

2023 年臺灣國際科學展覽會 優勝作品專輯

作品編號 130013
參展科別 行為與社會科學
作品名稱 Application of Technical Analysis to the
Stock Market
得獎獎項 一等獎

國家 Czech Republic
就讀學校 Ricany Grammar School
指導教師 Tomas Dosedel
作者姓名 Pavel Kletecka

關鍵詞 Investing、Market simulation、Technical
analysis indicators

作者照片



Abstract

This paper aims to create a program that allows training and testing technical analysis strategies on historical data. The paper describes the basic knowledge needed to understand technical analysis, its indicators and the strategies constructed from them. The steps that lead to the creation of such a program are also described, and the basic operation of said program is described. Finally, the paper presents two new proposed indicators and shows the outputs that may arise from the results of testing with this program.

Introduction

Money drives the world and in times of crisis, money is more important than ever. Most political, economic and, as we have seen recently, health crises are accompanied by pressure on the economy. With such pressure, ordinary citizens are worried about their money, which is at risk either from inflation or from an uncertain economic outlook. In such times, some people resort to appreciating their money by investing. It is wise investments that can protect savings from inflation, or at least mitigate the effects of inflation. Investing in the stock market is among the most popular ways of investing. During the global coronavirus crisis, the number of small investors more than doubled (Galik and Brody [2022]), and in 2021, Covid (new) investors accounted for 15 % of investors in the U.S. market (Schwab [2021]). It is this phenomenon that inspired the author to write this paper.

Investing in the stock market is one of the riskiest forms of investment, which means that there is a high probability of losing the originally invested capital. Some companies that allow retail investors to invest state that more than 90 % of their users lose their capital when investing in the stock market. On the other hand, investing in the stock market has one of the greatest potentials for profit. In the case of a long-term investment in index funds, for example, an average annual appreciation of between 8 % and 12 % can be expected (the average appreciation of the S&P 500, the most famous US index). However, it is necessary to wait several years for stable results. To see appreciation in a long-term portfolio, it is recommended to wait at least 12 years. This is because if a recession or a simple market correction comes, said portfolio can lose up to 50 % of its value in a year. But as history shows, markets do rise over the long term, and that is what long-term investors bet on, waiting for their capital to appreciate over time.

But the market does not grow all the time. The market often changes price trends, and in some rare cases a long-term trend can reverse for up to several tens of months. This phenomenon, volatility, is attempted to be exploited by so-called speculators. A speculator is an investor who sells and buys in short periods of time and thus speculates on price movements. The speculator therefore needs a way to determine the likely behaviour of the price in the foreseeable future in order to make their trades profitable. The two main ways of analysing price behaviour are fundamental and technical analysis. And since fundamental analysis incorporates a large amount of diverse information into its predictions, this paper concentrates on working with technical analysis.

1. Theoretical part

1.1 Stock market

Since the whole paper revolves around stocks and trading in the stock market, it would be good to first clarify what a stock is, where it gets its value and how it is traded.

A stock is the name for a security that represents a stake in a joint stock company. Share is a unit of stock and one share is the smallest fraction of stock company. Joint stock company issues shares to raise capital by selling shares to the stockholders. By purchasing a stock, they acquire a stake in the company, which is the right to vote at general meetings, a cut of the liquidation proceeds and the ability to share in the profits of the company (if the company chooses to do so, which is decided at a general meeting). It is not only the absolute number of shares that is important, but also the relative number to all the shares of the company. The more shares one has in a company, the more influence he or she has. It is these essentials that give a share its value. (Teweles and Bradley [2008])

People can buy shares on the stock market. But what most people think of as the stock market is only one part of it. Generally, the secondary stock market is referred to as the stock market, where shares are traded between individual investors on the principle of supply and demand. However, for shares to enter this market, they must first be issued by a company on the primary stock market, where investors buy shares directly from the company, where the price is predetermined by authorised investment bank. (Teweles and Bradley [2008])

However, accurately determining the value of stocks is very difficult. To determine the value of a stock as accurately as possible, fundamental analysis is used to find the true value of the stock. This is determined based on many factors such as macroeconomic factors (inflation, exchange rate, political situation in the country, etc.) and microeconomic factors (level of profits, earnings, costs, debt, etc.), risk, liquidity, competitive position, etc. This price is independent of the market price, but often the market price tries to match the true value of the stock. Fundamental analysis is used in determining the price of a stock in the primary stock market and by some investors in the secondary stock market. In the secondary stock market, the predominant method of determining future price behavior is through technical analysis. (Murphy [1999])

1.2 Technical analysis

Murphy defines technical analysis as “the study of market movements, primarily using charts, for the purpose of predicting future price trends.” (Murphy [1999]). Therefore, it can be said that technical analysis is a way of predicting the likely behaviour of a price based on market data. It is based on the idea that there is no need to observe the factors that fundamental analysis deals with, as they are all projected onto the market through other investors who follow them (Mur-

phy [1999]). Technical analysis only looks at information from the market and uses knowledge from econometrics (a combination of mathematics, statistics and economics) to process it.

1.3 Indicators

Technical analysis uses so-called indicators to process market data. These indicators receive information from the market about the opening, minimum, maximum and closing prices as well as the number of trades executed over a selected time period. Bauer and Dahlquist [1999]

The indicator acts as a market simplification tool by generating buy and sell signals. Each indicator thus contains theoretical information about when it generates which signal. Many indicators are defined by a mathematical formula into which market data are fed. Examples of mathematically expressed indicators are Moving Averages and examples of non-mathematical indicators are Supports and Resistances (these can also be expressed using mathematics, but this is not a common practice). This paper works specifically with indicators that can be expressed mathematically, because they can then be easily transcribed into the code of a testing program that evaluates them faster than an investor could.

