MODELING THE STOCK MARKET : THE ROLE OF MARKOV CHAINS AND DOUBLE GAMMA DISTRIBUTION

by

HOLMES CHARLES PATRICK BANDA

B.Ed(Science), BSc (Hons)

A thesis
submitted to the University of Malawi
in partial fulfillment of the
requirements for the degree of
Master of Science
in Mathematical Sciences

UNIVERSITY OF MALAWI

Chancellor College

December 2007

Declaration by the student

I declare that this is my own work and reference has been made to wherever other people's work has been used. I also declare that this work has not been presented or submitted for any award elsewhere.

.... **Date:**.................................

Holmes Charles Patrick Banda.

CERTIFICATION OF APPROVAL

We declare that this thesis is the result of the candidate's own work, and that where assistance has been sought, this has been acknowledged and that this thesis has been submitted with our approval.

.... **Date:**.................................

Dr J.J. Namangale (Supervisor).

.... **Date:**.................................

Dr L. Kazembe (Internal Examiner).

.... **Date:**.................................

Dr R. Mangani (Internal Assessor).

Dedication

To the most wonderful and beautiful lady that I shared a short but very productive and exciting life, my late wife Klezia; late father, Patrick Gwamanda Banda for the seed of life and upbringing, and the centroid of my joy and comfort - daughters Latanze and Lungile.

Acknowledgments

I would like to thank my supervisor Dr. Jimmy J. Namangale for his excellent assistance and invaluable direction and encouragement that he gave me throughout my studies, in general, and during all undertakings that culminated into the eventual production of this thesis, in particular.

Special thanks to Management of the Malawi Polytechnic for giving me the necessary support and according me the opportunity to complete my studies.

To my late wife and daughters, many thanks for the endurance, love and encouragement while I excused myself from family chores in pursuit of academic glory. I am particularly heavily indebted to my late wife who never lived to enjoy the fruits of her contribution towards my studies and life. May her soul rest in eternal peace.

Abstract

Determination of values of stocks and indices is very challenging and a very important aspect in finance. Knowing the value of a stock price, for instance, can be very vital in pricing various financial derivatives such as options. On the other hand, indices are useful tools for tracking stock market trends. By studying the pattern of index values over time, investors might gain insight that would help them make better investment decisions.

Various attempts have been made to predict future stock prices and index values but these have yielded mixed results due to the stochastic nature of financial markets. Despite the fact that there has been growing academic interest in the stock market, it still remains elusive as to what the next day's price of a stock, in particular, and value of an index, in general, will be even when the prices of the present and previous days are known. One of the popular approaches of pricing options, for instance, has been through the use of the Black-Scholes model. This and various other approaches have placed normality at the centre of the stochastic modeling. In this thesis, a statistical analysis on different indices and stocks traded on the world's major financial markets is performed and demonstrates, through simulation, that Markov chains and the double gamma distribution play a central role in the stochastic behaviour of prices on the stock market.

Contents

CHAPTER 1

Introduction

1.1. The Behaviour of Stock Prices

One of today's most vibrant financial institutions is probably the stock market. Over the last few decades, what was once an exclusive club of the rich has swiftly turned into every investor's domain for growing wealth. The growing interest and advances in trading technology have opened up markets such that today nearly anybody can own stocks. It is no wonder, therefore, that major media publications dedicate entire pages or airtime to report news on how major companies perform on the stock market. It has become an institution of great interest to investors as well as the public at large (Granger and Morgenstern (1970)).

While there has been growing interest and popularity in the stock market, it remains elusive what the next day's price of a stock will be even when the previous or present day's price is known. The determination of a stock price is a very important aspect in finance. The most basic motivation for predicting stock prices is financial gain. Every stakeholder in the world of finance seeks a position of advantage over competitors. It is not surprising therefore that nearly all investors continuously look for opportunities that will earn them high returns.

Since the price of derivatives is closely related to that of the underlying asset, the value of the underlying asset such as a stock price is used in pricing derivatives (Hull (1989)). A very popular and typical example is the pricing of options using the Black-Scholes model and binomial option pricing model. Both involve the price of stock in their formulae as shall be reviewed later. Other examples include swaps and futures/forwards (Hull (1989), Joshi (2004) and Neftci(1996)).

Many attempts have been made to predict future stock prices but have failed due to the unpredictable nature of financial markets. Fundamental and technical analyses of predicting stock prices have used statistical modelling techniques. None of the techniques, however, has proven to predict stock prices consistently thereby casting a shadow on the usefulness of many of the approaches.

Today most models do not focus on daily predictions of stock prices but other attributes. For instance, the Log-Asymmetric Conditional Duration (Log-ACD) model of Bauwens and Giot assists in finding the probability of a price increase (or decrease) at the time of the last quotes announcement given the past information (Chou (2001)).

In 1900, Louis Bachelier used the Brownian motion concept to model prices of stocks and commodities at the Paris Bourse. However the Brownian motion modelling has some flaws when used to model stock prices since, firstly, as the stock price is assumed to be a normal random variable, it can theoretically become negative. Secondly, stock prices often change in proportion to their size but Brownian modelling does not take this property into account (Ross (1996)).

In the next sections, we review basic definitions and concepts that are central to the study of finance and financial mathematics. The work in this thesis embraces three major disciplines which are finance, financial mathematics and stochastic processes. All the concepts and definitions reviewed in this chapter can be found in standard texts in financial mathematics and finance (see Hull (1989), Joshi (2004), Wilmott (1995), Ross (1999) and Neftci(1996)).

1.2. Financial Markets

As stocks are traded on the world's financial markets, we will highlight important aspects of this crucial medium in the world of finance. As commonly used in economics, we define a *financial market* as a setup which allows people to trade money for securities or commodities such as gold or other precious metals. Any commodity market, in general, may be viewed as a financial market provided the traders' objective is not immediate consumption of the commodity but as a means of controlling consumption over time. The market provides a medium through which funds are transferred from those who have excess funds (savers, lenders) to those who have a shortage (borrowers). Financial markets are classified according to the type of commodity being traded.

1.2.1. Capital Markets. We use the term *long-term capital* to refer to capital that is invested or lent and borrowed for long periods of time spanning over five years in most developed economies. *Capital markets* are markets for long-term capital and mainly consists of stock markets and bond markets. Stock markets provide financing through the issuance of shares or common stock. On the other hand, bond markets provide financing through the issuance of bonds. These are dealt with separately below.

1.2.1.1. *Stock Market.* In these markets the most commonly traded asset is the share or stock in a company. A *share* in a company is a fraction of ownership in a particular company and thus holders of shares own a fraction of a company (see Joshi (2004)). The markets that facilitate the trading of shares are called *stock* or *equity markets.* As shares are bought publicly on the equity market, companies traded on the stock market are public limited companies (plc). While a shareholder owns part of a company, his or her liabilities are limited to the amount invested, that is, the shareholder has no liability for its debts in case it goes bankrupt. Further, the shareholder may make money in two ways. Firstly, if the share price goes up, the shareholder may decide to sell the shares at a profit. Secondly, the company pays dividends to shareholders, dispensing the profits made by the company during a specified period.

The stock market is used to describe the totality of all stocks excluding bonds, securities and derivatives. It is one of the most crucial areas of a market economy as it is an avenue through which most companies raise their capital while providing income to investors. A stock market is, however, different from a Stock Exchange. The latter involves bringing buyers and sellers of stocks and securities together. Thus a stock exchange is a marketplace where buyers and sellers meet and agree on a price. The oldest stock exchange is the New York Stock Exchange having been formed in 1792 (Granger and Morgenstern (1970)). Since then there has been a boom of stock exchanges with the most popular ones being the European stock exchanges like the London Stock Exchange, the Tokyo Stock Exchange in Asia and the Johannesburg Stock Exchange in Africa. The Malawi Stock Exchange was inaugurated in 1995 but opened for business for the first time in November 1996 when it first listed National Insurance Company Limited (NICO).

A stock market *index*, on the other hand, is a listing of stock and a statistical measure that reflects the performance of a specific "basket" or portfolio of stocks considered to represent a particular market or sector of the economy (Hull (1989)). Indices (or indexes) often serve as barometers against which financial or economic performance is measured. For example, the S&P 500 Composite Stock Price Index is an index of 500 stocks representing major companies in leading industries within the U.S. economy. Stocks in the index are chosen for market size, liquidity, and industry group representation. We describe other major indices and stocks studied in this thesis in subsection 3.1.1.

1.2.1.2. *Bond Market.* In general terms, a *bond* is a contract through which an investor loans money to an entity (such as a company or government) that borrows the funds for a defined period of time at a specified interest rate. The original sum of money borrowed, called *principal*, is returned to the holder or investor at the expiration of a preagreed period, called the *maturity date*. The holder receives interest payments, called *coupons*, as compensation for the investor's release of money that is borrowed from the public by government or corporate institutions (see Joshi (2004)). If there is no coupon the bond is known as a *zero-coupon bond*. These instruments are traded in bond markets.

1.2.2. Money Markets. The money market is the financial market whose aim is to facilitate the lending and borrowing of money on a short-term basis. Money markets are operated mainly by banks and other financial institutions. These provide short term debt financing and investment. Money market instruments are a form of short-term debts that mature in less than one year. Examples of instruments that are traded on money markets include drafts or bills of exchange, treasury bills, short-term loans also technically called *repurchase agreements* and *certificates of deposits*. Certificates of deposits are issued by a bank acknowledging that a certain amount of money has been deposited with it for a certain period of time.

While money markets largely involve borrowing and lending by banks, other large companies and nationalised industries as well as the government are also involved in money market operations. Due to the liberalisation of building societies in most developed economies, building societies have lately become major participants in the money market as well. In Malawi, what was formerly called New Building Society has now turned into a fully fledged commercial bank.

