rule and the Simple-Moving-Average-Change-of-Direction rule.[2] In particular, note that

$$SMA_t(k-1) - SMA_{t-1}(k-1) = \frac{P_t - P_{t-k}}{k} = \frac{MOM_t(k)}{k}.$$

Therefore

$$\mathrm{Indicator}_t^{\Delta\mathrm{SMA}(k-1)} \equiv \mathrm{Indicator}_t^{\mathrm{MOM}(k)}, \tag{4}$$

where the symbol $\equiv$ means *equivalence*. The equivalence of two technical indicators follows from the following property: *the multiplication of a technical indicator by any positive real number produces an equivalent technical indicator.* This is because the trading signal is generated depending on the sign of the technical indicator. The formal presentation of this property:

$$\mathbf{1}_+\left(a \times \mathrm{Indicator}_t(k)\right) = \mathbf{1}_+\left(\mathrm{Indicator}_t(k)\right), \tag{5}$$

where a is any positive real number. Using relation (4) as an illustrating example, observe that if $SMA_t(k-1) > SMA_{t-1}(k-1)$ then $MOM_t(k) > 0$ and vice versa. In other words, the Simple-Moving-Average-Change-of-Direction rule, $\Delta\mathrm{SMA}(k-1)$, generates the Buy (Sell) trading signal when the Momentum rule, $MOM_t(k)$, generates the Buy (Sell) trading signal.

What else can we say about the relationship between different market timing rules? The ultimate goal of this section is to answer this question and demonstrate that all market timing rules considered in this paper are closely interconnected. In particular, we are going to show

[2] See, for example, `http://en.wikipedia.org/wiki/Momentum_(technical_analysis)`.

that the computation of a technical trading indicator for every market timing rule can be interpreted as the computation of the weighted moving average of monthly price changes over the lookback period. We will do it sequentially for each trading rule.

2.3.2 Momentum Rule

The computation of the technical trading indicator for the Momentum rule can be equivalently represented by

$$\begin{aligned}\text{Indicator}_t^{\text{MOM}(k)} &= MOM_t(k) = P_t - P_{t-k} \\ &= (P_t - P_{t-1}) + (P_{t-1} - P_{t-2}) + ... + (P_{t-k+1} - P_{t-k}) = \sum_{i=1}^{k} \Delta P_{t-i},\end{aligned} \tag{6}$$

where $\Delta P_{t-i} = P_{t-i+1} - P_{t-i}$. Consequently, using property (5), the computation of the technical indicator for the Momentum rule is equivalent to the computation of the equally weighted moving average of the monthly price changes over the lookback period:

$$\text{Indicator}_t^{\text{MOM}(k)} \equiv \frac{1}{k} \sum_{i=1}^{k} \Delta P_{t-i}. \tag{7}$$

2.3.3 Price-Minus-Moving-Average Rule

First, we derive the relationship between the Price-Minus-Moving-Average rule and the Momentum rule:

$$\begin{aligned}\text{Indicator}_t^{\text{P-MA}(k)} &= P_t - MA_t(k) = P_t - \frac{\sum_{j=0}^{k} w_{t-j} P_{t-j}}{\sum_{j=0}^{k} w_{t-j}} = \frac{\sum_{j=0}^{k} w_{t-j} P_t - \sum_{j=0}^{k} w_{t-j} P_{t-j}}{\sum_{j=0}^{k} w_{t-j}} \\ &= \frac{\sum_{j=1}^{k} w_{t-j}(P_t - P_{t-j})}{\sum_{j=0}^{k} w_{t-j}} = \frac{\sum_{j=1}^{k} w_{t-j} MOM_t(j)}{\sum_{j=0}^{k} w_{t-j}}.\end{aligned} \tag{8}$$

Using property (5), the relation above can be conveniently re-written as

$$\text{Indicator}_t^{\text{P-MA}(k)} \equiv \frac{\sum_{j=1}^{k} w_{t-j} MOM_t(j)}{\sum_{j=1}^{k} w_{t-j}}. \tag{9}$$

Consequently, the computation of the technical indicator for the Price-Minus-Moving-Average rule, $P_t - MA_t(k)$, is equivalent to the computation of the weighted moving average of technical indicators for the Momentum rules, $MOM_t(j)$, for $j \in [1,k]$. It is worth noting that the weighting scheme for computing the moving average of the momentum technical indicators, $MOM_t(j)$, is the same as the weighting scheme for computing the weighted moving average $MA_t(k)$.

Second, we use identity (6) and rewrite the numerator in (9) as

$$\sum_{j=1}^{k} w_{t-j} MOM_t(j) = \sum_{j=1}^{k} w_{t-j} \sum_{i=1}^{j} \Delta P_{t-i} = w_{t-1}\Delta P_{t-1} + w_{t-2}(\Delta P_{t-1} + \Delta P_{t-2}) + \ldots$$
$$+ w_{t-k}(\Delta P_{t-1} + \Delta P_{t-2} + \ldots + \Delta P_{t-k}) = (w_{t-1} + \ldots + w_{t-k})\Delta P_{t-1} \tag{10}$$
$$+ (w_{t-2} + \ldots + w_{t-k})\Delta P_{t-2} + \ldots + w_{t-k}\Delta P_{t-k} = \sum_{i=1}^{k}\left(\sum_{j=i}^{k} w_{t-j}\right)\Delta P_{t-i}.$$

The last expression tells us that the numerator in (9) is a weighted sum of the monthly price changes over the lookback period, where the weight of ΔP_{t-i} equals $\sum_{j=i}^{k} w_{t-i}$. Thus, another alternative expression for the computation of the technical indicator for the Price-Minus-Moving-Average rule is given by

$$\text{Indicator}_t^{\text{P-MA}(k)} \equiv \frac{\sum_{i=1}^{k}\left(\sum_{j=i}^{k} w_{t-j}\right)\Delta P_{t-i}}{\sum_{i=1}^{k}\left(\sum_{j=i}^{k} w_{t-j}\right)} = \frac{\sum_{i=1}^{k} x_i \Delta P_{t-i}}{\sum_{i=1}^{k} x_i}. \tag{11}$$

where

$$x_i = \sum_{j=i}^{k} w_{t-j} \tag{12}$$

is the weight of the price change ΔP_{t-i}. In words, the computation of the technical indicator for the Price-Minus-Moving-Average rule is equivalent to the computation of the weighted moving average of the monthly price changes over the lookback period.

It is important to note from equation (12) that the application of the Price-Minus-Moving-Average rule usually leads to overweighting the most recent price changes as compared to the original weighting scheme used to compute the moving average of prices. If the weighting scheme in a trading rule is already designed to overweight the most recent prices, then as a rule the trading signal is computed with a much stronger overweighting the most recent price changes. This will be demonstrated below.

Let us now, on the basis of (11), present the alternative expressions for the computation of Price-Minus-Moving-Average technical indicators that use the specific weighting schemes described in the beginning of this section. We start with the Simple Moving Average which uses the equally weighted moving average of prices. In this case the weight of ΔP_{t-i} is given by

$$x_i = \sum_{j=i}^{k} w_{t-j} = \sum_{j=i}^{k} 1 = k - i + 1.$$

Consequently, the equivalent representation for the computation of the technical indicator for the Price-Minus-Simple-Moving-Average rule:

$$\text{Indicator}_t^{\text{P-SMA}(k)} \equiv \frac{\sum_{i=1}^{k}(k-i+1)\Delta P_{t-i}}{\sum_{i=1}^{k}(k-i+1)} = \frac{k\Delta P_{t-1} + (k-1)\Delta P_{t-2} + \ldots + \Delta P_{t-k}}{k + (k-1) + \ldots + 1}. \quad (13)$$

This suggests that alternatively we can interpret the computation of the technical indicator for the Price-Minus-Simple-Moving-Average rule as the computation of the linearly weighted moving average of monthly price changes.

We next consider the Linear Moving Average which uses the linearly weighted moving average or prices. In this case the weight of ΔP_{t-i} is given by

$$x_i = \sum_{j=i}^{k} w_{t-j} = \sum_{j=i}^{k}(k-j+1) = \frac{(k-i+1)(k-i+2)}{2},$$

which is the sum of the terms of arithmetic sequence from 1 to $k-i+1$ with the common difference of 1. As the result, the equivalent representation for the computation of the technical indicator for the Price-Minus-Linear-Moving-Average rule

$$\text{Indicator}_t^{\text{P-LMA}(k)} \equiv \frac{\sum_{i=1}^{k} \frac{(k-i+1)(k-i+2)}{2} \Delta P_{t-i}}{\sum_{i=1}^{k} \frac{(k-i+1)(k-i+2)}{2}}. \quad (14)$$

Then we consider the Exponential Moving Average which uses the exponentially weighted moving average or prices. In this case the weight of ΔP_{t-i} is given by

$$x_i = \sum_{j=i}^{k} w_{t-j} = \sum_{j=i}^{k} \lambda^j = \frac{\lambda}{1-\lambda}\left(\lambda^{i-1} - \lambda^k\right), \quad (15)$$

which is the sum of the terms of geometric sequence from λ^i to λ^k. Consequently, the equivalent presentation for the computation of the technical indicator for the Price-Minus-Exponential-

Moving-Average rule

$$\text{Indicator}_t^{\text{P-EMA}(k)} \equiv \frac{\sum_{i=1}^{k} \left(\lambda^{i-1} - \lambda^k\right) \Delta P_{t-i}}{\sum_{i=1}^{k} \left(\lambda^{i-1} - \lambda^k\right)}. \tag{16}$$

If k is relatively large such that $\lambda^k \approx 0$, then the expression for the computation of the technical indicator for the Price-Minus-Exponential-Moving-Average rule becomes

$$\text{Indicator}_t^{\text{P-EMA}(k)} \equiv \frac{\sum_{i=1}^{k} \lambda^{i-1} \Delta P_{t-i}}{\sum_{i=1}^{k} \lambda^{i-1}} = \frac{\Delta P_{t-1} + \lambda \Delta P_{t-2} + \ldots + \lambda^{k-1} \Delta P_{t-k}}{1 + \lambda + \ldots + \lambda^{k-1}}, \text{ when } \lambda^k \approx 0. \tag{17}$$

In words, the computation of the trading signal for the Price-Minus-Exponential-Moving-Average rule, when k is rather large, is equivalent to the computation of the exponential moving average of monthly price changes. It is worth noting that this is probably the only trading rule where the weighing scheme for the computation of moving average of prices is identical to the weighing scheme for the computation of moving average of price changes.

