The Analysis of High Frequency Data of FTSE 100 on Brexit Referendum Day

Busara Jindayen

04 October 2019

Contents

Abstract . . . 2
Introduction . . . 3
- High Frequency Data . . . 3
- Volatility . . . 4
- FTSE 100 . . . 6
- Bid and Ask Price . . . 7
- Brexit Referendum . . . 7

Literature Review . . . 8
- Brexit and Market Volatility . . . 8
- Stochastic Process (Brownian Motion) . . . 9
- Instantaneous Volatility . . . 11

Methodology . . . 14
- Manipulation of High Frequency Data . . . 14
- Volatility with time-independent . . . 16
- Volatility with time-dependent . . . 17

Analysis . . . 21
Conclusion . . . 27
References . . . 28
Appendix . . . 31

Abstract

This paper is concerned with the estimation of volatility using Brownian Motion from a Stochastic Process. The model is followed the form of time-dependent Brownian Motion: $dP_t = \mu_t dt + \sigma_t dB_t$, where P_t denotes the log prices, dP_t is the return and σ_t is Semi-martingale. From the research of (Alvarez et al. 2011), the instantaneous volatility is calculated by Semi-martingale $\sigma_t^{(p=2)}$ where $p = 2$ is a quadratic variation. The time-dependent Brownian Motion model includes the volatility, in which volatility is provided by the estimators when the microstructure noise has shown in the stock price, especially in bid and ask bounce. The instantaneous volatility is built from the time-dependent Brownian Motion model and the instantaneous volatility has an estimator. The finding of Realised Volatility in high frequency data such as the bid and ask prices from UK stock index, FTSE 100 in different time interval, the result gives that the realised volatility is a good estimator of the instantaneous volatility.

The result of this paper answers the following questions:

1. Can the estimator of instantaneous volatility of FTSE 100 specifically analyse the events on three dates of the Brexit referendum day? The graphs of estimator indicate the high volatility when the significant events occured which has effected FTSE 100.

2. Can the estimator of instantaneous volatility of FTSE 100 describe the risk on the Brexit referendum day? The volatile graph of estimator decribes how much risk of FTSE 100 has in different time interval precisely.

3. Is the realised volatility a good estimator of instantaneous volatility to determine the volatility in specific time interval? The realised volatility is definitely a good estimator which estimates the volatility in precise time scale such as 5 minutes and 1 minute. The realised volatility estimator nearly matches to the instantaneous volatility.

Introduction

The log return of stock prices can be modelled, the extension of the Standard Brownian Motion: $dP_t = \mu dt + \sigma dB_t$ that has a volatility parameter (σ) which is a Stochastic Process can be estimated using quotes data such as bid and ask prices of stock index, FTSE 100. This paper is focusing on how to use high-frequency price data to understand better time-varying risk properties of the investment. There are two types of risk: the average volatility risk and the jump risk. The estimation of volatility, but not average volatility, is the main topic of this paper, the volatility of returns measures how much the return moves up and down around its average value.

High Frequency Data

High Frequency Data is the collection of data based on frequency in every millisecond or microsecond. The millisecond is a thousandth (10^{-3}) of a second, the microsecond is a millionth (10^{-6}) of a second. The London Stock Exchange (LSE) services the live data in microseconds to analysts, the real time data is costly because it is valuable for any financial instituition. The data can be collected at a sufficient frequency such as milliseconds or microseconds, which is mostly used in financial analysis, algorithm trading and high frequency trading. Even though, the milliseconds data is less frequent than the microseconds data but the milliseconds can be useful, it can be downloaded from an online source such as Swiss Banking Group (Dukascopy Bank SA 2016), therefore the milliseconds data is going to be used in this paper instead. High frequency data provides intraday prices that can apply for a better understanding of market behaviours. (Aldridge 2013) The forms of high-frequency data that are going to be used in this paper, are bid and ask price. The details of bid and ask price are in the section below.

The tick size was introduced which is the minimum variation in the price, it is the smallest movement in quoting or trading a security. For example the stocks in FTSE 100 can be traded on LSE, there are many tick sizes in which the stocks value between £1 and £5. The tick size is 0.00005 and there are 10 different bands. The FTSE 100 stocks can trade at £6200.00 or £6200.50 pence but not £6200.0001, each stock markets have established different tick sizes. The London Stock Exchange has a more complicated regime in which tick sizes vary not only by price but also by the type of stock. Tick size's rules protect the investors who have displayed trading interest to the market by making other investors pay a large amount more than them and the rules also have an impact on the trading costs paid by investors as well as on the overall market. (Angel 2012)

The tick size represents in the trading process, it brings about creating a floor for bid and ask spread. (see section the Bid and Ask Price) The tick size also impacts on the price time priority rules of the limit order book. A limit order is an order to trade that sets at specific price, in the past the orders were written in the book. (Angel 2012) For example, an investor places the order to purchase 100 shares of FTSE 100 at £5 which means the investor is not going to pay more than £5 and the order cannot be operated at higher price but it can be operated at lower price. The collection stocks of FTSE 100 is a liquid stock so it has lower tick sizes. (Angel 2012) The stock price or stock index, such as FTSE 100 are discrete random variable because they can take on certain values with certain probabilities, such as £6262.16 and £6262.17 but not £6262.10008. However, the stock returns are considered to be continuous random variable. (Morgan 2009) The Brownian Motion is the solution of Stochastic Differential Equation (SDE) which in SDE also includes the random variable but the stock returns are nonrandom growth because the stock return depends on the previous return, whereas the log of stock return is continuous random variable. This property of stock prices dictates the choice of mathematical model.

High Frequency Data is used in investment banks and hedge funds. The markets encourage liquidity by setting a bid and ask quotes. Those institutions usually have computer controlled high-speed, trading programs are never completely error-free.(Linton et al. 2011) The risk and manipulations are harmful to the market participants and cannot be compensated for the benefits. For example, the first crash of 2010 when the US stock market collapsed within a few minutes. (Trotman 2015) The flash crash is the security prices falling in the short time, since then mini flash crashes have repeatedly been observed in every stock exchange such as an issue the German stock market experienced its first flash crash in early 2014. (Maiello 2017)

The European Stock Exchange acts were adopted in the spring of 2014 which the introduction of automated trading stops if flash crashes occur trading will be interrupted briefly flash crashes cannot be prevented, but the damages can at least be limited.

Volatility

Volatility quantifies risk, it can be defined mathematically as the standard deviation of log returns of underlying asset calculated over a time period. The volatility of bid and ask prices as underlying asset has a parameter estimation. The mathematical models can be used where the asset prices are treated as random variables. In this paper, the method of measuring the stock volatility has $\hat{\sigma}$ as unbiased estimator of standard deviation of the log return prices. The estimator is unbiased because the mean of the estimator sampling distribution equals to the true value for the parameter. With an assumption of standard time series when time is evenly spaced which means the spacing of time of stock prices are constant. The estimation of volatility of stock prices where the stock prices cannot be negative. The stock returns move up and down which means the positive and negative returns. The stock prices behave similarly to exponential function so the stock prices can take the natural logarithm. Statistically, the practical way of estimating volatility uses the log return prices.

The one period of log return is a discrete time stochastic process;

The price series are not independent on each other, the random variable is normally the log return prices. Suppose the price P_t over time $t = 1, .., n$. The measurement of log-return prices r_t can be defined by the geometric returns which is in the term of the logarithm of the current price P_t divided by the previous price P_{t-1} as follow:

$$r_t = ln\frac{P_t}{P_{t-1}}$$

where $r_1 = 1$

The multiple period of log return is a continuous time stochastic process;

The k period log return where $r_t(k) = ln\frac{P_t}{P_{t-k}}$ is the sum of the k period log return:

$$r_t(k) = r_t + r_{t-1} + ... + r_{t-k+1}$$

The collection of stock prices as random variables is not independent and identically distributed (i.i.d) because the previous prices do influence on the current price and the chances of having the same prices are not identical. However, in some models, such as GBM are i.i.d because they only take a positive variation.

The mean $\overline{r}$ of the log-return prices is given by:

$$\overline{r} = \frac{1}{n}\sum_{t=1}^{n} r_t$$

where n is the number of returns.

The estimator of the standard deviation of the r_t is given by:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}(r_t - \overline{r})^2}$$

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n} r_t^2 - \frac{1}{n(n-1)}\left(\sum_{t=1}^{n} r_t\right)^2}$$

As it is mentioned in Literature Review section, the time-independent Standard Brownian Motion:

$$dr_t = \mu dt + \sigma dB_t$$

where μ is the expected of return

B_t is the Wiener Process (Brownian Motion)

The mathematical technique that is used in Brownian Motion is the stochastic calculas. The sample standard deviation (an estimator) is a statistical measure. The volatility is defined as the standard deviation (σ) of the log returns.

In finance, Realised Volatility (RV) is the popular volatility measurement, also known as Historical Volatility is the calculation of volatility for log returns from high-frequency data over short time interval during the trading day. (Barndorff-Nielsen and Shephard 2002) The realised volatility is going to be calculated for the Brexit data in the analysis. The formula is shown below:

$$RV = \sum_{t=1}^{n} r_t^2$$

where r_t is log return prices of underlying asset

n is the total number of day of underlying asset

$$r_t = ln \frac{P_t}{P_{t-1}}$$

where the units of time t is $t = 1, .., n$

The market microstructure noise is the noise from the price bounce which makes high frequency data estimates parameter, such as realised volatility unstable. Taking into an account the microstructure noise existing in the high-frequency data:

$$Y_t = r_t + \epsilon_t$$

is observed over the interval [0,1] in discrete time over a grid $t_i = \frac{i}{n}$ for $i = 0, 1, ..., n$

ϵ_t is independent of the r_t process. (Zu and Boswijk 2014)

The realised volatility considers the market microstructure noise because the small number of time points in the calculation of realised volatility, the microstructure noise is over represented in the realised volatility estimator. In this paper, the microstructure noise is not included in the analysis, the effects of noise are small, especially in bid and ask bounce effects when the price of stock rapidly bounce within the range between bid and ask price.(see section the Bid and Ask Price)

The discussion about nonparametric estimation of instantaneous volatility (σ_t) with high frequency data where nonparametric means it is not based on probability distribution that provides the probabilities of different outcomes in the experiment. There are three models with all the log return prices $r_1, r_2, ..., r_t$ that can be estimated volatility as follow:

$$dr_t = \mu_t dt + \sigma dB_t$$

$$dr_t = \mu_t dt + \sigma dB_t$$

$$dr_t = \mu_t dt + \sigma_t dB_t$$

The last model was chosen, it is the time-dependent Brownian Motion with Semi-martingale. The instantaneous volatility has been involved with the Brownian Semi-martingale which was extended from Brownian Motion Process and it is the parametric estimation unlike the instantaneous volatility σ_t it is the nonparametric estimation:

Let X_t be log price process

$$dX_t = \mu_t dt + \sigma_t dB_t$$

where B_t is a Standard Brownian Motion

μ_t is the instantaneous drift process

σ_t is the instantaneous volatility process

when $\mu_t = 0$, it derives this in Literature Review section

since X is a Brownian Semimartingale, it has a continuous sample paths and the process σ_t is positive, its quadratic variation ($p = 2$) can be determined from σ_s^p satisfies:

$$\langle X, X \rangle_t = \int_0^t \sigma_s^2 ds$$

as parametric estimation where $t \in [0, 1]$

The instataneous volatility satisfies

$$\sigma_t^2 = \frac{d\langle X, X \rangle_t}{dt}$$

(Zu and Boswijk 2014)

