Intellectual Asset

Original initiative

# An example of research oriented skill acquisition and knowledge discovery process for developing leadership in entrepreneurial initiatives

Ankit Gupta (B.E. | P.G.D.B.A. | S.C.C.B.M.)

2021

**Abstract**

Indian enterprises lack research orientation and more so at being original from a global leadership perspective. Through one of my efforts over last few years, I demonstrate that through research it is possible to acquire new skills, knowledge and develop leadership in business initiatives. Although at an individual level this only requires time, effort and perseverance. At an institutional level it will require much more.

1. ABB India Limited is used for all examples related to a specific stock as it is the first stock in-terms of alphabetical order on BSE. Its scrip code is 500002.

2. The words, “Stock”, “Scrip”, “Counter” and “Equity” are used interchangeably to refer to common company stock listed on BSE.

3. “Investing” and “Trading” are used interchangeably. As the final strategy can be seen as long term trading or short term investing. However in all most all of the cases the time between buying and selling is more than 1 day (But theoretically it can be within one day as well).

# Contents



[1] Internet Link

[2] Stock return versus stock prices: stock return = (stock price in the future - stock price today) / stock price today

[3]Not needed in version 8.0.0 and above
[4]Not needed in version 8.0.0 and above

# List of Algorithms

# List of Figures

# List of Tables

# Part I

# How do stock prices fluctuate ?

## Abstract

**Everyone knows why stock prices fluctuate**, but only a few people know how they fluctuate. A quick search with the term <How do stock prices fluctuate?> on Google shows 176 top results out of 69,30,000. Not a single result answers the question. A quick search with the term <"how" do stock prices fluctuate? -why -cause -reason> on Google shows 179 top results out of 25,30,00,000. Few of these results come close to identifying the question accurately, but none answer it. Few who came close talked about predicting stock prices. These results described potential methods that can be used to predict stock prices and how they failed to do so. So, I infer three things from this, one, all those people who have tried to explain how stock prices fluctuate have failed to do it and shared their findings publicly; and second, those who have succeeded have not disclosed it in public. Lastly, there is a lot of interest in this topic or question. Either way, this makes the pursuit of answering this question contemporary and possibly both risky and rewarding. I investigated further and deducted that these **initiatives fail as they try to predict instead of just explaining.** As you may be aware, **to explain any phenomenon, a model is usually created.** Depending on the purpose and end-use, a model can be extremely simple or highly complex. Also, the nature of a model is the same as that of the phenomenon. Therefore, since stock price fluctuations have a quantitative nature, its model is also quantitative in nature. That is to say that the model is composed of numbers. **This part explains an original approach to model price fluctuations** of more than 700 stocks listed on Indian stock exchanges.

As far as finance, accounting and economics is concerned a quantitative model is known as an econometric[5] model[6]. Econometric models for explaining price fluctuations of stocks[7] and other assets based on stocks, can be built through two different analytic approaches. First being fundamental analysis[8] , this approach explains the reasons behind price fluctuations. Second being **technical analysis** [9] , this approach explains how the actual price fluctuation happened.

If the stock market was perfectly efficient[10] , fundamental analysis would have been enough for making sound investments. On the other hand, If one knows about the nature of price fluctuations, one does not need to do fundamental analysis in order to profit from the fluctuation. **This part also explains this peculiar nature of stock price fluctuations in the Indian stock market.**


---

[5] Econometrics is the application of statistical methods to economic data in order to give empirical content to economic relationships. More precisely, it is "the quantitative analysis of actual economic phenomena based on the concurrent development of theory and observation, related by appropriate methods of inference". A basic tool for econometrics is the multiple linear regression model. Econometric theory uses statistical theory and mathematical statistics to evaluate and develop econometric methods.

[6] Econometric models are statistical models used in econometrics. An econometric model specifies the statistical relationship that is believed to hold between the various economic quantities pertaining to a particular economic phenomenon. An econometric model can be derived from a deterministic economic model by allowing for uncertainty, or from an economic model which itself is stochastic. However, it is also possible to use econometric models that are not tied to any specific economic theory.

[7] A stock (also known as equity) is a fungible and negotiable package of capital that can be traded with itself or with other financial assets as it holds some monetary value. It represents the ownership of a fraction of a corporation. This entitles the owner of the stock to a proportion of the corporation's assets and profits equal to how much stock they own. Units of stock are called "shares."

[8] Fundamental analysis, in accounting and finance, is the analysis of a business's financial health, competitors and markets. It also considers the overall state of the economy and factors including interest rates, production, earnings, employment, GDP, housing, manufacturing and management.

[9] In finance, technical analysis is an analysis methodology for forecasting the direction of prices through the study of past market data, primarily price and volume.

[10] Market efficiency - championed in the efficient market hypothesis (EMH) formulated by Eugene Fama in 1970, suggests that at any given time, prices fully reflect all available information on a particular stock and/or market. According to the EMH, no investor has an advantage in predicting a return on a stock price because no one has access to information not already available to everyone else.

# Chapter 1

# Industry significance

After having worked on this initiative for over two years I came across a competition on Kaggle[1].

e-mail from Kaggle to Me:

How efficient is the financial market[a]? Computer-based algorithms and models have long been used to identify inefficiencies. Now you get to try it for yourself. In this **competition hosted by Jane Street**[b], you will build a model to determine which of a number of potential trading opportunities to execute in order to maximize trading profit. Use your data science skills to see how quantitative researchers make more informed choices about which trading opportunities to execute. In the first three months of this challenge, you will **build your own quantitative trading model** to maximize returns using market data from a major global stock exchange. Next, your models will be tested against future market returns that will be reflected on the leaderboard.
**Total Prizes: $100,000** | Entry Deadline: February 22, 2021
It is impossible to perfectly "solve" the financial market, but you can look for an edge while exploring a real-world data science problem that quantitative researchers face every day.
Good luck,
Will Cukierski
Kaggle Competitions

[a]See 2.2 on page 3

[b]Jane Street Capital, typically referred to as Jane Street, is a global proprietary trading firm **with around 1,200 employees**. The company is one of the world's largest market-makers, trading more than $17 trillion worth of securities in 2020. In July 2020, S&P Global Ratings affirmed Jane Street on capital growth. The rating agency noted that Jane Street is a "highly profitable trading business," that "the company has generated very strong earnings so far in 2020 and that its trading has benefited from the market volatility related to the COVID-19 pandemic. Jane Street's trading is based on its own proprietary models. Quantitative analysis and insights into related markets enable it to make competitive markets in even the most complicated products. Technology is at the core of how it approaches trading, and it considers itself as much a technology company as a trading firm.

[1]A **subsidiary of Google** LLC, is an online community of data scientists and machine learning practitioners. Kaggle **allows users to** find and publish data sets, explore and **build models** in a web-based data-science environment, work with other data scientists and machine learning engineers, and **enter competitions to solve data science challenges**.

**"Buy low, sell high." It sounds so easy. . . .**

**In reality, trading for profit has always been a difficult problem to solve**, even more so in today's fast-moving and complex financial markets. Electronic trading allows for thousands of transactions to occur within a fraction of a second, resulting in nearly unlimited opportunities to potentially find and take advantage of price differences in real time. In a perfectly efficient market, buyers and sellers would have all the agency and information needed to make rational trading decisions. As a result, products would always remain at their "fair values" and never be undervalued or overpriced. However, financial markets are not perfectly efficient in the real world.

**Developing trading strategies to identify and take advantage of inefficiencies is challenging**. Even if a strategy is profitable now, it may not be in the future, and market volatility makes it impossible to predict the profitability of any given trade with certainty. As a result, it can be hard to distinguish good luck from having made a good trading decision. In the first three months of this challenge, you will build your own quantitative trading model to maximize returns using market data from a major global stock exchange. Next, you'll test the predictiveness of your models against future market returns and receive feedback on the leaderboard. Your challenge will be to use the historical data, mathematical tools, and technological tools at your disposal to create a model that gets as close to certainty as possible.

**If one is able to generate** a highly predictive model which selects **the right trades to execute**, they would also be playing an important role in sending the market signals that push prices closer to "fair" values. That is, a better model will mean the market will be more efficient going forward.

However, **developing good models will be challenging** for many reasons, including a very low signal-to-noise ratio, potential redundancy, strong feature correlation, and difficulty of coming up with a proper mathematical formulation.

Jane Street has **spent decades developing their own trading models** and machine learning solutions to identify profitable opportunities and quickly decide whether to execute trades. These models help Jane Street trade thousands of financial products each day across 200 trading venues around the world. Admittedly, this challenge far oversimplifies the depth of the quantitative problems Jane Streeters work on daily, and Jane Street is happy with the performance of its existing trading model for this particular question. However, there's nothing like a good puzzle, and this challenge will hopefully serve as a fun introduction to a type of data science problem that a Jane Streeter might tackle on a daily basis.

So if a Google subsidiary is conducting a competition of this scale to find new models and is sponsored by Jane street at a tune of $100,000/- then you can understand how significant this initiative is to the industry.

# Chapter 2

# Concepts used

## 2.1 Theoretical asset pricing models

The fundamental value of an asset is different from its observable prices in the market. The price that can be observed is the market price. The market price is determined by demand and supply of the asset and can therefore deviate from the fundamental value, but in the long run should converge to the fundamental value. However, this is seldom the case, by the time market price catches to the fundamental value either from below or from the above, the fundamental value changes. For example, lets say for a traded common stock of a listed company named XYZ Limited; the fundamental value is estimated to be Rs.100/- on a day when it traded at a market price of Rs.300/-, lets assume this day to be First January 2021. Now on First February 2021, the same stock is traded at a market price of Rs.250/- while its fundamental value is unchanged. Assuming a scenario, that on First March 2021, the same stock is trading at Rs.100/-; but its fundamental value now is estimated to be Rs.40/-. Keeping in mind the above, a myriad set of theoretical asset pricing models have been developed by economists. These models try to estimate / evaluate the price of an asset based on different parameters / factors / reasons.

