Simulating Stock Prices Using Geometric Brownian Motion: Evidence from Australian Companies

Krishna Reddy[1] **and Vaughan Clinton**[2]

Abstract

This study uses the geometric Brownian motion (GBM) method to simulate stock price paths, and tests whether the simulated stock prices align with actual stock returns. The sample for this study was based on the large listed Australian companies listed on the S&P/ASX 50 Index. Daily stock price data was obtained from the Thomson One database over the period 1 January 2013 to 31 December 2014. The findings are slightly encouraging as results show that over all time horizons the chances of a stock price simulated using GBM moving in the same direction as real stock prices was a little greater than 50 percent. However, the results improved slightly when portfolios were formed.

JEL Classification:

Keywords: Geometric Brownian motion, Simulation, CAPM, Mean absolute percentage error (MAPE), Industry, Australia

[1] University of Waikato, New Zealand
[2] University of Waikato, New Zealand

1 Introduction

There is an abundance of literature surrounding the pricing of securities in corporate finance; however there is still a lot of debate as to which method is the most reliable. Financial managers and investors are interested in simulating the price of stock, options, and derivatives in order to make important investment and financing decisions. Simulating the price of a stock means generating price paths that a stock may follow in the future. We talk about simulating stock prices because future stock prices are uncertain (called stochastic), but we believe that they follow, at least approximately, a set of rules that we can derive from historical data and our knowledge of stock prices (Sengupta, 2004). A simulation will be realistic only if the underlying model is realistic. The model must reflect our understanding of stock prices and conform to historical data (Sengupta, 2004).

In this study we focus on the geometric Brownian motion (hereafter GBM) method of simulating price paths, and test the model using a sample of large Australian stocks employing a range of techniques to assess how well the simulated stock prices align with actual stock returns. Marathe and Ryan (2005) assert that because of the significant financial impacts resulting from decisions made using the GBM assumption, it must be subject to test. It is of utmost importance to verify that a time series follows the GBM process, before relying on the result of such an assumption (Marathe & Ryan, 2005).

The remainder of this paper is set out as follows: section (2) describes the literature related to GBM and other forecasting methods. Section (3) details the data and research method used to simulate stock prices and test the GBM model. Section (4) presents the findings of the paper, and section (5) concludes and presents areas for future research.

2 Literature Review

a. The validity of geometric Brownian motion

Brownian motion is often used to explain the movement of time series variables, and in corporate finance the movement of asset prices. Brownian motion dates back to the nineteenth century when it was discovered by biologist Robert Brown examining pollen particles floating in water under the microscope (Ermogenous, 2005). Brown observed that the pollen particles exhibited a jittery motion, and concluded that the particles were 'alive'. This hypothesis was later confirmed by Albert Einstein in 1905 who observed that under the right conditions, the molecules of water moved at random. A common assumption for stock markets is that they follow Brownian motion, where asset prices are constantly changing often by random amounts (Ermogenous, 2005). This concept has led to the development of a number of models based on radically different theories.

Two common approaches to predicting stock prices are those based on the theory of technical analysis and those based on the theory of fundamental analysis (Fama, 1995). Technical theorists assume that history repeats itself, that is, past patterns of price behaviour tend to recur in the future. The fundamental analysis approach assumes that at any point in time an individual security has an intrinsic value that depends on the earning potential of the security, meaning some stocks are overpriced or under-priced (Fama, 1995). Many believe in an entirely different approach; the theory that stock market prices exhibit random walk. The random walk theory is the idea that stocks take a random and unpredictable path, making it near impossible to outperform the market without assuming additional risk. This theory casts serious doubts on the other methods of describing and predicting stock price behaviour. The GBM model incorporates this idea of random walks in stock prices through its uncertain component, along with the idea that stocks maintain price trends over time as the certain component. Brewer, Feng and Kwan (2012) describe the uncertain component to the GBM model as the product of the stock's volatility and a stochastic process called Weiner process, which incorporates random volatility and a time interval.

Sengupta (2004) claims that for GBM model to be effective one must imply that:

- The company is a going concern, and its stock prices are continuous in time and value.
- Stocks follow a Markov process, meaning only the current stock price is relevant for predicting future prices.
- The proportional return of a stock is log-normally distributed.
- The continuously compounded return for a stock is normally distributed.

As discussed in section 3, each of these assumptions has an effect on the GBM model and its inputs.

GBM has two components; a certain component and an uncertain component. The certain component represents the return that the stock will earn over a short period of time, also referred to as the drift of the stock. The uncertain component is a stochastic process including the stocks volatility and an element of random volatility (Sengupta, 2004). Brewer, Fend and Kwan (2012) show that only the volatility parameter is present in the Black-Scholes (BS) model, but the drift parameter is not, as the BS model is derived based on the idea of arbitrage-free pricing. For Brownian motion simulations both the drift and volatility parameter are required, and a higher drift value tends to result in higher simulated prices over the period being analysed (Brewer, Feng and Kwan, 2012).

Although the GBM process is well-supported, there is a growing amount of literature that focus on testing the validity of the model and accuracy of forecasts using Brownian motion. For example Abidin and Jaffar (2014) use GBM to forecast future closing prices of small sized companies in Bursa Malaysia. The study focuses on small sized companies because the asset prices are lower and more affordable for individual investors. The study looks into the accuracy of forecasts made using the model over different horizons, and also at the time horizon needed for data inputs into the model, that is, past stock prices. According to Abidin and Jaffar (2014), GBM can be used to forecast a maximum of two week closing prices. It was also found that one week's data was enough to forecast the share prices using GBM.

Marathe and Ryan (2005) discuss the process for checking whether a given time series follows the GBM process. They also look at methods to remove seasonal variation from a time series, which they claim is important because the GBM process does not include cyclical or seasonal effects. They found that of the four industries they studied, the time series for usage of established services met the criteria for a GBM process; while the data form growth of emergent services did not.

b. Modifications to the GBM model

One of the caveats to GBM is that it does not account for periods of constant values (Gajda & Wylomanska, 2012). Gajda and Wylomanska (2012) observed periods where prices stay on the same level, particularly true for assets with low liquidity. Gajda and Wylomanska propose an alternative approach based on subordinated tempered stable geometric Brownian motion, combining the conventional GBM model with inverse tempered stable subordinator. They tested the effectiveness of their modified method using Monte Carlo simulations, and reported that the calculated values closely reflect the theoretical ones. They also successfully applied the new model to data describing German Inter-Bank Rates.

Ladde and Wu (2009) also develop modified linear models of GBM by employing classical model building techniques. They have utilised stock price data to examine the accuracy of the existing GBM model under standard statistical tests. They then demonstrated the development of the modified GBM model under different data partitioning, with and without jumps. They compared the constructed models and the GBM models via a Monte Carlo technique. The findings suggest that data partitioning improves the results and the models with jumps are much better than the ones without jumps.

c. Other methods of forecasting stock prices

Overtime a number of models have been developed with the objective of forecasting stock prices and pricing options. Some of these models are summarised by Granger (1992), with a particular emphasis on non-linear models. Higgins (2011) demonstrate a simple model to forecast stock prices using analyst earnings forecasts based on the residual income model (RIM). Higgins shows how to implement the RIM and explains how to adjust for auto-correlation to improve forecast accuracy. The RIM uses a combination of fundamental accounting data and mechanical analysis of trends in time series data to derive a valuation of the firm.

Hadavandi, Shavandi and Ghanbari (2010) present an integrated approach based on genetic fuzzy systems (GFS) and artificial neural networks (ANN) for constructing a stock price forecasting expert system. To evaluate the capability of their proposed approach they apply the model to stock price data gathered from IT and Airlines sectors, and compare the outcomes with previous stock price forecasting methods using mean absolute percentage error (MAPE). The results they obtained show that the proposed approach outperforms all previous methods, and so it can be considered a suitable tool for stock price forecasting. However, a drawback to the MAPE approach is that it can only be used to predict the next day closing prices, making it less relevant to managers making strategic decisions.

Hsu, Liu, Yeh and Hung (2009) used a combination of the grey model (GM), Fourier series and Markov state transition matrices to produce a new integration prediction method called the Markov-Fourier grey model (MFGM). The hybrid model was used to predict the turning time of Taiwan weighted stock index (TAIEX) in order to improve forecasting accuracy. According to Hse et al. (2009), MFGM method can predict accurately but it is only suitable for long-term operation.

d. Contribution of this paper

Despite an abundance of literature on the application and modifications to the GBM model, there is an apparent lack of research on the accuracy of forecasts made with the model, and thus the validity of the GBM assumption. Therefore, this study contributes to the literature in the following ways:

- The sample chosen is large Australian stocks, representing a market with very little research on GBM.
- The paper focusses on the validity of the GBM assumption over a range of holding periods, namely, one week, two weeks, one month, six months and twelve months. This tests the GBM model on its validity in the long-term as well as the short-term, while most literature only tests the accuracy of forecasts over the short-term.

- Sample is subdivided into portfolios to analyse the effect of stock volatility, expected returns and industry on the accuracy of the model. Other research tends to focus on individual stocks and to the best of our knowledge no one to date has tested the validity of GBM on portfolios.