FTSE 100

The FTSE 100 stands for the Financial Times Stock Exchange 100 index. A financial index is a measurement of a section of the overall stock market. The FTSE 100 contains the top 100 UK publicly listed companies by market capitalisation. These are grouped, and their combined value tracked. The index is seen as a good gauge on how the overall UK economy is doing, although with many of the companies on the FTSE 100 being international. It is debated how accurate a measurement it offers. (Partridge 2016) For example, TECHMARK 100 is a UK index that tracks technology companies. By grouping stocks, the idea is to get an overall picture of a more significant segment of the market. That could be a country's economic health, a technology sector, the industrial sector etc. It started in January 1984 with a start value of 1000, and it is now wholly owned subsidiary of the London Stock Exchange (*LSE*). Companies on the FTSE 100 must also meet a free float requirement, which means that the general public should hold at least 25% of the companies' shares. It is the 100 largest companies in terms of market capitalisation that are permitted into the FTSE 100. (*FTSE Uk Index Series* 2019)

For example, a company in FTSE 100 has 1000 shares outstanding, and they are valued at £10 each at closing. That means the market capitalisation would be £10,000. If companies were ejected as soon as they fell below 100, they would be constant changes and unnecessary volatility. This depends on the market capitalisation of any companies and the company can rise back to the index before the FTSE100 will be updated. (Hardy 2019) Members are decided at quarterly reviews, when market capitalisation determined, companies above position 90 usually gets into the FTSE 100 and below 111 usually takes out, dropouts typically move into the FTSE 250. The LSE has the high performance market data system which its provides real time tick-by-tick data for anyone trading on their market. The real time data is available for the specific needs of individual customers. On the other hand, the FTSE 100 is only published every 15 seconds on weekdays between 8 am and 4.30 pm GMT (excluding holidays) which is maintained by LSE. (*Real Time Data* 2019) Fortunately, there is the real time data of FTSE 100 index is publicly available which is updated in milliseconds. This is valuable for the analysis of high frequency data of FTSE 100.

Bid and Ask Price

All markets have two prices, a price that people buy which is the Bid price and a price people sell out which is the Ask price. First, there are the closing price of the share and the markets typically have a spread a bid price and ask price. The spread is the difference between the price that is sold at and the amount that is bought. Likewise, the spread will vary for markets that see a lot of volumes so broader shares indices, currencies and commodities. The financial instruments tend to have very tight spread but the share that doesn't trade that much, they will have a wider spread. It shows how much the market has to move to be break-even.(Crux 2013) The different shareholders have different prices, they are prepared to sell.

For example, if a company's shares are selling at £2.50 that is the ask price, the lowest price anyone is prepared to sell the company's shares now. If the bid price is at £2.45, which is the highest figure anyone is prepared to pay for the company's shares right now. The difference between these two figures is the spread which is 5p. For new small companies, the spread can be wide because their share is not accessible yet. There is less competition for them. It might be harder to find a buyer if people wanted to sell quickly.(Crux 2013) Consistently, shares are highly liquid because shares are sold and bought all the time. However, if bad news hit company durables, there will be less competition for shares, and it is spread may widen. The bigger the spread, the greater the uncertainty the market has about that share, and the more people could lose if they needed to sell suddenly.(Crux 2013) An approximation of intraday bid and ask spreads provides accurate measures and computional saving. (Chung and Zhang 2014 ; Fong, Holden, and Trzcinka 2017)

The use of milliseconds time stamps is more precise whereas the usage of monthly trade and quote data which might conduct to incorrect statistical conclusion. The high frequency data consists bid and ask quotes and their sizes. The bid and ask data arrive discrepant and they are concerned of the quote process. The higher the bid and ask spread, there is an increase of the security to cover the spread. In tick data, the increment of the price can be comparable or smaller than the bid-ask spread and the bid and ask spread typically increase during the market uncertainty. The bid and ask bounce is different from traditional data that is based in just the closing prices. Tick data bring additional supply and demand information in the bid and ask prices and offering sizes. The bid and ask quotes can carry useful information about the maket behaviours which can be managed to the researcher's advantages. It is sometimes difficult to use bid and ask spread because continuous movement from the bid to ask and back introduces a jump process, cumbersome to deal with many models. (Aldridge 2013)

The construction of the main high-frequency benchmark can be calculated by the mid-range of effective price which the equation is defined by the B_t is the bid price and A_t is the ask price at time t as shown below:

$$M_t = \frac{B_t + A_t}{2}$$

M_t is the mid-point of bid and ask price.(Abdi and Ranaldo 2017)

Brexit Referendum

On the 23rd of June 2016, the British people were asked to choose whether United Kingdom's should continue to be a member of the European Union. The choice was simple: Remain, and continue membership as usual; or Leave, and withdraw from the EU.

European Economic Community (EEC) was created in 1957. The UK joined the EEC in 1973. The 2016 Referendum was a third referendum EEC. Joining was done under the Conservative government of Edward Heath, but after a change of leadership, Labour Prime Minister, Harold Wilson, disagreed with the decision to join and put it to the British people to vote on the issue. The referendum was decisively in favour of remaining, as 67% of the people voted to stay. In 1992, the European Union (EU) was created, as the organisation shifted to become more of a political union, rather than merely an economic one. (*European Union – Guide* 2019)

The United Kingdom has always enjoyed a somewhat special status within the EU. Shortly after the EU's creation, the Euro was introduced, a single shared currency to be used throughout Europe, across different nations. Most countries dropped their local currency and adopted the Euro (€). Countries which use the Euro are part of what was known as the Eurozone, from which the UK has an opt-out, and the country has always stuck to its currency, the Pound Sterling (£). Likewise, the Schengen Area, which abolished border controls between member states, is an agreement which allows freedom of travel between all other Schengen Area countries, without the need to show a passport. The most telling sign of the referendum was the rise of the "UK Independence Party", or "UKIP", and their somewhat controversial leader, Nigel Farage. In 2014, European Parliamentary election, UKIP came out on top, winning the most seats. It was the first time since 1906 that an election had been won by anyone other than the Conservatives and Labour. This was a European election, not a general election, which came the year later. UKIP managed to obtain 12.7% of the popular vote although this only won them 1 of the 650 seats at Westminster because of how the British voting system works.

The 2015 general election was also the reason this referendum came to be, current Prime Minister at the time, David Cameron, had promised in the lead-up to the election, that should the Conservatives win, and he remains Prime Minister, the UK will hold a referendum on EU membership. The Conservatives did win the general election, Cameron kept his promise, and the referendum date was set. On the 23rd of June, the referendum day arrived, so who was eligible to vote. British and Commonwealth citizens living in the UK were eligible to vote, while EU citizens residing in the UK, were not. Very early the next morning, it was announced that the UK had voted to leave the EU, with 52% to 48%. (*EU Referendum* 2016) The result shows a clear geographical divide, Scotland, Northern Ireland, and London voted to Remain, but the rest of England as well as Wales, voted to leave. Voting turnout was 72% with over 33% million votes. By demographics as well, there is a clear correlation between age and voting preference, with older people generally being more likely to vote to leave while young voters opted for staying. There have been calls for the capital city, London, to become its city-state within the EU as the capital also voted to remain, with only 40% voting to leave. Overnight, as the results were still coming in, the value of the pound plummeted by 10.4% as it fells low against the Euro from €1.3017 on 23 June to €1.1663 on 6 July 2016. (Barton 2019) Even though the result of the vote was to leave the EU, it is not something that goes into effect immediately and the UK is still part of the EU.(*Brexit Date: When Will the Uk Actually Leave the Eu?* 2019)

On 23rd June 2016, the voting on EU membership starting at 7:00 until 22:00. (The Shares Team 2016) The FTSE 100 was set for a flat open at the beginning of trading hour. There were also gains for supermarkets and miners. However, the stadium group's shared significantly decreased after losing their customers. On 24th June 2016, the EU referendum caused the pound sterling to crash to a 30 year low and sending shockwaves throughout Asian markets. The pound was down 9% against the US dollar and down around 7% against the Euro. The FTSE 100 failed more than 7% when the markets opening following Britain's decision to leave the EU. The FTSE100 bounded as Britain decided to leave EU and the Prime Minister Cameron resigned. The FTSE 100 plunged lower in the reaction to Britain voting to leave the European Union. The London's equity markets attempted to stability after Britain voted to leave the European Union, however the FTSE 100 continued recovery in aftermath of Brexit shock. (The Shares Team 2016)

Literature Review

On 23 June 2016, the EU referendum took place, it was a big shock to financial markets. The result from voting shows that the UK leaves EU. There is also the debate about which countries and companies would be affected the most by the Brexit. There is a lot of uncertainty in the European markets. This is unclear that UK would leave the EU or UK might leave EU with deal or no deal.

Brexit and Market Volatility

The report of (Raddant 2016) shows the breif of how the European markets reponded to the Brexit. The European stock indices, such as DAX30, FTSE100, FTSEMIB, IBEX35 and CAC40, increased slightly in the first half of the year 2016 until the beginning of June. The UK has the significant change in its currency pound(£)/euro(€) and pound(£)/dollar($) exchange rate that means investing in UK cheaper for euro and dollor currency. This indicates the stocks in the UK market adjusted by exchange rate. The (Raddant

2016) shows his analysis in stock volatility. There was an increase in the European markets after the Brexit referendum, but the volatility decresed in three weeks before the vote especially in Italy and UK. He observed that the difference stock price impacts on different sector, the stock from industrial sector has recovered fast but the stock from consumer performed weakly and the stock from financial sector reacted even worse. However, the UK stock prices from the financial sector have only recovered partly and even though the UK leaves EU but parts of UK market behave European.

The (Kurecic and Kokotovic 2018) mentioned microeconomic data shows the UK is more dependent on the EU. The UK economy might collapse if the UK take the no deal and the trade relations will return to World Trade Organization (WTO) regulation. The UK stock exchange, its currency and real estate market definitely have the negative effect on them. The UK is expected to have more negative impact than the EU, the UK financial has the most to lose with the UK leaving the European market. The analysis result of (Kurecic and Kokotovic 2018) shows that the investor quickly adapt to the political uncertainty. The consequence of Brexit has a little impact to the value of the markets.

However, the article of (Sita 2017) argued that the reaction to the UK spreaded worldwide, the capital markets and the UK politics were swayed. The investor did not rely on the FTSE 100 on 24th June the aftermath. The referendun was held on 23rd June 2016, the FTSE 100 was effected by the uncertainty not only by the perspective of UK leaving the EU, but also the disadvantange of international economic environment. The FTSE 100 dropped from 6338 to 5806 in the first 10 minutes after the markets opened on 24th June 2016. By the end of the day, the FTSE 100 increased to 6163. The withdraw of UK from EU, the economic costs and benefits of the brexit have been estimated. The extreme event like the Brexit referendum has the huge impact to the markets. The investor would imgrate the portfolio or some investors would remain the portfolio unchanged.

The (Smales 2016) mentioned the political uncertainty is very important to the investor confident when the voting date was calling and the polls of the outcome were every close. The investor made the choices during the time of political events. The political uncertainty impacts the financial market which reduces the growth of prospective return and makes the investor worried about the change of the tax regime that will effect the net returns. The empirical evidence of (Smales 2016) shows that the political uncertainty has a positive and relationship with financial market uncertainty. The implied volatility increses as the polling result increases. The higher the volatility will rise the cost of capital and the investors consider to make the pofolio choices.

The beginning of the year 2016, the FTSE100 hits the highest level after sharp rebound in Chinese exports over the state of the global economy. The reason is the demand for China-made goods. Mining companies were between the top FTSE risers that demands for commodities such as iron, copper and zinc would rise the back of stronger demand in china. There was a 25% fall in February 2016 when Chinese new year celebrations distributed the trade. Chinese exports returned to increase for the first time in nine months and it went up to 11.5% in March 2016, this compared with a year before. It was the massive increase since February 2015. (Monaghanand and Wearden 2016) It can be argued that the FTSE100 might be increased after the Brexit Referendum which is going to be part of the analysis of this paper.

