Journal Of Investment Management, Vol. 19, No. 3, (2021), pp. 81–98

www.joim.com

A MARKET SIGNAL-BASED ALTERNATIVE TO BUY-AND-HOLD INVESTING

Atanu Saha[a] *and Yong Xu*[b]

We propose a simple, hindsight-free, rule-based method of entry and exit into the stock market, with the goal of improving returns by averting large losses. Using data from 1928 through March 2020, we demonstrate that the proposed strategy delivers statistically significant outperformance over the S&P 500 total return index. Several robustness checks, including a Monte Carlo analysis, confirm the strategy's outperformance in various sub-sample periods and investment horizons. These results hold after accounting for reasonable transaction costs for in and out trades. The strategy's outperformance is explained by the non-normality and asymmetric persistence of market returns.

1 Introduction

As the Covid-19 pandemic reared its head, February and March of 2020 were brutal to equity investors. The S&P 500 index suffered a nearly 34% drop in just 23 trading days, after reaching its peak in mid-February. In March, the net flow out of U.S. equity funds and into money market funds reached record amounts, dwarfing the outflows during the 2008–2009 financial crisis or the 2001–2002 market downturn.[1] Despite many influential academics' (Malkiel, 1996; Sharpe, 2002) and practitioners' (Bogle, 1994) repeated proclamations that trying to time the market is futile, investors' actual behavior evidenced by the flight to safety during this and other past market downturns displays a strong preference for loss avoidance, which is our paper's motivation.

The objective of this paper is to propose a simple, hindsight-free, rule-based method of entry and exit into the stock market, with the goal of improving returns by averting large losses. Using data from 1928 through March 2020, we demonstrate that the proposed market timing strategy delivers robust and statistically significant outperformance over the benchmark S&P 500 total return index. We undertake several robustness checks, including a Monte Carlo analysis, to confirm the strategy's outperformance, including generating positive alpha, in various sub-sample periods and over different lengths of investment horizon. The strategy also passes the Henriksson

StoneTurn, New York, NY.
[a]E-mail: asaha@stoneturn.com.
[b]E-mail: yxu@stoneturn.com.

and Merton (1981) test of successful market timing, demonstrating that the strategy's ability to time the market is robust. These results hold even after accounting for reasonable transaction costs for the strategy's in and out trades.

The strategy's entry and exit signal obviously does not avoid all negative market returns; it also misses many positive ones. However, the strategy outperforms its benchmark because of an important characteristic of the market's return distribution: asymmetry of outliers. That is, the outlying negative returns are considerably larger in absolute value than the outlying positive ones. As a result, the gain from avoided worst returns far outweighs the loss from missed best returns when the investor stays out of the market. Another important reason why the proposed strategy works is the persistence of negative returns; that is, a large down month is, more often than not, followed by another month with a negative return. Because our exit signal in any month is based on observing large negative returns in the *preceding* month, by exiting the market the strategy avoids substantial losses that often follow a large down month.

The rest of this paper is organized as follows. In the next section, we review the relevant literature, focusing on the studies that have proposed various market timing strategies with the goal of outperforming buy-and-hold investment portfolios. Section 3 describes this study's dataset. In Section 4, we motivate the paper by discussing the asymmetric impact of avoiding worst market returns versus missing an equal number of best ones. In Section 5, we lay out the proposed strategy and discuss its results, as well as various checks we undertake to verify that our results are robust. In Section 6, we provide explanations for the strategy's outperformance by examining the characteristics of market returns, focusing on negative skewness and negative return persistence, the key factors explaining the strategy's success. In Section 7, we demonstrate that the strategy continues to outperform its benchmark even after accounting for transaction costs. The last section concludes.

2 The Relevant Prior Literature

Many academic studies have shown that avoiding the worst returns of down markets while staying fully invested at other times greatly improves the long-term performance of an investment portfolio (Shilling, 1992; Stowe, 2000; Estrada, 2008). Although these studies focused on different time periods—the post-war era from 1946 to 1991 in Shilling (1992), the 1990s in Stowe (2000), and over a hundred-year period between 1899 and 2006 in Estrada (2008)—they share a common theme: avoiding the worst months, days, or 'black swan' events has an enormous impact on the long- term performance of an investor's portfolio. Avoiding the worst returns during selected trading days yields an annualized improvement in return of 7.8% in Shilling (1992) and 6.2% in Estrada (2008) over a buy-and-hold portfolio of Dow Jones Industrial Average (DJIA). In Stowe (2000) the outperformance is more pronounced at 15.9% over the Center for Research in Securities Prices (CRSP) index.

However, none of these foregoing studies have attempted to construct or propose a practical method that would allow an investor to avoid the worst market returns. In other words, it is only known *ex post* when the worst returns were, and not *ex ante*. There is, however, an extensive body of literature that has investigated various signals to potentially predict *ex ante*, and thus avoid, large negative returns. The approaches in these studies can be characterized as market timing strategies. The *ex ante* signals on which the market timing strategies rely fall into two groups: fundamental ones, that is, signals related to the market, macroeconomic factors, or company valuation indicators; and technical, wherein the signals are derived from patterns of past market returns. Both streams of studies have examined whether market timing strategies can outperform the returns of a passive, buy-and-hold benchmark portfolio.

The studies relying on fundamental signals include Vergin (1996), Resnick and Smith (2002), Shen (2003), Cooper and Chieffe (2004), and Feldman *et al.* (2015). Each of these studies measures the relative performance of market timing versus a buy-and-hold portfolio of the S&P 500 index, adjusted for dividends. There is, however, considerable variety in the fundamental signals examined; they include, company valuations (Shen, 2003); asset prices (Vergin, 1996; Resnick and Smith, 2002); and economic indicators (Cooper and Chieffe, 2004; Feldman *et al.*, 2015). Shen (2003) relied on the spread between the earnings-to-price (E/P) ratio of the S&P 500 index and interest rates as the market timing signal. Resnick and Smith (2002) used the shape of the yield curve to forecast impending economic recession and a market downturn. The timing signals used in Cooper and Chieffe (2004) are turning points from economic expansion to contraction and investments are timed according to the phases of business cycles. Feldman *et al.* (2015) examined a variety of signals and concluded that the best performer was a market timing strategy based on Conference Board's Leading Economic Indicator. With the exception of Vergin (1996), all the aforementioned studies show that fundamental signal-based market timing strategies produce some incremental returns over a buy-and-hold market portfolio. However, a major limitation of these studies is that, with the exception of Shen (2003), none account for transaction costs or report results net of such costs. Shen demonstrated that the market timing strategy based on the spread between market's E/P ratio and interest rates outperforms the market even after a reasonable transaction cost per trade is accounted for. However, his statistical tests show that the outperformance, net of transaction costs, is not statistically significant.

