Using Geometric Brownian Motion Model to Predict Future Stock Prices

Liew Fu Teck A0167014M

Introduction

The geometric Brownian motion model is often used to build a statistical model when modelling the financial market. It is the most widely used model of stock price behaviour. In this course project, we attempt to model the stock return of a chosen stock on the financial market using the geometric Brownian motion model, and compare it against the actual stock return to ascertain if the model is a good fit.

Problem of Interest & Methodology

We are interested in finding the return of the stock price S_t, given any S_i for $i = 0, 1, 2..., i-1$, where S_i is defined as the stock closing price at the end of the i^{th} trading period. Under the geometric Brownian motion model, the analytical solution of the geometric Brownian motion model can be expressed as

$$S_t = S_0 exp((\mu - \frac{\sigma^2}{2})t + \sigma B_t)$$

where B_t is a Brownian motion.

From the above equation, we note that given the two constants μ and σ, we are able to produce a geometric Brownian motion model through the corresponding time interval.

The geometric Brownian motion model is preferred due to the following reasons:

- The model only assumes positive values (stock price at any given time is positive)
- Expected returns of the model are independent of stock prices at each time point

However, there are also some disadvantages to using this model.

- Stock price volatility tends to change over time, but volatility is assumed constant (constant σ) under the model.
- Stock prices often show sharp jumps / dips due to unpredictable events or news, but path is continuous under the model.

The second disadvantage is particularly hard to overcome, without incorporation of market sentiment through financial news into the statistical model. Hence, we will choose a stock and a corresponding time period unlikely to encounter major financial updates (eg. release of end-of-year financial results) that may affect stock prices to a large extent.

Data Collection

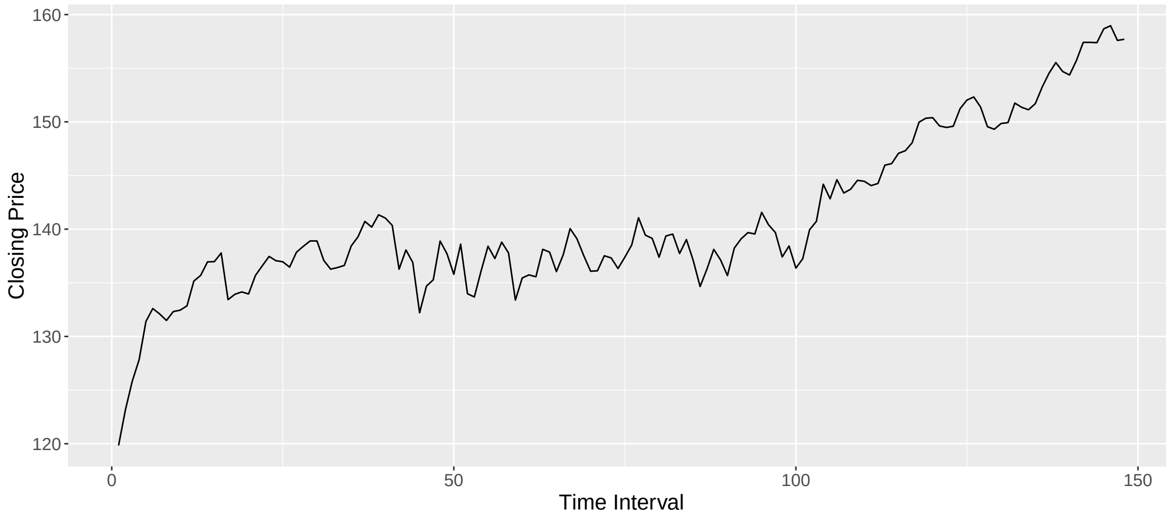
For data collection of this model, we will select a blue chip stock. Blue chip stocks are stocks that belong to large companies that have been in operation for a long period of time, and the companies are often dominant leaders in their respective industries. Stock prices of blue-chip comapnies are not as volatile as compared to smaller companies, and are less likely to be affected by financial news which might otherwise cause large deviations in stock prices of smaller companies. In this case, we will be using Microsoft for our primary analysis.