Stochastic Process (Brownian Motion)

The volatility (σ_t) of the log return prices is a Stochastic Process rather than a constant Standard Brownian Motion, which models the stock index more accurately. The sequences of stock prices that are independent of each other which is appropriate for modelling random samples of a population. There is most popular model for modelling some dependence between random variables which is so-called Stochastic Process. The stock price process changes over time, the price behaviour follows the Stochastic Process called Brownian Motion was introduced by Scottish Botanist Robert Brown. He was trying to describe the motion indicated by particles immerse in the gas or liquid, the particles were bombarded by moleclues present causing displacement or movement.(Brown 1828) The Brownian Motion is the mathematical models that describe the random movements, the random walk tends to Brownian Motion when the number of steps increases. Therefore, the Brownian Motion is an important continuous-time Stochastic Process.

Model 1: it follows (Dmouj 2006), the Brownian Motion (B_t) is a Stochastic Process which has the following properties:

1. B_t has a continuous path and $B_{t=0} = 0$
2. The increment of the Brownian process in the interval of time of the different between length of $t + s$ and s is $B_{t+s} - B_s$ This increment is normally distributed with mean zero and variance equal to the time interval t, so $B_{t+s} - B_s \sim N(0, t)$

The random walk is is the concept of taking steps which gives in understanding the Brownian Motion. The random walk satisfies the following properties:

For an integer n, n $>$ 0 and it can be defined the random walk process at the time t:

$$\{W_n(t), t > 0\}$$

1. The initial value of the process is: $W_n(0) = 0$
2. The layer spacing between tow sucessive jumps is equal to $\frac{1}{n}$
3. The up and down jumps are equal and of size $\frac{1}{\sqrt{n}}$ with equal probability.

The value of the random walk at the $i - th$ step is defined recursively as follows:

$$W_n(\frac{i}{n}) = W_n(\frac{i-1}{n}) + \frac{X_i}{\sqrt{n}}$$

for all $i \geq 1$

where X_i is independent binomial variable taking value 1 and -1 with equal probability 1/2

In the probability theory, the process $W_n(t)$ has a normal distribution with mean zero and variance t

The random walk at time t after n steps starting from initial value at time zero $W_n(0)$ as follows:

$$W_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{nt} X_i$$

Multiply the numerator and denominator by $\sqrt{t}$ satisfies:

$$W_n(t) = \sqrt{t} \left(\frac{\sum_{i=1}^{nt} X_i}{\sqrt{nt}} \right)$$

Hence the random walk process $W_n(t) \to N(0, t)$ in distribution when n gets large. The random walk explains the property of stock prices in which the stock prices take more steps (n gets larger) but the interval of time becomes smaller.

Model 2: the generalised random walk is known as Brownian Motion with Drift. It is the Stochastic Process B_t with constant μ and σ, the process is given by

$$B_t = \mu t + \sigma W_t$$

where t is time

W_t is the Wiener Process (Brownian Motion),

The increment of process is normally distributed with mean μt and variance $\sigma^2 t$

However, there is difficulty about Brownian Motion as a model for stock behaviour because the process is normally distributed with mean zero from model above whereas the price of stock normally grows at some rate and with a particular variance. The stock price can be modelled as a sum of deterministic time and Brownian Motion which the modelling of a stock price can be made through the changes of time. The modelling of the stock returns makes through the SDE

Let the stock price be P_t at the time t

where μ is the expected of rate of return

dt is the changes of price during the next period of time

There are two parts of the process which are:

1. The expected return from holding the stock during a period of time dt is $\mu P_t dt$
2. The stochastic and unexpected part such as unexpected news on the stock which reflects the random changes in stock price during the interval of time dt with a constant σ and a random walk dB_t. The unexpected return is $\sigma P_t dB_t$ (Dmouj 2006)

This definition of the daily log return is extended to the Stochastic Differential Equation followed by the stock price:

$$dP_t = \mu P_t dt + \sigma P_t dB_t$$

The Stochastic Differential Equation above is the Brownian Motion with Drift followed by the stock price P_t

$$\frac{dP_t}{P_t} = \mu dt + \sigma dB_t$$

The equation above is the instantaneous rate of return P_t

It can be written as:

$$\frac{dP_t}{P_t} = d(lnP_t) \approx ln(P_t) - ln(P_{t-1}) = ln(\frac{P_t}{P_{t-1}})$$

$$ln(\frac{P_t}{P_{t-1}}) \approx \mu dt + \sigma dB_t$$

Instantaneous Volatility

The research of (Andersen 2001) that has been referred in highfrequency package in R which was the Modelling of Realised Volatility. This research provides how to model the appropriate volatility estimators which is based on squared intraday return. The asset returns are the characteristics for financial instruments. The most important feature time-varying of the distribution. The high frequency stock prices can be used for modelling of daily returns volatility. It has become apparent that Realised Volatility model is used for forecasting the daily historical stock prices (Andersen and Bollerslev 1997) The Realised Volatility is a financial term and it is a estimator of Instantaneous Volatility.

The research of (He Wang yan 1992) indicates the Semi-martingale which is truely the form of log prices X_t which can be mathematically defined:

$$X_t = M_t + A_t$$

where M is a local martingale

A is a right continuous with left limits adapted process of locally bounded variation

The Semi-martingale form above is the same as:

$$X_t = \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \sum_{l=1}^{N_t} L_l$$

where $t \in [0, 1]$ (Fan and Wang 2007)

The two terms on the right hand side are dift and diffusion, repectively, and the far right term is the jump part which can be ignored. It can be seen that W_t is the Brownian Motion and σ^2 is the instantaneous volatility. Continuous-time Brownian Semi-martingale model where volatility σ can vary i.e $\sigma = \sigma_t$, is a Stochastic Volatility models that estimates the Instantaneous Volatility.(Alvarez et al. 2011) This means that the Semi-martingale has the quadratic variation:

$$\langle X, X \rangle_t = \int_0^t \sigma_s^2 ds$$

The semi-martingale has the convergence with quadratic variations. Likewise, it is better to understand what parameter it actually estimates as it is important to have model for the intraday price evolution.

Following the estimation volatility σ_t, (Alvarez et al. 2011) shows that

Consider a Stochastic Process with stock price P_t as the observed discrete increments is given by:

$$dP_t = \mu_t dt + \sigma_t dB_t$$

where B_t is the Brownian Motion Process

σ is assumed to be a positive process

This paper is going to estimate σ_t from using the increments of P, where the process of power variations of order p, the Stochastic Process can be defined by:

$$\hat{B}(p, \Delta_n)_t = \sum_{i=1}^{[\frac{t}{\Delta_n}]} |\Delta_i^n P|^p$$

where $t \in [0, T]$

$\Delta_i^n P = P_{i\Delta_n} - P_{(i-1)\Delta_n}$; the observed discrete increments

$i\Delta_n$ for all $i = 0, 1, ..., [\frac{T}{\Delta_n}]$ when T is a positive number

The asymptotic property of $\hat{B}$ is followed (Jacod 2006) which in the asymptotic setting:

The $\Delta = \Delta_n$ depends on the number of observations and it goes to zero as $n \to \infty$

The realised $r - th$ power variation at stage n

$$\{X\}_t^{r,n} = \sum_{i=1}^{\frac{t}{\Delta_n}} |X_{i\Delta_n} - X_{(i-1)\Delta_n}|^r$$

where X is the observations

when the process $\{X\}_t^{2,n}$ converges (as $\Delta_n \to 0$) to $\langle X, X \rangle$ (Semi-Martingale) (Jacod 2006)

The asymptotic properties of $\hat{B}(p, \Delta_n)_t$ is:

Let p be a positive number, set $m_p = \mathbb{E}[|U|^p]$ where $U \sim N(0, 1)$, so $m_p = 1$

Then locally uniformly in t:

$$\Delta_n^{1-\frac{p}{2}} \hat{B}(p, \Delta_n)_t \underset{n \to +\infty}{\overset{\mathbb{P}}{\to}} m_p A(p)_t$$

where $A(p)_t = \int_0^t \sigma_s^p ds$

The tool for estimation of the volatility is the use of the power variations of order p that has the volatility process as:

$$\int_0^t \sigma_s^p ds$$

(Alvarez et al. 2011) shows that the sequence of $\hat{B}(p, \Delta_n)_t$ is an estimator of $\int_0^t \sigma_s^p ds$

For simplicity it can be looked at $p = 2$ case, it is the convergence of the quadratic variations to the continuous time Semi-martingale, so it can be assumed that $p \geq 2$

The σ_t^2 can be estimated by following statistics:

$$\sum(p, \Delta_n, h_n)_t = \frac{\Delta_n^{1-\frac{p}{2}} (\hat{B}(p, \Delta_n)_{t+h_n} - \hat{B}(p, \Delta_n)_t)}{m_p h_n}$$

subsitute $p = 2$ where $m_p = \mathbb{E}(|U|^p)$, $U \sim N(0, 1)$, $|U|^p = |U|^2 = U^2$

is Chi-Square with 1 degree of freedom, therefore $m_p = \mathbb{E}(|U|^2) = 1$

$$\sum(2, \Delta_n, h_n)_t = \frac{\Delta_n^{1-\frac{2}{2}} (\hat{B}(p, \Delta_n)_{t+h_n} - \hat{B}(p, \Delta_n)_t)}{1 h_n}$$

$$\sum(2, \Delta_n, h_n)_t = \frac{\hat{B}(p, \Delta_n)_{t+h_n} - \hat{B}(p, \Delta_n)_t}{h_n}$$

where $h \to 0$

(Alvarez et al. 2011) derives an estimator for σ_t^2:

$$\sigma_t^2 \approx \sum_{i=1}^{[\frac{t+h_n}{\Delta_n}]} |\Delta_i^n P|^p - \sum_{i=1}^{[\frac{t}{\Delta_n}]} |\Delta_i^n P|^p$$

Consequently

$$\sigma_t^2 \approx \frac{1}{n} \sum_{i=[\frac{t}{\Delta_n}]}^{[\frac{t+h_n}{\Delta_n}]} |\Delta_i^n P|^2$$

This estimator is the mean of p-variations of length h_n where h_n is assumed to be a non-increasing sequence of positive numbers such that $h_n \to 0$, e.g. $h_n = \frac{1}{n}$

The highfrequency package is an implementation of research that has been published on the analysis of high-frequency financial data in R, which is the highfrequency package that offers the essential functions for the study of high-frequency data. The function that is used for analysis is the aggregation which reduced the massive number of observations that can reach the volume of millions of observations per stock per day. (Boudt, Cornelissen, and Payseur, n.d.) However, the implement of R code will produce the estimator of Instantaneous Volatility which is the same as RV function in the highfrequency package.

Methodology

R is a programming language is the important tool for statistical computing which I am going to use mainly for the analysis of high frequency data in this paper. The functions that come with R's library which help to manipulate the data for the analysis.