- Finally, a variety of methods are used to compare actual and simulated prices, specifically, the correlation coefficient, percentage of correct directional predictions, and mean absolute percentage error techniques. Although these methods have been applied by other researchers, most of them focus on only one method.

Based on the above, following hypotheses are proposed:

H1: There is no significant difference between the actual stock prices and the simulated prices using GBM over the sample period.

H2: There is no significant relationship between stock volatility and the difference between actual and simulated stock prices.

H3: There is no significant relationship between expected stock returns and the difference between actual and simulated stock prices.

H4: There is no significant relationship between industry and the difference between actual and simulated stock prices.

3 Data and Research Method

a. Data

Data was collated for the large listed Australian companies listed on the S&P/ASX 50 Index. Daily stock price data was obtained from the Thomson One database over the period 1 January 2014 to 31 December 2014. The start date for the simulations was 1 January 2014 which was chosen to avoid any effects of seasonality in stock prices. Benjamin and Bin (2011) reported that there is no evidence of a 'January effect' in the returns of the top 50 stocks in Australia, nor were there any significant effects in other months. Based on the findings of Benjamin and Bin (2011), it is concluded that there is no significant seasonality in the returns for the top 50 Australian stocks.

Information on the industry sector for each stock in the sample was taken from the ASX website and the current market capitalisation for each stock was taken from the Thomson One database. As the constituent stocks of the S&P/ASX 50 Index have changed since the beginning of the simulation period, it had to be adjusted for any changes taking place after 1 January 2014. Stocks that were added to the index after 1 January 2014 were removed and any companies with insufficient price data over the sample period were also removed. For example, South 32 (S32-AU) was removed as it was only recently formed as the result of a demerger with BHP Billiton Limited (BHP-AU), and consequently prices only exist for a short period. To replace this stock we added CIMIC Group Limited (CIM-AU) formerly known as Leighton Holdings Limited (LEI-AU), as this was the next largest company at the time. Similarly, Scentre Group (SCG-AU) was also removed as it was formed in June 2014 as a demerger from Westfield Corporation (WFD-AU). To replace this stock, Alumina Limited (AWC-AU) was added as this was the next largest company, after LEI.

To compute expected stock returns data was required for the risk-free rate, stock beta and expected market return. For the risk-free rate we used the current yield on Australian two-year government bonds, obtained from Bloomberg. Stock betas were taken from the Thomson One database, and to calculate the expected market return, prices of the S&P/ASX 50 Index from 1 January 2011 to 31 December 2013 were extracted from the McGraw Hill Financial website.

b. Application of GBM

Equation 1 below shows the formula for the proportional return of a stock:

$$\frac{\Delta S}{S} = \mu\Delta t + \sigma\varepsilon\sqrt{\Delta t} \tag{1}$$

First component shows the expected rate of return μ that a stock will earn over a short period of time Δt, this component is often referred to as the certain component. The second component follows a random process where σ is the expected volatility of the stock and $\varepsilon\sqrt{\Delta t}$ represents the random volatility which magnifies as the period of time increases.

GBM assumes that stock prices are log-normally distributed with a mean of the certain component and a standard deviation of the uncertain component, shown in equation 2 below:

$$\ln\frac{S_T}{S_0} \sim \phi\left[\left(\mu - \frac{\sigma^2}{2}\right)T, \sigma\varepsilon\sqrt{T}\right] \tag{2}$$

Where

S_0 is the stock price now and S_T is the price at time T. Notice that μ has been replaced with $(\mu-\sigma^2/2)$ to superimpose an uncertainty component to generate a fluctuating stock price. As T represents any time interval it is possible to simulate the price of a stock at time $t+\Delta t$ given its price at t, where Δt is a short time interval, using the lognormal distribution as shown in the following equation:

$$\ln\left(\frac{S_{t+\Delta t}}{S_t}\right) = \left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma\varepsilon\sqrt{\Delta t} \tag{3}$$

Finally, rearranging equation 3 results in the final equation I used in our stock price simulations, shown below:

$$S_{t+\Delta t} = S_t \exp\left[\left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma\varepsilon\sqrt{\Delta t}\right] \tag{4}$$

To recap, S_t is the stock price at time t, Δt is the time interval for prediction, μ is the expected annual rate of return, σ is the expected annual volatility, and ε is a randomly drawn number from a normal distribution with a mean of zero and a standard deviation of one, representing random volatility.

The time interval for prediction we use is one day, as we are interested in predicting daily prices over the period 1 January 2014 to 31 December 2014. We used the capital asset pricing model (CAPM) to calculate the expected annual return for each stock using the following formula:

$$\mu = r_f + \beta_m(r_m - r_f) \tag{5}$$

Where

r_f is the risk-free rate of return, β_m is the beta of stock against the market, and r_m is the expected return of the market portfolio. CAPM is used because it is simple to calculate and the input variables are easily accessible. Sengupta (2004) state that there are numerous measures of the expected rate of return that can be used for the model, such as the expected rate based on historical returns and estimates from analysts. Abidin and Jaffar (2014) used the mean percentage change in stock prices over one month because they were looking at forecasting stock prices from an investor's perspective, rather than from a corporate finance standpoint.

To estimate the annual volatility of each stock we have used the similar method to that proposed by Sengupta (2004) of taking the daily returns for each stock over a one year period. For simulation, we used the stock price data from 1 January 2013 to 31 December 2013. We calculated the daily standard deviation from the daily returns and used the formula given below (Equation 6) to calculate the annualised volatility of each stock where s represents the daily standard deviation and τ is the intervals measured in years. As daily price data is used, we assume 250 trading days per year, the time interval used was 1/250 (Sengupta, 2004).

$$\hat{\sigma} = \frac{s}{\sqrt{\tau}} \tag{6}$$

c. *How the hypotheses were tested*

This paper involves four tests of the GBM assumption, comparing simulated and actual stock prices for individual stocks, portfolios based on volatility, portfolios based on expected returns and portfolios of each industry. Each test corresponds to the hypotheses set in section 2 and were carried out as follows:

- Test one – Individual simulations were conducted for each constituent stock of the S&P/ASX 50 Index over the period 1 January 2014 to 31 December 2014. This test involves an analysis of the accuracy of forecasts over a one week, two week, one month, six month, and twelve month period.

- Test two – Stocks were ranked in terms of their annual volatility and grouped into quintile portfolios, portfolio 1 containing low volatility stocks and portfolio 5 containing high volatility stocks. The portfolios were value-weighted based on an investment of $1000 in each stock. Again the validity of the model was tested by comparing simulated and actual stock prices over a one week, two week, one month, six month and twelve month period.

- Test three – Stocks were ordered according to their expected returns and grouped into quintile portfolios. As in test 2, the portfolios were value-weighted to prevent the risk of higher-priced stocks skewing the data. The simulated and actual prices were compared based on holding periods of one week, two weeks, one month, six months, and twelve months.

- Test four - All 50 stocks were broken down into their respective industry sector to form portfolios containing stocks from the same industry. For this test only industries with three or more stocks were considered to prevent any bias from individual companies. Portfolios were based on an investment of $1000 in each stock, resulting in different sized portfolios; however this does not affect any of the metrics used to compare the simulated and actual results. The same time horizons were considered for this test as the other three.

d. *Comparing actual and simulated prices*

We have used three methods to compare simulated and actual prices. First, the correlation coefficient (r) is considered to measure the linear correlation between simulated and actual prices. The correlation coefficient produces a value between negative one and positive one, where negative one is perfect negative correlation, zero is no correlation, and positive one is perfect positive correlation. The following formula was used where x and y are different variables (simulated vs actual prices) and n is the number of observations:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}} \quad (7)$$

Second, we calculate the mean absolute percentage error (MAPE) between the actual and simulated values. This is the same method used by Abidin and Jaffar (2014), who also offer a range scale to assess the MAPE value against. MAPE values of less than 10 percent are considered highly accurate forecasts, 11 percent to 20 percent represents good forecasts, 21 percent to 50 percent being reasonable forecasts, and greater than 51 percent deemed to be an inaccurate forecast. We follow the same scale when analysing our results in the next section. MAPE was calculated using the following formula where A_t is the actual price, F_t is the forecasted price and n is the number of forecasts:

$$MAPE = \frac{\sum_{t=1}^{n} \left| \frac{A_t - F_t}{F_t} \right|}{n} \quad (8)$$

Third, we look at the percentage of correct directional predictions in the simulated prices benchmarked against the actual prices.

4 Results

4.1 Descriptive statistics

Table 1 reports the descriptive statistics for each stock including the industry sector which it belongs to, current market capitalisation at the time of writing, initial stock price, expected annual return, and expected annual volatility. A total of 50 stocks were analysed from twelve different industry sectors as defined on the ASX website. The largest sector by number of stocks was financials, with ten stocks, representing 20 percent of the total sample. In contrast, the sectors with the least stocks were consumer discretionary, information technology and telecommunications services, each with one stock.

The largest individual company by market capitalization is the Commonwealth Bank of Australia (CBA-AU) with over \$132 billion in total capital. This is significantly higher than the mean and median market capitalization at \$24 billion and \$10.7 billion, respectively. This suggests that the distribution of stocks in terms of market capitalization is skewed towards the smaller stocks, with a couple of very large stocks that increase the value of the mean. Data on the initial stock price at 1 January 2014 tells a similar story, with a range of \$76.74 between the highest-priced stock, the Commonwealth Bank of Australia (CBA-AU), and the lowest-priced stock, Qantas Airways Limited (QAN-AU). The distribution of initial stock prices among firms has a mean of \$18.75 and a standard deviation of \$19.00, also indicating a skewed distribution.

