UNIVERSITY OF AGDER

True Performance of Market Timing with Simple Moving Average

Jakob Ask

Supervisor
Valeriy Zakamulin

University of Agder, 2014
School of Business and Law

Abstract

In this paper we investigate the performance of the market timing strategy based on Simple Moving Average (SMA), looking at its performance when tested both in and out-of-sample. We examine whether or not its popularity is a result of the inherent flaws of in-sample testing or if its based on actual superiority. We find that the SMA strategy outperforms the market only when tested in-sample, and that there are too many uncertain factors to be able to conclude that it is effective when tested out-of-sample. As such, we do not find evidence of the SMA strategy conflicting with the efficient market hypothesis.

Contents

1 Introduction

Throughout time the financial markets have experienced some pretty serious turmoil, they have gone through periods of rapid decline and at times even crashing down to a mere factor of what they previously were, to soaring through the roof again after a few years. Examples of this are plentiful, most notable are probably the Great Depression of the 1930s and the crash of 2008, whom both affected markets world wide and lead to long periods of general economic decline which in the latter case turned into a global crisis with countries such as Greece and Iceland filing for bankruptcy.

People have tried to circumvent this by coming up with different trading strategies that can protect them from the harsh downfalls of the market, while still allowing them to experience the booms in the market. This has lead to some strategies, such as the market timing strategy, becoming highly popular.

The market timing strategy is an active strategy that attempts to outperform the passive buy-and-hold strategy by relying on previous market movements to predict the movement of future prices. In other words it relies on the trends you can find in the security prices in order to send a signal to either buy or sell a security before it experiences a major shock. A market timing strategy is reliant on the predictability of future prices, if the stock prices are truly random it would not be effective, as it has to be able to take advantage of trends in order to outperform the buy-and-hold strategy.

Over the course of the last 10 years multiple research papers have praised the market timing strategy for its ability to outperform the market, giving a significantly higher risk-adjusted return than the passive strategy. This is a clear contradiction to the efficient market hypothesis that states it is impossible to beat the market as stock prices incorporate and reflect all relevant information. One can not predict the future stock prices by looking for trends in the market. The efficient market hypothesis is a highly contested hypothesis, but is still considered to be a

cornerstone in modern financial theory.

The most popular of the market timing strategies is the moving average strategy. The moving average strategy computes a mean value of the past k months security prices, weighted in different manners depending on which of the moving average strategies is used. This average is then compared to the stock prices for the current period in order to create either a buy or sell signal depending on whether the stock prices are higher or lower than the average. Amongst the moving average strategies the simple moving average (SMA) strategy is the most popular of them, weighing all the previous stock prices equally.

There are some important problems with the previous research about the efficiency of the SMA strategy, first of all many of these papers only use an in-sample test. A major issue with in-sample testing is what we call data-mining bias, the fact that you base your results entirely on a strategy tested on the data set used to create it means the results are far less meaningful. A key point in getting reliable statistics is not to test a hypothesis with the data used to create it, as every data set will contain some correlations simply by random chance. This again will lead to a significant decline in the reliability of the data you get, say if you for example flip a coin 4 times and get 1 tails and 3 heads. It would seem like you have a 3/4 chance of getting heads. If you then tested this on your original data set you would see that you did indeed have a 3/4 chance of getting tails, which obviously is not correct.

Another part of the data-mining bias is that when testing a strategy on a in-sample data set you are privy to the results other authors have gotten while testing the same period, and as such will be influenced by them regarding what strategy is the best. As such you may end up not testing all the options available to you and only end up replicating what others have done before. Secondly, problems such as out-of-sample performance deterioration means that the results one gets while performing a in-sample test do not reflect the results one can expect to get once you apply the same trading rules to an out-of-sample data set (Aronson 2006,chapter 6). The third prominent problem with the previous research done is that they all examine the same time period, most going no further back than 1926 with some only going to 1950 or 1970.

The goal of our paper is to examine the effectiveness of the most popular moving average strategy, a 10 month in-sample SMA strategy, by testing it on a data set that expands further back in time than previous research, with a time horizon from 1857 to 2011. In addition to this we will also perform a out-of-sample test for the SMA strategy in order to see if it is actually possible to beat the market using a market timing strategy, or if this has only been achievable due to the flaws of in-sample testing. Out-of-sample testing should alleviate most of the data-mining problems associated with in-sample testing, and as such be a better representation of the real life performance of the market timing strategy. The SMA strategy has been tested on out-of-sample before, but then only for very short periods of time. We will have out-of-sample periods exceeding 100 years for most of this paper.

We find that the SMA strategy tested in-sample was able to generate the same returns as the buy-and-hold strategy while severely reducing both volatility and drawdowns, leading to a clearly superior risk-adjusted return. This matches the previous research done. The out-of-sample market timing strategy was able to generate Sharpe ratios that were higher for the active strategy for all sub-periods we tested. The performance of the market timing strategy improved the further back in time we went, suggesting that the efficiency of the market timing strategy is deteriorating. The markets are become more and more efficient. The moving average strategy would have been effective in the past when the markets were less efficient. However, when you take into account transaction costs, taxes and the decline of the efficiency of the model we do not believe the SMA strategy would be able to consistently outperform the passive strategy in the present. It is too fragile and too dependent on factors you have no control over to be able to say with confidence that it will perform better than the passive strategy when you take all factors into account. As such we do not find conclusive evidence that speaks against the efficient market hypothesis.

The remainder of the paper is organized as follows. Section 2 will go through relevant literature on the subject in order to give perspective on our research and how we are contributing to the subject. In Section 3 we will describe our data and our sources, as well as providing the descriptive statistics of the data sets. The empirical models used will be presented in the Section 4. We will provide tables and figures from our study in Section 5 as well as discussing these results in light of what others have have been able to produce. Section 6 will conclude our paper.

2 Literature Review

Even though technical trading analysis has been used by practitioners for a long time, it has been held in widespread skepticism by academics until most recently. For technical trading to be able to outperform the market it has to break with the efficient market hypothesis, a hypothesis held in high regards by most academics. One of the first papers that marked this change in the academic viewpoint was the paper released by Brock, Lakonishok and LeBaron (1992) which showed that by using the most popular trading rules they were able to get significantly better returns over a simple buy-and-hold strategy. This paper did not include transaction costs, but the results were so favorable for the active strategies that it was deemed likely that it would not have changed the bottom line, active strategies could outperform a simple buy-and-hold strategy.

Many of the popular papers released since then are tested using a in-sample testing procedure. By basing your results solely on a in-sample test you run a high risk of data-mining basis. Sullivan, Timmermann and White (1999) argued that the technical trading rules created and tested in an in-sample study performs poorly once applied to an out-of-sample data-set. However, they tested their strategies over a short duration, 1987-1996, and in this short duration the market was in a long boom period which would have negatively effected the efficiency of the out-of-sample test when compared to the passive strategy.

Once we entered into the new decade the financial markets hit quite a rough patch, we recently had two major crashes, the dot-com bobble of 2000 and the crash of 2008. These events have influenced the way many academics think, and have forced them to reevaluate their thoughts on market timing. If investors had followed a market timing strategy as they entered the new millennium, they would have been able to avoid many of the losses they experienced in these crashes. This is reflected in the literature released in the new decade, such as Park and Irwins paper of 2007 which went through numerous other works in order to try and answer once and for all if technical analysis was profitable. They found that even though there were a lot of positive evidence for the profitability of technical trading there were also an issue of the testing

procedures done, mentioning examples such as data snooping, ex post selection of trading rules and the variance in how risk and transaction costs are handled.

One of the most quoted technical trading strategies is the simple moving average (SMA) strategy, which dictates that once a security drops below the average value it has had for the previous "k" period it is sold and one moves into treasury bonds.

Siegel (2002) tested a simple moving average strategy in which he based his average on a 200-day average. He tested this strategy on the Dow Jones Industrial Average (DJIA) from 1886 to 2006 and discovered in his simulation that even when adjusted for transaction costs he was able to increase the risk-adjusted returns, though not the total returns, over a simple buy-and-hold strategy. It is fair to believe that if he had been able to include the returns of 2008 and forward he would have gotten even more favorable results for the timing strategy. It is also worth mentioning if that you exclude transaction costs his strategy would have given a higher absolute return as well as a higher risk-adjusted one. Using a signaling structure where the buy signal is produced if the index is over 1% of the moving average and a sell signal if under 1% he was able to produce returns 4% higher returns than the buy-and-hold strategy while having a 25% lower volatility.

Faber (2007) is one of the most famous of the newer papers that showed the efficiency of the SMA strategy. He went from using daily returns to monthly returns while basing his strategy on a lookback period of 10 months. The purpose of his paper was that he wanted to present a strategy that was as simple as possible, as well as fully mechanical, that would outperform the buy-and-hold strategy. He applied this strategy to multiple equity markets as well as other publicly traded assets such as the commodity market. For the stock markets he applied this strategy from 1900 to 2008, for the others he looked at 1970 to 2008. By simulating how the strategy would have performed if applied in the past Faber was able to not only increase the compounded returns from 9.21% to 10.45% over the period of 1900-2008, but he was also able to reduce the volatility by over 500 basis points and reduce the maximum drawdown from 83.66% to 50.31%. The SMA strategy using a 10 month average has later been confirmed as being efficient in multiple other papers such as Kilgallen (2012) and Gwilym, Clare, Seaton and Thomas (2010).

Gwilym, Clare, Seaton and Thomas (2010) expanded upon the work of others and found that a trend following strategy such as the SMA strategy and a momentum system such as the time-series momentum are not mutually exclusive. The inclusion of a trend-follwing filter to a portfolio of momentum winners managed to reduce volatility without impeding on the returns and as such gave a higher risk-adjusted return. This filter stated that if a portfolio of 6 momentum winners only had 3 markets that are trending higher, or positively, a split is done with 50% into equity and 50% in risk-free. They also concluded that the effectiveness of the momentum strategy used on the international equity markets has decreased over the past two decades, but that the emerging markets have helped to reduce this decrease.

Moskowitz, Ooi and Pedersen (2011) found that they were able to create substantial abnormal returns over multiple different assets types as well as different markets by utilizing what they called a "new asset pricing anomaly". While they might have found what they called a "new asset pricing anomaly" they used old trading rules for their tests. This asset pricing anomaly is what we often refer to as a time-series momentum. They found that there was a strong positive correlation between the securities past returns and the future returns over a time period of 12 months as well as a significant relationship between time-series and cross sectional momentum. These results were consistent over 25 years of data and over nearly 25 dozen future contracts as well as several other major asset classes. Interestingly they also found that their results were consistent for many markets were the investors would be of completely different mindsets and as such presented a challenge for many behavioral theories.

As stated, most of the popular papers on marketing timing are based on in-sample tests, we however will also be testing the efficiency of market timing on an out-of-sample test. An out-of-sample data set is simply put a data set that is not included in the data set in which the optimal trading rules are created. As an example, if a technical trading rule is created based on the time period 1970-1990 and then tested from 1990-2011 or even from 1940-1960 it has been tested on an out-of-sample data set. It is commonly believed that the performance of a strategy created in an in-sample period, and tested on an out-of-sample dataset will provide us with an unbiased estimate of its real-life performance. Though the results we get may not reflect the reality

perfectly, they are in other words a good approximation of what we can expect to get in the future if we were to apply our trading rules today.

In order to be able to do an out-of-sample test while only having historical returns we have divided our data set into two separate subsets, one is defined as the in-sample period and one as the out-of-sample period. The point in time which this partition happens is often arbitrarily chosen and as such is also a topic of discussion. We will explore in this paper the effects of switching this split-off point. Once the split is done, the strategy is used to determine what the best rule in the in-sample period would be, then this rule is used in the out-of-sample period to evaluate the strategy.

3 Data

We will be utilizing two different stock market indices in order to analyze the effectiveness of our active strategy, the Standard and Poor's Composite stock price index and the Dow Jones Industrial Average index. Our primary focus will be on the S&P 500 index, we will be taking advantage of the Dow Jones index as a means to insure that the results we achieve are applicable to more than just one index, ensuring its validity. The data set we will be using for the S&P 500 returns is a combination of two data sets. We have used the returns provided by William G. Schwerts from 1857 to 1925 and combined them with Amit Goyals from 1926 to the end of 2011 providing us with a total of 1860 months.

The Center for Research in Security Prices at the University of Chicago created the first comprehensive database of historical security prices in 1963, and has since been extensively used in empirical research in finance. They started their database in 1926 and as such the quality of the data from that time and forward has been closely scrutinized, and must be considered to be of a very reliable nature. As such the data for the period from 1857 to 1926 is of a lesser quality than the rest of the data set, though this is only natural as the further back in time you go the less reliable the information you get is. It is however still considered to be of high quality, for more information on the construction of this data see *Journal of Business*, July 1990. The reason for

choosing to extend our time period to further back than the CRSP data base is mainly that we wish to expand upon the research others have done by testing our strategy further back in time. We do this in order to provide additional validity to our results and give us more insight into the efficiency of our market timing strategy. The Standards and Poor's composite index is a value-weighted index based on the market capitalization of the 500 largest companies listed either on the New York Stock Exchange or on NASDAQ, chosen by a committee based on multiple factors such as their market size, liquidity and industry grouping. The S&P 500 is considered to be the best indication of the American stock market as a whole, outperforming the Dow Jones index due to its larger size of 500 companies compared to Dow Jones 30.

The Dow Jones Industrial Average index is a price weighted index, but in contrast to the S&P 500 index the 30 companies selected are not only chosen for their size but are meant to be a representation or cross section of US industry. It was first introduced in 1896 by Charles Dow, Wall Street Journal editor, and has since become one of the benchmarks for the indices for the US stock market. A total of 48 changes have happened to its roster, with companies such as Alcoa, Bank of America and Hewlett-Packard leaving in 2013 for Goldman Sachs, Nike and Visa. These changes are made to reflect the changes in the American economy and in the companies themselves. The index values for the total period of 1896,6 to 2011,12 are provided by the S&P Dow Jones Indices LLC while the dividends I've gotten from Barron's.

Some descriptive statistics of the data are reported in Tables 1 and 2.

Table 1 – Descriptive statistics of the S&P 500 data set.

S&P 500	ETB	Return	CAR
Number Values	1860	1860	1860
Min	0.001%	-29.42%	-29.94%
Max	1.62%	42.91%	42.22%
Median	0.33%	0.89%	0.51%
Mean	0.34%	0.83%	0.46%
Standard Error Mean	0.005%	0.11%	0.11%
Confidence Interval mean.0.95	0.01%	0.22%	0.22%
Standard Deviation	0.22%	4.89%	4.89%
Coefficient of Variable	64.08%	586.16%	1062.63%

The ETB column is the Treasury bill returns, while the Returns are the total returns from holding the market portfolio and CAR is the capital appreciation returns, or the returns without dividends.

