S&P500 Trading Strategy

2020 UCSB Hull Tactical ERP Contest

(Creative Category)

Michael Kwok (Team Lone Warrior)

3rd year Actuarial Science B.S., UCSB

Introduction

From experimenting with day-trading in my own time, I often found myself timing a stock's price movements, buying the stock when I think it'll go up within the next few minutes, and then watching the price drop moments later, which then prompted me to cut losses. However, the stock price would later come back up either during the day or within the next few days, and I would've been in-the-money had I just held the stock instead of selling it for a loss.

Such experiences led me to develop the research question: if I buy a stock at a *completely random time*, what are the chances that I'll *never* be in-the-money, i.e. what are the chances that I bought at the *highest price possible*? Naturally, the probability should be low given that stock prices generally fluctuate enough both intraday and inter-day to make it likely that you did not buy at the highest price possible, thereby allowing you to sell at a profit sooner or later. By answering this research question, I sought to profit off random market noise instead of actual market direction. Such a strategy could produce positive returns regardless of stock market trends and economic conditions.

With the goal of trading in mind, the next question was: "how *much* in-the-money should I be before I sell the stock?" I explored this idea with various stocks/ETFs by running statistical tests and back-tests in Python using Jupyter Notebooks. I discovered the rather obvious relationship: the higher you set your target in-the-money sell-price to be, the longer you'll have to wait until that goal is reached. Hence, there is a tradeoff between the size of your returns and the amount of time needed to wait on average before the stock can achieve those returns. Furthermore, as this target sell-price increases, the probability of buying and getting your money "stuck" out-of-the-money for a long time, if not forever, increases. If this scenario arises, would it be better to have a systematic way of cutting losses, or should this scenario be avoided at all costs to begin with?

For this contest, I applied my tests and findings to the daily historical price data of the S&P500 index, sometimes known as ticker symbol SPX. The initial results were mediocre, and positions that were taken right before the 2000 dotcom and 2008 housing bubbles were prone to getting stuck in the market for over a year. As a result, I incorporated a technical indicator, the simple moving average (SMA), into my strategy in hopes that it could operate as a signal for when to buy/sell positions without getting money stuck in the market for long periods of time. As the name implies, this indicator takes the arithmetic average of stock prices from a pre-specified trailing time frame, such as the past 180 days. After experimenting with the length of the SMA time frame as well as a target goal-sell, I eventually found a consistently profitable method of buying and selling the market. This report presents an original trading strategy that yields 32.7% annualized returns upon back-testing it on the S&P500 Index's (SPX) historical daily prices from 1962 to 2018, a period of 57 years. Implementation of the strategy in the real-world setting is discussed under "Implementation." Note that the strategy requires real-time monitoring of intraday prices, so it could not be used for the Live Trading Contest.

Parameters

The strategy's only parameter, the SMA, requires selection of the length of the trailing time frame from which to take the average on. Below, Equation 1 describes how to calculate the SMA that is used for the strategy. Note that the average is calculated using daily closing prices since the strategy is designed to trade at market open, so it makes sense to incorporate the most recent information possible (the previous day's closing price).

$$\text{SMA} = \frac{A_1 + A_2 + A_3 + \ldots + A_n}{n} \quad \textit{(Equation 1)}$$

where n = length of time frame (in days), A_i = ith day's closing price, i = # of days prior to today

When plotted as a line on a stock price chart, the SMA works as a lagging price-indicator that tracks the overall price movement without being heavily impacted by price noise. The SMA is updated every day, and its smoothness increases as the time frame increases in length.

For SPX, it was determined through testing that the optimal SMA time frame length is 1-day. So, for Equation 1, n=1. This SMA is just the previous day's closing price. However, the term "SMA" will be used throughout most of the report so that the strategy can generalize to other securities that may require longer SMA time frame settings to be optimal.

Trading Strategy

The strategy is simple:

1) Buy SPX* at market open if the opening price equals or exceeds the SMA
2) Sell SPX* when the price crosses below (or opens below, whichever comes first) the SMA.