In view of the COVID-19 situation which has caused the stock prices of many companies to be unstable in recent months, we will be excluding the analysis of stock prices in 2020. Taking these considerations into account, we will be fitting Microsoft's stock across a 6 month period from 3rd June 2019 to 29th November 2019, to simulate the stock price in the month of December and come up with a stock price prediction on 31st December 2019 .

The data is collected from Yahoo Finance and downloaded into a spreadsheet, containing information such as the daily opening, closing, highest and lowest stock prices. For our project, we will be using the daily closing price for our analysis.

We can first plot the stock price for this time period to see the general trend of the stock price to ensure that there are no points of sharp and sudden movement (usually attributed to financial news), since the geometric Brownian motion model is unable to model these effects. From the graph below, we see that this assumption is fulfilled.

Figure 1: Microsoft Stock Price Trend



We also need to check 2 other assumptions. They are:

- Daily stock returns (change in stock price when compared to previous day) are normally distributed.

Figure 2: Microsoft's Daily Return

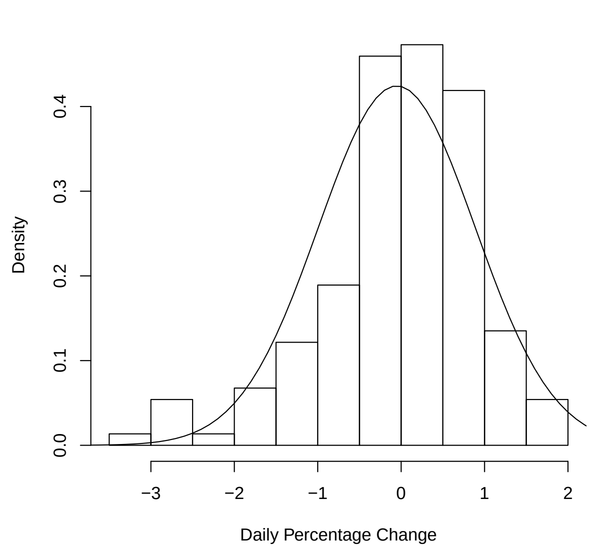
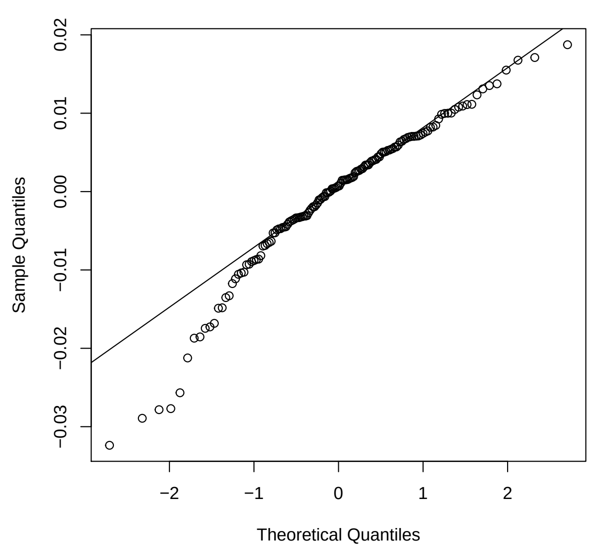


Figure 3: Normal Q−Q Plot



- Simulated price levels on 31st December 2019 are log-normally distributed.