1.2.3. Derivatives Markets. A *financial derivative*, also called *contingent claim* is a security whose value depends on the values of other underlying variables or assets. An example of an underlying asset is the value of stock while some examples of financial derivatives are highlighted below. The derivatives markets are markets for financial derivatives. Financial derivatives provide instruments for the management of financial risk (Hull (1989)).

Three common examples of derivatives are: futures and forwards, swaps and options. Futures and forwards are contracts to buy or sell an asset at a specified price at a known future date. On the other hand, *swaps* are agreements where parties agree to exchange cash flows involving various currencies, interest rates and other financial assets at a future agreed date. A definition of options and a thorough treatment of the theory of options can be found in section (1.4). Financial markets on which options and futures/forwards are traded are called *options markets* and *forwards/futures markets* respectively.

1.2.4. Insurance Markets. Insurance is a form of risk management that is primarily used to hedge against the risk of potential financial or material loss. Insurance is defined as the equitable transfer of the risk of a potential loss, from one entity to another, in exchange for a premium and duty of care. The insurance markets facilitate the redistribution of various risks through the trading of various insurance products normally referred to as policies.

There are many types of insurance depending on the type of risk they are supposed to hedge. Common insurance types include motor or car insurance which covers claims against the driver and loss of or damage to the vehicle itself; property insurance which provides protection against risks to property, such as fire, theft or weather damage; and financial loss insurance which protects individuals and companies against various financial risks such as protection from loss of sales if a fire in a factory prevented a company from carrying out its business for a time. Other well known insurance types are life insurance, health insurance, casualty insurance, travel insurance, professional indemnity insurance and marine insurance. The most common and well known insurance in Malawi is the car insurance since it is a legal requirement that all motor vehicles must be insured before they are certified roadworthy.

1.2.5. Foreign Exchange Markets. This is probably the largest financial market in terms of trading volume and the number of participants involved in the market. Foreign exchange markets are found wherever one currency is traded for another. In Malawi, in addition to the regular commercial banks, all major towns have foreign exchange markets commonly known as *forex bureaus*. The primary role of foreign exchange markets is the facilitation of the trading of foreign exchange.

Brief summaries of financial markets can be found in Wilmott et al (1995) and Joshi (2004).

The main question for investors and other market participants is what drives prices in any financial market. The answer to such a question is not obvious, however, an attempt to answer it may be provided by the Efficient Market Hypothesis which is reviewed in the next section.

1.3. Efficient Market Hypothesis

The Efficient Market Hypothesis (EMH) is a concept that claims that the present price of an asset incorporates and reflects all the information presently available including historical information (see Ross (1999), Joshi (2004) and Cockraine (2001)). As a consequence old information cannot be used to foretell future price movements. Since the acquisition of new information is highly competitive, it is not easy to make quick profits (Cockraine (2001)). However critics of the concept claim that past price movements reflect information that has not been universally recognized but will affect future prices. The general belief of the critics is that there is no prior reason why future price movements should be independent of past movements (see Ross (1999)).

There are three forms of the efficient market hypothesis and these are *Weak form* of market efficiency, *Semi-strong form* of market efficiency and *Strong form of market efficiency*. These are reviewed in the next subsection.

1.3.1. Weak form of market efficiency. The "weak" form states that all past market prices and data are fully reflected in securities prices. Any information contained in previous prices has been analysed and acted on by market forces and consequently securities such as stocks are neither under-valued nor over-valued. Proponents of weak form of market efficiency claim that a study and analysis of trends in historical prices, known as *technical analysis*, cannot help in the determination of future market prices.

1.3.2. Semi-strong form of market efficiency. The "semi-strong" form asserts that all publicly available information is fully reflected in securities prices. In a "semi-strong" market, current prices efficiently adjust to information that is publicly available. Since all publicly available information has been thoroughly analysed, assessed and acted upon by a large number of market players, both fundamental and technical analyses are ineffective. *Fundamental analysis* is the analysis of financial information such as company earnings and asset values to help investors select undervalued stocks (see Malkiel (2003)).

In addition, fundamental analysis involves studying the prospects for a company's business (Brealey (1983)).

1.3.3. Strong form of market efficiency. The "strong" form asserts that all information is fully reflected in asset prices. In a "strong form" efficient market, any attempt to make profitable use of monopolistic access to information would be fruitless since any such information has already been incorporated into the market price of the asset. Thus no one even with insider information could have any advantage over other investors.

It is a difficult task to predict values of asset prices. The historical prices are, however, there as a financial time series (Wilmott etal (1995)). The financial time series can be examined to suggest the likely jumps in asset prices, their mean and variance, and the likely distribution of assets. These qualities may be determined by a statistical analysis of historical data. Malkiel (2003) presents the arguments for and against the efficient market hypothesis but concludes that stock markets are far more efficient and far less predictable than what some recent academic papers suggest. On the other hand, while prices on average adjust quickly to firm-specific information a common finding in event studies is that the dispersion of returns increases around information events. Previous work done by various researchers shows that during mergers, stock prices of acquiring firms do not often react to merger announcements but later drift slowly down. Other studies suggest that stock prices do not react swiftly to specific information (see Fama (1991)).

Samuelson (1973) suggests that expected future price must be approximately equal to present price otherwise the present price would be different from what it is. If there were profits to be made, which all market participants could recognise, this would be acted upon quickly thereby raising or lowering the present price. As market participants' expectations of the future are different, they guess differently and this turns out to be the major reason why there are transactions in the marketplace in which one individual is buying and another is selling (Samuelson (1973)).

The one major aspect of asset prices that can be discerned from the last two sections is that asset prices are generally difficult to predict. The uncertainty in future asset prices poses a big problem for investors as any investment decision will harbour some level of risk. To manage risk, some investors use options (see Joshi (2004)). In the next section, we hightlight the main aspects of the theory of options.

1.4. The Theory of Options

In this section we look at the theory of options. The major motivation for reviewing options is that the values of some financial instruments, such as stock, go up and down. Due to their erratic behaviour stocks may be viewed as assets that harbour risk. In the world of finance, derivative products, such as options, are financial instruments that are used to hedge against risk. However, the values of all financial derivatives are contingent on the value of the underlying asset, in our case, stock. It is not surprising therefore that major option pricing models involve the price of stock in their formulae. Hence as the central theme of this thesis is stock pricing, it is imperative that we briefly review the theory of options.

Financial markets, as avenues where buyers and sellers transact business, have become more sophisticated as more complex transactions are being introduced (Wilmott et al (1995)). However, all investment decisions harbour risk and hence require an assessment and diversification of risk (see Joshi (2004)). In order to curtail risk, various financial instruments are used. These include swaps, futures and forwards and options. A thorough treatment of swaps and futures and forwards can be found in Hull (1989) and Neftci (1996). Our review of options is based on Joshi (2004), Wilmott et al (1995) and Cox et al (1979).

1.4.1. The language of options. An *option* is a contract that gives the holder the right, but not the obligation, to buy or sell some quantity of an underlying asset at a prearranged price on or before a certain date. In our case the underlying asset is stock. The price of an option is called the *premium.* The act of using the option is referred to as *exercising* the option. The prearranged price is called the *strike* or *exercise* price while the given date is termed the *expiration* or *exercise* or *maturity* date.

There are two basic types of options: call options and put options. A *call option* gives the holder the right to buy while a *put option* gives the holder the right to sell some quantity of the underlying asset at the prearranged price on or before a certain date. For call options, the option is said to be *in-the-money* if the value of the underlying asset is above the strike price, otherwise it is said to be *out-of-the-money*. On the other hand, a put option is in-the-money if the value of the underlying asset is below the strike price. The value by which an option is in-the-money is called the *intrinsic value*.

Options are classified into two major groups depending on the time they are exercised. Those options that are exercised on the maturity date itself are referred to as *European* options while those that are exercised on any date before the specified date are called *American* options. Option buyers are referred to as *holders* while option sellers are called *writers*. The option buyer is also said to have taken the *long position* and the option seller is said to have taken the *short position*.

1.4.2. The use of options. In general, there are two primary reasons why an investor would want to use options. These are speculation and hedging. An option is a major attraction in the management of risk since the maximum loss that can be incurred is the initial premium (see Joshi (2004). *Hedging* is a means of cushioning an investment against any risk or possible loss. In this way, options can be viewed as an insurance against any adverse movements in the underlying asset. If the value of the underlying asset is less than the exercise price, it does not benefit the call option holder to pay more for an asset that can be purchased for less. On the other hand if the value of the underlying asset is more than the strike price, the call option holder can exercise the option for a profit, that is, buying the asset at the exercise price and selling it at the current market value which is more than the price paid by the option holder. Thus, call option buyers hope that the value of the underlying asset will increase substantially before the option expires. However, buyers of put options hope that the value of the undelying asset plummets in order to make a profit (Wilmott et al (1995)).

Investors who believe that the price of an asset can rise buy stocks in that company. If the price of the asset rises, the investor makes money, otherwise the investor loses money. Such investors are said to be *speculating*. In the case where the investor speculates that the price will fall, one may opt to sell the asset or buy puts. When an investor sells shares that he or she does not own he or she is said to be selling *short* and will thus profit from a fall in shares. Explicit examples can be found in Hull (1989) and Joshi (2004).

1.4.3. Major Option Pricing Models. In this subsection, we review briefly two discrete-time models for valuing options. We specifically review models for valuing European options. The underlying asset is stock. Cox et al (1979) present simpler and more straightforward derivations of the two option pricing models that we review in this thesis.