The average expected annual return over the 50 stocks was 4.89 percent, which appears to be quite low. This is possibly the result of a generally low beta across all stocks, having an impact on the calculation of expected returns using CAPM. The average beta was 0.835, which casts some doubt on the calculation of beta given that the sample is comprised of the constituent stocks of the market index. On the other hand, as already discussed, the stocks in the sample appear to be concentrated in a small number of industries and may not truly reflect the market portfolio.

The results obtained for the volatility of the stocks were interesting as they appear to be very low overall. Sengupta (2004) suggest that stocks normally have volatilities in the range of 20 percent to 60 percent; however our sample produced a median volatility of 22.4 percent, and the highest volatility observed was only 48.3 percent. A possible explanation of this result is that the sample consists of only large, well-established companies that are generally at a later stage in their growth cycle than smaller, younger firms. Further, there may be country effects to the level of volatility, in particular, when Sengupta (2004) published his work there may have been more economic turbulence in the US than there is in today's Australian market. Another interesting finding is that the firm with the lowest volatility, Telstra Corporation Limited (TLS-AU) of 14.6 percent, was only marginally less than the median level of 22.4 percent, suggesting a clustering of firms between these values.

Table 1: Descriptive statistics

Company symbol	Industry Sector	Market capitalisation (millions)	Initial stock price (S_0)	Expected annual return (μ)	Expected annual volatility (σ)
AGL-AU	Utilities	10,478	14.39	0.0320	0.1790
AIO-AU	Industrials	6,369	5.77	0.0447	0.2440
AMC-AU	Materials	12,812	10.59	0.0269	0.2402
AMP-AU	Financials	18,131	4.43	0.0679	0.2480
ANZ-AU	Financials	89,506	32.22	0.0672	0.1897
APA-AU	Utilities	9,516	5.81	0.0457	0.2010
ASX-AU	Financials	7,802	36.85	0.0481	0.1772
AWC-AU	Metals & Mining	3,467	1.13	0.0583	0.3691
AZJ-AU	Industrials	11,183	4.92	0.0389	0.1806
BHP-AU	Metals & Mining	109,732	38.20	0.0645	0.2038
BXB-AU	Industrials	13,023	9.19	0.0368	0.2270
CBA-AU	Financials	132,719	77.84	0.0518	0.1663
CCL-AU	Consumer Staples	7,185	12.10	0.0317	0.2387
CIM-AU	Industrials	7,738	16.21	0.0836	0.3829
CPU-AU	Information Technology	5,255	11.33	0.0395	0.2301
CSL-AU	Health Care	32,162	68.82	0.0313	0.1984
CTX-AU	Energy	8,659	19.99	0.0611	0.3193
CWN-AU	Consumer Discretionary	9,251	16.93	0.0498	0.2439
DXS-AU	Real Estate	7,230	6.12	0.0467	0.2332
FDC-AU	Real Estate	4,254	2.32	0.0412	0.2145
GMG-AU	Real Estate	10,904	4.78	0.0498	0.2230
GPT-AU	Real Estate	7,768	3.43	0.0344	0.1967
IAG-AU	Financials	12,934	5.84	0.0498	0.2011
ILU-AU	Metals & Mining	3,509	8.76	0.0741	0.3703
IPL-AU	Materials	6,552	2.69	0.0590	0.2561
LLC-AU	Real Estate	9,055	11.18	0.0662	0.2723
MGR-AU	Real Estate	7,081	1.70	0.0484	0.2487
MQG-AU	Financials	26,007	55.08	0.0563	0.2934
NAB-AU	Financials	90,815	34.34	0.0700	0.1842
NCM-AU	Metals & Mining	10,386	8.45	0.0491	0.4834
ORG-AU	Energy	14,037	14.06	0.0389	0.2321
ORI-AU	Materials	8,048	24.07	0.0484	0.3029
OSH-AU	Energy	8,683	8.17	0.0501	0.2095
QAN-AU	Industrials	7,094	1.10	0.0693	0.3678
QBE-AU	Financials	14,291	11.70	0.0505	0.3886
RHC-AU	Health Care	12,256	42.58	0.0324	0.2244
RIO-AU	Metals & Mining	78,339	68.71	0.0679	0.2410
SGP-AU	Real Estate	9,848	3.65	0.0477	0.2082
SHL-AU	Health Care	7,839	16.45	0.0221	0.2277
STO-AU	Energy	8,090	14.67	0.0553	0.2245
SUN-AU	Financials	17,453	13.10	0.0426	0.1960
SYD-AU	Industrials	11,702	3.79	0.0361	0.1882
TCL-AU	Industrials	18,819	6.72	0.0255	0.1568
TLS-AU	Telecommunication Services	72,865	5.26	0.0324	0.1456
WBC-AU	Financials	99,950	32.34	0.0710	0.1992
WES-AU	Consumer Staples	46,692	44.83	0.0460	0.1752
WFD-AU	Real Estate	14,529	5.72	0.0358	0.1728
WOR-AU	Energy	2,530	16.85	0.0529	0.4403
WOW-AU	Consumer Staples	34,743	33.89	0.0382	0.1624
WPL-AU	Energy	22,902	38.56	0.0594	0.1974

4.2 Test one – individual stocks

This test attempts to prove whether or not there is a significant difference between actual stock prices and those simulated using the GBM model by looking at the stocks on an individual basis. As a starting point, we look at the simulation results of four stocks that are evenly distributed in the sample in terms of their market capitalization. The four stocks we focus on are the Commonwealth Bank of Australia (CBA-AU), Origin Energy Limited (ORG-AU), Santos Limited (STO-AU) and Worley Parsons Limited (WOR-AU). Table 2 shows the results for the first month of simulation, January 2014. From Table 2 we can see that the results after the first week of predictions (five working days) do not differ significantly from the actual prices of ORG-AU or STO-AU, but there are some differences for CBA-AU and in particular WOR-AU, which closed on 8 January with a $1.19 difference between the forecasted price.

Comparing the simulated and actual prices at the end of the month yields different results. Clearly the simulated price for ORG-AU was a close estimate to the actual closing price on 31 January, with a difference of only 6 cents. Also the simulated price for CBA-AU was not far off considering the value of the stock is almost five times greater than the other three. WOR-AU showed a 91 cent discrepancy, while STO-AU was much less accurate, with a MAPE of just below 13 percent.

Figure 1 provides a visual representation of the results over a much longer forecast horizon. What are instantly noticeable are the similarities in the behaviour between simulated and actual stock prices in terms of volatility. CBA-AU illustrates this point well, as the actual stock price can be observed fluctuating between $72 and $86 over the simulation period. The simulated prices can be seen exerting a similar level of variability between $72 and $84. Looking at the STO-AU chart shows a significant deviation from the actual prices, particularly after the month of April where the difference between simulated and actual prices reaches almost 40%. Despite this, the closing price on 31 December was very close to the simulated value. The chart for ORG-AU shows a very different trend, almost mirroring the actual prices, resulting in a moderately strong negative correlation over a twelve month period. WOR-AU displays a very similar long-term trend to the actual prices and only deviates from this trend in the last two months.

Based on these four stocks, it is difficult to infer whether or not there is a significant difference between simulated and actual stock prices. In order to make any conclusive observations, the whole sample of stocks must be considered.

Table 3 provides a summary of the results for test one using each of the methods to compare simulated and actual stock prices. Panel A looks at the correlation coefficient (r) between the two variables over the varying time horizons. The mean correlation over the short-term is slightly negative and grows positive as the simulation period is increased. This means that for one week and two week predictions simulated prices move in the opposite direction to that of the real prices. For periods of 1 month or longer they correct and begin to follow the prices more accurately. This could be a result of the certain component of the GBM model compensating for negative random fluctuations, as stock prices tend to increase over time. Looking at the median correlation leads to similar results as this is also negative for short periods and becomes positive after one month. The absolute mean and standard deviation are lowest for one month predictions suggesting that there is less variability over this prediction horizon.