Table 2 – Descriptive statistics of the S&P 500 data set.

Dow Jones	Return
Number Values	1068
Min	-44.30%
Max	28.66%
Median	0.88%
Mean	0.31%
Standard Error Mean	0.17%
Confidence Interval mean.0.95	0.33%
Standard Deviation	5.49%
Coefficient of Variable	1783.14%

Returns are the total returns from holding the market portfolio

In our analysis we have had to estimate our own risk free rate of return from 1857 to 1885 simply because there exists no record of the Treasury bill returns before 1885. This technique is similar to the one used in the paper by Welch and Goyal (2008) where they performed a regression with commercial paper returns. To test the validity of this regression they tested it towards a different periods risk free rate of return and ended up with a R squared value of 95.7% as well as a correlation between the observed risk free rate and the estimated one of 99.8%.

$$Treasury\ bill\ rate = -0.004 + 0.886 \times Commercial\ Paper\ Rate$$

Some considerations that should be made for the results are the massive increases in the availability of information, it is now easier than ever to get information on stock markets and movements, as well other peoples research into trading strategies. From about 1970 and forward there was a massive influx of technological advances that meant that information was more readably available to the investors. Popular economics papers such as the one from Brock, Lakonishok and LeBaron (1992) started to really popularize the idea of technical trading which had previously been seen as ineffective. We can see that as a general trend that from 1980 and to today the effectiveness a moving average has declined substantially.

4   Methodology

There are two strategies that have been extensively used in other research papers regarding technical trading, it is the simple moving average (SMA) rule and the time series momentum rule (MOM). The common ground for both of these rules is that they rely on either trends or momentum to be effective. They imply that a stock will keep its momentum and as such signaling a clear trend in the market as well. A rising stock will keep rising and a falling stock will keep falling, even for longer than its intrinsic value would predict. Both of these technical trading strategies follow the same basic principle, you keep your investments in risky assets for as long as a buy signal is generated and move into risk free assets as soon as a sell signal is generated.

4.1 Technical Trading Rules

In the in-sample simple moving average test we have chosen to include a look at the trading rules performance if based on a capital appreciation return instead of a total return. Capital appreciation returns do not include the returns of dividends or interest income and as such reflects the "true" market value increase of a security. We've included these returns as we believe that they might provide us with a better signal than total returns as it takes no considerations to a firms dividend policy.

For the MOM strategy the Buy signal is generated once the security rises above the value it had in the previous k-month period. If we let ($P_1, P_2, ..., P_T$) be observations of the security prices over a given period both adjusted and not adjusted for dividend returns, then a formalization of computation of momentum will be:

$$MOM_t(k) = P_t - P_{t\text{-}k}$$

As such the generation of the trading signals in the momentum strategy for month t+1 will be:

BUY:

$$MOM_t(k) > 0$$

SELL:

$$MOM_t(k) \leqslant 0$$

In a moving average strategy a signal is created by looking at the average value of the security over the k-month period. How this average is created depends on what type of moving average strategy you employ as the security prices within the lookback period are weighted differently

11

for different strategies. A buy signal is generated at the end of the month if the security price exceeds the k-month average, or a sell signal is created if the security price is lower than the k-month average. If a sell signal is generated you move into risk-free assets.

The weighted moving average is computed as

$$MA_t(k) = \frac{w_t P_t + w_{t-1} P_{t-1} + \ldots + w_{t-k+1} P_{t-k+1}}{w_t + w_{t-1} + \ldots + w_{t-k+1}} \quad ,$$

where the w_i represents the weights of P_i in the computation of the moving average.

For a general moving average strategy the signals for time t+1 are:

BUY:

$$P_t > MA_t(k)$$

SELL:

$$P_t \leqslant MA_t(k)$$

The most popular of the moving average strategies is the simple moving average, or SMA. In the SMA strategy all of the security prices in the lookback period are weighted equally. As such the SMA at time t is:

$$SMA_t(k) = \frac{1}{k} \sum_{j=0}^{k-1} P_{t-j}$$

The signals generated by the SMA strategy at time t+1:

BUY:

$$P_t > SMA_t(k)$$

SELL:

$$P_t \leqslant SMA_t(k)$$

In the MOM strategy you react on signals based on whether or not the security price right now is higher than it was exactly k-months ago, whereas in the SMA strategy you look at whether or not the security prices right now are higher than the average over the k-month period. There are some similarities between the momentum strategies and the moving average strategies. If the average value has increased over the k-month period it is not unlikely that the final value is higher than the starting value and as such both strategies will often produce a buy and a sell signal at the same time for the same periods.

There are a few other moving average strategies, such as the Linearly Weighted Moving Average (WMA) and the Exponential Moving Average with the difference being that they overweight the most recent prices in the computation of a moving average.

The Linearly Weighted Moving Average (WMA) gives different weights to the previous data that it bases itself on, the data from further back gets weighted less than the more recent data. It weighs the security prices linearly. In our study this means that the stock prices from for example 9 months ago are less influential in what signal gets produced than the stock prices of last month. Formally in an k-month WMA the last month has a weight of [k], the second last of $[k-1]$, and so on going to one:

$$WMA_t = \frac{kP_t + (k-1)P_{t-1} + \ldots + 2P_{t-k+2} + P_{t-k+1}}{k + (k-1) + \ldots + 2 + 1}.$$

In the Exponential Moving Average strategy the data from earlier in the lookback period gets

13

weighted less and less on an exponential scale. The weighting factor gets decreased exponentially the further back in time you go, never reaching zero.

The EMA is computed as

$$EMA_t(k) = \frac{1P_t + \alpha P_{t-1} + \alpha^2 P_{t-2} + ... + \alpha^{k-1} P_{t-k+1}}{1 + \alpha + \alpha^2 + ... + \alpha^{k-1}} \quad ,$$

where the alpha is a coefficient that represents the degree of weighting decrease in the model, between 0 and 1 with a lower value leading to a more severe decrease in the weighting of older observations.

We have chosen to focus on using a moving average rule as we believe that looking at the average values of stocks will provide us with a better signal than simply a start and end value. We have chosen to focus on the SMA over the WMA or the EMA, and the reason for this is simple. The SMA strategy is the most popular strategy of the moving average strategy and as such is the one with the most support behind it. There are several significant papers concerning the SMA strategy which we should be able to compare our results to.

The purpose of the moving average strategy is to filter out the noise in the market and only react to longer trends. It relies on the predictability of the market by saying that a security price will continue to follow a general trend. What one uses as the lookback period for the moving average strategy has a great influence on its effectiveness. A too short lookback period and you risk not being able to find the actual trends in the market and too long of a lookback period and you will get a sluggish signal that reacts to slowly to the market. It is a fine balance to be able to ensure that you do not react to noise in the market while also being able to enter and leave markets at the opportune time. Figure 1 shows an example of noise reduction due to a moving average strategy.

Figure 1. S&P 500 plotted against a 10 month simple moving average on a logarithmic scale

.
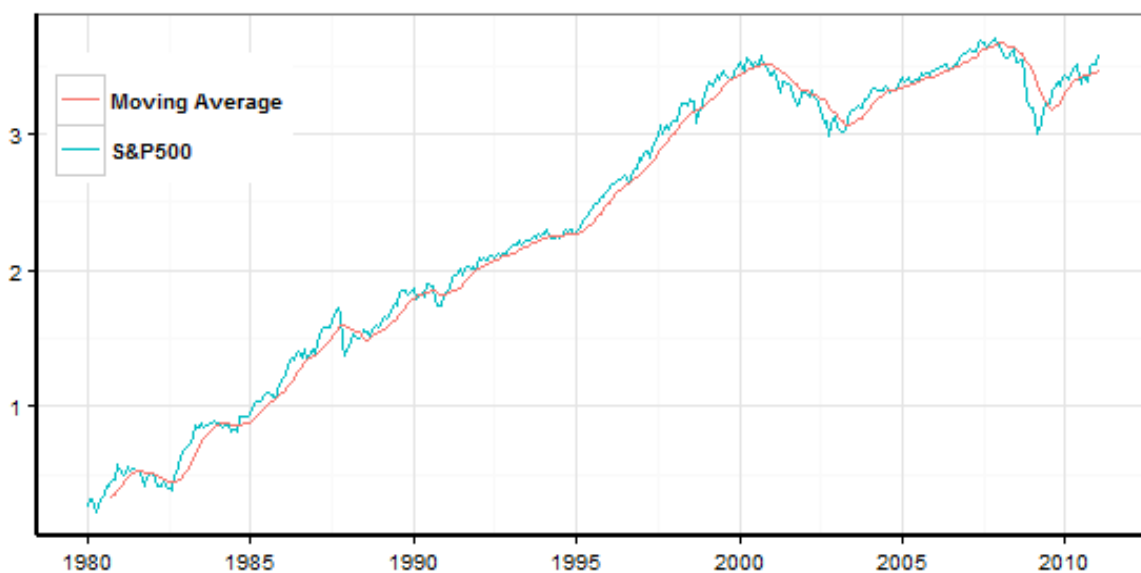
4.2 In-sample Test

In the first part of this paper we will perform in-sample tests of the 10 month SMA strategy from 1857 all the way to 2011 and as such we will extend upon the time period that previous authors have used. We hope that with this we will able able to extend upon the works of others and provide some new insight into the efficiency of a SMA strategy compared to that of a buy-and-hold strategy.

We will try to see if we can replicate the results other authors have gotten testing the 10-month SMA strategy even if we extend the time period, or if the previously reported results on the efficiency of the strategy are exaggerated.

When you collect results from a trading rule by applying this trading rule to the same data set which you used to optimize the trading rule, it is referred to as an in-sample test. You search through your own data, testing multiple different strategies by applying them to you underlying data set before you are able to find what you consider the optimal strategy. Once this optimal

strategy is found you test it on the same data set in order to produce a set of results. With doing this you are able to find out what type of strategy would have been best to use in the past, but the question is how useful is this information when trying to find out what strategy is going to work best in the future. There is no guarantee that the markets will continue to behave like that did in the past, and as such what was optimal before may not be optimal in the future.

4.3 Out-of-sample Test

For the second part of this paper we will be using an out-of-sample performance test, the benefits of this are two fold. Most importantly it severely reduces the amount of data-mining as we now test our best found trading rule on a different data set to that which we used to create it and secondly it will allow us to better try and simulate the real-life choices a trader has to make on which trading rule to use.

A trader will at all times have the information available about the historical performance of different trading rules, and as such should be able to chose which rule that would have been able to efficiently increase his profits in the past based on a given optimization criterion. In order to separate our data set between the in-sample period, the "past" for our trader, and the out-of-sample period, the "future" for our trader, we set a split-off point. The split point between the in-sample and the out-of-sample period is denoted by "t" and as such $1 < t < T$ will denote the entire time line, with T being the total amount of months in the data set. As such the initial time period $[1, t]$ is used to determine which trading rule will increase the optimization criteria in the most efficient way, this reflects the way a trader will always have the previous periods information available. The optimization criterion is given as $O(r_1, r_2, .., r_t)$ and is based on the returns of the market timing strategy for the entire in-sample period.

The SMA rule is based on selling or buying assets once the asset prices fall below or rises over a k month moving average. In the out-of-sample performance test its important to find which k month lookback period to use each month in order to produce the best possible results, which k to used is chosen by finding the best trading rule in the past. This is done by either using a rolling

or expanding window estimation scheme.

In the expanding window scheme the initial in-sample period $[1, t]$ is constantly expanded upon each month by an additional month each month that goes by. This means that the in-sample period constantly expands, creating a longer and longer in-sample period and as such adds to the validity of the results one achieves. The optimal k for any given period is given based on the results from the now extended time period. Formally this new in-sample period given by $[1, t+1]$, which then determines the trading signal for the period $(t+2)$ given by the $SMA(k_{t+1}^{*})$ rule.

For the expanding window scheme the optimal k is given by :

$$\underset{k \in [2,24]}{max} \; O(r_1, r_{2,}.., r_t)$$

It seeks to maximize the optimization criterion by choosing the length of look back period that would produce the best returns, within the boundaries of $[2,24]$ months.

This is repeated until we end up calculating the optimal k for the final T period of the time window, and as such the signal for the last month can be calculated. Then the returns for the entire original out-of-sample time period, $(r_{t+1}, r_{t+2}, \ldots, r_T)$ is measured and the market timing strategy is evaluated based on chosen optimization criterion.

For the rolling window scheme the original in-sample period of $[1, t]$ does not get expanded as it does for the expanding window. It however, is moved by one increment for each month that goes by. The optimal k to use for each new month is determined by looking at the last n observations and determining which k would have been optimal during those n months, with the n representing the original gap between the first month and the last month in the in-sample period, $[1, t]$. The signal to either buy or sell is then produced. This procedure is then repeated as with the expanding window, until the signal for the final month of the time series, T, is calculated. Maximizing the optimization criterion is then defined as:

$$\max_{k \in [2,24]} O\left(r_{\text{s-n+1}}, r_{\text{s-n+2}}, \dots, r_s\right)$$

The rolling-window analysis is often used when you suspect that the parameters of the model are changing, and as such is used in order to assess the stability of the model over time. As the rolling window has a shorter time frame, and a time frame that is constantly changing for its back-testing period (in-sample period) it is used when you have reason to believe that the parameters of the previous periods have changed compared to the current period. We will primarily be focusing on using the expanding-window scheme and only supplement this with the rolling-window scheme in order to ensure the stability of our parameters as well as the validity of our results as we have no reason to believe that the parameters are instable.

This should be able to eliminate most of the data mining bias one would experience using a in-sample SMA rule, as the choice of look back period is constantly tested and changed and is not influenced directly by the previous works of other authors. It should also allow us to find a out if our strategy would be able to outperform the simple buy-and-hold strategy if applied to new data as it comes in, and not just if it would work in a historical setting.

There are some concerns considering the out-of-sample performance method as well, the two major ones being :

- It is based on a single underlying data set
- The split-off point between in and out-of-sample is chosen arbitrarily

In order to try and combat these problems we will alter the different starting and ending points of our data set. By exploring the results we get by setting the start of the data set from 1857 to for example 1926, and still holding the end point at 2011, we will be able to shed some light on the effects of the strategy if we had had a different data set. Primarily we will divide the data set into an early and late period, 1857-1926, 1926-2011, as well as changing the split-off points between in- and out-of-sample for these periods

The second, and maybe most important factor, when it comes to the quality of the results we get is the split-off point. We will test the effects of changing the split-off points and as such examine how much of a factor this plays in the results we get, it might be that the split-off point it self is the deciding factor between whether the passive or active strategy is superior.

4.4 Optimization Criteria

There are a lot of different optimization criteria within the world of finance that would be suitable for our study, there is however one that sticks out when to comes to technical trading analysis. Most papers, including Faber (2007) and Park and Irvin (2007), lay a heavy focus on the Sharpe ratio when it comes to evaluating the performance of the different trading rules.