The weight of ΔP_{t-i} for the Reverse Exponential Moving Average is given by

$$x_i = \sum_{j=i}^{k} w_{t-j} = \sum_{j=i}^{k} \lambda^{k-j} = \frac{1 - \lambda^{k-i+1}}{1 - \lambda},$$

which is the sum of the terms of geometric sequence from 1 to λ^{k-i}. Consequently, the equivalent representation for the computation of the technical indicator for the Price-Minus-Reverse-Exponential-Moving-Average rule

$$\text{Indicator}_t^{\text{P-REMA}(k)} \equiv \frac{\sum_{i=1}^{k} \left(1 - \lambda^{k-i+1}\right) \Delta P_{t-i}}{\sum_{i=1}^{k} \left(1 - \lambda^{k-i+1}\right)}. \tag{18}$$

2.3.4 Moving-Average-Change-of-Direction Rule

The value of this technical trading indicator is based on the difference of two weighted moving averages computed at times t and $t-1$ respectively. We assume that the length of the lookback period is $k-1$ months, the reason for this assumption will become clear very soon. The straightforward computation yields

$$\begin{aligned}\text{Indicator}_t^{\Delta\text{MA}(k\,-\,1)} &= MA_t(k-1) - MA_{t-1}(k-1) = \frac{\sum_{i=0}^{k-1} w_{t-i}P_{t-i}}{\sum_{i=0}^{k-1} w_{t-i}} - \frac{\sum_{i=0}^{k-1} w_{t-i}P_{t-i-1}}{\sum_{i=0}^{k-1} w_{t-i}} \\ &= \frac{\sum_{i=0}^{k-1} w_{t-i}\left(P_{t-i} - P_{t-i-1}\right)}{\sum_{i=0}^{k-1} w_{t-i}} = \frac{\sum_{i=1}^{k} w_{t-i+1}\Delta P_{t-i}}{\sum_{i=1}^{k} w_{t-i+1}}.\end{aligned}$$

Similarly to the alternative representation for the computation of the technical indicator for the Price-Minus-Moving-Average rule (given by (11)), the computation of the technical indicator for the Moving-Average-Change-of-Direction rule is equivalent to the computation of the weighted moving average of monthly price changes:

$$\text{Indicator}_t^{\Delta\text{MA}(k\,-\,1)} \equiv \frac{\sum_{i=1}^{k} w_{t-i+1}\Delta P_{t-i}}{\sum_{i=1}^{k} w_{t-i+1}}. \tag{19}$$

Note that the weighting scheme for the computation of the moving average of monthly price changes is the same as for the computation of moving average of prices. From (19) we easily recover the relationship for the case of the Simple Moving Average where $w_{t-i+1} = 1$ for all i

$$\text{Indicator}_t^{\Delta\text{SMA}(k\,-\,1)} \equiv \frac{\sum_{i=1}^{k} \Delta P_{t-i}}{k} \equiv \text{Indicator}_t^{\text{MOM}(k)}.$$

In the case of the Linear Moving Average, where $w_{t-i+1} = k - i + 1$, we derive a new relationship:

$$\text{Indicator}_t^{\Delta\text{LMA}(k\,-\,1)} \equiv \frac{\sum_{i=1}^{k}(k-i+1)\Delta P_{t-i}}{\sum_{i=1}^{k}(k-i+1)} \equiv \text{Indicator}_t^{\text{P-SMA}}(k),$$

where the last equivalence comes from (13). Putting it into words, the Price-Minus-Simple-Moving-Average rule prescribes investing in the stocks (moving to cash) when the Linear Moving Average of prices over the lookback period of $k-1$ months increases (decreases).

In the case of the Exponential Moving Average and Reverse Exponential Moving Average, the resulting expressions for the Change-of-Direction rules can be written as

$$\text{Indicator}_t^{\Delta\text{EMA}(k\,-\,1)} \equiv \frac{\sum_{i=1}^{k} \lambda^{i-1}\Delta P_{t-i}}{\sum_{i=1}^{k} \lambda^{i-1}}, \quad \text{Indicator}_t^{\Delta\text{REMA}(k\,-\,1)} \equiv \frac{\sum_{i=1}^{k} \lambda^{k-i}\Delta P_{t-i}}{\sum_{i=1}^{k} \lambda^{k-i}}.$$

Observe in particular that if k is rather large, then, using result (17) we obtain yet another new relationship:

$$\text{Indicator}_t^{\text{P-EMA}(k)} \equiv \text{Indicator}_t^{\Delta\text{EMA}(k-1)}, \text{ when } \lambda^k \approx 0.$$

In words, when k is rather large, the Price-Minus-Exponential-Moving-Average rule is equivalent to the Exponential-Moving-Average-Change-of-Direction rule. As it might be observed, for the majority of weighting schemes considered in the paper, there is a one-to-one equivalence between a Price-Minus-Moving-Average rule and a corresponding Moving-Average-Change-of-Direction rule. Therefore, the majority of the moving-average-change-of-direction rules (and may be all of them) can also be expressed as the moving average of Momentum rules.

Finally it is worth commenting that the traders had long ago taken notice of the fact that, for example, very often a Buy signal is generated first by the Price-Minus-Moving-Average rule, then with some delay a Buy signal is generated by the Moving-Average-Change-of-Direction rule. Therefore the traders sometimes use the trading signal of the Moving-Average-Change-of-Direction rule to "confirm" the signal of the Price-Minus-Moving-Average rule (again, see Murphy (1999), Chapter 9). Our analysis provides a simple explanation for the existence of a delay between the signals generated by these two rules. The delay naturally occurs because the Price-Minus-Moving-Average rule overweights more heavily the most recent price changes than the Moving-Average-Change-of-Direction rule computed using the same weighting scheme. Therefore the Price-Minus-Moving-Average rule reacts more quickly to the recent trend changes than the Moving-Average-Change-of-Direction rule.[3]

2.3.5 Double Crossover Method

The relationship between the Double Crossover Method and the Momentum rule is as follows (here we use result (8))

$$\begin{aligned}\text{Indicator}_t^{\text{DCM}(s,k)} &= MA_t(s) - MA_t(k) = (P_t - MA_t(k)) - (P_t - MA_t(s)) \\ &= \frac{\sum_{j=1}^{k} w_{t-j}^k MOM_t(j)}{\sum_{j=0}^{k} w_{t-j}^k} - \frac{\sum_{j=1}^{s} w_{t-j}^s MOM_t(j)}{\sum_{j=0}^{s} w_{t-j}^s}.\end{aligned}$$

[3]Assume, for example, that the trader uses the simple moving average weighting scheme in both the rules. In this case our result says that the Price-Minus-Simple-Moving-Average rule is equivalent to the Linear-Moving-Average-Change-of-Direction rule. As a consequence, it is naturally to expect that the Price-Minus-Simple-Moving-Average rule reacts more quickly to the recent trend changes than the Simple-Moving-Average-Change-of-Direction rule.

Different superscripts in the weights mean that for the same subscript the weights are generally not equal. For example, in case of either linearly weighted moving averages or reverse exponential moving averages $w^k_{t-j} \neq w^s_{t-j}$, yet for the other weighting schemes considered in this paper $w^k_{t-j} = w^s_{t-j}$. In order to get a closer insight into the anatomy of the Double Crossover Method, we assume that one uses the exponential weighting scheme in the computation of moving averages (as it most often happens in practice). In this case the expression for the value of the technical indicator in terms of monthly price changes is given by (here we use results (10) and (15))

$$\begin{aligned}\text{Indicator}_t^{\text{DCM}(s,\,k)} &= \frac{\sum_{j=1}^{k}\lambda^j\sum_{i=1}^{j}\Delta P_{t-i}}{\sum_{j=0}^{k}\lambda^j} - \frac{\sum_{j=1}^{s}\lambda^j\sum_{i=1}^{j}\Delta P_{t-i}}{\sum_{j=0}^{s}\lambda^j} = \frac{\sum_{i=1}^{k}\left(\sum_{j=i}^{k}\lambda^j\right)\Delta P_{t-i}}{\sum_{j=1}^{k}\lambda^j} \\ &- \frac{\sum_{i=1}^{s}\left(\sum_{j=i}^{s}\lambda^j\right)\Delta P_{t-i}}{\sum_{j=1}^{s}\lambda^j} = \frac{\sum_{i=1}^{k}\left(\lambda^i-\lambda^{k+1}\right)\Delta P_{t-i}}{1-\lambda^{k+1}} - \frac{\sum_{i=1}^{s}\left(\lambda^i-\lambda^{s+1}\right)\Delta P_{t-i}}{1-\lambda^{s+1}}.\end{aligned} \tag{20}$$

If we assume in addition that both s and k are relatively large such that $\lambda^s \approx 0$ and $\lambda^k \approx 0$, then we obtain

$$\text{Indicator}_t^{\text{DCM}(s,\,k)} \approx \sum_{i=1}^{k}\lambda^i\Delta P_{t-i} - \sum_{i=1}^{s}\lambda^i\Delta P_{t-i} = \sum_{i=s+1}^{k}\lambda^i\Delta P_{t-i}.$$

The expression above can be conveniently re-written as

$$\text{Indicator}_t^{\text{DCM}(s,\,k)} \equiv \frac{\sum_{i=s+1}^{k}\lambda^{i-s-1}\Delta P_{t-i}}{\sum_{j=s+1}^{k}\lambda^{i-s-1}} \text{ when } k > s, \lambda^s \approx 0, \lambda^k \approx 0.$$

In words, the computation of the trading signal for the Double Crossover Method based on the exponentially weighted moving averages of lengths s and $k > s$, when both s and k are rather large, is equivalent to the computation of the exponentially weighted moving average of monthly price changes, ΔP_{t-i}, for $i \in [s+1, k]$. Note that the most recent s monthly price changes completely disappear in the computation of the technical trading indicator. When the values of s and k are not rather large, the most recent s monthly price changes do not disappear in the computation of the technical indicator, yet the weights of these price changes are reduced as compared to the weight of the subsequent $(s+1)$-th price change.

