Research on Volume-Order Imbalances in Equity-Futures Limit Order Book and Analysis of the Dynamics of Implied Volatility Curve of BANKNIFTY Option Chain

A PROJECT REPORT

Submitted in partial fulfillment of the
requirements for the award of the degrees

of
BACHELOR OF TECHNOLOGY

in
METALLURGY ENGINEERING AND MATERIALS SCIENCE

Submitted by:
ABHIJEET SINGH TOMAR

Guided by:
DR. ABHIJIT GHOSH
Assistant Professor
Department of Metallurgy Engineering & Materials Science

INDIAN INSTITUTE OF TECHNOLOGY INDORE
NOVEMBER 2019

<u>CANDIDATE'S DECLARATION</u>

I hereby declare that the project entitled **"Research on Volume-Order Imbalances in Equity-Futures Limit Order Book and Analysis of the Dynamics of Implied Volatility Curve of BANKNIFTY Option Chain"** submitted in partial fulfillment for the award of the degree of Bachelor of Technology in **Metallurgy Engineering and Materials Science** completed under the supervision of **Dr. ABHIJIT GHOSH, Assistant Professor, Department of Metallurgy Engineering and Materials Science, IIT Indore** is an authentic work.

Further, I declare that I have not submitted this work for the award of any other degree elsewhere.

Abhijeet Singh Tomar

CERTIFICATE by BTP Guide

It is certified that the above statement made by the students is correct to the best of my knowledge.

Dr. Abhijit Ghosh

PREFACE

This report on "**Research on Volume-Order Imbalances in Equity-Futures Limit Order Book and Analysis of the Dynamics of Implied Volatility Curve of BANKNIFTY Option Chain**" is prepared under the guidance of **Dr. Abhijit Ghosh**.

This report is the result of my work at IRageCapital Advisory Pvt. Ltd., Mumbai where I worked as a Quantitative Analyst Intern for a course of six months. I worked on a total of five projects out of which the two major ones are discussed in this thesis. The other were more relevant to inhouse data insights generation and were tightly bound to the proprietary and secret data sets owned by the firm and therefore cannot be discussed here. The thesis explains the work done on two very important fundamentals of the High Frequency Trading of Derivatives, the Implied Volatility Curve of Options and the Limit Orders Books. The work involves an attempt to solve the Implied Volatility Curve using spline regression methodology and its subsequent testing on live market data. It also attempts to verify the presence of Volume-Order Imbalance based signals in Indian markets which had a highly pronounced presence in the USA markets.

I have tried to the best of our ability and knowledge to explain the content in a lucid manner. I have added various figures and graphs to make it more illustrative.

ABHIJEET SINGH TOMAR

B.Tech. IV Year

Discipline of Metallurgy Engineering and Materials Science

IIT Indore

Acknowledgements

I would like to thank my B. Tech. Project Supervisor **Dr. Abhijit Ghosh** for his valuable guidance and support throughout the duration of the project and for giving me this opportunity to work in this interesting domain. Without your support, this report wouldn't have been possible.

I would like to express my sincere gratitude to **Mr. Ankit Soni**, **Senior Associate** at **Options Market Making Team** at **IRageCapital Advisory Pvt. Ltd**. He provided me with constant guidance and helped me understand various financial concepts and the relevant technologies. I would also like to thank all the members of the firm for always being open to discussions, solving problems and for helping me in every possible way, and thus creating a very productive environment to work in.

I would also like to express our deepest gratitude to our Respected Director sir, **Prof. Pradeep Mathur** for providing all the facilities and giving me the opportunity to work in a nice and conducive environment.

I am really grateful to our respected HOD **Dr. Parasharam M. Shirage**, DUGC Convener **Dr. Santosh S. Hosmani**, and all the respected faculty members of the **Department of Metallurgy Engineering and Materials Science, IIT Indore** for their kind approval and guidance, without which I wouldn't have been able to pursue this opportunity.

I would also like to thank my friends and family who provided constant encouragement and enabled me to perform at best of my abilities.

ABHIJEET SINGH TOMAR

B.Tech. IV Year

Discipline of Metallurgy Engineering and Materials Science
IIT Indore

ABSTRACT

The work done focuses on the High Frequency Trading, specifically the Market Making of Future and Future based Options in Indian Markets. We first study the dynamics of Volatility Curve for Options, which plays a very crucial role in the Market Making of Options. We tried to solve the Implied Volatility Curve using Spline Regression Methodologies and tested it for performance. This thesis also aims to investigate the performance of a Volume-Order Imbalance based trading strategy which was originally found to work for the USA Markets by Darrel Shen [15] in Indian Markets. We try to understand how the strategy performs on different futures contracts and its relationship with trading volume with respect to Indian Markets.

NON-DISCLOSURE AGREEMENT

I would also like to throw light on the fact that being bound under the industry standard non-disclosure and non-compete agreement, which was signed by me as a compulsive measure in order to be able to work at the firm, I am forced to not disclose some of the technical details of related to work, which include implementation methodologies, source code and data description, of my work at the discretion of the firm. I regret to inform that I cannot void the agreement at any point of time and through any means, as it would have criminal implications on me and will have severe consequences on my professional career. The final implementation and integration of my work into the firm's trading model will be done as per the discretion of the stake holders and directors of the firm.

LIST OF FIGURES

LIST OF TABLES

Table of Contents

Contents

(this page is intentionally left blank)

CHAPTER 1

INTRODUCTION

HFT IN INDIA: PAST, PRESENT AND FUTURE:

High-frequency trading, also recognized as HFT, is a method of trading that uses powerful computer programs to transact a large quantity of orders in fractions of a second. It makes use of complex algorithms to analyze multiple markets and execute orders primarily based on market prerequisites [17]. Typically, the participants with the quickest execution speeds are greater profitable than participants with slower execution speeds. In addition to the high velocity of orders, high-frequency trading is additionally characterized through high turnover costs and order-to-trade ratios. In the US and other developed markets, High-Frequency Trading and Algorithmic buying and selling debts for an estimated 70% of equities market share. In India, the percentage with admire to the whole turnover has expanded up to 49.8%. In India, the two famous exchanges, specifically National Stock Exchange, the NSE and Bombay Stock Exchange, the BSE, each facilitate High Frequency Trading with the aid of supplying the purchasers with splendid statistics aid and technological infrastructure to work with.

On April 3rd 2008, Securities & Exchange Board of India (SEBI), allowed algorithmic buying and selling by way of permitting Direct Market Access facility to institutional clients. In short, DMA allows brokers to furnish their infrastructure to clients and gives them access to the exchange buying and selling gadget except any intervention from their part. Initially, it was once furnished solely to institutional customers and now not retail traders. Nevertheless, the facility brought down prices for the institutional investor as nicely as assist in higher execution with the aid of reducing down the time spent in routing the order to the dealer and issuing the critical instructions.

On April 29th 2008, some of the top global players signing up for the DMA facility had already made this facility famous. The companies awaiting approval are FI's & FII's including UBS, Morgan Stanley, JP Morgan, and DSP Merrill Lynch. Among other things, Edelweiss Capital, India Infoline and Motilal Oswal Securities submitted their application to the stock exchanges.

It is worth noting that since 24 February 2009, Foreign Institutional Investors (FIIs) have been allowed to use DMA facility by investment managers appointed by them.

By July 31, 2008, the ground for operationalization of Direct Market Access (DMA) was prepared by leading brokerages along with stock exchanges. Brokerages including Citi, Merrill Lynch, Morgan Stanley, JP Morgan, Goldman Sachs, CLSA and Deutsche Equities have started checking their DMA technology in an effort to synchronize it with stock exchange systems.

NSE'S CONTRIBUTION TO THE INDUSTRY:

In June 2010, the National Stock Exchange (NSE) began offering additional 54' racks' on-lease colocation servers to broking firms in an effort to improve trading speed [17].

Among the international broking companies that used the facility are Deutsche Bank, Citi, Morgan Stanley, Goldman Sachs, and MF Global. Among the prominent domestic firms that signed up for the racks are Motilal Oswal Securities, JM Financial and Edelweiss Capital. The facility had also been chosen by local brokerages such as Globe Capital, SMC, Global Vision, East India and iRageCapital. Not surprisingly, with a few weeks of offering this facility, there was a long period of waiting up to 6 months to get a space on the server racks!

It was obvious to the Indian exchanges and regulatory bodies that Algorithmic Trading is well received by the country's institutional customers and banks and their demand will continue to rise. This was the time when exchanges in the automated trading field started to develop their offerings, financial technology companies began to provide automated trading platforms and SEBI continued to control markets.

On 12 May 2010, NSE moved to allow its trading platform's Financial Information Exchange (FIX) protocol to improve transaction speed for direct market access for overseas investors. Simply put, the FIX protocol helps to convert the language of orders issued by the Foreign Institutional Investors (FII) into the language the NSE understands, effectively reducing the time taken to execute the transaction.

CHANGES TO THE BROKERAGE INDUSTRY:

Broker commissions have started to decline as a result of a growing number of institutional clients warming up to the idea of Direct Market Access (DMA). They started offering automated software to consumers to keep up with the times. The new entrants to this market are discount brokers who are simply brokers with very small brokerage charges offering services. They can do this by offering only basic services, unlike full-service brokers who typically provide their clients with support and training programs.

REGULATIONS IN INDIAN STOCKS MARKETS:

Every year SEBI comes up with regulations to be implemented by traders and brokers in order to secure and risk-controlled the trading industry. That's why exchanges allow a business to undergo a series of rigorous checks if it wants to transact by algorithmic trading for any algorithm to be accepted by the markets. Such measurements include the number of orders placed per second, the average order price of any order placed, and the maximum amount exchanged during a single trading day.

A brief overview of the latest SEBI circular (SEBI / HO / MRD / DP / CIR / P/2018/62) dated April 09, 2018 is given below[17]:

Managed colocation service-

It is suggested that exchanges should change the pricing structure of their co-location rent to make it accessible to small and medium-sized members as the current practice of renting the entire server rack to one entity leads to high costs.

Latency measurement-

In order to make the reporting of latency for colocation and proximity hosting more transparent, it was suggested that the exchanges should provide minimum and maximum latencies along with the 50th and 99th percentile latencies.