Table 2: Simulated and actual prices for four selected stocks

	CBA-AU		ORG-AU		STO-AU		WOR-AU	
Date	Simulated price	Actual price	Simulated price	Actual price	Simulated price	Actual price	Simulated price	Actual price
2/01/2014	77.84	77.84	14.06	14.06	14.67	14.67	16.85	16.85
3/01/2014	78.00	77.58	13.80	14.01	14.66	14.56	16.54	16.68
6/01/2014	79.77	77.60	13.84	13.97	14.89	14.33	17.32	16.64
7/01/2014	78.81	77.72	13.65	13.89	15.17	14.35	17.70	16.32
8/01/2014	78.16	77.88	13.84	14.02	15.09	14.29	17.89	16.35
9/01/2014	77.21	77.98	13.99	14.07	15.02	14.50	17.77	16.58
10/01/2014	76.79	77.60	13.96	13.99	15.04	14.37	16.71	16.45
13/01/2014	76.19	77.00	14.00	13.81	14.80	14.26	16.68	16.84
14/01/2014	75.68	75.80	13.93	13.54	15.12	14.04	16.65	16.87
15/01/2014	73.85	76.19	13.69	13.50	15.42	14.17	16.61	17.11
16/01/2014	76.03	75.80	13.57	13.83	15.27	14.37	16.15	17.45
17/01/2014	76.53	75.47	13.56	13.83	15.12	14.51	16.32	17.28
20/01/2014	77.57	75.32	13.64	14.09	15.06	14.45	16.23	17.17
21/01/2014	78.68	76.10	13.57	14.20	14.98	14.41	15.58	16.98
22/01/2014	77.43	75.91	13.95	14.25	14.77	14.26	15.71	17.11
23/01/2014	77.11	75.13	13.84	13.90	14.49	14.14	15.45	16.92
24/01/2014	77.07	74.74	14.02	13.88	14.98	13.85	15.49	16.73
28/01/2014	76.89	74.13	14.27	13.74	14.52	13.69	15.72	16.61
29/01/2014	75.90	74.92	13.80	13.98	14.59	13.50	15.91	16.64
30/01/2014	75.43	74.20	13.63	13.94	14.71	13.32	15.38	16.49
31/01/2014	75.07	74.23	13.92	13.98	15.07	13.34	15.54	16.45

Figure 1: Line charts showing simulated and actual prices for four selected stocks

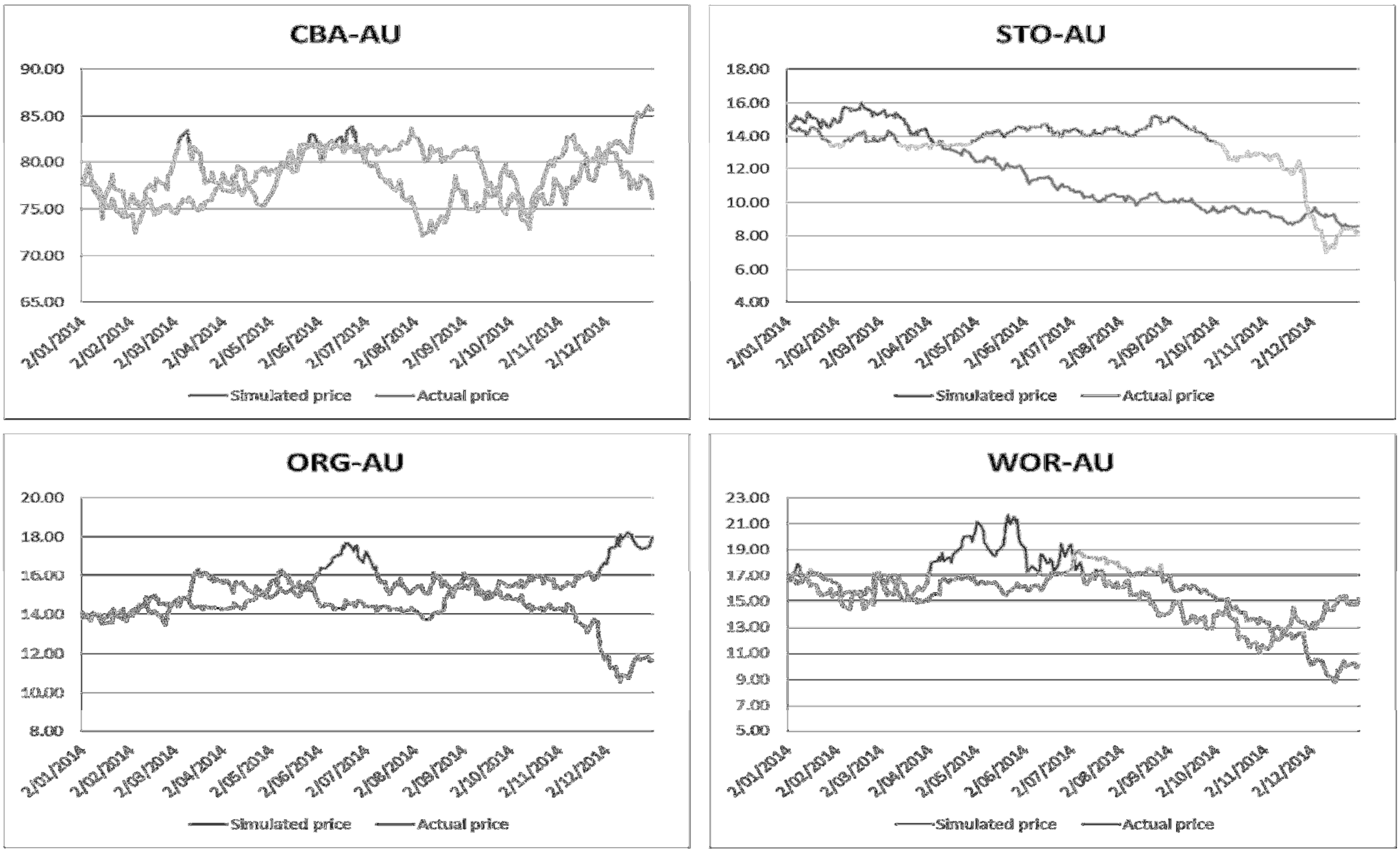

Table 3 Panel B presents the results for the MAPE of the forecasts against actual prices over the differing simulation periods. There is an obvious pattern emerging where longer periods result in a higher MAPE, evident from both the mean and median MAPE across all stocks. Based on the results it appears that the GBM model is most effective at forecasting stock prices over a one week, two week and one month period. However, based on the MAPE scale of judgement of forecasting accuracy proposed by Abidin and Jaffar (2014), the average MAPEs calculated in this paper would tend to suggest that model is highly accurate in forecasting actual stock prices. All of the median MAPEs fall below the 10 percent threshold and qualify as highly accurate forecasts, and based on mean MAPE, periods up to six months are highly accurate and a 12 month period is still a good forecast.

The percentage of correct directional predictions over the entire sample of stocks are shown in Table 3 Panel C. This measure counts the number of daily stock price predictions moving in the same direction as the actual prices over the period 1 January 2014 to 31 December 2014. The overall result is disconcerting, as the mean percentage of correct stock price movements for the simulation was only marignally greater than 50 percent. This would imply that on any given day, flipping a coin would be almost as effective a price prediction model. Although it seems that for shorter time periods the model may still be accurate, predicting a median of 60 percent correct directional movements over one week.

Table 3: Summary of results for test one

	Panel A: Correlation				
	1 week	2 weeks	1 month	6 months	12 months
Mean	-0.0515	-0.0470	0.0587	0.0956	0.1842
Mean - absolute	0.5257	0.4220	0.4200	0.4346	0.4365
Median	-0.0732	-0.0776	0.1965	0.1637	0.2129
Median - absolute	0.6306	0.3827	0.3813	0.4425	0.4685
Range	1.9468	1.8029	1.6985	1.7869	1.7718
Min	-0.9582	-0.9580	-0.8215	-0.9184	-0.8361
Max	0.9886	0.8450	0.8770	0.8685	0.9357
Standard deviation	0.6120	0.5007	0.4762	0.4923	0.4791

	Panel B: Mean absolute percentage error (MAPE)				
	1 week	2 weeks	1 month	6 months	12 months
Mean	2.46%	3.35%	4.36%	9.39%	12.03%
Median	1.95%	2.52%	2.90%	8.81%	9.49%
Range	6.33%	12.27%	20.66%	22.74%	29.94%
Min	0.52%	0.90%	1.43%	2.61%	3.65%
Max	6.85%	13.16%	22.09%	25.35%	33.59%
Standard deviation	1.54%	2.66%	4.03%	4.97%	6.90%

	Panel C: Direction prediction accuracy				
	1 week	2 weeks	1 month	6 months	12 months
Mean	52.40%	49.80%	50.60%	52.38%	52.10%
Median	60.00%	50.00%	50.00%	51.64%	51.98%
Range	100.00%	60.00%	65.00%	26.23%	18.65%
Min	0.00%	20.00%	20.00%	40.98%	42.06%
Max	100.00%	80.00%	85.00%	67.21%	60.71%
Standard deviation	24.54%	17.67%	12.60%	5.30%	3.96%

From an analysis of the accuracy of individual stock price forecasts, this test has provided mixed results. On one hand the correlation coefficient implies a negative relationship between simulated stock prices and actual prices in the short run. In contrast, the MAPE and directional prediction accuracy method provide support that over short periods the GBM model is accurate. The MAPE findings are of greater value, however, because the other methods look at the daily change in stock prices, which does not adequately test the GBM assumption.

4.3 Test two – volatility portfolios

The second test looks at portfolios formed on the basis of the stock's annual volatility to investigate the hypothesis that there is no significant relationship between stock volatility and the difference between actual and simulated stock prices. Figure 2 below shows the simulated prices for each portfolio based on a $1000 investment in each stock. Portfolio 1 represents stocks with a low volatility and portfolio 5 contains stocks with the highest annual volatility. The volatility of a stock affects the uncertain component of the GBM model, and when volatility increases it effectively magnifies any variation caused by random volatility. The chart below shows that as volatility increases the forecasted stock prices tend to stray further from their mean value. For instance, portfolio 1 shows a relatively small amount of fluctuation, staying within the range of $9,700 to $10,400. Portfolio 5 on the other hand displays much greater fluctuation, ranging between $9,300 and $11,200. I now turn my attention to the impact of this increased volatility on simulated prices compared to actual prices.

