Developing non-Machine Learning Algorithms for determining when to buy and sell on the stock market

Grant Proposal

Andrew Kalashnikov

Massachusetts Academy of Math and Science

Worcester, MA

Executive Summary (Eng)

The stock market is a difficult place for smaller investors to be successful, with constant competition against far larger entities. A particular issue that many smaller investors struggle with is the lack of access to effective, understandable, and flexible algorithms. Active investment for minor/individual players is difficult in a stock market where large investors have access to incredibly complex machine-learning algorithms to aid investing decisions. To address this issue, I will be attempting to develop an algorithm geared toward smaller investors that will outperform the basic algorithms available today and give smaller investors a chance to participate in the market on an equal footing. Previously, investors relied on a buy-and-hold strategy that has subsequently been shown to be flawed in an age of AI-guided investment strategies. We will attempt to develop a number of different algorithms using historical data, based on data trends such as moving averages and price changes to determine optimal times to buy and sell. The goal is to combine these techniques to create an algorithm that can result in higher returns than buy and hold by a statistically significant margin over a large number of stocks, ETFs and funds. We will use data from a very wide variety of securities to attempt to develop multiple algorithms that are widely applicable.

How do game theory algorithms compare to conventional algorithms for investment in the stock market?

The stock market is a place where fortunes are won or lost depending on how accurate the predictions of price movement are. An important part of these predictions is determining when to buy and sell stocks. Predictions are often done through algorithms, as these can counteract the natural human desire to develop storylines. For example, a study found that investors were influenced by both news information and price trends but relied more heavily on news items than trends in the data (Sobolev et al., 2017).

Table 1: Taken from (Sobolev et al., 2017): This table shows the reaction to news and data, demonstrating the overreliance on news

Table 1. Results of Experiment 1

			Trend	
			Positive	Negative
Panel A : Western group, N = 30				
Trading latency	News valence	Good	48.33 (48.89)	60.00 (61.84)
		Bad	45.67 (38.28)	36.67 (24.75)
Share number	News valence	Good	1.35 (0.92)	0.83 (0.96)
		Bad	1.03 (1.01)	0.4 (0.81)
Returns	News valence	Good	3.14 (2.00)	−3.31 (2.22)
		Bad	2.12 (1.38)	−2.68 (1.16)
Panel B: Eastern group, N = 30				
Trading latency	News valence	Good	92.33 (71.98)	106.34 (70.78)
		Bad	66.00 (55.27)	73.66 (61.45)
Share number	News valence	Good	1.13 (0.96)	1.15 (0.917)
		Bad	0.72 (0.96)	0.60 (0.87)
Returns	News valence	Good	4.24 (2.25)	−4.35 (2.25)
		Bad	2.94 (1.89)	−3.44 (2.02)

Although taking current events and news into account is an important component of stock prediction, looking at historical trends tends to provide a better mathematical base for investments as proven by the effectiveness of investments by large investment firms that utilized algorithmic trading. Most algorithms currently in use today are complex and utilize complicated machine learning methods, meaning that they are developed and kept proprietary by very large investors. Because they were developed by these entities, these algorithms to only the uber-rich and investment firms, giving them a leg up over smaller investors for whom these machine learning algorithms are too cumbersome to utilize effectively. To this end, small investors are mostly left to use very basic strategies such as buy and hold, leaving them at a significant disadvantage. Developing an algorithm that is more complicated than buy and hold but simpler than the machine learning algorithms used by large firms will help fill a noticeable gap in the industry.