Tick-by-tick data feed-

SEBI suggested that the exchange members be provided with tick-by-tick data free of charge.

Specific algorithm identifiers-

SEBI has instructed that all algorithmic orders entering their network should be tagged with the unique identifier assigned for approval when the particular algorithm has been submitted.

FUTURE OF ALGORITHMIC TRADING IN INDIA:

Over the years, India offers a good opportunity for algorithmic trading due to a number of factors such as colocation facilities and sophisticated technology at both major exchanges; a smart order routing system; and well-established and liquid stock exchanges [17].

Considering HFT and Algorithmic Trading's rapidly growing trend and demand in developing economies & emerging markets, numerous exchanges have made efforts to educate their members and build the skill sets needed for this technology-driven sector.

With many different trading platforms and resources available on the market, each claiming to be better than the other, a person testing the Algo trading might be spoiled and confused by default. Therefore, we've compiled a list of some of the most popular platforms and algo trading technology used in today's market (specifically for Indian equity markets) to level the playing field and give users a clear picture.

STOCK EXCHANGE:

The capital market's secondary level is what we call the stock market or the stock exchange. The stock exchange is a virtual market where existing securities are traded by buyers and sellers. It is a market operated by an entity or any government body that exchanges shares, securities, debentures, bonds, futures, options, etc.[16].

To buyers and sellers, a stock exchange is a meeting place. These may be brokers, agents, people. The commodity price is determined by the demand and supply rules. The Bombay Stock Exchange is the most prominent stock exchange in India. India has a total of twenty-one stock exchanges.

FUNCTIONS OF THE STOCK EXCHANGE:

- ***Liquidity and Marketability***: One of the stock exchange's key selling points is that it provides for high liquidity. At a time, the securities can be sold and converted to cash. It is a continuous market, and as per their wishes, investors can easily divest and reinvest.

- ***Price Determination***: The only way to determine the value of shares in a secondary market is through the rules of supply and demand. This process is enabled by a stock exchange through constant valuation of all securities. These stock prices of different companies can be monitored through the Sensex index.

- ***Safety***: The government regulates and administers the stock exchanges strictly. India's Securities Board is the governing body in the case of BSE. All transactions take place within the context of the law. This provides assurances and a safe place for securities transactions to the investor.

- ***Contribution to the Economy***: As we know, stock exchange deals in securities that have already been issued. But these securities are being sold and resold continuously, and so on. It allows mobilizing and channelizing the funds rather than sitting idle. This will boost the economy.

- ***Spreading of Equity***: The stock exchange ensures greater securities ownership. In reality, it educates the public on the safety and benefits of investing in the stock market. This guarantees improved payment performance and smooth operation. The aim is to draw more public investors and expand stock ownership for everyone's benefit.
- ***Speculation***: One often learns that a financial market is the stock exchange. And while this is true, the speculation remains within the framework of the law. A healthy dose of speculative trading is required for the liquidity and price determination stake, and the stock exchange provides us with such a platform[16].

TRADING AND SETTLEMENT PROCEDURE:

1. Selecting a Broker or Sub-broker

If a person wishes to trade in the stock market, they are unable to do so in their individual capacity. Only a broker or a sub-broker can make the transactions. Therefore, a broker must be named according to one's requirement. Now such a broker can be a person or a partner, a company or a financial institution (like banks). They have to be licensed with SEBI. Once such a broker has been hired, on the stock exchange you will buy / sell stocks.

2. Opening a Demat Account

Both securities are now available in electronic format since the reforms. External shares / securities issues no longer exist. A shareholder must therefore open a dematerialized account, i.e. a Demat account to hold these digital securities and trade them in them. So, with the depository participant, you or your broker will open a Demat account. There are actually two depository members in India, namely Central Depository Services Ltd. (CDSL) and NDSL.

3. Placing Orders

And then the investor actually places an order for shares to be bought or sold. The order will be placed with his broker or if the broker provides such services, the individual will be able to transact online. One important thing is that the order / instructions should be very simple. Example: Buy 100 XYZ Co. shares at or below Rs. 140/-. The broker will then function on the basis of your transactions and place an order for the shares at the price listed or, if appropriate, an even better price. The broker must send a slip to the shareholder for verification of order.

4. Execution of the Order

Once the investor's order is issued by the broker, he executes it. Within 24 hours, a contract notice must be given by the broker. The document contains all the transaction information, such as the number of shares transacted, the purchase cost, date and time, the sum of the brokerage, etc.

An important document is the Contract Note. It is proof of the contract in the case of a legal dispute. It also includes the Special-Order Code that the stock exchange assigns to it.

5. Settlement

It moves the real securities from the buyer to the seller. And the cash is also going to be exchanged. The broker will also handle the change here. There are two types of settlements,

- On the spot settlement: here we instantly swap the funds and the transaction follows the pattern of T+2. A transaction that takes place on Monday will therefore be resolved by Wednesday (by the second working day)
- Forward Settlement: Simply means that both parties have decided to settle on some future date. It may be T+7, T+9 and so on.

TYPE OF ORDERS:

The most common types of orders are market orders, limit orders, and stop-loss orders.

- A **market order** is an order to immediately buy or sell a security. Such form of order guarantees the execution of the order, but does not guarantee the execution value. In general, a market order will be executed at or close the current bid (for a sell order) or request (for a purchase order) value. For traders, however, it is important to remember that the last-traded price is not always the price at which a trading order is executed.
- A **limit order** is an order to buy or sell a security at or better than a specific price. A sales limit order can only be executed at or below the limit price and can only be executed at or above the limit price. Example: An investor wants to buy ABC stock shares for a total of $10. For this amount, the shareholder must request a limit order and this order will only be executed if the ABC stock price is $10 or less.
- A **stop order**, also known as a stop-loss order, is an order to buy or sell a stock once the stock price exceeds the price stated, known as the stop price. A stop order becomes a market order when the stop value is hit.

- At a **stop price** above the current market price, a buy-stop order is entered. Usually, investors use a buy-stop order to reduce a loss or make a gain on a stock they have sold short. At a stop price below the current market price, a sell stop order is entered. Generally speaking, investors use a sell stop order to limit a loss or protect a profit on a stock they own.

FUTURES AND OPTIONS:

Futures are financial derivative contracts that allow the parties to trade an asset at a predetermined future date and value. The buyer must purchase or the seller must sell the underlying asset at the set price, regardless of the current market price at the expiry date. Natural commodities or other financial instruments are the underlying assets. Future contracts describe the value of the underlying asset and are structured to promote a futures exchange trading. Futures can be used to hedge or speculate about trade. Futures — also called futures contracts— allow traders to lock up an asset or commodity underlying value.

These contracts have dates of expiration and set up front-known prices. Futures are known by their month of expiration. A gold futures contract in December, for instance, expires in December. The word futures appear to represent the market as a whole.

Options are financial instruments based on the value of underlying securities such as shares that are derivatives. A contract of options gives the purchaser the opportunity to buy or sell the underlying asset, depending on the type of contract they possess. Unlike futures, if they choose not to buy or sell the asset, the holder is not required.

- Call options allow the holder to buy the asset at a stated price at or before a specific time.
- Put options allow the holder to sell the asset at a stated price at or before a specific timeframe.

Each contract option will have a specific expiry date by which the holder will have to exercise its option. The price stated on an option is known as the price of the strike. Usually, options are purchased and sold through online or retail brokers.

CHAPTER 2

DESCRIPTION OF DATA USED

As stated in introductory texts, High Frequency Trading is all about high speed processing and computational decision making, hence to perform fast algorithmic calculations by powerful computers, it needs to be provided with data at low latencies. The exchanges play a key role in providing the data to its clients and brokerage houses by using a globally used colocation service which was explained earlier. The data provided by exchanges as constant feeds to the clients needs to be handled very carefully and smartly at the client's end in-order to cope up with the competing latencies as High Frequency Trading is an arms race at the end of the day. Different firms use different in-house technologies and methodologies to store and process the data. A brief description of the different types of data used is given below.

STANDARD DATA-FORMS PROVIDED/USED BY EXCHANGES:

1. **Tick-by-tick Data:**

   As the name suggests, this is the data feed that gets triggered each time the there is some change in the limit order book. The limit order book or L2, as stated earlier is nothing but a record of all the standing orders i.e. all bid and ask quotes along with their respective sizes. It also contains other relevant information required to characterize the order like the instrument id which marks the instrument to which the order/state is referring to, timestamp which is the EPOCH time when the particular order/action occurred, exchange IDs and matching exchange IDs for trade which map the orders with a unique exchange generated id at the time of logging and is henceforth to unique trace orders in case of any modifications and cancellations. The LOB is essentially a matching engine for buyers and sellers in the market [6]. Within a LOB, the best bid (ask) price is the highest (lowest) price a market maker is willing to buy (sell) the asset at to market takers. The maximum number of contracts that the market makers are willing to buy (sell) at the bid (ask) price is called the best bid (ask) volume. Any market taker who wishes to buy (sell) at the counterparty price can submit a market order to trade at the best ask (bid) price up to the ask (bid) volume available. If the market order to buy (sell) is larger than the ask (sell) volume, then they will walk the book; the market taker will continue buying (selling) at the next-best ask (bid) price until their entire market order is filled.

This data feed is the of the most granularity of all the data types used in the capitals market. The time difference between subsequent ticks is of the order of 10^{-9} seconds i.e. the order of nanoseconds. This is the level of speed at which the industry works and in order to be competent, the participants have to work towards optimizing and improving their algorithms and system infrastructure in order to cope up.

In simple terms, tick-by-data, also known as tbt data is the most granular and simple form of information about an instrument that is traded at the exchanges. Hence most High Frequency Trading algorithms use this data as an input for their decision making and subsequent trading. Due to higher granularity and complexity (less time difference between subsequent ticks) in terms of latency, it takes a relatively more time to process and perform computations on tick-by-tick data than snapshot data.