*During back-testing, all transactions utilized the entire account balance on every trade to obtain the 32.7% annualized return since this strategy has statistically proven itself to be consistently profitable under nearly all market conditions (elaborated later in the report). However, in practice, one may of course choose not to trade their entire account balance if they are not comfortable with doing so.

There are two possible outcomes after buying:

1. If the sell-criteria is met on the same day of the buy, then SPX is sold as a day-trade and the result is a loss. This loss is the difference between the opening price and the SMA.
2. Otherwise, if the sell-criteria is *not* met on the same day of the buy, then the position is held until the sell-criteria is met. This will take place within the next few days either at market open (via an inter-day price gap-down) or during the day.

Essentially, the SMA acts as a stop-loss on the day of the buy. If the position lasts longer than a day, then the SMA becomes a trailing stop-sell that is updated daily at market close. Below are two visual examples of how the strategy can play out; each candlestick represents 1 day.

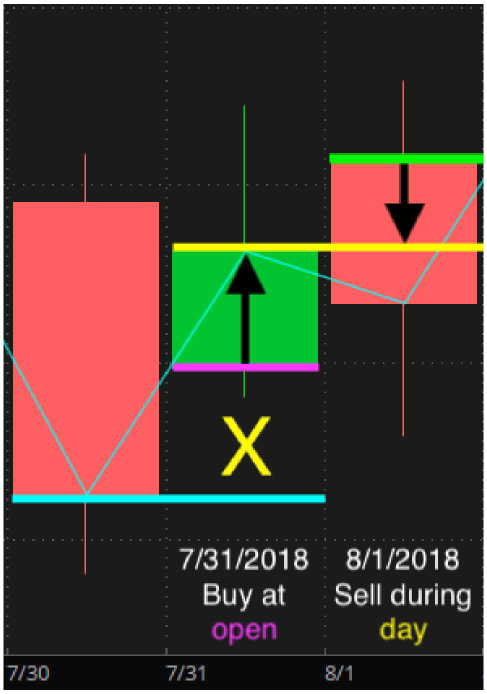

Figure 1: profitable trade

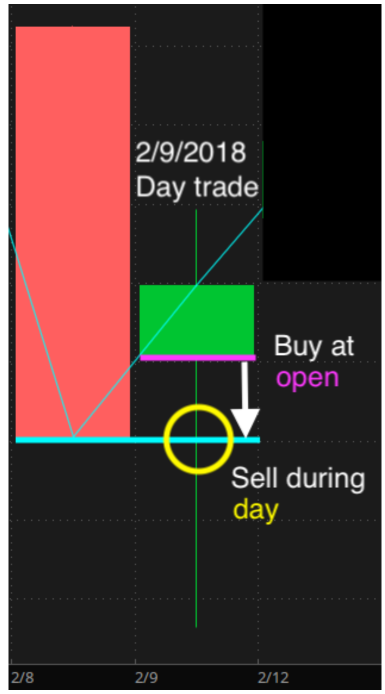

Figure 2: unprofitable trade

Figure 1

i) On 7/30/2018, the price closes at the blue line. This becomes the new SMA.

ii) On 7/31/2018, the strategy buys SPX at market open (magenta line) because the price is greater than the SMA (blue line). As indicated by the yellow "X", the price does *not* go below the SMA during the day, so the SMA stop-loss is not triggered. At market close, the SMA gets updated (yellow line), and the position is held until the next day (8/1/2018).

iii) On 8/1/2018, the price opens at the green line. During the day, the price goes below the updated SMA (yellow line). When it does so, the position is sold via the SMA stop-sell, and the profit is the difference between the yellow and magenta lines.

Figure 2

i) On 2/8/2018, the price closes at the blue line. This becomes the new SMA.

ii) On 2/9/2018, the strategy buys SPX at market open (magenta line) because the price is greater than the SMA (blue line). During the day, the price *does* go below the SMA (blue line). This is evident in the yellow circle that depicts where the price crosses below the SMA. As a result, the SMA stop-loss is triggered, and the position is sold on the spot. The loss is the difference between the blue and magenta lines.

Overall Returns

The graphs below compare the performances of the strategy and the S&P index.