Indicators can be divided into two basic types, trend indicators and oscillators. The trend indicators show the direction, strength and possible reversal of a trend. The oscillators are often used in a non-trending environment and move within a range of set values. An important difference between a trend indicator and an oscillator is the different way they react to price behaviour. While oscillators detect a change in price before the change occurs, trend indicators show the price change after the change has occurred. Hence, it is advisable to use trend indicators to determine the strength and direction of the trend and to start using an oscillator when a reversal is imminent, which predicts the likely behaviour of the price after the trend ends in a better way. Therefore, it is general practice to use multiple indicators when analyzing a single asset. (Bauer and Dahlquist [1999])

The essential choice in analysis is not only the right combination of indicators, but also their setting by means of parameters. The concept of indicator sensitivity is much related to this. Indicator sensitivity describes how accurately an indicator reacts to price changes. Highly sensitive indicators detect even the slightest, sometimes irrelevant, changes, but low-sensitivity indicators can then sleep on the signal and generate it too late. Different indicators have different sensitivities with the same settings, but it is the parameter settings that affect sensitivity the most. Therefore, the correct setting of each indicator is crucial for the best possible analysis.

Sensitivity is also related to the length of the position holding. Every investor changes positions, buys and sells, after a different length of time. For positions held for a few days, it is recommended to use sensitive indicators, while for investing over months, less sensitive indicators are used. According to Murphy, it is possible to change the sensitivity of indicators when a change in trend is about

to occur, which the investor needs to analyse as accurately as possible (Murphy [1999]).

An example of a mathematically expressed indicator together with the necessary theory for its use is described in chapter 3.2.

1.3.1 Trading strategy

A trading strategy is a set of activities leading to the achievement of business goals. Not all investors follow clear strategies, thereby increasing the riskiness of their investment. However, it is also important to note that trading according to a strategy is risky and in some cases unprofitable. Therefore, it is important to find the optimal trading strategy, which is different for each investor according to their objectives and the level of risk they are willing to take.

A stock market trading strategy must contain two key pieces of information. The strategy must include a clearly defined time to open a position and a clearly defined time to close a position. The term buy and sell shares is deliberately not used here.

Types of positions

Whether opening a position means selling or buying a stock depends on the type of position. There are two types of positions, a short position and a long position.

A short position is a way of profiting from a fall in price. The investor borrows shares, which he sells when he opens the position, and buys the same number of shares back to return them when he closes the position. The investor thus makes a profit if he closes the position at a lower price than he opened it. This method is significantly riskier than a long position.

A long position makes a profit based on the rise in price. The investor buys shares when the position is opened and sells them when the position is closed. In this case, the investor wants the price to rise and wants to exit the position at a higher price than the price at which he opened the position. This type of position is used more commonly.

Signal generation

The individual signals are generated by the mentioned indicators. Each indicator, in addition to its expression, must contain theoretical information about when it generates what signals. The theory usually gives signals for a long position, but these can be easily converted into signals for a short position by interchanging the opening and closing signals (the opening signal of a long position equals the closing signal of a short position and vice versa).

The investor does not have to trade only according to the signals from one indicator. Combining different indicators filters out false signals and reduces risk.

According to Krutina, the most important moment of trading is closing the position, because not closing the position in time leads to a loss of invested capital, while missing the signal to open costs the investor only theoretical profit. Therefore, a proper strategy should include a tool to control potential losses (Krutina [2021]).

Stop-loss

A stop-loss is a trader's tool that allows them to set a threshold that, when the price crosses, closes the position. It is a tool against a steep fall in price. A similar tool is the take-profit, which, like the stop-loss, sets a threshold that, when crossed, closes the position. The difference is that the take-profit is positioned to result in a controlled profit, while the stop-loss controls the loss.

Some indicators in their theory also generate a stop-loss value in addition to the opening signals, which is appropriate according to this indicator. For simple investing this can be useful information, but when combining multiple indicators the stop-loss values can vary and the investor has little means of adjusting the level of risk.

An interesting way that solves all these problems of how to set the stop-loss is the Modified Standard Exit Strategy (*MSES*), which is a modified version of the Standard Exit Strategy (*SES*). The *MSES* determines risk by multiplying the value of the Average True Range (*ATR*) indicator. The value of the *MSES* is calculated according to formula (1.1)

$$SES_{t+1}(n, m) = x_t \pm m \cdot ATR_t(n), \quad (1.1)$$

where

x_t is the closing price at time t ,
 n is the length of period,
 m is a selectable parameter,
 $ATR_t(n)$ is the value of the Average True Range indicator with parameter n at time t calculated using the formula (1.2)

$$ATR_t(n) = \frac{1}{n} \cdot \sum_{i=0}^{n-1} \max\left\{h_{t-i} - l_{t-i}; |h_{t-i} - x_{t-i-1}|; |l_{t-i} - x_{t-i-1}|\right\}, \quad (1.2)$$

where

x_t is the closing price at time t ,
 h_t is the maximum price at time t ,
 l_t is the minimum price at time t ,
 n is the length of period,
 $\max\{\dots\}$ is the function whose value is equal to the value of the element with the highest value.

Whether to add or subtract the *ATR* multiple (in 1.1) depends on the type of position. For a long position the multiple is subtracted, for a short position it is added.

The *SES* indicator has set a stop-loss only when a position is opened and this stop-loss does not change for that position. This determines the maximum loss for the position, but it does not at all address the possibility that the profit of the open position increases and the stop-loss needs to be readjusted.

This problem is solved by *MSES*. The *MSES* value sets the stop-loss for each following time period after the position is opened. The stop-loss value is only changed if it reduces the loss. If the *MSES* value should set the stop-loss so that it increases the loss, the stop-loss boundary does not change (Krutina [2021]).

2. Methodology

This section of the paper explains how the practical part of the analysis is conducted. Since the main contribution of this paper is the testing program itself and the results from it, the idea that leads to its programming is described here. The strategies are trained and evaluated using a set of historical data as described in the following text. The implementation into the program is described as well. But the most important thing for the investor is that the program tells them which strategies are the most profitable. This is what the section on strategy evaluation is for. The program also allows investors to track the riskiness of their strategies, which is just as, if not more, important than the profitability. Finally, there is a section on why and how an investor should compare their strategies with other types of investing.