Figure 2: Line chart showing share price simulations for portfolios of varying volatility

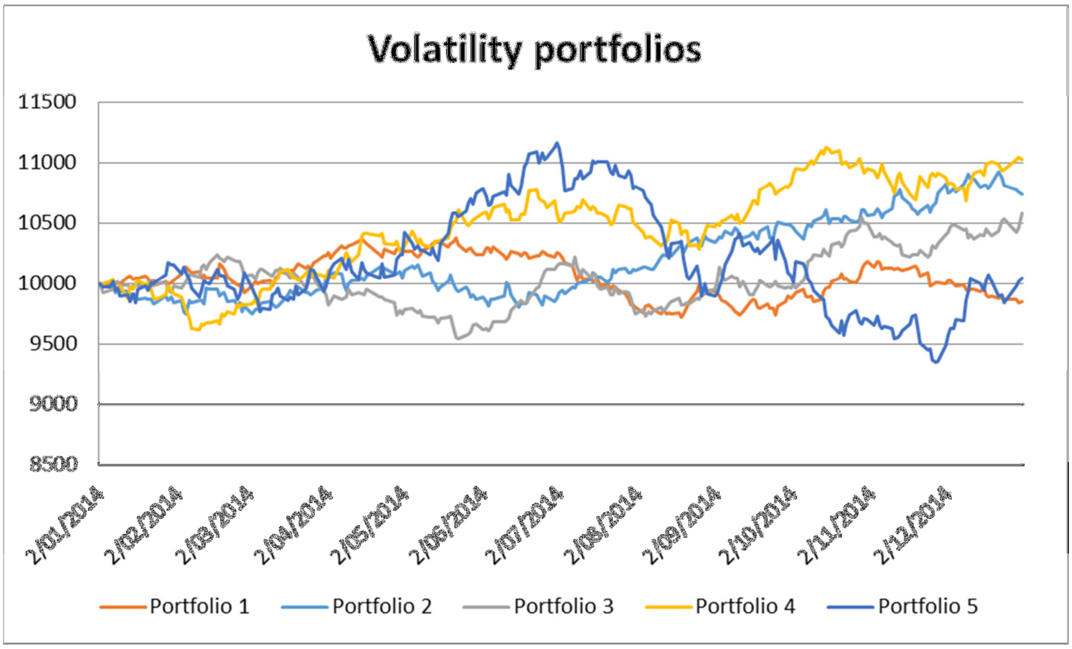

Table 4 shows the results of each method used to compare simulated and actual stock prices for each portfolio. Based on the correlation coefficients shown in panel A there are no clear patterns shown between the different portfolios. The average correlation between simulated prices and actual prices appears to be higher for portfolios 2 and 3, and negative for portfolios 1 and 5. This could suggest that there is an optimal amount of volatility in a stock for the GBM assumption to apply, which makes sense because the model relies on both certain and uncertain components when making forecasts. In the absence of one of these components they model has no applicability. Another observation that can be made is that the correlation between simulated and true share prices seems independent of the time period being used, as each differs for the different portfolios.

When the stocks are grouped into portfolios it seems that the absolute percentage errors decline overall, particularly for short simulation periods. As more stocks are added to a portfolio this diversifies the risk and consequently the percentage of errors in the forecasts decrease. This effect was also observed by Sengupta (2004) who found that by holding stocks for longer periods and diversifying investments, investors were able to reduce their risk. Portfolio 1 seems to have a lower MAPE in longer forecasting periods than the other stocks and portfolio 5 with much greater volatility seems to have a higher MAPE. This provides support for the theory that by holding stocks for longer you reduce your risk.

Panel C reports the accuracy of the model in predicting the direction of movement in stock prices. Grouping stocks into portfolios based on volatility leads to some interesting results, as it appears that portfolio 3 outperformed the other portfolios in both the short-term and long-term. In three of the five simulation timeframes, portfolio 3 had the highest rate of predicting the direction of stock price movements consistent with the market. This result is important because the stocks that make up this portfolio have a moderate level of volatility, reaffirming the results found using the correlation method. Based on the sample of firms used for this paper, stocks with a moderate level of volatility tend to provide more accurate forecasts using the GBM model than firms with high volatility or low volatility.

Table 4: Summary of results for test two

	1 week	2 weeks	1 month	6 months	12 months
Panel A: Correlation					
Portfolio 1	0.2134	-0.6524	-0.2377	0.6095	-0.2631
Portfolio 2	0.1013	0.1282	0.0609	0.3650	0.6824
Portfolio 3	0.9086	0.8196	0.4635	-0.5316	-0.0973
Portfolio 4	-0.3206	-0.0276	0.3263	0.1404	0.5198
Portfolio 5	-0.4105	-0.4993	0.3165	0.2984	-0.0968
Panel B: Mean absolute percentage error (MAPE)					
	1 week	2 weeks	1 month	6 months	12 months
Portfolio 1	0.18%	0.67%	1.03%	1.06%	3.04%
Portfolio 2	0.25%	0.44%	0.74%	3.51%	3.71%
Portfolio 3	0.43%	0.54%	0.74%	4.49%	5.14%
Portfolio 4	0.48%	0.62%	0.86%	2.69%	3.01%
Portfolio 5	0.44%	0.76%	1.15%	4.26%	7.95%
Panel C: Direction prediction accuracy					
	1 week	2 weeks	1 month	6 months	12 months
Portfolio 1	40.00%	40.00%	30.00%	40.98%	45.24%
Portfolio 2	60.00%	60.00%	50.00%	51.64%	53.57%
Portfolio 3	80.00%	80.00%	55.00%	52.46%	52.78%
Portfolio 4	60.00%	40.00%	65.00%	51.64%	52.38%
Portfolio 5	40.00%	60.00%	50.00%	50.00%	47.62%

4.4 Test three – expected return portfolios

This test uses a similar approach to the previous test whereby stocks are aggregated into portfolios based on their expected annual returns. The aim of this test is to examine the third hypothesis that there is no significant relationship between expected stock returns and the difference between actual and simulated stock prices. One would intuitively assume that stocks with a higher expected return would perform better than those with lower expected returns. However, as shown in figure 3 below, this may not always be the case. Although portfolio 5 and portfolio 3 perform well, their performance is matched by portfolio 2 containing firms with moderate-to-low expected returns. Forecasts for portfolio 4 show that over the simulation period these stocks performed poorly, despite having moderate-to-high expected returns. The figure also shows that all five portfolios exhibit roughly the same amount of volatility; this is because only the certain component of the GBM model is being affected by grouping stocks on their respective return expectations.

Figure 3: Line chart showing share price simulations for portfolios of varying levels of expected returns

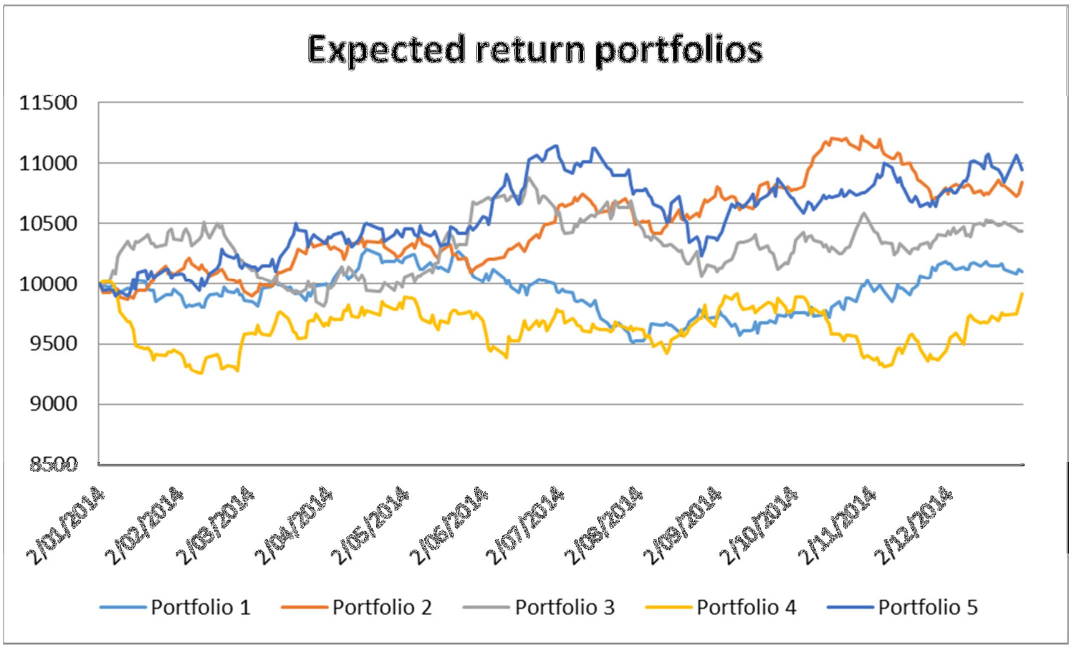

Table 5 summarises the results obtained under each of the measures of the accuracy of the GBM model for simulating stock prices. From the correlation coefficients given in Panel A, the same conclusions can be made as in test 1. Specifically, across all the portfolios the strength of relationship between forecasted stock prices and actual stock prices for portfolio 2 and 5 grows when longer time periods are considered. Overall, there are no discernible trends in the data suggesting that expected returns have an influence on the relationship between simulated prices and actual prices.