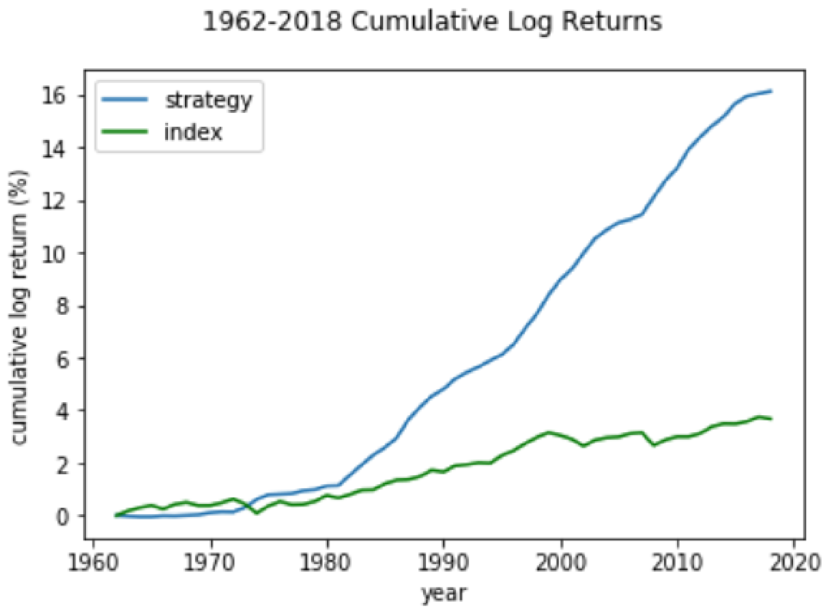

Figure 3: Comparison of cumulative log-returns (%) of index vs strategy

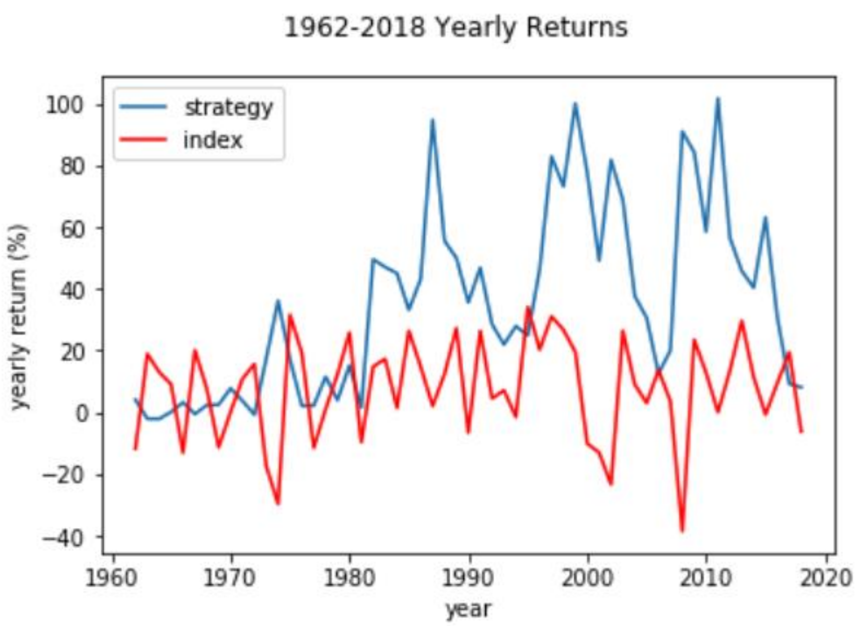

Figure 4: Comparison of yearly returns (%) of index vs strategy

Figure 3 compares the strategy's cumulative log-returns (blue) with the S&P index's cumulative historical log-returns (green) from 1962 to 2018 with the assumption that strategy's returns are re-invested during each trade. Clearly, the strategy outperforms the S&P index.