2.1 Testing program

The main aim of the project is to develop a program for creation of strategies based on technical analysis and their further training and testing on historical data.

The testing program is written in the C# programming language and uses a MySQL relational database to store the data. The program is written as a console application for which a GUI (Graphics User Interface) is being created at this time. This initial console application is therefore used only for computation and verification of the functionality of all aspects of the program.

One of the key features of this program, compared to similar existing programs, is the variability of strategies and the possibility of creating your own strategy. A strategy is composed of two conditions that generate signals (opening and closing). The individual conditions further consist of indicators. The value of the indicator is calculated in the strategy for each day and generates a signal for the next day based on the set condition. The signals are generated after the market closes and the positions are adjusted when the market opens. The high variability is achieved by the fact that the investor can choose up to 14 commonly used indicators, set the individual indicators using parameters and then combine them without limitation. When using a large number of indicators for one strategy, the only disadvantage is an increase in computation time. Such strategies are then trained and tested on historical data through two cycles (training and evaluation).

2.2 Data

In order to train and evaluate strategies on historical data, the program first needs to obtain these data. Historical data are drawn from the Yahoo Finance site (Yahoo! [2022]) by a special function that has been implemented in the program. This project uses this program to invest in the stock market, but the

special feature mentioned above allows investor to download historical data from both the commodity market and the forex market. This allows the investor to create strategies based on technical analysis not only for the stock market.

This historical data are stored in a database under an assigned code, which consists of a code letter and a number. The historical data can be uploaded into two groups of test periods. The first group consists of training periods, here the goal of the program is to find the strategy that behaves most optimally on the selected training periods. Based on the result from this group, the investor can make changes in their strategies. However, in order not to create overtrained strategies (strategies that have high efficiency only on a particular combination of training periods) they should still evaluate their strategies.

Evaluation is a process whereby strategies are tested on evaluation periods that have been chosen randomly. The investor should only take the results of the evaluation of the chosen strategy as a guide, because the evaluation set of test periods is intended to simulate the use of the chosen strategy in the current market by its random selection. Thus, the investor should not adjust the strategy based on the results of the evaluation, but only decide whether he wants to use the strategy.

In order to be able to subdivide the test periods within groups, the investor can mark the periods they want to follow together with the same code letter. This can be done both in the training group and in the evaluation group.

2.3 Evaluation of strategies

The evaluation of the strategies takes place in both sets of historical data. This way, the program determines the efficiency of the strategy, which is the ratio between return and risk, which is described in a separate chapter (3.1.4). When comparing two strategies, the more efficient strategy should be better. However, the problem here is that the more efficient strategy may be very risky, more than the investor wants, and so efficiency is not the only thing an investor should consider.

The key assumption

For strategy testing (simulating trading by strategy on historical data) to be accurate, the program must satisfy one important assumption. This assumption is that the program computes the indicators sequentially in sense that at time t only data known up to this time t are used. This results in the program not taking into account data that investor would not know in actual trading. Some similar existing programs have this problem and do not address it, so the investor gets distorted results that do not match reality.

The testing process

Testing a strategy on a test period results in information about the return rate, riskiness and number of trades executed.

The return rate is calculated in a standard way, it is the fraction of the return (the difference between the value of the simulated portfolio before and after the test period) and the original value of the portfolio.

The program can calculate the risk in three ways. The first two ways are similar. These ways measure the riskiness of a strategy using the standard deviation or mean absolute deviation of the percentage gain of each test period. Both methods work on a similar principle and their results resemble. The third way is to use the coefficient of variation, which works with the standard deviation. However, the standard deviation is divided by the average gain, which gives riskiness that can be compared even between strategies with very different percentage appreciations. However, the standard deviation was used in the control testing, which is the author's preferred method.

The last important piece of information for each test is the number of trades made. This information is important because of the fees that different brokers charge for trade execution. The current version of the program does not allow recalculation of the test results based on the broker's fees. In a future version of the program, it is planned to allow investors to enter the amount of fees as an input parameter (or directly the broker through which the investor wants to trade) and the program would include this information in the testing.

The automatic evaluation

The automatic evaluation process is impractical with small number of strategies tested. While the program can evaluate average profit and risk, it cannot determine which strategy is the most optimal for a given investor. Such a decision must be made by the investor.

However, with a large number of strategies being tested, an investor cannot manually evaluate all strategies. At this point it is necessary to use automatic evaluation. The program ranks the strategies according to their effectiveness and the investor then manually evaluates those that performed best in the automatic evaluation.

This program allows for three types of automatic evaluation, resulting from three types of risk calculation. As with riskiness, the first and second evaluation models (models using standard deviation and mean absolute deviation) lead to similar results, while the evaluation using the coefficient of variation gives different results.

For an appropriate manual evaluation, it is advisable to select strategies that have been best evaluated by all evaluation methods.

2.4 Comparative types of investment

Just because an investor finds the most optimal strategy based on technical analysis does not necessarily mean that it is suitable for them. Most strategies are not as profitable over the long term as investing in, for example, index funds or commodities such as gold. Therefore, an investor should compare their found strategy against the most commonly used ways of investing applying the found strategy in practice.

Buy & Hold

The most common trading strategies in the stock market include the Buy & Hold strategy. This strategy is different from those explored in this paper. The idea behind this strategy is that the investor enters a long position as soon as they have the funds to invest and closes the position as soon as they want to move the money into another investment or out of investing altogether.

Despite the fact that this strategy is very simple to use and that only a few individuals manage to beat a Buy & Hold strategy invested in, for example, the US S&P 500 index, it does have serious downsides. The biggest disadvantage is the timing of buying and selling. Without indicators, an investor cannot tell whether they buy the peak (locally high price) or the bottom (locally low price). The same applies to situation when investor exits a position. Here the problem is made even greater by the fact that some investors cannot afford to wait until their investment is profitable if the price is currently low (due to a crisis, correction, or other reasons). However, this problem can be solved by investing for the long term, which overcomes crises and market corrections. (Murphy [1999])

Investing outside the stock market

Nowadays, an investor has many options where to invest outside the stock market. The closest thing to the stock market is the commodities market, futures or forex market. However, an investor can also invest in bonds, mortgages or, if they have enough capital, real estate.