Background Information

In this paper, the focus will be on three different investment tools: stocks, mutual funds and exchange-traded funds. A stock is a fractional part of a company, and the owner of a stock is a shareholder in that company, meaning that they own some portion of that company. A mutual fund is an actively managed investment portfolio that takes money from many small investors and uses it to invest in stocks according to invest in securities in a way that aligns with the fund's overall strategy (Picardo, 2024). Finally, exchange-traded funds (ETFs), are a hybrid of stocks and mutual funds, where they can be traded at any time and usually track a single index while still owning small parts of different stocks. The strategy usually recommended for small investors is buying and holding a single stock or fund for an extended period, as this generally flattens out market fluctuations over time and usually leads to growth in the long term. However, the issue with these strategies is that the maximum return on investment is quite low, and it also may prevent the investor from selling stocks that are losing value (Sanderson & Lumpkin-Sowers, 2018). This strategy is very far from the more robust and active management strategies used by larger investors, as these strategies tend to be somewhat riskier but usually significantly more effective. These strategies aren't even viable for the smaller investors that are willing to tolerate risk, as they will need to spend large amounts of time that they don't have analyzing news and price trends (Taylor, 2024).

Previous Research

Most previous research conducted in this field has focused on machine learning algorithms that used the previous trends of a specific investment to attempt to predict its future movement (Ayyildiz & Iskenderoglu, 2024).

There has been relatively little research focused on making algorithms without machine learning due to the assumption that a machine learning algorithm would be more accurate and the fact that there was no need to avoid using machine learning. It was also assumed that the buy and hold approach should be sufficient for those smaller investors whom the machine learning algorithms didn't suit. However, recent research (Table 2) has shown that the buy and hold strategy may not be as viable in the new era of stock market volatility (Sanderson & Lumpkin-Sowers, 2018). The worsening outcomes of the buy and hold approach suggest that a new strategy will be necessary for low-volume investment.

Table 2 (Taken from Sanderson & Lumpkin-Sowers, 2018): This table is a demonstration of the results of a 19 year buy and hold strategy applied to ETFs, and it shows that only .07 % survive due to the quick turnover, and thus a simple buy and hold is almost 100% losing over long periods of time

Table 2. Success Rates for All NASDAQ Exchange Traded Funds (ETF's).

Holding Period (Years)	Mean Prob. of a Gain	Min	Max	Range	Std Dev	Skew	CV	Number of ETF's	Survival Rate
1	59.49%	1.30%	100.00%	98.70%	24.03%	−0.5792	0.404	1374	
2	67.12%	1.16%	100.00%	98.84%	27.43%	−0.6458	0.409	1228	89.37%
3	72.71%	1.22%	100.00%	98.78%	26.86%	−0.8895	0.369	1055	76.78%
4	75.18%	1.18%	100.00%	98.82%	26.85%	−0.9540	0.357	944	68.70%
5	78.29%	2.74%	100.00%	97.26%	25.83%	−1.1139	0.330	799	58.15%
6	83.73%	4.55%	100.00%	95.46%	24.76%	−1.5002	0.296	673	48.98%
7	87.11%	6.67%	100.00%	93.33%	22.88%	−1.8630	0.263	571	41.56%
8	87.24%	3.19%	100.00%	96.81%	23.60%	−2.0863	0.271	452	32.90%
9	93.20%	9.00%	100.00%	91.00%	16.12%	−3.0983	0.173	332	24.16%
10	95.97%	10.00%	100.00%	90.00%	11.97%	−5.0068	0.125	207	15.07%
11	97.02%	40.79%	100.00%	59.21%	8.81%	−4.0153	0.091	166	12.08%
12	96.90%	43.75%	100.00%	56.25%	8.80%	−3.7230	0.091	138	10.04%
13	97.71%	55.77%	100.00%	44.23%	7.71%	−4.0150	0.079	112	8.15%
14	98.01%	50.00%	100.00%	50.00%	7.59%	−4.4286	0.077	101	7.35%
15	97.38%	36.36%	100.00%	63.64%	10.29%	−4.2739	0.106	82	5.97%
16	97.98%	66.67%	100.00%	33.33%	7.28%	−3.8001	0.074	32	2.33%
17	99.17%	83.33%	100.00%	16.67%	3.35%	−4.2810	0.034	30	2.18%
18	98.96%	79.17%	100.00%	20.83%	4.66%	−4.4721	0.047	20	1.46%
19	99.12%	83.33%	100.00%	16.67%	3.82%	−4.3589	0.039	19	1.38%
20	100.00%	100.00%	100.00%					2	0.15%
21	100.00%	100.00%	100.00%					1	0.07%
22	100.00%	100.00%	100.00%					1	0.07%
23	100.00%	100.00%	100.00%					1	0.07%