A second stream of studies on market timing strategies have relied on technical indicators to generate trading signals. These studies include Han *et al.* (2013), Siegel (2014), Feldman *et al.* (2015), and Glabadanidis (2016). They all examined signals based on a simple moving average (SMA) of past market returns. The SMA time windows examined in these studies varied from 5 to 200 days and a trading signal was triggered whenever a crossover occurred between the portfolio's close price and its SMA. Although these SMA-based timing strategies were found to outperform the buy-and-hold benchmark portfolio before transaction costs, the evidence on outperformance after these costs were included is not robust. Siegel (2014) applied a 1% band over the 200-day moving average of DJIA to minimize the number of trades, but net excess returns remained barely positive or became negative in different sub-periods. Feldman *et al.* (2015) did not report excess returns net of transaction costs for their timing strategies. Han *et al.* (2013) and Glabadanidis (2016) estimated the Breakeven Transaction Costs (BETC) per trade; that is, the minimum level of transaction cost per trade at which the SMA strategies' average returns equal those of the factor-based decile benchmark portfolios. The BETC estimates range from 0.29% to 0.64% per trade in Han *et al.* (2013) and from 0.00% to 0.15% in Glabadanidis (2016). These ranges of BETC appear to be around or below the level of 0.5% transaction costs assumed in Siegel (2014), implying that these strategies' outperformance was not robust after considering appropriate transaction costs. One of the principal reasons for the lack of consistent outperformance of SMA-based strategies over buy-and-hold returns is that these strategies typically generate a considerable number of in and out trading signals and thus, even a reasonable transaction cost per trade eliminates or substantially reduces any potential outperformance.

In this paper, we propose a simple market timing strategy that does not rely on either fundamental factors or SMA-based technical indicators. To the best of our knowledge, our study is the only one to demonstrate robust and statistically significant outperformance over the benchmark even after accounting for reasonable transaction costs. The inspiration for our paper comes from the body of studies which have shown avoiding the worst returns in down markets greatly improves the long-term performance of a buy-and-hold portfolio. The key driver of our strategy's outperformance is the fairly reliable timing signal that emanates from the existence of a 'fat' left tail (negative skewness and excess kurtosis) of market returns, which has been well-documented in the academic literature (see, for example, Cont, 2001). Estrada (2008) examined 15 international equity markets, including the U.S., and demonstrated that, with the exception of Thailand, they all exhibit significant negative skewness and excess kurtosis in returns. Therefore, avoiding the worst returns contributes far more than the diminution in performance by missing some of the best returns when staying out of the market.

3 The Dataset

Our dataset consists of two variables: the S&P 500 total return index and the three-month U.S. Treasury bill rates. The S&P 500 total return index reflects the returns from both capital gains and dividend payments of the underlying large-cap stocks. This is a better proxy for the buy-and-hold portfolio than the commonly-quoted S&P 500 index because any evaluation of a market timing strategy should take into account the loss of dividend earnings when an investor is out of the market. As the S&P 500 total return index was not launched until January 4, 1988, we estimated the index's total returns prior to that date by making the necessary adjustments to the index for dividend payments. In particular, we downloaded the data on monthly dividend payments per share for the S&P 500 index for the period 1936–1987 and then recreated the S&P 500 total return index by capturing both capital gains and dividends in total returns for the years prior to 1988.[2] Our data on the S&P 500 total return index contains dividend-adjusted monthly index values from January 1928 to March 2020. These index values are used to compute the monthly logarithmic returns, which we use throughout this paper and simply call it 'market return.' Data for the S&P 500 total return index and the dividend stream were sourced from Bloomberg.

We obtained the monthly data on the three-month US Treasury bill rates from St. Louis Fed for the period of January 1934 to March 2020. To fill in the rates for the years 1928 through 1933, we assumed that the monthly T-bill rate during these years to be the same as the average T-bill rate in the 12 months of 1934.

4 Importance of Avoiding Large Down Months

Trying to time the market is futile, buy-and-hold proponents argue, because it often leads to missing out on the best days of market gains. Empirical evidence does support this proposition. Our examination of monthly market returns since 1928 shows that missing only the best 10% of all monthly returns (i.e., eliminating all returns above the 90th percentile) reduces the average annual return of the S&P 500 total return index to −3.1%, which is 11.3% lower than the average buy-and-hold (B&H) return of the index of 8.2%. However, this is only half the story. Table 1 shows that avoiding the worst monthly returns (i.e., returns below the bottom

Table 1 Return and Balance Under Alternative Scenarios.

		Starting balance $100,000	
	Average rate of return (%)	Balance after 40 years	Gain/loss
Buy-and-hold throughout	8.2%	$2,622,755	$2,522,755
Miss best 10% of all months	−3.1%	$29,007	($70,993)
Avoid worst 10% of all months	22.2%	$705,319,793	$705,219,793

Note: The analysis uses S&P 500 Total Return data from 1928 through March 2020.

10th percentile) increases the index's return to 22.2%, which is 14% higher than the average B&H return. The consequences of these rates of returns compounded over a 40-year time horizon and a starting balance of $100,000 are also shown in Table 1. These figures underscore the importance of being able to avoid the worst returns; the benefit of averting the worst returns substantially outweighs the lost opportunity of missing out the best ones.

In the above table, the best or the worst 10% returns are based on the data for the entire sample period of 1928–2020. However, a more realistic approach would be to examine the time-varying 10th percentiles in defining the worst months *without* the benefit of hindsight. For example, in January 1950, the cut-off to define the worst months (i.e., the bottom 10th percentile) could only be based on the data up through December 1949. More generally, at any time τ, we look back