All these methods have their own advantages and disadvantages, their own values of return and risk. However, before an investor starts to use their strategy based on technical analysis, they should make sure that there is no better investment alternative for them.

The function of comparing a strategy against other types of investing is not yet available in the program, but it is able to calculate the value of a portfolio using a Buy & Hold strategy, and not only for the stock market, but also for the commodity or forex market.

3. Practical part

3.1 Creation of the testing program

As already mentioned, the created program is a testing program that allows quick testing of technical analysis strategies on historical data. This program does not allow to execute trades, it only serves as a learning tool for investors, giving them very quick feedback on their chosen strategy.

3.1.1 Historical data

The program allows to retrieve, store and subsequently use historical data. Historical data are represented as market information on price and volume of trades. This data are obtained on the basis of a unique ticker (the designation of a given asset on the exchange). For each ticker, historical data are stored separately.

Arrangement

Before the program starts collecting historical data, the investor must choose which tickers they want to use for testing. The investor selects so-called test periods, which are selected time periods of one ticker with fixed boundaries.

The investor divides these periods into two groups, training and evaluation. Each test period is identified by a certain unique code, which consists of a code letter and a number. If the investor wants to group test periods within a group, they can label them with the same code letter (together with a different number, for example c1 and c2). This grouping allows the investor to see the results of the particular groups instead of either the overall group result or the individual test period results.

The training group is designed to tune the strategies and adjust them based on the results of this testing. This group of data can be either drawn randomly, or chosen specifically if a strategy for a particular market is of interest (for example, a strategy profiting in a declining market). The best strategies from the training phase, and indeed all candidate strategies, should still be tested by the investor using the evaluation group.

The evaluation group of test periods is used to test the strategies in a market that simulates their behaviour in a real market. The program thus avoids the creation of an “over-trained strategy”. An “overtrained strategy” is a strategy that exhibits extremely good results only in a particular market period and its application in the current market would be significantly different. The results obtained in the evaluation phase serve as a tool which helps the investor to decide which strategy they choose for market investment. No modification in strategy definitions should be made after this phase.

The selection of the test periods is made prior to testing and changing the test periods requires the elimination of all testing performed to date.

Retrieving

To retrieve historical data for the purpose of the program, there are so-called APIs, which are already pre-programmed functions that can quickly provide data. However, the APIs for historical market data are pay-per-scope and an API that meets the requirements of this project would be too expensive.

So, for the purpose of this project, an own function, which can retrieve historical data, was created. This function downloads the data from Yahoo Finance and it is capable to obtain data that are up to 50 years old. The speed is $45 \text{ sym} \cdot d/s$ (the number of days for one ticker that the program loads per second). This speed is limited by the Yahoo Finance page that does not allow more than 100 data to be loaded per page, which slows down the download.

Since many indicators need data that precede the selected time period length (one of the parameters) to calculate the value of an indicator for the selected moment, the program must also load data for some preceding time span. Therefore, the program downloads historical data for all trading days in the test period plus historical data for trading days one year before the start of the test period.

3.1.2 Strategy notation

Computers and humans do not think alike. For an investor, using a strategy means creating signals based on charts. For a computer, this is inelegant because the computer works with numbers and does not need a chart to generate signals. Therefore, it was necessary to create a unique notation that allows arbitrary strategies to be created using indicators that can be modified as if they were strategies used by person.

Detailed explanation

A generated strategy notation is enabled by so-called “strategy sentences” (hereafter referred to as sentences), which are explained in the following.

As mentioned earlier, a strategy must include exact moments when to open and close the position. Therefore, the sentences consist of two sides separated by an underscore:

$$\textit{OpeningSide_ClosingSide}$$

These two sides are the conditions that, when they hold, perform the given operation on the position (either open or close it). A change in the position occurs only if such a change is possible. This means that if the position is already open and the program receives a signal to open, nothing changes. If both conditions are true at the same time, the signal to close overrules the signal to buy and the position is closed if it has been open.

Each condition (side) consists of an inequality and a continuity coefficient (here labeled k):

$$\textit{OpeningInequality}\{k_{open}\}_ \textit{ClosingInequality}\{k_{close}\}$$

The continuity coefficient is a tool that tries to make the machine-written strategy closer to a strategy that a person would use. The value of the coefficient determines the number of time periods for which the inequality must be satisfied without interruption for the side to generate a signal.

The inequality is a classical inequality that uses - values of - indicators (*IND*) that form the inequality instead of numerical values. A page does not necessarily contain only one inequality, but can contain multiple inequalities that are joined using the logical operations AND (&&) and OR (||). An example of theoretical notation:

$$IND_1 > IND_2 \&\& IND_3 < IND_2\{k_{open}\} _ IND_1 <= IND_2\{k_{close}\}$$

The notation of the value of each indicator also has its own requirements. Each indicator value that enters the inequality starts with a hashtag (#) and ends with square brackets ([*n*]) with a number (*n*) indicating which indicator value to use (some indicators generate multiple values, for example Bollinger Bands generate a value for the upper and lower bound). The hashtag is followed by the written designation of the indicator, and the data series (*DS*) from which the indicator is calculated and the parameters (*p*) of the indicator are given in brackets. The end of bracket is followed by a square brackets containing the number *n*. Example of theoretical notation of an indicator value:

$$\#BB(DS, p_1, p_2, \dots)[n]$$

The data series is the label for the data source from which the indicator is calculated. The data source may be direct market information (price or trading volume information) or the values of another indicator. The data series is always terminated by square brackets (implying that even an indicator that is in an inequality is a data series), which in the case of indicators that contain the number *n*, or if it is market information, the square brackets are without a number. If the indicator is used as a series of data, its entry is no longer marked with a hashtag, because the hashtag indicates the numbered information for the inequality. Example of specific indicator value notations are:

$$\#BB(x[], SMA, 20, 2)[2]$$

$$\#MZB(RSI(x[], SMA, 14)[1], 50, 0.95)[1]$$

The value of the indicator can be replaced by the value of market information, so for example the closing price. Everything described so far is a summary of the newly created strategy notation, so that the program is able to read and use the strategy correctly. It is necessary to follow the rules of strategy writing (every character, every bracket has a meaning and in case of absence of some character the program may not be able to apply the strategy correctly). A simple example has been chosen to show the overall notation:

$$\#x[] >= \#EMA(x[], 10)[1]\{1\} _ \#RSI(x[], SMA, 14)[1] < \#K(x[], 50)[1]\{1\}$$

A strategy written in this way opens its position if the closing price (*x*) is greater than or equal to the Exponential Moving Average (*EMA*) of the closing prices. In addition, the strategy closes its position if the Relative Strength Index (*RSI*) for closing prices is less than a constant value (*K*) of 50.