Section II: Specific Aims

This proposal's objective is to develop a non-machine learning algorithm to determine when to buy and sell on the stock market, compare it to previous such algorithms, and see improved performance.

Our long-term goal is to develop algorithms that will outperform previous non-machine learning algorithms and be usable for small investors. Developing such algorithms will hopefully level the playing field for small investors in comparison to larger firms and make the stock market a place where investors have similar chances of profitability or loss, no matter their size.

Specific Aim 1: The first aim of the project was to develop a number of algorithms for determining when to buy and sell.

Specific Aim 2: The second aim of the project was to determine how well these algorithms performed against conventional algorithms and compare them to each other in order to find the strongest algorithm to hopefully compare it to ML algorithms.

The expected outcome of this work is to create a non-ML algorithm that is relatively simple to understand for those with a high-school level understanding of mathematics and a solid understanding of the stock market in general. Additionally, this algorithm is ideally meant to outperform the buy and hold algorithm and be somewhat comparable to machine learning algorithms, which make correct decisions roughly 70% of the time (Ayyildiz & Iskenderoglu, 2024).

Section III: Project Goals and Methodology

Relevance/Significance

Solid investments are critical for building middle class wealth and reducing the income inequality that has been growing rapidly in some developed countries including the US (Saez & Zucman, 2020). However, the stock market has long been dominated by larger institutional investors that have access to greater resources than smaller investors. These investors find themselves in a very unfriendly market in direct competition with larger investors that have access to ML algorithms that are extremely complex, hard to develop, and difficult to implement with the resources at their disposal. Even to this day, algorithmic trading is mainly restricted to large firms and institutional investors (Chen, 2024), meaning that most small investors still rely on the same decision-making methods as their predecessors a hundred years ago. This issue has the possibility to become even more serious as ML continues to increase in accuracy and complexity, widening the gap between those utilizing algorithms and those that fail to do so.

Another important issue to note is that many smaller investors lack an understanding of the function of ML algorithms and thus be reluctant to utilize them due to their lack of faith in an algorithm whose methods are new and unknown to them. Developing an algorithm without the use of advanced computer science techniques will make it much more understandable for investors and give them a basis for building their own algorithms to account for specific needs. Clearly demonstrating the development process of this algorithm can also help regulators and policymakers by increasing the transparency of investment algorithms, as opposed to the many “black box” algorithms in use today (Chen, 2024).

Innovation

This project aims to target an underserved group of investors, small investors with a base level of technical knowledge, which previous algorithms have failed to adequately serve. This algorithm will also utilize a variety of different factors in decision-making without machine learning and be adjustable for variable risk tolerance among the diverse group of users it is attempting to support. It will also hopefully become one of the first easy to understand and publicly available algorithms to compete with machine learning algorithms on performance. Additionally, this algorithm will be able to work across very different market segments and types of investments with minimal changes, adding to its flexibility and ease of use. It would also allow investors to continuously change the exact parameters of the algorithm as they see fit, allowing them to adjust to a changing world at the personal level, giving the individual more influence on the outcome. Additionally, human input may improve performance by accounting for factors beyond the algorithms control, such as changes in the financial or political climate (Huang et al., 2024).

Methodology

Specific Aim #1:

The first objective of this work was to develop baseline algorithms for determining when to buy and sell, ideally having a few with promising performance on initial testing before moving onto aim #2.