Manipulation of High Frequency Data

The R base package has an options settings function which allows user to set and examine the global options, R computes and displays its result. (R Core Team 2018) The options base is needed to allow setting the Local Time columns. The options base scipen is set to 999 which applies the milliseconds to prevent scientific notation e.g. exponential notation (e). The digits secs base controls the total number of digits, in this case the milliseconds are manipulated to 3 decimal places.

```
options(scipen = 999)
options(digits.secs = 3)
```

The R utils package has a read.csv function which imports the high frequency data from chosen directory, read a file in table format and it also creates a data frame from it. (R Core Team 2018) The bid and ask prices of FTSE100 are included the date and time as the column name, Local Time which is shown the milliseconds time of bid and ask prices. For example, the FTSE100 on the date 23/06/2016 shows the first 5 rows using head function as shown below:

```
FTSE100_23_06_2016 <- read.csv("FTSE100_23_06_2016.csv")
head(FTSE100_23_06_2016, 5)
```

```
##                          Local.time     Ask     Bid
## 1 23.06.2016 08:00:00.242 GMT+0100 6306.41 6306.30
## 2 23.06.2016 08:00:00.361 GMT+0100 6306.41 6305.41
## 3 23.06.2016 08:00:00.412 GMT+0100 6306.41 6305.41
## 4 23.06.2016 08:00:00.613 GMT+0100 6312.19 6311.19
## 5 23.06.2016 08:00:00.931 GMT+0100 6312.48 6311.48
```

The FTSE100 on the date 23/06/2016 shows the last 5 rows using tail function of the observations as shown below:

```
tail(FTSE100_23_06_2016, 5)
```

```
##                              Local.time     Ask     Bid
## 70345 23.06.2016 16:29:58.704 GMT+0100 6336.51 6335.51
## 70346 23.06.2016 16:29:59.006 GMT+0100 6336.53 6335.53
## 70347 23.06.2016 16:29:59.395 GMT+0100 6336.34 6335.34
## 70348 23.06.2016 16:29:59.810 GMT+0100 6336.99 6335.99
## 70349 23.06.2016 16:30:00.199 GMT+0100 6337.24 6336.24
```

The manipulation of quotes database need to be in xts-objects format from using extensible time-series (xts) in R library which is a simple and reliable way to convert many different objects into a uniform format for use within R. (Ryan and Ulrich 2018) Using the strptime function from xts library converts the factor class in Local time column to objects of class POSIXlt representing calendar dates and times.. The stock index FTSE100 on 23/06/2016 is converted the Local time from 23.06.2016 08:00:00.242 GMT+0100 to 2016-06-23 08:00:00.242 BST (POSIXlt format) then using as.xts function to organise the sequence of ask and bid prices by Local time column.

```r
library(xts)

FTSE100_23_06_2016$Local.time <- strptime(FTSE100_23_06_2016$Local.time,
                                          "%d.%m.%Y %H:%M:%OS")
stock_index_23_06_2016 <- as.xts(cbind(FTSE100_23_06_2016$Ask,FTSE100_23_06_2016$Bid),
                                 order.by=FTSE100_23_06_2016$Local.time)
```

At this point, the quote data is required to create column names Ask and Bid using colnames function because zoo series that comes with xts library is only possible with column names.

```r
colnames(stock_index_23_06_2016) <- c("Ask","Bid")
head(stock_index_23_06_2016,3)
```

```
##                             Ask     Bid
## 2016-06-23 08:00:00.242 6306.41 6306.30
## 2016-06-23 08:00:00.361 6306.41 6305.41
## 2016-06-23 08:00:00.411 6306.41 6305.41
```

The midpoint of bid and ask prices can be defined by the half of the bid and ask price together then using names function to create midpoint column.

```r
mid_stock_index_23_06_2016 <- (stock_index_23_06_2016$Ask + stock_index_23_06_2016$Bid)/2
names(mid_stock_index_23_06_2016) <- "midpoint"
head(mid_stock_index_23_06_2016,3)
```

```
##                         midpoint
## 2016-06-23 08:00:00.242 6306.355
## 2016-06-23 08:00:00.361 6305.910
## 2016-06-23 08:00:00.411 6305.910
```

The highfrequency package is the tools for high frequency data analysis which provides functionality to manage, clean and match quotes and trades data. (Boudt, Cornelissen, and Payseur 2019) The aggregation of high frequency data which forces the stock prices to show different time step. This can be done using aggregates function from highfrequency package. For example, FTSE 100 on 23/06/2016 manipuates 08:02:00, 08:03:00 and 08:04:00 that show the changes of time every 1 minute and 08:00:01, 08:00:02 and 08:00:03 that show the changes of time every 1 second. The high frequency data has the large observations which needs to put together because it is more tracable but it can lose the observations. Using the aggregates function does not negatively effect on the results.

```r
library(highfrequency)

mid_1m_23_06_2016 <- aggregatets(mid_stock_index_23_06_2016, on="minutes", k=1)
mid_1s_23_06_2016 <- aggregatets(mid_stock_index_23_06_2016, on="seconds", k=1)
```

The returns of 1 minute and 1 second can be computed by taking log function and use diff function with taking away the first row as it produces the null (NA).

```r
return_1m_23_06_2016 <- diff(log(mid_1m_23_06_2016))[-1]
names(return_1m_23_06_2016) <- "return1m"
return_1s_23_06_2016 <- diff(log(mid_1s_23_06_2016))[-1]
names(return_1s_23_06_2016) <- "return1s"
```

FTSE 100 on 23/06/2016 indicates the evenly spacing in 1 minute and the log return prices in return 1 minute column.

```
head(return_1m_23_06_2016,3)
```

```
##                        return1m
## 2016-06-23 08:02:00 -0.0003058031
## 2016-06-23 08:03:00 -0.0011551314
## 2016-06-23 08:04:00  0.0004425484
```

FTSE 100 on 23/06/2016 indicates the evenly spacing in 1 second and the log return prices in return 1 second column.

```
head(return_1s_23_06_2016,3)
```

```
##                        return1s
## 2016-06-23 08:00:02 0.00010297348
## 2016-06-23 08:00:03 0.00005544287
## 2016-06-23 08:00:04 0.00001108820
```

After the manipulations of high frequency data, the stock index FTSE 100 consists of 1 minute and 1 second log return prices on 22/06/2016, 23/06/2016 and 24/06/2016. Moreover, the previous Fridays on 03/06/2016, 10/06/2016 and 17/06/2016 of FTSE 100 consists of 1 minute and 1 second log return prices are considered for analysis which are able to compare with the analysis of FTSE100 on 24/06/2016. At this point, the implement of volatility with time-independent (σ) and Realised Volatility (RV) are going to be computed in R as a function.

Volatility with time-independent

The implement of volatility with time-independent (σ) from Brownian Motion is equivalent to the estimation of standard deviation of log returns for each day where N is the total number of log returns using length function and μ is the mean of log returns.

```
Sigma <- function(logReturn)
{
  N <- length(logReturn)

  mu <- (1/N)*sum(logReturn)

  sqrt((1/N) * sum((logReturn - mu)^2))
}
```

For example, the function is added the log returns of 1m and 1s on 23/06/2016. The time-independent volatility is computed by Sigma function using the 1m and 1s returns.

```
Vol_return1m_23_06_2016 <- Sigma(return_1m_23_06_2016)
Vol_return1m_23_06_2016
```

```
## [1] 0.0004834859
```

```
Vol_return1s_23_06_2016 <- Sigma(return_1s_23_06_2016)
Vol_return1s_23_06_2016
```

```
## [1] 0.00006859413
```

Volatility with time-dependent

The implement of volatility with time-dependent (σ_t) is equivalent to the estimation of realised volatility of log returns for each day. This can be computed by creating the Realised Volatility (RV) function. The function needs two variables which are log return price, such as 1m return and 1s return, and the time interval when time changes every 1 hour (60 minutes), 30 minutes, 15 minutes, 10 minutes, 5 minutes and 1 minute. The function is treated with the two if-statements, then having six if-statements insides each two if-statements. The realised volatility function computes the sum of absolute of log returns in different time interval power two as the quadratic variation ($p = 2$)

```
lenReturn1m <- length(return_1m_03_06_2016) # Set the number of the length of return1m
lenReturn1s <- length(return_1s_03_06_2016) # Set the number of the length of return1s

RV <- function(return, m)
{

  # 1 minute return
  if (length(return) == lenReturn1m)
  {

        # Realised Volatility every 1 hour
        if (m == 60)
        {
          RV <- 0
          for(t in 1:8)
          {

           # For example, the realised volatility function computes the sum
           # of absolute of log returns in different time interval power two.
           RV[t] <- sum( abs( return[ (m*(t-1)+1):(m*t) ] )^ 2 )

          }
         RV
        }


        # Realised Volatility every 30 minutes
        else if (m == 30)
        {
          RV <- 0
          for(t in 1:16)
          {
           RV[t] <- sum( abs( return[ (m*(t-1)+1):(m*t) ] )^ 2 )
          }
          RV
        }
```

```
        # Realised Volatility every 15 minutes
        else if (m == 15)
        {
          RV <- 0
          for(t in 1:32)
          {
           RV[t] <- sum( abs( return[ (m*(t-1)+1):(m*t) ] )^ 2 )
          }
          RV
        }


        # Realised Volatility every 10 minutes
        else if (m == 10)
        {
          RV <- 0
          for(t in 1:48)
          {
           RV[t] <- sum( abs( return[ (m*(t-1)+1):(m*t) ] )^ 2 )
          }
         RV
        }


        # Realised Volatility every 5 minutes
        else if (m == 5)
        {
          RV <- 0
          for(t in 1:96)
          {
           RV[t] <- sum( abs( return[ (m*(t-1)+1):(m*t) ] )^ 2 )
          }
          RV
        }


        # Realised Volatility every 1 minute
        else if (m == 1)
        {
          RV <- 0
          for(t in 1:509)
          {
           RV[t] <- sum( abs( return[ (m*(t-1)+1):(m*t) ] )^ 2 )
          }
          RV
        }

  }

  # 1 second return
  else if (length(return) == lenReturn1s)
  {
```

```
# Realised Volatility every 1 hour
if (m == 60)
{
  RV <- 0
  for(t in 1:8)
  {

   # For example, the realised volatility function computes the sum
   # of absolute of log returns in different time interval power two.
   RV[t] <- sum( abs( return[ (60*m*(t-1)+1):(60*m*t) ] )^ 2 )

  }
  RV
}


# Realised Volatility every 30 minutes
else if (m == 30)
{
  RV <- 0
  for(t in 1:16)
  {
   RV[t] <- sum( abs( return[ (60*m*(t-1)+1):(60*m*t) ] )^ 2 )
  }
  RV
}


# Realised Volatility every 15 minutes
else if (m == 15)
{
  RV <- 0
  for(t in 1:32)
  {
   RV[t] <- sum( abs( return[ (60*m*(t-1)+1):(60*m*t) ] )^ 2 )
  }
  RV
}


# Realised Volatility every 10 minutes
else if (m == 10)
{
  RV <- 0
  for(t in 1:48)
  {
   RV[t] <- sum( abs( return[ (60*m*(t-1)+1):(60*m*t) ] )^ 2 )
  }
  RV
}


# Realised Volatility every 5 minutes
```

```
        else if (m == 5)
        {
          RV <- 0
          for(t in 1:96)
          {
           RV[t] <- sum( abs( return[ (60*m*(t-1)+1):(60*m*t) ] )^ 2 )
          }
          RV
        }


        # Realised Volatility every 1 minute
        else if (m == 1)
        {
          RV <- 0
          for(t in 1:509)
          {
           RV[t] <- sum( abs( return[ (60*m*(t-1)+1):(60*m*t) ] )^ 2 )
          }
          RV
        }
  }
}
```

For example, on 23rd June 2016, the volatility with time-dependent (σ_t) indicates the realised volatility of FTSE 100 in different time interval. When the power variation ($p = 2$) is the quadratic variation, the log returns of 1m and 1s are divided to every hour (60 minutes), 30 minutes, 15 minutes, 10 minutes, 5 minutes and 1 minute, respectively. It can be seen below as the volatility in the different time interval.