The Sharpe ratio is a reward-to-risk performance measure that calculates the risk adjusted returns of a portfolio by subtracting the risk free returns, such as the Treasury bill returns, from the risky returns of the portfolio and dividing the results on the standard deviation of the portfolio returns. The greater the the Sharpe ratio the greater the risk adjusted returns are. Sharpe ratio seems like a natural fit in our situation, were we have two mutually exclusive strategies in that our active and passive strategies are not combinable. You can not keep all our money in a buy-and-hold strategy as well employing an active strategy that switches out between risk-free and risky assets.

One of the major critiques of the Sharpe ratio is that while it does take risk into account it takes both the upper and lower ends of the standard deviation spectrum into its calculations. Few investors would consider the "risk" of their stocks increasing higher than expected to be a negative thing. The Sharpe ratio takes both the upper and lower price movements into account when measuring the performance of the portfolio. There are a few different ratios such as the Sortino ratio which does not take the upper price movements into account, but only the downside risks, or the Calmar ratio where only the maximum drawdowns of the portfolio are considered. Studies done by Eling and Schumacher (2007) and Eling (2008) show that there is a strong positive correlation between all these performance measures, meaning that they all move in the same patterns. The recently released paper by Valeriy Zakamulin (2013) showed that when all of

these, and several other, optimization criteria were used in the trading rules the comparative differences between the passive and active market timing strategy were the same. All these papers have concluded that the choice of performance measure does virtually not have an impact on the comparative differences between the passive and the active strategy. We have chosen to use the Sharpe ratio for this paper, as it is the most commonly used and easily comparable performance measure to use.

A formalization of the optimization criterion, with SR being Sharpe ratio, will then be:

$$\max_{k \in [2,24]} SR(r_1, r_2, .., r_t)$$

4.5 Other Performance Measures

In the active strategy the Sharpe ratio will be used to determine which lookback period would have been optimal for each time period, but it is not the only performance measure we will use in this paper. The most important thing at the end of the day for an investor is maximizing the utility of his initial investment, to get the most out of the money they have invested, because of this we will also include a final wealth measure which will tell us how the growth of wealth has been. In order to achieve this there are two problems an investor has to consider, the first being which risky portfolio it is optimal to invest in, and secondly how to allocate his wealth between the risky and the risk-free portfolio. The first issue is solved by evaluating the portfolios based on their Sharpe ratio, which is easy enough.

The second issue however is rather unsolvable as it is reliant on each individual investor. What the optimal allocation of wealth between risky and risk-free is completely reliant on the amount of risk the investor is willing to take on, a risk averse individual will maximize his utility by investing mostly into the risk-free while a risk seeking investor will maximize it by going mostly into the risky assets. Add to this the existence of market imperfections, such as restrictions on short selling and borrowing, the issue becomes quite unsolvable. We hope that by including the final wealth of each strategy that we might be able to give some insight to how investors might

expect the growth of wealth will unfold, even though it takes no considerations to risk preferences directly.

4.6 Transaction Costs and Taxes

Transaction costs are generally viewed as the costs incurred due to either selling or buying securities, but they may also include other non-monetary values such as time investments and effort. For our paper we have chosen not to include the transaction costs as their implementation would severely complicate the methodology required to write this paper. Trying to set a real value to things which have no real monetary value is often difficult to do and is highly subjective, this as well as the differing real transaction costs based on which brokers you use or what the spread is at any given time means that including transaction costs would be hard. For the same reasons as they we have mentioned above will we not include taxes in this paper, simply due to the practical limitations they imply. We will however temper our results with transaction costs and taxes in mind, and as such hopefully be able to achieve some meaningful results.

4.7 Capital Appreciation Return

The Miller-Modigliani theorem states that a firms capital structure, as well as its dividend policy, has no effect on a firms true value, however this theorem is based on the assumption of not only a tax free society but also an efficient market. By removing dividend payouts from the return calculations we hope that we should be able to generate a signal that more efficiently predicts the major market movements as we will be able to avoid the noise generated by different dividend policies. Dividend policies vary a lot and can delude our results as for example some firms keep issuing dividend payouts when their true-value is dropping in order to keep up appearances for their stock holders. This means that we get an inflated stock price, and by excluding dividends and other such factors we hope that our model will be able to better predict market movements and as such provide us with a better signal for out trading rule. By excluding the dividends we are only looking at the speculative part of the returns.

5 Empirical results

The empirical results will be split into two major parts. For the first part of the empirical results we will be testing the simple moving average strategy with a 10 month lookback period using an in-sample test. This strategy we will base on both total returns as well as capital appreciation returns. In the second part of the empirical results we will take this type of SMA strategy one step further and use it in an out-of-sample as well as in-sample data set. This we will do by utilizing both a rolling and expandingbwindow scheme. We will not be look at the capital appreciation returns for the out-of-sample test.

For both parts we will apply the strategy in three different main time periods:

- A look at the total efficiency of the strategy from 1857 to 2011
- Dividing the period into two major time periods, 1857-1926 and 1926-2011

5.1 In-Sample Test

The discussion of the empirical results are mainly a comparison between the timing strategy and the buy-and-hold strategy, times when the timing strategy based on the capital appreciation returns is used will be made clear.

5.1.1 S&P 500 1857-2011

We find it natural to first apply the strategy to the entirety of the data set, before breaking it into smaller parts and analyzing the individual pieces. If we look at the period from 1857 all the way to 2011 we can see that if you had invested a single dollar in stocks using a single buy-and-hold strategy you would be worse off in 2011 than if you had followed the active strategy. Figure 2

shows the growth of wealth from 1857 to 2011 on a logarithmic scale.

Figure 2 - Logarithmic wealth 1857-2011
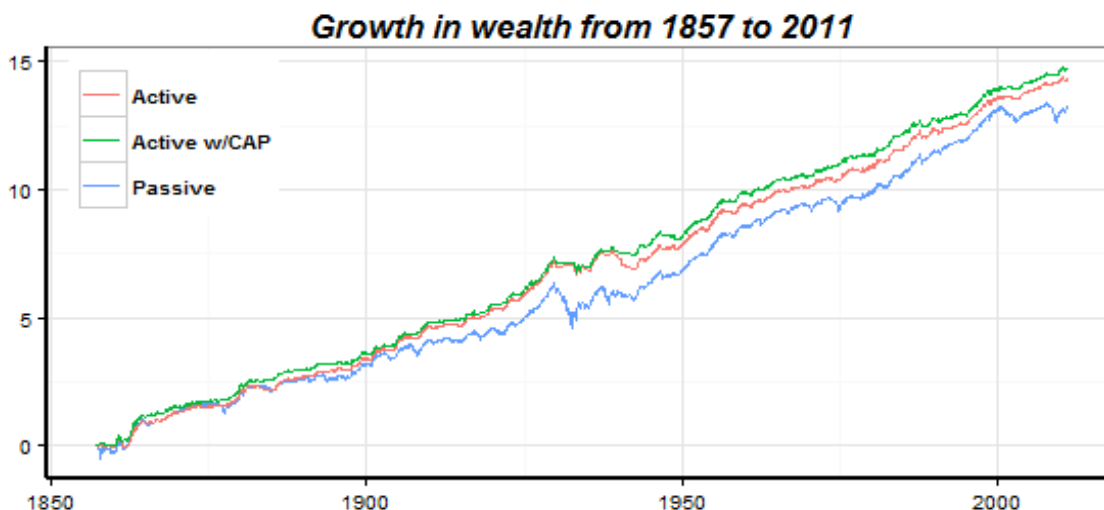
The Active line is the active 10-month SMA strategy based on total returns, while the Active w/CAP is the active strategy with capital appreciation returns.

Looking at Table 3 we can see that the active portfolio, and especially the one based on the capital appreciation returns, outperforms the buy-and-hold strategy. It has a increase in Sharpe ratio of about 45%, a decrease in volatility by 46% and a decrease in drawdown from 83% to 50%. We see that the timing portfolio outperforms the simple buy-and-hold strategy on all the criteria, and the active portfolio based on the capital appreciation returns performs even stronger for this time period. The active portfolio based on total returns experiences an increase over the passive strategy of about 45% for the Sharpe ratio, a 200% increase in final wealth as well as a decrease in the maximum drawdown from 83% to about 50%.

Table 3 – Key figures 1857-2011

	Passive	Active	Active w/CAP
Volatility	16.94%	11.59%	10.77%
Sharpe	0.354	0.514	0.569
Max Drawdown	83.10%	50.48%	46.76%
Average Return	10.54%	10.50%	10.70%
Compound Return	8.99%	9.77%	10.07%
Final Wealth	569402	1713378	2601284
Relative Increase over Passive		Active	Active w/CAP
Sharpe Ratio	-	45.31%	60.99%
Drawdown	-	-39.26%	-43.73%
Wealth	-	200.91%	356.84%
Volatility	-	-31.57%	-46.65%

Average Return are arithmetic returns while the
Compound Return are geometric returns.

The decrease in volatility has a lot to say for the final wealth of the portfolio as a sharp decrease in value takes a long time to recover from. If the stock drops by 20% one day and increase by 20% the next day it will still net you a rather large loss. As Faber (2007) so nicely put it, "The unfortunate mathematics of a 75% decline require an investor to realize a 300% gain just to get back to even – the equivalent of compounding at 10% for 15 years.". This can clearly be seen in the graph from Figure 2, the buy-and-hold strategy falls behind because as soon as it experiences the rough drawdowns of the 1930s it can not recover compared to the active strategy. The active strategy largely avoids these drawdowns and because of this ends up with a larger total wealth at the end of the period, the buy-and-hold strategy is not able to come back after such a massive drawdown.

Even though the buy-and-hold strategy gives a higher average annual return, the final wealth of the active portfolio is 200% higher than that of the buy-and-hold strategy. This is due to the fact that the annual average return is based on arithmetic returns and not geometric ones. For the annual returns, they are simply added together before dividing them on the number of years. As an investor this is not as interesting, as an arithmetic return of 90%, 50%, 30% and -90% is still averaged out to 20%, while an investor would actually be left with a return of nearly negative 22%. As such the compounded returns are far more interesting, as they are calculated using

geometric returns. Geometric returns is the average of a set of products, showing us the actual % change in wealth from point A to point Z. We can see that the compounded returns for the timing strategy are higher than that of the buy-and-hold strategy. The compounded returns give us a better view of the effectiveness of the strategy, but they do not account for the level of risk you get in each strategy. This is why the risk-adjusted returns, or the Sharpe ratio, is the most important of our key figures.

With an increase in Sharpe ratio of 45% it is clear here that the timing portfolio out-performs the simple buy-and-hold strategy over a longer period of time, and even more so the timing strategy using the capital appreciation return.

We have tested the active strategy on a time span of 154 years, but most investors do have such an extensive time horizon. We shall now test the strategy on a shorter time period in order to see how it would have performed in the past as well as how it would have performed with a shorter time span.

5.1.2   S&P 500 1857-1925

During this time period the world experienced quite a few major changes, we went back to the gold standard and experienced the first world war. In 1914 wall street was shut down for a period of three months, so the data from that period has been set to zero, however this is such a small part of the data set that it should not have any significant effect on the results.

Figure 3 - Logarithmic wealth increase 1857-1925
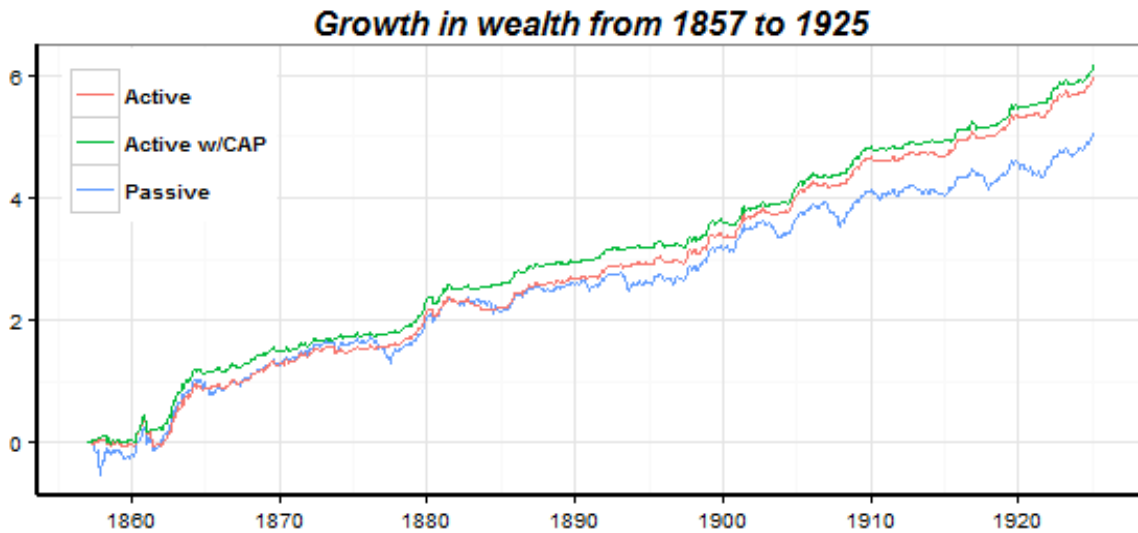
From Figure 3 we can see that the even during shorter time periods, though to be fair it is still a fairly long time span, the active portfolio still out-performs the buy-and-hold strategy when it comes to wealth gain. The buy-and-hold strategy is fairly even with the active strategy for the earlier years, from about 1863 and to the bear market in the late 1870's they produce about the same wealth, but as soon as the market has a short down period the active strategy sells out into risk free and avoids the major downsides. This ends up giving the active strategy a much higher wealth at the end of the period due to the cumulative effects of gains.

Table 4 – Key figures 1857-1925

	Passive	Active	Active w/CAP
Volatility	13.66%	10.04%	9.11%
Sharpe Ratio	0.278	0.470	0.539
Max Drawdown	43.01%	36.92%	22.26%
Average Return	8.71%	9.72%	9.92%
Compound Return	7.71%	9.18%	9.48%
Final Wealth	156	393	474
Relative Increase over Passive		Active	Active CAP
Sharpe Ratio	-	68.97%	93.58%
Drawdown	-	-14.16%	-48.25%
Wealth	-	151.61%	203.60%
Volatility	-	-26.52%	-53.05%

We can see that Table 4 still shows us the same picture as before, the volatility has decreased by 26.5%, the Sharpe ratio has increased by nearly 70% and the maximum drawdown you will suffer is about 37% versus the previous of 43%. There is now also an even clearer advantage to basing our model on the capital appreciation returns over the total returns, as it gives better results on all of our key figures. The changes in wealth gain are still quite substantial, an increase of respectably 150% and 203% shows us a truly great increase in total returns. The active strategy would have once again been able to increase both total as well as risk adjusted returns.