The final cautionary note on the use of the Double Crossover Method is as follows. A trader might be tempted to use different types of weighting schemes in the shorter and longer moving averages. Yet in this case there is absolutely no guarantee that the weights of the most recent prices changes remain positive. Negative weights of the most recent price changes in the computation of the trading signal can lead to "unforeseen" consequences. One potential consequence in this case is that the time lag between the market action and the generation of a trading signal may increase substantially. Another possibility is that the trading indicator starts generating completely inappropriate Buy and Sell signals.

2.3.6 Discussion

Summing up the results presented above, we reveal that all technical trading indicators considered in this paper are computed in the same general manner. We find, for instance, that the computation of every technical trading indicator can be interpreted as the computation of a weighted average of the momentum rules computed using different lookback periods. Thus, the momentum rule might be considered as an elementary trading rule on the basis of which one can construct more elaborate rules. The most insightful conclusion emerging from our analysis is that the computation of every technical trading indicator can also be interpreted as the computation of the weighted moving average of monthly price changes over the lookback period. This allows us, for example, to establish a one-to-one equivalence between a price-minus-moving-average rule and a corresponding moving-average-change-of-direction rule.

Our main conclusion is that, despite being computed seemingly different at the first sight, the only real difference between miscellaneous rules lies in the weighting scheme used to compute the moving average of monthly price changes. Figure 1 illustrates clearly distinctive weighting schemes for the computations of technical trading indicators based on moving averages. In particular, this figure illustrates the weighting schemes for the Momentum rule, the Price-Minus-Reverse-Exponential-Moving-Average rule (with $\lambda = 0.8$), the Price-Minus-Simple-Moving-Average rule, the Price-Minus-Linear-Moving-Average rule, the Reverse-Exponential-Moving-Average-Change-of-Direction rule (with $\lambda = 0.9$), and the Double Crossover Method (based on using two exponential moving averages with $\lambda = 0.8$). For all technical indicators we use $k = 10$ which means that to compute the value of a technical indicator we use the most recent price change, ΔP_{t-1}, denoted as Lag0, and 9 preceding

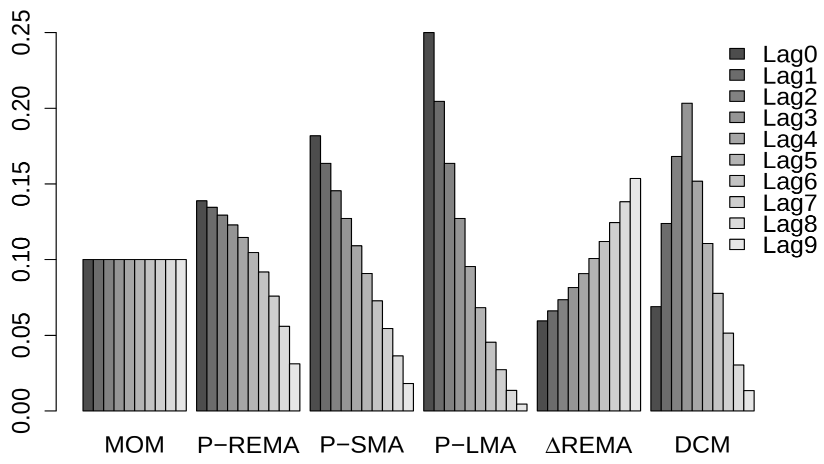

Figure 1: Weights of monthly price changes used for the computations of the technical trading indicators with $k = 10$. **MOM** denotes the Momentum rule. **P-REMA** denotes the Price-Minus-Reverse-Exponential-Moving-Average rule (with $\lambda = 0.8$). **P-SMA** denotes the Price-Minus-Simple-Moving-Average rule. **P-LMA** denotes the Price-Minus-Linear-Moving-Average rule. **ΔREMA** denotes the Reverse-Exponential-Moving-Average-Change-of-Direction rule (with $\lambda = 0.9$). **DCM** denotes the Double Crossover Method (based on using two exponential moving averages with $\lambda = 0.8$ and $s = 3$). **Lag**$(i-1)$ denotes the weight of the lag ΔP_{t-i}, where Lag0 denotes the most recent price change ΔP_{t-1} and Lag9 denotes the most oldest price change ΔP_{t-10}.

lagged price changes up to lag ΔP_{t-10}, denoted as Lag9. In addition, in the computation of the technical indicator for the Double Crossover Method we use $s = 3$.

Apparently, the Momentum rule assigns equal weights to all monthly price changes in the lookback period. The next three rules overweight the most recent price changes. They are arranged according to increasing degree of overweighting. Whereas the Price-Minus-Simple-Moving-Average rule employs the linear weighting scheme, the degree of overweighting in the Price-Minus-Reverse-Exponential-Moving-Average rule can be gradually varied from the equal weighting scheme (when $\lambda = 0$) to the linear weighting scheme (when $\lambda = 1$), see property (2). Formally this can be expressed by

$$\lim_{\lambda \to 0} \text{Indicator}_t^{\text{P-REMA}(k)} = \text{Indicator}_t^{\text{MOM}(k)}, \quad \lim_{\lambda \to 1} \text{Indicator}_t^{\text{P-REMA}(k)} = \text{Indicator}_t^{\text{P-SMA}(k)}.$$

Comparing to the Price-Minus-Simple-Moving-Average rule, a higher degree of overweighing can be attained by using the Price-Minus-Exponential-Moving-Average rule. When $\lambda \approx$

0.85 the degree of overweighing the most recent price changes in the Price-Minus-Exponential-Moving-Average rule is virtually the same as in the Price-Minus-Linear-Moving-Average rule. Therefore, we demonstrate only the weighing scheme in the Price-Minus-Linear-Moving-Average rule. In contrast to the previous rules, the Reverse-Exponential-Moving-Average-Change-of-Direction rule underweights the most recent price changes. Finally, the weighting scheme in the Double Crossover Method underweights both the most recent and the most old price changes. In this weighing scheme the price change $\Delta P_{t-s-1} = \Delta P_{t-4}$ has the largest weight in the computation of moving average.

Our alternative representation of the computation of technical trading indicators by means of the moving average of price changes, together with the graphical visualization of the weighting schemes for different rules presented in Figure 1, reveals a couple of paradoxes. The first paradox consists in the following. Many analysts argue that the most recent stock prices contain more relevant information on the future direction of the stock price than earlier stock prices. Therefore, one should better use the $LMA(k)$ instead of the $SMA(k)$ in the computation of trading signals. Yet in terms of the monthly price changes the application of the Price-Minus-Simple-Moving-Average rule already leads to overweighting the most recent price changes. If it is the most recent stock price changes (but not prices) that contain more relevant information on the future direction of the stock price, then the use of the Price-Minus-Linear-Moving-Average rule leads to a severe overweighting the most recent price changes, which might be suboptimal.

The other paradox is related to the effect produced by the use of a shorter moving average in the computation of a trading signal for the Double Crossover Method. Specifically, our alternative representation of the computation of technical trading indicators reveals an apparent conflict of goals that some analysts want to pursue. In particular, on the one hand, one wants to put more weight on the most recent prices that are supposed to be more relevant. On the other hand, one wants to smooth the noise by using a shorter moving average instead of the last closing price (as in the Price-Minus-Moving-Average rule). It turns out that these two goals cannot be attained simultaneously because the noise smoothing results in a substantial reduction of weights assigned to the most recent price changes (and, therefore, most recent prices). Figure 1 clearly demonstrates that the weighting scheme for the Double Crossover Method has a hump-shaped form such that the largest weight is given to the monthly price change at lag s. Then, as the lag number decreases to 0 or increases to $k-1$, the weight of the lag decreases. Consequently, the use of the Double Crossover Method can be justified only when the price change at lag s contains the most relevant information on the future direction of the stock price.

3 Historical Performance of Trading Rules

3.1 Data

In our empirical study we use the capital appreciation and total return on the Standard and Poor's Composite stock price index, as well as the risk-free rate of return proxied by the Treasury Bill rate. Our sample period begins in January 1860 and ends in December 2009 (150 full years), giving a total of 1800 monthly observations. The data on the S&P Composite index comes from two sources. The returns for the period January 1860 to December 1925 are provided by William Schwert.[4] The returns for the period January 1926 to December 2009 are computed from the closing monthly priced of the S&P Composite index and corresponding dividend data provided by Amit Goyal.[5] The Treasury Bill rate for the period January 1920 to December 2009 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.[6] First, we run a regression

$$\text{Treasury-bill rate}_t = \alpha + \beta \times \text{Commercial Paper Rate}_t + e_t$$

over the period from January 1920 to December 1971. The estimated regression coefficients are $\alpha = -0.00039$ and $\beta = 0.9156$; the goodness of fit, as measured by the regression R-square, amounts to 95.7%. Then the values of the Treasury Bill rate over the period January 1860 to December 1919 are obtained using the regression above with the estimated coefficients for the period 1920 to 1971.

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

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

[6] http://research.stlouisfed.org/fred2/series/M13002US35620M156NNBR

	CAP	**MKT**	**RF**
Mean	6.03	10.46	3.91
St. dev.	17.39	17.40	0.73
Skewness	0.18	0.23	0.93
Kurtosis	8.15	8.38	2.26
Shapiro-Wilk	**0.93**	**0.93**	**0.94**
	(0.00)	(0.00)	(0.00)
AC_1	**0.08**	**0.08**	**0.98**
	(0.00)	(0.00)	(0.00)

Table 1: Descriptive statistics of data used in the study. **CAP**, **MKT**, and **RF** denote the capital appreciation return, the total market return, and the risk-free rate of return respectively. Means and standard deviation are annualized and reported in percents. **Shapiro-Wilk** denotes the value of the test statistics in the Shapiro-Wilk normality test. The p-values of the normality test are reported in brackets below the test statistics. $\mathbf{AC}_1$ denotes the first-order autocorrelation. For each AC_1 we test the hypothesis $H_0 : \text{AC}_1 = 0$. The p-values are reported in brackets below the values of autocorrelation. Bold text indicates values that are statistically significant at the 5% level.