2. **Broadcast Data:**

Broadcast data, better know as snapshot data is a somewhat less granular and less informative version of the tick-by-tick data stream. It is simply just a snapshot of limit order book at fixed intervals of time (in our case 200ms). It tells us the state of limit order book at fixed and regular intervals of time. As the time difference between subsequent data feeds is fixed and not very small (i.e. of milliseconds rather than nanoseconds in tick-by-tick data), it is easier and faster to process snapshot data than tick-by-tick data. However, the snapshot data might not contain continuous information for a particular instrument as the instruments in which the state remains same for the given time interval, it skips getting the snapshot of those instruments and hence we need to use forward filling of data in order to get continuous data for any given interval.

The above two data-forms are not explicitly provided by the exchanges in the stated format, rather the data is listened by firms through a standard colocation protocol and is further stored and processed according to the firm's discretion.

RELEVANT DATA-FORMS USED/STORED BY IRAGECAPITAL:

1. **Contract Headers:**

   Contract header is a simply a mapping of instrument ID's with the relevant information of that instrument, such as the name, expiry date, strike price (in case of Options), lot size etc. This is a standardized table made to list down all instruments and store standard information about them which remain fixed throughout their lifetime.

2. **FIX Logs:**

   The Financial Information Exchange (FIX) Protocol has revolutionized the trading environment, making it important to promote many of the developments in electronic trading that have arisen over the past decade. FIX has become the language used extensively by buying and selling companies, trading platforms and even regulators to communicate trade information on the global financial markets. The non-proprietary, free and open model is being continuously improved to meet emerging market and regulatory needs and is being used every day by thousands of companies to complete millions of transactions. FIX is the way the world trades and it becomes an essential ingredient in lowering the price of trade, optimizing efficiencies and achieving greater transparency. FIX offers significant benefits for businesses interested in exploring new investment opportunities; it lowers market entry costs for participants able to communicate easily both domestically and internationally, as well as significantly reducing switching costs. The vocabulary of the FIX Protocol consists of a series of messaging standards used in commercial communications. Originally developed to support pre-trade and trading stock trading, it is now undergoing rapid expansion into post-trade space, supporting straight-through processing (STP) from interest signals (IOI) to allocations and confirmations. It is also seeing a significant increase in fixed income, foreign exchange and derivative listed markets. The FIX messaging standard is owned, maintained and developed through FIX Trading Community ™ member firms ' collaborative efforts, which include many of the world's leading financial institutions. Many companies are working together to ensure that the standard continues to meet evolving trading standards and facilitate its increased adoption, which offers enormous potential benefits throughout the financial community.

3. **Strategy Logs:**

   Strategy Logs, also known as strat logs, are records of the company's in-house strategies for trading. All decisions taken, actions taken and their subsequent impact on parameters such as statements on profit and loss, Greek options, open positions, etc. are all stored in strategy logs. Strategy plays a key role in determining the performance parameters and in further analyzing any strategy the company is implementing. Strategy logs make debugging and subsequent development much easier and allow any trade arbitrage to develop. Strategy Logs, also known as strat logs, are records of the company's in-house strategies for trading. All decisions taken, actions taken and their subsequent impact on parameters such as statements on profit and loss, Greek options, open positions, etc. are all stored in strategy logs. Strategy plays a key role in determining the performance parameters and in further analyzing any strategy the company is implementing. Strategy logs make debugging and subsequent development much easier and allow any trade arbitrage to develop

CHAPTER 3

IMPLIED VOLATILITY CURVE AND MARKET MAKING

MARKET MAKING:

The pace and ease of buying and selling stocks is often taken for granted. Use your broker to place an order and it will be executed in seconds. Market makers are a major reason why such transactions can occur so quickly. Whenever an asset is bought or sold, the other end of the exchange must be someone.

Market makers make stock markets literally, hence their name. These are usually banks or brokerage firms that are ready with firm request-and-bid rates every second of the trading day. Without market makers, matching buyers and sellers would take considerably longer, lowering liquidity and potentially increasing trading costs as it would be more difficult to enter or exit positions. Since investors and traders prefer to buy and sell quickly, financial markets need to operate smoothly. There would probably be fewer transactions without market makers and the overall markets would slow down. This, in turn, would reduce the amount of money the companies have available.

Market makers need to quote prices on an ongoing basis and their volumes are willing to buy and sell at. This helps to maintain market consistency. In times of volatility, the buying and selling process helps market makers willing to buy and sell at established prices to maintain normalcy. Buyers might find it hard to get into a hot stock without their presence, or sellers might find themselves unable to sell a stock if their price goes south.

IRageCapital is a very old and prominent player in the Options Market Making market, being involved in the business since the advent of High Frequency Trading in India (2009). It was said to have more than 80% market share in the Options Market Making Regime then, however nowadays due to significant increase in the number of competitors, it now has roughly 50% market cap in the regime.

IMPLIED VOLATILITY AND THE CURVE:

Volatility Smiles are patterns of implied volatility that arise in financial options for pricing. It matches finding a single parameter (implied volatility) that needs to be modified to fit market prices for the Black–Scholes formula. For a given expiration in particular, options whose strike price differs significantly from the price command of the underlying asset higher prices (and thus implied volatility) than suggested by standard option pricing models. It is said that these options are either in-the-money or out-of - the-money.

Implied volatility can be defined as the ambiguity of the underlying stock of an option and the adjustments caused at the trading rates of different options. IV is the dominant view of the market that the underlying asset is likely to reach a given value. In-, at-and out-of - the-money refers to an options contract's strike price as it relates to that asset's current market price. Volatility skew is important to watch when buying and selling options because the volatility implied increases as the uncertainty surrounding the underlying stock increases. Instead of the expected flat surface, graphing implied volatility against strike prices for a given expiry yields a skewed "smile." The trend varies across different markets. Before the crash of 1987, equity options traded on American markets did not show a smile of volatility, but started to show one afterwards. Investor reassessments of the fat-tail probabilities are believed to have led to higher out - of-the-money options prices. This phenomenon suggests shortcomings in the default pricing model of the Black–Scholes option which assumes constant uncertainty and log-normal distribution of underlying returns on capital. Empirical asset returns distributions, however, tend to exhibit fat-tails (kurtosis) and skew. Modelling the volatility smile is an active area of research in quantitative finance, and better pricing models such as the stochastic volatility model partially address this issue. A related concept is that of term structure of volatility, which describes how (implied) volatility differs for related options with different maturities. An implied volatility surface is a 3-D plot that plots volatility smile and term structure of volatility in a consolidated three-dimensional surface for all options on a given underlying asset.

MARKET MAKING AND THE CURVE:

As stated above, market makers are essentially liquidity providers. In the case of Options Market Making, and specifically at iRageCapital, the use of Implied Volatility Curve is of utmost importance. The Implied Volatility curve is used as a primary tool to determine the prices of options, all across the option chain. It gives a simplified method to determine the price of the entire option chain, by simply using the value of curve at points of desired strike price and also giving an easy way to understand the distribution of prices across the entire option chain and the relevant information like that associated with the forward skew and the backward skew. The entire price quoting on both the sell side and the buy side of the Limit Order Book is done using the Implied Volatility Curve, along with some other parameters that are used to facilitate the quoting. Hence, for market makers, solution of volatility curve at all points during their trading session is very important. Therefore, the analyses of dynamics of the Volatility Curve is very helpful in order to attain better efficiency and increased profitability. As stated earlier, in High Frequency Trading, speed plays a key role in order to be competitive, hence faster and accurate algorithms of fitting the Implied Volatility Curve are desirable.

OBJECTIVES OF PROJECT:

The objective of the project was to analyse and study the characteristics, behaviour and the factors governing the Implied Volatility Curve (aka Volatility Smile/IV Curve/Vol Curve/Volatility Skew) of the **BANKNIFTY** index of the Nation Stock Exchange of India Ltd. (NSE).

Bank Nifty represents the 12 most liquid and large capitalized stocks from the banking sector which trade on the National Stock Exchange. It provides investors and market intermediaries a benchmark that captures the capital market performance of Indian banking sector.

Being a sectoral index, BANKNIFTY only represents the banking sector, thus a lot of numerical and non-numerical information can be drawn from studying the nature of its Vol-Curve on daily basis. It is also one of the most actively traded index and option in the F&O market, hence information about the market/investor sentiments, performance of the banking sector and the effects of national & international events on Indian banking sector as a whole can be drawn out from the analysis of the BANKNIFTY Vol-Curve. There are many possible approaches for the curve analysis, the one used in this project was **Spline Regression.**

Composition of the BANKNIFTY index:

As mentioned earlier, BANKNIFTY is a sectoral index, diversified across 12 banking stocks listed on the National Stock Exchange. The constituents are listed below:

Company's Name	Weight (%) as on Apr 29th, 2016
HDFC BANK	32.41 %
ICICI BANK	20.01%
AXIS BANK	11.44%
KOTAK MAHINDRA BANK	10.70%
STATE BANK OF INDIA	8.53%
INDUSIND BANK	7.43%
YES BANK	4.49%
BANK OF BARODA	2.17%
FEDERAL BANK LTD.	1.15%
PUNJAB NATIONAL BANK	0.94%

Table 1: Composition of BANKNIFTY index

BENEFITS OF SPLINE REFRESSION:

Linear regression is a widely used statistical technique in public health to investigate the relationship between continuous dependent and independent variables. It is important to use models that closely match the data for any regression technique. Transformations of the response variable will boost fit and correct violations of design assumptions such as constant variance of errors. Predictor variables can be divided into logical classes or additional terms can be added which are functions of existing predictors such as quadratic or cubic terms. Certain methods, such as spline modeling, provide a better fit, taking into account the variability in the relationship between the predictor variable and the variable response, both within and between predictor variable rates. Nevertheless, there is no best approach, as some modeling methods, depending on the data, that produce better results for expected values (e.g. smaller confidence intervals) than other methods. Spline analyzes are often simplistic and definitions are often nuanced.

METHODOLOGY**:**

There are two major steps involved in the analysis of Vol-Curve:

1. Gathering data and calculating Implied Volatilities using appropriate formulae, and
2. Plotting the Volatilities for the entire Option Chain and studying the curve.