Figure 4 compares the strategy's yearly returns with the S&P index's yearly returns. The correlation coefficient is 0.00647, indicating that there is *no correlation* between the strategy's returns and the market's returns. Thus, the strategy performs well regardless of what the market's condition is. Note that the strategy has periods of particularly high returns (above 30%) in the years 1982-1991, 1996-2005, and 2008-2016. This potentially indicates a cyclical trend that can be further investigated. Furthermore, the strategy sees a negative return in 4 of the first 11 years, and it takes 11 years (to 1973) before the strategy begins yielding above-average results (10%+ annually). The "Additional Notes" section may explain the reason for these two oddities.

See Tables 7 and 8 in the "Appendix" section at the end of this report for the specifics of how the annual returns of the strategy compare to those of the index.

Examining the Trades

Upon back-testing this strategy over the course of the 57 years using compounding returns, the 32.7% annualized return is produced through 10,636 trades, which translates to an average of 186 trades per year. See Tables 1 and 2 under "Appendix" for the statistics of these trades.

To examine why the strategy is successful, let X be the payoff (return) of a trade in a world of 2 events {gain, loss}, G be the random variable representing the amount gained in the outcome "gain", and L be the random variable representing the amount lost in the outcome "loss". Then the expected return per trade, E[X], is as follows:

$$E[X] = P(gain) \times E[G|gain] \; + \; P(loss) \times E[L|loss] \quad \textit{(Equation 2)}$$

E[G|gain] is estimated as the average return of all positive trades while E[L|loss] is estimated as the average return of all negative trades. Meanwhile, P(gain) and P(loss) are estimated by the frequency of each type of trade (gain, loss). Using Equation 2 and the values from Table 1, the strategy's expected return per trade is calculated as:

$$E[X] = (0.13106)*(0.01282) + (0.86894)*(-0.000165) \approx 0.154\% \text{ return per trade}$$

Interpretation: only 13.11% of the trades are gains, but the average gain is significantly larger than the average loss: 1.28% versus -0.017%, respectively. As a result, the expected return per trade is a positive 0.154%. Hence, the strategy is profitable in the long run.

Explanation

The strategy's success implies that the buy- and sell-criteria produces trades at a profitable profit/loss ratio. From comparing the average gain with the average loss, it can be inferred that for trading the S&P index in particular, using the yesterday's closing price as a sell-point keeps losses small but also provides ample room for growth. However, why exactly does this occur? Several factors may be at play.

First, the S&P index as a whole is so diversified that it usually does not move much overnight. Thus, if the open price opens at or above yesterday's close (i.e. the buy criteria is satisfied), it cannot open *that* much higher above yesterday's close. This is proven as follows: on days where the buy criteria is met, the average % difference between yesterday's close and today's open is 0.0219%, very close to zero. From the strategy's point of view, this is a -0.0219% potential margin of loss.

Second, on days where the buy criteria holds and the price does *not* go below yesterday's close on the same day (i.e. the sell criteria is *not* triggered), SPX's average change for the day (measured from open to close) is 1.054%, a *relatively large amount*. From the strategy's point of view, this is a 1.054% potential margin of profit.

Third, on the days where the buy criteria holds but the price *does* go below yesterday's close on the same day (i.e. the sell criteria *is* triggered), SPX's average change for the day (from open to close) is -0.0622%, a *drastic decrease* from 1.054%. Thus, if the sell criteria is triggered, then it is likely that SPX will not make a significant gain for the day.

For comparison, SPX's average change for *any* given day (from open to close) is 0.0284%, close to zero.

From the above information, we can conclude the following:

i) If the buy criteria holds and the sell criteria is *not* triggered on the same day, then it is *statistically likely* for SPX to experience a *relatively large gain* (1.054% on average) on that day. The strategy proactively takes advantage of this fact, and the gains are rigorously protected by the trailing SMA stop-sell that's updated daily at market close.

ii) If the buy criteria holds but the sell criteria *is* triggered on the same day, then it is *statistically likely* for SPX to *not* experience a significant gain (-0.0622% on average) for that day. The strategy defensively avoids this scenario by selling at the predetermined stop-loss (yesterday's closing price) in advance during the day to achieve an average loss of -0.0219%. This cuts losses small before they can get any bigger.