Automatic creation

In case the investor decided to rewrite a strategy using the aforementioned notation, it would be called “a manual entry” - the investor manually enters the strategy and it gets tested. However, if he wants to test combinations of multiple indicators with different combinations of parameters, manual entry would be tedious. Therefore, the program includes the ability to automatically create a large number of strategies according to the information entered and then automatically test them.

This way of creating strategies is composed of two parts: Entering individual sides and entering parameters into each side.

In automatic strategy creation, the investor enters the individual opening and then closing sides. Each opening side is combined with each closing side.

However, the investor may not only create combinations of different sides, but also variations of a particular side. A variation of a particular side is created by the investor by omitting a place in the side notation where a parameter would normally be. There may be more than one such a place, and such places are indicated in the notation by compound parentheses ($\{\}$). For each such place, the investor then specifies a list of parameters that can be inserted. In the case of numeric parameters, the list can be specified as an arithmetic series.

The program first produces variations of each side, where each parameter is combined with each other. Once all sides are created, all opening sides are combined with all closing sides. This step produces a huge number of strategies that correspond to the strategy notation described above.

Calculation principle

Now that the strategy notation has been described, it is useful to understand how the program calculates the individual indicators, the values of the inequalities, and finally the individual signals.

The program calculates the signals for each test period separately. These signals are calculated before the simulation process itself. The principle of the calculation is as follows: first, the values of the individual indicators used in the strategy are calculated, then the inequalities are evaluated, then logical operations are performed on the truth values of the inequalities, and lastly the signals are generated from the results of the logical operations.

The most complex step of the whole process is the process of calculating the individual indicators. The result of this process is a series of values, but their computation requires calculation of additional series of values, the data series (the values from which the indicators are calculated). At this point, it is necessary to use a recursive function that is able to calculate both the indicator values and their series of data. The principle of the function is that if the data series is unknown, the function calls itself and inserts only the entry of the data series it

does not know into the new calculation. This process is repeated until only the entry requiring market information remains. This must always happen because all indicators work with market data, only this data go through a different number of calculations.

Once the values of the indicators are known, they are lined up against each other so that the values for the same point in time are opposite each other. A function is then used to determine the truth values for the inequalities and this again produces a series of values, this time truth values. These are concatenated using logical operations so that each side has only one truth value.

The final step is to create the signals. These are formed from the truth values for each side. Here, the continuity coefficient is already included. It works on the principle that the signal is generated only once in an unbroken series of truth values of 1, at the time point that corresponds to the numerical value of the coefficient. If the unbroken series of truth values of 1 is not long enough for a member corresponding in order to the coefficient value to exist at all, the series does not generate a signal. This generates a series of signals that are further used in the simulation process.

3.1.3 Trading simulation

The process of simulating market trading is very simple. The simulation process uses signals generated by the strategy and applies them to the historical market. From the results of this application, the investor is then able to judge the quality of the strategy.

At the beginning of the simulation process, the initial capital and the percentage of capital allocated to each test period is determined. The amount of capital cannot be neglected. Even though the resulting appreciation is a percentage, the initial capital plays a crucial role in the purchase of shares. In fact, the number of shares purchased is always an integer number and is rounded down. Therefore, if the initial capital is low and expensive shares are purchased, the corresponding value of shares may not be purchased, which corresponds to the realistic application of the strategy to the market if no fraction of shares can be purchased. The percentage of capital is used for money management, which causes all the capital not to be invested in a single asset without any diversification. Murphy states that the capital to open one position should not exceed 5 % Murphy [1999]. The recommended setting by author for the percentage of capital is between 3 % and 5 %.

The simulation process is carried out on a day-by-day basis for each test period. The test period starts and ends on the selected days that were determined when the test period was selected. Each day it is verified whether a signal is generated to open or close a position. If such a signal is generated and it makes sense to make a position change based on the signal (no point in opening a position when the position is already open, etc.), this change is made at the opening of the upcoming time period. In case the simulation should end with an open

position, the program does not include the last open position in the result.

The program also allows for the creation of stop-losses and if the strategy under test has a stop-loss, then when the position is closed using the stop-loss, the program does not wait for the upcoming time period and closes the position at the price corresponding to the stop-loss value.

At the end of the simulation, the output of this process is the total valuation for the test period and the number of trades that were executed. This information is further used to evaluate the strategy and assess its effectiveness.

3.1.4 Numerical evaluation

Numerical evaluation is a technique for evaluating candidate strategies. The main purpose of numerical evaluation is to determine the riskiness and effectiveness of individual strategies. The effectiveness is used to facilitate automatic objective comparison of a large number of strategies. It is necessary to distinguish between numerical evaluation and evaluation testing, which simulates the application of a strategy to the actual market by applying the strategy to random historical data.

The program calculates the risk and, based on the risk, the effectiveness of strategies in three ways. The method for computation the efficiency is the same for all calculation methods. Efficiency is always calculated as the ratio of the average annual appreciation and riskiness (as shown in the section about the coefficient of variation).