1.4.3.1. *The Binomial Option Pricing Model.* The Binomial Option Pricing Model is premised on the assumption that the stock price S_t at any time t follows a multiplicative binomial process over discrete periods without paying dividends and transaction costs. Under the binomial process the price of stock in the next period may be in one of the two states, "up" or "down". The movement of stock is observed over n periods.

Let S be the current price of stock and K be the exercise price. Suppose that in each period the price of the stock can go up by u with probability p or down by d with probability *1-p*. Thus, if S is the current stock price, then during the next period the stock price may move from S up to uS or down to dS. Then the price of stock at the end of n periods will be

$$u^j d^{n-j} S, \tag{1.4.1}$$

where j is the number of times the stock is in an "up" state. If the option expires out-of-money, that is,

$$u^j d^{n-j} S < K, \tag{1.4.2}$$

then the stock can be purchased for $u^j d^{n-j} S$ since it is cheaper. Thus the call option has no value. However, the option will have some value if value of stock is greater than the exercise price, that is,

$$u^j d^{n-j} S > K. \tag{1.4.3}$$

(see Netfci (1996)). In this case one can buy the stock at K and sell it at a higher price of $u^j d^{n-j} S$ to make a profit of $u^j d^{n-j} S - K$. In view of this, market participants would place a value of $u^j d^{n-j} S - K$ on the option, in particular,

$$C = max[u^j d^{n-j} S - K, 0]. \tag{1.4.4}$$

Using the assumption that there are to be no arbitrage opportunities and the fact that the call must finish in-the-money, the value of a call option is given by

$$C = \frac{\Sigma_{j=a}^{n} (\frac{n!}{j!(n-j)!}) p^j (1-p)^{n-j} [u^j d^{n-j} S - K]}{r^n}, \tag{1.4.5}$$

where:

C = the current value of the call,

S = the current stock price,

K = the exercise price,

n = the number of periods remaining to expiration,

r = the one plus the riskless rate of interest,

a = the minimum number of upward movements required for the option to finish in-the-money while p and $1-p$ are defined as follows: $p \equiv \frac{r-d}{u-d}$ and $1-p \equiv \frac{u-r}{u-d}$. $\frac{n!}{j!(n-j)!}$ represents the number of paths the stock can take to reach a certain point in a binomial tree (see Cox et al (1979)).

A relationship between the underlying asset and its options, called **put-call-parity**, is used to find the corresponding value of a put option. If P is the value of a put option, then using the put-call-parity

$$S + P - C = Ke^{(T-t)}, \tag{1.4.6}$$

the value of the European put option is found to be

$$P = \frac{\Sigma_{j=a}^{n}(\frac{n!}{j!(n-j)!})p^j(1-p)^{n-j}[K - u^j d^{n-j} S]}{r^n} \tag{1.4.7}$$

1.4.3.2. *The Black-Scholes Model.* Most texts present a derivation of the Black-Scholes formula for calculating options by using the concept of arbitrage and the lognormal model of asset price movements (see Willmott et al (1995), Joshi 2004, Hull (1989) and Ross (1999)). The assumptions used and eventual derivation of the model using stochastic differential equations can be found in Willmott et al (1995) and Hull (1989).

However, other authors derive the Black-Scholes formula as a limiting case of the Binomial Option pricing formula in equation (1.4.5) (see Cox et al (1979)). Equation (1.4.5) can be rewritten as

$$C = S[\frac{\Sigma_{j=a}^{n}(\frac{n!}{j!(n-j)!})p^j(1-p)^{n-j}u^j d^{n-j}}{r^n}] - Kr^{-n}[\Sigma_{j=a}^{n}(\frac{n!}{j!(n-j)!})p^j(1-p)^{n-j}] \tag{1.4.8}$$

Let the terms in the closed parenthesis be B_1 and B_2 and noting that

$$\frac{p^j(1-p)^{n-j}u^j d^{n-j}}{r^n} = [(\frac{u}{r})p]^j[(\frac{d}{r})(1-p)]^{n-j} \tag{1.4.9}$$

This can be written as $p_*^j(1-p_*)^{n-j}$ where $p_* = (\frac{u}{r})p$ and $1 - p_* = (\frac{d}{r})(1-p)$. Then the equation can now be written as

$$C = SB_1 - Kr^{-n}B_2. \tag{1.4.10}$$

By letting the number of periods approach infinity, B_1 and B_2 converge to $\Phi(d_1)$ and $\Phi(d_2)$ respectively. Thus

$$C = S\Phi(d_1) - Kr^{-n}\Phi(d_2), \tag{1.4.11}$$

where

$$d_1 = \frac{\ln(\frac{S}{K}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}$$

and

$$d_2 = \frac{\ln(\frac{S}{K}) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} = d_1 - \sigma\sqrt{T-t}.$$

$\Phi(\cdot)$ is the cumulative probability distribution function for a standardised normal variable, where $C =$ the current value of the call,
$S =$ the current stock price,
$K =$ the exercise price,
$r =$ risk-free interest rate,
$\sigma =$ volatility of stock,
$t =$ current time,
and $T =$ expiration date of the option.

It can also be shown using equation (1.4.6) that the value of the European put option, P, is

$$P = -Kr^{-n}\phi(-d_2) - S\phi(-d_1) \tag{1.4.12}$$

Evaluation of values of American call and put options is thoroughly covered in Hull (1989) and Wilmott et al (1995).

In this thesis, we develop a model that can be used to model price changes using the double gamma probability distribution. With the model, we forecast prices of different stocks and the world's major financial indices. Lastly, we compare the results of this model with results obtained when a conventional log-normal distribution of stock prices is assumed.

1.5. Structure of the thesis

The rest of the thesis is structured as follows. Chapter 2 reviews major concepts in stochastic processes that are cardinal to work related to this thesis. It provides a synopsis of relevant literature in the area of stochastic processes as well as a highlight of some asset pricing models. Most of the work in this chapter can be found in Bhat (1984), Medhi (1982), Ross (1999), Kao (1997), Joshi (2004) and other relevant research articles. In Chapter 3, the research methodology used and the proposed model are presented. In Chapter 4, results are presented, analysed and conclusions are made.

CHAPTER 2

Review of Some Stochastic Processes

In this chapter literature in areas relevant to the study is reviewed. In particular stochastic processes and some stock price models are reviewed at length. Parameter estimation is also reviewed.

2.1. Stochastic Processes

Many phenomena may be observed as random realisations over time. This is true in finance. Hence the study of collections of random observations over time called *stochastic processes* is very crucial. In this chapter we review some stochastic processes that have been used in modeling finance dynamics.

Definition 2.1.1. A *stochastic process* is a family of random variables $\{X(t) : t \geq 0\}$ where $t \in T$.

The values $X(t)$ assumed by the process are called *states* while the set of all possible values is called the *state space*. On the other hand the set of all possible values of the indexing parameter is called the *parameter space* or *index set* T. The index t is often viewed as a time parameter while the index set T is viewed as the set of all possible time points. A typical example is the price of a stock in a financial market at time t, say $S(t)$. The states would be the values $S(t)$ assumes at any time t. When the index set T is countable the process $\{X(t)\}$ is said to be a *discrete-time stochastic process* while when the process is defined at every instant over a finite or infinite interval, then $\{X(t)\}$ is said to be a *continuous-time stochastic process*.

Every empirical data or stochastic process has a theoretical probability distribution behind it. There are various techniques of modeling an unknown probability density through parameter estimation. If the model turns out to be a good fit, the properties of the stochastic process can be approximated by the known properties of the distribution. In a similar way, if a real-life process, such as a stock price process, is observed to have the attributes of some stochastic process, then the behaviour of the real stock prices can be easily modeled (cf Bhat (1984)).

In this thesis we focus on stock prices $\{S(t)\}$ as our stochastic process. An analytical study of historical stock data can be used to estimate its essential characteristics. Where it is not possible to explicitly deduce the model of the stochastic behaviour analytically due to its complexity, provided there is a starting point, the derivation of an estimate of a model may be obtained through simulation techniques. This is achieved by mimicking the process several times and averaging the sample characteristics so obtained. These techniques are reviewed in greater detail in subsection (2.2.3).

In the next section, we briefly review some concepts that are central to the thesis.

2.1.1. Distribution of Stochastic Processes. For any stochastic process $\{X(t)\}$ it is customary to attempt to fit a probability distribution in order to understand the characteristics of the process. Although a stochastic process $\{X(t) : t \in T\}$ has a corresponding probability distribution, in practice the specific information on the process $\{X(t) : t \in T\}$ may not be easily described by a simple distribution (see Bhat (1984)). The common approach is to define a joint distribution by studying the process at discrete time points.

Let $(t_1, t_2, \cdots, t_n)$, where $t_1 < t_2 < \cdots < t_n$, be a set of discrete time points. Then the joint distribution for the process $X(t)$ at these points is defined as

$$P[X(t_1) \leq x_1, X(t_2) \leq x_2, \cdots, X(t_n) \leq x_n] \tag{2.1.1}$$

This distribution assumes its simplest form when the random variables are independent as it is thus given as the product of individual marginal distributions. However, in most practical situations the processes are more complex because of the existence of dependencies among the random variables. Although it is desirable to have a joint distribution of the form (2.1.1) some conditional probability distribution functions, called *transition probabilities*, are defined based on some information of the stochastic process available for any specific value of the time parameter.

Let t_0 and t_1 be two points in the index set T such that $t_0 \leq t_1$. Then the conditional transitional function may be written as

$$F(x_0, x_1; t_0, t_1) = P[X(t_1) \leq x_1 | X(t_0) = x_0] \tag{2.1.2}$$

For a stochastic process with discrete parameter and state spaces, the transition probabilities are defined as

$$P_{ij}^{(m,n)} = P(X_n = j | X_m = i), n \geq m \tag{2.1.3}$$

We shall revisit these in subsubsection (2.1.2.1) under some general properties.