Table 1 summarizes the descriptive statistics for the data used in our study. The results of the Shapiro-Wilk test reject the normality in all data series. In addition, all data series exhibit a statistically significant positive autocorrelation.

3.2 Technical Trading Rules

The goal of this section is to measure and compare the out-of-sample performance of six clearly distinct market timing rules based on moving averages. The weighting schemes for these rules are presented in Figure 1 and include the Momentum rule, Price-Minus-Reverse-Exponential-Moving-Average rule, Price-Minus-Simple-Moving-Average rule, Price-Minus-Linear-Moving-Average rule, Reverse-Exponential-Moving-Average-Change-of-Direction rule, and the Double Crossover Method. In order to compute the value of the trading indicator for the first, third, and forth rule, we need to specify the length of the lookback period k. The value of k is determined dynamically by the out-of-sample simulation procedure that will be described below. The equivalent representations for the computation of the value of the trading indicators for these rules

$$\text{Indicator}_t^{\text{MOM(k)}} = P_t - P_{t-k} \equiv \frac{1}{k}\sum_{i=0}^{k-1} \Delta P_{t-i} \equiv SMA_t(k-1) - SMA_{t-1}(k-1).$$

$$\text{Indicator}_t^{\text{P-SMA(k)}} = P_t - SMA_t(k) \equiv \frac{\sum_{i=0}^{k-1}(k-i)\Delta P_{t-i}}{\sum_{i=0}^{k-1}(k-i)} \equiv LMA_t(k-1) - LMA_{t-1}(k-1).$$

$$\text{Indicator}_t^{\text{P-LMA(k)}} = P_t - LMA_t(k) \equiv \frac{\sum_{i=0}^{k-1} \frac{(k-i)(k-i+1)}{2}\Delta P_{t-i}}{\sum_{i=0}^{k-1} \frac{(k-i)(k-i+1)}{2}}.$$

The value of the technical indicator for the second and fifth rule depends on the length of the lookback period k and the value of the decay factor λ which determines the degree of over-weighting (or under-weighting) the most recent prices. In order to avoid over-optimization in out-of-sample testing, we perform the optimization with respect to k only; the value of λ is held constant through time. In the Price-Minus-Reverse-Exponential-Moving-Average rule we use $\lambda = 0.8$ which provides the degree of overweighing the most recent price changes somewhere in between the degrees provided by the equally weighted and linearly weighted schemes. In the Reverse-Exponential-Moving-Average-Change-of-Direction rule we use $\lambda = 0.9$ which provides a moderate degree of underweighting the most recent price changes. The computation of the value of the trading indicators for these two rules

$$\text{Indicator}_t^{\text{P-REMA(k)}} = P_t - REMA_t(k) \equiv \frac{\sum_{i=0}^{k-1}\left(1-\lambda^{k-i}\right)\Delta P_{t-i}}{\sum_{i=0}^{k-1}\left(1-\lambda^{k-i}\right)} \text{ for } \lambda = 0.8.$$

$$\text{Indicator}_t^{\Delta\text{REMA(k)}} = REMA_t(k) - REMA_{t-1}(k) \equiv \frac{\sum_{i=0}^{k}\lambda^{k-i}\Delta P_{t-i}}{\sum_{i=0}^{k}\lambda^{k-i}} \text{ for } \lambda = 0.9.$$

The value of the technical indicator for the Double Crossover Method depends on the choice of the weighing schemes in two moving averages and the lengths of shorter and longer averages. Since most often analysts use two exponentially weighted moving averages in the Double Crossover Method, we also decide in favour of using two exponentially weighted moving averages with $\lambda = 0.8$. Among analysts, one of the most popular combination is to use 50-day and 200-day averages in this timing rule. Therefore we fix the length of the shorter average to be $s = 2$, whereas the length of the longer average $k > 2$ is determined by the dynamic optimization procedure. Consequently, in our tests the computation of the value of the trading indicator for the Double Crossover Method goes according to

$$\text{Indicator}_t^{\text{DCM(2,k)}} = EMA_t(2) - EMA_t(k) \text{ for } \lambda = 0.8 \text{ and } k > 2.$$

3.3 Transaction Costs

In order to assess the real-life performance of a market timing rule, we need to account for the fact that rebalancing an active portfolio incurs transaction costs. Transaction costs in capital markets consist of the following three primary components: half-size of the quoted bid-ask spread, brokerage fees (commissions), and market impact costs. In addition, there are various taxes, delay costs, opportunity costs, etc. (see, for example, Freyre-Sanders, Guobuzaite, and Byrne (2004)). In our study we consider the average bid-ask half-spread as the only determinant of the one-way transaction costs, and we neglect all other components of transaction costs. Berkowitz, Logue, and Noser (1988), Chan and Lakonishok (1993), and Knez and Ready (1996) estimate the average one-way transaction costs for institutional investors to be in the range of 0.23% to 0.25%. Therefore in our study, we assume that the one-way transaction costs in the stock market amount to 0.25%. Denoting by γ the one-way transaction costs, the return to the market timing strategy over month t is given by

$$r_t = \begin{cases} r_{Mt} & \text{if } (\delta_{t|t-1} = \text{Buy}) \text{ and } \delta_{t-1|t-2} = \text{Buy}), \\ r_{Mt} - \gamma & \text{if } (\delta_{t|t-1} = \text{Buy}) \text{ and } (\delta_{t-1|t-2} = \text{Sell}), \\ r_{ft} & \text{if } (\delta_{t|t-1} = \text{Sell}) \text{ and } (\delta_{t-1|t-2} = \text{Sell}), \\ r_{ft} - \gamma & \text{if } (\delta_{t|t-1} = \text{Sell}) \text{ and } (\delta_{t-1|t-2} = \text{Buy}). \end{cases} \tag{21}$$

3.4 Methodology for Out-of-Sample Testing of Trading Rules

To simulate the returns to the market timing strategy that are given by (21), for each market timing rule we need to compute the value of the technical indicator which provides us with Buy and Sell signals. It is crucial to observe that in order to compute the value of the technical indicator we need to specify the length of the lookback period k. One approach to the choice of k is to use the full historical data sample, simulate the returns to the market timing strategy for different k, and pick up the value of k which produces the best performance. Yet this approach is termed as "data-mining" and the performance of the best trading rule in a back test (that is, in-sample performance) generally severely over-estimates the real-life performance.

It is widely believed that the out-of-sample performance of a trading strategy provides a much more reliable estimate of it's real-life performance as compared with the in-sample performance (see Sullivan et al. (1999), White (2000), and Aronson (2006)). The out-of-sample performance measurement method is based on simulating the real-life trading where a trader has to make a choice of what length of the lookback period k to use given the information about the past performances of the market timing strategy for different values of k. Specifically, the out-of-sample testing procedure begins with splitting the full historical data sample $[1, T]$ into the initial in-sample subset $[1, p]$ and out-of-sample subset $[p + 1, T]$, where T is the last observation in the full sample and p denotes the splitting point. Then the best rule discovered in the mined data (in-sample) is evaluated on the out-of-sample data.

The out-of-sample performance can be evaluated with either a rolling- or expanding-window estimation scheme. A common belief among traders is that, regardless of the choice of historical period, the same specific value of k is optimal for using in a given technical indicator. For example, the majority of traders believe that in the Price-Minus-Simple-Moving-Average trading rule the optimal value of k equals to 10. If the length of the optimal lookback period is constant through time, then it is natural to use the expanding-window estimation scheme to determine the value of k. Out-of-sample simulation of a market timing strategy using an expanding-window estimation of k is performed as follows. The in-sample period of $[1, t]$, $t \in [p, T - 1]$, is used to complete the procedure of selecting the best trading rule given some optimization criterion $O(r_1, r_2, \ldots, r_t)$ defined over the returns to the market timing strategy up to month t.[7] Formally, in our study the choice of the optimal k_t^* is given by

$$k_t^* = \arg \max_{k \in [k^{\min}, k^{\max}]} O(r_1, r_2, \ldots, r_t),$$

where $k^{\min}$ and $k^{\max}$ are the minimum and maximum values for k. Subsequently, the trading signal for month $t+1$ is determined using the lookback period of length k_t^*. Then the in-sample period is expanded by one month, and the best trading rule selection procedure is performed once again using the new in-sample period of $[1, t+1]$ to determine the trading signal for month $t + 2$. This procedure is repeated, by pushing the endpoint of the in-sample period ahead by one month with each iteration of this process, until the trading signal for the last month T is determined.

[7] We follow closely the methodology employed by Lukac, Brorsen, and Irwin (1988), Lukac and Brorsen (1990), and Zakamulin (2014) among others. Note that this methodology has a dynamic aspect, in which the trading rule is being modified over time as the market evolves.

The rolling-window estimation scheme is used when parameter instability is suspected. That is, when the length of the optimal lookback period varies through time. In the rolling-window estimation scheme the choice of k is done using the most recent n observations. In this case the choice of the optimal k_t^* is given by

$$k_t^* = \arg \max_{k \in [k^{\min}, k^{\max}]} O(r_{t-n+1}, r_{t-n+2}, \dots, r_t).$$

We set the value of $k^{\min}$ to be the minimum possible length (measured in the number of lagged prices) of the lookback period for a given trading rule. For the Double Crossover Method $k^{\min} = 3$, for all other rules $k^{\min} = 1$. To select the appropriate value for $k^{\max}$, we studied the most popular recommendations of technical analysts for the choice of the optimal lookback period. In practice, the recommended value for the length k virtually never exceeds 12 months. To be on the safe side, in our empirical study we set $k^{\max} = 24$.

3.5 Choice of performance measure

There is a big uncertainty about what optimization criterion to use in the determination of the best trading rule using the past data. To limit the choice of optimization criteria, we consider an investor who decides whether to follow the passive buy-and-hold strategy or to follow the active market timing strategy. Since the two strategies are supposed to be mutually exclusive, it is natural to employ a reward-to-total-risk performance measure as the optimization criterion. That is, our investor chooses the value of k which maximizes some portfolio performance measure in a back test, that is, using the past (in-sample) data.