The data used for the project was the Broadcast Data provided by the National Stock Exchange to its paid subscribers and brokers. This data essentially contains the below mentioned fields on a Nano-second level for every trading session i.e. it is logged and broadcasted to subscribers by the National Stock Exchange at frequency of one Nano-second on every trading day between 9:15 IST and 15:30 IST. The fields are:

- **Epoch Time-Stamp**: Time in EPOCH format for the given tick of the data
- **Instrument**: Name of the financial instrument to which the data belongs (e.g. BANKNIFTY).
- **Instrument Type**: The type of instrument i.e. either of Futures (XX), Call Option (CE) or Put Option (PE)
- **Expiry**: The expiry date of the instrument
- **Strike Price** (for Options): The Strike Price of the Option at which it will be executed.
- **Underlying Instrument**: Since F&O are derivatives, they always have a underlying asset associated with them, which is the actual asset on which they function.
- **L2 Information**: L2 refers to the Limit Order Books maintained by the exchanges in order to maintain a record of all the orders (bid and ask) registered for the instrument. The information is essentially the quoted price and the quoted quantity upto 5 levels, which the market participants wish to buy or sell.
- There are many other fields in the Broadcast Data such as Open Interest, Current Open Interest, Closing IV etc. which are not relevant to the project and hence not mentioned

The aforementioned data is obtained in a very crude form from the National Stock Exchange and needs to be cleaned in order to be used for further processing. Many data cleaning and handling procedures had to be done for in order to enable further computations and relevant processing using the data, details of which are not mentioned here.

After the data is ready, Implied Volatilities of each and every option had to be calculated using some known formula. The most commonly used model for Implied Volatility calculation is the **Black-Scholes Model.**

Brief description of The Black Scholes Model:
A mathematical model for pricing a contract option is also known as the Black-Scholes-Merton (BSM) model. The equation calculates the variance over time of financial instruments such as shares, in particular, and using the underlying asset's implied volatility extracts the value of a call option. The model assumes that the value of highly traded capital follows a Brownian geometric motion with continuous uncertainty and drift. The model involves the constant price variance of the stock, the time value of money, the strike price of the option, and the period to the expiry of the option when applied to a stock option. It was the first commonly used model of option pricing, also called Black-Scholes-Merton. It is used to calculate the theoretical value of options based on current stock prices, expected dividends, the strike price of the option, expected interest rates, expiry time and expected volatility.

The formula, developed by three economists, Fischer Black, Myron Scholes and Robert Merton, is perhaps the most well-known model of pricing options in the world. It was introduced in the Journal of Political Economy in their 1973 paper, "The Pricing of Options and Corporate Liabilities." Two years before Scholes and Merton were awarded the 1997 Nobel Prize in Economics for their work in finding a new method for determining the price of derivatives (the Nobel Prize is not granted posthumously; however, the Nobel Committee acknowledged Black's role in the Black-Scholes model).

The Black-Scholes model makes certain assumptions:

- The option is European and can only be exercised at expiration.
- No dividends are paid out during the life of the option.

- Markets are efficient (i.e., market movements cannot be predicted).
- There are no transaction costs in buying the option.
- The risk-free rate and volatility of the underlying are known and constant.
- The returns on the underlying are normally distributed.

While the original Black-Scholes model did not consider the effects of dividends paid during the lifetime of the option, by determining the ex-dividend date value of the underlying stock, the model is often adapted to account for dividends.

The formula of Black-Scholes Model:

The formula's mathematics are complicated and can be intimidating. The calculation of the Black Scholes call option is determined by the cumulative standard normal probability distribution function multiplying the stock price. Then, from the resulting value of the previous calculation, the net present value (NPV) of the strike price multiplied by the cumulative standard normal distribution is subtracted.

As it can be clearly seen that the formula was devised originally for calculating the actual price of the option given the Volatility (and of course other information) for it, the formula was reverse engineered in order to get Volatility given the price of the option. Since the price input is the actual price taken from real time data, it is intuitive to call the Volatility thus obtained as Implied Volatility (IV), because it is the Volatility implied by the current standing prices in the market. The IV calculations were done for each and every instrument present in the broadcast data at every time instant.

In mathematical notation:

$$c = SN(d_1) - Xe^{-rT}(d_2)$$
$$p = Xe^{-rT}N(-d_2) = SN(-d_1)$$

where

$$d_1 = \left(ln\,(S \div X) + (r + \sigma^2 \div 2)\,T\right) \div \sigma\sqrt{T}$$
$$d_2 = \left[\left(ln\,(S \div X) + (r - \sigma^2 \div 2)\,T\right) = d_1 - \sigma\sqrt{T}\right] \div \sigma\sqrt{T}$$

c = call (European style)
p = put (European style)
S = Stock price
X – Strike price of the option
r = Risk-free interest rate
T = Time to expiration (in years)
σ = Volatility of the relative price change of the underlying stock price
N(x) = The cumulative normal distribution function

CURVE FITTING:

The curve plotting methodology and implementation cannot be disclosed under the non-compete and non-disclosure agreement which was signed by me as a compulsory requirement before joining as an Intern. Hence the organisation doesn't authorize disclosing any information regarding the same, therefore the details are not mentioned.

CHAPTER 4

VOLUME-ORDER IMBALANCE IN LIMIT ORDER BOOK

Since traders submit limit orders to buy (sell), they affect the volumes of the limit order book bid (ask) and thus give us a view of the intentions of the traders. Classifying the volume of trade as either bid (ask) would enable us to gain insight into the direction of future price changes. To measure this trade purpose, we look at the difference between the volume of the offering and the volume of the offer, called the imbalance of the order. Order imbalance is an essential descriptor that enables us to understand the market's general sense and direction. When knowledgeable traders have data that has not yet been included in the asset price, due to the positive (or negative) news, they will take a long (or short) position and then increase the asset imbalance. Certain market participants, who merely observe this trend in the LOB, could use this knowledge to develop a strategy for generating positive returns. The next section will examine carefully the relationship between order imbalance and mid-price changes and decide whether it can be used to predict future high-frequency price changes. We also examined the statistical properties of the order imbalance and how to apply them to a trading strategy in order to produce statistically significant positive returns on a daily basis.

VOLUME ORDER IMBALACE:

The order imbalance is defined using the algorithm of Lee and Ready [9] to classify trades as either initiated by the buyer or initiated by the seller. This is done by checking if the price of trade is closer to the quoted price (selling) or asking (buying). Instead, our description is more analogous to the that we will call Volume Order Imbalance (VOI):

$$OI_t = \delta V_t^B - \delta V_t^A$$

where

$$\delta V_t^B = \begin{cases} 0, & P_t^B < P_{t-1}^B \\ V_t^B - V_{t-1}^B, & P_t^B = P_{t-1}^B \\ V_t^B, & P_t^B > P_{t-1}^B \end{cases}, \qquad \delta V_t^A = \begin{cases} V_t^A, & P_t^A < P_{t-1}^A \\ V_t^A - V_{t-1}^A, & P_t^A = P_{t-1}^A \\ 0, & P_t^A > P_{t-1}^A \end{cases}$$

where V^B_t and V^A_t are the bid and ask volumes at time t respectively and P^B_t and P^A_t are the best bid and ask prices at time t respectively. If the new bid price is lower than the previous bid price, which means either the trader cancelled his purchase limit order or a P^B_{t-1} order was completed. Since we don't have a more granular order or message book, we can't be sure of the purpose of the trader, so we're setting V^B_t=0. If the current price of the bid is the same as the previous price, we take the difference in the volume of the bid to represent the incremental pressure of the last period. Finally, if the current price of the bid is higher than the previous price, this can be viewed as an upward demand trend due to the desire of the seller to buy at a higher price. Downward price momentum and sell pressure can be interpreted analogously from the current and previous ask prices [7] [8] [15].

ORDER IMBALANCE RATIO:

The VOI calculates only the extent of the imbalance that is not enough to explain the market traders ' conduct. For instance, if the current volume of bid change is 300 and the current volume of request change is 200, the VOI is 100, which is considered to be a strong buy signal. However, this does not take into account the ratio between the volume of the bid and the volume of the request indicating the strength of the market's potential buyers.

We formulate a new parameter knowns as the Order Imbalance Ratio (OIR), given by:

$$\rho_t = \frac{V_t^B - V_t^A}{V_t^B + V_t^A}$$

This aspect complements the disparity in volume order by helping us to distinguish cases where the difference is high but the ratio is small. In the above example, the OIR is only 0.2, indicating that after all[3][4][9] may not be as strong as the original signal to buy.

OTHER PARAMETERS:

Apart from the VOI and OIR, we also include a way of classifying trades as approved buyers or sellers. We can determine the average trade price between two time stages using the traded volume and turnover information in the data set. We define the Average Trade Price, TP_t from $(t - 1, t]$ as:

$$\overline{TP}_t = \begin{cases} M_1, & t = 1 \\ \frac{1}{300} \frac{T_t - T_{t-1}}{V_t - V_{t-1}}, & V_t \neq V_{t-1} \\ \overline{TP}_{t-1}, & V_t = V_{t-1} \end{cases}$$

Where Tt is the turnover (CNY trading volume) and V_t is the t-time transaction size. This mechanism represents the average value paid by other market participants in their transactions to be viewed as a proxy for trade imbalance. We may classify trades as more buyer (seller) initiated by checking whether T_P is closer to the offer (bid) value. Rather than a binary classification, though, we define the factor as the distance of the average trade price from the average mid-price over the time-step $(t - 1, t]$:

$$R_t = \overline{TP} - \frac{M_{t-1} + M_t}{2} = \overline{TP}_t - \overline{MP}_t$$

Where M_t is the t-time mid-price. Due to its mean reversive properties, the factor R_t, which we call the mid-price basis (MPB), is an important predictor of price change. It provides an ongoing classification of whether trades were initiated by buyers or sellers. A large positive (negative) quantity means that, on average, the trades were closer to the price of the bid [1][2][6].

STATISTICAL ANALYSIS OF PARAMETERS:

Positive autocorrelation of the order imbalance as shown in Figure. The autocorrelation imbalance of the order is significant up to lag 15 for most days.