Standard deviation

One of the most common ways of determining risk is the standard deviation. Standard deviation is a statistics that which measures the variation of a distribution around its mean in terms of quadratic distances. It is calculated from the average of the annual returns and their mean as

$$\sigma = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (r_i - \bar{r})^2},$$

where

- r_i is the average of the annual return of test period i ,
- n is the number of test periods,
- \bar{r} is the mean of averages of the annual returns of test periods.

Mean absolute deviation

The mean absolute deviation is very similar to the standard deviation but differs in the way it measures the distances from the mean.

$$A = \frac{1}{n} \cdot \sum_{i=1}^n |r_i - \bar{r}|,$$

where

r_i is the average of the annual return of test period i ,
 n is the number of test periods,
 \bar{r} is the mean of averages of the annual returns of test periods.

The risk values are very similar and this method is presented here as an alternative to the standard deviation.

Coefficient of variation

The coefficient of variation is calculated as the standard deviation relative to the mean:

$$v = \frac{\sigma}{\bar{r}},$$

where

σ is the standard deviation (risk),
 \bar{r} is the mean of averages of the annual returns of test periods.

Thus, calculating the efficiency of a strategy using the coefficient of variation is similar to calculating the efficiency for the standard deviation, but the value of the mean is squared. This method gives more weight to the profit than the previous methods. Hence, the efficiency is calculated in this way after substituting for the coefficient of variation:

$$E = \frac{\bar{r}^2}{\sigma},$$

where

σ is the standard deviation (risk),
 \bar{r} is the mean of averages of the annual returns of test periods.

3.2 Proposal of new indicators

While working on this project, it is necessary to understand the functioning of technical indicators and then implement them into the program. Some of the indicators inspire the author to create his own new indicators, which are also implemented in the program and tested by the program. The obtained results show that these new indicators could be successfully used in the strategy building as they often lead to better results compared to the original commonly used indicators. Two new indicators are proposed in the succeeding two sections.

3.2.1 Middle Zone Bands

The first indicator is inspired by the Bollinger Bands and it extends them for non-normal distributions. This new indicator is mainly meant as an assistive indicator to other indicators that require the creation of boundaries where signal generation occurs when they are crossed.

Bollinger Bands

Bollinger Bands (BB) is a banded indicator that uses the idea that most of the time prices move between two bands, which are defined by price being there 95 % of the time. The indicator uses addition and subtraction of standard deviation to find these bands. If data come from a normal (Gaussian) distribution then approximately 95 % of the values lie between bounds computed as the mean minus plus the standard deviation. Economic theory then says that if the price falls outside these boundaries, it tries to get back into this range, and it does so over several time periods (Murphy [1999]). The value of the boundaries of the Bollinger Bands is then calculated using the formula:

$$BB_t(n, m) = SMA_t(n) \pm m \cdot \sigma_t(n), \quad (3.1)$$

where

n is the length of period,
 m is a tuning parameter,
 $SMA_t(n)$ is the value of the SMA (by formula 3.3) at time t with parameter n ,
 $\sigma_t(n)$ is the standard deviation at time t with parameter n calculated as

$$\sigma_t(n) = \sqrt{\frac{1}{n-1} \cdot \sum_{i=0}^{n-1} (x_{t-i} - SMA_t(n))^2}, \quad (3.2)$$

where

x_t is the closing price at time t ,
 n is the length of period,
 $SMA_t(n)$ is the value of the SMA at time t with parameter n calculated as

$$SMA(n) = \frac{1}{n} \cdot \sum_{i=0}^{n-1} x_{t-i}, \quad (3.3)$$

where

x_t is the closing price at time t ,
 n is the length of period.

As mentioned, Bollinger Bands are based on the idea that twice the standard deviation creates boundaries between which include 95 % of the data. However, this assumption is only valid if the data distribution is normal. Bollinger Bands can also be used for non-normal symmetric distributions, but the parameter m must be adjusted (if the goal is to have 95 % of the data in the band). If the distribution is asymmetric, these values are not accurate. This is the first big problem. The second problem is that if an investor wants to change the percentage of values to be between the boundaries, the investor must calculate what multiple of the standard deviation corresponds to their desired percentage of values in the band.

Proposed modification

The problems described in the previous section lead to the development of a modification of this indicator that extends its accurate use only when the data are symmetrically distributed to data that are also asymmetrically distributed. This modification has been called Middle Zone Bands and it aims to simplify the notation of the percentage of values in the band and to include asymmetric price distributions.

The whole modification is based on the creation of boundaries that are equidistant from the mean and that progressively move away from the mean until the percentage of the data lying between these boundaries is greater than or equal to an investor-specified percentage value (p). The value of both bands is calculated for each time period from the last n values, which is the value chosen by the investor. The indicator value is calculated according to the following formula:

$$MZB_t(n, p) = MA_t(n) \pm \delta, \quad (3.4)$$

where

- n is the length of period,
- p is the proportion of desired values from the interval $\langle 0; 1 \rangle$,
- $MA_t(n)$ is the value of the selected Moving Average (MA) at time t with parameter n ,
- δ is the smallest number such that

$$\frac{\#i \left(x_{t-i} \in \left(MA_t(n) - \delta, MA_t(n) + \delta \right) \right)}{n} \geq p; i \in \{0, \dots, n-1\}, \quad (3.5)$$

where

- x_t is the closing price at time t ,
- n is the length of period,
- p is the proportion of desired values from the interval $\langle 0; 1 \rangle$,
- $MA_t(n)$ is the value of the selected Moving Average (MA) at time t with parameter n .

3.2.2 Divergence

The second proposed indicator is completely new in a sense that it is not a modification of existing commonly used indicator. This new indicator is referred to as Divergence, after the phenomenon that inspired it.

When generating signals, an investor does not have to rely only on the intersection of two indicators, but can also track their mutual movement and generate signals based on it. This method of signal generation is called divergence. The investor looks at two series of data and how they move together. The series of data can be the value of an indicator or even just the price.