In most real-life situations stochastic processes exhibit some form of dependence (Bhat (1984) and Medhi (1982)). Hence stochastic processes may be broadly described according to the nature of dependence relationship existing among members of the family.

DEFINITION 2.1.2. If for all $t_1, t_2, \cdots, t_n \in T$ and $t_1 < t_2 < \cdots < t_n$, $X(t_2) - X(t_1), X(t_3) - X(t_2), \cdots, X(t_n) - X(t_{n-1})$ are independent, then $\{X(t) : t \in T\}$ is said to be a *process with independent increments*.

This implies that in a process with independent increments the magnitudes of state change over non-overlapping intervals are mutually independent (Kao (1997)). A related property is the stationary increment property.

DEFINITION 2.1.3. A stochastic process $\{X(t) : t \in T\}$ is said to possess the *stationary increment property* if the random variable $X(t+s) - X(t)$ possesses the same probability distribution for all t and any $s > 0$.

This implies that the probability distribution of the magnitude of state change depends only on the difference in the lengths of the time indices and is independent of the time origin used for indexing parameter (Kao (1997) and Ross (1996)). Various authors have assumed these properties for changes in stock prices.

2.1.2. Some Common Stochastic Processes. In this section we review some common stochastic processes that are encountered in financial mathematics. We shall further explore some important properties of such processes, most specifically we review processes with discrete time and parameter spaces. The bulk of the content in this section can be found in Ross (1996), Moran (1968), Kao (1997), Bhat (1984) and Medhi (1982).

2.1.2.1. *Markov Chains.* An asset such as a stock traded on the stock market either increases, decreases or does not change in price each time the market opens and closes. In this way the stock can, apart from being regarded as a physical system with three possible states, be viewed as a stochastic process. Stock prices are assumed to have a special property which we define below.

DEFINITION 2.1.4. The stochastic process $\{X_n, n = 0, 1, 2, \cdots\}$ is a *discrete-time Markov chain*, if, for all $j, i, j_1, j_2, ..., j_{n-1} \in \mathcal{N}$,

$$P[x_n = j|X_{n-1} = i, X_{n-2} = j_1, ., X_0 = j_{n-1}] \;\; = \;\; P[X_n = j|X_{n-1} = i] = P_{ij}$$

For the continuous-time Markov chain, we adopt the definition given by Ross (1996).

DEFINITION 2.1.5. The stochastic process $\{X(t), t \geq 0\}$ is a *continuous-time Markov chain* if for all s, $t \geq 0$ and nonnegative integers i, j, $x(u)$, $0 \leq u \leq s$,

$$P[X(t+s) = j|X(s) = i, X(u) = x(u), 0 \leq u \leq s] \;\; = \;\; P[X(t+s) = j|X(s) = i]$$

The j values in the two definitions are referred to as *states* of the Markov chain. Thus if X_n has outcome j, the process is said to be at state j at the *nth* trial. The probability P_{ij}, called *one-step transition probability*, represents the probability that the process will make a transition into state j given that it was previously in state i. The transition probabilities share similar properties as those of ordinary probabilities such that $P_{ij} \geq 0$, $i, j \geq 0$, $\sum_{j=0}^{\infty} P_{ij} = 1$, $i = 0, 1, \cdots\cdots, \infty$.

Markov chains are widely used in the modeling of problems in many application areas of economic systems. The Markov chains are classified in accordance with some fundamental properties of the states of the system (see subsubsection 2.1.2.1).

A Markov property may alternatively be interpreted as stating that the conditional distribution of any future state X_{n+1} given the past states $X_0, X_1, \cdots, \cdots, X_{n-1}$ and the present state X_n is independent of the past states and depends only on the present state (Ross (1996)). Stock prices are assumed to follow the Markov process because of the weak form of market efficiency which states that the present price reflects all the information of previous prices (Hull (1989)). This implies that only the present state is relevant for predicting the future, hence if S_t is the price of stock at time t, then

$$P[S_{t+1} = s|S_t = s_t, \ldots, S_1 = s_1, S_0 = s_0] = P[S_{t+1} = s|S_t = s_t].$$

The transition probability P_{ij} is referred to as a *one-step transition probability* while the *n-step transition probability*, denoted P_{ij}^n, is defined as

$$P_{ij}^n = P[X_{n+m} = j \mid X_m = i, \quad n \geq 0, \quad i,j \geq 0].$$

This represents the probability that a process in state i will be in state j after n additional transitions. The one-step transition probability shall be written P_{ij} instead of P_{ij}^1. In order to compute n-step transition probabilities, the *Chapman-Kolmogorov Equations* defined below are normally used

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m, \quad \forall n,m \geq 0, \quad \forall i,j.$$

(cf. Ross (1996)).

The transition probabilities are usually presented in matrix form as

$$\underline{P} \quad = \quad \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1i} & \cdots & p_{1n} \\ p_{21} & p_{22} & p_{23} & cdots & p_{2i} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \cdots & p_{ni} & \cdots & p_{nn} \end{pmatrix}.$$

While the Chapman-Kolmogorov equations can be used in the computation of transition probabilities, an illuminating approach is to undertake a classification of the states. The next subsection highlights the various classes in which states of a Markov chain may fall.

Classification of States

The states of any Markov chain can be classified according to certain basic properties. The classification is based on definitions given in this subsection.

DEFINITION 2.1.6. State j is said to be *accessible* from state i if j can be reached from i in a finite number of steps. If two states i and j are accessible to each other, then they are said to *communicate*. We denote this by $i \leftrightarrow j$.

It can be shown that communication is an equivalence class (cf. Ross (1996)). In this way, two states that communicate are said to be in the same *class*. Further a Markov chain is said to be *irreducible* if there is only one class, in other words, if all states communicate. Mathematically, the following properties of the communication relation hold (cf. Bhat (1984), Ross (1996)):

(i) $i \leftrightarrow i$ (Reflexivity)

(ii) if $i \leftrightarrow j$, then $j \leftrightarrow i$ (Symmetry)

(iii) if $i \leftrightarrow j$ and $j \leftrightarrow k$, then $i \leftrightarrow k$ (Transitivity).

While the equivalence classification of states takes into account the external relationship between the states, another closely related classification takes into account the internal nature of each state. This form of classification is considered through the following definitions (cf. Bhat (1984), Ross (1996)).

We denote the probability that, starting from i, the process moves to state j for the first time in the $n-th$ step by f_{ij}^n. More formally, let

(2.1.4) $$f_{ij}^n = P[X_n = j, X_k \neq j, k = 1, 2, \cdots, n-1 \mid X_0 = i]$$

and

(2.1.5) $$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^n$$

With this notation, we give the following definitions.

Definition 2.1.7. A state i is said to be *recurrent* if and only if, starting from state i eventual return to state i is certain.

In terms of the probabilities given by (2.1.5) this implies that the state i is *recurrent* if and only if $f_{ii} = 1$. However, it is possible that the process may not return to the state it originally started from. This situation solicits another classification as given in the following definition by Bhat (1984).

Definition 2.1.8. A state is said to be *transient* if and only if, starting from state i, there is a positive probability that the process may not eventually return to state i, that is, $f_{ii} < 1$.

At times it is important to consider the number of moves required for the process to return to a specified state. Accordingly, let μ_{ii} denote the expected number of transitions needed to return to state i assuming the process started from state i. When state i is recurrent, the mathematical expectation of the number of transitions required for the first return to state i in n steps is given by

$$\mu_{ii} = \sum_{i=1}^{\infty} n {f_{ii}}^n$$

The number of transitions required for the first return to the same state is called the *recurrence time* and consequently μ_{ii} is called the *mean recurrence time* of state i. In view of the definition of μ_{ii}, a recurrent state can be further classified as null recurrent or positive recurrent.

DEFINITION 2.1.9. A recurrent state i is said to be *null recurrent* if and only if $\mu_{ii} = \infty$. A recurrent state is said to be *positive recurrent* if and only if $\mu_{ii} < \infty$, that is, the mean recurrence time is finite.

The states of a Markov chain can also be classified as transient or recurrent using transition probabilities P_{ii}^n, the probability that the process occupies state i after n steps given that it was initially in state i. This is different from f_{ii}^n which refers to the probability of the first return to state i in n steps. This classification is captured in the following proposition and corollary found in Ross (1996) with some slight modification adapted from Bhat (1984).

PROPOSITION 2.1.10. *A state i is*

(i) *recurrent if and only if* $\sum_{n=0}^{\infty} P_{ii}^n = \infty$.
(ii) *transient if and only if* $\sum_{n=0}^{\infty} P_{ii}^n < \infty$

COROLLARY 2.1.11. *If i is recurrent and $i \leftrightarrow j$, then j is also recurrent.*

The proofs of the proposition and the corollary can be found in Ross (1996). The corollary shows that recurrence is a class property. Further, since all states in an equivalence class communicate, they are all either recurrent or transient. This implies that the class of states as a whole can therefore be considered as being either recurrent or transient.

The description of recurrent states given in proposition 2.1.10 provides another way to characterise a recurrent state.

DEFINITION 2.1.12. Let $d(i)$ denote the greatest common divisor of all integers $n \geq 1$ for which $P_{ii}^n > 0$. The integer $d(i)$ is called the *period* of state i. When the period is 1, the state is called *aperiodic*.

Further classification of states can be discerned from the definitions that follow.

DEFINITION 2.1.13. A state i is said to be an *absorbing state* if and only if $P_{ii} = 1$.