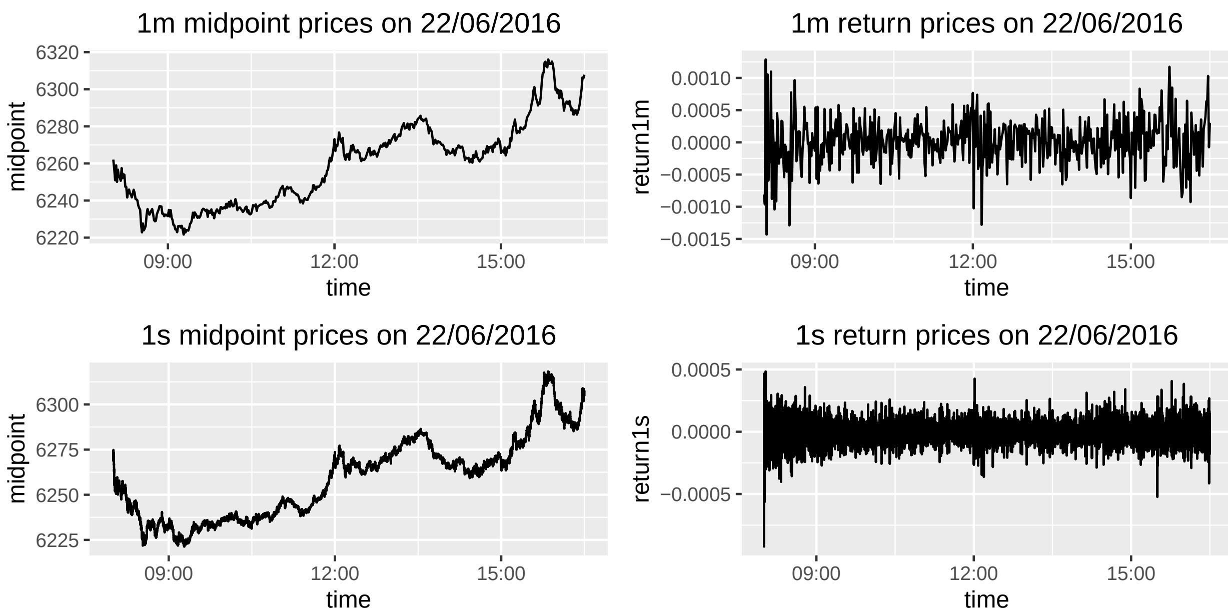
```
# Realised Volatility of 1 minute returns in every hour,
# 30 minutes, 15 minutes, 10 minutes, 5 minutes and 1 minute

RV1hr_return1m_23_06_2016 <- RV(return_1m_23_06_2016, 60)
RV30m_return1m_23_06_2016 <- RV(return_1m_23_06_2016, 30)
RV15m_return1m_23_06_2016 <- RV(return_1m_23_06_2016, 15)
RV10m_return1m_23_06_2016 <- RV(return_1m_23_06_2016, 10)
RV5m_return1m_23_06_2016  <- RV(return_1m_23_06_2016, 5)
RV1m_return1m_23_06_2016  <- RV(return_1m_23_06_2016, 1)

# Realised Volatility of 1 second returns in every hour,
# 30 minutes, 15 minutes, 10 minutes, 5 minutes and 1 minute

RV1hr_return1s_23_06_2016 <- RV(return_1s_23_06_2016, 60)
RV30m_return1s_23_06_2016 <- RV(return_1s_23_06_2016, 30)
RV15m_return1s_23_06_2016 <- RV(return_1s_23_06_2016, 15)
RV10m_return1s_23_06_2016 <- RV(return_1s_23_06_2016, 10)
RV5m_return1s_23_06_2016  <- RV(return_1s_23_06_2016, 5)
RV1m_return1s_23_06_2016  <- RV(return_1s_23_06_2016, 1)
```

Analysis

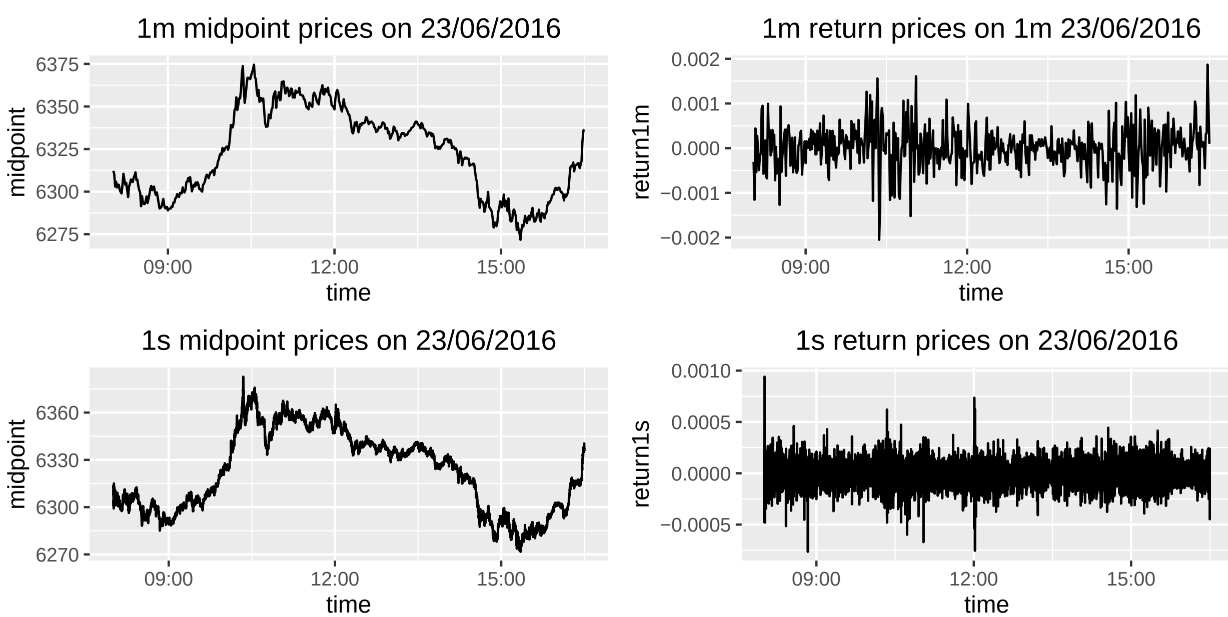
The plot of midpoint and return prices are produced by using the autoplot function in ggplot2 library. This library is based on the grammar of graphics, the graph can be built from the components, for example, a data set, a set of geoms which is visual marks that represent data points, and a coordinate system. The autoplot function is a single command that draws a particular plot for an object of a particular class (Wickham 2016) The midpoint and return prices are the data frame with time column which allows the quick plots. The gridExtra library provides a number of functions to work with grid graphics and arranges multiple grid plots on page. (Auguie 2017) The grid.arrange function from the gridExtra library combines the 4 plots within 2 columns.

```r
library(ggplot2)
library(gridExtra)

a1 <- autoplot(mid_1m_23_06_2016) +
  labs(x ="time", title = "1m midpoint prices on 23/06/2016") +
  theme(plot.title = element_text(hjust = 0.5))

a2 <- autoplot(return_1m_23_06_2016) +
  labs(x ="time", title = "1m return prices on 1m 23/06/2016") +
  theme(plot.title = element_text(hjust = 0.5))

a3 <- autoplot(mid_1s_23_06_2016) +
  labs(x ="time", title = "1s midpoint prices on 23/06/2016") +
  theme(plot.title = element_text(hjust = 0.5))

a4 <- autoplot(return_1s_23_06_2016) +
  labs(x ="time", title = "1s return prices on 23/06/2016") +
  theme(plot.title = element_text(hjust = 0.5))

grid.arrange(a1, a2, a3, a4, ncol=2)
```

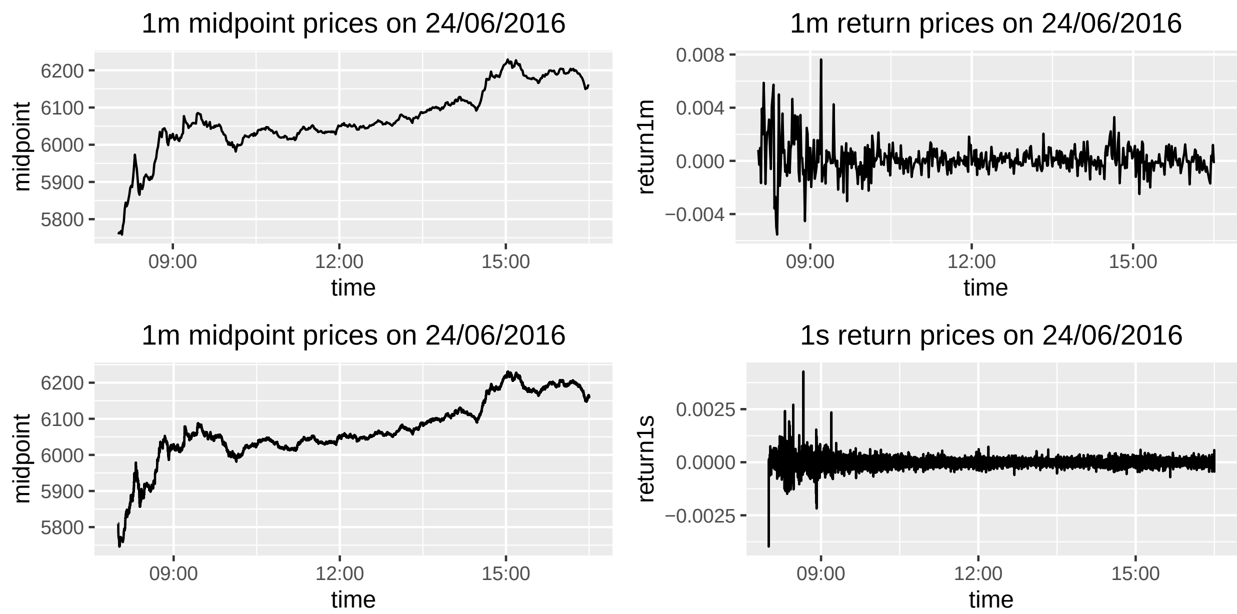
The midpoint of bid and ask prices reduces the observations of high frequency data. The 1m and 1s return prices are computed by the 1m and 1s midpoint prices, it will clearly indicate the movement of the price when there is a sudden fluctuation. The return prices from high frequency data are efficiently important data to estimate the volatility. On 22/06/2016, the plot of 1m return prices shows the sharp increased at 12:00, it was steady until 15:00, then it was risen sharply until 16:00 which is half an hour before the market closed. Similarly, the plot of 1s return prices obviously shows the sudden movement of the price at 12:00 and 16:00. The 1s returns are similar to 1m returns but 1s returns indicate the dense return prices with explicit high prices at specific time.



On 23/06/2016, the beginning of the day, the price was constant until 09:00, then it was slightly going upward to around 10:30. After the lunch time, the price was still steady, then it continuously decreased until 15:00. Today is the EU referendum day, the price rised up while the result was still counting towards the end of the day.



On 24/06/2016, the price sharply went upward after the market opened within 30 minutes, it was steady until around 14:45, then the minor price went up until 15:00, it was steady until the end of the trading hour. Besides, the plots of 1m and 1s return prices on 03/06/2016, 10/06/2016 and 17/06/2016 can be seen in an appendix. The graphs of midpoint and return prices ralates to specific events during that time when the price was changed. The movement of the return prices clearly show the high fluctuations. This answers the half of the first question that the significant events on 3 dates are important to FTSE 100 which can be assumed to be similarly to the realised volatility estimation.