Figure 4: Microsoft Sample GBM

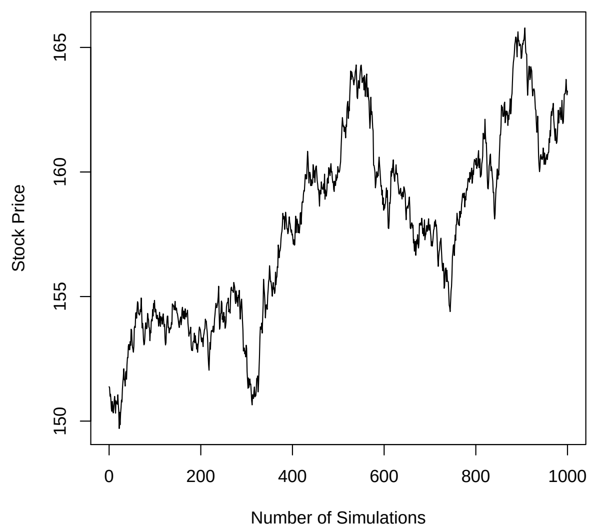
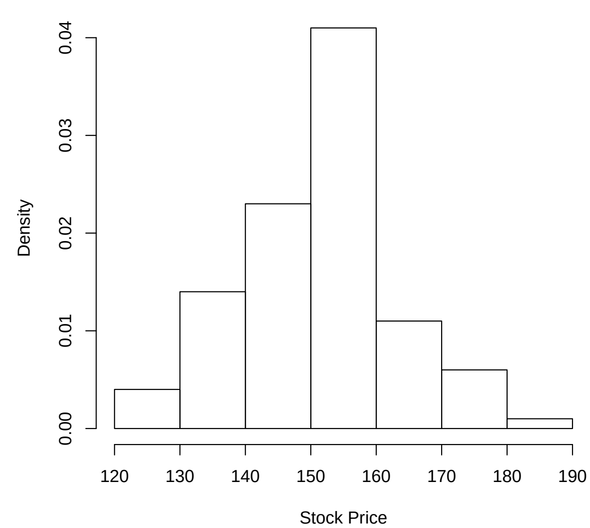


Figure 5: Histogram GBM Simulations



From the Q-Q plots (see Appendix for explanation on Q-Q plots) and histograms, we see that the 2 assumptions are generally satisfied. Hence, we can proceed to use the geometric Brownian motion model to conduct our analysis.

Problem of Interest

As a baseline, we first investigate if a geometric Brownian motion model with no drift (i.e $\mu = 0$) is a good model for the stock price of Microsoft, taking the volatility σ to be the annualized volatility of the Microsoft stock based on historical data.

Given that the simulated stock prices on 31st December follow a lognormal distribution, we take the natural log of all the simulations to calculate the corresponding 95% CI. The 95% CI is given as (148.74, 153.28). Given that the actual Microsoft stock price (157.70) is not within this 95% CI, we see that a geometric Brownian motion model with no drift is not a good model to model the movement of stock prices. This is a reasonable finding, since stock prices encounter strong drift following financial news, especially bad news [1]. Given that our model is unable to capture the effect of financial news in the real world, it is therefore essential to account for drift in such a non-stationary process.

We investigate if a geometric Brownian motion model with drift is a better model. To do so, we go back to the definition and formula to calculate estimated drift and volatility under the geometric

[1]Chan, Wesley. (2003). Stock Price Reaction to News and No-News: Drift and Reversal after Headlines. Journal of Financial Economics. 70. 223-260. 10.1016/S0304-405X(03)00146-6.

Brownian motion model, using the stock price data from 3rd June to 29th November 2019.

Volatility

We denote S_i as the stock closing price at the end of i^{th} trading period and τ as the length of time interval between two consecutive trading periods expressed in years, $\tau = t_i$ - t_{i-1}, where i > 0. If u_i is the logarithm of the daily return on the stock over the short time interval τ i.e.

$$u_i = ln(\frac{S_i}{S_{i-1}})$$

for i = 1,2,. . . ,n, then the unbiased estimtor $\bar{u}$ of the logarithm of the returns is given by:

$$\bar{u} = \frac{1}{n}\sum_{i=1}^{n} u_i$$

The estimator of the standard deviation of the u_i's is then given by:

$$v = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(u_i - \bar{u})^2}$$

where $\bar{u}$ is the unbiased estimator of the logreturn u_i. We can then estimte σ by:

$$\sigma^* = \frac{v}{\sqrt{\tau}}$$

with a standard error $\frac{\sigma^*}{\sqrt{2n}}$.