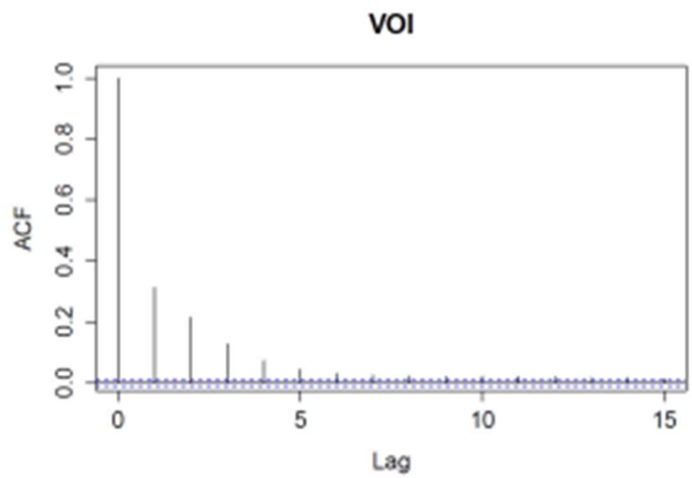

Figure 1: Plot of Auto-correlation of VOI against different Lags

Their first difference has a significant negative self-correlation lag-1 and is consistent with the results. This indicates that positive (negative) imbalances are often accompanied by recurrent positive (negative) imbalances as traders split their orders over multiple periods. See below in Figure 2.1.

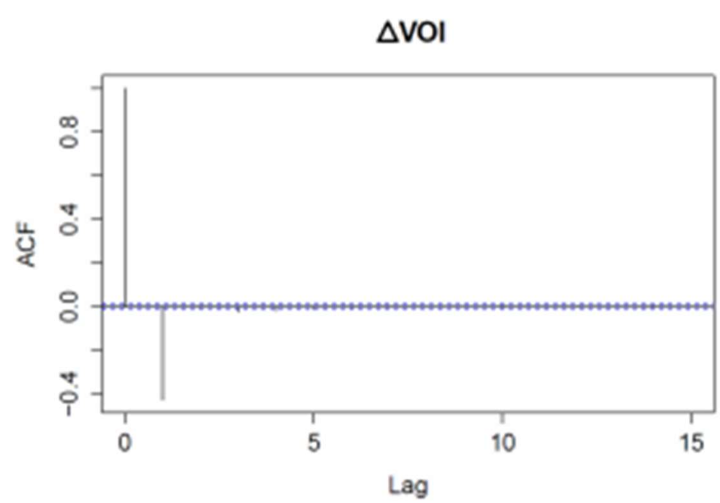

Figure 2: Plot of Auto-correlation of change in VOI against different Lags

We also find a positive link between the VOI and contemporary price changes. That is, the $OI_t = M_t - M_{t-1}$ correlation is 0.40, although the relationship is not as good. This positive relationship is shown in figure 2.2 below.

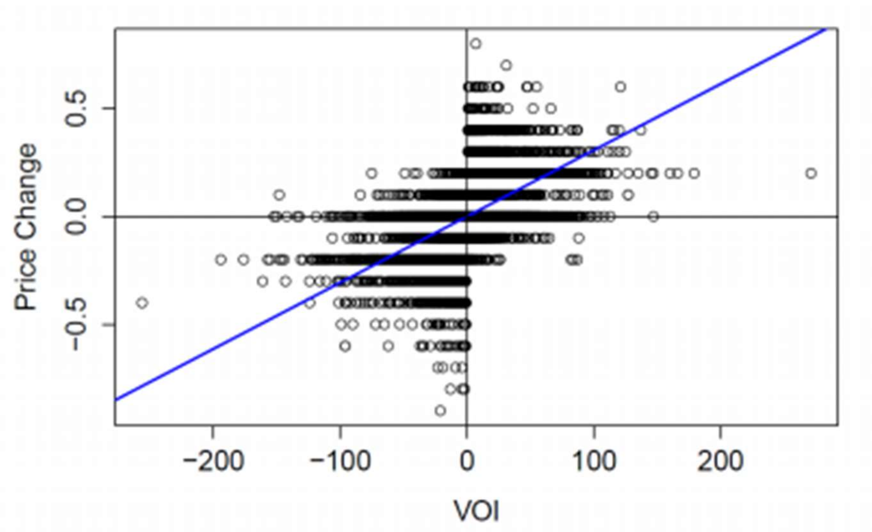

Figure 3: Scatter Plot of Instantaneous Mid-Price Change for different VOI values

In addition, fitting a contemporary linear model an average daily R2 of 0.155= $\alpha + \beta OI_t + \pi t$. Even if we change the VOI definition to match the order flow imbalance definition, the average R2 is only improved to 0.294. Since studies are carried out in the space of high frequency, the difference in R2 is probably due to the different time scales used in the analysis. By following the 10th time interval, we can get a 0.6537 daily average R2 that is consistent [10].

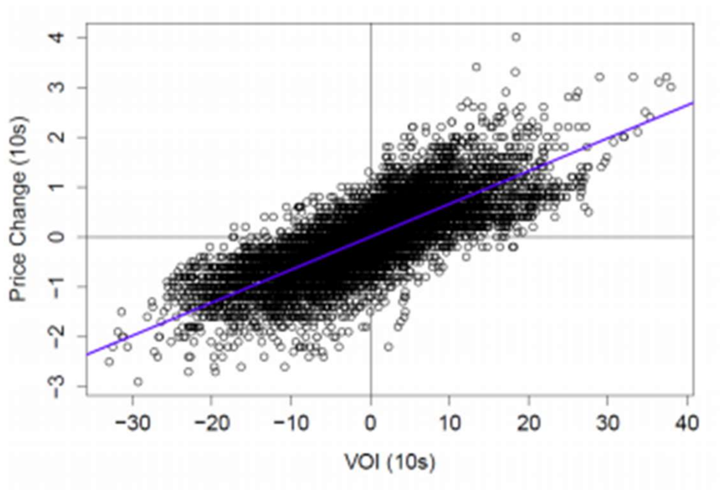

Figure 4: Scatter Plot of Mid-Price Change after 10s for VOI values

The OIR is another measure of the imbalance in order and should share similar statistical characteristics with VOI. Figure 3.1 displays the autocorrelation below and they share the same signs and identical magnitudes with the VOI autocorrelation in Figure 2.1.

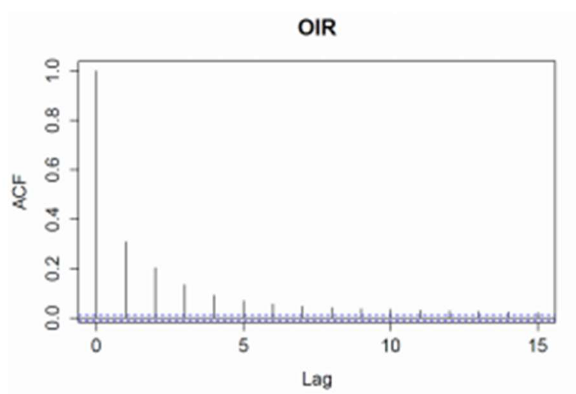

Figure 5: Plot of Auto-correlation factors against different lag values

Nevertheless, we find that the OIR-contemporary price change relationship is actually the opposite of VOI. The correlation between −0.4126 is the -0.4126 correlation. Because πt's autocorrelations are also large and positive as VOI was for the first 5 lags, we can conclude that this is an analogous reflection of traders ' order splitting behaviour. Through default, a high OIR means that at a given time, the volume of the bid is much greater than the volume of the offer, meaning that many traders intend to buy, and very few intend to sell. But as a large OIR is associated with a negative price shift, it means that more traders are willing to buy when prices drop. This result shows how the orders are separated over time as opposed to the autocorrelations which suggest only the existence of order-splitting [13].

We expect R_t to return to mean 0, so if $R_t > 0$, the mid-price eventually rises and returns to the average price of trade and if $R_t < 0$, then we would expect the mid-price to fall back to the average price of trade. So, when $R_t > 0$ and when $R_t < 0$ sells the signal, we have a buy signal. The positive R_t (MPB) relationship with the average mid-price (response) shift is shown in Figure 3.3 below.

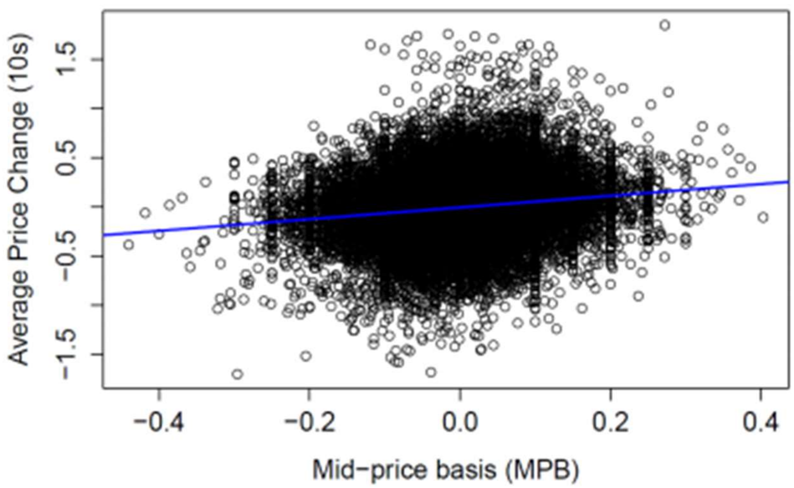

Figure 6: Scatter Plot of Average Mid-Price change with mid-price basis

PARAMETER SELECTION AND MODELLING:

We integrate the new OIR and MPB features as described in our linear model respectively. The spread adjustment will also be included in each feature by dividing the spread. Below is the final linear model[14].

$$\overline{\Delta M}_{t,k} = \beta_0 + \sum_{j=0}^{L} \beta_{OI,j} \frac{OI_{t-j}}{S_t} + \sum_{j=0}^{L} \beta_{\rho,j} \frac{\rho_{t-j}}{S_t} + \beta_R \frac{R_t}{S_t} + \varepsilon_t$$

where $\Delta M_{t,k} = (1/h)*\sum^{k}_{j=1} (M_{t+j} - M_t)$ is the k-step average mid-price change, OI_{t-j} is the j-lag Volume Order Imbalance from the previous strategy, ρ_{t-j} is the j-lag Order Imbalance Ratio, R_t is the instantaneous mid-price basis, and S_t is the instantaneous bid-ask spread.