One of the many reasons

at-least in theory

discounting the fact that market prices are in multiples of fundamental value

a multiple of 3

a multiple of 2.5

a multiple of 2.5

## 2.2 Market efficiency[1]

Market efficiency was championed in the efficient market hypothesis (EMH) formulated by Eugene F Fama[2] in 1970, suggests that at any given time, prices fully reflect all available information on a particular stock and/or market. According to the EMH, no investor has an advantage in predicting a return on a stock price because no one has access to information not already available to everyone else.

### 2.2.1 Random variables

Statistically speaking, a random variable is a variable whose future value is to some extent unknown. Only way to describe the future value of a random variable is in terms of a probability distribution that governs the said random variable. Future stock prices are random variables.

[1] Internet Link
[2] Internet Link

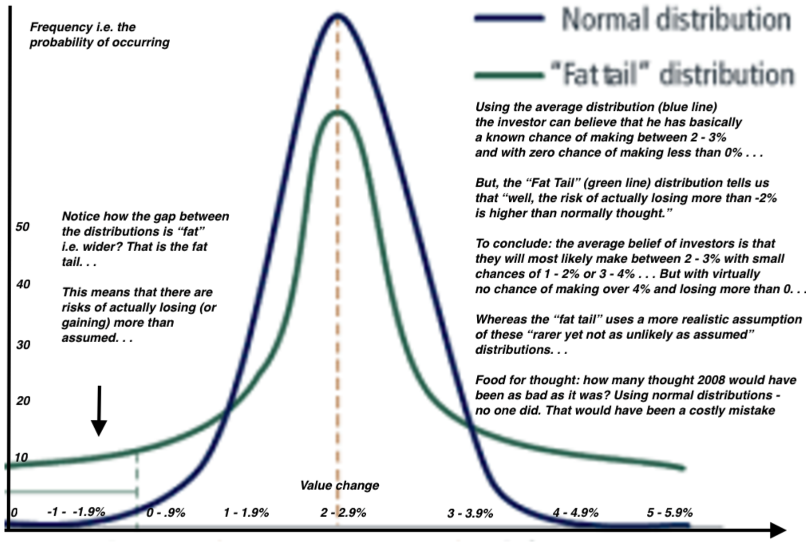
### 2.2.2 Normal distribution

It is common the see the phrase "the normal distribution", the term "normal" in fact refers to a whole family of probability distributions. The two parameters used to distinguish one normal distribution from another normal distribution are their means and standard deviations. Normal distributions are two-parameter distributions; knowledge of the mean and the standard deviation of a normal distribution is sufficient to completely characterize the distribution.

### 2.2.3 Stock prices and returns[3]

In real-world data analysis, the mean and standard deviation of a random variable are almost never known, but rather must be estimated from a sample. Furthermore, the type of distribution that generated a sample is also unknown. Distributions of stock returns are "fat-tailed" relative to a standard normal distribution. That is, the distribution is characterized with frequencies of large positive and large negative returns are higher than what would be expected from a standard normal distribution.

Figure 2.1: Probability distributions




In more vivid terms, if daily returns[4] of any stock were from a standard normal distribution, then a daily return of more than four standard deviations would be observed once every 50 years. However, such returns are observed is stocks four times in every five years[5].

### 2.2.4 Efficient market hypothesis

The efficient market hypothesis presented by Eugene F Fama assumes the following:

- The proposed model description* assumes that the "market" assesses the probability distributions and the "market" sets prices. Which in-turn will represent a completely accurate view of the world if all the individual participants in the market

*Of price formation for a particular stock in an efficient market

[3] Stock return versus stock prices: stock return = (stock price in the future - stock price today) / stock price today

[4] Percentage change in price over single trading day

[5] On an average in a study that explored such returns in the US stock market

- have same information
- agree upon the implications of this information on the joint distribution of future prices

He, further states that neither of the above assumptions are completely descriptive nor are they realistic. This model, he states, is a simplistic view of the market and should be used to understand complex phenomenon.

## 2.3 Random walk theory

Random walk theory believes that it is impossible to outperform the market without assuming additional risk. This theory considers technical analysis undependable because a technical analyst would only buy or sell a security after a trend has fully developed. Likewise, the theory also finds fundamental analysis undependable due to poor quality of information collected which is often misinterpreted. Critics of the theory contend that stocks do maintain price trends over time – in other words, that it is possible to outperform the market by carefully selecting entry and exit points for equity investments.

The random walk theory raised many eyebrows in 1973 when author Burton Malkiel coined the term in his book "A Random Walk Down Wall Street."[6] This would preclude anyone from exploiting mispriced stocks consistently because price movements are mostly random and driven by unforeseen events.

This theory concluded that, due to the short-term randomness of returns, investors would be better off investing in a passively managed, well-diversified fund. A controversial aspect of Malkiel's book theorized that "a blindfolded monkey throwing darts at a newspaper's financial pages could select a portfolio that would do just as well as one carefully selected by experts."[7]

### 2.3.1 Random walk theory in action

The most well-known practical example of random walk theory occurred in 1988 when the Wall Street Journal sought to test Malkiel's theory by creating the annual Wall Street Journal Dartboard Contest, pitting professional investors against darts for stock-picking supremacy. Wall Street Journal staff members played the role of the dart-throwing monkeys.[8] After more than 140 contests, the Wall Street Journal presented the results, which showed the experts won 87 of the contests and the dart throwers won 55. However, the experts were only able to beat the Dow Jones Industrial Average (DJIA) in 76 contests. Malkiel commented that the experts' picks benefited from the publicity jump in the price of a stock that tends to occur when stock experts make a recommendation. Passive management proponents contend that, because the experts could only beat the market half the time, investors would be better off investing in a passive fund that charges far lower management fees.

## 2.4 Variance and covariance

The expected value of a random variable gives a crude measure of the "center of location" of the distribution of that random variable. To refine the picture of a distribution distributed about its "center of location" we need some measure of spread (or concentration) around that value. The simplest measure to calculate for many distributions is the variance.

[6] Norton. "A Random Walk Down Wall Street." Accessed Jan. 22, 2021.

[7] Forbes. "Any Monkey Can Beat the Market." Accessed Jan. 22, 2021.

[8] The Wall Street Journal. "Journal's Dartboard Retires After 14 Years of Stock Picks." Accessed Nov. 17, 2020.

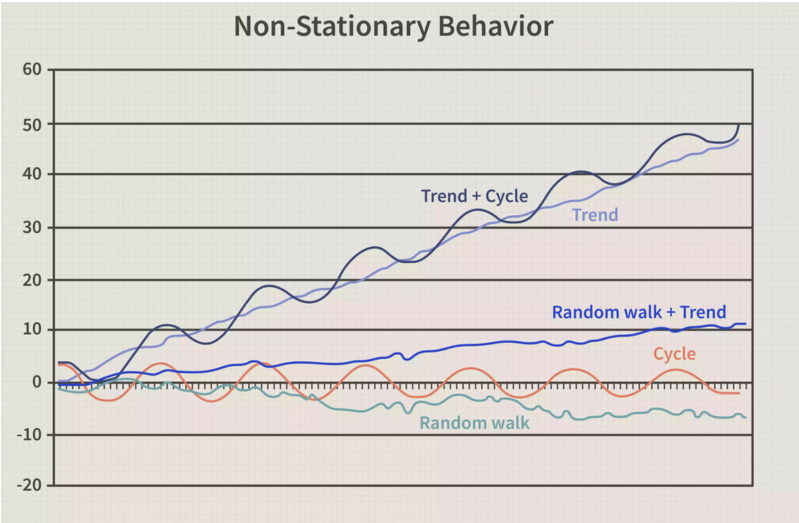
Therefore, Variance is used in statistics to describe the spread between a data set from its mean value. So the larger the variance, the larger the distance between the numbers in the set and the mean. Conversely, a smaller variance means the numbers in the set are closer to the mean. Along with the above statistical definition, the term variance can also be used in a financial context. Many stock experts and financial advisers use a stock's variance to measure its volatility. A covariance, on the other hand, refers to the measure of how two random variables will change when they are compared to each other. In a financial or investment context, though, the term covariance describes the returns on two different investments over a period of time when compared to different variables. These assets are usually marketable securities in an investor's portfolio, such as stocks. A positive covariance means both investments' returns tend to move upward or downward in value at the same time. An inverse or negative covariance, on the other hand, means the returns will move away from each other. So when one rises, the other one falls.

## 2.5 Stationary versus non-stationary data

Stock market data points are non-stationary[>]. Non-stationary data points can have trends, cycles, random walks, or combinations of the three. Non-stationary data, as a rule, are unpredictable and cannot be perfectly modeled or have accurate forecast. The results obtained by using non-stationary time series may be spurious in that they may indicate a relationship between two variables where one does not exist. In order to receive consistent, reliable results, the non-stationary data needs to be transformed into stationary data. In contrast to a non-stationary process which generates non-stationary data, a stationary process generates data which reverts around a constant long-term mean and has a constant variance independent of time.