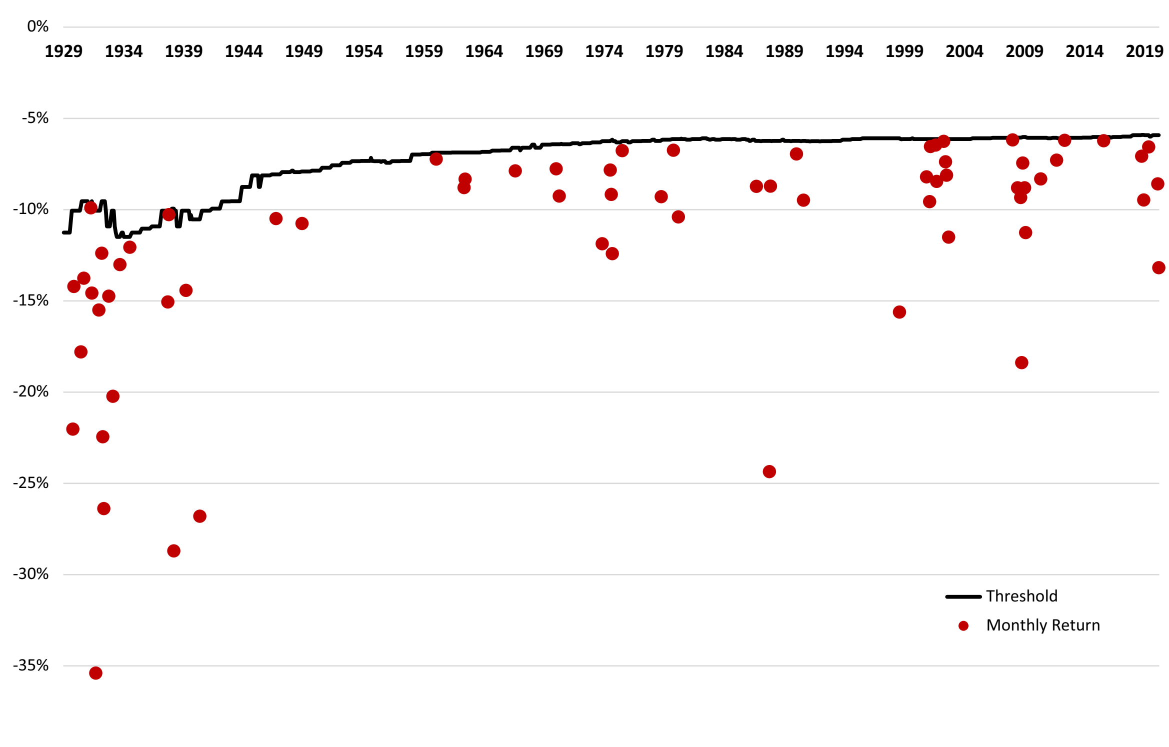

Figure 1 Time-varying monthly thresholds and return on large down months.

at all the data from 1928 through time τ in creating time-varying thresholds that define the cutoffs for the worst months. By doing this rather than over a rolling window, we also eliminate the need to choose the length of the window, as is done in SMA strategies. Figure 1 shows our estimated time-varying thresholds and the red dots denote the monthly returns below their corresponding thresholds, that is, they are the returns during the 'large down months.'[3]

The frequency of these large down months and the market returns during these months, over various 10-year intervals starting from 1928, are shown in Table 2. The number of large down months in the 10-year periods ranges from one (during 1948–1957) to 17, with the latter residing in the 10-year period 1928–1937, which spanned the Great Depression. There is a total of 67 large down months from 1928 to March 2020.

Table 2 Frequency and Average Return During Large Down Months.

	Interval		No. of large down months	Average return in large down months
1	1928	1937	17	−17.0%
2	1938	1947	4	−20.1%
3	1948	1957	1	−10.8%
4	1958	1967	4	−8.1%
5	1968	1977	7	−9.3%
6	1978	1987	6	−11.4%
7	1988	1997	2	−8.2%
8	1998	2007	10	−8.8%
9	2008	2017	11	−8.9%
10	2018	2020	5	−9.0%
			Average	−11.2%
Total Number of Months (Jan 1928–Mar 2020):				1,107
Total Number of Large Down Months:				67
% of Large Down Months:				6.1%

Two facets of the findings shown in Table 2 are noteworthy. First, the average return during large down months is considerably worse (i.e., larger in absolute value) than the bottom 10th percentile of returns for the entire sample period (which is −5.2%). In other words, there are many months with much worse returns than the thresholds, a fact that is also evident from the red dots in Figure 1. Second, the number of large down months relative to the total number of months is fairly low (only 6.1%). Both these market return characteristics enable us to design a strategy, which aims at avoiding the large down months while keeping the number of trades and their associated transaction costs under control. The strategy and its implementation are discussed next.

5 Implementation and Results of the In and Out Strategy

In this section, we first describe the proposed market timing strategy, which involves coming in or out of the market based on entry or exit signals. We then discuss the results and the various robustness checks that we undertook to ensure that the strategy's outperformance over a B&H portfolio holds under a variety of market environments and across different time periods.

5.1 The description of the in and out strategy

Our in and out market timing strategy is simple, rule-based, and long-only.[4] It does not rely on sophisticated algorithms or optimized parameters; nor does it suffer from any hindsight bias. At any point in time, to decide whether to exit the market (or to re-enter) we rely on the information available only at that point in time. The objective of this strategy is to generate better returns than the benchmark B&H portfolio's (i.e., S&P 500 total returns index), on both absolute and risk-adjusted bases.

The strategy's portfolio begins by allocating 100% capital to the B&H portfolio. The strategy fully sells out of this portfolio and invests the proceeds in three-month Treasury bills if and when the sell signal is triggered. At the beginning of each month, we determine whether the sell signal is triggered if the *preceding* month's market return is below a time-varying threshold.[5] Specifically:

$$\begin{aligned} S_t &= 1 \quad \text{if } Rm_{t-1} < Th_{t-1}; \\ S_t &= 0 \quad \text{otherwise} \end{aligned} \tag{1}$$

where S_t takes a value of zero or one, with one denoting a sell signal and zero a not-sell signal; Rm_{t-1} denotes the market's return in month $t-1$, and Th_{t-1} denotes the threshold value in month $t-1$. The time-varying threshold is defined as: $Th_t = -P(90)_t$, where $P(90)_t$ denotes the 90th percentile of market returns over the time period spanning the interval $[0, t]$, with time zero being January of 1928.

Thus, at the beginning of each month, the threshold value is updated based on the data for the 90th percentile of market returns *up through the end of the preceding month*.[6] If the threshold is not triggered and the sell signal is turned off, the portfolio either switches back to or stays fully invested in the market, depending on whether one was invested in Treasury bills or the market in the preceding month. Conversely, if the sell signal is turned on, the portfolio either switches out of the market or stays in Treasury bills.

Recalling from Equation (1) that S_t takes a value of one when invested in Treasury bills and a value of zero when staying invested in the market, the monthly return of the in and out strategy's portfolio, denoted by Rp_t, can be expressed as:

$$Rp_t = S_t \cdot Rf_t + (1 - S_t) \cdot Rm_t \tag{2}$$

where Rf_t denotes the monthly risk-free rate (i.e., three-month T-bill rate). In the remainder of the paper, we will use the annualized versions of Rp_t, defined above, to measure the in and out strategy's performance.

Table 3 Performance of Buy-and-Hold and In and Out Strategies Jan 1928–March 2020.

	Buy and hold	In and out
Annualized return	8.2%	9.2%
Annualized sigma	18.7%	17.2%
Sharpe ratio	0.26	0.34
Sortino ratio	0.37	0.51

5.2 The results

In Table 3 below, we report the raw and risk-adjusted return of the in and out strategy, and the benchmark B&H portfolio.

Table 3 shows that the strategy outperforms the B&H portfolio in each of the four risk–return performance measures over the sample period, 1928 through March 2020. On average, the in and out strategy outperforms the B&H portfolio by 1.0% per year and lowers the portfolio's yearly volatility (annualized standard deviation) by 1.5%. As a result, the in and out strategy has higher Sharpe and Sortino ratios than the B&H portfolio.[7]

5.3 Robustness checks

To evaluate the robustness of the outperformance of the proposed strategy, we undertake three different analyses: (a) relative performance with different start years; (b) relative performance with randomly chosen contiguous time periods of varying lengths using a Monte Carlo simulation; and (c) regression analysis of the strategy's returns to estimate their alpha and the 'up' and 'down' betas.