Justification and Feasibility.

The algorithms will be developed using historical data from a large variety of different stocks, ETFs and mutual funds such as the S&P 500 (Qin, 2018). It will use a variety of techniques, notably MA, day-to-day change and, the derivative of this change to determine the direction of price movement. A series of independent algorithms will be developed for further testing, and the entire project will utilize historical data from Investing.com which will be cross verified with other sources. The code to execute the testing of algorithms will be run using python and specifically the Pandas module, as the stocks and their associated data will be inserted into a Pandas data frame for processing. The testing method will allow for rapid testing of strategies on data and also give more flexibility in the type of output we receive. Non-machine learning algorithm development has been previously attempted in a method that was vaguely similar to this project in a paper by Kuo and Chou in 2021, and this is a graph of the performance of their algorithms, which demonstrates the ability to outperform buy and hold with a certain algorithm.

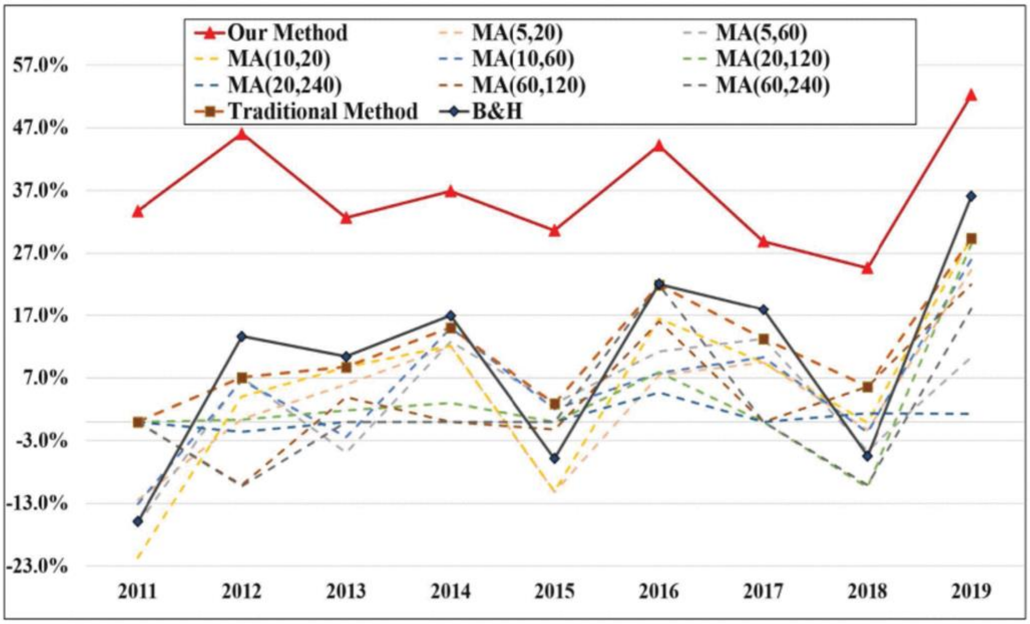

Figure 1(Taken from Kuo & Chou 21): Performance of various developed algorithms as well as Buy and Hold

Summary of Preliminary Data.

A single algorithm has been developed, which has slightly outperformed buy and hold in initial testing. This algorithm is based mainly on moving averages; however, some tolerance margins are added that slightly alter its function from a simple MA approach.

Expected Outcomes.

The overall outcome of this part of the project is to have several algorithms that perform as well or better than buy and hold in initial testing to move to further testing.

Potential Pitfalls and Alternative Strategies.

Some potential issues may be a failure to develop algorithms that sufficiently outperform buy and hold or failing to develop a sizeable number of these algorithms. Another potential issue is having these algorithms be too rigidly defined to be altered easily by the user to adapt to their needs.