Drift

Using the original geometric Brownian motion equation $S_t = S_0 exp((\mu - \frac{\sigma^2}{2})t + \sigma B_t)$, we can obtain the drift μ as:

$$\mu = \frac{\bar{u}}{\tau} + \frac{1}{2}\sigma^2$$

With the formula of the 2 parameters, we can now proceed to find a better estimate of the estimated Microsoft stock price on 31st December 2019 under the geometric Brownian motion model.

From the formulae, we find that $\sigma^* = 0.0173$ and $\mu = 0.00386$. Hence, the updated geometric Brownian motion formula for Microsoft is given as:

$$S_t = S_0 exp((0.00386 - \frac{0.0173^2}{2})t + 0.0173B_t)$$

Running the geometric Brownian model simulation with the two updated parameters, we attain a corresponding 95% confidence interval of (168.10,173.70).

However, we see that our 95% CI using the updated parameters also does not contain the actual stock price on 31st December as well. We conclude that the geometric Brownian motion model is not a good model to predict Microsoft's stock price.

Extension of Discussion

Using the same framework as above, we explore the feasibility of using such a trading strategy to predict the stock prices of companies across different industries. The top 5 largest companies and their corresponding industries are shown in the following table.

Table 1: Top 5 Companies across Different Sectors

Technology	Consumer Services	Finance	Healthcare	Telecommunications
Apple	Coca-Cola	Berkshire Hathaway	Abbott	AT&T
Microsoft	Nike	Bank of America	Johnson & Johnson	BCE Inc.
Amazon	PepsiCo	JP Morgan	Merck & Co.	China Mobile
Alphabet	Proctor & Gamble	Visa	Novartis	T-Mobile
Facebook	Philip Morris	Mastercard	Pfizer	Verizon

These industries were chosen because companies within these industries were well-represented within the top percentile of stocks on the USA stock market, when filtered by market capitalization. We then filter for the top 5 companies in these 5 industries and aim to carry out the same geometric Brownian process on the same time period, and find the corresponding 95% confidence interval.

It is important to note that there may have been industry-specific financial news that may have affected the aforementioned 20 stocks in some way, which is a limitation of the geometric Brownian motion model as discussed earlier. However, we expect that the impact and fluctuations in stock prices for these stocks would not have been as large, in comparison to smaller companies within the same industry.

We note that the normality assumption is not well-satisfied for some companies, as seen from the Q-Q plots and histograms in the appendix. However, we relax these conditions for our extended discussion, as the main aim is to understand how well the geometric Brownian motion model can predict the stock price at a future time point for blue-chip companies in general. We then proceed to calculate the corresponding μ, σ^*, 95% CI and compare it to the actual stock price of the companies on 31st December.

Table 2: Top 5 Companies in Technology Sector

Company Name	μ	σ^*	95% CI without drift	95% CI with drift	Actual Stock Price
Apple	0.0071	0.0209	(262.08,271.71)	(325.19,338.52)	293.65
Microsoft	0.0039	0.0173	(149.01,153.55)	(167.54,173.18)	157.70
Amazon	0.0011	0.0167	(1773.79,1825.79)	(1839.81,1899.98)	1847.84
Alphabet	0.0039	0.0205	(1279.70,1326.77)	(1442.10,1500.17)	1337.02
Facebook	0.0035	0.0212	(197.67,205.05)	(220.13,229.31)	205.25

The corresponding tables for the other 4 industries and the respective companies are included in the appendix for brevity purposes.

From the 95% CIs and the stock prices, we observe a point worthy of mention.