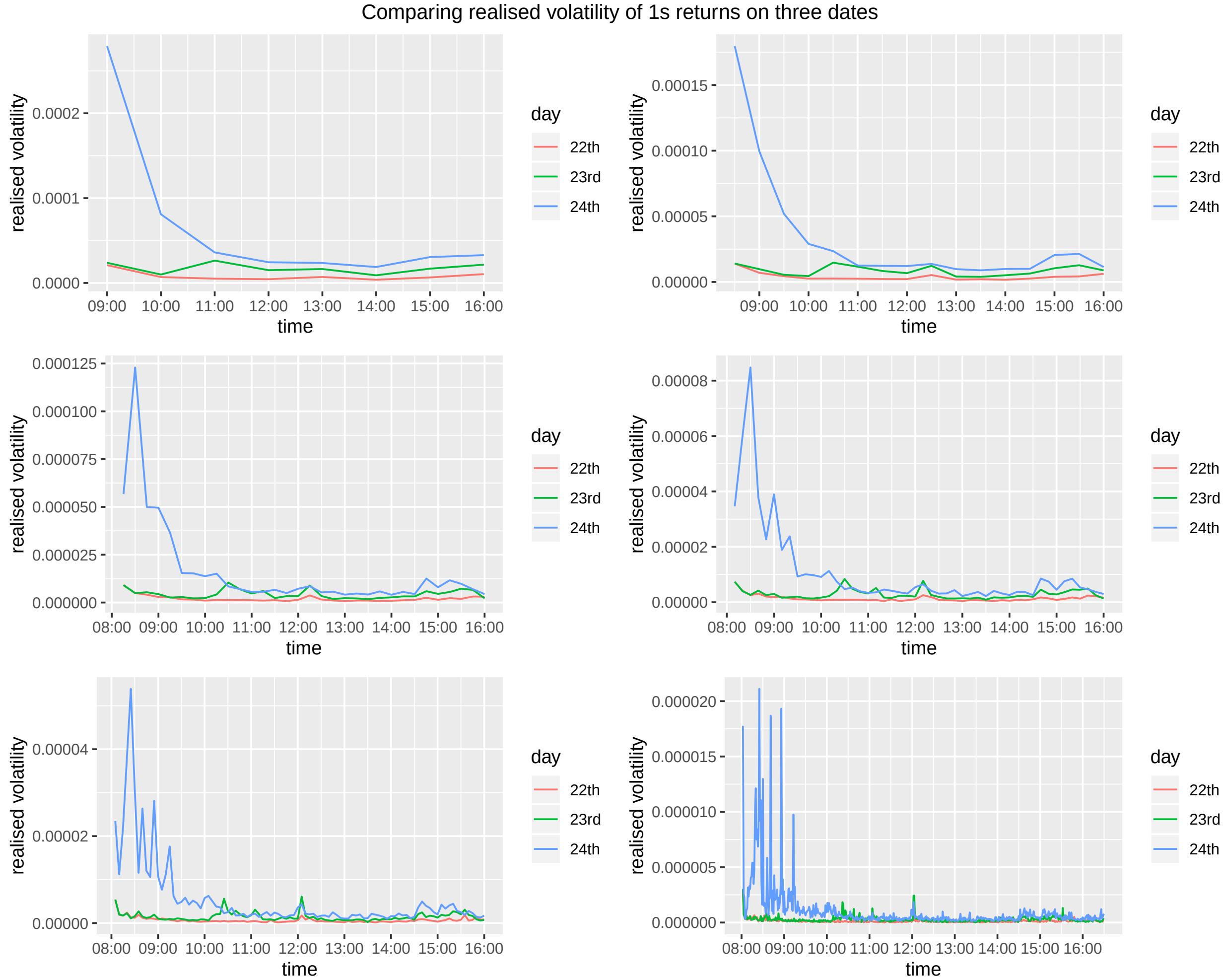
The function data.frame creates the data frames with collection of variables that share properties of matrices and lists. (R Core Team 2018) It basically means the class of the rows and columns, in this case, the realised volatility column is added into the data.frame function, because the data needs to be in a column to plot graphs in ggplot function. The graphs of realised volatility of 1m and 1s returns are created by using ggplot function. For example, the plot of 1 second return price is made for each day (22/06/2016, 23/06/2016, 24/06/2016) The plot aggregates every hour to calculate the realised volatility estimator. The ggplot function has a comparion data frame as a data, an x-axis is time in every 1 hour change, the y-axis is the realised volatility and color is the dates that represent in the graph with the legend on the side. All is in an aesthetic function of ggplot, the line of ggplot can be produced by adding geom_line() and x, y and color name labels are added in label function. The scale of date and time is needed to break in x axis every hour and the date labels create the time format for the x axis.

```
# This library is needed for data_breaks function
library(scales)

# Create the data frame
RV1hr_return1s_22_06_2016 <- data.frame(RV1hr_return1s_22_06_2016)
RV1hr_return1s_23_06_2016 <- data.frame(RV1hr_return1s_23_06_2016)
RV1hr_return1s_24_06_2016 <- data.frame(RV1hr_return1s_24_06_2016)

# Create the calender time for x axis by computing the time interval of every hour
start_1hr <- as.POSIXct(strptime("2016-06-22 09:00", format="%Y-%m-%d %H:%M"))
end_1hr   <- as.POSIXct(strptime("2016-06-22 16:00", format="%Y-%m-%d %H:%M"))
time_1hr  <- seq(start_1hr, end_1hr, by="1 hour")

# Repeat the index of the time interval to fit for three dates
comptime_1hr    <- rep(time_1hr, 3)

# Get the numeric data from the data frame
comprelvol_1hr <- c(RV1hr_return1s_22_06_2016$RV1hr_return1s_22_06_2016,
                    RV1hr_return1s_23_06_2016$RV1hr_return1s_23_06_2016,
                    RV1hr_return1s_24_06_2016$RV1hr_return1s_24_06_2016)

# Determine the three dates to have the same length with the numeric data
```

```r
compday_1hr     <- c(rep("22th",length(index(RV1hr_return1s_22_06_2016))),
                     rep("23rd",length(index(RV1hr_return1s_22_06_2016))),
                     rep("24th",length(index(RV1hr_return1s_22_06_2016))))

# Create the new data frame for ggplot function
comparison_1hr <- data.frame(comptime_1hr, comprelvol_1hr, compday_1hr)

# Plot the graph of the comparion of relised volatility on three dates
p <- ggplot(comparison_1hr, aes(x=comptime_1hr,
                                y=comprelvol_1hr,
                                color=compday_1hr)) +
     geom_line() +
     labs(x = "time", y = "realised volatility", color="day") +
     scale_x_datetime(breaks=date_breaks("1 hour"), date_labels = "%H:%M")
```

At this step, the events during EU referendum on 22/06/2016, 23/06/2016 and 24/06/2016 have been introduced for analysis. (The Shares Team 2016) The realised volatility of 1s returns on three dates is an appropriate estimator when time changes every hour, 30 minutes, 15 minutes, 10 minutes, 5 minutes and 1 minute, respectively. Besides, the realised volatility of 1m returns on three dates can be seen in an appendix.

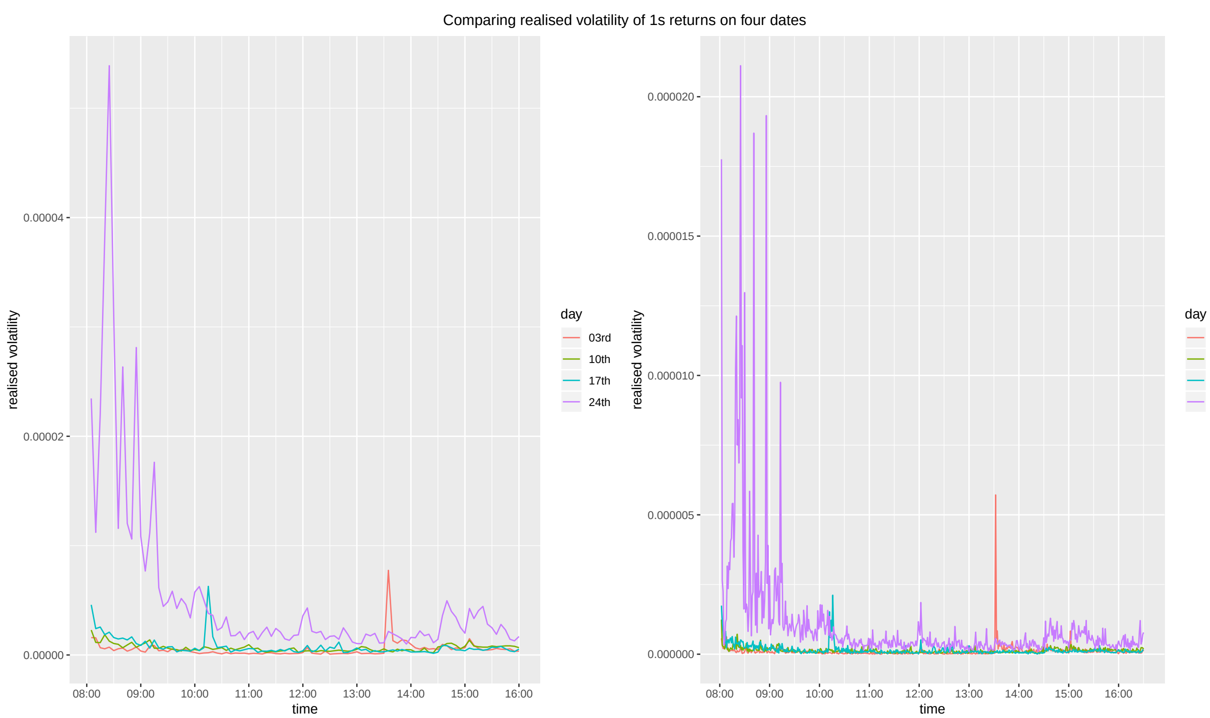
Comparing realised volatility of 1s returns on three dates



The volatility of 1 second returns on 22nd June 2016 is as minor as on 23rd, but the volatility of 1 second returns on 24th is obviously higher than 2 days before. However, the stock index FTSE 100 on 22nd opened midly higher before the EU referendum day. According to BBC poll survey (*EU Referendum Poll Tracker* 2016) which showed the Remain side by side with the Leave as the campaign of both sides attempted to sway the votes. The graphs of 1 second returns show the volatility on 22nd a lot less volatile than the volatility on 23rd, then its volatility on 22nd sharply fell down, when the Debenhams retailer simultaneously released a trading statement, while shareholder meetings were being hosted. The polls remained tight for both Remain and Leave, this can be seen as the minor volatility in the graph at 12:00 because of referendum uncertainty and the software solution provider Mirada released the commercial for the new product in Mexico, and its shares increased, this was a little increase on investors portfolios. After lunch break, the volatility of FTSE 100 was consistency because there was uncertainty and unpredictable during this time.

On 23rd June 2016, the voting on EU membership starting at 7:00. The FTSE 100 was set for a flat open at the beginning of trading hour. The graphs of realised volatility estimation of 1 second returns indicate the fall after an hour as the FTSE 100 raised as the nation started voting on whether to remain a member of European Union. There were also gains for supermarkets and miners. However, the stadium Group's shared significantly decreased after losing their customers. This can be seen in the graphs represent the minor volatility before lunch hour. The volatility fluctuated slightly, and it was consistenly increasing toward the end of the day. However, the volatility decreased when the market closed despite heading to EU referendum result tomorrow, the polling stations opened until 22:00.