The most widely recognized reward-to-risk measure is the Sharpe ratio. Thus, the Sharpe ratio represents the natural optimization criterion to find the best trading rule. The Sharpe ratio uses the mean excess returns as a measure of reward, and the standard deviation of excess returns as a measure of risk. Specifically, the Sharpe ratio of trading strategy i with excess returns $r_{it}^e = r_{it} - r_{ft}$ is computed as (according to Sharpe (1994))

$$SR_i = \frac{\mu(r_i^e)}{\sigma(r_i^e)},$$

where $\mu(r_i^e)$ and $\sigma(r_i^e)$ denote the mean and standard deviation of r_{it}^e respectively.

For the Sharpe ratio of each market timing strategy we report the p-value of testing the null hypothesis that it is equal to the Sharpe ratio of the market portfolio (denoted by SR_M). For this purpose we apply the Jobson and Korkie (1981) test with the Memmel (2003) correction. Specifically, given SR_i, SR_M, and ρ as the estimated Sharpe ratios and correlation coefficient over a sample of size T, the test of the null hypothesis: $H_0 : SR_i - SR_M = 0$ is obtained via the test statistic

$$z = \frac{SR_i - SR_M}{\sqrt{\frac{1}{T}\left[2(1-\rho^2) + \frac{1}{2}(SR_i^2 + SR_M^2 - 2\rho^2 SR_i SR_M)\right]}},$$

which is asymptotically distributed as a standard normal.

As with any reward-to-risk ratio, the use of the Sharpe ratio has some inconveniences. In particular, its value is difficult to interpret, and to decide whether the timing strategy outperforms the market, one also needs to compute the Sharpe ratio of the market portfolio and compare one to the other. To facilitate performance measurement with the Sharpe ratio, we closely follow the method presented by Modigliani and Modigliani (1997) and employ the M^2 measure (Modigliani-Modigliani measure or Modigliani-squared measure). The idea is to mix the active portfolio with a position in the risk-free asset so that the complete portfolio has the same risk as the passive market. The returns to the complete portfolio are

$$r_{it}^* = a(r_{it} - r_{ft}) + r_{ft},$$

where a is the proportion invested in the active portfolio. When $a > 1$ ($a < 1$), it means that the complete portfolio represents a levered (unlevered) version of the original portfolio. The value of a that equates the risk of the complete portfolio with the risk of the market portfolio is

$$a = \frac{\sigma(r_M^e)}{\sigma(r_i^e)}.$$

In a similar manner to Bodie, Kane, and Marcus (2007), we compute the M^2 measure as the difference between the return to the complete portfolio and the return to the market portfolio. As the result, the expression for M^2 measure is given by

$$M_i^2 = \mu(r_i^{*e}) - \mu(r_M^e) = (SR_i - SR_M)\,\sigma(r_M^e).$$

Note that this M^2 measure produces the same ranking of risky portfolios as the Sharpe ratio, but it has the significant advantage of being in units of percent return, which makes it dramatically more intuitive to interpret. Specifically, this measure tells us by how much, in basis points, portfolio i outperformed (if $M_i^2 > 0$) or underperformed (if $M_i^2 < 0$) the market portfolio on a risk-adjusted basis.

Because the Sharpe ratio is often criticized on the grounds that the standard deviation appears to be an inadequate measure of risk, as a robustness test, we also used the Sortino ratio (due to Sortino and Price (1994)) and a few other popular reward-to-risk ratios as the optimization criterion instead of the Sharpe ratio. The results of these tests showed that regardless of the reward-to-risk ratio used, the comparative performance of the active market timing strategy and the passive market strategy remains virtually the same.

3.6 Time-Variations in the Length of the Optimal Lookback Period

As it was mentioned earlier, in the literature on market timing one usually supposes that there is some specific length of the lookback period, k, which is optimal for using in a given technical indicator. Yet there is a big controversy among technical analysts about the optimal value of k. For instance, for the Price-Minus-Simple-Moving-Average rule the recommended value of k varies from 50 to 200 days (see Brock et al. (1992), Sullivan et al. (1999), and Okunev and White (2003)). This common belief, that the optimal lookback period is constant, justifies the use of the expanding-window estimation scheme in the out-of-sample simulation of the trading strategy. And that is why the rolling-window estimation scheme is practically never used. The goal of this section is to check whether this common belief is fallacious or not.

In order to find out how stable the length of the optimal lookback period is for each trading rule, we use a rolling window of n months. For each specific window $[t, t+n]$, for $t \in [1, T-n]$, we find the value of k which maximizes the in-sample performance of a trading rule. Specifically, the optimal k_t^* is given by

$$k_t^* = \arg \max_{k \in [k^{\min}, k^{\max}]} SR(r_t, r_{t+1}, \ldots, r_{t+n}),$$

where $SR(\cdot)$ is the Sharpe ratio computed using the returns $r_t, r_{t+1}, \ldots, r_{t+n}$ to a trading strategy under investigation. Then we report the descriptive statistics of k_t^* for each trading

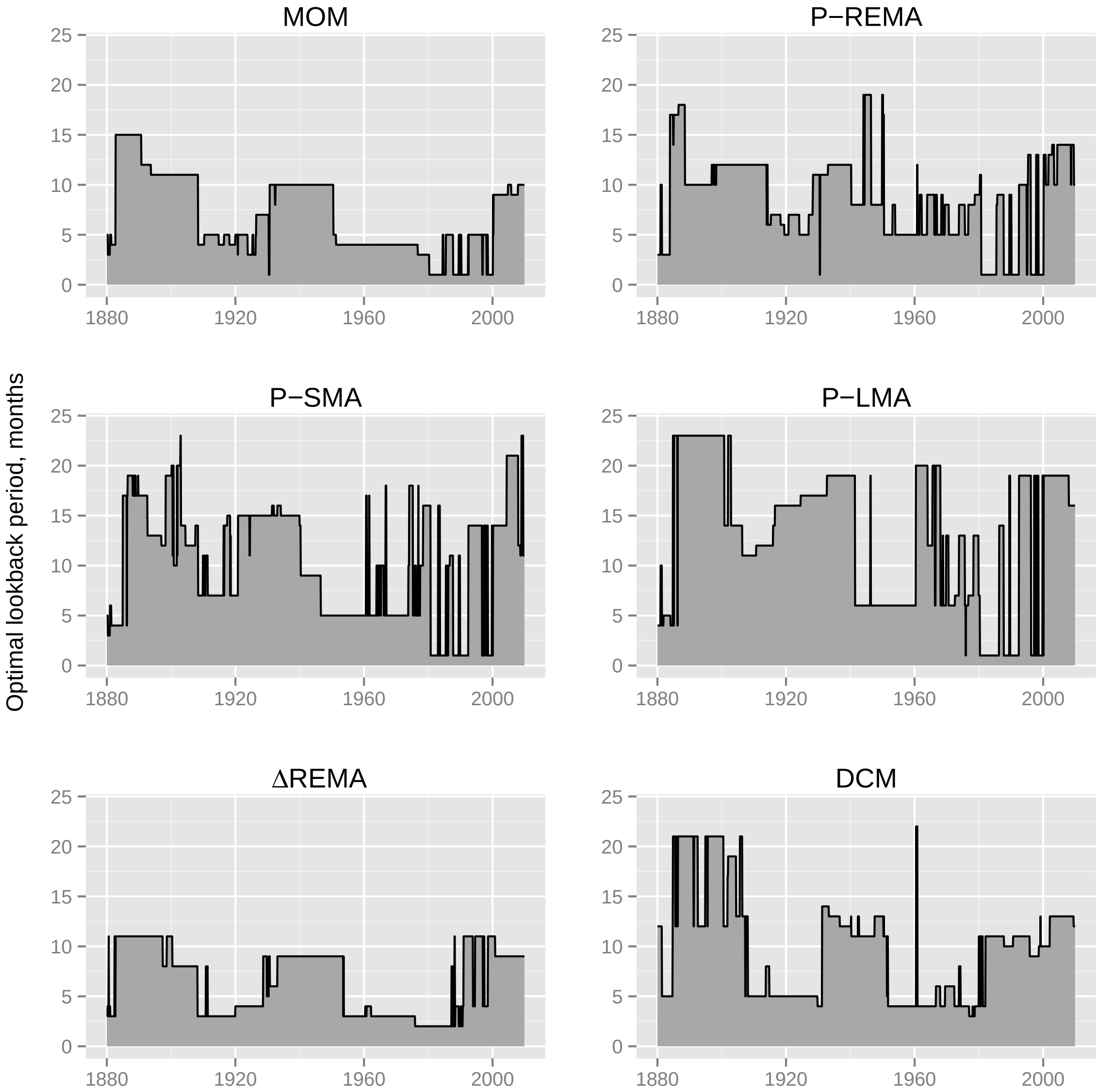

Figure 2: The optimal lookback period, measured in months, for different technical trading rules over a rolling window of 20 years. The first reported value for the optimal lookback period in the graphs is for the period from January 1860 to December 1879. **MOM** denotes the Momentum rule. **P-REMA** denotes the Price-Minus-Reverse-Exponential-Moving-Average rule (with $\lambda = 0.8$). **P-SMA** denotes the Price-Minus-Simple-Moving-Average rule. **P-LMA** denotes the Price-Minus-Linear-Moving-Average rule. **ΔREMA** denotes the Reverse-Exponential-Moving-Average-Change-of-Direction rule (with $\lambda = 0.9$). **DCM** denotes the Double Crossover Method (based on using two exponential moving averages with $\lambda = 0.8$).

rule used in our empirical study.