[>] Their means, variances, and co-variances change over time

Figure 2.2: Non-Stationay Process




### 2.5.1 Trend and difference stationary

A random walk with or without a drift can be transformed to a stationary process by differencing and then the process becomes difference-stationary. The disadvantage of differencing is that the process loses one observation each time the difference is taken. A non-stationary process with a deterministic trend becomes stationary after removing the trend, or detrending. No observation is lost when detrending is used to transform a non-stationary process to a stationary one. In the case of a random walk with a drift and deterministic trend, detrending can remove the deterministic trend and the drift, but the variance will continue to go to infinity. As a result, differencing must also be applied to remove the stochastic trend. Sometimes the non-stationary series may combine a stochastic and deterministic trend at the same time and to avoid obtaining misleading results both differencing and detrending should be applied, as differencing will remove the trend in the variance and detrending will remove the deterministic trend.

## 2.6 Stochastic modeling

Stochastic modeling is a form of mathematical modeling. This type of modeling uses random variables. Stochastic modeling presents data that account for certain levels of unpredictability or randomness. Stochastic modeling is opposite to deterministic modeling. Stochastic modeling, is inherently random, and the uncertain factors are built into the model.

## 2.7 Spectral / Fourier analysis

Over the past 200 years, Fourier analysis has made fundamental contributions to fields ranging from signal processing, communications, and neuroscience, to partial differential equations, astronomy, and geology. However, in contrast to its modern ubiquity, its origins stem from a very specific problem—modeling the orbits of celestial bodies. The seeds of the theory were sown in the mid-18th century when the mathematicians Leonhard Euler, Joseph-Louis Lagrange, and Alexis Clairaut observed that orbits could be approximated as linear combinations of trigonometric functions, i.e., sines and cosines. In fact, to estimate the coefficients from the data, Clairaut published the first explicit formulation of the Descrete Fourier Transform (DFT) in 1754, 14 years before Jean-Baptiste Joseph Fourier was born. While studying the orbit of the asteroid Pallas in 1805, Carl Friedrich Gauss discovered a computational shortcut. His calculation, which appeared posthumously as an unpublished paper in 1866, was the first clear use of the Fast Fourier Transform (FFT)—an efficient way to compute the DFT. Gauss' algorithm was largely forgotten until it was independently rediscovered in a more general form almost a century later by James Cooley and John Tukey in 1965.

Nourished by this half century of progress, the theory blossomed when Fourier presented his seminal paper on heat conduction to the Paris Academy in 1807. In his treatise, Fourier claimed that **any arbitrary function could be represented by the superposition of trigonometric functions**. This broader claim was initially received with much skepticism, and it would take another 5 years before the Paris Academy awarded his paper the grand prize in 1812. Despite the award, the Academy's panel of judges, which included Lagrange, Laplace and Legendre, still held reservations about the rigor of his analysis, especially in relation to the challenging question posed by convergence. Further advances by Dirichlet, Poisson, and Riemann addressed these subtle issues, and provided the foundation for today's Fourier transform, upon which many modern mathematical applications are based[9].

### 2.7.1 The Fourier transform

In particular, one of the most structurally revealing analyses that can be performed on a time series is to express its values as a linear combination of trigonometric functions.

[9] Briggs and Henson, 1995

This procedure relies on the Discrete-Time Fourier Transform (DTFT), and **allows the data to be transformed to the frequency domain**. Since the Fourier transform is a unitary operator that changes the basis function representation of a time series from impulses to sinusoids, Parseval's theorem states that, when represented as a vector, the Euclidean length of the time series is preserved under the transformation (with proper normalization). This observation forms the foundation of **spectral decomposition**, and provides a method to visualize the data in the frequency domain. This representation, known as the power spectrum, characterizes **how much of the variability in the data comes from low- versus high-frequency fluctuations**.

### 2.7.2 The power spectrum

In many situations, a time series can be modeled as the realization of a stochastic process, which can often be characterized by its first and second moments. The DTFT of the auto- and cross-covariance functions can then be interpreted as the frequency distribution of the power contained within the variance and covariance of these time series, respectively. Similarly, the inverse DTFT can be used to find the lagged second moments as functions of the auto- and cross-power spectra.

### 2.7.3 The business cycle

One of the most natural applications of spectral analysis is to measure the business cycle, which many studies have done[10].

## 2.8 Elliott wave principle

The Elliott Wave Theory was introduced by Ralph Nelson Elliott[11] which was inspired by the Dow Theory[12] and by observations found throughout nature. Elliott concluded that the movement of the **stock market could be predicted by observing and identifying a repetitive pattern of waves**. In fact, Elliott believed that all of man's activities, not just the stock market, were influenced by these identifiable series of waves. Elliott based part his work on the **Dow Theory, which also defines price movement in terms of waves**, but Elliott discovered the fractal nature of market action. Thus Elliott was able to analyze markets in greater depth, identifying the specific characteristics of wave patterns and making detailed market predictions based on the patterns he had identified. The Elliott Wave Theory describes the stock market's behavior as a series of waves up and another series of waves down to complete a market cycle. Those cycles are grouped into eight waves, with five of those following the main trend, and three being corrective trends. After the eight moves are made, the cycle is complete.

## 2.9 Fractals

### 2.9.1 Fractal mathematics

Fractal, in mathematics, is a class of complex geometric shapes that commonly have "fractional dimension," a concept first introduced by the mathematician Felix Hausdorff in 1918. Fractals are distinct from the simple figures of classical, or Euclidean, geometry[>]. They are capable of describing many irregularly shaped objects or spatially nonuniform

[>]the square, the circle, the sphere, and so forth

[10]Granger and Hatanaka, 1964; King and Watson, 1996; Baxter and King, 1999
[11]Elliott, R.N. 1938.
[12]Nelson,S. The A B C of stock speculation, 1902. Burlington, VT: Fraser Publishing Co., 1964 reprint.

phenomena in nature such as coastlines and mountain ranges. The term fractal, derived from the Latin word fractus ("fragmented," or "broken"), was coined by the Polish-born mathematician Benoit B. Mandelbrot. Although the key concepts associated with fractals had been studied for years by mathematicians, and many examples, such as the Koch or "snowflake" curve were long known, Mandelbrot was the first to point out that fractals could be an ideal tool in applied mathematics for modeling a variety of phenomena from physical objects to the behavior of the stock market. Since its introduction in 1975, the concept of the fractal has given rise to a new system of geometry that has had a significant impact on such diverse fields as physical chemistry, physiology, and fluid mechanics. Many fractals possess the property of self-similarity, at least approximately, if not exactly. A self-similar object is one whose component parts resemble the whole. This reiteration of details or patterns occurs at progressively smaller scales and can, in the case of purely abstract entities, continue indefinitely, so that each part of each part, when magnified, will look basically like a fixed part of the whole object. In effect, a self-similar object remains invariant under changes of scale—i.e., it has scaling symmetry. This fractal phenomenon can often be detected in such objects as snowflakes and tree barks. All natural fractals of this kind, as well as some mathematical self-similar ones, are stochastic, or random; they thus scale in a statistical sense. Fractal geometry with its concepts of self-similarity and noninteger dimensionality has been applied increasingly in statistical mechanics, notably when dealing with physical systems consisting of seemingly random features. For example, fractal simulations have been used to plot the distribution of galaxy clusters throughout the universe and to study problems related to fluid turbulence. Fractal geometry also has contributed to computer graphics. Fractal algorithms have made it possible to generate lifelike images of complicated, highly irregular natural objects, such as the rugged terrains of mountains and the intricate branch systems of trees[13].

### 2.9.2 Fractal Analysis

Overall, the fractal dimension gives an indication of the average irregularity series. However, there are various fractal dimensions. They are calculated from autocorrelations, variogram values, energy spectra of the Fourier decomposition or the wavelet decomposition, and by using many other techniques. Each calculation algorithm presents both constraints and advantages, which are set out in specialist works. Although the fractal dimension is an overarching indicator of irregularity of a series that is arranged over time, the multi-fractal approach supplies both an overall view and a local view of all irregularities, at every level. Multi-fractals do not summarize information relating to "average" irregularity, but they provide full information about the singularities of a curve. It is therefore possible to follow how singularities emerge, intensify or disappear at a given scale. Various algorithms, set out within specialist works, make it possible to plot several spectra. We should simply remember that it becomes possible to observe the time distribution of irregularities. They are sometimes clear on both small and large scales, which is attested by a spectrum of asymmetrical singularities. On the other hand, a symmetrical spectrum indicates that singularities are also distributed within the various scales[14].

### 2.9.3 Fractal Indicator

The fractal indicator is based on a simple price pattern that is frequently seen in financial markets. Outside of trading, a fractal is a recurring geometric pattern that is repeated

[13] Britannica, The Editors of Encyclopaedia. "Fractal". Encyclopedia Britannica, Invalid Date, https://www.britannica.com/science/fractal. Accessed 5 June 2021.

[14] André Dauphiné, Chapter 2, Geographical Models with Mathematica, Pages 19-72, ISBN 9781785482250

on all time frames. From this concept, the fractal indicator was devised. The indicator isolates potential turning points on a price chart. It then draws arrows to indicate the existence of a pattern. The bullish fractal pattern signals the price could move higher. A bearish fractal signals the price could move lower. Bullish fractals are marked by a down arrow, and bearish fractals are marked by an up arrow. The fractal indicator will generate signals frequently. The existence of a fractal isn't necessarily important since the pattern is so common. The fractal is indicating the possibility of a trend change. This is because fractals are essentially showing a "U-shape" in price. A bearish fractal has the price moving upward and then downward, forming an upsidedown U. A bullish fractal occurs when the price is moving down but then starts to move up, forming a U. Because fractals occur so frequently, and many of the signals aren't reliable entry points, fractals are typically filtered using some other form of technical analysis. Bill Williams also invented the alligator indicator which isolates trends. By combining fractals with trend analysis, a trader may decide to only trade bullish fractals signals while the price trend is up. If the trend is down they may take only short trades on bearish fractal signals, for example. Fractals could also be used with other indicators, such as pivot points or Fibonacci retracement levels. A fractal is only acted on if it aligns with one of these other indicators and potentially the longer-term price direction[15].