To verify the robustness of our results, we examined the performance metrics of the two strategies over a shorter sample period by excluding the first

10 years. The new time-varying thresholds were computed using data strictly from this shorter time period. The key results presented in Table 3 continue to hold using the data from 1938 to 2020: the in and out strategy yields higher annualized return, lower volatility, and higher Sharpe and Sortino ratios than the B&H portfolio. We adopted the same robustness check of our results also through a Monte Carlo analysis using both sample periods, 1928–2020 and 1938–2020. The detailed results are shown in Table 5.

In Table 4, we compare the relative performance of B&H and the in and out strategy across different sub-periods, generated by restricting the dataset to different start years, with five-year increments starting in 1928.[8] The in and out strategy outperforms the benchmark in 17 out of 18 periods, with the annualized outperformance ranging from 0.3% in 1933–2020 to 1.9% in 2008–2020. The only time period during which the in and out strategy does not outperform the B&H benchmark is the last subperiod, 2013–2020; however, the underperformance is miniscule, less than 0.035%. This is expected from the nature of the in and out strategy which is designed to avoid large down months; consequently, the in and out strategy is expected to converge to the B&H returns in prolonged bull markets, as was largely the case during the years 2013–2020.

As a second robustness check of the strategy's outperformance, we conducted a Monte Carlo analysis, in which we randomly chose start and end points, creating investment periods of random

Table 4 Relative Performance of Buy and Hold vs In and Out for Different Start Years.

Start year	End year (through Mar)	Annualized return: Buy and Hold	Annualized return: In and Out	Difference	Beat	No. of out months
1928	2020	8.2%	9.2%	1.0%	1	66
1933	2020	9.6%	9.9%	0.3%	1	54
1938	2020	9.5%	9.6%	0.1%	1	49
1943	2020	9.9%	10.3%	0.4%	1	46
1948	2020	9.8%	10.2%	0.4%	1	45
1953	2020	9.3%	9.8%	0.5%	1	44
1958	2020	9.1%	9.6%	0.5%	1	44
1963	2020	8.9%	9.5%	0.5%	1	41
1968	2020	8.8%	9.3%	0.6%	1	40
1973	2020	9.0%	9.6%	0.6%	1	38
1978	2020	10.3%	10.7%	0.4%	1	33
1983	2020	10.1%	10.8%	0.7%	1	30
1988	2020	9.5%	10.4%	0.9%	1	27
1993	2020	8.5%	9.4%	0.9%	1	25
1998	2020	6.3%	7.4%	1.1%	1	25
2003	2020	8.3%	9.6%	1.3%	1	15
2008	2020	6.7%	8.6%	1.9%	1	15
2013	2020	10.3%	10.2%	0.0%	0	5
			Average	**0.67%**	**94.4%**	

Table 5 Monte Carlo Analysis Using Random Time Periods.

	Panel A: Jan 1928–Mar 2020		Panel B: Jan 1938–Mar 2020	
	Buy and Hold	In and Out	Buy and Hold	In and Out
Annualized return	9.0%	9.6%	9.2%	9.8%
Annualized sigma	15.2%	13.8%	14.6%	13.0%
Sharpe ratio	0.38	0.46	39.5%	48.9%
Sortino ratio	0.57	0.71	59.1%	75.3%
Annualized Alpha		1.5%		1.7%
%Beat Return		81.9%		83.7%
%Beat Sigma		99.8%		100.0%
%Beat Sharpe		93.0%		94.0%
%Beat Sortino		92.9%		94.3%
%Positive alpha		95.3%		97.3%

Note: In the Monte Carlo simulation contiguous periods of time, with random lengths from 61 months to 1107 months, are chosen 50,000 times.

lengths spanning different time periods between January 1928 and March 2020.[9] Specifically, we simulated 50,000 investment windows ranging in length between 61 months and 1,107 months (which is the maximum number of months in our sample), with the starting and ending points randomly selected from month 1 (i.e., January 1928) to month 1,047 (i.e., March 2015), and ensuring that the end month is at least 60 months after the start month. Within each of the 50,000 investment windows, we computed the performance metrics for the in and out and B&H strategies.

The results of the Monte Carlo simulation are presented in Table 5. It shows that the in and out strategy, on average, beats the B&H benchmark returns by 0.6%, during both sample periods, starting from 1928 and from 1938. Table 5 also shows that the in and out strategy outperforms the benchmark in more than 80% of the 50,000 randomly selected investment periods.

For each random investment period, we also computed the strategy's 'alpha' using the capital asset pricing model (CAPM) equation:

$$Rp_t - Rf_t = \alpha + a_1(Rm_t - Rf_t) + \varepsilon_t \quad (3)$$

where the variables have been defined in Equation (2), and ε_t is the residual term. Table 5 shows, in over 95% of the randomly chosen investment horizons, the strategy generates a positive alpha, and the annualized average alpha across 50,000 iterations is at least 1.5%. Table 5 also demonstrates the superior risk-adjusted performance of the strategy, as reflected in higher Sharpe and Sortino ratios. Finally, statistical tests confirm that all performance metrics (annualized return, Sharpe and Sortino ratios) are statistically significantly higher (with higher than 99% confidence) for the in and out strategy than the B&H portfolio.[10] The robust results from our Monte Carlo analysis indicate that the in and out strategy will likely outperform the B&H portfolio, on average, across a wide range of market environments.

For our final robustness check, we applied the parametric test proposed by Henriksson and

Merton (1981) to evaluate whether the strategy's market timing ability is superior and statistically significant. The regression model for this parametric test is specified as follows:

$$Rp_t - Rf_t = \beta_0 + \beta_1 \cdot \max(0, Rm_t - Rf_t) + \beta_2 \cdot \min(0, Rm_t - Rf_t) + \varepsilon_t \quad (4)$$

with the variables being the same as defined in Equations (2) and (3). Henriksson and Merton (1981) shows that successful market timing requires the difference between the 'up-market' beta, β_1, and the 'down- market beta', β_2, to be positive and statistically significant.

Table 6 Regression Analysis of Strategy's Performance.

Start year	End year (through Mar)	Difference between β_1 and β_2	Annualized α
1928	2020	0.15**	1.8%**
1933	2020	0.06*	1.0%*
1938	2020	0.01	0.8%
1943	2020	0.12**	1.2%**
1948	2020	0.12**	1.2%*
1953	2020	0.12**	1.3%*
1958	2020	0.11**	1.3%*
1963	2020	0.11*	1.3%*
1968	2020	0.11*	1.3%*
1973	2020	0.11*	1.4%*
1978	2020	0.13*	1.4%*
1983	2020	0.13*	2.0%**
1988	2020	0.19**	2.3%**
1993	2020	0.17*	2.4%*
1998	2020	0.15	2.2%*
2003	2020	0.30**	3.5%**
2008	2020	0.27*	4.0%*
2013	2020	0.16	2.3%
	Average	**0.14**	**1.8%**

Notes: The β_1 and β_2 coefficients are from the regression equation set out in Equation (4).
The α coefficient is from the regression equation set out in Equation (3).
**Denotes statistically significant at 99% confidence level.
*Denotes statistically significant at 95% confidence level.