Specific Aim #2:

The next step is to compare the performance of the algorithms to each other and to other conventional algorithms, such as buy and hold, based on the returns generated over the 19-year period examined in the study when utilizing each algorithm.

Justification and Feasibility.

The work done in this section is focused on finding the best of the algorithms that were developed to make sure that we accurately represent the optimal solution that was found. It's also crucial to test the performance of this algorithm against already existing algorithms to find issues and demonstrate a strong reason for its use by having it outperform the buy and hold method. It is also important to note that one of the studies that showed how some ETFs completely collapse in a similar 19-year period (Figure 2) (Sanderson & Lumpkin-Sowers, 2018).

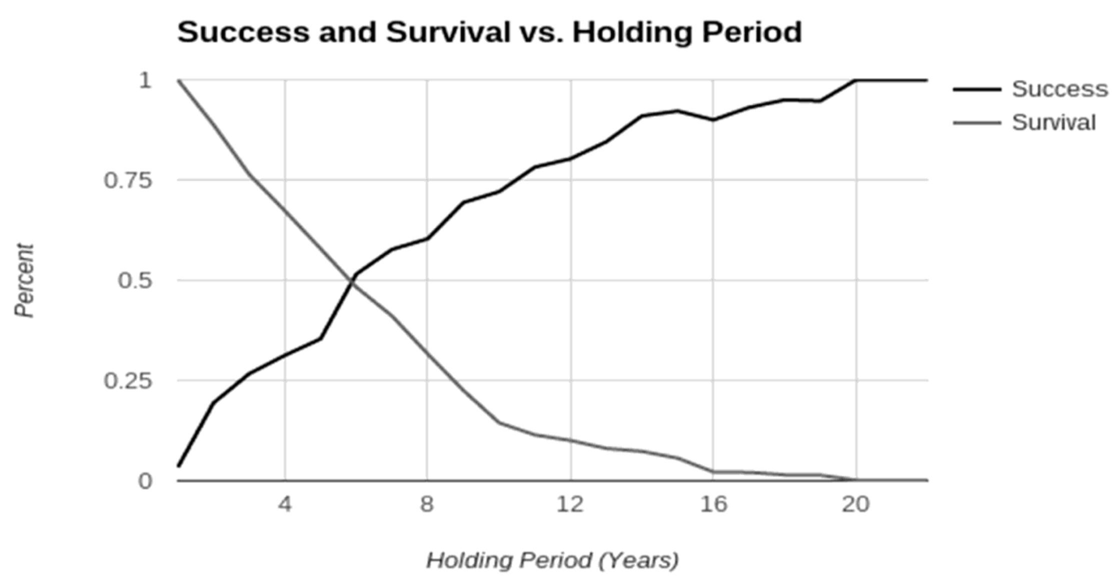

Figure 2(Taken from Sanderson & Lumpkin-Sowers, 2018): Graph of the survival of ETF's over different time periods, shows the disaster of buy and hold and what we have to avoid in our algorithms

Expected Outcomes.

The overall goal of this aim is to discover the most effective algorithm for the stock market, to attempt to provide a solution for smaller investors.

Potential Pitfalls and Alternative Strategies.

We expect a lot of issues with developing algorithms that perform well enough to be statistically significantly better than buy and hold.

Section IV: Preliminary Data

The focus will be on the results of the first algorithm that was developed, which was built around the moving average of the stock price exclusively. This algorithm showed strong results in initial testing with only the Apple stock, demonstrating 13% higher returns over buy and hold over the full 20 year period (Figure 3).