[15] https://www.investopedia.com/terms/f/fractal.asp

# Chapter 3

# Technical analysis

Technical analysis is an attempt to understand price movements of securities based on historical prices and volumes rather than underlying company fundamentals, political events, and economic factors. Technical analysts believe in chart analysis to look for some significant information. Several chart analysis techniques are considered as analysis tools. There are three main popular charting techniques:

- Bar charts,
- Point-and-figure charts, and
- Candlestick charts[1].

These charts are based upon numbers which are analyzed mathematically, statistically, scientifically and pseudo-scientifically. The quest for a quantitative statistical description of stock prices began more than a century ago with Bachelier's thesis[2] in which he described price movements by a random walk[3]. Although Bachelier's random walk has since been modified, the effort to model stock price movements by a stochastic process[4] continues[5] . Jack D. Schwager, a board member of Fund Seeder and author of several books on technical analysis, uses the term "normalized" to describe stochastic oscillators that have predetermined boundaries, both on the high and low sides. An example of such an oscillator is the relative strength index (RSI)—a popular momentum indicator used in technical analysis—which has a range of 0 to 100. It is usually set at either the 20 to 80 range or the 30 to 70 range.

which depended on price action

[1]Person, J.L., A complete guide to technical trading tactics : how to profit using pivot points, candlesticks & other indicators. 2004, Canada: John Wiley & Sons.

[2]L. Bachelier, 'Théorie de la spéculation' [Ph.D. thesis in mathematics],Annales Scientifiques de l'Ecole Normale Supérieure III-17,21-86 (1900)

[3]Random walk theory suggests that changes in stock prices have the same distribution and are independent of each other. Therefore, it assumes the past movement or trend of a stock price or market cannot be used to predict its future movement. In short, random walk theory proclaims that stocks take a random and unpredictable path that makes all methods of predicting stock prices futile in the long run.

[4]In the late 1950s, George Lane developed stochastics, an indicator that measures the relationship between an issue's closing price and its price range over a predetermined period of time. The premise of stochastics is that when a stock trends upwards, its closing price tends to trade at the high end of the day's range or price action. Stochastics is used to show when a stock has moved into an overbought or oversold position.

[5]Price action refers to the range of prices at which a stock trades throughout the daily session. For example, if a stock opened at Rs.10, traded as low as Rs.9.75 and as high as Rs.10.75, then closed at Rs.10.50 for the day, the price action or range would be between Rs.9.75 (the low of the day) and Rs.10.75 (the high of the day). Conversely, if the price has a downward movement, the closing price tends to trade at or near the low range of the day's trading session.

## 3.1 Transforming time into frequency

It is somewhat intuitive to relate to the fact that stock prices˃ are usually reported with respect to time. Example, *"SENSEX jumped 2.13% this week"*, *"analysts believe SENSEX to touch 52000 within next six months"* and so on. But˃, what if they were reported like *"SENSEX's volatility is currently 0.27 Hertz"*, *"analysts believe that SENSEX has a future half yearly volatility of 0.94 Hertz"* and so on. Analysis based on this assumption is what transforming price time series into a frequency domain is all about˃.

˃As they are reported

˃Counter intuitive to imagine a frequency based reporting

˃From the author's perspective

The frequency domain has long been part of economics[6] , and spectral theory[7] has also been used in finance to derive theoretical pricing models for derivative securities[8] . However, econometric[9] and empirical applications of spectral analysis have been less popular in economics and finance, in part because economic time series are rarely considered stationary[10]. However, there has been a recent rebirth of interest in economic applications in response to modern advances in nonstationary signal analysis[11] .

Spectral and co-spectral power[12], often calculated using either the Fourier[13] or wavelet[14] transform, provide a natural way to study the cyclical components of variance and covariance[15], two important measures of risk in the financial domain. Specifically, spectral power decomposes the variability of a time series resulting from fluctuations at a specific frequency, while cospectral power decomposes the covariance between two real-valued time series, and measures the tendency for them to move together over specific time horizons.

In a recent empirical study, Chaudhuri and Lo (2015) perform a spectral decomposition of the U.S. stock market and individual common stock returns over time. They noticed that measures related to risk and co-movement varied not only across time, but also across frequencies over time. Recently, wavelets[16] and other transforms[17] have also been used to study financial data in the time-frequency domain, and depending on the specific context, these alternative techniques can provide substantial benefits in terms of implementation.

The work done by Agrawal et al.[18] utilizes the Discrete Fourier Transform(DFT)[19] to reduce the time series dimensions[20]. However, other techniques are suggested including Singular Value Decomposition[21] (SVD)[22] and the Discrete Wavelet Transform[23]

[6] Granger and Hatanaka, 1964; Engle, 1974; Granger and Engle, 1983; Hasbrouck and Sofianos, 1993

[7] **Spectral analysis** determines how variance or covariance distributes over frequency for a finite time series data set.

[8] Linetsky, 2002; Linetsky, 2004a; Linetsky, 2004b; Linetsky, 2008

[9] The application of statistical methods to the study of economic data and problems

[10] See section 2.5 on page 6

[11] Baxter and King, 1999; Carr and Madan, 1999; Croux, Forni, and Reichlin, 2001; Ramsey, 2002; Crowley, 2007; Huang, Wu, Qu, Long, Shen, and Zhang, 2003; Breitung and Candelon, 2006; Rua, 2010; Rua, 2012

[12] A spectrum refers to one variable, and a cospectrum is for the combined result of two variables. The integral of a spectrum over all the frequencies (i.e., the sum of power density of a spectrum) is equal to the variance of the time series and the integral of a cospectra over all the frequencies gives you the covariance. Therefore, the spectrum breaks the sample variance of time series up into pieces across different frequencies, and cospectra breaks the sample covariance across different frequencies.

[13] See section 2.7 on page 7

[14] see section 2.8 on page 8

[15] See section 2.4 on page 5

[16] Ramsey, 2002; Crowley, 2007; Rua, 2010; Rua, 2012

[17] Huang, Wu, Qu, Long, Shen, and Zhang, 2003

[18] R. Agrawal, C. Faloutsos, and A. Swami: Efficient Similarity Search in Sequence Databases, Proc. Int'l Conf. Foundations of Data Organiz ations and Algorithms, pp. 69-84, Oct, 1993.

[19] The Discrete Fourier Transform (DFT) can be understood as a numerical approximation to the Fourier transform.

[20] With time series data, the data often contain more than a single series, and is a set of multiple time series.

[21] In linear algebra, the Singular Value Decomposition (SVD) of a matrix is a factorization of that matrix into three matrices. It has some interesting algebraic properties and conveys important geometrical and theoretical insights about linear transformations.

[22] Korn, F., Jagadish, H & Faloutsos. C.: Efficiently supporting ad hoc queries in large datasets of time sequences. Proc. of SIGMOD '97, Tucson, AZ, pp 289-300, 1997

[23] A discrete wavelet transform (DWT) is a transform that decomposes a given signal into a number of sets, where each set is a time series of coefficients describing the time evolution of the signal in the

(DWT)[24]. Keogh et al.[25] introduce a novel transforming technique for time series dimensionality reduction call Piecewise Aggregate Approximation[26] (PAA)[27] so-called symbolic aggregate approximation[28] (SAX), and Extended SAX by Battuguldur Lkhagva, Yu Suzuki, Kyoji Kawagoe[29].

## 3.2 Challenges

There are technical analysis techniques that, for some reason, seem to have always been around. Why? Most probably because the trivial logic that governs these methods and their application is seemingly so hard-and-fast that almost anyone trading the markets sooner or later gives them a try. A case in point is the legendary technique of cycle trading, extolled by classics (such as, for example, the rocket scientist Mr. James Hurst), revered by some, and ridiculed by others. So what is so attractive about this method that made it one of the methods of choice for a great many of our fellow-traders? At first blush, if everything is calculated correctly, the technique seems to be so simple that any more or less thinking person who gets familiar with it for the first time can hardly help asking himself how come that there's still some money left out there. Indeed, what can be simpler - the markets move in cycles. The universe is boundless. Buy at rock-bottom, sell at the crests, and count the bucks. Simple, isn't it?

It's pretty obvious to anyone who has as much as had a brush with the market that there is defintely something cyclical to its nature. However, the **cyclical nature of the market is not governed by any clearly defined law - cycles form with constantly changing oscillating amplitude, frequency and phase displacements**. Given the cyclical nature of the world around us, the human brain inadvertently looks for a "cyclical" explanation of the different aspects of life: the changes of seasons and governments, droughts and bumper crops, downturns and periods of economic growth. The recurring processes in nature have been studied and used by humans since time immemorial. Our distant ancestry were able to use their observations of these phenomena to create things that are still capable of impressing us today.

Sound scientific basis for the modern spectral analysis was provided in the 17-th century by Sir Isaac Newton. The field was contributed to by a number of other luminaries, such as Daniel Bournoulli, Norbert Wiener, John Tukey. As far as trading is concerned, one cannot but mention the scientific legacy of the esteemed Jean Baptiste Joseph Fourier, the renowned French Egyptologist and Mathematician, whose paper Théorie analytique de la chaleur (1822) gave birth to the scientific notion of what is now known as the Fourier series. Another major, if the greatest, contributor to the field was Mr. John Burg, who viewed the issue from the angle of maximum entropy in his 1975 doctoral thesis. This approach laid the foundation for cyclic analysis in trading owing to the very small amount of data it required to provide spectral estimates. In theory, to profit from the cyclic nature

corresponding frequency band.