Table 6 presents the results from the Henriksson and Merton market timing test applied to the full sample (the first row) and for the various subperiods. Over the entire sample, the difference between the 'up' and 'down' beta is statistically significant at the 99% level of confidence and so is the annualized alpha of 1.8% (estimated using Equation (3)).

In 17 out of the 17 subperiods, with the starting point moving up in five-year increments from 1933, the difference between the betas is positive and is statistically significant in 14 of the 17 subperiods. Furthermore, in 15 of the 17 subperiods the estimated annualized alpha is statistically significant, and they are all positive, ranging between 0.8% and 4.0%. Table 6 further demonstrates the strategy's successful market timing ability and its outperformance relative to the B&H benchmark.

6 Explanation of the Results

In this section, we provide explanations for the in and out strategy's outperformance over a B&H portfolio. The two key reasons are non-normality and persistence of market returns. However, before discussing these two features, we examine some simple statistics on the correctness of the exit and entry signals generated by the strategy.

6.1 Correct and incorrect exit and entry signals

In the preceding sections we have already discussed the results of the Henriksson and Merton market timing test for the strategy. Here, we discuss an alternative way to examine the performance impact of correct and incorrect signals generated by the strategy.

We characterize an exit signal to be correct if the strategy exits the market in any given month *and* the market return for that month is negative; and thus, that month's negative return is the loss avoided by the correct exit signal. Conversely, we characterize an exit signal to be incorrect if the strategy exits the market in any given month *and* the market return for that month is positive. That month's positive return is the gain missed due to the incorrect exit signal. As noted in Table 4, for the entire sample, there are 66 months in which the strategy stays out of the market. Of these, exactly 50% are correct signals. However, the performance impact of these signals (that is, the loss avoided and the gain missed) is not symmetric. We summarize below the gain and loss due to the 33 correct and 33 incorrect exit signals:

Cumulative gain (i.e., market loss avoided) due to correct exit signal: 262.3% Cumulative loss (i.e., market gain missed) due to incorrect exit signal: −182.6%.

The difference is 79.8%, which translates to an average gain of 1.2% across the total of 66 months during which the strategy stays out of the market. This explains the strategy's outperformance over a B&H portfolio.

We now turn to the entry signals. There are exactly 50 months in which the strategy exits Treasury bills and re-enters the market. The total number of entry months is fewer than the number of months that the strategy stays out of the market because in many cases the strategy stays out of the market in two or more consecutive months.[11]

There are 31 correct (the strategy entered the market and the return that month was positive) and 19 incorrect (the strategy entered the market and the return that month was negative) entry signals. We summarize below the gain and loss due to the 31 correct and 19 incorrect entry signals:

Cumulative gain due to correct entry signal: 178.9% Cumulative loss due to incorrect entry signal: −84.2%.

The difference is 94.8%, which translates to an average gain of 1.9% across the total of 50 months during which the strategy re-enters enters the market.

The foregoing analysis shows that the proposed in and out strategy outperforms a B&H portfolio because the losses avoided far outweighs the gains missed by staying out of the market. This asymmetry is explained by the characteristics of the distribution of market returns, discussed next.

6.2 Non-normality of market returns

In Table 7, we present the results of the analysis on the characteristics of the market returns. In Panel A of this table, we apply the standard tests for skewness and kurtosis on the market returns data.[12]

The null hypotheses of zero skewness and zero excess kurtosis are rejected in the entire sample and in every subperiod. Market returns are found to be highly negatively skewed (i.e., the distribution of returns have longer left tails) and the tails are significantly fatter than that of the normal distribution (since the kurtosis is significantly larger than 3). Not surprisingly, the null hypothesis of normality[13] is strongly rejected for the entire sample and for each of the 17 subperiods.

These findings show that market returns are characterized by the presence of very large (in absolute value) negative returns; and they are bigger in size than large positive ones, as evidenced by the negative skewness. This observation is confirmed by the results shown in Panel B of Table 7. Here we report the actual annualized average of market returns for the entire sample and for each of the

Table 7 Test of Normality of Returns and Tail-Trimmed Annualized Average Return.

Panel A: Normality tests					Panel B: Untrimmed and trimmed means		
					Annualized mean return of market		
Start year	End year (through Mar)	Skewness	Kurtosis	Test of normality *P*-Value	Actual mean	2.5% and 97.5% Tail-trimmed mean	Difference
1928	2020	−0.62**	10.35**	0.0000	8.2%	9.5%	1.3%
1933	2020	−0.57**	8.19**	0.0000	9.6%	10.6%	1.0%
1938	2020	−0.89**	8.00**	0.0000	9.5%	10.5%	1.0%
1943	2020	−0.68**	5.24**	0.0000	9.9%	10.9%	1.0%
1948	2020	−0.68**	5.34**	0.0000	9.8%	10.7%	0.9%
1953	2020	−0.68**	5.44**	0.0000	9.3%	10.2%	0.9%
1958	2020	−0.73**	5.59**	0.0000	9.1%	10.1%	0.9%
1963	2020	−0.72**	5.64**	0.0000	8.9%	9.9%	1.0%
1968	2020	−0.71**	5.51**	0.0000	8.8%	9.8%	1.0%
1973	2020	−0.75**	5.69**	0.0000	9.0%	10.1%	1.1%
1978	2020	−0.93**	6.04**	0.0000	10.3%	11.6%	1.3%
1983	2020	−1.05**	6.52**	0.0000	10.1%	11.9%	1.8%
1988	2020	−0.83**	4.76**	0.0000	9.5%	11.1%	1.6%
1993	2020	−0.92**	4.84**	0.0000	8.5%	10.1%	1.6%
1998	2020	−0.89**	4.62**	0.0000	6.3%	8.3%	2.0%
2003	2020	−1.07**	5.76**	0.0000	8.3%	10.3%	2.0%
2008	2020	−1.02**	4.91**	0.0000	6.7%	8.5%	1.8%
2013	2020	−1.16**	5.34**	0.0001	10.3%	12.3%	2.0%
	Average	**−0.83**	**5.99**		**9.0%**	**10.3%**	**1.3%**

Notes: **Denotes statistically significant at 99% confidence level.
*Denotes statistically significant at 95% confidence level.

17 subperiods. We then recomputed the means by removing the data points that reside in the two tails of the distribution of market returns; specifically, we remove the returns below the 2.5th percentile and above the 97.5th percentile and then compute the mean of these tail-trimmed returns. Thus, this trimmed dataset has 5% fewer observations than the untrimmed one. For the entire sample, and for each of the 17 subperiods, the tail-trimmed means are considerably higher than the actual (untrimmed) means. In constructing the tail-trimmed mean, although we are removing equal number of data points from the right and left tails of the returns distribution, the returns that reside in the far left tail (i.e., below the 2.5th percentile) are much larger in absolute value than the returns in the far right tail; as a result, the trimmed means are higher than the untrimmed ones. This important market return characteristic is a key reason why the proposed strategy works: the gain from the avoided large negative returns far outweighs the loss from missed positive ones while staying out of the market. However, there is another important reason for the strategy's outperformance, which we explore next.