Table 5 Panel B shows the MAPE calculated for a holding of each portfolio over the differing time periods. The findings related to the absolute percentage error under this test are very similar to those of test 1 and 2. Again, there seems to be a strong relationship between time periods for prediction and the MAPE value, where longer periods such as six months and twelve months tend to increase MAPE. Also the diversification of individual stocks into portfolios leads to a reduction in the amount of risk bared by an investor, and consequently this has flow-on effects reducing the amount of prediction errors.

In terms of the accuracy of GBM in predicting directional share price movements, grouping firms by their respective return expectations tends to improve the model's forecasting power over a longer period. In earlier tests it was found that based on the accuracy of direction predictions the GBM assumption was only relevant in short time periods of two weeks or less. By forming portfolios based on expected returns this brought the average direction prediction accuracy percentage to 60% over a one month simulation period.

Table 5: Summary results for test three

	Panel A: Correlation				
	1 week	2 weeks	1 month	6 months	12 months
Portfolio 1	-0.7248	0.3521	0.2783	0.2491	0.0706
Portfolio 2	0.2714	0.3790	-0.1633	0.5156	0.6399
Portfolio 3	0.7571	0.4935	0.2191	0.0372	0.0915
Portfolio 4	-0.8032	-0.7816	-0.6117	0.2770	0.1974
Portfolio 5	0.4832	-0.2810	-0.2805	0.6886	0.6720
	Panel B: Mean absolute percentage error (MAPE)				
	1 week	2 weeks	1 month	6 months	12 months
Portfolio 1	0.90%	0.99%	1.29%	2.25%	6.16%
Portfolio 2	0.30%	0.53%	0.94%	2.23%	2.25%
Portfolio 3	0.76%	1.60%	2.28%	2.61%	2.27%
Portfolio 4	1.15%	2.33%	4.22%	6.94%	9.24%
Portfolio 5	0.51%	1.27%	2.18%	2.41%	2.78%
	Panel C: Direction prediction accuracy				
	1 week	2 weeks	1 month	6 months	12 months
Portfolio 1	40.00%	70.00%	65.00%	51.64%	48.02%
Portfolio 2	80.00%	70.00%	50.00%	50.82%	50.79%
Portfolio 3	40.00%	60.00%	70.00%	48.36%	46.83%
Portfolio 4	40.00%	40.00%	60.00%	51.64%	49.60%
Portfolio 5	60.00%	50.00%	55.00%	50.82%	49.21%

4.5 Test four – industry portfolios

For this final test, the individual stocks from the sample were aggregated into portfolios based on their industry sector. This test examines whether there are any industry effects when comparing stock prices simulated using GBM with actual prices. Figure 4 provides a chart for each industry sector to allow visual comparison between simulated stock prices and actual stock prices. All simulations were conducted over the period 1 January 2014 to 31 December 2014, based on an investment of $1000 in each company from their respective industry, resulting in different sized portfolios.

The charts show that in the short-term simulated stock prices tracked closely to actual prices, however after around 3 to 4 months they start to deviate significantly. The Healthcare portfolio demonstrates this phenomenon very well as the prices don't diverge until around June, where they steadily grow further apart until the end of the simulation period. As a general rule, the process of geometric Brownian motion observed in these charts tends to follow a slight upward trend which prices of stocks don't tend to deviate significantly from this path. This can make forecasts difficult in an industry with cyclical patterns, seasonality, or during periods of economic downturn. These effects would account for many of the deviations seen in the charts in Figure 4. Based on the trends shown in actual prices, GBM would be most suited in forecasting prices in industries with predictable and stable growth, such as industrials, healthcare and real estate sectors.

Table 6 shows the measures used to compare simulated and actual prices, and summarises the results for each industry portfolio. From Panel A we can deduce that the financials portfolio displayed the most correlation between simulated prices and actual prices over each of the time horizons considered. It is also evident that simulated stock prices are largely negatively correlated with real stock prices over short forecast horizons, but over longer horizons they become more positively related, though the relationship is not strong.

Figure 4: Line charts showing simulated and actual prices for industry portfolios

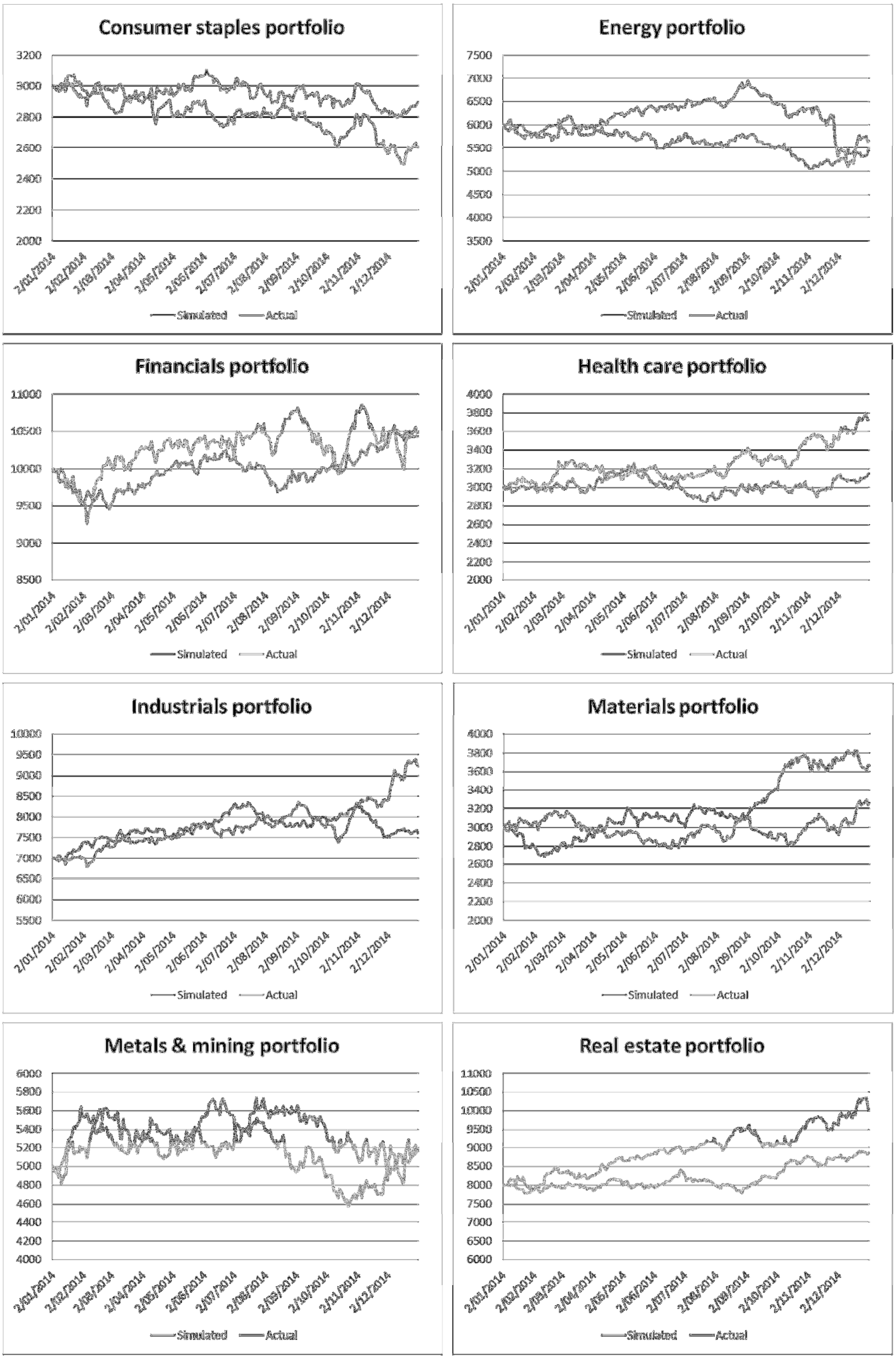

Table 6: Summary of results for test 4

	Panel A: Correlation				
	1 week	2 weeks	1 month	6 months	12 months
Consumer Staples	-0.0441	-0.1413	0.1553	-0.2712	0.4744
Energy	-0.2698	-0.3080	-0.0219	-0.5245	0.0854
Financials	-0.2879	0.4796	0.7903	0.6918	0.4907
Health Care	0.6640	0.0497	0.3206	0.1669	0.1410
Industrials	-0.7960	-0.0563	0.6324	0.8371	0.4142
Materials	-0.7498	-0.5144	-0.5686	-0.8356	-0.0146
Metals & Mining	0.0760	0.4145	0.8165	0.0645	0.2410
Real Estate	-0.3958	-0.4800	0.3472	0.6867	0.7637
	Panel B: Mean absolute percentage error (MAPE)				
	1 week	2 weeks	1 month	6 months	12 months
Consumer Staples	0.59%	0.95%	1.64%	3.74%	5.32%
Energy	2.54%	2.18%	2.14%	5.48%	9.39%
Financials	0.63%	0.73%	0.85%	2.70%	3.29%
Health Care	0.75%	1.72%	1.97%	3.40%	7.24%
Industrials	1.27%	1.28%	2.05%	2.69%	4.24%
Materials	1.71%	2.08%	4.60%	7.15%	11.22%
Metals & Mining	0.92%	2.26%	3.54%	3.89%	5.14%
Real Estate	1.15%	1.56%	1.98%	5.32%	8.34%
	Panel C: Direction prediction accuracy				
	1 week	2 weeks	1 month	6 months	12 months
Consumer Staples	60.00%	50.00%	40.00%	49.18%	48.81%
Energy	60.00%	40.00%	40.00%	48.36%	49.21%
Financials	60.00%	70.00%	50.00%	45.90%	46.83%
Health Care	20.00%	50.00%	40.00%	44.26%	48.02%
Industrials	0.00%	40.00%	45.00%	53.28%	53.57%
Materials	20.00%	30.00%	45.00%	45.90%	48.41%
Metals & Mining	80.00%	70.00%	55.00%	50.00%	51.19%
Real Estate	20.00%	30.00%	55.00%	51.64%	50.79%