The definition implies that once the process enters state i it remains in that state. Thus when state i is absorbing $f_{ii}^1 = P_{ii} = 1$ and hence $\mu_i = 1$ which shows that i is positive recurrent. In this way, a Markov chain may also be classified as *absorbing* if it has at least one absorbing state, and if from every state it is possible to go to an absorbing state (not necessarily in one step). On the other hand, in an absorbing Markov chain, a state which is not absorbing is called *transient*.

In other types of Markov chains it is possible that all states belong to the same equivalence class. Since communication is an equivalence relation, any two classes may either be disjoint or the same. If no states outside of an equivalence class can be reached from any state within the class, the class is said to be *closed*.

DEFINITION 2.1.14. A Markov chain is *irreducible* if its only closed class is the set of states in its state space S.

Thus all the states of an irreducible Markov chain belong to one equivalence class.

There are many situations that are modeled as Markov chains in particular and stochastic processes in general. We briefly review some examples of stochastic processes that possess the Markov property and hence modeled as Markov chains.

2.1.2.2. *Martingales.* Although the origin of martingales lies in the history of games of chance, they are powerful tools for analysing a variety of stochastic processes. We adopt the definition given by Ross (1996) with a slight modification on notation.

DEFINITION 2.1.15. A stochastic process $\{S_t, t \geq 1\}$ is said to be a *martingale* process if $E[|S_t|] < \infty$ for all t and $E[S_{t+1} \mid S_t, S_{t-1}, \cdots, S_2, S_1] = S_t$.

On the other hand if $E[S_{t+1} \mid S_t] \geq S_t$ for all t it is called a *submartingale* while if $E[S_{t+1} \mid S_t] \leq S_t$ it is called *supermartingale*. A similar definition of martingales as a continuous time process can be found in Neftci (1996).

The stochastic process $\{S_t, t \geq 1\}$ could be the price process of a security whose price at any time t is S_t. The martingale property implies that the best forecast of unobserved future values is the last observation on S_t (Neftci (1996)). Thus at any one time the current price fully represents all the information. In this sense, efficient markets are equated to the existence of a martingale.

The definition also implies that future movements in martingales are impossible to forecast (Neftci (1996)). If S_t is a martingale and consider the forecast change in S_t over a time interval length $\Delta > 0$, then

$$E[S_{t+\Delta} - S_t] = E[S_{t+\Delta}] - E[S_t].$$

Since S_t is a martingale and $E[S_t]$ is a forecast of the martingale that is already revealed, then

$$E[S_{t+\Delta} - S_t] = 0.$$

Thus a fundamental characteristic of martingales is the impossibility to forecast their future movements. However, stock prices are not completely unpredictable and hence are generally not martingales (Neftci (1996)). Although most financial assets are not martingales, they can be converted into martingales. The advantage of this is that properties of martingales can be used to analyse financial data. For instance, a probability density P' can be identified such that a financial asset such as stock discounted by the risk-free rate r become martingales. Then an equality such as

$$E^{P'}[e^{-r\Delta} S_{t+\Delta}] = S_t$$

for all $\Delta > 0$ can be used in pricing derivative securities. Methods for converting other processes into martingales can be found in Neftci (1996).

2.1.2.3. *Birth and Death Processes.* The birth-death process is a special case of continuous-time Markov process where the states represent the current size of a population and where the transitions are limited to states 'increase' and 'decrease'. Although death and birth processes are more relevant in problems related to populations, it is also used in economics. Birth and death process models are used in inventory systems if replenishment of stock is accompanied only by placing orders. In such a system, if demands for items occur in a Poisson process, then the inventory in between replenishments can be modeled as a pure death process (Bhat (1984)).

Definition 2.1.16. A *birth and death process* $\{X(t), t \geq 0\}$ is a continuous-time discrete space (with state-space N) Markov process such that

(a) $P[X(t+h) = n+1|X(t)] = n = \lambda_n h + o(h)$, for $n > 0$
(b) $P[X(t+h) = n-1|X(t)] = n = \mu_n h + o(h)$, for $n \geq 1$

(c) $P[X(t+h) = n|X(t)] = n = 1 - (\lambda_n + \mu_n)h + o(h)$, for $n > 0$,

where any function $f(\cdot)$ is said to be $o(h)$ if $\lim_{h \to 0} \dfrac{f(h)}{h} = 0$.

It can be seen from the definition that three types of transitions are possible: one birth, or one death, or no birth nor death. In other words, three states can be distinguished and may be interpreted as states 'increase', 'decrease' or 'no change'. The state of the process is usually interpreted as the size of the population (Ross (1996) and Kao (1997)). When the state increases by 1, it is said that a birth has occurred while when it decreases by 1 it is said that a death has occurred. In other words, when a birth occurs, the process goes from state n to $n+1$. When a death occurs, the process goes from state n to state $n-1$. When neither death nor birth occurs, the process remains in the same state n.

The birth and death process is specified by birth rates $\{\lambda_n\}_{n=0,1,\cdots\infty}$ and death rates $\{\mu_n\}_{n=1..\infty\}}$. If $\lambda_n = 0$ for all n, the process is said to be a *pure death* process while if $\mu_n = 0$ for all n the process is said to be a *pure birth* process.

An example of a pure birth is a *Yule* process. This is the case where in a population each member acts independently and gives birth at an exponential rate λ. If no single member of the population ever dies, then, if $X(t)$ represents the population at time t, the process $\{X(t), t \geq 0\}$ is a pure birth process with $\lambda_n = n\lambda$.

A detailed treatment of birth and death processes can be found in Ross (1996), Kao (1997), Medhi (1982) and Bhat (1984).

2.1.2.4. *Counting Processes.* In the study of a number of phenomena, it may be useful to consider the number of occurrences during a period of time or space. For instance, one may be interested in the number of times the stock price has gone up or down in any given period. This is used in option pricing using the Binomial option pricing model reviewed in subsubsection 1.4.5. Such type of stochastic process is called a *counting process.* We adopt the definition given in Ross (1996).

Definition 2.1.17. A stochastic process $\{N(t), t \geq 0\}$ is said to be a *counting process* if $N(t)$ represents the total number of events that have occurred up to time t.

In view of the definition, a counting process $N(t)$ must satisfy the following:

(i) $N(t) \geq 0$.

(ii) $N(t)$ is integer valued.

(iii) If $s < t$, then $N(s) \leq N(t)$.

(iv) For $s < t$, $N(t) - N(s)$ equals the number of events that have occurred in the interval $(s, t]$.

Counting processes that possess independent increments and stationary increments are a special type of a well known stochastic process called the *Poisson process.* Using the four requirements that a counting process must satisfy as outlined above, the following theorem is cited.

THEOREM 2.1.18. *Let* $\{N(t), t \geq 0\}$ *be a counting process with independent increments such that* $N(0) = 0$. *Then there exists a constant* $\lambda > 0$ *such that the transition probability distribution of the stochastic process* $\{N(t), t \geq 0\}$ *has a Poisson distribution given by*

$$P[N(t+s) - N(s) = n] = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \cdots,$$

where s, $t \geq 0$.

The proof of this theorem can be found in Medhi (1982), Bhat (1984), Kao (1997) and Ross (1996).

The expected number of events, called the *rate* of the process, that have occurred up to time t can be found as $E[N(t)] = \lambda t$. We use this theorem and associated results in our proposed model in chapter 3.

In the next three subsubsections, we review two stochastic processes in continuous time which are widely used in finance and other fields. These fall broadly under the notion of Brownian motion.

2.1.2.5. *Geometric Brownian Motion.*

DEFINITION 2.1.19. Suppose $P(t)$ is the price of a security at a time t from the present. The set of prices $\{P(t)\}$, with $0 \leq t < \infty$, is said to follow a *Geometric Brownian motion* with drift parameter μ and volatility parameter σ if for all $t \geq 0$ and $s \geq 0$, the random variable $\dfrac{P(t+s)}{P(t)}$ is independent of all prices up to time t and $log\dfrac{P(t+s)}{P(t)}$ is a normal random variable with mean μs and variance $s\sigma^2$.

The geometric brownian motion implies that it is only the present price, not past history of prices, that affects the movements of future prices. In addition, probabilities of the ratio of the price $P(t)$ at a future time t to the present price $P(t_0)$ will not depend on the present price $P(t_0)$. A more thorough coverage of this concept is given in Ross (1996).

2.1.2.6. *Brownian Motion.* In this section the basic definition of the Brownian motion is given and is related to stock prices. We also briefly trace its origins and note its similarities to the geometric Brownian motion. A thorough introduction to the concept of Brownian motion can be found in Ross (1999), Willmott et al (1995), Joshi (1989), Neftci (1996) and Kao (1997).

Definition 2.1.20. Let $\{P(t)\}$ be a set of prices for $0 \leq t < \infty$. The set of prices $\{P(t)\}$ is said to follow a *Brownian motion* with drift parameter μ and variance parameter σ^2 if for all $t \geq 0$ and $s \geq 0$, the random variable $P(t+s) - P(t)$ is independent of all prices up to time t and is a normal random variable with mean μs and variance $s\sigma^2$.

In 1827 the Brownian motion was used to describe the unusual motion displayed by a small article totally immersed in a liquid or gas. Later in 1925, Albert Einstein showed mathematically that Brownian Motion could be explained by assuming that the immersed particle was continually being bombarded by the smaller particles surrounding it. It was, however, independently introduced in 1900 by Louis Bachelier to model price movements of stocks and commodities (Ross (1996)).

Stock price movements seem to display behaviour similar to Brownian motion. The immersed particle may be viewed as the stock price and the smaller particles as the trades that move the stock price. Each trade moves the price up or down and each trade is independent from other trades.

It is worthwhile noting that the geometric Brownian motion and the Brownian motion share the property that the price at a future date depends only on the present price. The only difference between the two concepts is that in the Brownian motion it is the difference in prices that has a normal distribution whereas in the geometric Brownian motion it is the logarithm of their ratio that has a normal distribution.