Table 8 Market Return on Large Move Months and the Next Month.

	Average monthly return	
	Large down	Large up
Month of	−11.9%	10.2%
Next month	−2.8%	−0.9%
Average monthly return of all months	0.68%	

Notes: A large down month is defined as one with worse than 10th percentile return; a large up month is defined as one with better than 90th percentile return.

In Panel B, the annualized mean returns are computed without tails being trimmed and with the tails trimmed, i.e., observations below 2.5th percentile and above 97.5th percentile are removed.

6.3 *Persistence of market returns*

Table 8 shows that a month with large (in absolute value) negative returns is often followed by a month of negative returns; but that is not the case for months with large positive returns. This asymmetry of the persistence of market returns is an important contributing factor to the strategy's outperformance. If a month with a large negative return was more often than not followed by a positive return month, then the strategy would have failed, because the strategy's exit signal is generated by observing a large negative return in the preceding month.

To confirm the observation summarized in Table 8, we conducted a regression analysis to test whether the persistence of positive or negative returns is statistically significant. Table 9 shows the estimated coefficients from the regression of the current month's return on the subsequent month return. Three different regressions are estimated using data: (a) for all months; (b) for only the subset of months with negative returns and the subsequent months; and (c) for only the subset of months with positive returns and the subsequent months. In the first regression, the positive and statistically significant coefficient for the next-month return variable confirms the persistence (i.e., positive autocorrelation) of market returns that has been documented in many prior studies.[14] However, the estimated coefficient from the second and third regressions tells an interesting story and confirms what we observed in Table 8. The positive and significant coefficient for the following-month return when the dataset is restricted to months with negative returns (and the subsequent months) confirms the high likelihood for a negative-return month to be followed by another negative one. By contrast, when the same regression is run with the sample restricted to months with positive returns (and the subsequent months), the coefficient is not significantly different from zero (and is also much smaller).

Table 9 Regression Analysis of Monthly Return on One-Month Ahead Return.

Sample restricted to	Coeff. estimate (γ_1)	*T*-Stat
All months	0.085	2.86
Only negative and month after	0.165	3.03
Only positive and month after	0.064	1.90

Notes: Estimated Equation: $Rm_t = \gamma_0 + \gamma_1 \cdot Rm_{t+1} + \epsilon_\tau$
Rm_t denotes market return of month t

Taken together, the results in Tables 7–9 support the two key reasons for the success of the in and out strategy: non-normality and asymmetric persistence of market returns.

7 Transaction Costs

In this section, we examine the impact of transaction costs on the strategy's performance. Needless to say, a strategy's performance net of transaction costs not only depends on the size of transaction costs per trade, but also on the number of in and out trades. As we demonstrated earlier, our proposed strategy involves a fairly low percentage of in and out signals, which as we discuss next, explains its success even after transactions costs are accounted for.

Transaction costs for any market timing strategy, including ours, include brokerage and other fees for each transaction, bid ask spreads, and capital gains taxes when one sells the portfolio. Furthermore, when an investor buys or sells a large portfolio, transaction costs may also involve the market impact of buying or selling large positions, which has been characterized in the academic literature as execution cost[15] or 'slippage.'

Taxes are unlikely to be a large component of transaction costs in evaluating the strategy's relative performance versus a B&H portfolio. This is because: (a) capital gains taxes will be borne only for gains when unwinding the portfolio and capital losses can be carried forward to partially or full offset future capital gains taxes; (b) the holding period, i.e., when the investor is fully invested in the market is often more than a year,[16] which means the capital gains would be taxed at a lower long-term rate; and (c) a B&H portfolio will also incur taxes whenever a distribution is made or when the portfolio is liquidated at the end of the investment horizon.

Consistent with several prior studies, we will first assume a fixed transaction cost of 0.5% per trade[17]; that is, the portfolio value is reduced by a hundred basis points for each round-trip trade. However, we believe a fixed transaction cost of 0.5% over a 92-year period is not realistic. To reflect the recent downward trend in transaction costs associated with decimalization, electronic trading, and more recently commissions-free

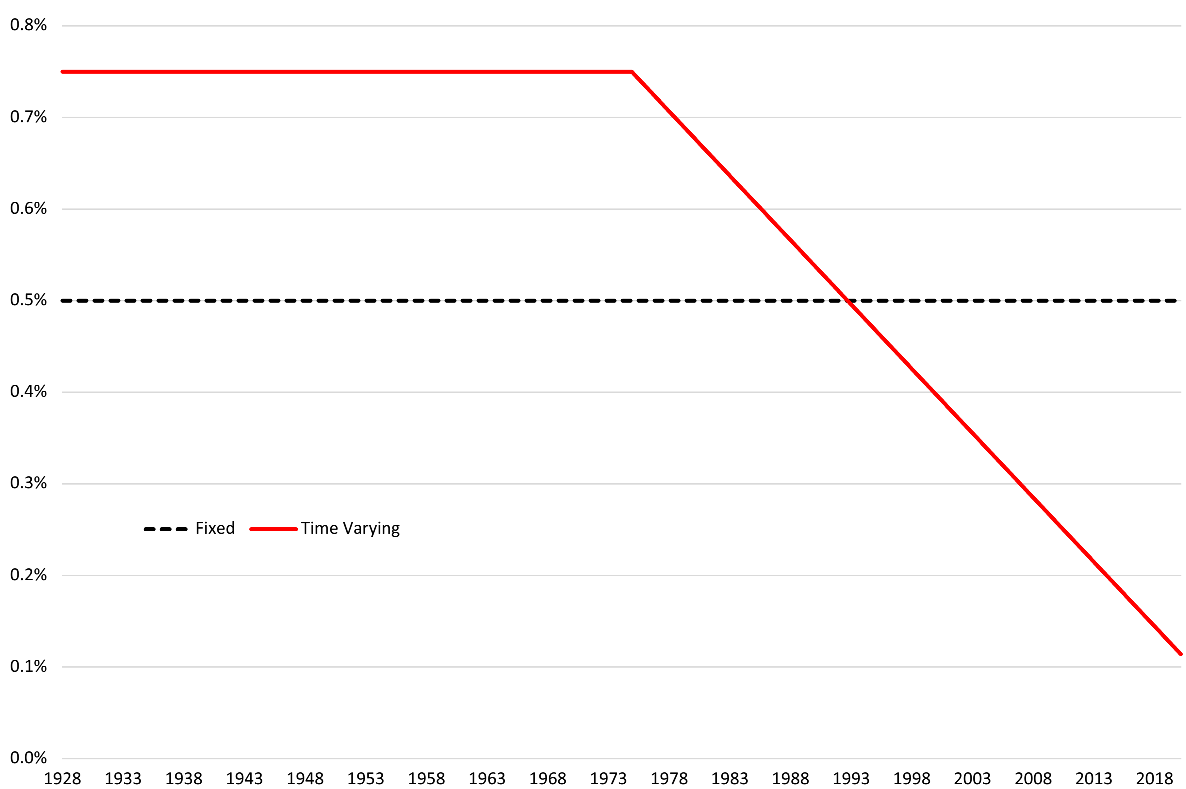

Figure 2 Fixed and time-varying transaction costs per trade.