On 24th June 2016, there was a fluctuation of volatility at the beginning of trading hour as the EU referendum caused the pound sterling to crash to a 30 year low and sending shockwaves throughout Asian markets. The pound was down 9% against the US dollar and down around 7% against the Euro. The FTSE 100 fall more than 7% when the markets opening following Britain's decision to leave the EU. The realised volatility of 1 second returns sharply decreases around 08:00 to 11:00 when the FTSE100 bounded as Britain decided to leave EU and the Prime Minister Cameron resigned. The FTSE 100 plunged lower in the reaction to Britain voting 51.9% to 48.1% to leave the European Union. The London's equity markets attempted to stability after Britain voted to leave the European Union which shows in the graphs of realised volatility continuously falling towards to end of the day which means the FTSE 100 continued recovery in the aftermath of Brexit shock.



However, the volatility of 1 second return prices in every 5 minutes and 1 minute on 24th June 2016 has been focusing in the timeline of key moments during the EU referendum count. (The BT Team 2016) The comparison of the volatility on 24th with the volatility on the previous Fridays: 03th, 10th and 17th are reasonable for analysis. The graphs of volatility on 24th indicates the high oscillation while the UK pound dropped against the US dollar as a possible leave which is greater than the Black Wednesday crash in 1992 (The Shares Team 2016) The volatility on 24th is apparently more volatile than on 03th, 10th and 17th. However, the FTSE100 was set to higher than usual on 03th as the oil prices pushed higher, unfortunately there was clearly mildly high volatility after the lunch break (around 13:30) as a mob of mining, oil crude majors and gold prices were disappeared by a property and financial issues as traders examined to the UK, European and US economic group data.

The FTSE 100 opened a shade lower because it was almost the weekend, according to financial websites (The Shares Team 2016), with US and Asian markets was moving back overnight on falling of oil prices, wider commodity prices and the Brexit uncertainty calling ahead of the EU referendum. The equity's London started on a positive note with miners, property and banks providing positive advice ahead of UK inflation data out on 17th as crude oil bounded ahead. This indicates the clear volatility on 17th in mid-morning. On 23th June 1016 at 22.25 (The BT Team 2016), Gibraltar is the first area to declare the result, with a predictable for Remain at 96% of the vote, the result impacts on the stock returns in the next day morning. As it can be seen that, the realised volatility is more volatile on 24th June 2016 from 8:00 to 10:00. The big result was declared to Leave at midday, in Newcastle for Remain with 50.7% followed by Sunderland voted to Leave by a significant margin in favour of Brexit. (The BT Team 2016) The graph of realised volatility on 24th shows the unpredictable volatility. There was not much volatile in the graph of realised volatility after the lunch hour as it recovered from the shock of Britain's decision to leave the European Union to end 3.15% lower at 6,138.69 which is higher than last Friday's close.

Conclusion

The volatility estimation indicates the high volatility which shows the high risk in stock price, in the opposite, the low volatility shows the low risk in stock price. The time-independent volatility only provides the volatility in each day but not in precise time intervel. On the other hand, the time-dependent volatility allows to estimate the volatility in time interval which is appropriate for the high frequency data. The time-dependent volatility is the σ_t from Brownian Motion (with drift), the σ_t is the semi-martingale which is the convergence of realised volatility. The instataneous volatility has a realised volatility estimator which produces the nearest volatility in very precise time interval. The events during Brexit referendum, it clearly shows higher volatility on 24th than 22nd and 23rd. However, the volatility has recovered in the Brexit aftermath. The volatility on 24th is more volatile than the previous Fridays, 03rd, 10th and 17th. The graphs of realised volatility makes it clear that the Brexit referendum effected to stock index FTSE 100. From the results of realised volatility estimation, the realised volatility is a good estimator of instantaneous volatility. The further research could find the instantaneous volatility with microstructure noise which is even more accurate for the high frequency data, and it could be compared to different models such as GARCH volatility model.

References

Abdi, Farshid, and Angelo Ranaldo. 2017. "A Simple Estimation of Bid-Ask Spreads from Daily Close, High, and Low Prices." *The Review of Financial Studies* 30 (12): 4437–80. https://doi.org/10.1093/rfs/hhx084.

Aldridge. 2013. *High-Frequency Trading 2e + Ws.* John Wiley & Sons. https://www.ebook.de/de/product/18566832/aldridge_high_frequency_trading_2e_ws.html.

Alvarez, Alexander, Fabien Panloup, Monique Pontier, and Nicolas Savy. 2011. "Estimation of the Instantaneous Volatility." *Statistical Inference for Stochastic Processes* 15 (1): 27–59. https://doi.org/10.1007/s11203-011-9062-2.

Andersen, T. 2001. "The Distribution of Realized Stock Return Volatility." *Journal of Financial Economics* 61 (1): 43–76. https://doi.org/10.1016/s0304-405x(01)00055-1.

Andersen, Torben G., and Tim Bollerslev. 1997. "Intraday Periodicity and Volatility Persistence in Financial Markets." *Journal of Empirical Finance* 4 (2-3): 115–58. https://doi.org/10.1016/s0927-5398(97)00004-2.

Andersen, Torben G, Tim Bollerslev, and Francis X Diebold. 2010. "Parametric and Nonparametric Measurements of Volatility: Volume 1: Tools and Techniques." In *Handbook of Financial Econometrics*, edited by Yacine Aït-Sahalia and Lars Peter Hansen, 67–137. North-Holland.

Angel, James J. 2012. *Tick Size Regulation: Costs, Benefits and Risks.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/ attachment_data/file/289037/12-1068 -eia7-tick-size-regulation-costs-benefits.pdf.

Auguie, Baptiste. 2017. *GridExtra: Miscellaneous Functions for "Grid" Graphics.* https://CRAN.R-project.org/package=gridExtra.

Barndorff-Nielsen, Ole E., and Neil Shephard. 2002. "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64 (2): 253–80. https://doi.org/10.1111/1467-9868.00336.

Barton, Charlie. 2019. *How Is Brexit Affecting the Pound?* https://www.finder.com/uk/brexit-pound.

Boudt, Kris, Jonathan Cornelissen, and Scott Payseur. 2019. *Highfrequency: Tools for Highfrequency Data Analysis.* https://CRAN.R-project.org/package=highfrequency.

———. n.d. "Highfrequency: Toolkit for the Analysis of Highfrequency Financial Data in R."

Brexit Date: When Will the Uk Actually Leave the Eu? 2019. https://www.theweek.co.uk/brexit/73810/when-will-brexit-happen.

Brown, Robert. 1828. "XXVII. A Brief Account of Microscopical Observations Made in the Months of June, July and August 1827, on the Particles Contained in the Pollen of Plants and on the General Existence of Active Molecules in Organic and Inorganic Bodies." *The Philosophical Magazine* 4 (21): 161–73. https://doi.org/10.1080/14786442808674769.

Chung, Kee H., and Hao Zhang. 2014. "A Simple Approximation of Intraday Spreads Using Daily Data." *Journal of Financial Markets* 17 (C): 94–120. https://EconPapers.repec.org/RePEc:eee:finmar:v:17:y:2014:i:c:p:94-120.

Crux, James. 2013. *How to Beat the Bid/Offer Spread.* https://www.sharesmagazine.co.uk/article/how-to-beat-the-spread-3.

Dmouj, Abdelmoula. 2006. *Stock Price Modelling: Theory and Practice.* https://beta.vu.nl/nl/Images/werkstuk-dmouj_tcm235-91341.pdf.

Dukascopy Bank SA. 2016. *Historical Data Feed.* https://www.dukascopy.com/swiss/english/marketwatch/historical.

EU Referendum. 2016. https://www.bbc.co.uk/news/politics/eu_referendum/results.

EU Referendum Poll Tracker. 2016. https://www.bbc.co.uk/news/uk-politics-eu-referendum-36271589.

European Union – Guide. 2019. https://www.politics.co.uk/reference/european-union-guide.

Fan, Jianqing, and Yazhen Wang. 2007. "Multi-Scale Jump and Volatility Analysis for High-Frequency Financial Data." *Journal of the American Statistical Association* 102 (480): 1349–62. http://www.jstor.org/stable/27639985.

Fong, Kingsley Y L, Craig W Holden, and Charles A Trzcinka. 2017. "What Are the Best Liquidity Proxies for Global Research?" *Review of Finance* 21 (4): 1355–1401. https://doi.org/10.1093/rof/rfx003.

FTSE Uk Index Series. 2019. https://www.ftserussell.com/index/spotlight/ftse-uk-index-series.

Hardy, Emily. 2019. *M&S and Easyjet Facing Drop from the Ftse100: Who Else Is at Risk, and Which Firms Could Ascend to the Stock Market Premier League*. https://www.thisismoney.co.uk/money/markets/article-7082167/M-S-Easyjet-fall-FTSE-100-risk-ascend.html.

He Wang yan. 1992. *Semimartingale Theory and Stochastic Calculus*. CRC Press. https://www.amazon.com/Semimartingale-Theory-Stochastic-Calculus-Wang/ dp/0849377153?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xm2& camp=2025&creative=165953&creativeASIN =0849377153.

Jacod, Jean. 2006. "Asymptotic Properties of Realized Power Variations and Related Functionals of Semimartingales" 118 (April).

Kurecic, Petar, and Filip Kokotovic. 2018. "Empirical Analysis of the Impact of Brexit Referendum and Post-Referendum Events on Selected Stock Exchange Indexes." *South East European Journal of Economics and Business* 13 (1): 7–16. https://doi.org/10.2478/jeb-2018-0001.

Linton, Oliver, Maureen Hara, Jean-Pierre Zigrand, Dave Cliff, and Terrence Hendershott. 2011. *The Future of Computer Trading in Financial Markets*.

Maiello, Michael. 2017. *How One Man Crashed the Stock Market*. https://www.thedailybeast.com/how-one-man-crashed-the-stock-market.

Monaghanand, Angela, and Graeme Wearden. 2016. *FTSE 100 Hits Highest Level in 2016*. https://www.theguardian.com/business/2016/apr/13/ftse-100-hits-highest-level-in-2016.

Morgan, Frank. 2009. "Stochastic Calculus and the Nobel Prize Winning Black-Scholes Equation." *Math Horizons* 17 (2): 16–18. https://doi.org/10.4169/194762109x468337.

Partridge, Dr Matthew. 2016. *FTSE100 or Ftse250: Which Is the Best Gauge of the Uk's Health?* https://moneyweek.com/444633/ftse-100-of-ftse-250-which-is-the-best-gauging-to-the-health-of-uk-plc/#.

Raddant, Matthias. 2016. "The Response of European Stock Markets to the Brexit." Kiel Policy Brief 100. Kiel: Kiel Institute for the World Economy (IfW). http://hdl.handle.net/10419/156089.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Real Time Data. 2019. https://www.londonstockexchange.com/products-and-services/ market-data/realtimedata/realtimedata.htm.

Ryan, Jeffrey A., and Joshua M. Ulrich. 2018. *Xts: EXtensible Time Series*. https://CRAN.R-project.org/package=xts.

Sita, Bernard Ben. 2017. "Volatility Patterns of the Constituents of FTSE100 in the Aftermath of the U.k. Brexit Referendum." *Finance Research Letters* 23 (November): 137–46. https://doi.org/10.1016/j.frl.2017.02.011.

Smales, Lee A. 2016. ""Brexit": A Case Study in the Relationship Between Political and Financial Market Uncertainty." *International Review of Finance* 17 (3): 451–59. https://doi.org/10.1111/irfi.12100.

The BT Team. 2016. *Timeline: EU Referendum*. http://home.bt.com/news/uk-news/ timeline-eu-referendum-june-23-2016-11364069679096.