Significance

Since this strategy's annual returns from 1973 onwards are all positive and have zero correlation with the index's annual returns, it is significant that the strategy's profitability is not hampered by recessions (namely, the 2000 dot-com bubble and the 2008 housing crisis). Upon investigating individual trades during these unfavorable economic periods, it was observed that the strategy's bullish buy-criteria prevents it from trading on bearish days. Additionally, the strategy is able to catch and profit off of the often-volatile price action that takes place during a steep bear market, thereby strategically reaping above-average profits during these times. To conclude, the strategy profits off both favorable and unfavorable market conditions.

Additional Notes

A crucial component of this strategy is the comparison of the current day's opening price with the previous day's closing price. Thus, it is important to point out that prior to the mid-1980s, the daily historical price data is structurally different than that of the most recent decades. In particular, the previous day's closing price was almost always *equal* to the next day's opening price. This may be explained by the fact that markets were not electronically accessible until the 1980s, thereby limiting how much the price can change between market close and the following day's market open.

This phenomenon would then limit how well the strategy can perform, for the buy criteria would no longer be meaningful since it would get triggered every day, and the sell criteria would be triggered very easily as well, causing the strategy to experience an abnormally high amount of small losses during these pre-1980 years. So, a new back-test was ran for 1982-2018 (37 years) to avoid the pre-1980s data, and the annualized return increased dramatically to 49.885%.

Implementation

Since the SPX index itself is not tradable, this strategy must be implemented with tradable market-tracking stocks/ETFs, such as SPY. For SPY, back-testing revealed that several "optimal" SMA time frame lengths included 35 days, 150 days, and 210 days. In addition, a "target goal sell" was necessary to be set; these optimal goals ranged from 4% to 9% above the buy price. However, the best annualized returns that could be achieved (from 1993 to 2018) were around 6% to 7%, and there were relatively few trades per year, suggesting an undesirably high amount of correlation between the strategy and the market.

The strategy's differing success between SPY and SPX may be due to the fact that as an index, SPX is less volatile. For example, on 6/4/2020, the difference between the daily high and daily low of SPX and SPY were 1.2458% and 1.2683%, respectively. These small but consistent differences in volatility occur daily and may be the determining factor in whether or not the strategy succeeds. Also, note that SPX's price cannot actively change during aftermarket and premarket hours the way it can for SPY since SPX cannot be traded.

To determine if these index-ETF differences were truly influential in the strategy's performance, the strategy was tested on the Dow Jones Index, commonly known as DJI. From 1985 to 2020, the back-test yields a 20% annualized return with an optimal SMA time frame length of 1-day just like SPX. However, no such results could be reproduced for its ETF counterpart DIA.

This problem may have workarounds that are yet to be found. For example, additional SMAs of varying lengths can be added and candlestick patterns can be analyzed to create a more sophisticated strategy. However, a takeaway for now is that strategies should be built on tradable securities such as SPY rather than indices like SPX to avoid accidental exploitation of price behavior (specifically, lower volatility) that may be unique to indices.

Conclusion

The trading strategy described in this report presents two novel findings. First, it is historically capable of achieving 32% annualized returns by trading the S&P index from 1962 to 2018; this is *triple* of the historical 10% annualized return of the index. Second, the S&P's price in relation to yesterday's close (the 1-day SMA) can serve as a predictor as to whether or not the S&P will make a significant gain during the day. Although these are exciting findings, an implementable method has still yet to be developed since the current strategy's success might be largely, if not fully, attributed to the fact that indices are generally lower in volatility.

Literature Review

My ideas are original, and I have never deliberately researched elsewhere for inspiration for the fear of losing my own ingenuity. To check if similar ideas exist already, however, I searched up "algorithmic S&P trading strategy" on Google.

One of the top results was from "AlgorithmicTrading.net", which provides "robust trading systems to the general public." Their "flagship algorithm" combines seven strategies that trade S&P futures and options based on basic concepts such as momentum and mean-reversion. Another top result was from "Toptal.com", which is a platform for companies to hire freelancers. One article by a data scientist, Andrea Nalon, provides a tutorial on how to train different machine learning models on the S&P data. Upon skimming through both articles, I did not see anything that particularly resembles my own strategy, findings, or trading philosophies.