This method of signal generation is widely used by investors, but until now it

has not been described mathematically and thus its implementation in a program would not be possible. Therefore, for the program, this phenomenon has been simplified. First, the program calculate the ratio of the average values of the two series being monitored. Then, said ratio is multiplied by the inverse of the ratio of the current values of the two series. This has the effect that if two series being monitored behave similarly in the past then the indicator oscillates (it is an oscillator) around 0 (for simplicity one is subtracted to move the oscillating line to 0). If the two series are in the same trend, the indicator also oscillates around 0, but if the two series being tracked move in opposite directions, the indicator forms peaks, indicating a moment of divergence. These peaks are the signals that this indicator generates. Based on the theoretical signal generation of the selected series, the investor then determines whether a positive or negative peak corresponds to the opening position and vice versa. The Divergence value is calculated using a simple formula:

$$D_t(G, H, n) = \frac{MA_t^H(n)}{MA_t^G(n)} \cdot \frac{G_t}{H_t} - 1, \quad (3.6)$$

where

G, H	are data series,
G_t, H_t	are the values of the individual series at time t ,
n	is the length of period,
$MA_t^G(n), MA_t^H(n)$	are the values of the selected moving average for the values of the individual series with parameter n at time t .

3.3 Program outputs

Once the program is completed and all indicators are verified to give the correct pre-calculated results on the model data, it is necessary to test the program as a whole. To do this, one comprehensive test is carried out and a narrower evaluation from that.

This testing uses all 14 reprogrammed indicators and 2 newly designed indicators. Thus, the testing includes 22 unique opening pages and 23 unique closing pages. These produce 355 variations of opening pages and 359 variations of closing tents. After removing the sides that generate no signal, all combinations of opening and closing sides left almost 97,000 testable strategies. To give an idea, this testing process takes 21 days of net time on a single processor thread of an ordinary home computer.

The best strategies from the testing are subjectively selected by the author for evaluation. After the evaluation, the strategy that perform best during the evaluation is selected and compared against the selected investment methods described in chapter 2.4.

For example these outputs are the result of the final evaluation of different investment methods:

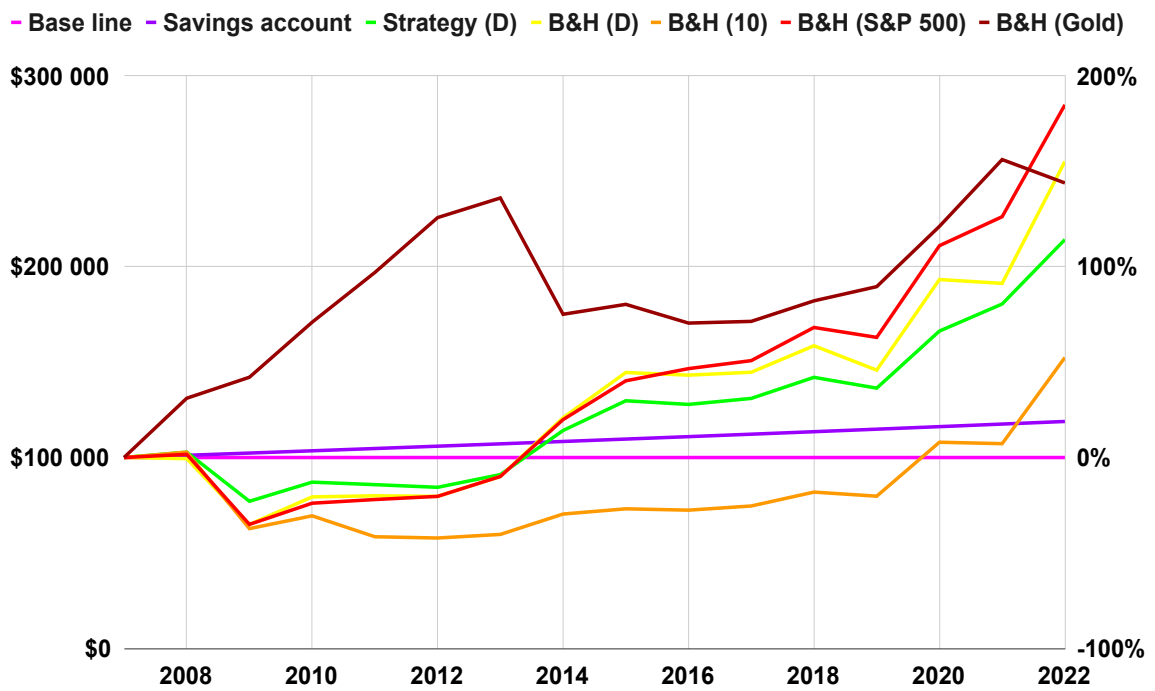


Figure 3.1: Return rates of chosen investment methods over time (in New Taiwan dollars, always as of 1 January).