The Shares Team. 2016. *FTSE 100 Market & Company News*. https://www.sharesmagazine.co.uk/indices/index/UKX/news/market?page=16.

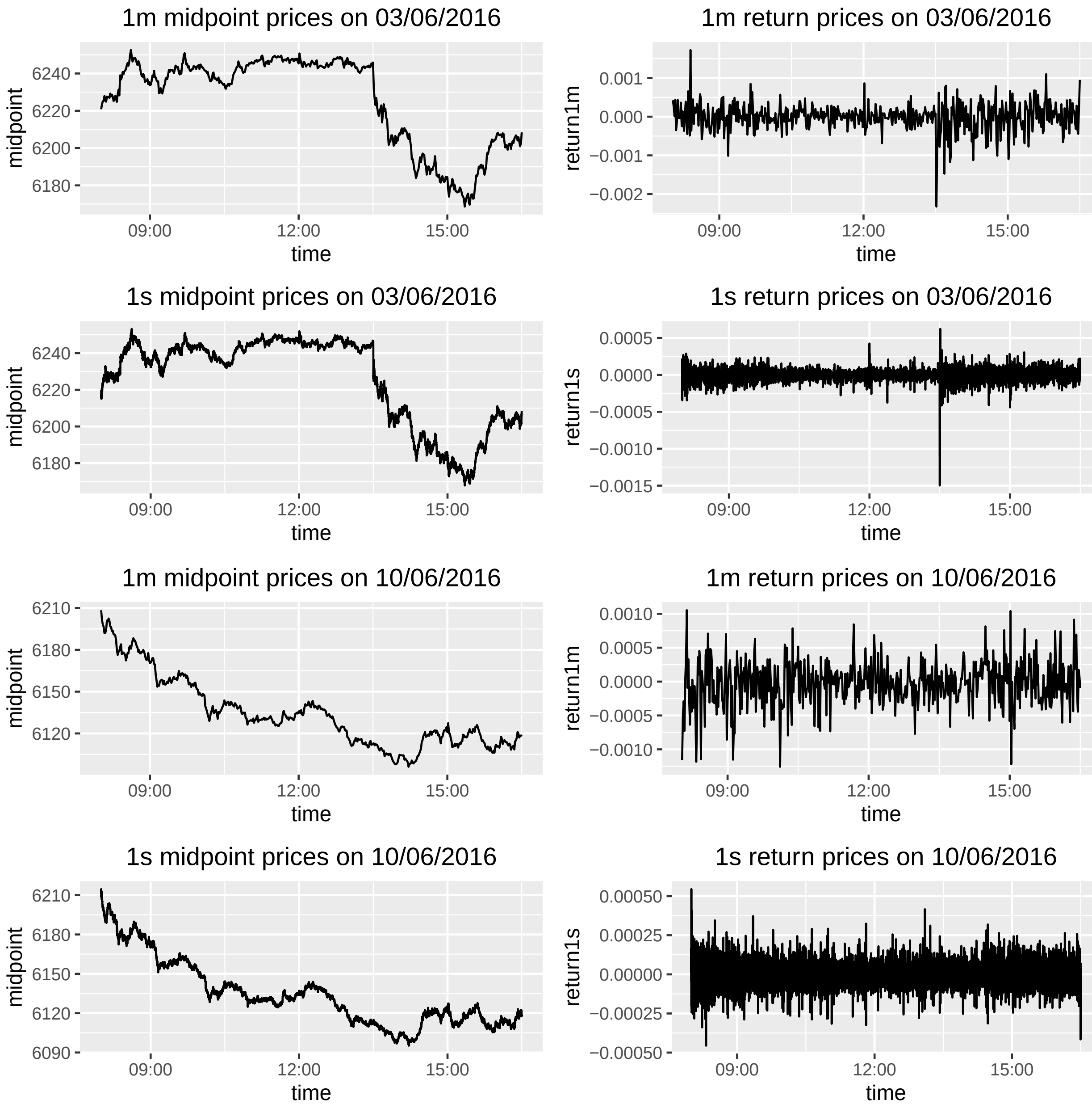
Trotman, Andrew. 2015. *What Happened During the Flash Crash?* https://www.telegraph.co.uk/finance/financial-crime/11553696/ What-happened-during-the-Flash-Crash.html.

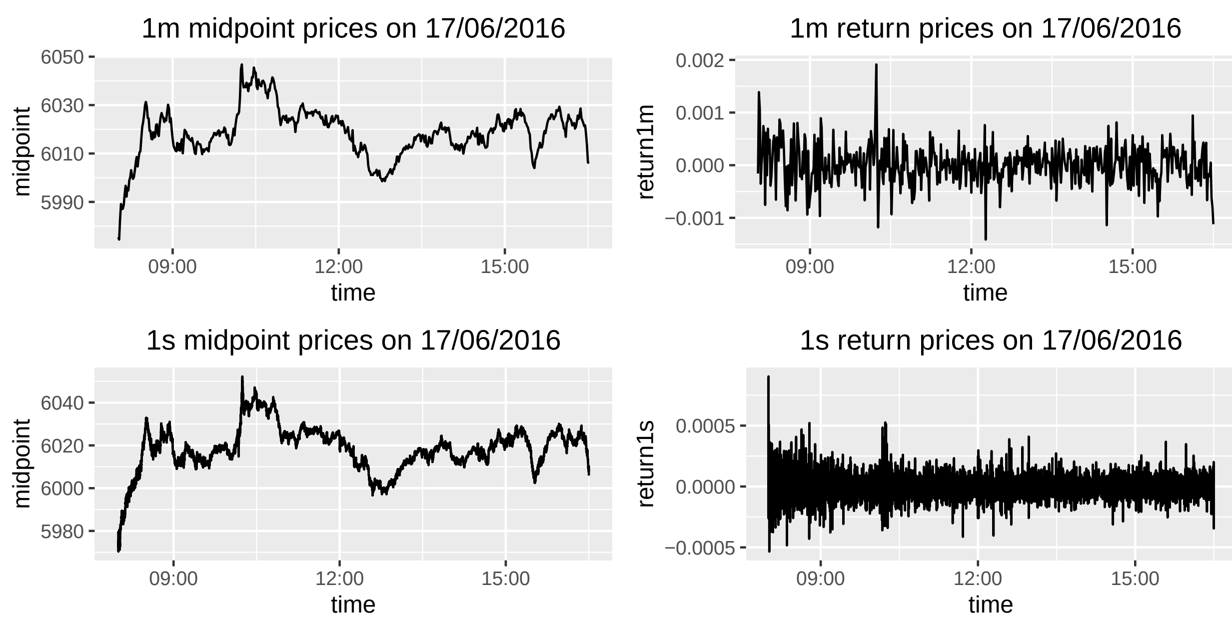
Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

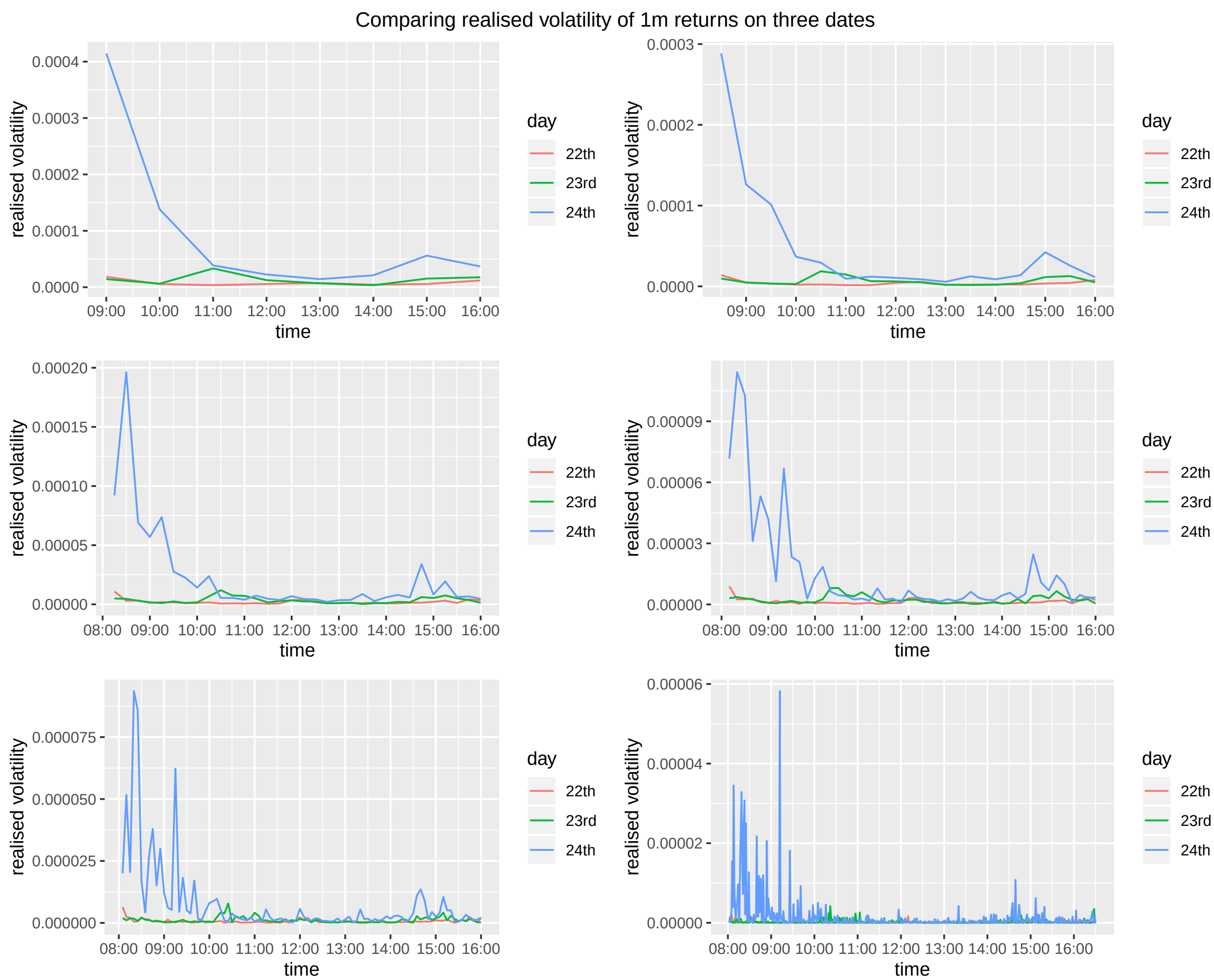
Yihui (RStudio, Inc. Boston, and M. A. USA) Xie. 2016. *Bookdown*. Taylor & Francis Ltd. https://www.ebook.de/de/product/28474987/yihui_rstudio_inc_boston_ma_usa_xie_bookdown.html.

Zu, Yang, and H. Peter Boswijk. 2014. "Estimating Spot Volatility with High-Frequency Financial Data." *Journal of Econometrics* 181 (2): 117–35. https://doi.org/10.1016/j.jeconom.2014.04.001.

Appendix


1m midpoint prices on 03/06/2016
midpoint
time
1m return prices on 03/06/2016
return1m
time
1s midpoint prices on 03/06/2016
midpoint
time
1s return prices on 03/06/2016
return1s
time
1m midpoint prices on 10/06/2016
midpoint
time
1m return prices on 10/06/2016
return1m
time
1s midpoint prices on 10/06/2016
midpoint
time
1s return prices on 10/06/2016
return1s
time


1m midpoint prices on 17/06/2016
1m return prices on 17/06/2016
1s midpoint prices on 17/06/2016
1s return prices on 17/06/2016
midpoint
return1m
return1s
time


Comparing realised volatility of 1m returns on three dates
realised volatility
time
day
22th
23rd
24th