We need to choose a suitable period length, n, that covers a series of alternating bull and

	MOM	**P-REMA**	**P-SMA**	**P-LMA**	**ΔREMA**	**DCM**
Mean	7.0	8.6	10.4	13.0	6.5	10.0
Median	5	9	11	14	8	11
Std. Deviation	3.9	4.3	5.6	7.1	3.4	5.4
Minimum	1	1	1	1	2	3
Maximum	15	19	23	23	11	22

Table 2: Descriptive statistics of the optimal lookback period (the number of lagged monthly prices) for different technical trading rules over a rolling window of 20 years. **MOM** denotes the Momentum rule. **P-REMA** denotes the Price-Minus-Reverse-Exponential-Moving-Average rule (with $\lambda = 0.8$). **P-SMA** denotes the Price-Minus-Simple-Moving-Average rule. **P-LMA** denotes the Price-Minus-Linear-Moving-Average rule. **ΔREMA** denotes the Reverse-Exponential-Moving-Average-Change-of-Direction rule (with $\lambda = 0.9$). **DCM** denotes the Double Crossover Method (based on using two exponential moving averages with $\lambda = 0.8$).

bear markets. Our choice is $n = 240$ (20 years) and is motivated by the results reported by Lunde and Timmermann (2003). In particular, these authors studied the durations of bull and bear markets using virtually the same dataset as ours. The bull and bear markets are determined as a filter rule θ_1/θ_2 where θ_1 is a percentage defining the threshold of the movements in stock prices that trigger a switch from a bear to a bull market, and θ_2 is the percentage for shifts from a bull to a bear market. Using a 15/15 filter rule, Lunde and Timmermann find that the mean durations of the bull and bear markets are 24.5 and 7.7 months respectively, with the longest bull and bear market durations of about 10 and 2 years respectively. Therefore with the lookback period of 20 years we are guaranteed to cover several alternating bull and bear markets.

The results of our investigation are visualized in Figure 2 and the descriptive statistics of the optimal lookback period (which is the number of lagged monthly prices) for different technical trading rules are reported in Table 2. Specifically, Table 2 reports the descriptive statistics of the optimal lookbacks for our total historical sample from January 1860 to December 2009. This sample sample covers the period of 150 years (1800 months) and, with a 20-year window, includes 1561 different values for the optimal k_t^*, where the first value is for the period from January 1860 to December 1879, the second value is for the period from February 1860 to January 1880, etc.

Apparently, the results suggest that for each technical trading rule there is no single optimal lookback period. On the contrary, the results indicate that there are substantial time-variations in the length of the optimal lookback period. For example, for the most popular Price-Minus-Simple-Moving-Average rule the optimal lookback period varies from 1 to 23 months. Nevertheless, over the total historical sample, the mean value of the optimal lookback period for this rule amounts to 10.4 months which is very close to the most often recommended value of 10 months (200 days). In contrast, in our study the mean value of the optimal lookback period for the Momentum rule is 7.0 months which is substantially lower than the most often recommended value of 12 months (see Moskowitz et al. (2012)).

The results reported in this section have two important implications. First of all, these results challenge the common belief on the constancy of the length of the optimal lookback period. As an immediate consequence, these results advocate that the trading strategy, simulated with the rolling-window estimation of the optimal lookback period, might produce better performance than that with the expanding-window estimation. Secondly, these results propose a simple explanation for the existing big diversity of the popular recommendations concerning the choice of the optimal value for k. Specifically, different recommendations for the value of k appear as the results of finding the best trading rule in the back-test using different historical periods.

3.7 Empirical Results of Performance Measurement

Despite many advantages of the out-of-sample performance measurement method, it has one unresolved deficiency that may seriously corrupt the estimation of the real-life performance of a market timing strategy. The primary concern is that no guidance exists on how to choose the split point between the in-sample and out-of-sample subsets. One possible approach is to choose the initial in-sample segment with a minimum length and use the remaining part of the sample for the out-of-sample test (see Marcellino, Stock, and Watson (2006) and Pesaran, Pick, and Timmermann (2011)). Another potential approach is to do the opposite and reserve a small fraction of the sample for the out-of-sample period (as in Sullivan et al. (1999)). Alternatively, the split point can be selected to lie somewhere in the middle of the sample. In any case, according to conventional wisdom, the out-of-sample performance of a trading strategy provides an unbiased estimate of its real-life performance.

Yet recently, the conventional wisdom about the unbiased nature of traditional out-of-sample testing has been challenged. In the context of out-of-sample forecast evaluation, Rossi and Inoue (2012) and Hansen and Timmermann (2013) report that the results of out-of-sample forecast tests depend significantly on how the sample split point is determined. Zakamulin (2014) also demonstrates that the out-of-sample performance of market timing strategies depends critically on the choice of a split point. The primary reason why the choice of split point sometimes dramatically affects the out-of-sample performance of the market timing strategy lies in the fact that the performance of market timing strategies is highly non-uniform. Generally, a market timing strategy under-performs the passive strategy during bull markets and shows a superior performance during severe bear markets. According to Lunde and Timmermann (2003), the mean bull market duration exceeds the mean bear market duration by a factor of 1.5-2.3, depending on the filter size in the bull-bear detection algorithm. As a result, one has to expect that, over short-term horizons, most of the time a market timing strategy under-performs the market to some extent, but occasionally it delivers an extraordinary out-performance. In addition, a bull market might last over the course of a decade (again, this number depends on the filter size). Therefore, as argued in Zakamulin (2014), in the out-of-sample testing one has to choose the initial in-sample segment to have a minimum length. Otherwise, as in the tests performed by Sullivan et al. (1999), the whole out-of-sample period may be qualified as a bull market which leads to an erroneous conclusion that market timing does not work at all.

Motivated by the discussion above, we choose the length of the initial in-sample period to be 10 years. Specifically, in our tests the initial in-sample period is from January 1860 to December 1869. Consequently, our out-of-sample period is from January 1870 to December 2009 which spans 140 years. We perform the simulation of the returns to six market timing rules using both expanding- and rolling-window estimation schemes to determine the lookback period length. In the latter case the length of the rolling window is of 10 years. However, it is worth noting that, in principle, the performance of the market timing rule implemented with a rolling-window estimation scheme depends on the length of the rolling window. As a matter of fact, we tested different lengths of the rolling window (in the interval $n \in [2, 20]$ years) and our experiments showed that the performance of a market timing strategy varies insignificantly as long as the length of the rolling window exceeds 5 years. That is, our experiments indicated that decreasing the length of a rolling window to a period shorter than 5 years usually substantially deteriorates the performance of a market timing strategy.

We begin presenting the results of simulations and performance measurement using the traditional method. Table 3 reports the descriptive statistics and performances of the active trading rules and the passive market portfolio as well. Specifically, this table reports the means, standard deviations, skewness, and minimum and maximum of monthly returns. In addition, this table reports the Sharpe ratio of each strategy. For each market timing strategy we test the hypothesis that its Sharpe ratio is equal to the Sharpe ratio of the passive market strategy.

Our first observation is that, judging by the Sharpe ratios, practically every market timing strategy outperforms the passive market strategy on the risk-adjusted basis. Only the performance of the market timing strategy based on the Reverse-Exponential-Moving-Average-Change-of-Direction rule, implemented using the expanding-window estimation scheme, is worse than that of the passive strategy. Yet only two trading strategies exhibit performances that are statistically significantly different from the passive market performance. These strategies are based on the Momentum rule and the Price-Minus-Reverse-Exponential-Moving-Average rule, and both of the strategies are simulated using the rolling-window estimation scheme.

Our second observation is that, contrary to the common belief, neither over-weighting nor under-weighting the recent price changes improves the performance of a market timing strategy based on moving averages. We find that the Momentum rule, where the price changes are equally weighted, produces the best performance in the out-of-sample tests. An instructive way of paraphrasing this result is to say "the simplest solution is often the best solution".

Our third observation is that we find indications that the use of the rolling-window estimation scheme, to determine the length of the optimal lookback period, produces better performance than the use of the expanding-window estimation scheme. This is valid for 4 out of 6 trading rules used in our study. As a matter of fact, this finding is not surprising given the fact that there is no single optimal lookback period for a trading rule. Yet the difference between performances produced by the rolling- and expanding-window estimation schemes are not substantial.

Our forth observation is about the descriptive statistics of the returns to market timing strategies. All market timing strategies are virtually equally risky, the standard deviation of monthly returns varies between 3.2% and 3.5%. We observe a significant risk reduction compared to the riskiness of the passive market portfolio. However, the reduction of risk is not

	MKT	MOM		P-REMA		P-SMA		P-LMA		ΔREMA		DCM	
		Rol	**Exp**	**Rol**	**Exp**	**Rol**	**Exp**	**Rol**	**Exp**	**Rol**	**Exp**	**Rol**	**Exp**
Mean	0.84	0.83	0.77	0.79	0.78	0.74	0.78	0.75	0.75	0.71	0.66	0.73	0.75
St. dev.	5.03	3.42	3.41	3.23	3.25	3.27	3.31	3.29	3.29	3.31	3.47	3.29	3.30
Skewness	0.29	0.81	0.69	-0.12	-0.44	-0.39	-0.49	-0.39	-0.36	-0.61	0.49	-0.54	-0.51
Minimum	-29.42	-23.52	-23.52	-21.54	-23.52	-23.52	-23.52	-23.52	-23.52	-23.52	-23.52	-23.52	-23.77
Maximum	42.91	42.66	42.66	16.10	16.10	16.10	16.10	16.10	16.10	13.46	42.66	13.46	16.10
Sharpe ratio	0.36	**0.52**	0.46	**0.50**	0.49	0.44	0.49	0.46	0.45	0.40	0.33	0.43	0.45
p-value	-	(0.02)	(0.13)	(0.05)	(0.07)	(0.22)	(0.07)	(0.16)	(0.18)	(0.49)	(0.75)	(0.28)	(0.18)

Table 3: Descriptive statistics and performances of the trading strategies. For each trading rule and estimation scheme we simulate real-life technical trading over the period January 1870 to December 2009; then we compute the descriptive statistics of the passive market strategy and each active trading strategy. The descriptive statistics is for monthly returns, means and standard deviation are reported in percentages. The Sharpe ratios are annualized; the p-values of testing the null hypothesis of their equality to the Sharpe ratio of the market portfolio are reported in brackets. The bold text indicates values that are statistically significant at the 5% level. **MKT** denotes the passive market strategy. **MOM** denotes the Momentum rule. **P-REMA** denotes the Price-Minus-Reverse-Exponential-Moving-Average rule (with $\lambda = 0.8$). **P-SMA** denotes the Price-Minus-Simple-Moving-Average rule. **P-LMA** denotes the Price-Minus-Linear-Moving-Average rule. **ΔREMA** denotes the Reverse-Exponential-Moving-Average-Change-of-Direction rule (with $\lambda = 0.9$). **DCM** denotes the Double Crossover Method (based on using two exponential moving averages with $\lambda = 0.8$). **Rol** and **Exp** denote the rolling- and expanding-window estimation schemes respectively.

surprising because virtually in any market timing strategy about 1/3 of the time the money are held in cash. The mean returns to a market timing strategy are also below the mean returns to the market portfolio. Yet for the majority of timing rules the decrease in mean returns is lesser than the decrease in risk which improves the risk-adjusted performance of a market timing strategy. We also observe that whereas the returns to the market portfolio are positively skewed, the returns to the majority of market timing strategies are negatively skewed. This suggests that, while the passive strategy has a higher variation on gains, the market timing strategy typically has a higher variation on losses. The comparison of the maximum and minimum monthly returns to the market portfolio and those to the market timing strategies allows us to explain the reason for negative skewness. Note that the majority of the market timing rules lets the "big downward mover" months pass through, but misses the "big upward mover" months.