2.1.2.7. *The Wiener Process.* The Wiener process is similar to the Brownian motion process with the exception that the Wiener process is appropriate for continuous stochastic processes. We paraphrase the definition found in Medhi (1982).

Definition 2.1.21. A stochastic process $\{W(t), t \geq 0\}$ is said to be a *Wiener process* if it satisfies the following conditions:

(a) $\{W(t), t \geq 0\}$ has stationary independent increments.

(b) Every increment $W(t) - W(s)$ is normally distributed with mean $\mu(t - s)$ and variance $\sigma^2(t - s)$.

The first part of the definition implies that the Wiener process is a Markov process with independent increments while the second part implies that a Wiener process is Gaussian. A Wiener process in which $W(0) = 0$, $\mu = 0$ and $\sigma = 1$ is called a *standard Wiener process.* The Wiener process has many applications. It is used to model the movement of particles immersed in a liquid or gas in quantum mechanics. In finance, the Wiener process is used to model price fluctuations in stock and commodity markets. A detailed treatment of the Wiener process can be found in Bhat (1984), Neftci (1986) and Willmott (1995).

As has been mentioned in section (2.1.1), to understand the characteristics of a stochastic process it is desirable to fit a probability distribution. The characteristics of such probability distributions need to be estimated. One such procedure for estimation is the method of maximum likelihood and this is the object of discussion in the next section.

2.2. Review of Parameter Estimation

The study of properties of stochastic processes is very crucial in finance. Historical data is usually used to estimate the essential characteristics of stochastic processes (see subsection 2.1.1). In some cases, a probability distribution may be fitted. The characterstics or parameters of such probability distributions need to be estimated. One such procedure for estimation is maximum likelihood estimation. In this chapter, we review maximum likelihood estimation and related concepts.

2.2.1. Maximum Likelihood Estimation. Given a data set taken from any population, it is customary to inquire about the characteristics of the population from which it was taken. One way of making such inferences is to assume some kind of probabilistic model that would describe the population. But since the parameters of such a model are not known, statistical inferences of the population would not be easy without employing estimation. One single most popular method of estimation is Maximum Likelihood Estimation (MLE).

DEFINITION 2.2.1. Let $\{X_1, X_2, X_3, \cdots, X_n\}$ be a set of random variables with a joint density function $f(x_1, x_2, \cdots, x_n)$. Given observed values $X_i = x_i$ for $i = 1, 2, \cdots, n$, the *likelihood* of a function of $x_1, x_2, \cdots, x_n$ is defined as

$$lik(\vartheta) = f(x_1, x_2, \cdots, x_n|\vartheta).$$

The likelihood function gives the probability of observing the given data as a function of the parameter. The aim of maximum likelihood estimation is to find the parameter value(s) that makes the observed data most likely. Instead of maximising $lik(\vartheta)$ it is easier to maximise its logarithm $l(\vartheta) = \Sigma log[f(x_i|\vartheta)]$ since the logarithm of a product of variables simplifies into the sum of logarithm of the individual variables. Then the maximum likelihood estimators of $\vartheta_1, \vartheta_2, \cdots, \vartheta_n$ are solutions to the simultaneous equations given by $\dfrac{\partial(l)}{\partial\vartheta_i} = 0$, where $i = 1, 2, \cdots, n$.

We briefly review selected examples of maximum likelihood estimators for some common probability distributions. Among the distributions, we highlight the gamma distribution which is central to this study.

Example 2.2.2. **Normal Distribution**

Under the Normal distribution, there are two parameters that would be estimated. These are mean μ and standard deviation σ. Since the natural logarithm of the likelihood function is

$$l(\mu, \sigma) = -n\text{ln}\sigma - \frac{n}{2}\text{ln}2\pi - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2,$$

then using partial differentiation it can be shown that the maximum likelihood estimators for μ and σ are $\mu = \overline{x}$ and $\sigma = \sqrt{\dfrac{1}{n}\Sigma(x_i - \overline{x})^2}$ respectively.

Example 2.2.3. **Poisson Distribution**

Suppose a random variable follows a Poisson distribution with parameter λ, then

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}.$$

If $x_1, x_2, \cdots, x_n$ are independently and identically distributed and Poisson, then the logarithm of the corresponding likelihood function is

$$l(\lambda) = \text{ln}\lambda\sum_{i=1}^{n} x_i - n\lambda - \sum_{i=1}^{n}\text{ln}x_i!.$$

Upon differentiating partially and solving the equation $\dfrac{\partial(l)}{\partial\lambda} = 0$, the maximum likelihood estimator obtained is $\widehat{\lambda} = \overline{x}$.

Example 2.2.4. **Gamma Distribution**
Let $x_1, x_2, \cdots, x_n$ be a random sample taken from a gamma distribution with parameters α and β, then its density function is given by

$$f(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)}\beta^{\alpha}x^{\alpha-1}e^{-\beta x},$$

where $0 < x < \infty$. The parameter α is called a *shape parameter* for the gamma function and β is called a *scale parameter*.

Although the logarithm of an independently and identically distributed sample that follows the gamma distribution is

$$l(\alpha,\beta) = n\alpha\ln\beta + (\alpha-1)\sum_{i=1}^{n}\ln x_i - \beta\sum_{i=1}^{n}x_i - n\ln\Gamma\alpha,$$

it is impossible to obtain maximum likelihood estimators using the procedure used in the above examples. Instead numerical methods may be used. However, using the method of moments, the maximum likelihood estimators are found to be

$$\widehat{\alpha} = \frac{\overline{x}^2}{s^2} \quad and \quad \widehat{\beta} = \frac{s^2}{\overline{x}}$$

where s^2 is sample variance while $\overline{x}$ is sample mean.

In our model we use these two maximum likelihood estimators. A more rigorous treatment of maximum likelihood estimation is provided in Myung (2003), Hogg et al (1978) and Rice (1988).

Example 2.2.5. **Double Gamma Distribution** The general form of the double gamma distribution has the following probability density function:

$$f(x) = \frac{1}{2}\frac{\left(\frac{|x-\mu|}{\beta}\right)^{(\alpha-1)}e^{-\left(\frac{|x-\mu|}{\beta}\right)}}{\beta\Gamma(\alpha)}$$

where μ and β are the positive location and scale parameters respectively. The distribution is also referred to as the *reflected gamma distribution*. Our proposed model is based on the double gamma distribution and its model parameters shall be estimated using the maximum likelihood estimators given above.

When $\mu = 0$ and $\beta = 1$ then the distribution is referred to as the standard form of the gamma distribution given by the following probability density function

$$f(x) = \frac{1}{2}\frac{|x|^{\alpha-1}e^{-|x|}}{\Gamma(\alpha)}$$

where α is a positive number that is the shape parameter and Γ is the standard gamma function.

The following are graphs of the double gamma probability density function for different values of the shape parameter.

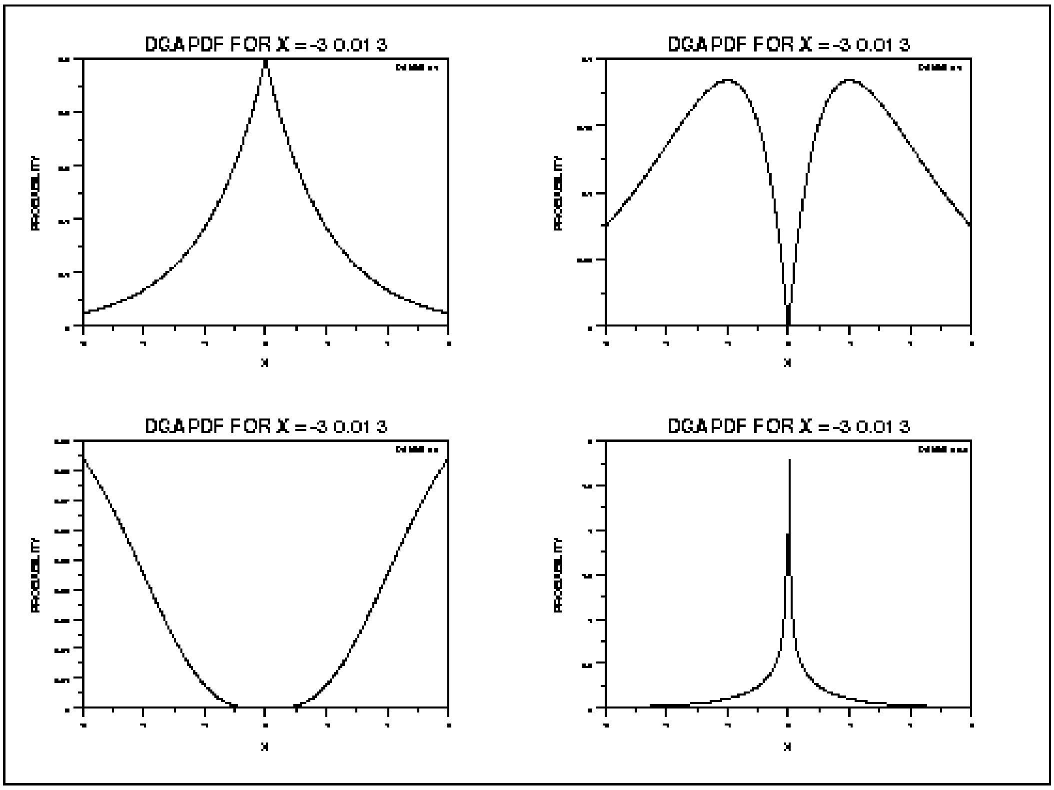

FIGURE 2.1. Graphs of the double gamma probability density function with shape parameters 1, 2, 5 and 0.5

From left to right, the top row exhibits the graphs of the double gamma probability density function with shape parameters 1 and 2 respectively while the bottom row displays graphs corresponding to shape parameters 5 and 0.5 respectively.