- Out of all 25 companies, we have 13 companies whose actual stock price lies between the upper bound of the 95% CI without drift and the lower bound of the 95% CI with drift. It is particularly interesting for the technological services and finance industry, as this is the case for 4 technological companies and all 5 finance companies. We may consider carrying out further analysis to ascertain if the formula for μ and σ under the geometric Brownian motion model can be modified to better fit certain industry profiles. **However, this is out of the scope of the project and will not be discussed further.**

We then do a comparison to see how many companies' stock prices have been correctly predicted using the geometric Brownian motion model with drift, and without drift.

Table 3: Accuracy of Prediction of Stock Prices

Industry	GBM without drift	GBM with drift
Technology	0	1
Consumer Services	1	3
Finance	0	0
Healthcare	0	2
Telecommunications	1	0

From table 3, we see that with the exception of the telecommunications industry, the geometric Brownian motion model with drift performed as well / outperformed the one without drift, in terms of predicting the stock price. However, the geometric Brownian model itself as a whole does not seem to predict stock prices very well, given its accuracy across all sectors.

We now consider a different simulation approach, where we use the closing price data of last n data points and carry out a 1-stage geometric Brownian motion simulation to predict the stock price on 31st December 2019. In this case, we let n = 150 and calculate the parameters μ and σ^* using the formulae discussed above. We exclude the calculation of the 95% CI without drift given its lack of usefulness as shown from our results above.

Table 4: Top 5 Companies in Technology Sector

Company Name	μ	σ^*	95% CI with drift	Actual Stock Price
Apple	0.00690	0.0202	(292.81,295.17)	293.65
Microsoft	0.00330	0.0169	(157.79,158.84)	157.70
Amazon	0.00036	0.0177	(1843.64,1856.60)	1847.84
Alphabet	0.00260	0.0207	(1336.28,1347.32)	1337.02
Facebook	0.00180	0.0228	(204.21,206.06)	205.25

The corresponding tables for the other 4 industries and the respective companies are included in the appendix for brevity purposes.

We then check if there is any improvement in prediction as compared to the original geometric Brownian motion model with drift simulated over 6 months.

Table 5: Accuracy of Prediction of Stock Prices

Industry	6 month simulation	1 day simulation
Technology	1	4
Consumer Services	3	3
Finance	0	5
Healthcare	2	3
Telecommunications	0	5

From the above table, we can conclude that this 1-day simulation prediction has a much higher accuracy level.

Practical Usage

Due to the lack of fluctuation that can take place in a single day as compared to 6 months, it might seem trivial to do a one-period simulation of geometric Brownian motion, considering that it is more likely that the stock price of the next day (in this case 31st December 2019) will fall within the 95% CI, as compared to a 6-month simulation period. However, a practical usage would be to utilise such an approach to create short-term investment strategies by investing in companies whose 95% CI of stock price is higher than the current market stock price, and update the CI daily based on the last n days to decide if any actions should be taken to maximise returns to investment (eg. to sell, to buy more).

Limitations & Improvements

The analysis was done on the 25 companies under the assumption that their stock returns all follow a normal distribution. However, this does not hold true in real life. In reality, stock returns tend to be skewed towards a positive return in the long run[2]. Furthermore, as discussed at the start of the report, the geometric Brownian model does not take the impact of financial news into account. This may impact the feasibility and accuracy of the geometric Brownian model even further, considering that the assumptions on which the model are based on are not fulfilled in the first place. Hence, the geometric Brownian motion model is an oversimplication in these regards.

Based on our analyis, some areas to consider for follow-up include the use of varying time frames to predict the stock price.

1) In this report, we have only considered two separate time frames of 6-months and a 1-day simulation period. There might be other periods which give rise to a 95% CI of higher accuracy.

2) As discussed earlier, we can also consider the modification of parameters for different industries (using a bias) if we see an observable trend for the stock price and their respective 95% CIs across all / most companies in the particular industry.

3) We can also consider improving the accuracy of the geometric Brownian motion model by incorporating market sentiment, or modelling the parameters under the geometric Brownian model as a function of other macroeconomic factors (eg. market interest rate, consumer confidence etc.)