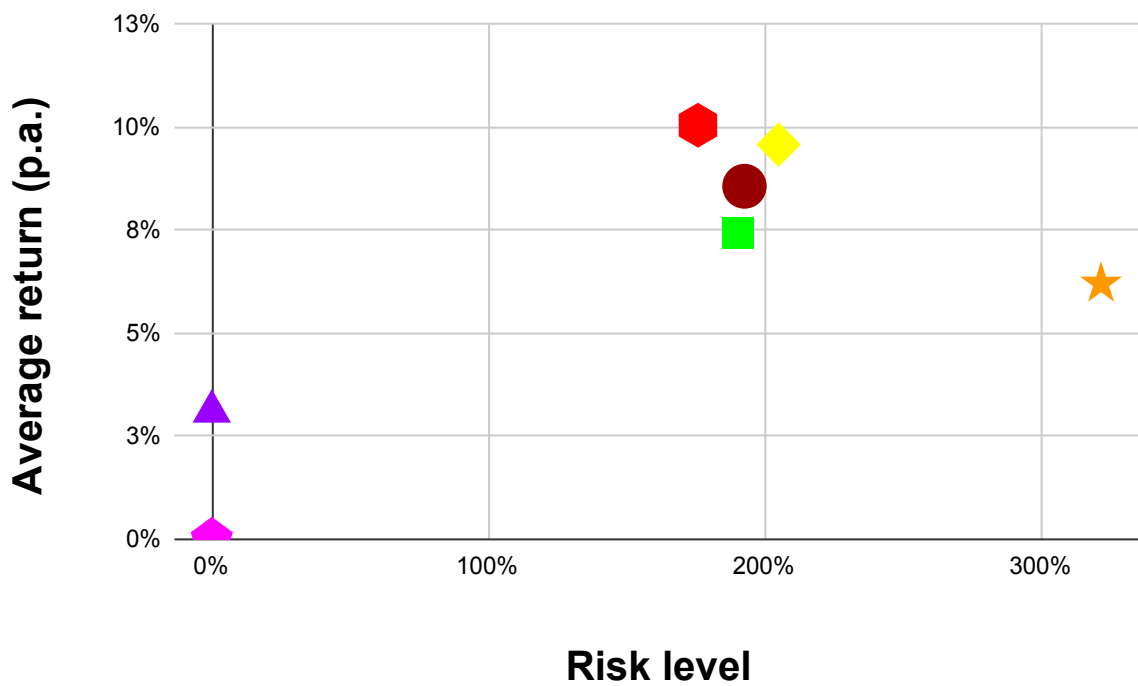


Figure 3.2: Average return rates of chosen investment methods relative to risk (after 15 years).

Conclusion

The main goal of this paper is to create a program that allows investors to quickly and accurately test strategies based on technical analysis.

This paper describes the basic information that an investor who wants to invest in the stock market needs to know. The basic idea of technical analysis and what a trading strategy based on technical analysis should look like is described.

Furthermore, the thought process that leads to the creation of the mentioned program is explained. The method of obtaining and sorting historical data is depicted. The special notation that allows the transcription of technical analysis strategies is explained. So that the program can understand these strategies thanks to the special notation. The way in which the strategies are automatically evaluated and the ways in which the program calculates the risk of each strategy are also described. Lastly, alternative methods of investing on and off the stock market are presented for comparison with the strategies.

In addition to the practical contribution of the paper, which is the program itself, the paper also contains a theoretical contribution in the form of the proposal of two new indicators that have been described here. The results of the testing of these indicators show that they are equally if not more effective than the commonly used indicators.

The last part of the paper is a demonstration of the outputs that may emerge from the said program as a result of the testing. These outputs were part of extensive testing that is not described in this paper, but the results of this testing are attached in the Attachments section.

Discussion

While the goal of creating a testing program has been met, it is not sufficient for investor education purposes. A graphical user interface (GUI) would need to be created to make the program easier to understand for investors. In this interface, the investor would be able to display the historical data in the form of a graph and to insert individual monitored indicators (which would interleave the graph) into this graph. The investor would also be able to better understand the functioning of the stock market by visualising the simulation process, which would lead to a graphical representation of the automatic trading using the chosen strategy.

The author is currently working on the next version of this program. Work is being done on the aforementioned GUI and the creation of separate server and client applications. The server application would do the automatic testing and this application would be controlled by the client (client application). For the server application, the author is also improving the computational and simulation functions, which should now support multi-threaded calculations. The client application, on the other hand, will get the possibility to graphically display historical data, individual indicators and even the progress of the simulation process.

Building on this project, a broader exploration of the new indicators introduced in this paper could also be included. To these, the author now adds another indicator that is based on the statistical phenomenon of quantiles.

The author considers the general goal of technical analysis to be the creation of a universal strategy that would be able to have a higher return and lower risk in the long run than the Buy & Hold strategy applied to indices (e.g. the S&P 500). The search for such a strategy is something the author intends to pursue in the future.

Bibliography

Richard J. Bauer and Julie R. Dahlquist. *Technical Market Indicators*. Wiley, 1999. ISBN 9780471197218.

Milan Galik and Paul K. Brody. Annual Report 2021, 2022. URL https://investors.interactivebrokers.com/download/2021_IBG_AR.pdf.

Tomáš Krutina. Analýza vybraných indikátorů technické analýzy v době propadů na akciových trzích. Master's thesis, University of West Bohemia in Pilsen, Faculty of Applied Sciences, 2021.

John J. Murphy. *Technical Analysis of the Financial Markets*. New York Institute of Finance, 1999. ISBN 0735200661.

Charles Schwab. The Rise of the Investor Generation, 2021. URL <https://www.aboutschwab.com/generation-investor-study-2021>.

Richard J. Teweles and Edward S. Bradley. *The Stock Market*. Wiley Sons, Incorporated, John, 2008. ISBN 9780470342428.

Yahoo! Yahoo Finance, 2022. URL <https://finance.yahoo.com/>.

List of Figures

3.1	Return rates of chosen investment methods over time (in New Taiwan dollars, always as of 1 January).	22
3.2	Average return rates of chosen investment methods relative to risk (after 15 years).	22

Attachments

Due to the large amount of data, strategies, indicators and evaluations of all testing, tables containing results of the calculations are not suitable to be inserted directly into this report. To view (and possibly download) all the tables contained in the database, the database itself, the program used and its source code, it is necessary to access Google Drive, where all the mentioned parts of the paper are stored and made available. On the aforementioned Google Drive, the table containing the values from which the resulting graphs are created is also attached. For this reason, a link to the relevant Google Drive is provided - see below. For easier transcription in case the paper is in a printed form and the link cannot be copied, a shortened link to the mentioned Google Drive is also included here. If the shortened link does not work, please use the original link.

Google Drive link:

<https://drive.google.com/drive/folders/1VNZjkytSEQxmmxw7eevSyAia9V0uIT2Z?usp=sharing>

Short link to Google Drive:

<https://tinyurl.com/2p8ep8nn>

【評語】 130013

This is an excellent piece of work. The applicant has a talent in mathematical thinking. Notably , this project is well-motivated with a goal for everyday people to invest in stock market without losing money , by providing free-of-charge technical analysis to anyone who may be interested. Although there is a lack of specific scientific hypothesis in this project , the applicant provides a solid mathematic method to predict and prevent people losing money from the stock market. Another merit to this project is that the applicant's presentation style is very clear , well-rounded and enthusiastic. The applicant addressed the reviewers' questions very well.