In addition to maximum likelihood estimation, other estimation methods may be used to obtain estimators such as the sample mean. However, some estimators may not be very good estimators. For the sample mean, we appeal to the Law of Large Numbers. This is reviewed in the next section.

2.2.2. The Law of Large Numbers. The use of repeated experiments to model the behaviour of random variables is a standard practice in applied science. After a large number of experiments, it is possible to estimate, for instance, the relative frequency of a random variable. As most commonly computed statistics, such as averages, are expressed in terms of sums it is desirable to consider taking limits. We take advantage of two laws, the Weak Law of Large Numbers (WWLN) and the Strong Law of Large Numbers (SLLW), that we use in our thesis and we state them here. We adopt versions found in Ross (1984) but we modify them slightly.

THEOREM 2.2.6. *The Strong Law of Large Numbers*
Let $X_1, X_2, \cdots\cdots$ be a sequence of independent and identically distributed random variables, each having a finite mean $\mu = E[X_i]$. Then,

$$\mathrm{P}\left(\lim_{n\to\infty} \overline{X}_n = \mu\right) = 1$$

.

The Strong Law of Large Numbers asserts that the sample mean $\overline{X}$ converges to the population mean μ with probability 1. This justifies the use of $\overline{X}$ as an estimator for μ provided the sample is large.

THEOREM 2.2.7. *The Weak Law of Large Numbers*
Let $X_1, X_2, \cdots\cdots$ be a sequence of independent and identically distributed random variables, each having a finite mean $\mu = E[X_i]$, Then, for any $\varepsilon \geq 0$,

$$\lim_{n\to\infty} \mathrm{P}\left(\left|\overline{X}_n - \mu\right| < \varepsilon\right) = 1.$$

In both cases, $\overline{X}_n = (X_1 + \cdots + X_n)/n$. The proof of the Law of Large Numbers can be found in Rice (1988) while that of the Weak Law of Large Numbers is found in Ross (1984).

It is worth pointing out that these laws mean the same thing. The only difference is in the manner in which the sample mean converges to the population mean. The weak law states that as the sample size grows larger, the difference between the sample mean and the population mean will approach zero. The strong law states that as the sample size grows larger, the probability that the sample mean and the population mean will be exactly equal approaches 1.0. In essence, both laws imply that the sample mean $\overline{X}$ is increasingly likely to be close to the population mean μ as $n \to \infty$. This justifies the use of the sample mean $\overline{X}$ as an estimator for population mean μ.

While maximum likelihood estimation is used to estimate the basic statistics of a system, it also helpful to model the system through repeated experimentation. This is the focus of the next section.

2.2.3. Simulation. Simulation is one of the most widely used statistical approaches used to model the operation of a real system. It involves learning about a real system using a model that represents the real system. The model comprises mathematical expressions and logical relationships that are used to evaluate outputs for given values of inputs. The values obtained using the model are then compared with the real system or situation.

Basically any simulation model has two inputs called *controllable inputs* and *probabilistic inputs*. The values for the controllable inputs are selected by the analyst while the values for the probabilistic inputs are randomly generated by a computer. The model uses values of the controllable inputs and values of the probabilistic inputs to generate a value or values of the output. Data obtained from results of a series of similar experiments using a variety of values of the controllable inputs is analysed and reviewed. The analysis and review enables the analyst to make adjustments to the controllable inputs so that a desired result of the real system can be obtained. This procedure of modeling the real situation through repeated experimentation under the same conditions is known as *simulation*.

In this thesis, our random variable is the stock price S_t which we simulate as a Gamma random variable. In Chapter 3 we explain in greater detail how we carried out the simulation to obtain our results. Explicit examples of simulations of normal, exponential, gamma, binomial, geometric and Poisson random variables are given in Ross (1984).

2.3. Stock Price Models

The determination of future values of assets such as stocks is becoming increasingly relevant to investors as well as consumers and producers. Price forecasting is used for developing trading strategies and negotiation skills to maximise benefit. In finance, the underlying asset's price is used in the evaluation of risk and pricing of derivative assets (see section 1.4).

In this section we review some of the work that has been previously done in asset pricing in general and stock pricing in particular. This is intended to link work in this thesis and previous work done by other researchers.

There are different types of models that have been suggested for modeling financial data in general and stock prices in particular. These models can be classified into two basic categories: discrete time models and continuous time models. Most financial time series are observable at fixed discrete time points. For instance, indices and stocks trading on the world's financial markets consist of daily opening and closing prices. The discrete nature of such types of financial data is one of the motivations behind discrete time models. Thus the models that we review in this section belong to either one of the two categories and where necessary we specify the category that a particular model belongs to.

2.3.1. The Bachelier Model and Samuelson Model. Efforts to develop a mathematical model for stock price behaviour can be traced back to two centuries ago. In 1827, Robert Brown, while studying the random motion of a pollen on the surface of water, introduced the notion of Brownian motion. Bachelier seems to be the first to develop a mathematical theory of Brownian motion and used it to value stock options on the Paris stock market (Straja (2006)). The price S_t of stock at any time t in Bachelier's additive model takes the form, in modern technology,

$$S_t = S_0 + \mu t + \sigma B_t, t \geq 0 \tag{2.3.1}$$

where S_0 is the price at time $t = 0$, μ and σ represent drift and volatility respectively while B_t is a standard Brownian process (Shepp (2000)). In differential form equation 2.3.1 may be written as

$$dS_t = \mu dt + \sigma dB_t \tag{2.3.2}$$

The Bachelier model is premised on the assumption that the logarithm of price relatives $L(t,T) = ln[S(t+T) - S(t)]$ has the following four properties: random, statistically independent, identically distributed and that their marginal distribution is Gaussian with mean zero (see Mandelbrot (1963) and Fama (1970)). Due to advances in the theory of speculation Bachelier's four hypotheses have undergone various amendments as shown by Mandelbrot (1963). Fama (1963) also concluded that a better description of distributions of daily returns on common stocks is given by non-normal stable distributions other than the normal distribution.

In his work, Bachelier used the model to value a European option. One major weakness of his model is that prices can be negative. However, since then the theory has been modified by, among others, Samuelson (1965). Samuelson (1965) developed a similar model but in exponential form by replacing Brownian motion with the geometric Brownian motion and gave the pathwise solution

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

of the stochastic differential equation

$$dS_t = S_t \mu dt + S_t \sigma dW_t \tag{2.3.3}$$

where μ and σ are as specified in Equation (2.3.1) while W_t is a Wiener process and $t \geq 0$. The pathwise solution has an added advantage that the price S_t remains positive for all values of t. This model is widely used today, in particular in option pricing and hedging (Follmer and Schweizer (1993)).

In more recent times, there has been a diversification of methods for modeling stock price behaviour though a complete divorce from the Brownian motion concept is impossible. In a recent paper, Rydberg and Shephard (1998) use the compound Poisson process to model asset prices in addition to similar work independently done by Rogers and Zane (1998). Their model is based on the assumption that the non-stationary and non-linear price process follows

$$S_t = S_0 + \sum_{t=1}^{N(t)} Z_t, \tag{2.3.4}$$

$t \geq 0$ and $\{N(t)\}_{t \geq 0}$ is a counting process that counts the number of transactions up to time t while Z_t is the price process associated with the $t - th$ trade (see Rydberg and Shephard (1998)). Norvaisa (2000) uses real analysis to model stock prices. His work is found in Norvaisa (2000).

2.3.2. The Binomial Model. The Binomial Model for stock prices can be described as a tree. At any deterministic time points, the nodes split into two. Thus at the end of any time period the price can be in any two possible states, "up" or "down". The price of stock starts with a value S but moves up to uS with probability p and down to dS with probability $(1 - p)$ over a small interval of length Δt. Then at the end of $n\Delta t$ periods the price of stock will be

$$S_n = u^j d^{n-j} S. \tag{2.3.5}$$

At time n there are n possible values that the stock price can take and, in particular,

$$P[S_n = u^j d^{n-j} S] = \frac{n!}{(n-j)!j!} p^j (1-p)^{(n-j)}, j = 0, 1, 2, \cdots, n, \tag{2.3.6}$$

where: $u = \frac{1}{d}$, $d = e^{-\sigma\sqrt{\Delta t}}$, $p = \frac{e^{r\Delta t} - d}{u - d}$, r = rate of interest during each period and σ is stock volatility (see Rydberg (2000), Hull (1989) and Wilmott etal (1995)).

2.3.3. Lognormal Model for Stock Prices. The lognormal model for stock prices is widely covered in the literature, especially in standard financial mathematics texts. The bulk of the work in this section can be found in Joshi (2004) and Hull (1989). However, Ito's Lemma is adopted from Neftci (1996) and is restated below.

LEMMA 2.3.1. *Let $G(S_t, t)$ be a twice-differentiable function of t and of the random process S_t $dS_t = a_t dt + \sigma_t dW_t$, $t \geq 0$ with well behaved drift and diffusion parameters, a_t and σ_t. Then*

$$dG = \frac{\partial G}{\partial S_t} dS_t + \frac{\partial G}{\partial t} dt + \frac{1}{2} \frac{\partial^2 G}{\partial {S_t}^2} {\sigma_t}^2 dt. \tag{2.3.7}$$

A derivation of Ito's Lemma can be found in Appendix 4A in Hull (1989).