Conclusion

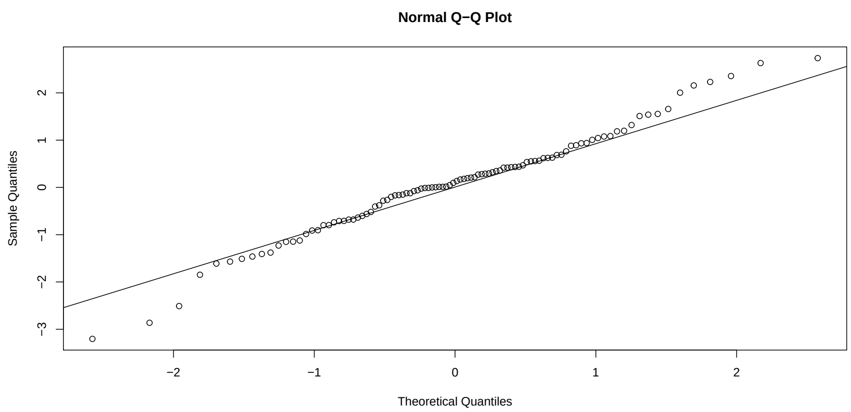
In this project, we modelled the stock price at a future point in time by considering the movement of stock prices as a stochastic process. From the comparison of the results and real-life data, we see that the geometric Brownian motion is not a good model to predict stock prices over a 6-month period, even though the assumptions required to use the model are largely met for the datasets (companies) for this project. However, its accuracy is improved when modelled over a shorter time period (1 day).

[2]Dierking, D. (2019). Benefits of Holding Stocks for the Long Term. Investopedia. Retrieved 24 April 2020, from https://www.investopedia.com/articles/investing/052216/4-benefits-holding-stocks-long-term.asp

Appendix

Explanation of Q-Q plots

Q-Q plots are scatterplots created by plotting two sets of quantiles against one another. If both sets of quantiles came from the same distribution, we expect to see the points falling on a straight line. In the report, we are plotting a Normal Q-Q plot, to ascertain if the daily stock returns for the companies follow a normal distribution. For example, the following Q-Q plot is derived from a N(0,1) distribution using the **rnorm** function.



References

- Ross, Sheldon M. Introduction to Probability Models. Academic Press, 2019.
- Chan, Wesley. (2003). Stock Price Reaction to News and No-News: Drift and Reversal after Headlines. Journal of Financial Economics. 70. 223-260. 10.1016/S0304-405X(03)00146-6.
- Dierking, D. (2019). Benefits of Holding Stocks for the Long Term. Investopedia. Retrieved 24 April 2020, from https://www.investopedia.com/articles/investing/052216/4-benefits-holding-stocks-long-term.asp
- Yang, Zhiyun. Geometric Brownian Motion Model in Financial Market
- Dmouj, Abdelmoula. Stock Price Modelling: Theory and Practice

Calculations of μ, σ^* and 95% CI based on geometric Brownian motion model with / without drift

Table 6: Top 5 Companies in Consumer Services Sector

Company Name	μ	σ^*	95% CI without drift	95% CI with drift	Actual Stock Price
Coca-Cola	0.00120	0.0146	(52.71,54.06)	(54.72,56.28)	55.35
Nike	0.00310	0.0199	(91.78,94.99)	(100.96,104.91)	101.31
PepsiCo	0.00091	0.0130	(134.30,137.34)	(138.35,141.86)	136.67
Proctor & Gamble	0.00270	0.0161	(120.31,123.70)	(130.83,134.95)	124.90
Philip Morris	0.00110	0.0242	(81.03,84.48)	(84.08,88.09)	85.09