For Volume Order Imbalance and Order Imbalance Ratio, we also parameterize lag L. β_0 is the constant term for coefficients, $\beta_{OI,j}$ corresponds to the spread-adjusted j-lag VOI, $\beta_{\pi,j}$ corresponds to the spread-adjusted j-lag OIR, and β_R corresponds to the spread-adjusted MPB. It is assumed that πt errors are independent and distributed with zero mean and constant variance in the same way as normal. This model will be built using ordinary least squares linear regression. The parameters are set to k = 20 and L = 5 even though they may not be optimal [11] [8].

Lastly, therefore, we will use the linear model of the previous day to predict today's 20-step average mid-price change and trade only if the change is above 0.2 (below -0.2). There is R2= 0.0815 in the linear model.

IMPLEMENTATION:

The final implementation was done by programming the logic in python language. Libraries like numpy, pandas, scikit learn were used rigorously for the programming.

Due to compliance reasons which were stated earlier, the source code cannot be disclosed.

CHAPTER 5

RESULTS AND ANALYSIS DISCUSSION

IMPLIED VOLATILITY CURVE:

The final output of the project is a model coded in Python Language, which takes as input the broadcast data and generated IV values plotted on a curve for the entire option chain of the BANKNIFTY index.

The model is being tested on regular basis, being subjected to all kinds of scenarios that emerge in the day-to-day market sessions. The model was also tested for days which had a lot of market activity, or were of national importance (e.g. vote counting day) and affected the market in a very heavy manner.

Snapshots and relevant interpretations of the resultant curve are mentioned hereafter.

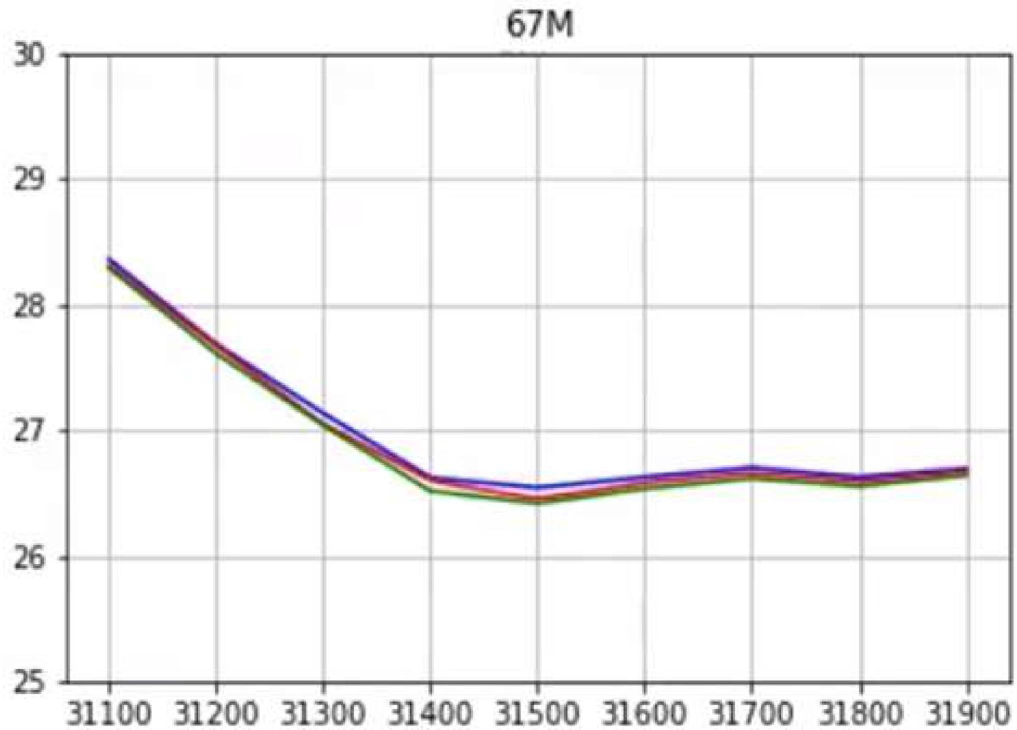

Figure 7: BANKNIFTY Vol curve as solved by our model on 29th May 2019

As we can see in above figure, the Vol-Curve was in a transition period from a Smile (U-shape) to a shape having negative slope, called as reverse skew.

The X axis denotes the strikes available in the option chain, whereas the Y axis denotes the IVs for the options.

The volatility smile skew pattern is commonly seen in near-term equity options and options in the forex market. Volatility smiles tell us that demand is greater for options that are in-the-money or out-of-the-money. The smile shows that the options that are furthest in- or out-of-the-money have the highest implied volatility. Not all options will have an implied volatility smile. Near-term equity options and currency-related options are more likely to have a volatility smile.

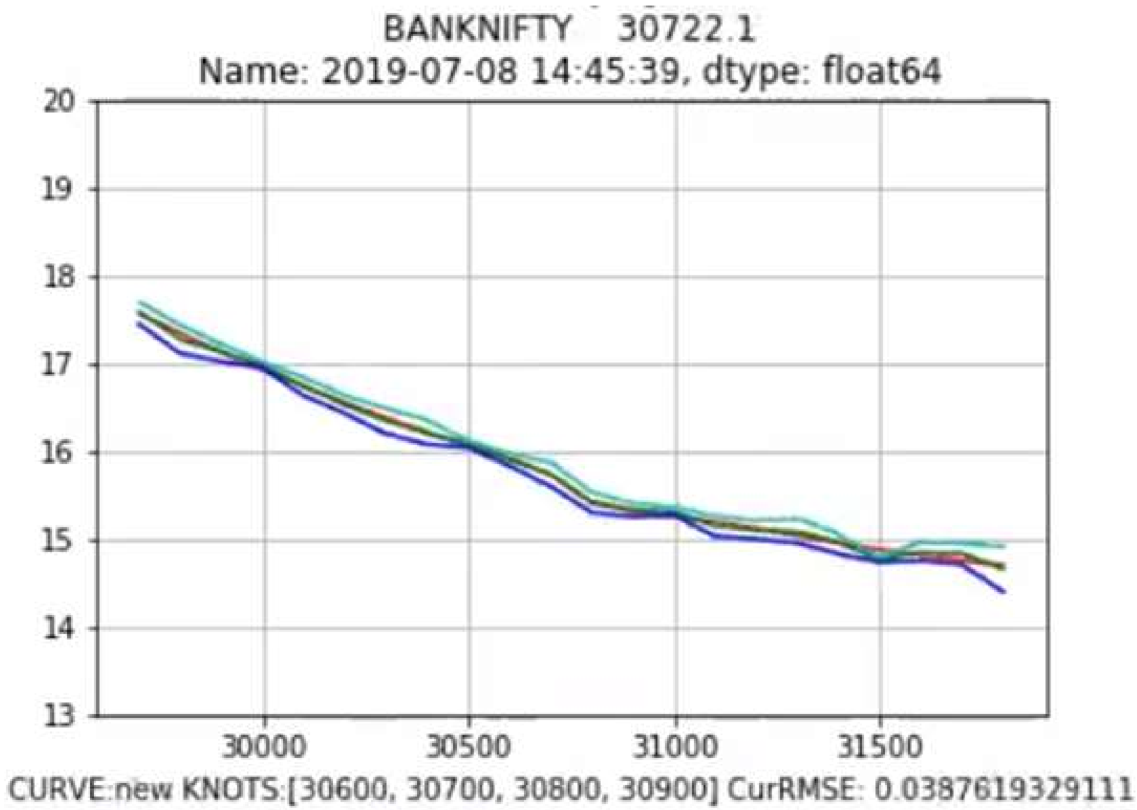

Figure 8: BANKNIFTY Vol curve as solved by our model on 8th July 2019

In the above figure the curve has started to show reverse skew in a pronounced manner and has a tendency to drop down even further.

In the reverse skew pattern, the IV for options at the lower strikes are higher than the IV at higher strikes. The reverse skew pattern suggests that in-the-money calls and out-of-the-money puts are more expensive compared to out-of-the-money calls and in-the-money puts.

The popular explanation for the manifestation of the reverse volatility skew is that investors are generally worried about market crashes and buy puts for protection. Another possible explanation is that in-the-money calls have become popular alternatives to outright stock purchases as they offer leverage and hence increased ROI. This leads to greater demands for in-the-money calls and therefore increased IV at the lower strikes.

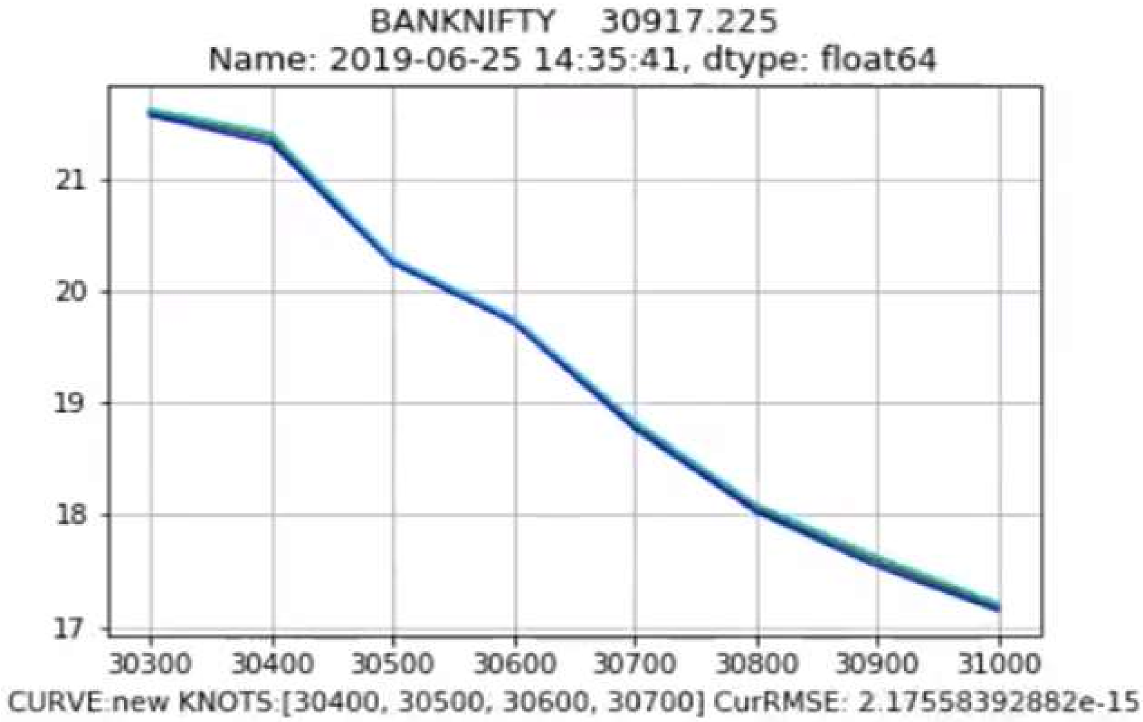

Figure 9: BANKNIFTY Vol curve as solved by our model on 25th May 2019

In the above figure, a huge reverse skew is overserved depicting very high volatilities of Puts than the Calls. Which means that the risk of market crashing which the investors were speculating before, which caused the distortion of smile, has actually come in action and now the market has actually started to crash. Volatility represents a level of risk present within a particular investment. It relates directly to the underlying asset associated with the option and is derived from the options price. The IV cannot be directly analyzed. Instead, it functions as part of a formula used to predict the future direction of a particular underlying asset. As the IV goes up, the price of the associated asset goes down. Hence, from the curve we can predict that in the market bears are in control and absolute shorting is occurring.