Both the average annual return as well as the compound returns are higher for the timing strategy, once again a clear indication of its effectiveness.

5.1.3   S&P 500 1926-2011

Let us now investigate the effects of applying the timing strategy to the period of 1926 to 2011.

Figure 4 shows once again that one of the major benefits of our timing strategy is that while it its slightly slower at re-entering the market it is quite quick at exiting it. During the major downward trend of the 1930's we have left the market and gone into risk free assets, thus avoid the worst of the drawdowns. This can also be seen from the dramatic decrease in maximum drawdown from a staggering 83% to a still high, but significantly more manageable, 50%. As a result of this we can see the wealth of the timing portfolio is quite stable during the 30's while the buy-and-hold goes far into the negatives.

During the 2nd world war we enter into a long bull period of increasing returns from the market and we can see that by about 1955 the timing portfolio and the buy-and-hold one are equal in wealth. This seems to imply that during long bull markets our portfolio can not fully generate the same level of returns as the buy-and-hold strategy.

Figure 4 – Logarithmic Wealth 1926-2011
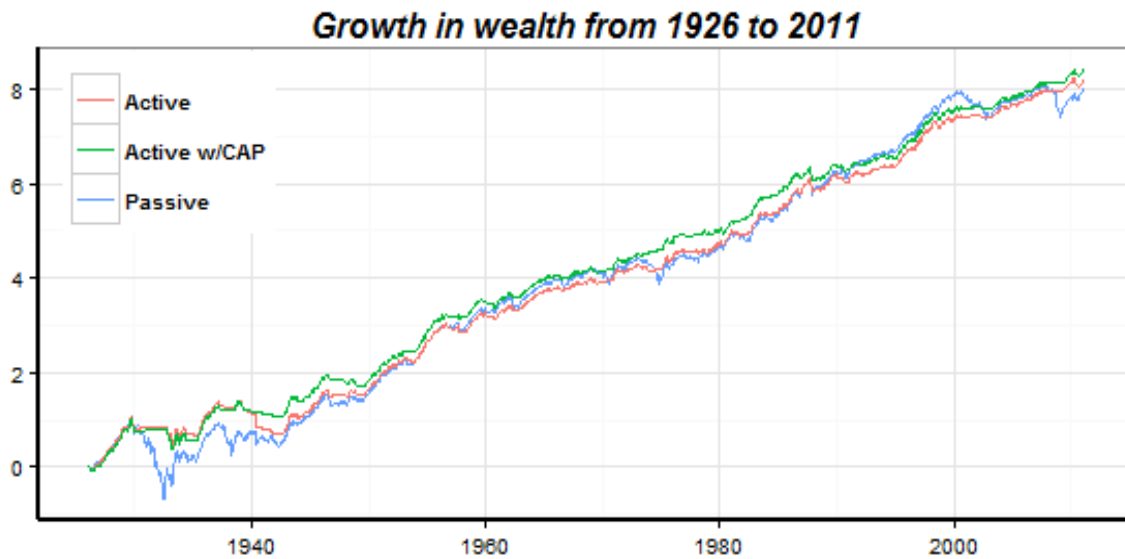
Looking at Table 5 we can see that active portfolio still provides us with a far less volatile investment strategy and as such a much better return for the risk you are taking. This is also evident by the Sharpe ratio increasing by 35%.

Table 5 – Key figures 1926-2011

	Passive	Active	Active w/CAP
Volatility	19.22%	12.73%	11.96%
Sharpe Ratio	0.395	0.533	0.581
Max Drawdown	83.10%	50.48%	46.76%
Average Return	11.86%	10.98%	11.17%
Compound Return	9.86%	10.09%	10.39%
Final Wealth	2953	3531	4443
Relative Increase over Passive		Active	Active CAP
Sharpe Ratio	-	35.09%	47.34%
Drawdown	-	-39.26%	-43.73%
Wealth	-	19.60%	50.48%
Volatility	-	-33.76%	-43.00%

For the 1926-2011 period the basic timing strategy barely produces a higher compounded return, had it not been for the dot-com bubble of 2000 and the financial crisis of 2008 the final wealth would have been higher with the buy-and-hold. Compounded returns is not the only factor in deciding what approach is the best, as the goal is to create the best return for the risk you are taking as such the risk-adjusted returns are more interesting. It is clear that once again the timing strategy outperforms the buy-and-hold strategy when looking at the significant increase in Sharpe ratio.

5.1.4 Comparing the two periods

When comparing the results from the two different periods there are a few important factors to take into account. The time periods are of a different length, the earlier period did not suffer nearly the same amount of financial distress that the later one did and the actual data is more reliable as well as more available.

The difference in time span is 17 years. Shorter time spans may give us to few observations to be statistically viable, but with the amount of observations we still have we do not feel that this is a realistic concern.

What we do believe can affect the results slightly is the fact that from 1857 to 1885 we have had to create an estimation of the risk free returns using the commercial paper rates, and even though it has been proven to be a quite good estimation it is still something to consider.

We can see that for both the periods the active strategy based on the capital appreciation returns performs the best, having a lower volatility, higher Sharpe, better returns and a greater reduction in drawdowns.

Table 6 – Comparison of key figures

	Volatility	Sharpe Ratio	Max Drawdown	Average Return	Compound Return
1857-1926					
Passive	13.66%	0.278	43.01%	8.71%	7.71%
Active	10.04%	0.470	36.92%	9.72%	9.18%
Active w/CAP	9.11%	0.539	22.26%	9.92%	9.48%
1926-2011					
Passive	19.22%	0.395	83.10%	11.86%	9.86%
Active	12.73%	0.533	50.48%	10.98%	10.09%
Active w/CAP	11.96%	0.581	46.76%	11.17%	10.39%

The volatility in the first period is lower for all the strategies compared to that of the second period, but mostly for the passive one. This indicates that when there are no major recessions in the market the active strategy does not give as large reductions in risk. This makes sense as the strategy will only trigger a sell signal if we have experienced a longer downward trend, and when there are few such periods it will follow the buy-and-hold strategy closer. Volatility in the first period is reduced by 26.5% while it is reduced by nearly 34% in the second period.

Using the moving average strategy provides us with a better increase in Sharpe ratios for the 1857-1925 period, there we experienced an increase of almost 69% while in the 1926-2011 we got an increase of about 35%. We can see here as well that the timing strategy performs the best when the market is volatile, as its major benefit is avoid the major downward trends in the market while still getting back into the market as soon as it starts trending upwards again.

We see that the drawdown for period one is only 43% versus the drawdown of the second period of 83%, this supports our earlier statements of the second period being a more volatile period as well as obviously having longer falls in the economy. In the second period the averaged annual returns for the passive strategy are higher than for the active portfolio, while it is the opposite for the first period. This shows us that while our model does best during volatile periods, or times when the market suddenly falls quickly, it starts to lag behind once we enter long periods of growth. During the first period we had a stable upwards trending market, with a few smaller down periods, in which the timing portfolio strives In the second period the timing portfolio did quite well during all of the major crashes, and as such was able to get a head in wealth, but

during the long bull periods it had a lower return leading to a lower averaged return for the period as a whole.

5.2   Out-of-sample Tests

For the out-of-sample tests we will not be using capital appreciation returns at all, and only focusing on the total returns. The out-of-sample performance test should be able to better replicate the results a trader would be able to get if he started using a technical trading strategy compared to the in-sample test. We will still be using a moving average strategy, but now we will divide our time period into two different periods. One will be classified as an in-sample period and one as an out-of-sample period. In the in-sample period we will find which trading rule performs the best based on a optimization criterion then we will apply this trading rule to the out-of-sample sample period. This should allow us to better reflect the choices a real trader would have if we were to examine the possibilities he had today, he would be able to look back at previous results and determine which strategy he would like to use for the future.

We will be using a expanding window and rolling window scheme for our in-sample period. In the expanding window scheme the in-sample period is extended by one month each time we have calculated the optimal look back period, while in the rolling window scheme the in-sample period length is constant, but it is shifted one month forward for each completion. The split-off point will be originally set at a fixed length from the starting point, but will be changed in order to explore the effects of doing so.

5.2.1   S&P 500 1857-2011

Once again we will start by examining the timing model using our entire data set, for these initial results the split-off point has been set to 1881. This split-off point was chosen rather arbitrarily, this is done in order to try and explore a often cited weakness of the out-of-sample test.

The out-of-sample performance model is often criticized due to the large impact many people

believe the choice of split-off point has on the results of the strategy. We have chosen to set our initial split-off points based on absolutely no concrete facts, we just picked a year and stuck to it. For the remainder of the paper the initial split-off point will be 24 years into the future from the starting point, unless otherwise stated.

Figure 5 - Logarithmic Wealth 1881-2011 Expanding Window
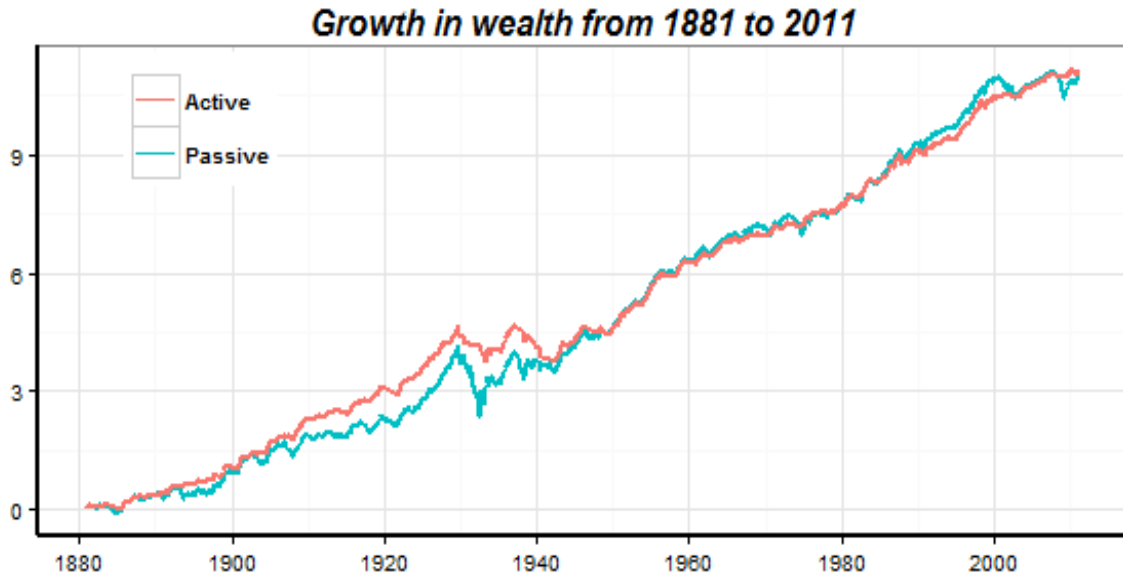
We can get some indications from Figure 5 about how our model performs compared to the standard buy-and-hold strategy. One of the clearest advantages it has it that it exits the market before the major crashes happens. We see that especially well in the 1930s, as when the stock markets crashed our timing strategy would have sold out and gone into risk-free thus avoiding the worst of the fall. The market experienced a drawdown of 83% while the active strategy got away with a loss of only 60%. While this still is a rather large decrease in wealth it is still a significant improvement over the passive strategy.

With this decrease in losses the active strategy is able to stay a head for a while, but once the markets start to sky rocket it shows its weakness. Due to sluggishness of the active strategy it is slower at reentering the market once its starts to climb back up again and already by about 1945 the two strategies are even again. From this point and forward the two strategies perform equally

well, with few differences between them.

Once we enter into the strong bull market of the 1990s we can see the passive strategy increasing over that of the active one, this implies that even with this augmentation to the standard market timing strategy used in the first part of the paper, the active strategy does not perform as well as the passive one during strong bull periods. At the same time we can see that as soon as the strong bear market of the early and late 2000s hits the active strategy strongly out performs the passive strategy.

Figure 5 implies a few things that we previously discovered to be true about the market timing strategies compared to the passive strategy when it comes to wealth gain.

- The higher the volatility in the market, the better it performs
- It outperforms the passive strategy during bear periods while staying relatively even during weak bull periods.
- During strong bull periods it loses out to the passive strategy due to the signaling delay.

Let us take a look at a few of the other performance criteria as the increase of wealth is not the only factor that should be taken into consideration.

Table 6 – Key figures for 1881-2011 Expanding Window

	Passive	Active
Volatility	16.96%	11.59%
Sharpe	0.365	0.477
Max Drawdown	83.10%	60.14%
Avg Yrl Return	10.40%	9.68%
Compound Ret	8.85%	8.94%
Final Wealth	61667	68188
Relative Increase over Passive		Active
Sharpe	-	30.71%
Drawdown	-	-27.63%
Wealth	-	10.57%
Volatility	-	-31.65%

Index	S&P500
Type	Expanding
Start	1857
In-sample	1881
End	2011
Mean k	6.63

It is quite surprising that the average look back period for the strategy is 6.63 months as the most commonly used k for moving average strategies is 10 months. (Brock et al.(1992),Siegel (2002) and Faber (2007)). This means that while setting the lookback period to a static 10 months may provide the best results for a simple in-sample test over the past century, it is far from the optimal average lookback period if you use an expanding-window scheme. While it might be expected that an average k may not perfectly reflect the optimal k for a static SMA strategy, it is unexpected that the difference is this noticeable

The reason for this might be that for the earlier periods, 1857-1925, the average k is very different from the rest of the data set and as such would not have been considered in the other studies who did not go as far back as 1857.

We see that the passive strategy has a higher volatility than the active strategy and that the active strategy manages to reduce this volatility by over 31%. This is a massive reduction in volatility and an attractive trait for a strategy. You can argue that this is to be expected as one of the main functions of the active strategy is to reduce volatility by working as a noise reduction for the market. Meaning that it provides a less volatile results at the cost of using more time to decide on a path of action. While the passive strategy experiences all the ups and downs of the market the active strategy only acts if there has been an increase or decrease in the market over a longer period. This leads to a decrease in volatility as the active strategy receives the rather stable risk-free rate of return while the passive gets the fluctuation of the market.

Our timing strategy provides a less volatile market strategy than the buy-and-hold strategy, and this leads to it ending up with a higher growth of wealth over the 154 year period. This is despite it having a lower average annual return than the passive strategy. The decrease in volatility leads to a higher growth of wealth than the decrease in annual returns takes away from the strategy. A reduction in volatility means a lot for the development of wealth as a 1% decrease requires more than a 1% increase to recover from.

Final wealth at the end of the time period has increased by 10.5% while still providing us with a

lower volatility. This is reflected in the difference of the Sharpe ratios as well, with the active Sharpe ratio being 30% higher than that of the passive strategy. We have been able to increase our risk adjusted returns as well as our actual returns, as the final wealth has also increased by 10.5%.