[24] Chan, K.& Fu, W.: Efficient time series matching by wavelets. Proc. of the 15th IEEE International Conference on Data Engineering, 1999.

[25] Keogh, E., Chakrabarti, K., Pazzani, M., Mehrotra, S., 2000. Dimensionality reduction for fast similarity Search in large time series databases. Journal of Knowledge and Information Systems 3 (3), 263–286.

[26] This technique approximates the dimensions by segmenting the sequences into equi-length sections and recording the mean value of these sections. The extended versions of PAA can be found by the works of Lin J., Keogh E., Lonardi S., Chiu B

[27] Lin J., Keogh E., Lonardi S., Chiu B. A Symbolic Representation of Time Series, with Implications for Streaming Algorithms. In proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery. (2003)

[28] The Symbolic Aggregate approXimation (SAX) algorithm bins continuous time series into intervals, transforming independently each time series (a sequence of floats) into a sequence of symbols, usually letters.

[29] Battuguldur Lkhagva, Yu Suzuki, Kyoji Kawagoe: New Time Series Data Representation ESAX for Financial Applications. ICDE Workshops 2006

of the market, one needs to divide the price movement into the trend, cycle and noise. Then the periods when the trend doesn't exist need to be identified, when the cyclical component carries enough weight "to drown out" the market noise. Ideally, the trend should only be traded with when the level of the cyclical activity is low and it is only during those periods, when the cyclical activity is very strong that cycle trading can be applied.

Just like any technique called to describe the nature of the market cycles are a simplified model of the market. As the basis for the creation of a market model, the sinusoid is used. Simple sinusoids are combined to model the market's cycle nature using a range of mathematical methods, from simple cycle finders (for example, the determination of the average distance between two lows) to Fourier transforms or Maximum Entropy Spectral Analysis, reckoned to be the most efficient of the methods. As the parameters for this kind of modeling the amplitudes, frequencies and phases of the primitive cycles being combined are used. The art of cyclic analysis consists in the ability to correctly select the required combinations while taking into account the resonances and objectively determining the parameters. The cyclic analysis approach can be used for describing classical chart patterns, determining trend channels, parameter determination methods for MAs and indicators, as well as for calculating stop loss signals.

### 3.2.1 Technical Pitfalls

Any more or less versed trader is fully cognizant of the fact that due to the nature of the market it is nearly impossible to identify the beginning and end of a cyclical period or correctly select the cycle frequency to trade with. Basically, a market cycle is a repetition of a stock's or currency's average fluctuations - a low, a rally, and a new low, dividing the period of time, occupied by the cycle, into the following four stages: the initial rock-bottom/recovery stage, upward move, distribution stage, and downward move. However, even if we do wholeheartedly embrace the elegant theory of the market's cyclical nature, how do we identify the beginning of a cycle and its end? Probably, this is both the greatest challenge and deepest pitfalls awaiting anyone who treads on the risky path of cycle-trading.

Besides, normally, the market rises and falls several times before any of the three major points is reached. This can confuse the trader and mislead him into entering or exiting a position too early, which, actually, happens rather often. Actually, any cycles-based trading offers an acceptable amount of risk only if the cycles are consistently repeated at least 85-90 % of the time and the smaller moves account for not more than the remaining 10-15 %. All those who attempt to come up with a viable classification of the different cycle stages and just another smart technique of how to identify them, either mentally or automatically with a specialized software, have so far failed. Attempting to picturesquely describe the different market stages, lulls, and upsurges with the help of tenacious psychological phenomena is one thing, identifying these phenomena as they happen is a completely different matter.

### 3.2.2 Fundamental Pitfalls

Today, cycles remain one of the most spoken about trading-related topics. As already mentioned above, the beginning and end of a cycle are hard to predict. This is highly disadvantageous, as this very divide is the cornerstone of cyclic analysis and even the simple realization of it's existence by a trader has a practical meaning. Often, the fluctuations of an amplitude fade just as the trend reaches the required height and the model needs to be adjusted with new parameters either manually or automatically. This takes a while and, normally, the kind of the model you receive after the adjustment can only provide a good picture of the market - alas, the trend can no longer be used. In other words, a cycle with other parameters can become dominant at any moment in time. Modeling the market's cyclic nature, despite its individual character, is based on the gargantuan task of approximating the price time movements to analytical functions.

# Chapter 4

# A direct stochastic numerical mathematical model of stock price volatility based on decomposition of weighted average price using additive fractal adaptive polynomial spline with multilevel knots

## 4.1 Model Description

This model can be described by six primary equations, seven secondary equations and three sets of ordered pairs. First primary equation defines that, weighted average price (WAP) of a particular stock "s" on day $d_n$ can be decomposed into a model. In turn whose values also vary and mimic the values of WAP with variation in "s" and $d_n$.

$$W_{(s,d_n)} \quad \approx \quad \fint(M_{(s,d_n)}) \tag{4.1}$$

Originality of this approach resides in the understating that this model does not attempt to predict future. A non-attempt of predicting future is due to following reasons:

- According to section 2.1 Theoretical asset pricing models became a necessity since asset values observed in a market are different from their fundamental or fair value. This is not only due to demand and supply factors but also due to speculation. In case of secondary equity markets, speculation is prevalent at daily, weekly, monthly and quarterly level also. Although, one of many objectives of setting up stock markets around the world was price discovery and this price discovery was assumed to be fair which was then used to provide liquidity to investors; This is not the case today, stock markets are more than just price discovery mechanisms. And, since stock markets now involve many actors with varying motivations and intentions, it is hard if not impossible to incorporate these "motivations" and "intentions" as variable inputs to predictive models. This indeed should be a primary reason why predictive models fail in the long run.

- EMH (see sections 2.2) if true, out-rightly rejects our ability to predict future prices of stocks. If on the other hand, EMH is false, then predicted prices will never be accurate since those will not incorporate consequences of all present˃ and future˃ information.

˃Unknown to some outsiders

˃Unknown to all outsiders

This is why the above equation has an approximation sign˃ instead of an equal sign. $W_{(s,d_n)}$ is the weighted average price of a particular stock "s" on day $d_n$ and, $M_{(s,d_n)}$ is the modeled decomposition of $W_{(s,d_n)}$ for a particular stock "s" on day $d_n$ also, [redacted] $d_n$ is the last trading day having value greater than 85. Second primary equation defines that, the above mentioned model decomposition is achieved by decomposing the model $M_{(s,d_n)}$ into four sub-models. In turn whose values also vary and mimic the values of WAP˃ with variation in "s" and $d_n$.

˃intentionally similar but not exactly equal

˃In part, depending on the fractal

$$M_{(s,d_n)} \approx \fint (m_{l(s,d_n)}, m_{q(s,d_n)}, m_{m(s,d_n)}, m_{w(s,d_n)}) \qquad (4.2)$$

Continuing our non-predictive approximation the above equation represents an equivalent of a ensemble predictive model. This model is a decomposition˃ as theorized by the Fourier transform (see sub-section 2.7.1). The above decomposition consists of [redacted]. **This "hybrid" nature of decomposing WAP, then using the decomposition to explain variation in WAP, while using multiple [redacted] splines[1] [2] with [redacted][3] knots and determining spline coefficients using [redacted][4] approach is what makes this model truly innovative. This innovation overcomes the primary challenge of converting non-stationary WAP into a stationary model without loosing information depth.**

˃In principal

$m_{l(s,d_n)}$ is a stochastic representing long duration volatility of the stock "s" on day $d_n$ and, $m_{q(s,d_n)}$ is a [redacted] spline representing [redacted] volatility at [redacted] period of the stock "s" on day $d_n$ and, $m_{m(s,d_n)}$ is a [redacted] spline representing [redacted] volatility at [redacted] period of the stock "s" on day $d_n$ and, $m_{w(s,d_n)}$ is a [redacted] spline representing [redacted] volatility at [redacted] period of the stock "s" on day $d_n$.

$$m_{l(s,d_n)} = \frac{\text{[redacted]}\left(\frac{W_{(s,d_n)}\ \text{[redacted]}}{\text{[redacted]}}\right)\text{[redacted]}}{\text{[redacted]}} \qquad (4.3)$$

The above equation yields a similar score as observed on BSE website as shown below (example ABB India Limited):

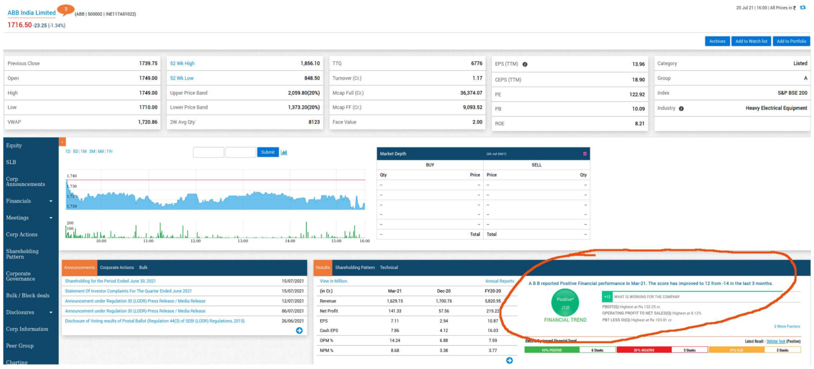
[1] **Instead of trigonometric function (see section 2.7.1)**

[2] **While preserving the essence of time to frequency transformation (see section 2.7 and 2.8)**

[3] **See section 2.9**

[4] **See section 2.6**

Figure 4.1: A similar stochastic currently available on BSE website



In the above figure, score of ABB India Limited has improved to 12 (after June 21) from -14 (after March 21) while from my analysis it has improved from 1 (as on 1st April 2021) to 25 (as on 1st July 2021). The difference in both cases is almost equal[>]. The **originality here is that I can report this stochastic on a daily basis whereas marketsmojo does it only on a quarterly basis**.