Now we turn to the alternative presentation of the performance of the market timing strategies. This alternative presentation is suggested by Zakamulin (2014) and motivated as follows. Zakamulin (2014) agues that the traditional performance measurement using a single number (for example, the value of a Sharpe ratio) estimated over very long out-of-sample period (which is beyond the investment horizon of most individual investors) is very misleading for investors with short- and intermediate-term investment horizons. This is because a single number for performance creates a wrong impression that performance is time-invariant, whereas in reality it varies dramatically over time. Thus, it is impossible to provide an accurate picture of market timing performance without taking into account the time-varying nature of performance. With this fact in mind, instead of providing a single number for the performance of a market timing strategy over a very long historical period, we measure the performance over shorter N-year disjoint periods, and then provide the descriptive statistics of the historical performance over these N-year periods.

Because our primer interest is to find out whether a market timing strategy can beat the market, we always need to compare the performance of a market timing strategy with that of the market. To simplify the performance comparison in this case, we employ the Modigliani-Modigliani measure. To illustrate the fact that the market timing performance is very uneven over time, Figure 3 plots the annualized M^2 performance measure computed over disjoint intervals of 5 years. The first period is from January 1870 to December 1874, the second period is from January 1875 to December 1879, etc. Then we plot the value of M^2 measure versus the historical period. The plots in this figure clearly indicate that the superior performance of market timing was generated mainly over relatively few particular historical episodes. Specifically, they are the severe bear markets of the decades of 1870s, 1900s, 1930s, 1970s, and finally 2000s. Many market timing rules consistently under-performed the market over the course of several decades. For example, the most popular Price-Minus-Simple-Moving-Average rule under-performed the market over the period from early 1930s to late 1960s, and then over the period from late 1970s to the beginning of 2000s.

The alternative presentation of the performance of the market timing rules is reported in Table 4. In particular, this table reports the descriptive statistics of the annualized M^2 performance measure over medium- to long-term investment horizons of 5 and 10 years. First of all, the table reports the mean value of M^2, which reflects the average performance of a market timing strategy over an N-year horizon. In addition to the mean, the table reports the standard deviation, which reflects the variability of M^2, as well as the minimum and maximum values, which define the range of possible values for M^2. The table also reports the quartiles of distribution of M^2. These quartiles are supposed to help investors to roughly estimate the frequency distribution of the performance of a market timing strategy. We remind the reader that quartiles are the three points that divide a ranked set of data values into four equal parts. The first quartile is the number below which lies the bottom 25% of data. Presumably, the probability that the performance of a market timing strategy over an N-year horizon will be below the first quartile equals 25%. The second quartile (the median) divides the range in the middle and has 50% of the data below it. Thus, the probability that the performance of a market timing strategy over an N-year horizon will be below the median equals 50%. The third quartile has 75% of the data below it and the top 25% of the data above it. In addition, this table reports the probability that a market timing strategy outperforms the passive strategy over an N-year investment horizon. Finally, the table reports two conditional mean performances: the mean out-performance, $E[M^2|M^2 > 0]$, which is the mean value of M^2 conditional on $M^2 > 0$, and the mean under-performance, $E[M^2|M^2 < 0]$, which is the mean value of M^2 conditional on $M^2 < 0$.

First, we interpret the descriptive statistics for the performance of market timing strategies over a 5-year horizon. Over this horizon, 4 out of 6 timing strategies exhibit a positive mean

	MOM		P-REMA		P-SMA		P-LMA		ΔREMA		DCM	
	Rol	**Exp**	**Rol**	**Exp**	**Rol**	**Exp**	**Rol**	**Exp**	**Rol**	**Exp**	**Rol**	**Exp**
Panel A : 5-year investment horizon												
Minimum	-8.39	-5.91	-6.82	-5.24	-7.37	-8.14	-7.07	-7.38	-16.64	-9.55	-17.48	-6.46
1st Quartile	-1.80	-4.04	-1.68	-1.79	-2.89	-2.55	-3.25	-2.27	-4.59	-4.41	-4.21	-3.10
Median	0.73	0.65	-0.49	-0.57	-0.42	-0.30	-0.71	-0.85	-2.20	-2.20	-1.39	-1.77
Mean	2.07	0.83	1.18	0.88	0.19	0.89	0.27	0.29	-0.74	-1.40	-0.25	0.31
3rd Quartile	3.35	2.70	2.23	4.07	3.28	3.81	3.35	1.47	2.93	0.83	3.90	4.45
Maximum	28.78	20.88	15.34	11.56	13.10	14.21	11.64	11.47	11.83	9.47	10.30	12.86
Std. Deviation	6.91	5.90	4.85	4.32	4.42	5.30	4.70	4.71	6.23	5.08	5.96	4.84
Outperf. Prob.	57.14	53.57	46.43	46.43	42.86	42.86	42.86	35.71	35.71	28.57	46.43	39.29
Mean Underperf.	-2.96	-3.76	-2.11	-2.38	-2.75	-2.60	-3.06	-2.42	-4.49	-3.98	-4.62	-2.96
Mean Outperf.	5.65	4.81	4.97	4.63	4.10	5.55	4.71	5.16	6.01	5.06	4.79	5.35
Panel B : 10-year investment horizon												
Minimum	-3.93	-4.35	-2.20	-2.92	-2.84	-4.10	-3.38	-4.33	-5.18	-5.81	-4.56	-4.50
1st Quartile	-0.57	-2.76	-0.43	-0.86	-1.06	-1.21	-1.49	-1.52	-1.54	-3.21	-2.58	-1.25
Median	1.59	0.92	1.08	0.93	-0.06	0.27	0.17	-0.18	0.02	-2.43	-0.37	-0.44
Mean	2.33	1.14	1.58	1.28	0.56	1.25	0.72	0.70	-0.01	-1.00	0.26	0.75
3rd Quartile	4.62	2.50	1.74	2.92	1.12	4.04	1.64	2.97	1.60	-0.07	2.08	3.11
Maximum	15.86	11.70	8.54	7.48	9.00	9.58	8.30	7.48	5.02	7.42	8.04	6.61
Std. Deviation	4.89	4.55	3.03	3.02	2.97	3.86	3.40	3.34	3.08	4.14	3.85	3.45
Outperf. Prob.	57.14	64.29	64.29	57.14	50.00	57.14	50.00	50.00	50.00	28.57	42.86	42.86
Mean Underperf.	-1.38	-3.30	-1.05	-1.40	-1.33	-1.88	-1.68	-1.86	-2.42	-3.11	-2.36	-1.66
Mean Outperf.	5.11	3.60	3.04	3.28	2.46	3.60	3.12	3.26	2.40	4.26	3.76	3.96

Table 4: Descriptive statistics of the annualized M^2 performance measure. For each trading rule and estimation scheme we simulate real-life technical trading and compute the descriptive statistics of M^2 measure over 5-year and 10-year non-overlapping periods. The values of all descriptive statistics are reported in percents. **Outperf. Prob.** denotes the historical probability of outperformance; **Mean Underperf.** denotes the expected performance if trading rule underperforms the market; **Mean Outperf.** denotes the expected performance if trading rule outperforms the market. **MOM** denotes the Momentum rule. **P-REMA** denotes the Price-Minus-Reverse-Exponential-Moving-Average rule (with $\lambda = 0.8$). **P-SMA** denotes the Price-Minus-Simple-Moving-Average rule. **P-LMA** denotes the Price-Minus-Linear-Moving-Average rule. **ΔREMA** denotes the Reverse-Exponential-Moving-Average-Change-of-Direction rule (with $\lambda = 0.9$). **DCM** denotes the Double Crossover Method (based on using two exponential moving averages with $\lambda = 0.8$).