Figure 6 - Logarithmic Wealth 1881-2011 Rolling Window
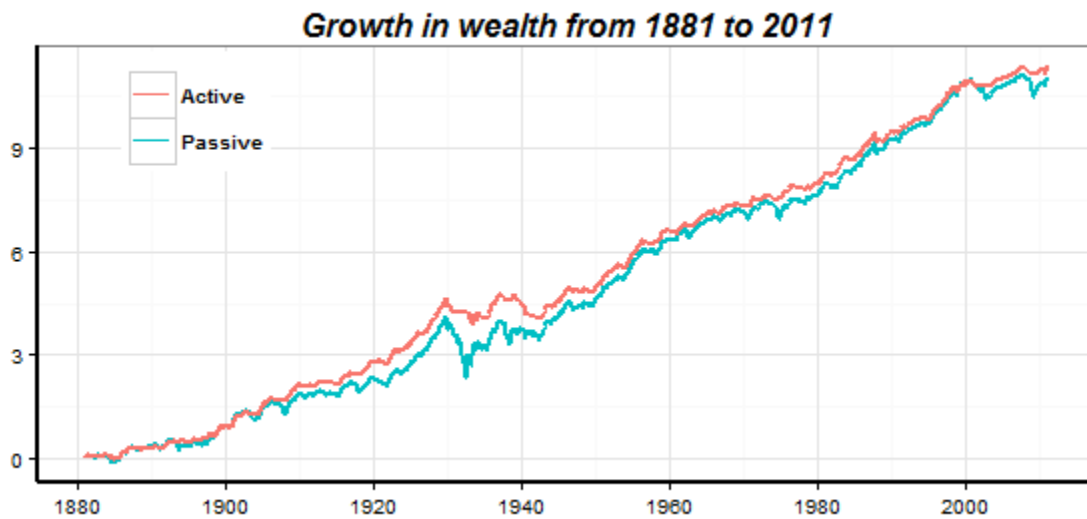
Taking a quick look at the results of the rolling window estimation show that in this case it has out performed the expanding window scheme.

Table 7 – Key figures for 1881-2011 Rolling Window

	Passive	Active
Volatility	16.96%	11.70%
Sharpe	0.365	0.489
Max Drawdown	83.10%	53.93%
Avg Yrl Return	10.40%	9.88%
Compound Ret	8.85%	9.13%
Final Wealth	61667	86109
Relative Increase over Passive		Active
Sharpe	-	33.85%
Drawdown	-	-35.10%
Wealth	-	39.64%
Volatility	-	-31.02%

Index	S&P500
Type	Rolling
Start	1857
In-sample	1881
End	2011
Mean k	9.33

It has a minuscule increase in volatility over the expanding windows volatility, but that seems to have been balanced out by a reduction in maximum drawdown. The decrease in maximum drawdown it has compared to the expanding window, which experienced a 60% drawdown compared to the rolling windows 53%, means that it were better able to handle the great depression in 1930, and as such got a competitive advantage. This advantage meant that over a long period of time it was able to continuously increase its wealth at a comparable return to that of the expanding window. Even though the rolling window experienced a lower average annual return the advantage it got by timing the crash in th 30s better than the expanding window gave it the head start it needed produce a better result.

The final wealth increase compared to the passive strategy for the rolling window was a staggering 39.6%, a major improvement to the expanding windows 10%. This is despise the Sharpe ratios and volatility being close to equal, the improvement can be attributed solely to the decrease in drawdown. It is important to emphasize though that since the Sharpe ratios are so even, the risk-adjusted returns for the expanding and rolling window are about the same.

While the results of the out-of-sample technical trading rules are by no means poor, we see immediately that they are not quite as good as the results we got when we used an in-sample model. This, however, comes as no surprise.

In the in-sample model we were able to look back at the entire period as one, and while having perfect information, choosing the optimal lookback period to use in our active strategy is quite simple. Now we try to more accurately reflect how these trading rules would have worked out for an investor if he had employed them at the split-off point, without having perfect information for the next 100 years. The fact that our out-of-sample performance test was able to produce comparable results at all speaks well to the results we are getting.

More on the split-off point, let us see what kind of effect pushing the original split-off point another 19 years into the "future" and move it from 1881 to 1900.

Table 8 – Key figures for 1900-2011 Expanding Window

	Passive	Active
Volatility	17.75%	12.12%
Sharpe	0.394	0.509
Max Drawdown	83.10%	60.14%
Avg Yrl Return	11.24%	10.33%
Compound Ret	9.53%	9.52%
Final Wealth	24554	24184
Relative Increase over Passive		Active
Sharpe	-	29.24%
Drawdown	-	-27.63%
Wealth	-	-1.51%
Volatility	-	-31.75%

Index	S&P500
Type	Expanding
Start	1857
In-sample	1900
End	2011
Mean k	7.25

We see from Table 8 that while many of the results reflect the same relationship between the passive and active strategy, the active strategy does now not produce a higher compounded return nor as a result, a higher final wealth than the passive strategy. The arbitrarily chosen split-off point has in other words played a role in which of the strategies provide us with the greatest wealth gain over the period. Moving the split-off point has changed which of the two strategies had the highest wealth gain. The passive buy-and-hold strategy now produces a very slightly higher final wealth then the active strategy.

It is worth noting however that the timing strategy still provides a sharp reduction in both volatility and maximum drawdown, and as such still has a significantly higher Sharpe ratio and a better risk-adjusted return.

If we now move the start of our in-sample period to 1890, and set our entire initial in-sample period to the ten years from 1890-1900, we see that once again the active strategy out performs the passive one by quite the margin. The key figures for the passive strategy remains the same, but the 10 years from 1890-1900 provides us with a initial signal that more efficiently predicts the market. Much of the increase in wealth this provides us with can be explained by the fact that it produces a more efficient signal for the stock market crash in 1930 and because of this the

active strategy only experiences a maximum drawdown of 50.48% compared to the 60% it had with the previous in-sample period.

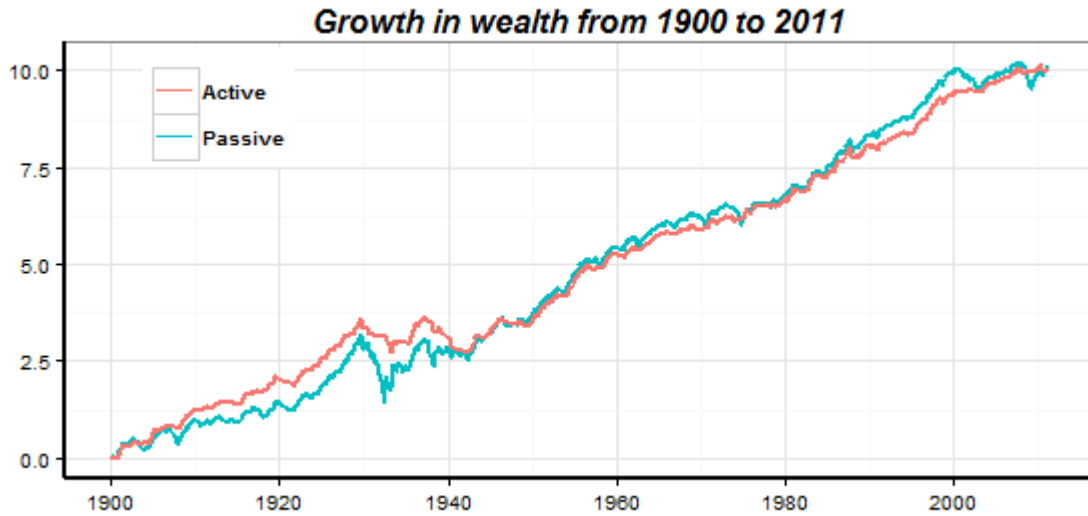*Figure 7 - Logarithmic Wealth 1900-2011 Expanding Window*

Growth in wealth from 1900 to 2011

Table 9 – Key figures for 1900-2011 with the in-sample period beginning in 1890 Expanding Window

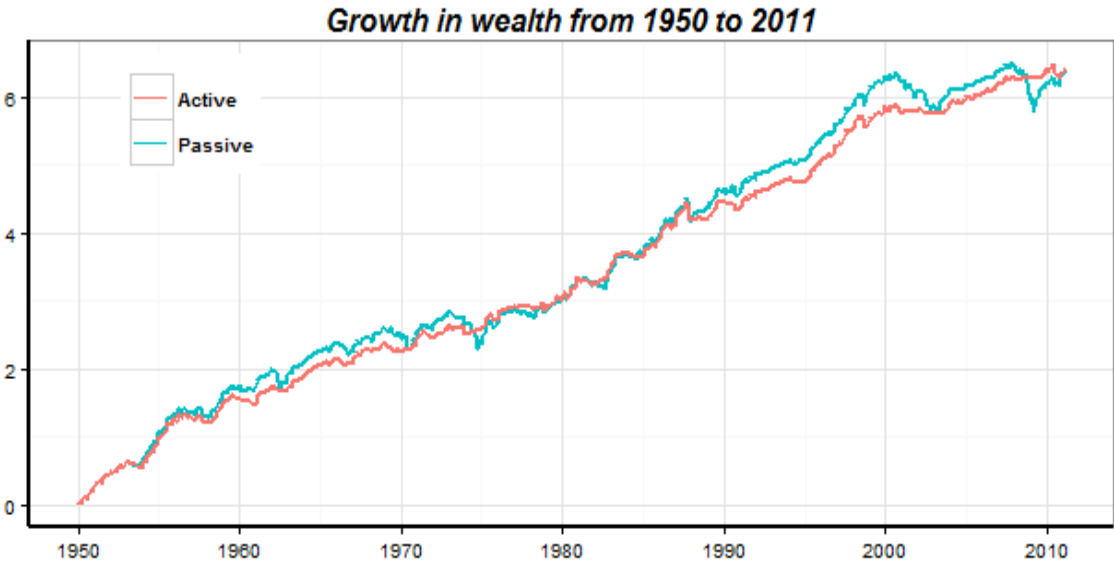|	Passive	Active
Volatility	17.75%	11.86%
Sharpe	0.394	0.544
Max Drawdown	83.10%	50.48%
Avg Yrl Return	11.24%	10.65%
Compound Ret	8.85%	9.88%
Final Wealth	24554	34859

 Table 9 shows us how much the results we get vary by the initial in-sample period, by changing the split-off point and the starting point we have been able to create quite different results which have lead to both the passive and active strategy coming up a head. We will continue to explore this as we separate our data set more and look at different lengths of time.

5.2.2   S&P 500 1926-2011

Keeping the interval the same as before we have set the split-off point between in-sample and out-of-sample at 1950 while starting the in-sample period in 1926, because of this the effects of the great depression will not show directly in our results but rather indirectly through the optimal trading rule chosen in the initial in-sample period.

Examining the graphical growth of wealth for the period, we see a lot of the same signs as before. During strong bull periods the passive portfolio has a stronger wealth gain than the active one, whereas the active one gains its advantages during sudden shocks in the markets. We see that for a long period of time, from about 1950 to 1975, the passive strategy is a head of the active one. It is not until the bull markets of the late 70s that the active strategy gets a head of the passive one. After that its a steady incline for the both of them, with the passive one getting a strong lead in the heavy bull markets of the 90s before crashing down again at the new century. Had it not been for the fall of the markets in the 2000s the passive strategy would have produced a far higher increase in wealth than the active strategy.

Figure 8– Logarithmic growth in wealth for 1950-2011 Expanding Window

Let us now look at some of the key figures for this period by checking Table 10, and see if find the same patterns as before:

Table 10 – Key figures for 1950-2011 Expanding Window

	Passive	Active
Volatility	14.61%	11.41%
Sharpe	0.470	0.568
Max Drawdown	50.96%	23.26%
Avg Yrl Return	12.22%	11.81%
Compound Ret	9.59%	9.74%
Final Wealth	598	615
Relative Increase over Passive		Active
Sharpe	-	20.96%
Drawdown	-	-54.35%
Wealth	-	2.91%
Volatility	-	-21.92%

Index	S&P500
Type	Expanding
Start	1926
In-sample	1950
End	2011
Mean k	9.68

We have had a big reduction in the volatility, the differences between the two strategies at almost 22% with the drawdown also decreasing by 54%. These are large differences that should have been able to tip the scale in the favor of the active portfolio. However, when we look at the differences in wealth gain we see that the active strategy only comes a head by a mere 2.91%, with a difference in compounded returns of only 15 basis points. This is because of the long periods of strong bull markets in which the passive strategy has been able to greatly increase its wealth while the active one has trailed behind. The few downwards periods we have had in this time period were short, but quite strong, with the crash of 2008 leading to a drawdown of almost 51% for the passive strategy.

One of the most popular lookback periods to use is the 10 month one, we however found that for the 1857-2011 period we averaged at 6.63 month period in our expanding window out-of-sample test. As we mentioned then, we thought this could be explained by the fact that we mostly had short lookback periods during the first 50 years of our strategy. We see further evidence of this now as we now have an mean k of 9.68, in other words when looking at the same time period that other authors such as Faber (2007) who ran his tests from 1900 and forward, we get about

the same optimal lookback period. While this in itself is not an important point, it shows that the optimal lookback period is dependent on which time period you look at, and that further strengthens the argument for using a fluent lookback period and not a fixed one.

In short the active strategy barely comes a head for this period in regards to the development of wealth, but as before an investor would get a drastic increase in his risk-adjusted returns by employing a market timing strategy. A market timing strategy would increase the Sharpe ratio by nearly 21% while still providing a 2.91% increase in wealth over the passive strategy.

Let us once again take a quick look at the results we would have gotten if he had used the rolling window scheme instead of the expanding one.

Table 11 – Comparing key figures for rolling and expanding-window strategies

Index	S&P500
Type	Expanding
Start	1926
In-sample	1950
End	2011
Mean k	9.68
% Sharpe	20.96%
% Drawdown	-54.35%
% Wealth	2.91%
% Volatility	-21.92%

| Index	S&P500
Type	Rolling
Start	1926
In-sample	1950
End	2011
Mean k	10.91
% Sharpe	18.61%
% Drawdown	-54.35%
% Wealth	-0.20%
% Volatility	-20.68%

We can see in Table 11 that the results are quite similar, but this time the expanding window comes out a head. This is due to the fact that this time they were equally good at predicting the major downward trends in the market and thus the rolling window scheme were not able to get such a strong start compared to the expanding window like it did for the entire time period. On the whole the expanding strategy was able to reduce volatility slightly better and as such was able to produce a better risk-adjusted return. The expanding window strategy was able to outperform the passive strategy on not only total returns but risk-adjusted returns as well, whereas the rolling strategy only did better than the passive one when it comes to risk-adjusted

returns.