[>]Since marketsmojo has not specified the exact date of scoring

$$m_{q(s,d_n)} = \frac{\sum_{j=0}^{\text{[redacted]}} p_{(d_n,s,j)}\text{[redacted]} + \sum_{j=0}^{\text{[redacted]}} q_{(d_n,s,j)}\text{[redacted]}}{2} \qquad (4.4)$$

where, $p_{(d_n,s,j)}$ is defined as the adaptive coefficients of [redacted] of the first fractal [redacted]

$$y_p = \text{[redacted]}$$

where, $q_{(d_n,s,j)}$ is defined as the adaptive coefficients of [redacted] of the first fractal [redacted]

$$y_q = \text{[redacted]}$$

Above mentioned, three equations are a result of many trial attempts to mix and match the actual variability with the model[>]. This approach also transforms time series data from a time domain into a frequency domain. Although many researchers use predefined and automated algorithms to achieve this[>], our approach is original as it is developed from scratch[>].

[>]In data science this is also known as curve fitting

[>]Another reason why they fail

[>]Its like re-inventing the wheel which fortunately results in a better wheel

on day $d_n$ and the spline is defined as

$$y_{d_n} = W_{(s,d_n)} \text{[redacted]} (d_n \text{[redacted]} \times (\frac{W_{(s,d_n)} \text{[redacted]}}{\text{[redacted]}})) - \text{[redacted]}$$

While considering different splines researchers do not do this at all. What we have done is that instead of applying a spline on the original curve, we have applied it on a [redacted]. In other words, we are using splines for explaining variability that is not explained [redacted] The same is continued till all the variability is explained.

having a knot definition of $x_{d_n} = \{1, 2, ..., \text{[redacted]}\}$ therefore the ordered pairs of the polynomial spline are

$$\{(x_{d_n}, y_{d_n})\} = \{(1, W_{(s,d_n}\text{[redacted]}) \text{[redacted]} (d_n \text{[redacted]} \times (\frac{W_{(s,\text{[redacted]})} \text{[redacted]}}{\text{[redacted]}})) - \text{[redacted]}), ..., (\text{[redacted]} W_{(s,d_n)} \text{[redacted]} (d_n \times (\frac{W_{(s,d_n)} \text{[redacted]})}{\text{[redacted]}})) - \text{[redacted]})\}$$

$$m_{m(s,d_n)} = \sum_{j=0}^{\text{[redacted]}} r_{(d_n,s,j)}\text{[redacted]} \tag{4.5}$$

where, $r_{(d_n,s,j)}$ is defined as the adaptive coefficients of the second fractal polynomial

$y_r =$ [redacted]

on day $d_n$ and the spline is defined as

$$y_{d_n} = W_{(s,d_n)} \text{[redacted]} (d_n \text{[redacted]} \times (\frac{W_{(s,d_n)} \text{[redacted]} W_{(s,\text{[redacted]})}}{\text{[redacted]}})) \text{[redacted]} W_{(s,\text{[redacted]})} \text{[redacted]} m_{q(s,d_n)}$$

having a knot definition of $x_{d_n} = \{1, 2, ..., \text{[redacted]}\}$ therefore the ordered pairs of the polynomial spline are

$$\{(x_{d_n}, y_{d_n})\} = \{(1, W_{(s,d_n\text{[redacted]}} ) \text{[redacted]} (d_n \text{[redacted]} \times (\frac{W_{(s,d_n\text{[redacted]})} \text{[redacted]} W_{(s\text{[redacted]})}}{\text{[redacted]}})) \text{[redacted]} W_{(s,\text{[redacted]})} \text{[redacted]} m_{q(s,d_n\text{[redacted]})}), ..$$

$$.,(\text{[redacted]}, W_{(s,d_n)} \text{[redacted]} (d_n \times (\frac{W_{(s,d_n)} \text{[redacted]} W_{(s,\text{[redacted]})}}{\text{[redacted]}})) \text{[redacted]} W_{(s,\text{[redacted]})} \text{[redacted]} m_{q(s,d_n)})\}$$

$$m_{w(s,d_n)} = \sum_{j=0}^{\text{[redacted]}} s_{(d_n,s,j)}\text{[redacted]} \tag{4.6}$$

where, $s_{(d_n,s,j)}$ is defined as the adaptive coefficients of the third fractal polynomial

$y_s =$ [redacted]

on day $d_n$ and the spline is defined as

$$y_{d_n} = W_{(s,d_n)} \text{[redacted]} (d_n \text{[redacted]} \times (\frac{W_{(s,d_n)} \text{[redacted]} W_{(s,\text{[redacted]})}}{\text{[redacted]}})) \text{[redacted]} W_{(s,\text{[redacted]})} \text{[redacted]} m_{q(s,\text{[redacted]})} \text{[redacted]} m_{m(s,\text{[redacted]})}$$

having a knot definition of $x_{d_n} = \{1, 2, \text{[redacted]}\}$ therefore the ordered pairs of the polynomial spline are

$$\{(x_{d_n}, y_{d_n})\} = \{(1, W_{(s,d_n\text{[redacted]}} ) \text{[redacted]} (d_n \text{[redacted]} \times (\frac{W_{(s,d_n\text{[redacted]})} \text{[redacted]} W_{(s,\text{[redacted]})}}{\text{[redacted]}})) \text{[redacted]} W_{(s,\text{[redacted]})} \text{[redacted]} m_{q(s,\text{[redacted]})} \text{[redacted]} m_{m(s,\text{[redacted]})}), ..$$

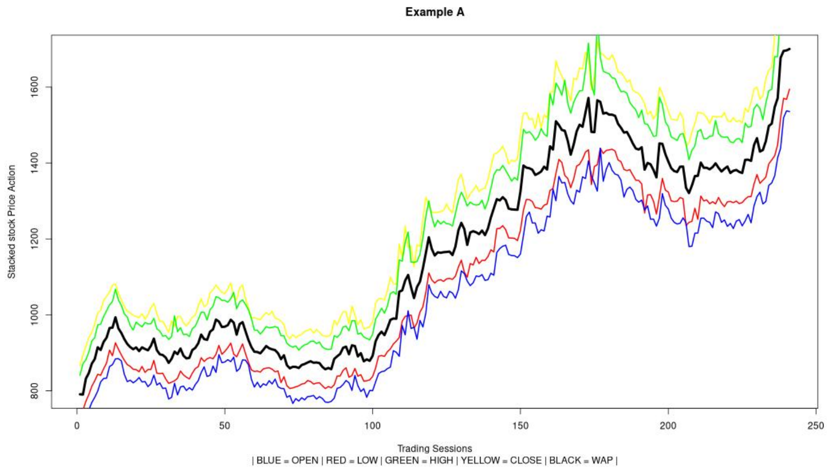
$$.,\text{[redacted]}, W_{(s,d_n)} \text{[redacted]} (d_n \times (\frac{W_{(s,d_n)} \text{[redacted]} W_{(s,\text{[redacted]})}}{\text{[redacted]}})) \text{[redacted]} W_{(s,\text{[redacted]})} \text{[redacted]} m_{q(s,\text{[redacted]})} \text{[redacted]} m_{m(s,\text{[redacted]})})\}$$

As is evident, the combination of [redacted] knots placed in such a way is unique as they are of different [redacted] and are placed at different [redacted] levels. Overall this modeling approach is both unique and original.

## 4.2 Envisaging Model's theory

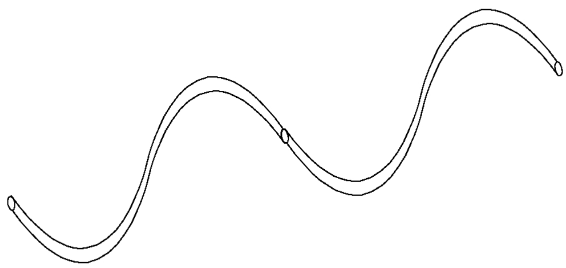
A typical counter shows the following price volatility across five basic variables vis-Ã -vis Open, Low, High, Close, WAP

Figure 4.2: Example A: Daily **Open,Low,High, Close & WAP** over a period of time for one counter




In above figure, it is impossible to distinguish them from each other through the naked eye. But mathematically they can be distinguished, at least three of them can be differentiated pretty easily. So, there has to be a very close relationship between them at least such a relation should exist separately for each counter. Following image depicts this relationship graphically (Not for this counter but in general). At first, visualize a pipe having cylindrical cross-section bended in a wave type form (as shown in below figure). Next, visualize the circular or any other cross-section representing that of the pipe (Here shown as a circle figure below).

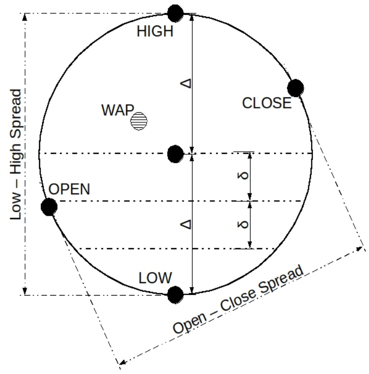
Figure 4.3: Concept: A pipe having a cylindrical cross-section bended in a wave type form (Price volatility of a counter overtime simplified in form and including all prices, the Highs, Lows, Opens and Closes within the pipe boundary)



I discovered a key feature of this pipe, no mater what the time period (i.e. whether we are observing prices over years, months, weeks or days) price volatility could be simply looked at in this way. I realized that there has to be multiple levels of this pipe. In other words, If you were to zoom-in into the price volatility trend with respect to time (i.e. from year to month, or from month to days etc.) you will find the same trend reoccurring again. In other words, the time resolution of volatility does not affect the basic nature of the trend[5]. Question then was at what resolution does the nature (or volatility) is insignificant to make profits.