5 Conclusions

This study explores the geometric Brownian motion model for simulating stock price paths, and provides three methods to test the validity of the model. The first method calculates the correlation coefficient between simulated stock prices and actual stock prices. Most of the prior studies have suggested that there is a weak relationship between the two variables. We have reported a negative correlation during short periods of simulation, which becomes positive with longer forecast horizons. Noise or volatility in the market makes simulated stock price and actual stock price to have a negative correlation in the short term, whereas stock prices stabilise to its mean value in the longer run causing a positive correlation between simulated and actual stock prices. However, the correlation coefficient still represents only a weak relationship at best.

The second method used is the mean absolute percentage error (MAPE) technique. Using this technique yields different results to the first method as the MAPE values are relatively low over all time periods. It was found that MAPE was lowest over simulation periods of one week, two weeks, and one month, but the error tended to increase when longer horizons were considered.

The third and final method used a simple process for checking whether simulated daily stock prices exhibited the same directional movement as actual stock prices. The findings were slightly encouraging as results show that over all time horizons the chances of a stock price simulated using GBM moving in the same direction as real stock prices did was just a little greater than 50 percent. However, it was later found that when portfolios were formed that these odds went up slightly.

However, this is the first study and the literature relating to the testing of the GBM assumptions are very limited. Therefore, there were some limitations which also provide potential areas for future research. For example, this study used sample from large listed companies from a single country. Therefore, including other countries may yield different results, and a larger sample would improve the validity of the conclusions. Future research involving different periods with different start dates could be considered. Other modifications could also be used to further the reliability of the model, such as a model incorporating jumps. Also, it would be interesting to compare the accuracy of the RIM and GBM model in predicting stock prices.

References

Abidin, S., & Jaffar, M. (2014). Forecasting share prices of small size companies in bursa Malaysia using geometric Brownian motion. *Applied Mathematics & Information Sciences, 8*(1), 107-112. http://dx.doi.org/10.12785/amis/080112

Benjamin, L., & Bin, L. (2011). Monthly seasonality in the top 50 Australian stocks. *Journal of Modern Accounting and Auditing, 7*(4), 380-390. Retrieved from http://www98.griffith.edu.au/dspace/

Brewer, K., Feng, Y., & Kwan, C. (2012). Geometric Brownian motion, option pricing, and simulation: some spreadsheet-based exercises in financial modelling. *Spreadsheets in Education, 5*(3), article 4. Retrieved from http://epublications.bond.edu.au/ejsie/

Ermogenous, A. (2005). *Brownian motion and its application in the stock market*. Retrieved from http://academic.udayton.edu/EPUMD/Volume%202_2006/brownian2.pdf

Fama, E. (1995). Random walks in stock market prices. *Financial Analysts Journal, 51*(1), 75-80. http://dx.doi.org/10.2469/faj.v51.n1.1861

Gajda, J., & Wylomanska, A. (2012). Geometric Brownian motion with tempered stable waiting times. *Journal of Statistical Physics, 148*(2), 742-752. http://dx.doi.org/10.1007/s10955-013-0754-4

Granger, C. (1992). Forecasting stock market prices: lessons for forecasters. *International Journal of Forecasting, 8*(1), 3-13. http://dx.doi.org/10.1016/0169-2070 (92)90003-R

Hadavandi, E., Shavandi, H., & Ghanbari, A. (2010). Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting. *Knowledge-Based Systems, 23*(8), 800-808. http://dx.doi.org/10.1016/j.knosys.2010.05.004

Higgins, H. (2011). Forecasting stock price with the residual income model. *Review of Quantitative Finance and Accounting, 36*(4), 583-604. http://dx.doi.org/10.1007/s11156-010-0187-y

Hsu, Y., Liu, M., Yeh, J., & Hung, H. (2009). Forecasting the turning time of stock market based on Markov-Fourier Grey model. *Expert Systems With Applications, 36*(4), 8597-8603. http://dx.doi.org/10.1016/j.eswa.2008.10.075

Ladde, G., & Wu, L. (2009). Development of modified geometric Brownian motion models by using stock price data and basic statistics. *Nonlinear Analysis, 71*(12), e1203-e1208. http://dx.doi.org/10.1016/j.na.2009.01.151

Marathe, R., & Ryan, S. (2005). One the validity of the geometric Brownian motion assumption. *The Engineering Economist: A Journal Devoted to the Problems of Capital Investment, 50*(2), 159-192. http://dx.doi.org/10.1080/00137910590949904

Sengupta, C. (2004). *Financial modelling using Excel and VBA*. New Jersey, United States: John Wiley & Sons.

Appendices

	Correlation				
	1 week	2 weeks	1 month	6 months	12 months
AGL-AU	0.3819	0.3127	0.2366	-0.0769	0.5966
AIO-AU	-0.9582	-0.5190	-0.4053	-0.3947	0.0559
AMC-AU	0.6035	0.8450	0.0967	0.1571	0.8460
AMP-AU	-0.6500	0.0720	0.3740	0.7509	0.3823
ANZ-AU	0.7305	0.6551	0.6722	0.2017	0.0172
APA-AU	-0.2973	-0.1523	0.2527	0.2727	0.8667
ASX-AU	0.6652	-0.0952	0.4326	-0.0847	0.0141
AWC-AU	-0.6062	-0.8442	-0.7580	0.6103	0.8635
AZJ-AU	0.0066	0.1306	0.3768	0.0638	-0.3928
BHP-AU	0.7417	-0.6118	-0.1184	-0.2105	-0.5117
BXB-AU	-0.2758	-0.5185	-0.1367	0.6256	-0.2492
CBA-AU	-0.6777	0.7393	0.5100	0.5820	0.1274
CCL-AU	-0.0900	0.4943	0.8072	0.6148	0.7873
CIM-AU	0.1604	-0.2982	0.1564	0.7275	0.6058
CPU-AU	-0.1825	0.4266	-0.1997	-0.7159	-0.0925
CSL-AU	-0.8990	0.1715	-0.7833	0.0361	-0.1199
CTX-AU	0.1535	0.3556	0.4500	-0.8167	-0.2669
CWN-AU	-0.1511	-0.2401	-0.4343	0.2145	0.4532
DXS-AU	-0.7517	-0.6851	-0.4871	0.4648	0.5922
FDC-AU	-0.0564	-0.4734	-0.5812	-0.4613	0.5982
GMG-AU	-0.8062	0.2877	0.4706	-0.5496	0.5639
GPT-AU	-0.9318	-0.9078	-0.5569	-0.6978	-0.8361
IAG-AU	-0.2753	0.3566	0.6817	0.2346	0.7870
ILU-AU	0.4537	0.1768	0.5370	0.4832	0.7008
IPL-AU	-0.8778	-0.9580	-0.8215	0.0173	-0.0659
LLC-AU	-0.7521	-0.8470	0.3696	0.8685	0.9357
MGR-AU	0.1436	-0.5028	0.1233	0.3891	-0.3170
MQG-AU	0.7305	0.2906	-0.3343	0.6665	0.6286
NAB-AU	0.9886	-0.4087	0.5164	-0.3618	-0.4838
NCM-AU	-0.8344	0.5451	0.8770	0.1419	0.2513
ORG-AU	0.9255	0.2317	-0.0954	0.3245	-0.4913
ORI-AU	-0.7276	-0.1334	0.2961	-0.3877	-0.5373
OSH-AU	0.7920	-0.0097	-0.3857	-0.9184	-0.2748
QAN-AU	0.6200	-0.0599	-0.3504	0.6897	0.2179
QBE-AU	-0.7360	-0.2956	0.2871	0.1387	0.2078
RHC-AU	0.7974	0.8224	0.2738	0.3994	0.6819
RIO-AU	-0.0340	-0.1065	0.0541	-0.4237	-0.0858
SGP-AU	0.8359	0.8246	0.6786	0.6599	0.8183
SHL-AU	-0.1256	-0.5920	-0.2918	0.4986	-0.3993
STO-AU	-0.6659	-0.4923	0.3405	-0.3692	0.3743
SUN-AU	-0.6412	-0.2488	-0.7915	0.6312	0.5564
SYD-AU	-0.0363	0.4376	0.4615	0.1241	0.5072
TCL-AU	0.5044	0.0707	0.2456	0.6945	0.0349
TLS-AU	0.9729	0.1631	0.7881	-0.7355	-0.4091
WBC-AU	0.2738	0.7234	0.3392	-0.5342	-0.2075
WES-AU	0.1848	-0.5789	-0.6383	-0.4660	-0.0212
WFD-AU	-0.2323	0.2429	0.2635	0.1703	-0.5459
WOR-AU	-0.7277	-0.7776	-0.3319	0.2355	0.4915
WOW-AU	0.1893	-0.2245	-0.0498	-0.2696	0.5857
WPL-AU	-0.4290	-0.1443	-0.4819	0.5646	0.3679