This shows us that while the two strategies produce fairly similar results the margins between whether an active strategy or the passive strategy preforms best are so small that even minor differences influence the end results.

We can tamper with the split-off point, but instead of moving the split-off point forward in relation to the starting point, let us move it backwards this time. Let us set it to 1940 and see how, or if, this affects our results.

Figure 9– Logarithmic growth in wealth for 1940-2011 Expanding Window
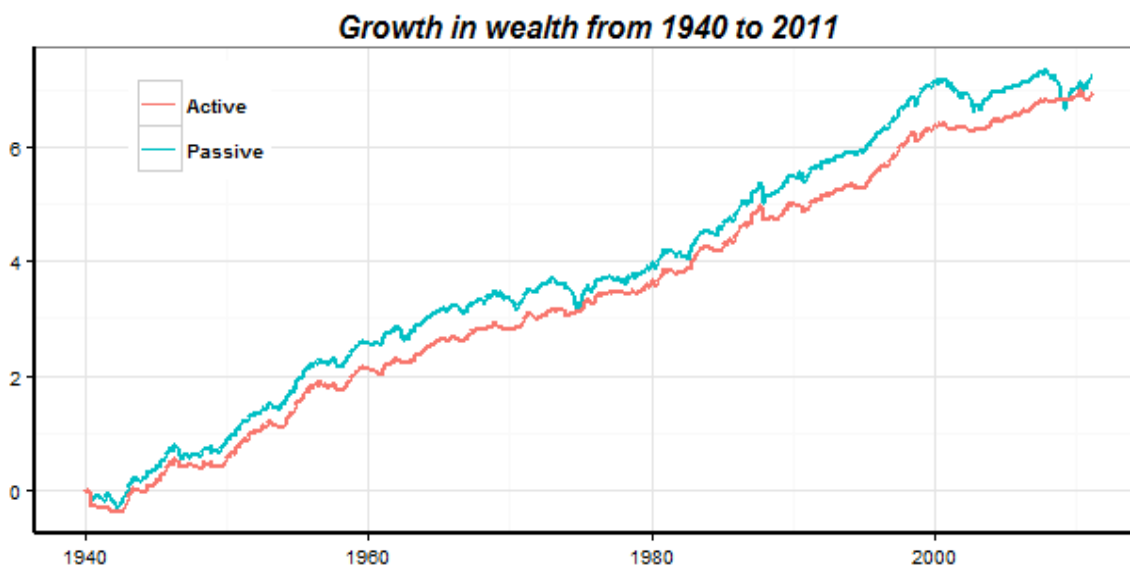
During the first three years both of the portfolios produce negative results, from 1940 to 1943 neither of the strategies are able to give us positive returns. This can be due to the uncertainty the markets experienced as a result of the second world war. As the positive economical effects of war start to kick into gear the American stock markets being to incline again and soon after our portfolios are back to their break even point. The active portfolio gets its advantage from being able to exit the risky markets before the major falls, it does this at the expense of being slower and as such it takes longer for it to reenter the market and experience the upward trends. Seeing as we now start our out-of-sample period directly into a downward period the active portfolio is

not able to get any immediate advantages over the passive strategy and as such the passive strategy gets a solid advantage in wealth gain for the entire period. Even the strong bear markets of the 2000s are not enough to get the strategies back on a even footing when it comes to pure wealth gain.

Once again though, the active strategy is still able to produce a higher risk-adjusted return than the passive strategy but this time the difference is a lot smaller, with an increase in Sharpe ratio of about 12%.

Table 12 – Key figures for 1940-2011 Expanding Window

	Passive	Active
Volatility	14.82%	11.72%
Sharpe	0.488	0.546
Max Drawdown	50.96%	31.94%
Avg Yrl Return	11.96%	11.05%
Compound Ret	10.75%	10.30%
Final Wealth	1408	1053
Relative Increase over Passive		Active
Sharpe	-	11.98%
Drawdown	-	-37.31%
Wealth	-	-25.17%
Volatility	-	-20.87%

| Index	S&P500
Type	Expanding
Start	1926
In-sample	1940
End	2011
Mean k	9.83

This once again supports the assumptions we previously made, our strategy performs best during volatile periods were it can reduce volatility by predicting the market based on signals created using momentum. It does not outperform the simple buy-and-hold during pure bull periods, but relies on the inevitability of the market crashing at some point. We also see that once again the choice of split-off point has changed the fact that our active strategy now only produces a higher risk-adjusted return and not a higher total return, as it did with the split-off point at 1950. Setting the split-off point even further back to 1935 provides us with the same conclusions, but the active strategy now has an even lower total return with a reduction in final wealth of 34% while actually increasing the difference in risk-adjusted rewards with an increase in Sharpe ratio to a difference of almost 14%. It is worth noting that the average lookback period for both of these

split-off points still lies around 10 months.

5.2.3 S&P 500 1857-1926

Looking at this earlier data set it is important to remember that the results we get here might not be as good of an representation of the results we would be able to get if we were to apply our active strategy as of today. This is due to many factors, but maybe the most important one is the great advances in technology that allows even the "average Joe" to get a hold of and utilize information and use this information in meaningful ways. If we were to try to manually do the calculations used in this paper it would take countless man hours that we now can avoid easily by using modern computers. Countless advancements in technology has lead to the markets being more efficient, and as such has reduced our ability to "beat" the market.

Figure 10– Logarithmic growth in wealth for 1881-1926 Expanding Window
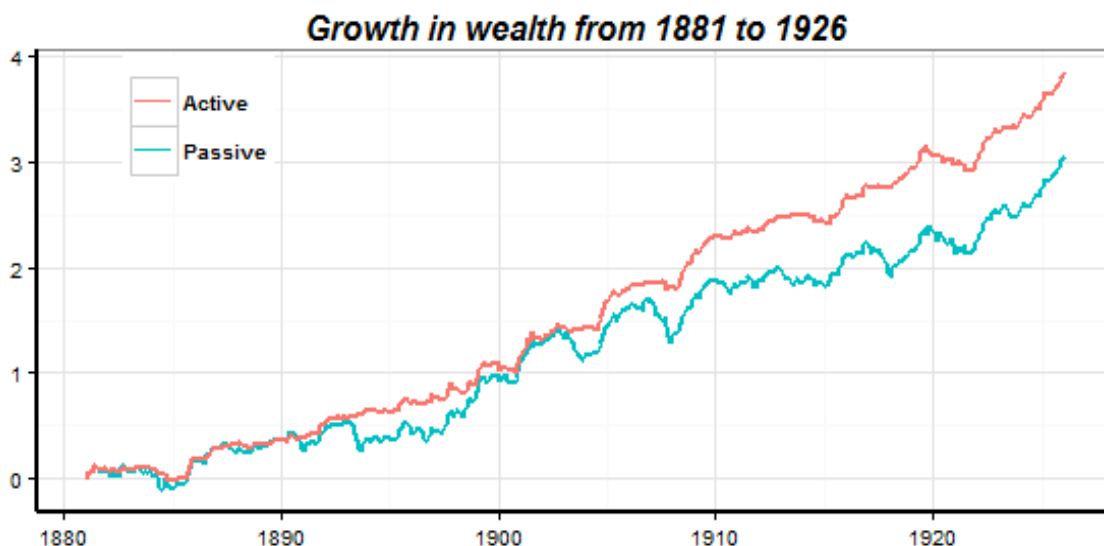
With a quick glance at Figure 10 we see that the passive strategy experiences a lot of drawdowns in this period, with the growth in wealth often dropping down quite a bit. This leads to our active strategy ending up with a massively increased wealth at the end of the period. Considering this is a logarithmic scale an increase in wealth from 3 to almost 4 is massive.

Looking at the key figures for this time period we can see that the increase of wealth is indeed significant, with a difference of 120% between the two strategies.

Table 13 – Key figures for 1881-1926 Expanding Window

	Passive	Active
Volatility	11.50%	8.46%
Sharpe	0.313	0.597
Max Drawdown	33.93%	22.37%
Avg Yrl Return	7.74%	9.31%
Compound Ret	7.05%	8.94%
Final Wealth	21.44	47.17
Relative Increase over Passive		Active
Sharpe	-	90.52%
Drawdown	-	-34.06%
Wealth	-	120.01%
Volatility	-	-26.47%

Index	S&P500
Type	Expanding
Start	1857
In-sample	1881
End	1926
Mean k	3.00

The active strategy has a solid decrease in volatility as well as drawdown, and as we could read from Table 13 it handles the lesser drawdowns better than the passive strategy does. The high volatility in this period and the amount of drawdowns are probably causes for the massive increase in efficiency for the active strategy, it has a higher compounded return of almost 2 percentage points. Sharpe ratios are based on volatility and returns, and as such it is no surprise that the active portfolio has a much higher Sharpe ratio than the passive one, a staggering 90% difference between them. A much higher Sharpe ratio than we have been able to generate for the other time periods.

We see, as we thought, that the reason we had such a low average lookback period when looking at the entirety of the data set was because of the earlier periods. For our current time period we are averaging at a lookback period of 3 months. When looking at the same figure for the rolling window though, we are getting an average of 7.15. This is rather odd, it is not expected that we would get such a big difference between the two models.

Table 14– Comparing key figures for rolling and expanding-window strategies

Index	S&P500
Type	Expanding
Start	1857
In-sample	1881
End	1926
Mean k	3.00
% Sharpe	90.52%
% Drawdown	-34.06%
% Wealth	120.01%
% Volatility	-26.47%

Index	S&P500
Type	Rolling
Start	1857
In-sample	1881
End	1926
Mean k	7.15
% Sharpe	76.06%
% Drawdown	-60.27%
% Wealth	85.45%
% Volatility	-26.44%

When examining the other results from the rolling window scheme we see that it has performed quite well for this period as well. It does not provide us with as good results which that of the expanding one, but they are still a great improvement over the passive strategy. An increase in risk-adjusted and total returns of this magnitude must be considered a success when it comes to our active strategy.

Let us push the split-off point another 14 years and set it to 1895 and examine the effectiveness of our strategy then.

Table 14 – Comparing key figures for rolling and expanding-window strategies

Index	S&P500
Type	Expanding
Start	1857
In-sample	1895
End	1926
Mean k	3.00
% Sharpe	64.60%
% Drawdown	-34.06%
% Wealth	71.40%
% Volatility	-23.64%

Index	S&P500
Type	Rolling
Start	1857
In-sample	1895
End	1926
Mean k	6.75
% Sharpe	64.80%
% Drawdown	-59.90%
% Wealth	70.16%
% Volatility	-24.04%

For the first time switching the split-off point has not actually changed whether or not the active strategy is better than the passive strategy when it comes to both total as well as risk-adjusted returns. It has however significantly reduced the results we got with the previous split-off point.

5.2.4 Dow Jones

The Dow Jones data set has a different time horizon to that of the S&P 500 one, it goes from June 1896 to 2011, to simplify the fact that our data starts in the middle of the year we have we have chosen to set our beginning to 1897 while still going to 2011. The initial split-off point has been set to 1926 which means that the initial in-sample period does not have the same amount of months in it as it does for the rest of the S&P 500 results. In order to effectively compare the results we get from the two indices we will also be using the S&P 500 for this part, but set its limits to that of the Dow Jones data. That means its beginning, split-off point, and ending will be at the same time as the Dow Jones data.

Figure 11 – Logarithmic growth in wealth for 1926-2011 Expanding Window
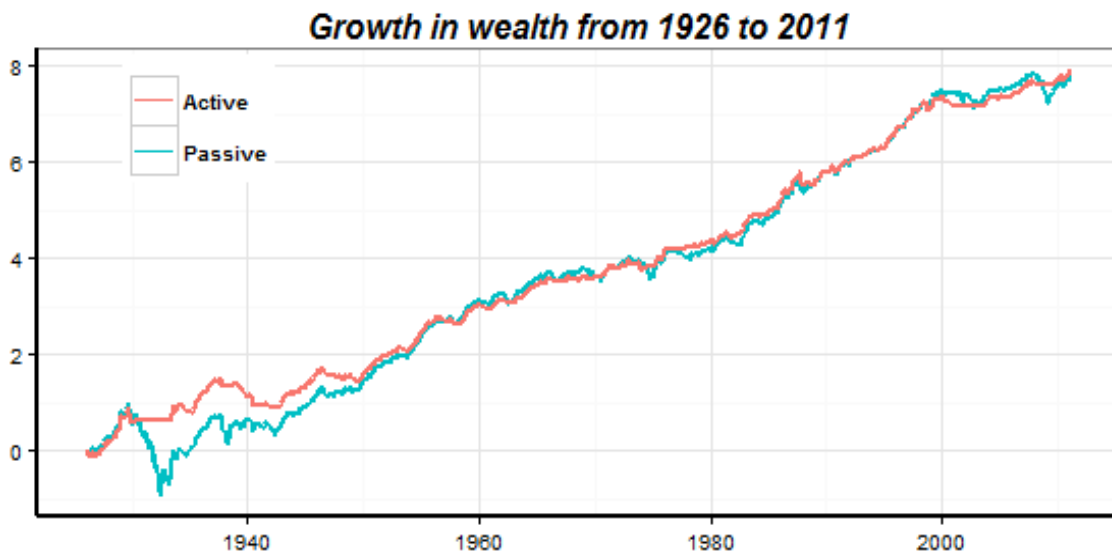
As we can see from Figure 11 our strategy behaves in the same patterns as it did for the S&P 500 data. It exits the market in 1930s and goes into risk free and because of this is able to avoid the worst of the great depression. After this it stays above the passive strategy until we enter the long and strong bull markets from the 50s and then they stay relatively even until the markets crashing again in 2008, ending up with the active strategy slightly a head of the passive one in terms of pure wealth development.

Table 15 – Key figures for 1926-2011 for the Dow Jones index
Expanding Window

	Passive	Active
Volatility	18.62%	12.66%
Sharpe	0.389	0.509
Max Drawdown	85.58%	44.30%
Avg Yrl Return	11.48%	10.61%
Compound Ret	9.59%	9.74%
Final Wealth	2393.12	2693.32
Relative Increase over Passive		Active
Sharpe	-	31.03%
Drawdown	-	-48.24%
Wealth	-	12.54%
Volatility	-	-32.01%

Index	DowJones
Type	Expanding
Start	1897
In-sample	1926
End	2011
Mean k	12.96

This is reflected in the results as well, we can see that the active strategy gives us a slightly higher ending wealth with a compounded return of 9.74% while also reducing volatility by a good third of that of the passive strategy. The active strategy was able to greatly reduce the maximum drawdown experienced in the 30s down to a 44.30% drop instead of a 86% drop, this is what lead to its early head start that we can clearly see in Table 15. An increase in the Sharpe ratio of 31.03% reflects the increase in risk-adjusted returns. For the Dow Jones index we see that the standard optimum of a 10 months lookback period does not hold true, averaging at almost 13 months. While this might not be abnormal it is still interesting, for the Dow Jones index it seems that it is profitable to extend your lookback period further than for the S&P 500 index.