"''Fractal patterns appear ... in the price changes of securities..." B. Mandelbrot, September 15, 2008, Scientific American

Figure 4.4: Concept: Pipe cross-section (counter position for a single day)




The above figure describes a stocks **price action in-terms of a cyclic quadrilateral**. Looking at both figures above, one can visualize how on some days the pipe would have a smaller diameter and on some days larger, one can also visualize the progression of relationships that exist between various price points of the counter overtime.

[5] A fractal is a geometric shape that can be separated into parts, each of which is a reduced-scale version of the whole. In finance, this concept is not a rootless abstraction but a theoretical reformulation of a down-to-earth bit of market folklore— namely, that movements of a stock or currency all look alike when a market chart is enlarged or reduced so that it fits the same time and price scale. An observer then cannot tell which of the data concern prices that change from week to week, day to day or hour to hour. This quality defines the charts as fractal curves .... A more specific technical term for the resemblance between the parts and the whole is self-affinity. This property is related to the better-known concept of fractals called selfsimilarity, in which every feature of a picture is reduced or blown up by the same ratio—a process familiar to anyone who has ever ordered a photographic enlargement. Financial market charts, however, are far from being self-similar. ... The existence of unchanging properties is not given much weight by most statisticians. But they are beloved of physicists and mathematicians like myself, who call them invariances and are happiest with models that present an attractive invariance property ... a fractal price model can be altered to show how the activity of markets speeds up and slows down—the essence of volatility. This variability is the reason that the prefix "multi-" was added to the word "fractal ... The discrepancies between the pictures painted by modern portfolio theory and the actual movement of prices are obvious. Prices do not vary continuously, and they oscillate wildly at all timescales. Volatility—far from a static entity to be ignored or easily compensated for—is at the very heart of what goes on in financial markets. In the past, money managers embraced the continuity and constrained price movements of modern portfolio theory because of the absence of strong alternatives. But a money manager need no longer accept the current financial models at face value. - B. Mandelbrot, September 15, 2008, Scientific American - https://www.scientificamerican.com/article/multifractals-explain-wall-street/

# Chapter 5

# Research approach and method

Firstly at an individual level, usually short term stock traders depend on market rumors/news/corporate actions etcetera for identifying trading opportunities. But for working professionals it is time consuming. Also, the time, news takes to reach an external retail investor is too late. Furthermore, people working in the industry or are close to the industry or have better networks have a faster idea about the trading opportunities. So, the idea was to identify stocks[1] that have a potential for upward movement beforehand the market or other traders do. Secondly [redacted] I try to capture the research undertaken, process involved, the model, experimental results and application of the model. Furthermore, I also suggest how the same model can be scaled and applied to different markets [redacted]
I explain in detail the experimental test results, their analysis and the emergent strategy for trading in order to maximize profits and minimize risks.

## 5.1 Research Objectives

Primary objective of this research was to enable [redacted] building platform [redacted] But, there were other objectives as well. Following points summarize them:

- Counter[>] trading should not depend on the market information/ knowledge/ intelligence of the user[>].
- Counter trading should take less than one hour for the user to use it including transactional operations[>] during market hours.
- Strategy for Counter trading should minimize the downward price movement risk or use it as an advantage
- The Model should not dependent on the number of counters which need to be tracked/watched for entry[>] and exit[>] opportunities
- The Model be sufficiently efficient and effective even if based on publicly and freely available data

[>]Counter is a technical name for stock

[>]of the software aka the retail investor

[>]Buying and selling of stocks

[>]Buying

[>]Selling

[1] Based purely on price movement and tracking volatility of a list of stocks

- The Model has to be self learning, i.e. its effectiveness should increase more with time
- The Model shall prohibit market operations on counters which exhibit non-fundamental˃ price movements ˃Manipulated
- The Model should suggest trade opportunities which have low turnaround times typically near a month
- The Model should have an efficiency and effectiveness based upon return on capital deployed per day such that the trades executed by this model should give a high annual effective rate of return on capital deployed compounded monthly˃ ˃AERRoCDCM
- The Model should provide a signal˃ for entry into a trade ˃Indicator
- The Model should provide a signal for exit to minimize downward risk
- The Model should be able to calculate the buy price˃ of any counter on a specific day i.e. its daily low and also calculate the sell price i.e. it's daily high. ˃Days Low value

## 5.2 Phase I unstructured learning

I had never invested in secondary capital markets not even thought of it. I did not know the market from a practical view point. However, being a management graduate I knew only of the theory. [redacted] From this point on I started experimenting, learning, analyzing, investing and trading all simultaneously. I started with observing and executing experimental trades in intra-day segment. During this time my main objective was to study price movements and profiting by predicting tomorrows high and low prices of a typical counter. Everyday I selected some 450 counters and started to track their daily price movements. As intra-day (even short term) price fluctuations of such counters were highly speculative and driven by sentiments rather than fundamentals, I failed but with limited success.

Success was a ray of hope. I had tried brute force prediction of next day high and low prices. By brute force, I mean, that I used more than 30 different methods (level 1 models) to predict possible highs and lows of counters. And, then strategically combining outputs (Level 2 model) from these methods to predict a final high and low prediction. Obviously this was extremely time-consuming for 450 counters using spreadsheets with advanced formulas. Overtime and after loosing some money I discovered the following:

- Out of all the level 1 models only one would give a close prediction
- I could not know in advance which this level 1 model would be
- None of the level 2 models was close to actual highs and lows
- Without having trade-level data (i.e. transaction data for each trade) I was unable to predict what will occur first, high or low
- I got to learn practicality in counter price movements and how the market worked on a daily basis
- I got ideas for developing higher level models
- I understood that spreadsheets were not the right tool for analyzing due to its slow and limited functionality

- I realized that I could not spend hours on the trading terminal and that too during working hours as this was only a stepping stone not the end

I know computer programming, but I am not a professional programmer, I know data, but I am not a professional equity analyst. With this low levels of confidence and already failed once I kept on working with spreadsheets and developing higher level models. Finally, I developed a working model but without complete success. By this, I mean, I was able to predict the next day high and low but the spreadsheet workbook become so large that I could not use it to track more than 10 counters, also it was still lacking a very large part of analysis; furthermore, I had to do a lot of data entry manually each day (Firstly, the data structure had more than 2 dimensions i.e. it could not be represented through a 2 dimensional table and Secondly, the daily outputs from this model had to become the input of another model as a time-series data).

## 5.3 Phase II structured learning

I searched for a tool that had to full-fill my data analysis requirements and it should:

- Be non-propriety
- Enable me to create my own data model from scratch
- Be faster than spreadsheets
- Have a command line interface and scripting support for complete flexibility

One such tool was the R Statistical Programming Environment, but learning it was difficult as per On-line reviews. Indeed, the basic programming semantics was easy, but, actual application was hard to learn through tutorials. So, I decided to learn by application and thus started programming the existing model into R environment. During programming, new ideas about the model were discovered and then implemented. The model development process involved:

1. Envisaging a theory for this model
2. Proving that the model explains actual real world behavior
3. Applying the theory i.e. developing a process and strategy for applying the model relevant to the practical situation
4. Proving the concept i.e. simulating the model application with current new situational data without undertaking actual transactions
5. Testing the model application by transacting in real world live markets
6. Fine tuning the model, application strategy and process with new-found test results

### 5.3.1 Proving the theory

There are thus, three aspects to the theory:

1. That there are multiple levels of the price volatility of a counter. In other words there is significant depth to price volatility.
2. Can the model mimic and explain this price fluctuation and signal as to which counters should be traded and when

3. Is there a significant relationship between different price points of a counter at the intra-day level (or depth)

After considerable analysis and its automation by coding the same in R, I found that at the first level (at highest resolution or depth zero), price volatility of a counter depends on only one dependent variable or factor. At the second level, price volatility of a counter depends on five factors. At the third level, price volatility of a counter depends on seven factors. The fourth level (at lowest resolution or depth of -3) which has significant effect on profits, price volatility of a counter depends on 9 factors. At lower levels the price volatility was insignificant in the range of $e^{-11}$ to $e^{-14}$ which is acceptably insignificant. These consistent results across counters were obtained over a period of two years thereby proving that the theory explains counter price volatility in sufficiency and depth. Which in-turn means that the model can be used to at-least identify where the price of a counter is headed and a prediction be made as to which counter is expected to increase in price within a given time frame. The question then remains as to how does one select counters which will potentially show an up trend within one month on any given day which can be bought the next day.

### 5.3.2 Applying the theory

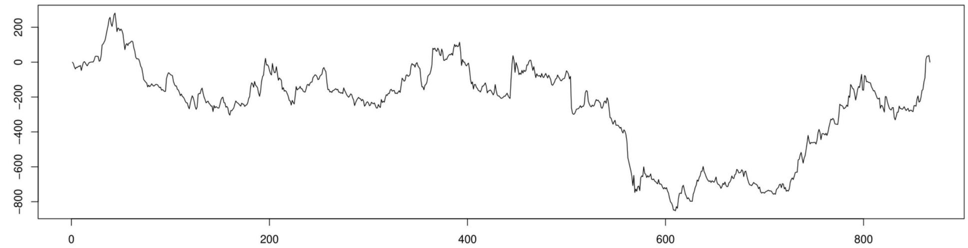
Once the theory proved itself reliable, next step was to identify or develop signals that helped in identifying trading opportunities i.e. signals that the model would throw up for the user which in-turn would help the user to identify among 144 counters, which ones to explore further for trading on any given day.