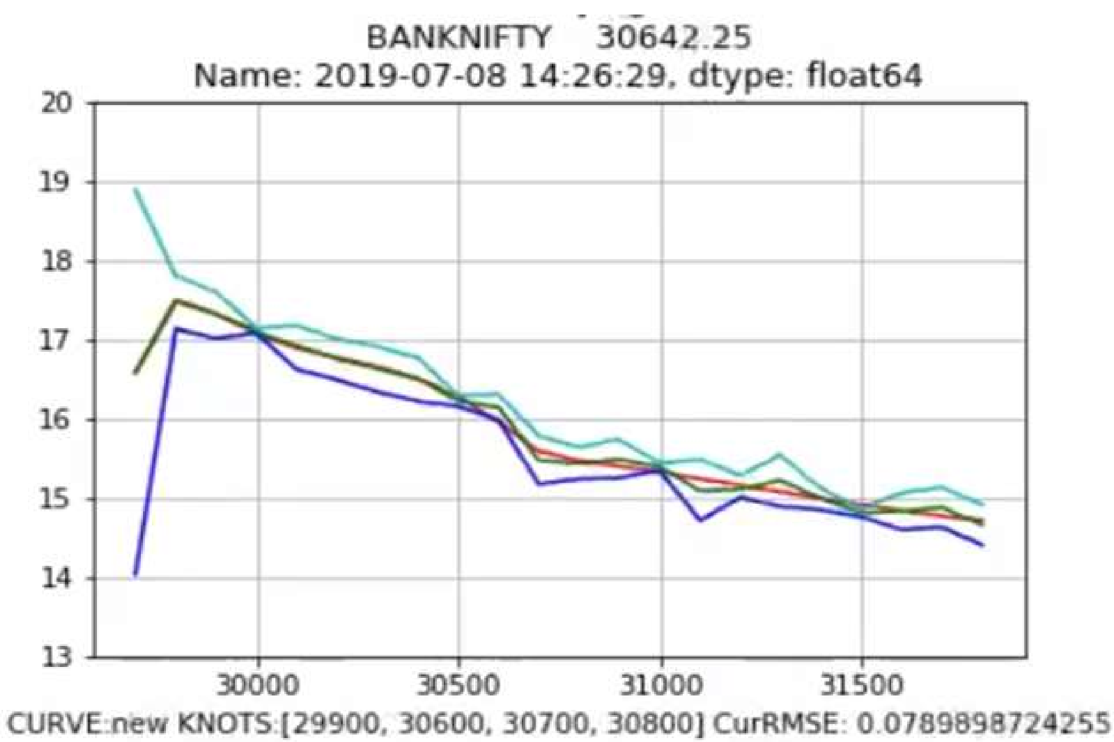

Figure 9: BANKNIFTY Vol curve as solved by our model on 8th July 2019

In the above figure, another type of market sentiment is depicted. Here a wide spread can be observed between the bid (bottom blue) and the ask IVs (top cyan), which depicts uncertainty and panic in the market. In such scenarios, the participants are not sure as to what is the actual/fair price in order to buy/sell the asset and hence in order reduce their risk and adverse cost of selection, they start trading at passive prices and hence the difference of price between buyers and sellers increases. This stage is generally observed in transitioning market conditions, as here also the skew is apparently approaching a reverse skew. Such patterns also arise before some news event that is expected to bring about a drastic change in the market, but in an unknown direction, and hence everyone start playing safe in their trades. Another case where this may occur is when there is extremely low liquidity in the market and hence the power to govern prices are in the hands of very few participants, and therefore they intend to monopolise the market with their own speculations.

VOLUME-ORDER IMBALANCE IN LIMIT ORDER BOOK:

Below is the table which contains the results of tested strategy. It is being tested for multiple portfolios on intraday basis. The results attached are for Punjab National Bank and Tata Steel Futures traded on the National Stock Exchange, under the symbols PNB and TATASTEEL respectively. These are results for 4 consecutive days.

The strategy was implemented on data by breaking it into 13-time wise equal intervals, so that any dependency on time or seasonality can be recognized.

Section of the day	Hit Percentage	Buy Count	Sell Count	PnL (INR)	Txn. Cost
1	85	42	42	35686	3514
2	80	44	44	34820	3680
3	96	63	64	572529	5321
4	87	83	82	-454236	6936
5	NA	0	0	0	0
6	47	4	4	1765	335
7	73	41	42	550215	3485
8	97	227	225	-897192	19042
9	49	256	256	174761	21589
10	70	251	251	169963	21137
11	14	237	237	158457	20043
12	81	68	70	1107839	5861
13	73	33	32	-512378	2778
NET	71	996	364	942230	113720

Table 2: Profit and Loss statement (PnL) of 16th July 2019 PNB's August expiry Futures data tested by our strategy.

In table 2, it can be seen that a greater number of triggers were achieved in the later half of the trading session. This is in agreement with the general observation of increased activity during the latter half of the trading session in general. Also, the transaction costs were significant at around 12% of the profit made. Also the average success rate of signals was relatively mediocre at around 71, along with a high volatility with peaks at as high as 97 and dips to 47 in some sections of the day, which is a very undesirable situation and should be rectified at all costs in order to attain better profitability.

Section	Hit %	Buy Count	Sell Count	PnL (INR)	Txn. Cost
1	77	41	40	-481093	3343
2	70	85	87	1093319	7081
3	65	78	78	45775	6374
4	69	52	52	30745	4254
5	62	35	34	-494216	2816
6	56	28	28	18007	2292
7	66	44	44	27192	3607
8	72	35	34	-492136	2835
9	69	46	48	1050700	3849
10	74	72	71	-465056	5855
11	57	82	82	46120	6729
12	73	91	91	50967	7482
13	67	81	81	42004	6645
NET	67	770	770	472331	63168

Table 3: Profit and Loss statement (PnL) of 17th July 2019 PNB's August expiry Futures data tested by our strategy.

In table 3, higher activity and triggers during the latter half of the day can be seen again. Here, the transaction seems to be increased to roughly 14% of the profit made. Futures trading are characterized by higher transaction costs, due to which only strong directional alphas are successful in this regime. Here the profit has also reduced significantly. Also the average success rate of signals was relatively low at around 67, but were seen to be consistent without much broader peaks and deeper dips, which further explains the decrease in revenue and the increase in transaction costs.

Section	Hit %	Buy Count	Sell Count	PnL (INR)	Txn. Cost
1	54	33	33	33243	2806
2	61	67	67	50354	5645
3	83	63	64	571097	5352
4	76	51	49	1018518	4217
5	85	41	42	549869	3480
6	89	60	61	552156	5043
7	58	53	52	-488052	4351
8	67	45	44	-482822	3672
9	66	72	74	1080769	5980
10	69	68	67	-463686	5536
11	44	77	77	52807	6342
12	88	87	87	66313	7186
13	51	63	63	41709	5190
NET	70	780	780	545242	64807

Table 4: Profit and Loss statement (PnL) of 18th July 2019 PNB's August expiry Futures data tested by our strategy.

In table 4, the seasonality of higher activity in the second half of the day is reverified and the profit has also slightly increased from the last session by roughly 15%. Here, the transaction cost has also reduced to roughly 10 percent of the profit made. Lower transaction costs are always desirable, but are hard to achieve in the case of future's trading. Also the average success rate of signals was relatively low at around 70, with dips to as low as 44 in some sections, which further explains the decrease in revenue and the increase in transaction costs.

Section	Hit %	Buy Count	Sell Count	PnL (INR)	Txn. Cost
1	77	41	40	-481093	3343
2	70	85	87	1093319	7081
3	65	78	78	45775	6374
4	69	52	52	30745	4254
5	62	35	34	-494216	2816
6	56	28	28	18007	2292
7	66	44	44	27192	3607
8	72	35	34	-492136	2835
9	69	46	48	1050700	3849
10	74	72	71	-465056	5855
11	57	82	82	46120	6729
12	73	91	91	50967	7482
13	67	81	81	42004	6645
NET	67	770	770	472331	63168

Table 5: Profit and Loss statement (PnL) of 19th July 2019 PNB's August expiry Futures data tested by our strategy.

In table 5, a new pattern of higher number of triggers during the first half of the day is observed. One possible reason for this maybe the press conference by our respected Financial minister Mrs. Nirmala Sitharaman the previous day, where a lot of important announcements regarding the taxation and slowdown in the Indian Economic markets were made. In the case above, the transaction cost is at 13% of the profit made. Also the average success rate of signals was relatively low at around 67, which further explains the decrease in revenue and the increase in transaction costs. The signal strength was however consistent throughout the day with much deviation from the average observed value.