Comparing the two indices using the same time horizon shows us that there are similar results from both of them. There are differences, like the fact that the Dow Jones index produces a slightly lower volatility, or that the S&P500 index ends up having a higher final wealth for its passive portfolio. The passive strategy used on the S&P500 index produces a higher final wealth than the active strategy does on either of the indices.

Figure 12 – Logarithmic growth in wealth for 1926-2011 Expanding Window
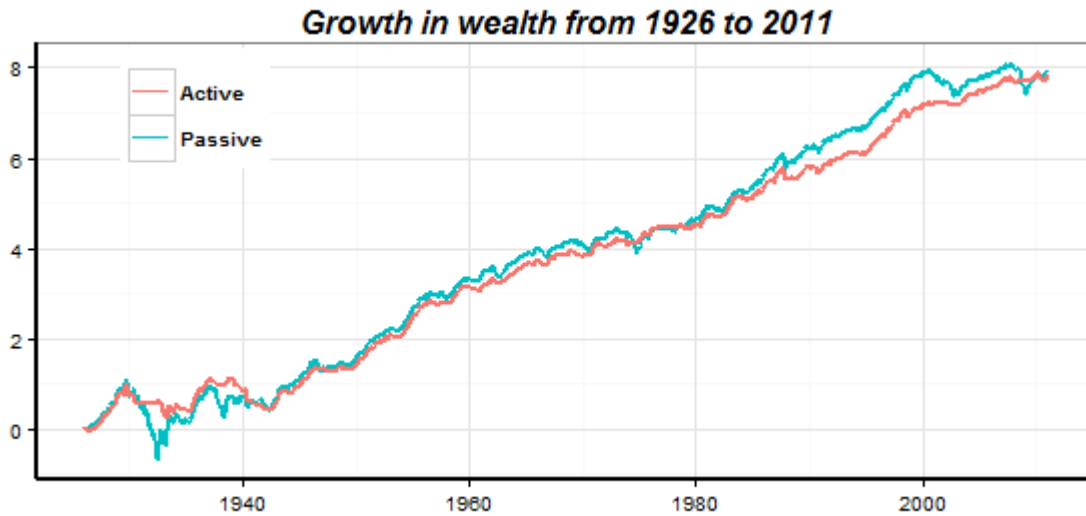
Table 16 – Key figures for 1926-2011 for the S&P 500 index Expanding Window

	Passive	Active
Volatility	19.22%	12.86%
Sharpe	0.395	0.498
Max Drawdown	83.10%	53.93%
Avg Yrl Return	11.86%	10.57%
Compound Ret	9.86%	9.66%
Final Wealth	2953	2538
Relative Increase over Passive		Active
Sharpe	-	26.32%
Drawdown	-	-35.10%
Wealth	-	-14.03%
Volatility	-	-33.09%

Index	S&P500
Type	Expanding
Start	1897
In-sample	1926
End	2011
Mean k	8.66

49

However, the results we have gotten using our active strategy with an expanding-window scheme on the Dow Jones Index are comparable to the results we have gotten using the S&P500 index. The results show the same patterns and seems to have the same mannerisms as the results we have gotten previously. This strengthens the results we have gotten and helps to such that they are not a result of random chance and as such helping to validate the conclusions we can draw from them.

5.3   Comparing the In-sample and Out-of-sample Results

Table 17 – Comparing the in-sample and out-of-sample results for 1881-2011

Relative increase over passive	Active	Active CAP
Sharpe	50.09%	62.51%
Drawdown	-39.26%	-43.73%
Wealth	195.99%	266.34%
Volatility	-32.64%	-46.03%

Index	S&P500
Type	Expanding
Start	1857
In-sample	1881
End	2011
Mean k	6.63
% Sharpe	30.71%
% Drawdown	-27.63%
% Wealth	10.57%
% Volatility	-31.65%

The time horizon for the two strategies are slightly different due to the out-of-sample model needing to have a in-sample period in order to produce the optimal trading rule. The in-sample model goes from 1881 to 2011 while the out-of-sample model is from 1857-2011 with a split-off point at 1881.

We see that the in-sample model produces far more positive results than the out-of-sample model, which is in no way surprising . The out-of-sample test is a far better representation of the how the SMA strategy would have performed had it been applied in 1881, and the large differences in the efficiency of the out-of-sample test and in-sample test show us how exaggerated the reports on the in-sample SMA strategy is.

6 Conclusion

We find that when using the SMA with a 10-month lookback period and applying it to an entirely in-sample data set we were able to consistently outperform the buy-and-hold strategy for a differing set of time periods. The active strategy severely reduced volatility while still giving about the same or higher average annual returns, which lead to higher compounded returns and a higher final wealth. Our active strategy produced both a higher total return as well as a higher risk-adjusted return and as such reflects the same results as that of others such as Faber (2007) and Siegel (2002).

However, it was not enough for us to just look at how a timing strategy works when based entirely on an in-sample data set. An in-sample data set does not provide us with results that are accurate enough to be able to provide a reliable enough conclusion. When performing a test based purely on in-sample data you are provided with perfect information and can make decisions based on what you know, instead of what you expect to happen in the future. The saying "hindsight is 20/20" reflects this issue well, it is easy enough to be able to optimize a strategy that outperforms the buy-and-hold strategy when we know exactly how the markets are going to evolve. Therefor we did a second set of test, were we split our time period into two, one in-sample period were we created the optimum trading rules, and one out-of-sample period were we then tested this rule and recorded the results.

Using this out-of-sample test we were able to get results that would more accurately predict the efficiency of the model if it were to be implemented today. The results showed that as expected the out-of-sample test gave us a worse risk-adjusted return than the in-sample one. On a whole we were able to reduce the volatility by about the same amount, but the in-sample test was able to better reduce the drawdowns and as such ended up with a much higher wealth gain and risk-adjusted return.

Most important however is how our active strategy did compared to the passive strategy when applied in this out-of-sample manner. Looking at the time period as a whole, with the split-off point set at 1881, we can see that the out-of-sample model gave us a higher total return on our

investment, as well as a higher risk-adjusted with an increase in Sharpe ratio of 30.71%. Once we started examining the effects of moving this split-off point we began to see the difficulties with the out-of-sample active strategy. Moving the split-off points by only a few years had large impacts on the results we got, with it often being the deciding factor of whether or not the active strategy gave us a higher total return than the passive strategy. The active strategy was always able to provide us with a higher Sharpe ratio, but for all but one time moving the split-off point tipped the bar on pure wealth gain into the passive strategies favor.

We saw that when starting the out-of-sample period in the 1940s when the markets were in a decline our strategy performed the worst comparatively to the passive strategy. This is because the downward period was so short, had the markets continued to behave bearish the active strategy would have stayed in risk-free while they buy-and-hold would experience severe losses. Since the markets so quickly started to incline again the active strategy lost out due to its delay in returning to risky assets. It show us how fragile the active strategy is, since it relies on being able to time when the markets will experience a decline it will not perform well if set directly into a short declining period. Adding onto this is the problem that the active strategy is not able to keep up with the passive one in long periods of strong bull markets. An investor that went into our active strategy in the early 1950s would not have been able to see a profit over the passive one until the mid 1970s, and then it would take another 30 years before he would see one again. Most people do not invest with such a long time span, and even though his risk-adjusted returns would have been higher with the active portfolio, the actual returns he would have gotten would have been lower for most of the period.

We have used a longer time window for both our in-sample and out-of-sample test than what previous authors have used, going as far back as 1857 compared to the most commonly used starting point of 1926. This has not played a significant role in the conclusion we can draw from the data set, we have however noticed a trend in the results. The active strategy performed better in the past than the present, the further back in time the better the risk-adjusted returns you are able to generate. This implies that the markets are getting more and more efficient and as such the moving average strategy is getting less effective.

In conclusion the SMA market timing strategy when tested on an out-of-sample data set was able to produce a higher risk-adjusted return than the passive strategy, but the results it produced were far to fragile and subject to chance in way of the split-off point for us to be able to say that it will reliably outperform the buy-and-hold strategy for all periods. It relies heavily on some factors in which we have no control, such as the way the markets develop. When taking into consideration the transaction costs and taxes that we so far have ignored, it does not seem likely that our market timing strategy provides us with an unreasonable return compared to the risk you get. You will consistently get a lower risk than with a passive strategy, but your returns will in most cases also suffer. We can only conclude that the SMA strategy only consistently outperforms the passive strategy when done in a heavily flawed in-sample test.

Acknowledgments

References

Amit Goyal Stock Returns
http://www.hec.unil.ch/agoyal/

Bodie, Kane and Marcus -"*Investments and Portfolio Management*" – 2010 Edition

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.

Eling and Schumacher(2007) «*Does the Choice of Performance Measure Influence the Evaluation of Hegde Funds?*» Journal of Banking and Finance, *31* (9), 2632 – 2247.

Eling (2008) «*Does Hedge Fund Performance Persist?Overview and New Empirical Evidence*» Financial Analysts Journal, 64 (3), 54–66

Faber, Mebane .T (2007) "*A Quantitative Approach to Tactical Asset Allocation*"-The Journal of Wealth Management, Spring 2007 , 69-79

Goyal, A. and Welch, I. (2008). "*A Comprehensive Look at the Empirical Performance of Equity Premium Prediction*", Review of Financial Studies, 21 (4), 1455–1508.

Gwlying, Owain ap. Clare, Andrew. Seaton, James . and Thomas, Stephen . "*Price and Momentum as Robust Tactical Approaches to Global Equity Investing*" Journal of Investing, 19 (3), 80–91.

Hunter, R. "*Which indicies best represent the market*" Hargreaves Lansdown (2012)

Kilgallen, T. (2012) "*Testing the Simple Moving Average across Commodities, Global Stock Indicies, and Courrencies*" Journal of Wealth Management, 15 (1), 82–100.

Moskowitz,Ooi and Pedersen (2011), "*Time Series Momentum*", Journal of Financial Economics (2012), 228-250

New York Univeristy – Stern School of Business
http://pages.stern.nyu.edu/~adamodar/New_Home_Page/invfables/dividirrelevance.htm

Schwert Stock Returns 1802-1925
http://schwert.ssb.rochester.edu/mstock.htm

Siegel , J.(2002) "*Stocks for the Long Run*" - McGraw-Hill Companies 2007 Edition

Sullivan, Timmermann and White (1999) "*Data-Snooping, Techincal Trading Rule Performance, and The Bootstrap*" Journal of Finance, 54 (5), 1647–1691.

Zakamulin, V. "*The Real-Life Performance of Market Timing with Moving Average and Time-Series Momentum Rules*" - Working Paper, University of Agder

Appendix

Main program used for calculating out-of-sample expanding window SMA for either S&P 500 or Dow Jones

```
rm(list=ls(all=TRUE))
setwd("C:\\New folder\\Dropbox\\Skole Master\\Master Thesis\\R")
source("Myfunctions.r")
library(ggplot2)
library(PerformanceAnalytics)
#Read data
datasp = read.table("S&P500.csv", header=TRUE)
datadj = read.table("DowJones.csv", header=TRUE)

#Defining variables from data
retsp=datasp$Return
retdj=datadj$Return
rf= datasp$ETB
date=datasp$Date

date=ts(date,start=c(1857,1),frequency=12)
retsp = ts(retsp, start=c(1857,1), frequency = 12)
rf = ts(rf, start=c(1857,1), frequency = 12)

#Time window
time.start=1897
time.end=2011
insample.endstart=1926
insample.endstartnr=(which(time(retsp)==insample.endstart))
index.end=(which(time(retsp)==time.end))
```

```
numberoftimesdone=(index.end-insample.endstartnr+1)

#Testing variables
r.time=rep(0,numberoftimesdone)
r.lets=rep(0,numberoftimesdone)

Q=14
ind.start = which(time(retsp) == time.start)

start.time <- Sys.time()

for (time in 1:numberoftimesdone){
 #Defining in sample
 ind.end=insample.endstartnr+time-2

 n = ind.end-ind.start +1
 r.port = rep(0,n)

 r.stocks = retsp[ind.start:ind.end]
 r.tbill = rf[ind.start:ind.end]

 Sharpe.testvariable=rep(0,Q)

 for (lookback in 2:Q){
  for (i in 1:n) {
   period.end = ind.start + i - 2
   period.start = period.end - lookback + 1
   if (period.start < 1) period.start <- 1
   signal = SMA(retsp[period.start:period.end])
   if (signal>0) r.port[i] = retsp[period.end+1] else r.port[i] = rf[period.end+1]
  }
  Sharpe.testvariable[lookback] = SR(r.port-r.tbill)
 }
```

```
bestk=which.max(Sharpe.testvariable)
r.lets[time]=bestk

# now we know the best lookback and need to define the signal for time "time"
period.end = ind.end
period.start = period.end - bestk + 1
if (period.start < 1) period.start <- 1
signal = SMA(retsp[period.start:period.end])
if (signal>0) r.time[time] = retsp[period.end+1] else r.time[time] = rf[period.end+1]

}

end.time <- Sys.time()
time.taken <- end.time - start.time
time.taken # to measure how long the program works

retsp = window(retsp, start=c(insample.endstart,1), end=c(time.end,1))
ret.passive = as.numeric(retsp)
rf = window(rf, start=c(insample.endstart,1), end=c(time.end,1))
r.tbill = as.numeric(rf)
date= window(date, start=c(insample.endstart,1), end=c(time.end,2))

ind.passive = cumprod(c(1,1+ret.passive))
ind.active = cumprod(c(1,1+r.time))

ind.ts = ts(cbind(ind.passive,ind.active), start=c(insample.endstart,1),frequency=12)
plot(log(ind.ts), plot.type = "single", col=c("blue","red"))
legend(x="topleft", legend=c("Passive","Active"), col=c("blue","red"), lty=1)

titlea=insample.endstart
titleb=time.end
titlec="Growth in wealth from"
titled="to"
```

```
test_data=data.frame(ind.passive,ind.active)

ind.df <- as.data.frame(ind.ts)
ind.df$date <- seq(as.Date('1926-01-01'), as.Date('2011-02-01'), by = "month")

ggplot(ind.df, aes(date)) +
  geom_line(aes(y = log(ind.passive), colour = "Passive",size="1")) +
  geom_line(aes(y = log(ind.active), colour = "Active",size="1"))+
  theme_bw()+
  theme(legend.position=c(0.15, 0.85))+
  theme(legend.text = element_text(colour="black", size = 10, face = "bold"))+
  labs(title=paste(titlec,titlea,titled,titleb,sep=" "))+
  theme(axis.line = element_line(size = 1))+
  guides(size=FALSE)+
  guides(colour=guide_legend(title=NULL))+
  theme(axis.title=element_text(face="bold.italic",
                    size="12", color="black"))+
  theme(plot.title = element_text(size = rel(1.3),face="bold.italic"))+
  labs(x="",y="")


#Growth of Wealth, with initial wealth = 100
n <- length(r.time)
wealth.a <- cumprod(c(1,1+r.time)) # from portfolio of stocks and bonds
wealth.p <- cumprod(c(1,1+ret.passive)) # from stocks


#Final wealth rounded to 2 decimal points
Wealth.Final.Passive <- round(wealth.p[n+1], digits=2)
Wealth.Final.Active <- round(wealth.a[n+1], digits=2)


Sharpe.passive=SR(ret.passive-r.tbill)
```

```
Sharpe.active=SR(r.time-r.tbill)

MtAr.Passive=MtA(mean(ret.passive))

MtAr.Active=MtA(mean(r.time))

volpassive=Vol(ret.passive)

volactive=Vol(r.time)

MeanK=mean(r.lets)

Drawdownpassive=maxDrawdown(ret.passive)

Drawdownactive=maxDrawdown(r.time)

Program="Expanding"

Indexnavn="S&P500"

Keyfigures=data.frame(volpassive,volactive,
                Drawdownpassive,Drawdownactive,
                Wealth.Final.Passive,Wealth.Final.Active,
                MtAr.Active,MtAr.Passive,
                Sharpe.active,Sharpe.passive,MeanK,
                time.start,
                time.end,
                insample.endstart,
                Program,
                Indexnavn)
View(Keyfigures)
```