Table 10 Relative Performance without and with Transaction Costs.

	Buy and Hold	In and Out strategy		
		Transaction cost:		
		None	Fixed 0.5%	Time-varying
Annualized return	8.17%	9.20%	8.65%	8.60%
Annualized sigma	18.71%	17.20%	17.20%	17.20%
Sharpe ratio	0.262	0.345	0.313	0.310
Sortino ratio	0.375	0.510	0.462	0.457
Breakeven Transaction Cost Per Trade:			0.94%	

trades, we have also assumed an alternative of time-varying transaction costs. Figure 2 shows that the time-varying transaction cost per trade takes the value of 0.75% until 1974, the mid-point of our sample period, and trends linearly downwards towards 0.1% per trade in 2020.[18] Over the entire sample period, the time-varying transaction costs average to approximately 0.6% per trade, which is slightly higher than the fixed amount of 0.5% per trade. In addition, we estimated a breakeven transaction cost at which an investor is indifferent between our proposed strategy and a B&H portfolio, since the average returns are equal. This is conservative; a risk-averse investor might prefer the proposed market timing strategy even though its return is the same as B&H because of the lower volatility of returns.

Table 10 presents the relative performance without and with transaction costs for the in and out strategy over the B&H portfolio. After accounting for a reasonable transaction cost, either fixed or time-varying, the in and out strategy continues to produce higher absolute and risk-adjusted returns than the B&H benchmark. We also find that the breakeven transaction cost is 0.94% per trade, or 1.98% per roundtrip trade.

As a robustness check for our strategy with transaction costs, we undertook another Monte-Carlo simulation analysis similar to the one for Table 5. Our analysis confirmed the absolute and risk-adjusted outperformance of our strategy with time-varying transaction costs. Furthermore, in 85% of the 50,000 randomly chosen investment horizons, the strategy yielded a positive alpha after accounting for transaction costs.

In Table 11, we report the results of the Henriksson and Merton (1981) test for successful market timing and the estimated annualized alpha with and without transaction costs. As is evident from the results in Table 11, the strategy passes the Henriksson and Merton test and produces a statistically significant alpha[19] even after transaction costs are accounted for.

Lastly, $1 invested in a portfolio starting in January 1928 would grow to a value of $1,877 as of March 2020 based on the returns of a B&H portfolio; but it would grow to $2,926, and $4,849 based on the proposed strategy's net[20] and gross returns, respectively. The cumulative growth of the investment portfolios is depicted in Figure 3 on a logarithm scale.

8 Concluding Comments

In this paper, we have proposed a simple, hindsight-free, rule-based method of entry and exit into the equity market. Using data from

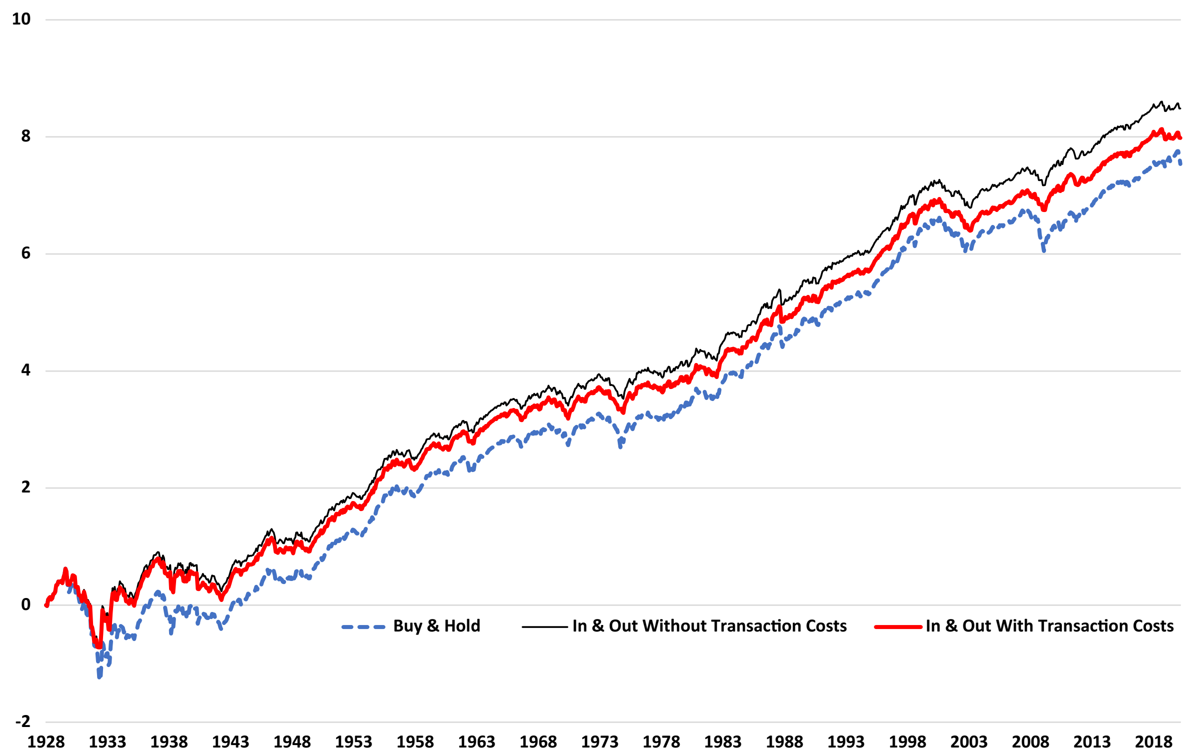

Figure 3 Growth of balances over time: 1928–2020 logarithmic scale.

Table 11 Regression Analysis of Strategy's Outperformance without and with Transaction Costs.

	Difference between β_1 and β_2	Annualized α
Without transaction costs	0.154**	1.78%**
With transaction costs (fixed 0.5%)	0.143**	1.22%*
With transaction costs (time-varying)	0.142**	1.17%*

Notes: The β_1 and β_2 coefficients are from the regression equation set out in Equation (4). The α coefficient is from the regression equation set out in Equation (3).
**Denotes statistically significant at 99% confidence level.
*Denotes statistically significant at 95% confidence level.

1928 through March 2020, we have demonstrated that the proposed market timing strategy delivers robust and statistically significant outperformance, net of transaction costs, over the S&P 500 total return index. These results are robust and hold up across various time periods characterized by different market environments.

The strategy's outperformance is explained by the characteristics of non-normality and asymmetric persistence of market returns, which will likely continue across different equity markets and over time. The recent market turbulence in the wake of the Covid-19 pandemic very well might increase investors' preference for timing strategies aimed at avoiding large losses and boosting long-term portfolio performance. We hope our proposed strategy, with its ease of execution and robust results, would address such a need.