Section	Hit %	Buy Count	Sell Count	PnL (INR)	Txn. Cost
1	49	93	94	422357	5810
2	86	99	97	-748667	6073
3	69	120	122	811323	7504
4	90	151	149	-732168	9255
5	93	143	143	35919	8802
6	92	0	1	383521	31
7	82	115	115	39264	7049
8	88	155	154	-333575	9439
9	83	130	132	799329	7933
10	87	194	192	-692767	11605
11	83	188	188	67775	11217
12	81	129	129	46273	7679
13	85	192	193	430408	11340
NET	84	1709	1709	528991	103737

Table 6: Profit and Loss statement (PnL) of 17th October 2019 TATASTEEL's October expiry Futures data tested by our strategy.

In table 6, it can be again seen that a greater number of triggers were achieved in the latter half of the trading session. This is in agreement with the observed pattern for the PnL sheets for PNB Futures. But here, the transactions have drastically increased to 20% of the earned profit. This can be attributed to the fact that the price of TATASTEEL futures were nearly 7 times to that of PNB at the time of observation, hence the cost of a round trade are meant to increase. However, here the average signal strength or hit ratio is at very high levels of around 84, which should normally give better yields than other cases, but again the transactions costs hindered away this effect. Also, at the start of the day, very few number of triggers and very poor signal performance is observed, which further adds to increase in transaction costs.

Section	Hit %	Buy Count	Sell Count	PnL (INR)	Txn. Cost
1	72	53	54	390917	3191
2	79	90	88	-715735	5289
3	81	141	143	782197	8460
4	80	136	135	-331149	8180
5	86	111	112	410277	6749
6	82	55	54	-365481	3308
7	84	95	96	405561	5789
8	85	117	117	32990	7115
9	68	141	141	31460	8593
10	82	93	92	-353063	5638
11	80	106	107	413119	6507
12	78	121	121	35459	7405
13	74	117	116	-355331	7111
NET	80	1376	1376	381223	83336

Table 7: Profit and Loss statement (PnL) of 18th October 2019 TATASTEEL's October expiry Futures data tested by our strategy.

In table 7, it can be seen that the number of triggers were consistent throughout the day and spread evenly. But here, the transactions have drastically increased to 22% of the earned profit. This can be attributed to the fact that the price of TATASTEEL futures were nearly 7 times to that of PNB at the time of observation, hence the cost of a round trade are meant to increase. However, here the average signal strength or hit ratio is at very good levels of around 80, which should normally give better yields than other cases, but again the transactions costs hindered away this effect. Also, at the start of the day, very few number of triggers but a good signal performance is observed, which kind of neutralizes the damage by increase in transaction costs.

Section	Hit %	Buy Count	Sell Count	PnL (INR)	Txn. Cost
1	81	60	59	-355819	3567
2	72	174	176	800717	10470
3	86	141	139	-709982	8343
4	90	156	156	32124	9255
5	89	158	159	406497	9415
6	85	93	92	-350809	5454
7	87	94	94	21470	5533
8	87	181	181	40250	10572
9	91	145	145	32290	8453
10	92	143	143	35652	8327
11	88	107	109	755318	6321
12	88	96	95	-345903	5588
13	64	85	85	18564	4990
NET	86	1633	1633	380367	96288

Table 8: Profit and Loss statement (PnL) of 19th October 2019 TATASTEEL's October expiry Futures data tested by our strategy.

In table 8, it can be again seen that the number of triggers were consistent throughout the day and spread evenly. But here, the transactions have drastically increased to 25% of the earned profit. This can be attributed to the fact that the price of TATASTEEL futures were nearly 7 times to that of PNB at the time of observation, hence the cost of a round trade are meant to increase. However, here the average signal strength or hit ratio is at very good levels of around 86, which should normally give better yields than other cases, but again the transactions costs hindered away this effect. Also, at the start of the day, very few number of triggers but a good signal performance is observed, which kind of neutralizes the damage by increase in transaction costs.

Section	Hit %	Buy Count	Sell Count	PnL (INR)	Txn. Cost
1	54	78	79	391965	4637
2	82	59	57	-722251	3423
3	61	104	106	763130	6201
4	52	171	171	50270	10314
5	56	261	259	-694037	16111
6	80	3	4	391769	217
7	45	240	241	472431	15258
8	51	186	186	65600	11906
9	59	156	156	55223	10028
10	57	182	182	57144	11768
11	68	198	196	-744965	12716
12	63	205	205	90079	12944
13	73	186	187	452429	11758
NET	59	2029	2029	628787	127281

Table 9: Profit and Loss statement (PnL) of 20th October 2019 TATASTEEL's October expiry Futures data tested by our strategy.

In table 9, a new pattern of higher number of triggers during the middle of the day is observed. One possible reason for this can be announcement of a critical news related to the corporate affairs of the TATA Steel industry during the trading session. In the case above, the transaction cost is at 19% of the profit made, which is apparently low as compared to the other datasets of TATASTEEL. Also the average success rate of signals was relatively very low at around 59, which further explains the decrease in revenue and the increase in transaction costs. The signal strength was also volatile during the day and can be seen to take extreme values.

CHAPTER 6

CONCLUSION AND FUTURE SCOPE OF WORK

By the work done on Volume-Order Imbalances in Limit Order Books for Indian Markets, it can be clearly seen that like the USA markets, Indian Markets also have a pronounced presence of Volume and Order Imbalances. A lot of information regarding the market sentiment, speculations about the direction and magnitude of the mid-price can be drawn out of the Volume-Order Imbalance parameters. This effect is however varying for different entities due to the fact that different stocks have different type of activity at a particular point of time, and hence the trading quantities are very different for them and therefore the information extracted will have different means and is a subject of further study in this area for Indian Markets. Looking at the results, it can also be concluded that despite having high strength signals at times, the performance of the strategy does not improve. This makes it difficult for a trader to completely rely on this strategy and idea to work as a standalone strategy or indicator. This suggests that this information can be used as a reinforcement signal to other indicators that one might already be having and hence work as a good collective strategy. The poor performance of the strategy can also be partly attributed to the fact that due to regulatory and the current policy structure, futures have high transaction cost associated with, thus it is difficult to maintain profitability at all times and therefore weak signals get beaten out very badly in the case of adverse situations. However, here the directional information achieved from the Volume-Order Imbalance parameters is used to take positions in the futures market. Other possibilities and variants like taking appropriate positions and performing signal execution in other derivative segments like Future-based Options still remain untouched and definitely will be one of the important points in any future extension of this work.

The choice of solving the Implied Volatility curve using Spline Regression methodology turned to be a promising one, as higher accuracy with little variances were attained during the testing of the model. Splines, as are known to have smoothing and simplifying characteristics when used for curve fitting, performed very similarly in the case of Implied Volatility Curve as well. The model showed very smaller number of bid-ask spread breaches and hence carried high value of sanity with it at all points of observation. It can be confidently concluded that Spline based Implied Volatility Curve solutions can be used for Market Making Purposes and we can expect it to show good performance in its application. It however remains a subject of future findings that this solution is applicable universally or just Indian markets specifically. If the data is made available in the required format, this work should be further verified for foreign markets as well. Another import aspect of Market Making of Option using Spline solved Volatility Curve still remains unexplored.

REFERENCES

[1] Irene Aldridge. *High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems.* John Wiley & Sons, Inc, 2010.

[2] Am´elie Charles and Olivier Darn´e. *Variance ratio tests of random walk: An overview. Journal of Economic Surveys,* 23(3):503–527, 2009.

[3] Tarun Chordia, Richard Roll, and Avanidhar Subrahmanyam. *Order imbalance, liquidity, and market returns. Journal of Financial Economics,* 65:111–130, 2002.

[4] Tarun Chordia and Avanidhar Subrahmanyam. *Order imbalance and individual stock returns: Theory and evidence. Journal of Financial Economics,* 72:485– 518, 2004.

[5] Rama Cont, Arseniy Kukanov, and Sasha Stoikov. *The price impact of order book events. Journal of Financial Econometrics,* 12(1):47–88, 2014.

[6] Martin D. Gould, Mason A. Porter, Stacy Williams, Mark McDonald, Daniel J. Fenn, and Sam D. Howison. *Limit order books. Quantitative Finance*, 13(11):1709–1742, 2013.

[7] Han-Ching Huang, Yong-Chern Su, and Yi-Chun Liu. *The performance of imbalance-based trading strategy on tender offer announcement day. Investment Management and Financial Innovations,* 11(2):38–46, 2014.

[8] Jonathan Karpoff. *The relation between price changes and trading volume: A survey. Journal of Financial and Quantitative Analysis*, 22(1):109–126, 1987.

[9] Charles M.C. Lee and Mark J. Ready. *Inferring trade direction from intraday data. The Journal of Finance,* 46:733–747, 1991.

[10] Xuan Liu. *Stationary processes and hft.* Master's thesis, University of Oxford, 2015. 36

[11] Rahul Ravi and Yuqing Sha. *Autocorrelated Order-Imbalance and Price Momentum in the Stock Market. International Journal of Economics and Finance*, 6(10):39–54, 2014.

[12] Ruey S. Tsay. *Analysis of Financial Time Series*. John Wiley & Sons, Inc, 2010.

[13] Geroge H.K. Wang and Jot Yau. *Trading volume, bid-ask spread, and price volatility in futures markets. Journal of Futures Markets,* 20(10):943–970, 2000.

[14] Zhaodong Wang and Weian Zheng. *High-Frequency Trading and Probability Theory.* World Scientific Publishing Co. Pte. Ltd, 2015.

[15] Darryl Shen, *Order Imbalance Based Strategy in High Frequency Trading*, Linacre College, University of Oxford.

[16] www.investopedia.com

[17] blog.quantinsti.com

[18] Sheldon Natenberg, *Option Volatility and Pricing (1994)*, Chapter 1-7, McGraw Hill Publications.

[19]: Sarves Verma, Gunhan Mehmet Ertosun, Wei Wang, Benjamin Ambruster, Kay Giesecke, *Implied Volatility Modelling*, https://web.stanford.edu/class/msande444/2008/ivts2report.pdf

[20]: Seung-Ho Yang et. Al., *Modeling Implied Volatility Surfaces Using Two-dimensional Cubic Spline with Estimated Grid Points*, IEMS Vol. 9, No. 4, pp. 323-338, December 2010