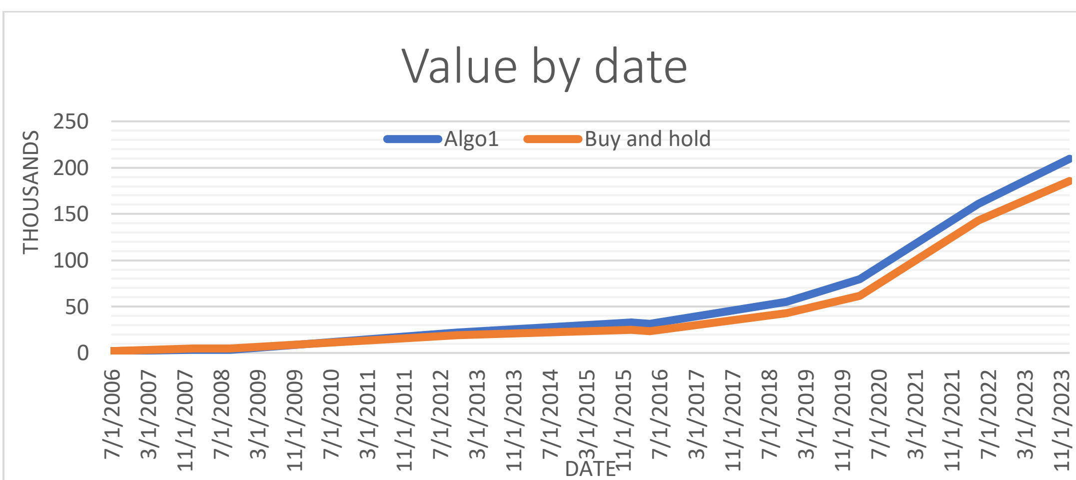

Figure 3: Graph of the total value of the apple stock using buy and hold and my algorithm by date for the entire time of the experiment

The next step of testing was with a group of six stocks, where the algorithm performed at an average of 10% over buy and hold, but results were very varied. Finally, the algorithm was tested with the entire group of 30 stocks, where it was discovered to perform worse than buy and hold with a mean return that was about 10% less (Figure 4).

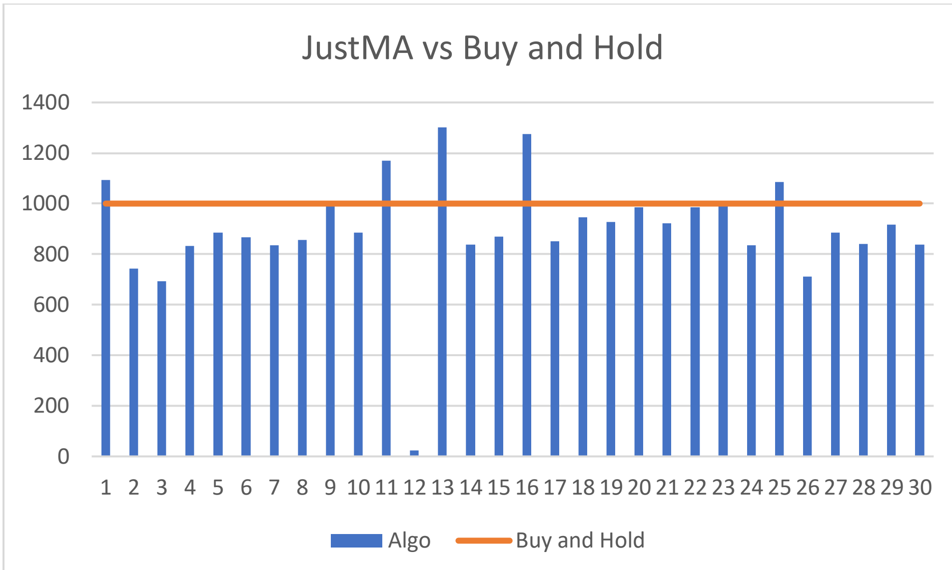

Figure 4: This figure shows the results of the MA algorithm when compared to buy and hold across all 30 stocks

Section V: Conclusion and Discussion

The data from the test of all 30 stocks is a bit disappointing since the developed algorithm performs notably worse than buy and hold across all 30 stocks. However, there are a couple of promising things that can be seen in this information. One of the most important ones is that the Apple data is clearly not strongly correlated with the rest of the market data meaning that overfitting isn't an issue as shown by these results. Another positive from this data is that this algorithm doesn't destroy the value of the shares, so if some of the deficiencies can be addressed through further development, there is potential for it to become better than buy and hold.