A potential avenue of future research would be to explore whether a similar strategy, when applied to a diversified portfolio of equity and fixed income funds, would outperform its suitable benchmark.

Acknowledgment

The opinions expressed in this paper are of the authors and not of StoneTurn or any of its members. The paper is purely academic and is not intended to provide any form of investment advice. We are grateful to Burton Malkiel, Arthur Havenner, and the anonymous referee for their helpful comments on an earlier version of the paper. However, any errors are our own.

Endnotes

1 Fund flow data from Investment Company Institute; data sourced from Bloomberg.

2 For the period from 1928 to 1935, we assume that the monthly dividend return is the same as the average monthly dividend return in 1936. As a validation of our method of index construction, we verified that the dividend-adjusted S&P 500 index that we created exactly matched the data for the S&P 500 total return index in the years after January 4, 1988 when the index was launched.

3 For reasons explained later in this paper, the time-varying thresholds in Figure 1 are based on the time-varying, negative 90th percentile, rather than the 10th percentile.

4 The in and out strategy can be easily extended to a long–short strategy. Instead of investing in Treasury bills when the sell signal is triggered, the investor could go short the market until the buy signal is turned on.

5 To minimize any gap risk or look-ahead bias between signal generation and trade execution, the trading signal could be generated one trading day before the month-end execution. We analyzed this adjustment and it did not affect any of the paper's findings.

6 For the time-varying thresholds, we chose to use the negative of the 90th percentile, as opposed to the 10th percentile of returns, because the 10th percentile is sensitive to large negative returns (especially because we are using logarithmic returns); by contrast the 90th percentile has less variability over time. However, all key results of this paper remain unchanged if we were to use the 10th percentile as the threshold.

7 As is evident from Table 2, the first 10 years of our sample period (i.e., 1928–1937) were marked by many large down months, which could have a non-trivial effect on the time-varying thresholds, since these years are included in the threshold calculation as one moves forward in time.

8 The first row of Table 4 shows that total number of months during which the strategy stays out of the market is 66, over the entire sample period of 1928 through March 2020. However, in Table 2, we reported that there were 67 large down months; this discrepancy is explained by the fact that there is a one-month lag between each of the large down month (Table 2) and each of the corresponding out-of-the-market months (Table 4). Both February and March of 2020 were large down months; but the strategy (based on the returns from the prior month) stays out of the market only in March and not in February 2020.

9 We also adopted a variant of the Monte Carlo analysis by restricting the sample period to Jan 1938–Mar 2020. Our conclusions stay the same.

10 Further statistical tests of significance of the strategy's results are shown in Table 6.

11 Out of a total of 66 months, there are 39 months where the strategy stays out of the market only for a single month; there are 18 months where the strategy stays out of the market for two consecutive months; and there are nine months where the strategy stays out of the market for three consecutive months.

12 See D'Agostino *et al*. (1990).

13 See Jarque and Bera (1987).

14 See, for example, Poterba and Summers (1986); Campbell (1990).

15 See, for example, Lee and Ready (1991); Bessembinder (2003).

16 The average length of time the strategy remains fully invested in the market is 28.5 months.

17 See, for example, Siegel (2014).

18 Specifically, the time-varying transaction cost after 1974 takes the following form: $TC_t = \max(TC_{t-1} - \frac{TC_0 - 0.1\%}{N}, 0.1\%)$, where $TC_0 = 0.75\%$ and $N = 543$, the number of months from January 1975 to March 2020.

19 The null hypothesis of the test is that alpha is zero, against the alternative hypothesis that it is positive.

20 In Figure 3, net returns are based on the fixed transaction cost of 0.5% per trade, or 1% per round-trip trade.

References

Bessembinder, H. (2003). "Issues in Assessing Trade Execution Costs," *Journal of Financial Markets* **6**(3), 233–257.

Bogle, J. (1994). *Bogle on Mutual Funds: New Perspectives for the Intelligent Investor*, Wiley.

Campbell, J. Y. (1990). "Measuring the Persistence of Expected Returns," *American Economic Review* **80**(2), 43–47.

Cont, R. (2001). "Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues," *Quantitative Finance* **1**, 203–226.

Copeland, M. and Copeland, T. (1999). "Market Timing: Style and Size Rotation Using the VIX," *Financial Analysts Journal* **55**(2), 73–81.

D'Agostino, R. B., Belanger, A. J., and D'Agostino, Jr, R. B. (1990). "A Suggestion for Using Powerful and Informative Tests of Normality," *American Statistician* **44**(4), 316–321.

Glabadanidis, P. (2016). "Market Timing with Moving Averages," *International Review of Finance* **15**(3), 387–425.

Estrada, J. (2008). "Black Swans and Market Timing: How Not to Generate Alpha: How Not to Generate Alpha," *Journal of Investing* **17**(4), 20–34.

Feldman, T. Jung, A., and Klein, J. (2015). "Buy and Hold versus Timing Strategies: The Winner Is. . .," *Journal of Portfolio Management* **42**(1), 110–118.

Han, Y., Yang, K. and Zhou, G. (2013). "A New Anomaly: The Cross-sectional Profitability of Technical Analysis," *Journal of Financial and Quantitative Analysis* **48**(5), 1433–1461.

Henriksson, R. and Merton, R. (1981). "On Market Timing and Investor Performance. II. Statistical Procedures for Evaluating Forecasting Skills," *The Journal of Business* **54**(4), 513–533.

Jarque, C. M. and Bera, A. K. (1987). "A Test for Normality of Observations and Regression Residuals," *International Statistical Review* **55**(2), 163–172.

Lee, M. C. Charles and Ready, M. J. (1991). "Inferring Trade Direction from Intraday Data," *The Journal of Finance* **46**(2), 733–746.

Malkiel, B. G. (1996). *A Random Walk Down Wall Street*, W. W. Norton.

Poterba, J. M. and Summers, L. H. (1986). "The Persistence of Volatility and Stock Market Fluctuations," *The American Economic Review* **76**(5), 1142–1151.

Resnick, B. and Smith, G. (2002). "Using the Yield Curve to Time the Stock Market to Time the Stock Market," *Financial Analysts Journal* **58**(3), 82–90.

Sharpe, W. F. (2002), *Indexed Investing: A Prosaic Way to Beat the Average Investor*, Presented at the Spring President's Forum.

Shen, P. (2003). "Market Timing Strategies that Worked," *Journal of Portfolio Management* **29**(2), 57–68.

Shilling, A. (1992). "Market Timing: Better Than a Buy and Hold Strategy," *Financial Analysts Journal* **48**(2), 46–50.

Siegel, J. (2014). *Stocks for the Long Run: The Definitive Guide to Financial Market Returns & Long-Term Investment Strategies*, 5th ed. New York: McGraw-Hill.

Stowe, J. (2000). "A Market Timing Myth," *Journal of Investing* Winter 2000, **9**(4), 55–59.

Vergin, G. (1996). "Market Timing Strategies: Can You Get Rich?" *Journal of Investing* **5**(4), 79–86.

Keywords: Market timing; in and out strategy; loss aversion; buy and hold.