Program used to calculate Rolling Window SMA

```
rm(list=ls(all=TRUE))
setwd("C:\\New folder\\Dropbox\\Skole Master\\Master Thesis\\R")
source("Myfunctions.r")
library(ggplot2)
library(PerformanceAnalytics)
#Read data
```

```r
datasp = read.table("S&P500.csv", header=TRUE)
datadj = read.table("DowJones.csv", header=TRUE)

#Defining variables from data
retsp=datasp$Return
retdj=datadj$Return
rf= datasp$ETB
date=datasp$Date

date=ts(date,start=c(1857,1),frequency=12)
retsp = ts(retsp, start=c(1857,1), frequency = 12)
rf = ts(rf, start=c(1857,1), frequency = 12)

#Time window
time.start=1857
time.end=2011
insample.endstart=1881
insample.endstartnr=(which(time(retsp)==insample.endstart))

index.end=(which(time(retsp)==time.end))
numberoftimesdone=(index.end-insample.endstartnr+1)

#Testing variables
r.time=rep(0,numberoftimesdone)
r.lets=rep(0,numberoftimesdone)

Q=24
index.start = which(time(retsp) == time.start)

start.time <- Sys.time()

for (time in 1:numberoftimesdone){
  #Defining in sample
```

```
ind.end=insample.endstartnr+time-2
ind.start=index.start+time
n = ind.end-ind.start +1
r.port = rep(0,n)


r.stocks = retsp[ind.start:ind.end]
r.tbill = rf[ind.start:ind.end]


Sharpe.testvariable=rep(0,Q)


for (lookback in 2:Q){
  for (i in 1:n) {
    period.end = ind.start + i - 2
    period.start = period.end - lookback + 1
    if (period.start < 1) period.start <- 1
    signal = SMA(retsp[period.start:period.end])
    if (signal>0) r.port[i] = retsp[period.end+1] else r.port[i] = rf[period.end+1]
  }
  Sharpe.testvariable[lookback] = SR(r.port-r.tbill)
}
bestk=which.max(Sharpe.testvariable)
r.lets[time]=bestk


# now we know the best lookback and need to define the signal for time "time"
period.end = ind.end
period.start = period.end - bestk + 1
if (period.start < 1) period.start <- 1
signal = SMA(retsp[period.start:period.end])
if (signal>0) r.time[time] = retsp[period.end+1] else r.time[time] = rf[period.end+1]


}


end.time <- Sys.time()
```

```
time.taken <- end.time - start.time
time.taken # to measure how long the program works


retsp = window(retsp, start=c(insample.endstart,1), end=c(time.end,1))
ret.passive = as.numeric(retsp)
rf = window(rf, start=c(insample.endstart,1), end=c(time.end,1))
r.tbill = as.numeric(rf)
date= window(date, start=c(insample.endstart,1), end=c(time.end,2))


ind.passive = cumprod(c(1,1+ret.passive))
ind.active = cumprod(c(1,1+r.time))




titlea=insample.endstart
titleb=time.end
titlec="Growth in wealth from"
titled="to"
test_data=data.frame(ind.passive,ind.active)


ind.ts = ts(cbind(ind.passive,ind.active), start=c(insample.endstart,1),frequency=12)


ind.df <- as.data.frame(ind.ts)
ind.df$date <- seq(as.Date('1881-01-01'), as.Date('2011-02-01'), by = "month")


ggplot(ind.df, aes(date)) +
  geom_line(aes(y = log(ind.passive), colour = "Passive",size="1")) +
  geom_line(aes(y = log(ind.active), colour = "Active",size="1"))+
  theme_bw()+
  theme(legend.position=c(0.15, 0.85))+
  theme(legend.text = element_text(colour="black", size = 10, face = "bold"))+
  labs(title=paste(titlec,titlea,titled,titleb,sep=" "))+
```

```
  theme(axis.line = element_line(size = 1))+
  guides(size=FALSE)+
  guides(colour=guide_legend(title=NULL))+
  theme(axis.title=element_text(face="bold.italic",
                    size="12", color="black"))+
  theme(plot.title = element_text(size = rel(1.3),face="bold.italic"))+
  labs(x="",y="")


#Growth of Wealth, with initial wealth = 100
n <- length(r.time)
wealth.a <- cumprod(c(1,1+r.time)) # from portfolio of stocks and bonds
wealth.p <- cumprod(c(1,1+ret.passive)) # from stocks


#Final wealth rounded to 2 decimal points
Wealth.Final.Passive <- round(wealth.p[n+1], digits=2)
Wealth.Final.Active <- round(wealth.a[n+1], digits=2)

Sharpe.passive=SR(ret.passive-r.tbill)
Sharpe.active=SR(r.time-r.tbill)
MtAr.Passive=MtA(mean(ret.passive))
MtAr.Active=MtA(mean(r.time))
volpassive=Vol(ret.passive)
volactive=Vol(r.time)
MeanK=mean(r.lets)
Drawdownpassive=maxDrawdown(ret.passive)
Drawdownactive=maxDrawdown(r.time)
Program="Rolling"
Indexnavn="S&P500"

Keyfigures=data.frame(volpassive,volactive,
```

Drawdownpassive,Drawdownactive,

            Wealth.Final.Passive,Wealth.Final.Active,

            MtAr.Active,MtAr.Passive,

            Sharpe.active,Sharpe.passive,MeanK,

            time.start,

            time.end,

            insample.endstart,

            Program,

            Indexnavn)

View(Keyfigures)

Functions used in the different programs

#Computes the annualized Sharpe ratio

SR <- function(er) {

  return(mean(er)/sd(er)*sqrt(12))

}

#Used to create the signal for the SMA

SMA <- function(mkt) {

  index <- cumprod(c(1,1+mkt))

  signal <- index[length(index)]-mean(index)

  return(signal)

}

#Creates the estimated tbillreturns

tbrestimate <- function(tscprates)

{-0.004+ 0.886*tscprates}

#From percentages to monlthy returns

atm <- function(cpr)

{(rates <-cpr/100)

{(1+rates)^(1/12)-1}}

```
#Computes the annulized volatility of returns
Vol <- function(vol) {
  return(sd(vol)*sqrt(12))
}
#Converts monlthy returns to annual returns
MtA <- function(mta) {
  return(((1+mta)^12)-1)
}
```

Program used to calculate in-sample SMA

```
rm(list=ls(all=TRUE))
setwd("C:\\New folder\\Dropbox\\Skole Master\\Master Thesis\\Innlevering")
source("Myfunctions.r")
library(ggplot2)
#Read data

data = read.table("ER1857to1885-3.csv", header=TRUE)
dataCPR = read.table("commercialpaperrates.dat", header=TRUE)

#Get the total returns, CAP and risk-free rate of return from the data set

mkt <- data$Return
rf  <- data$ETB
CAP=data$CaptialAdj

#Estimating T-bill returns for 1857 to 1971

cpr=dataCPR[,3]
n=length(cpr)
tbr=rep(0,n)
```

```
tbr=tbrestimate(cpr)
monthlycpr=atm(tbr)




#Timesetting the Variables
mkt <- ts(mkt, start=c(1857,1), frequency = 12)
CAP <- ts(CAP, start=c(1857,1), frequency = 12)
rf <- ts(rf, start=c(1857,1), frequency = 12)

#Defining Portfolio timeperiod

#Based on a program from the lectures, but modifed in order to allow changes to
#when the time period ends. Originaly final wealth and Sharpe ratio did not
#depend on period end.
time.start <- c(1881,1)
time.end=c(2011,12)
ind.start <- which(time(mkt) == time.start)
ind.end=which(time(mkt) == time.end)
nobs=(ind.end-ind.start)

# Portfolio management part of the program

#Based on a program from the lectures. Modified in order to react to changes
#time.end as well as in order to produce returns based on Capital
#appreciation returns.



lookback.period = 10
n <- nobs +1
r.port <- rep(0,n)
ws <- rep(0,n)
r.portCAP=rep(0,n)
```

```
wsC=rep(0,n)

for (i in 1:n) {
  period.end <- ind.start + i - 2
  period.start <- period.end - lookback.period + 2
  if (period.start < 1) period.start <- 1
  signal <- SMA(mkt[period.start:period.end])
  if (signal>0) ws[i] <- 1
  r.port[i] <- ws[i]*mkt[ind.start+i-1] + (1-ws[i])*rf[ind.start+i-1]
}
for (i in 1:n) {
  period.end <- ind.start + i - 2
  period.start <- period.end - lookback.period + 2
  if (period.start < 1) period.start <- 1
  signal <- SMA(CAP[period.start:period.end])
  if (signal>0) wsC[i] <- 1
  r.portCAP[i] <- wsC[i]*mkt[ind.start+i-1] + (1-wsC[i])*rf[ind.start+i-1]
}


#Performance Measures and creating graph for wealth growth

#Setting the time period for the returns

r.stocks = mkt[ind.start:ind.end]
r.tbill = rf[ind.start:ind.end]


Sharpe.Stocks = SR(r.stocks-r.tbill)
Sharpe.Port = SR(r.port-r.tbill)
Sharpe.PortCAP=SR(r.portCAP-r.tbill)

#Growth of Wealth, with initial wealth = 1
```

```r
n <- length(r.port)
wealth.p <- cumprod(c(1,1+r.port)) # from portfolio of stocks and bonds
wealth.s <- cumprod(c(1,1+r.stocks)) # from stocks
wealth.pC=cumprod(c(1,1+r.portCAP))


#Final wealth rounded to 2 decimal points
Wealth.Final.Passive <- round(wealth.s[n+1], digits=2)
Wealth.Final.Active <- round(wealth.p[n+1], digits=2)
Wealth.Final.ActiveCAP=round(wealth.pC[n+1],digits=2)



#Time setting growth of wealth, combining the three portfolios and plotting
#them to eachother
wealth <- ts(cbind(wealth.s,wealth.p,wealth.pC), start=time.start,
        frequency = 12)


start <- time.start
end <- time.end
wealth <- window(wealth, start, end)
#plot(wealth, plot.type = "single", col=c("red","blue","green"),xaxt="n")
#legend(x="topleft", legend=c("Stocks","Timing Portfolio","Timing with CAP"),
    #col=c("red","blue","green"), lty=1)
#axis(1, at=seq(1860,2012,by=10))


#Volatility compuations
volstock=Vol(r.stocks)
volport=Vol(r.port)
volportcap=Vol(r.portCAP)
voltbill=Vol(r.tbill)


#Portfolio returns from monthly to annual


MtAr.stocks=MtA(mean(r.stocks))
```

```
MtAr.port=MtA(mean(r.port))
MtAr.portCAP=MtA(mean(r.portCAP))


#Drawdowns from the package "Performance Analytics"
ind.passive=r.stocks
ind.active=r.port
ind.activecap=r.portCAP
library(PerformanceAnalytics)
maxDrawdown(r.stocks)
maxDrawdown(r.port)
maxDrawdown(r.portCAP)
maxDrawdown(r.tbill)



titlea=start
titleb=time.end
titlec="Growth in wealth from"
titled="to"



test_data=data.frame(ind.passive,ind.active)

ind.df <- as.data.frame(wealth)
ind.df$date <- seq(as.Date('1881-01-01'), as.Date('2011-02-01'), by = "month")

ggplot(ind.df, aes(date)) +
  geom_line(aes(y = wealth.s, colour = "Passive")) +
  geom_line(aes(y = wealth.p, colour = "Active"))+
  geom_line(aes(y = wealth.pC, colour = "Active w/CAP"))+
  theme_bw()+
  theme(legend.position=c(0.12, 0.80))+
  theme(legend.text = element_text(colour="black", size = 10, face = "bold"))+
  labs(title=paste(titlec,titlea,titled,titleb,sep=" "))+
```

```
theme(axis.line = element_line(size = 1))+

guides(size=FALSE)+

guides(colour=guide_legend(title=NULL))+

theme(axis.title=element_text(face="bold.italic",

                     size="12", color="black"))+

theme(plot.title = element_text(size = rel(1.3),face="bold.italic"))+

labs(x="",y="")




Sharpe.passive=SR(r.stocks-r.tbill)

Sharpe.active=SR(r.port-r.tbill)

Sharpe.activeCAP=SR(r.portCAP-r.tbill)

MtAr.Passive=MtA(mean(r.stocks))

MtAr.Active=MtA(mean(r.port))

MtAr.ActiveCAP=MtA(mean(r.portCAP))

volpassive=Vol(r.stocks)

volactive=Vol(r.port)

volactiveCAP=Vol(r.portCAP)

Drawdownpassive=maxDrawdown(r.stocks)

Drawdownactive=maxDrawdown(r.port)

DrawdownactiveCAP=maxDrawdown(r.portCAP)

Program="SMA"

Indexnavn="S&P500"

Keyfigures=data.frame(volpassive,volactive,volactiveCAP,

            Drawdownpassive,Drawdownactive,DrawdownactiveCAP,

            Wealth.Final.Passive,Wealth.Final.Active,Wealth.Final.ActiveCAP,

            MtAr.Active,MtAr.Passive,MtAr.ActiveCAP,

            Sharpe.active,Sharpe.passive,Sharpe.activeCAP,lookback.period,

            time.start,

            time.end,

            Program,

            Indexnavn)

View(Keyfigures)
```

```
test=SMA(mkt[period.start:period.end])
```