Section VIII: References

Ayyildiz, N., & Iskenderoglu, O. (2024). How effective is machine learning in stock market predictions? *Heliyon, 10*(2). https://doi.org/10.1016/j.heliyon.2024.e24123

Chen, J. (2024). *Algorithmic trading: Definition, how it works, pros & cons*. Investopedia. https://www.investopedia.com/terms/a/algorithmictrading.asp

Chen, C., Chen, C., & Liu, T. (2020). Investment performance of machine learning: Analysis of S&P 500 Index. *International Journal of Economics and Financial Issues, 10*(1). https://www.econjournals.com/index.php/ijefi/article/download/8925/pdf/21843

Huang, W., Satoru, G., & Nakamura, M. (2021). Decision-making for stock trading based on trading probability by considering whole market movement. *European Journal of Operational Research, 157*(1), 227–241. https://doi.org/10.1016/S0377-2217(03)00144-9

Huang, Y., Zhou, C., Cui, K., & Lu, X. (2024). Improving algorithmic trading consistency via human alignment and imitation learning. *Expert Systems with Applications, 253,* 124350. https://doi.org/10.1016/j.eswa.2024.124350

Kuo, S.-Y., & Chou, Y.-H. (2021). Building intelligent moving average-based stock trading system using metaheuristic algorithms. *IEEE Access, 9,* 140383–140396. https://doi.org/10.1109/ACCESS.2021.3119041

Ling, F., Ng, D., & Muhamad, R. (2014). An empirical re-investigation on the ‘buy-and-hold strategy' in four Asian markets: A 20 years’ study. *World Applied Sciences Journal, 30*(30). https://doi.org/10.5829/idosi.wasj.2014.30.icmrp.30

Picardo, E. (2024). *Investing explained: Types of investments and how to get started.* Investopedia. https://www.investopedia.com/terms/i/investing.asp

Qin, L. (2018). Game theory-based investment strategy vs. buy-and-hold: Which optimizes profits? *ISEF Abstracts.* https://abstracts.societyforscience.org/Home/FullAbstract?ISEFYears=0%2C&Category=Any%20Category&Finalist=Qin&AllAbstracts=True&FairCountry=Any%20Country&FairState=Any%20State&ProjectId=16428

Saez, E., & Zucman, G. (2020). The rise of income and wealth inequality in America: Evidence from distributional macroeconomic accounts. *Journal of Economic Perspectives, 34*(4), 3–26. https://doi.org/10.1257/jep.34.4.3

Sakhare, A., Mhaskar, N., Mishra, V., & Chavan, M. (2021). Algorithmic trading for a buy-sell platform: Study and comparison. *ITM Web of Conferences, 40,* 03020. https://doi.org/10.1051/itmconf/20214003020

Sanderson, R., & Lumpkin-Sowers, N. (2018). Buy and hold in the new age of stock market volatility: A story about ETFs. *International Journal of Financial Studies, 6*(3). https://doi.org/10.3390/ijfs6030079

Sikalo, M., Arnaut-Berilo, A., & Zaimovic, A. (2022). Efficient asset allocation: Application of game theory-based model for superior performance. *International Journal of Financial Studies, 10*(1), 20. https://doi.org/10.3390/ijfs10010020

Sobolev, D., Chan, B., & Harvey, N. (2017). Buy, sell, or hold? A sense-making account of factors influencing trading decisions. *Cogent Economics & Finance, 5.* https://doi.org/10.1080/23322039.2017.1295618

Taylor, B. (2024). *5 key investment strategies to learn before trading*. Investopedia. https://www.investopedia.com/investing/investing-strategies/