Table 7: Appendix - Correlation for each individual stock

	Mean absolute percentage error (MAPE)				
	1 week	2 weeks	1 month	6 months	12 months
AGL-AU	1.94%	2.43%	2.72%	3.02%	4.05%
AIO-AU	1.95%	2.61%	2.79%	11.25%	8.09%
AMC-AU	1.47%	1.25%	2.17%	7.62%	17.71%
AMP-AU	1.08%	1.24%	1.62%	7.70%	10.90%
ANZ-AU	0.52%	1.29%	1.65%	4.60%	4.43%
APA-AU	1.93%	1.69%	1.87%	5.01%	4.81%
ASX-AU	3.32%	4.89%	4.52%	4.56%	6.35%
AWC-AU	5.27%	11.90%	18.26%	15.96%	15.77%
AZJ-AU	0.88%	1.92%	1.93%	3.13%	7.00%
BHP-AU	0.67%	3.52%	5.04%	11.01%	16.13%
BXB-AU	3.27%	4.86%	6.42%	8.36%	8.71%
CBA-AU	1.22%	1.17%	1.75%	2.64%	3.65%
CCL-AU	1.68%	1.34%	1.56%	7.87%	6.12%
CIM-AU	1.21%	1.79%	5.28%	9.11%	7.46%
CPU-AU	3.16%	3.21%	2.93%	12.00%	8.18%
CSL-AU	2.77%	2.21%	3.60%	12.86%	19.18%
CTX-AU	2.02%	2.61%	4.36%	25.35%	33.59%
CWN-AU	3.34%	4.17%	4.41%	10.36%	7.27%
DXS-AU	2.04%	2.65%	2.29%	3.99%	5.08%
FDC-AU	1.76%	2.97%	3.21%	11.84%	10.08%
GMG-AU	1.66%	1.71%	2.35%	8.86%	9.66%
GPT-AU	6.33%	7.64%	9.00%	17.04%	26.77%
IAG-AU	1.79%	1.59%	1.43%	3.17%	9.43%
ILU-AU	1.67%	2.23%	1.95%	14.40%	12.66%
IPL-AU	2.34%	5.18%	7.95%	8.61%	8.66%
LLC-AU	3.38%	4.50%	3.76%	8.77%	10.00%
MGR-AU	1.10%	1.88%	1.78%	2.61%	7.37%
MQG-AU	1.93%	3.12%	2.89%	4.43%	5.92%
NAB-AU	3.18%	2.73%	4.19%	8.24%	6.65%
NCM-AU	5.97%	13.16%	22.09%	14.34%	22.04%
ORG-AU	1.19%	1.37%	1.73%	6.39%	11.54%
ORI-AU	2.93%	2.31%	4.86%	11.71%	17.10%
OSH-AU	1.27%	1.76%	3.63%	19.82%	24.58%
QAN-AU	3.18%	2.68%	3.61%	7.17%	13.35%
QBE-AU	6.85%	9.95%	11.62%	8.91%	8.77%
RHC-AU	2.36%	2.19%	2.04%	9.79%	7.21%
RIO-AU	4.86%	7.56%	10.60%	18.07%	19.63%
SGP-AU	1.14%	1.03%	2.51%	6.50%	6.38%
SHL-AU	2.42%	4.73%	3.78%	5.64%	9.55%
STO-AU	3.90%	5.08%	5.74%	10.95%	18.42%
SUN-AU	0.87%	0.90%	2.56%	10.80%	8.21%
SYD-AU	0.92%	1.02%	2.50%	9.57%	7.63%
TCL-AU	1.22%	2.32%	2.64%	3.47%	6.67%
TLS-AU	1.71%	1.87%	2.02%	5.29%	7.95%
WBC-AU	1.72%	2.14%	1.47%	13.61%	22.27%
WES-AU	2.48%	4.84%	8.03%	19.28%	27.64%
WFD-AU	1.31%	1.14%	1.68%	5.28%	19.02%
WOR-AU	6.00%	4.42%	5.49%	10.68%	13.58%
WOW-AU	2.31%	3.37%	2.91%	11.55%	15.23%
WPL-AU	3.66%	3.23%	2.65%	6.34%	13.24%

Table 8: Appendix - MAPE for each individual stock

	Direction prediction accuracy				
	1 week	2 weeks	1 month	6 months	12 months
AGL-AU	60.00%	60.00%	55.00%	56.56%	53.97%
AIO-AU	0.00%	20.00%	45.00%	53.28%	55.16%
AMC-AU	40.00%	60.00%	50.00%	49.18%	50.79%
AMP-AU	60.00%	60.00%	50.00%	54.92%	51.98%
ANZ-AU	80.00%	50.00%	45.00%	47.54%	45.24%
APA-AU	20.00%	30.00%	55.00%	53.28%	52.78%
ASX-AU	60.00%	60.00%	55.00%	54.10%	55.95%
AWC-AU	60.00%	40.00%	45.00%	54.10%	47.62%
AZJ-AU	60.00%	60.00%	50.00%	47.54%	47.22%
BHP-AU	80.00%	70.00%	65.00%	54.10%	57.94%
BXB-AU	20.00%	30.00%	40.00%	55.74%	53.57%
CBA-AU	20.00%	40.00%	50.00%	45.90%	49.21%
CCL-AU	80.00%	80.00%	65.00%	55.74%	54.37%
CIM-AU	40.00%	60.00%	60.00%	43.44%	50.79%
CPU-AU	60.00%	60.00%	45.00%	51.64%	49.21%
CSL-AU	40.00%	40.00%	45.00%	54.10%	51.19%
CTX-AU	60.00%	50.00%	55.00%	59.84%	53.57%
CWN-AU	60.00%	70.00%	75.00%	47.54%	49.21%
DXS-AU	40.00%	30.00%	45.00%	47.54%	52.38%
FDC-AU	80.00%	80.00%	70.00%	53.28%	56.75%
GMG-AU	60.00%	70.00%	65.00%	58.20%	51.19%
GPT-AU	40.00%	30.00%	40.00%	55.74%	56.75%
IAG-AU	60.00%	50.00%	50.00%	55.74%	57.94%
ILU-AU	60.00%	60.00%	60.00%	50.82%	54.37%
IPL-AU	40.00%	20.00%	25.00%	51.64%	55.56%
LLC-AU	40.00%	20.00%	40.00%	50.82%	52.78%
MGR-AU	40.00%	30.00%	50.00%	62.30%	57.14%
MQG-AU	20.00%	30.00%	20.00%	50.82%	50.00%
NAB-AU	100.00%	60.00%	50.00%	53.28%	52.38%
NCM-AU	20.00%	30.00%	45.00%	49.18%	51.19%
ORG-AU	80.00%	70.00%	65.00%	50.82%	53.57%
ORI-AU	0.00%	30.00%	40.00%	49.18%	52.38%
OSH-AU	40.00%	30.00%	35.00%	50.00%	50.00%
QAN-AU	80.00%	70.00%	65.00%	62.30%	58.33%
QBE-AU	20.00%	30.00%	40.00%	46.72%	47.62%
RHC-AU	60.00%	70.00%	65.00%	53.28%	55.16%
RIO-AU	40.00%	50.00%	55.00%	47.54%	48.02%
SGP-AU	80.00%	60.00%	65.00%	64.75%	60.71%
SHL-AU	80.00%	60.00%	40.00%	49.18%	50.00%
STO-AU	60.00%	50.00%	55.00%	50.82%	48.02%
SUN-AU	60.00%	60.00%	35.00%	50.82%	51.98%
SYD-AU	80.00%	80.00%	85.00%	67.21%	59.92%
TCL-AU	100.00%	60.00%	50.00%	60.66%	56.75%
TLS-AU	100.00%	80.00%	70.00%	54.10%	51.19%
WBC-AU	60.00%	50.00%	45.00%	52.46%	48.81%
WES-AU	40.00%	40.00%	40.00%	46.72%	47.22%
WFD-AU	40.00%	50.00%	50.00%	49.18%	47.22%
WOR-AU	40.00%	30.00%	45.00%	46.72%	48.41%
WOW-AU	40.00%	30.00%	35.00%	40.98%	42.06%
WPL-AU	20.00%	40.00%	40.00%	47.54%	49.21%

Table 9: appendix - Direction prediction accuracy for each individual stock