Table 7: Top 5 Companies in Financial Services Sector

Company Name	μ	σ^*	95% CI without drift	95% CI with drift	Actual Stock Price
Berkshire Hathaway	0.0017	0.0123	(217.97,222.62)	(230.11,235.60)	226.50
Bank of America	0.0037	0.0207	(32.68,33.87)	(36.67,38.16)	35.22
JP Morgan	0.0035	0.0173	(129.71,133.64)	(144.63,149.52)	139.40
Visa	0.0026	0.0176	(181.59,187.18)	(196.63,203.38)	187.90
Mastercard	0.0032	0.0211	(286.50,297.16)	(316.30,329.40)	298.59

Table 8: Top 5 Companies in Healthcare Sector

Company Name	μ	σ^*	95% CI without drift	95% CI with drift	Actual Stock Price
Abbott	0.0021	0.0175	(84.11,86.68)	(89.76,92.83)	86.86
Johnson & Johnson	0.0009	0.0171	(135.40,139.41)	(139.32,143.97)	145.87
Merck & Co.	0.0015	0.0177	(85.79,88.44)	(89.91,93.03)	90.95
Novartis	0.0009	0.0140	(91.17,93.40)	(93.96,96.52)	94.69
Pfizer	-0.0012	0.0182	(37.89,39.09)	(36.68,37.98)	39.18

Table 9: Top 5 Companies in Telecommunications Services Sector

Company Name	μ	σ^*	95% CI without drift	95% CI with drift	Actual Stock Price
AT&T	0.0031	0.0165	(36.83,37.89)	(40.49,41.79)	39.08
BCE Inc.	0.0010	0.0116	(47.60,48.55)	(49.11,50.22)	46.35
China Mobile	-0.0026	0.0155	(37.21,38.22)	(34.49,35.54)	42.27
T-Mobile	0.0011	0.0185	(77.22,79.74)	(79.99,82.89)	78.42
Verizon	0.0011	0.0125	(59.59,60.88)	(61.80,63.30)	61.40

Calculations of μ, σ^* and 95% CI based on geometric Brownian motion model with drift based on increment of 1 period

Table 10: Top 5 Companies in Consumer Services Sector

Company Name	μ	σ^*	95% CI with drift	Actual Stock Price
Coca-Cola	0.0018	0.0139	(55.28,55.59)	55.35
Nike	0.0035	0.0192	(100.92,101.69)	101.31
PepsiCo	0.0010	0.0124	(136.74,137.42)	136.67
Proctor & Gamble	0.0025	0.0155	(124.55,125.32)	124.90
Philip Morris	0.0010	0.0231	(85.54,86.33)	85.09

Table 11: Top 5 Companies in Financial Services Sector

Company Name	μ	σ^*	95% CI with drift	Actual Stock Price
Berkshire Hathaway	0.0018	0.0118	(225.85,226.92)	226.50
Bank of America	0.0034	0.0201	(35.18,35.46)	35.22
JP Morgan	0.0035	0.0168	(138.83,139.76)	139.40
Visa	0.0024	0.0203	(187.84,189.09)	187.90
Mastercard	0.0021	0.0167	(297.69,300.10)	298.59

Table 12: Top 5 Companies in Healthcare Sector

Company Name	μ	σ^*	95% CI with drift	Actual Stock Price
Abbott	0.00200	0.0166	(86.80,87.37)	86.86
Johnson & Johnson	0.00150	0.0161	(145.24,146.17)	145.87
Merck & Co.	0.00200	0.0167	(91.03,91.63)	90.95
Novartis	0.00140	0.0134	(94.53,95.03)	94.69
Pfizer	-0.00079	0.0172	(38.80,39.06)	39.18

Table 13: Top 5 Companies in Telecommunications Services Sector

Company Name	μ	σ^*	95% CI with drift	Actual Stock Price
AT&T	0.00280	0.0162	(39.08,39.33)	39.08
BCE Inc.	0.00035	0.0118	(46.19,46.41)	46.35
China Mobile	-0.00061	0.0155	(42.09,42.35)	42.27
T-Mobile	0.00051	0.0182	(77.95,78.51)	78.42
Verizon	0.00078	0.0140	(61.16,61.50)	61.40