Overall, the content online regarding algorithmic trading appears rather vague and sparse. This is likely explained by the fact that if someone can develop a high-yielding quantitative trading strategy, they are now in possession of a valuable asset and should have no reason to freely post it online for public use.

Here are the two websites mentioned above:

- https://algorithmictrading.net/project/algorithmic-trading-system-es-emini/
- https://www.toptal.com/machine-learning/s-p-500-automated-trading

Closing Remarks

At the time of this writing, I have developed additional trading strategies that historically yield 20%+ annualized returns when back-tested—this time, on tradable securities. I am currently using these strategies with my own money, and they incorporate some of the ideas listed in this report as well as new ideas. I am grateful that this competition has pushed my creative thinking and analytical skills to new heights, and I look forward to participating again next year.

Special thanks to Professor Shkolnik for providing feedback and guidance on writing this report.

Appendix

Strategy's Positive vs Negative Trades (1962-2018)

	Positive trades	Negative trades
Percentage of all 10,636 trades	13.106%	86.894%
Minimum	0.001%	-1.003%
25th percentile	0.565%	-0.011%
Median	1.011%	-0.009%
Mean	1.282%	-0.0165%
75th percentile	1.676%	-0.002%
Maximum	9.925%	0%
Standard deviation	1.057%	0.051%

Table 1: Summary statistics of profit/loss (%) of positive vs negative trades

Strategy's Trades (1962-2018)

	All trades
Minimum	-1.003%
25th percentile	-0.011%
Median	-0.008%
Mean	0.154%
75th percentile	-0.001%
Maximum	9.925%
Standard deviation	0.584%

Table 2: Summary statistics of profit/loss (%) for all 10,636 trades

Yearly Performance: Index vs Strategy (1962-2018)

Year	Index	Strategy	Year	Index	Strategy	Year	Index	Strategy
1962	-11.81	4.044	1981	-9.73	1.508	2000	-10.14	77.634
1963	18.89	-2.112	1982	14.76	49.497	2001	-13.04	49.128
1964	12.97	-2.147	1983	17.27	46.976	2002	-23.37	81.696
1965	9.06	0.083	1984	1.4	44.985	2003	26.38	68.694
1966	-13.09	3.206	1985	26.33	33.128	2004	8.99	37.636
1967	20.09	-0.596	1986	14.62	43.202	2005	3.0	30.556
1968	7.66	2.283	1987	2.03	94.625	2006	13.62	12.543
1969	-11.36	2.444	1988	12.4	55.486	2007	3.53	20.031
1970	0.1	7.719	1989	27.25	49.914	2008	-38.49	90.901
1971	10.79	3.702	1990	-6.56	35.542	2009	23.45	84.142
1972	15.63	-0.78	1991	26.31	46.781	2010	12.78	58.463
1973	-17.37	17.37	1992	4.46	28.449	2011	0.0	101.625
1974	-29.72	36.1	1993	7.06	21.92	2012	13.41	56.508
1975	31.55	17.188	1994	-1.54	27.867	2013	29.6	45.744
1976	19.15	2.072	1995	34.11	24.863	2014	11.39	40.325
1977	-11.5	2.183	1996	20.26	46.275	2015	-0.73	63.029
1978	1.06	11.501	1997	31.01	82.774	2016	9.54	30.559
1979	12.31	3.941	1998	26.67	73.139	2017	19.42	9.19
1980	25.77	15.135	1999	19.53	99.971	2018	-6.24	8.095

Table 3: Comparison of yearly returns (%) of index vs strategy

Yearly Performance: Index vs Strategy (1962-2018)

	Yearly index return	Yearly strategy return
Minimum	-38.49%	-2.15%
25th percentile	-0.73%	7.72%
Median	10.79%	30.56%
Mean	7.74%	35.03%
75th percentile	19.42%	49.91%
Maximum	34.11%	101.63%
Standard deviation	16.09%	30.15%

Table 4: Summary statistics comparison of yearly returns (%) of index vs strategy