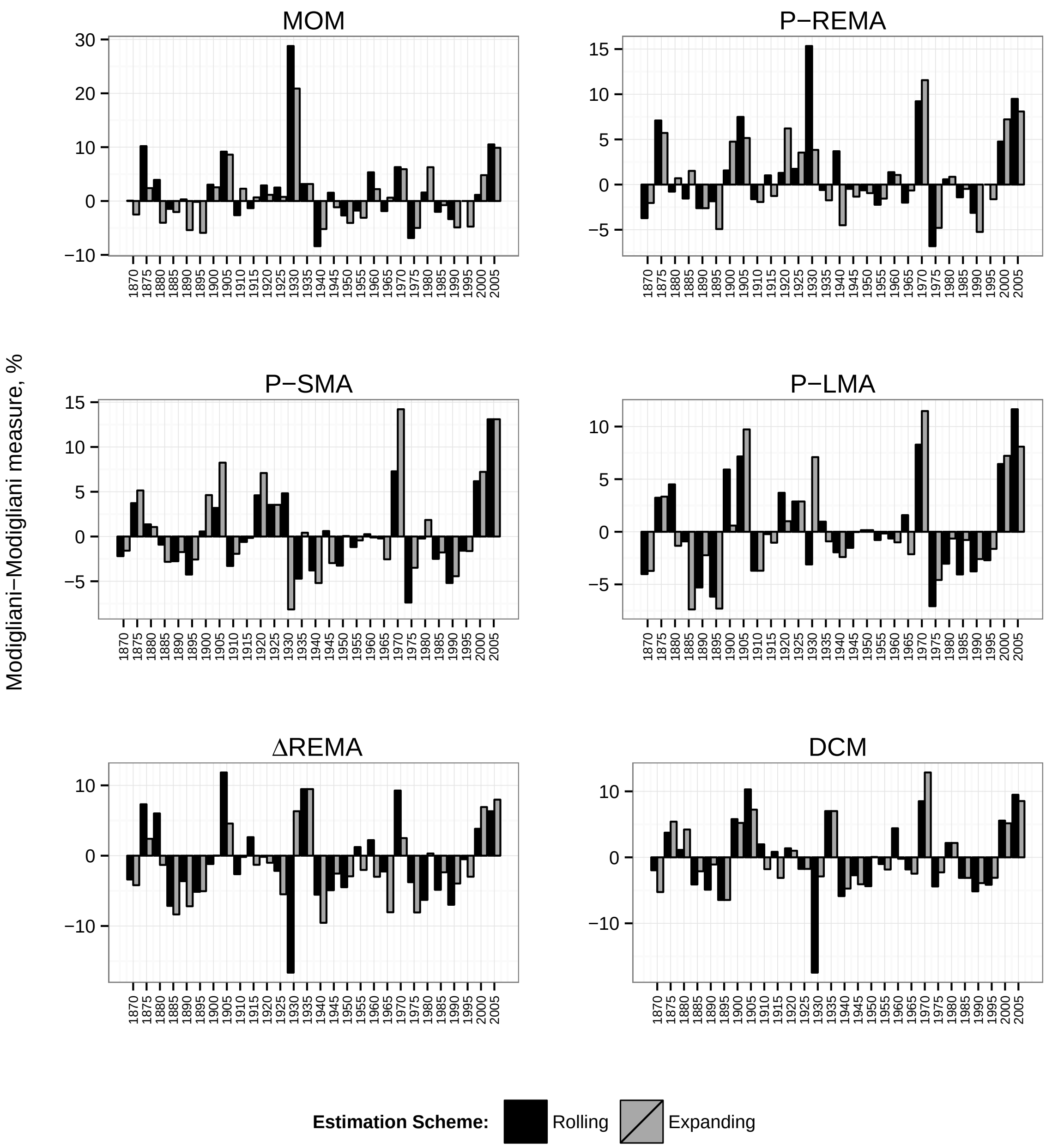

Figure 3: Annualized M^2 performance measure computed over disjoint intervals of 5-years. For each trading rule and estimation scheme we simulate real-life technical trading and compute the out-of-sample performance of market timing strategies over 5-year periods. The first period is from January 1870 to December 1874, the second period is from January 1875 to December 1879, etc. Then we plot the value of M^2 measure versus the historical period. The values of M^2 are reported in percents. **MOM** denotes the Momentum rule. **P-REMA** denotes the Price-Minus-Reverse-Exponential-Moving-Average rule (with $\lambda = 0.8$). **P-SMA** denotes the Price-Minus-Simple-Moving-Average rule. **P-LMA** denotes the Price-Minus-Linear-Moving-Average rule. **ΔREMA** denotes the Reverse-Exponential-Moving-Average-Change-of-Direction rule (with $\lambda = 0.9$). **DCM** denotes the Double Crossover Method (based on using two exponential moving averages with $\lambda = 0.8$).

value for M^2. That is, the majority of market timing rules outperform the market on average. Yet only the Momentum rule has a positive median value for M^2. In words, this means that only the Momentum rule outperforms the market more than 50% of time over a 5-year horizon. For all other trading rules the probability of outperformance is less than 50%. When the mean is larger than the median, this means that the probability distribution of M^2 is positively skewed. This implies that on average the out-performance is greater than under-performance. This is confirmed by the mean conditional performance values. Specifically, for every trading rule $E[M^2|M^2 > 0] > \left|E[M^2|M^2 < 0]\right|$, that is, the mean value of performance, conditional on the trading rule outperforms the market, is greater than the absolute value for the mean performance, conditional on the trading rule underperforms the market. The variability of the performance of the Momentum rule is the largest one. However, this is because the Momentum rule showed an extraordinary good performance over the period January 1930 to December 1934. As a matter of fact, the value of the performance measure for this period, in statistical terms, should be considered as an outlier because this value is very distant from the other values for the performance measure.

The comparison of the descriptive statistics for the performance of market timing strategies over a 10-year horizon with that over a 5-year horizon reveals that increasing the investment horizon increases the chances that a market timing strategy outperforms the market on a risk-adjusted basis. Specifically, in this case for 4 out of 6 market timing strategies the outperformance probability is equal or above 50%. And over this horizon, 5 out of 6 timing strategies exhibit a positive mean value for M^2. Yet the ranking of the trading rules, according to their performance, remains the same regardless of the length of the investment horizon. The best performance is delivered by the Momentum rule; the second best by the Price-Minus-Reverse-Exponential-Moving-Average rule; the worst one by the Reverse-Exponential-Moving-Average-Change-of-Direction rule.

4 Conclusion

In this paper, we presented the methodology to study the computation of trading indicators in many market timing rules based on moving averages and analyzed the commonalities and differences between the rules. Our analysis revealed that the computation of every technical trading indicator considered in this paper can be equivalently interpreted as the computation of the weighted moving average of price changes over the lookback period. Despite a variety of rules that look seemingly different at the first sight, we found that the only real difference between the rules lies in the weighting scheme used to compute the moving average of price changes. The most popular trading indicators employ either equal-weighing of price changes, over-weighting the most recent price changes, or a hump-shaped weighting scheme with under-weighting both the most recent and most distant price changes. The trading rules basically vary only by the degree of over- and under-weighting the most recent price changes.

We also performed the longest out-of-sample testing of a few distinct trading rules in order to find out whether the real-life performance of market timing strategies can support the existing myths and common beliefs about market timing. The results of this testing are as follows. First, contrary to the common belief, our results indicated that there is no single optimal lookback period in each trading rule. Second, we found no support for the common belief that over-weighting the recent prices allows one to improve the performance of a market timing rule. Our results suggested that equal weighing of price changes is the most optimal weighting scheme to use in market timing. Third, we did find support for the claim that one can beat the market by timing it. Yet the chances for beating the market depend on the length of the investment horizon. Whereas over very long-term horizons the market timing strategy is almost sure to outperform the market on a risk-adjusted basis, over more realistic medium-term horizons the market timing strategy is equally likely to outperform as to underperform. Yet we found that the average outperformance is greater than the average underperformance.

References

Aronson, D. (2006). *Evidence-Based Technical Analysis: Applying the Scientific Method and Statistical Inference to Trading Signals.* John Wiley & Sons, Ltd.

Berkowitz, S. A., Logue, D. E., and Noser, E. A. (1988). "The Total Costs of Transactions on the NYSE", *Journal of Finance*, *43*(1), 97–112.

Bodie, Z., Kane, A., and Marcus, A. J. (2007). *Investments.* McGraw Hill.

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.

Chan, L. K. C. and Lakonishok, J. (1993). "Institutional Trades and Intraday Stock Price Behavior", *Journal of Financial Economics*, *33*(2), 173–199.

Freyre-Sanders, A., Guobuzaite, R., and Byrne, K. (2004). "A Review of Trading Cost Model: Reducing Transaction Costs", *Journal of Investing*, *13*(3), 93–116.

Hansen, P. R. and Timmermann, A. (2013). "Choice of Sample Split in Out-of-Sample Forecast Evaluation", Working paper, European University Institute, Stanford University and CREATES.

Jobson, J. D. and Korkie, B. M. (1981). "Performance Hypothesis Testing with the Sharpe and Treynor Measures", *Journal of Finance*, *36*(4), 889–908.

Kirkpatrick, C. D. and Dahlquist, J. (2010). *Technical Analysis: The Complete Resource for Financial Market Technicians.* FT Press; 2nd edition.

Knez, P. J. and Ready, M. J. (1996). "Estimating the Profits from Trading Strategies", *Review of Financial Studies*, *9*(4), 1121–1163.

Lo, A. W., Mamaysky, H., and Wang, J. (2000). "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation", *The Journal of Finance*, *55*(4), 1705–1770.

Lukac, L. P. and Brorsen, B. W. (1990). "A Comprehensive Test of Futures Market Disequilibrium", *Financial Review*, *25*(4), 593–622.

Lukac, L. P., Brorsen, B. W., and Irwin, S. H. (1988). "A Test of Futures Market Disequilibrium Using Twelve Different Technical Trading Systems", *Applied Economics*, *20*(5), 623–639.

Lunde, A. and Timmermann, A. (2003). "Duration Dependence in Stock Prices: An Analysis of Bull and Bear Markets", *Journal of Business and Economic Statistics*, *22*(3), 253–273.

Marcellino, M., Stock, J. H., and Watson, M. W. (2006). "A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series", *Journal of Econometrics*, *135*(1-2), 499–526.

Memmel, C. (2003). "Performance Hypothesis Testing with the Sharpe Ratio", *Finance Letters*, *1*, 21–23.

Modigliani, F. and Modigliani, L. (1997). "Risk-Adjusted Performance", *Journal of Portfolio Management*, *23*(2), 45–54.

Moskowitz, T. J., Ooi, Y. H., and Pedersen, L. H. (2012). "Time Series Momentum", *Journal of Financial Economics*, *104*(2), 228–250.

Murphy, J. J. (1999). *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications.* New York Institute of Finance.

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.

Park, C.-H. and Irwin, S. H. (2007). "What Do We Know About the Profitability of Technical Analysis?", *Journal of Economic Surveys*, *21*(4), 786–826.

Pesaran, M. H., Pick, A., and Timmermann, A. (2011). "Variable Selection, Estimation and Inference for Multi-Period Forecasting Problems", *Journal of Econometrics*, *164*(1), 173–187.

Rossi, B. and Inoue, A. (2012). "Out-of-Sample Forecast Tests Robust to the Choice of Window Size", *Journal of Business and Economic Statistics*, *30*(3), 432–453.

Sharpe, W. F. (1994). "The Sharpe Ratio", *Journal of Portfolio Management*, *21*(1), 49–58.

Sortino, F. A. and Price, L. N. (1994). "Performance Measurement in a Downside Risk Framework", *Journal of Investing*, *3*(3), 59 – 65.

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.

White, H. (2000). "A Reality Check for Data Snooping", *Econometrica*, *68*(5), 1097–1126.

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.