In-order to obtain transaction signals following key decision parameters were defined:

- The rolling level 1 trend signal (RADLINE)
- The rolling level 2 trend signal (WAP1)
- The rolling level 3 trend signal (WAP2)
- The rolling level 4 trend signal (WAP3)
- The rolling next month price prediction i.e. what can the price of a counter be at the end of one month from today, if it followed the same volatility trend at all four levels (SIGNAL)

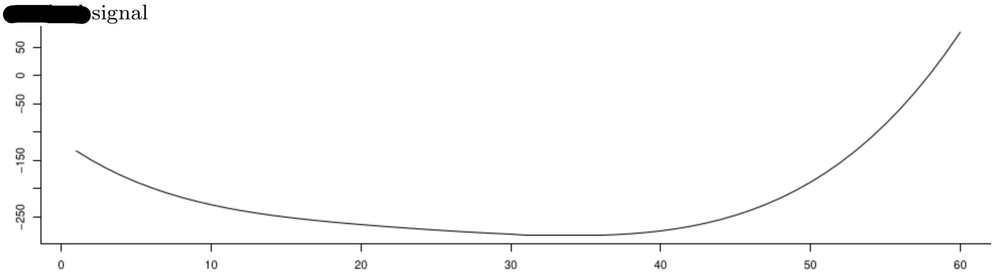
Firstly, WAP was taken as the basic unit of analysis. It was observed that other four price points (namely, Open, High, Low and Close) are a function of WAP as WAP not only depicts the overall market sentiments but is also a more accurate parameter for counter performance. Furthermore, WAP is hard to manipulate and is harder to get affected by transaction anomalies. Secondly, in-order to [redacted] among all counters it was necessary to extract the level 1 signal. Thirdly, the [redacted] level 1 signal was required for further analysis.

Figure 5.1: Example A (Continued): Level 1 WAP [redacted] trend signal [redacted] WAP)



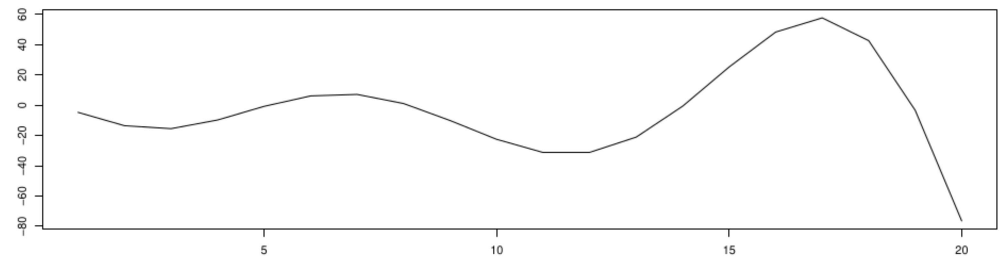

From this level 1 [redacted] signal, level 2 signal was [redacted]

Figure 5.2: Example A (Continued): The rolling level 2 signal obtained from level 1 [redacted] signal



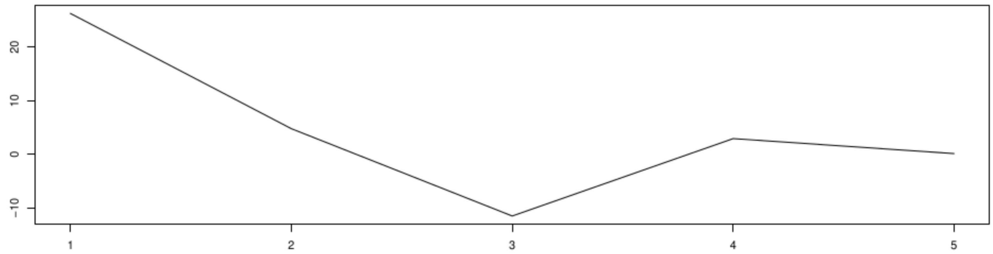

After [redacted] the level 2 signal, level 3 signal was [redacted]

Figure 5.3: Example A (Continued): The rolling level 3 signal obtained from level 2 [redacted] signal



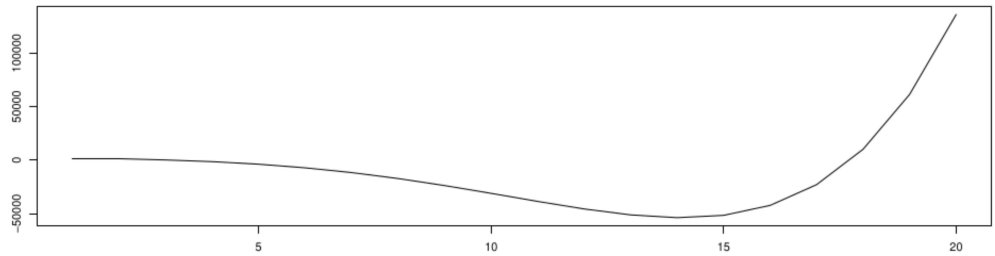

Finally, level 4 signal was [redacted]

Figure 5.4: Example A (Continued): The rolling level 4 signal obtained from level 3 [redacted] signal




Using all these four levels of signals a predictive signal over next month was calculated as shown in the following figure.

Figure 5.5: Example A (Continued): The rolling next month predictive signal




Remember, all the four levels and predictive signal is calculated daily for each counter. So the above figures are for a specific day for a specific counter only. With subtle differences in price movements during trend formation it became necessary to magnify these differences. In order to that following key parameters (Coefficients of Fourier Decomposition[2]) were defined and analyzed further to capture and identify a larger and more prominent trend:

- The rolling weekly change in direction, quantum of this change and momentum of this change
- The rolling monthly change in direction, quantum of this change and momentum of this change
- The rolling quarterly change in direction, quantum of this change and momentum of this change
- The rolling change in end of next month prediction its direction and quantum

These four stochastic[3] parameters were calculated daily [redacted] and finally four more parameters were defined as follows:

**SIGNAL** captured the change in direction and quantum of the rolling change in end of next month prediction

**PRIORITY** captured the change in direction and quantum of level 2 signal

[2]See 2.7 on page 7
[3]See 2.6 on page 7

**OPPORTUNITY** captured the change in direction and quantum of level 3 signal

**HISTORIC** captured the change in direction and quantum of level 1 signal

Now, coupled with the above stated parameters and recorded past price points for all the counters under research, I mined for counters which had shown accelerated price volatility in upward direction. Key objective for this exercise was to identify a pure phenomenon (a fractal[4]) only in numbers which explained this sudden upward price fluctuation for the above stated counters. After a lot of hits and trials one such phenomenon was discovered during analysis. Based on this phenomenon further research was carried out to identify and define the significant relationship between different price points of a counter at the intra-day level (or depth). Equipped with both the phenomenon and the relationships it was time to prove my concept.

### 5.3.3 Proving the concept

Till now except during the unstructured learning process I was not trading at all which means all the analysis was only on paper and theory none of it was used for actual trading in the market. Therefore, I decided to continue not to trade and first simulate my theory and concept not with past data but with live data. The major difference between working with past data (or events that had already happened[>]) and working with live data (or with events that had not yet happened) was that choosing a particular counter for trading would not be influenced by me knowing that this particular counter would show accelerated upward price volatility or not. Which is to say that the concept would prove itself if and only if it helped me to choose a counter which at that moment did not show any inclination of upward movement, but, later within one month would move upward according to the theory's expectation. There are few points worth noting during the simulation exercise:

[>]Process also known as backtesting

- Open to Open appreciation ranges from minus two percent to plus eighteen percent, since this is only the open to open appreciation, actual realization would be higher for each transaction as open might not be the lowest for the day nor the highest for the day. Also the counter might not be bought on the same next day and there might be a delay in buying at the call price.
- Majority of the counters have shown a positive appreciation
- Majority of the transactions would not have been executed due to the fact that call prices were either lower than the day's low or higher than the day's high
- Majority of the transactions were completed within close to thirty days
- Risk of loosing money has been reduced by far; it is far less than normal probability of fifty percent.

The concept of calculating the call price both for buy and sell transaction was to counter market behavior. The market behaved such that more often than not the defining high or low price for that specific day occurred at opening time and then never returned to that price during the day. So, it was imperative that the day's low/high be estimated before market opening to maximize profits. However, during simulation it was discovered that the call price was either too low (in case of buy transactions) or was too high (in case of sell transactions). This was the case with majority of the transactions. If this was to be implemented then the model would have failed to deliver profits. The key question then

[4]See 2.9 on page 8

was, whether the estimation was wrong or the other market participants never placed orders at my call price. This question could only be answered by participating actively in the market and thus live testing the model.

## 5.4 Testing the model

Testing the model involved executing market transactions based on analysis, suggestions and calculations done through the model. So I opened fresh accounts with a new broker namely [redacted] to execute my trades. The plan was to have a designated capital for each trade transaction so that all the trades can be compared in terms of percentages. The per trade capital was thus fixed and was equal for all transactions. If the model failed to deliver profits while testing, it would have been considered a complete failure even if above objectives were achieved. A startling discovery during the test was that majority of the trades got executed during testing as opposed from simulating. This only proves one point, that market participants do not place their orders at the prices which my model suggests. This in-turn means that my pricing model gives me an advantage in the market.