Ito's Lemma is used to derive the process followed by $G = \ln S_t$ to obtain

$$dG = (\mu - \frac{\sigma^2}{2}) dt + \sigma dW. \tag{2.3.8}$$

As μ and σ are constants, then G follows a Wiener process with constant drift rate $\mu - \dfrac{\sigma^2}{2}$ and constant variance rate σ^2. Thus

$$\ln S_T - \ln S_t \sim \phi[(\mu - \frac{\sigma^2}{2})(T - t), \sigma\sqrt{T - t}] \tag{2.3.9}$$

where S_T is the stock price at time T, S_t is the stock price at current time t while $\phi(\rho, \kappa)$ denotes a normal distribution with mean ρ and standard deviation κ for $T \geq t$.

Using properties of the normal distribution, it follows that

$$\ln S_T \sim \phi[\ln S_t + (\mu - \frac{\sigma^2}{2})(T - t), \sigma\sqrt{T - t}] \tag{2.3.10}$$

This shows that $\ln S_T$ has a lognormal distribution. The uncertainty about the logarithm of the stock price is

$$\sqrt{var[\ln S_T]} \approx \sqrt{T - t}. \tag{2.3.11}$$

Using equation (2.3.10) and properties of the lognormal distribution the expected value of S_T is given as

$$E(S_T) = S_t e^{\mu(T-t)}. \tag{2.3.12}$$

(see Hull (1989)).

Equation (2.3.12) can, therefore, be used to estimate the price of stock at time T. Although this is the case, empirical studies show that the distribution of stock returns is far from normal and that the logarithm of stock prices tend to have a distribution with log-linear tails (Bibby and Sorensen (1997). Further it is shown that after a sufficiently long time the logarithm of the stock price is approximately hyperbolically distributed (see Bibby and Sorensen (1997)). A hyperbolic diffusion model for stock prices is reviewed in the next subsection.

2.3.4. Hyperbolic Diffusion Model for Stock Prices. There is empirical evidence that stock returns are better modelled by distributions other than the normal distribution. Eberlein and Keller (1995) use a class of hyperbolic distributions to fit empirical returns with high accuracy. Hyperbolic distributions differ from normal distributions in that the log-density of the former is a hyperbola while the latter is a parabola. Hyperbolic distributions have been used in various scientific areas such as modelling of turbulence and sand deposits. One class of hyperbolic distributions is given by the hyperbolic density function

$$hyp(x) = \frac{\sqrt{\alpha^2-\beta^2}}{2\alpha\sigma K_1(\sigma\sqrt{\alpha^2-\beta^2})} e^{-\alpha\sqrt{\sigma^2+(x-\mu)^2}+\beta(x-\mu)}$$

where K_1 denotes the modified Bessel function of the third kind with index 1, α and β (with $\alpha > 0$ and $0 \leq \mid \beta \mid \leq \alpha$) determine the shape of the distribution while σ and μ are scale and location parameters respectively.

Eberlein and Keller (1995) analyse the prices of ten of the stocks that compose the German stock index , DAX. Maximum likelihood estimation is performed to estimate model parameters and after carrying out significant tests it is concluded that daily stock returns are best modelled by hyperbolic distributions.

Bibby and Sorensen (1997) use hyperbolic distributions and propose a diffusion process model for the logarithm of stock price. Due to empirical evidence that the logarithm of the stock price is a process with increments that are not independent, a model for the stock price S_t of the following form is suggested:

$$S_t = e^{\kappa t + X_t} \tag{2.3.13}$$

where

$$X_t = X_0 + \int_0^t v(X_s) dW_s$$

where X_t is the state variable and κt is the constant drift rate. On application of Ito's Lemma, the following is obtained

$$dS_t = S_t\{[\kappa + \frac{1}{2}v^2(logS_t - \kappa t)]dt + v(logS_t - \kappa t)dW_t\}. \tag{2.3.14}$$

This implies that the asset price S_t follows geometric Brownian motion provided $v(x)$ is constant. The parameters of the distribution may be calculated numerically. Predota (2006), however, suggests the use of asymptotic formulas for maximum-likelihood estimators of hyperbolic density functions while Bibby and Sorensen (1995) use martingale estimation functions. Further methods for estimating parameters can be found in Kessler (2000) and Bibby and Sorensen (2001).

Bibby and Sorensen (1997) obtained several statistical properties of the process S_t . For instance, they showed that the marginal distribution of $logS_t$ is hyperbolic and hence $logS_t$ is approximately hyperbolically distributed after a sufficiently long time period. Further, the distribution of increments over short intervals has thick tails while an increment over a long interval follows a distribution that is close to being hyperbolic. They also provided the theory in applying the hyperbolic diffusion model to option pricing.

2.3.5. Time Series Models. In most financial applications, the prediction of future values of assets is very crucial. Scientific forecasting, based on sound and statistical methods, is used to provide a likely or expected value for some outcome. One such method is regression. *Regression* is a technique for exploring relationships between variables of any discrete time series data such as stock data. There are many regression models in finance that attempt to establish relationships between variables with a view to forecast future values. For instance, *linear regression* explores linear relationships by fitting straight lines through data using the method of least squares. It attempts to fit a model of the form

$$\widehat{y}_t = a + bx_t. \tag{2.3.15}$$

However, stock price data do not exhibit linear relationships to warrant the use of simple linear regression models of the form 2.3.15. More rigorous time series analysis methods are often sought. These are reviewed in the following subsections.

2.3.5.1. *Autoregressive Model of order p.* Autoregressions are regression models that relate a variable to its past values. If $\{Y_t\}$ is a time series, then a general autoregressive model of order p, denoted AR(p), is a model of the form

$$Y_t = \sum_{i=1}^{p} \varphi_i Y_{t-i} + \varepsilon_t, \tag{2.3.16}$$

where $\varphi_1, \varphi_2, \cdots, \varphi_p$ are the parameters of the model and ε_t is an error term with zero mean and constant variance σ^2. An error term with zero mean and constant variance is referred to as a *white noise* error term. In the case where $p = 1$, the model relates the variable Y_t to Y_{t-1} and it is called *first order autoregressive model*, abbreviated by AR(1) and given by

$$Y_t = \varphi_1 Y_{t-1} + \varepsilon_t. \tag{2.3.17}$$

The autoregressive model represents a variable as a linear function of its past values. The AR(1) model is restrictive since it assumes that Y_t depends only on Y_{t-1}. However, in reality Y_t might depend on other variables, hence the need for an AR(p) model.

Although an autoregressive model of order p may be used in modeling financial time series, one major challenge is the choice of the value of p and related parameters. In practice, useful tools are the autocorrelation function (ACF) and partial autocorrelation function (PACF). A detailed treatment of these may be found in Anderson et al and Lutkepohl.

An improvement of the AR(p) model includes a moving average component. This model is the subject of the next subsection.

2.3.5.2. *Autoregressive Moving-Average (ARMA) Model.* One model that has proven to be extremely useful in the analysis of discrete-time random processes is the Auto-regressive Moving-Average (ARMA) model. The ARMA model has two parts, an autoregressive (AR) part and a moving average (MA) part. The model is usually referred to as the ARMA(p,q) model where p is the order of the autoregressive part and q is the order of the moving average part. In a moving average process, a variable is expressed in terms of current and previous white noise errors. Consequently an MA(q) is a moving average model of order q

and is written as

$$Y_t = \varepsilon_t + \sum_{i=1}^{q} \varphi_i \varepsilon_{t-i}. \tag{2.3.18}$$

Combining the AR(p) and the MA(q) parts leads to the ARMA(p,q) model of the form

$$Y_t = \varepsilon_t + \Sigma_{i=1}^{p} \phi_i Y_{t-i} + \Sigma_{i=1}^{q} \theta_i \varepsilon_{t-i}. \tag{2.3.19}$$

The error terms ε_t are assumed to be independently and identically distributed random variables from a normal distribution with zero mean and variance σ^2, that is, $\varepsilon_t \sim N(0, \sigma^2)$.

2.3.5.3. *Auto-Regressive Conditional Heteroscedasticity (ARCH) Model.* Mandelbrot (1963) observed that large price changes tend to be followed by other large changes, while small changes are usually followed by other small changes. This phenomenon, known as *volatility clustering*, is best modeled by a model developed by Robert Engle (2001) called Auto-Regressive Conditional Heteroscedasticity (ARCH) Model.

As the name suggests, the ARCH model has two properties: autoregression and conditional heteroskedasticity. Autoregression implies that it uses previous estimates of volatility to calculate subsequent (future) values while conditional heteroskedasticity implies that the volatility varies with time. The simplest ARCH model is the ARCH(1) model given by

$$Y_t = \sigma_t \varepsilon_t, \sigma_t^2 = \omega + \alpha y_{t-1}^2, t = 1, 2, \cdots, T, \tag{2.3.20}$$

where $\varepsilon_t's$ are independent identically distributed and $\varepsilon_t \sim N(0, 1)$. A general ARCH(p) model is given by

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 \tag{2.3.21}$$

(see Rydberg (2000) and Engle (2001)).

ARCH models were later generalised by Bollerslev (1986) and have become to be known as generalised ARCH (GARCH) models (see Bollerslev (1986)). The GARCH type of models are widely used to model market returns. Other autoregressive models include AutoRegressive Integrated Moving Average (ARIMA) models popularised by Box and Jenkins (1976) (see Bhat (1984)). ARIMA models have been applied to forecast the prices of electricity in Spanish and Californian markets (see Contreras et al (2003)). Other ARCH-type models that have been developed in the recent past include the Exponential GARCH, EGARCH, model and Heterogeneous ARCH, HARCH, model mentioned in Rydberg (2000).

2.3.6. General Random Walk Models and Other Models. Due to the efficient market hypothesis asset prices are generally assumed to move randomly (Wilmott etal (1995)). This implies that the past historical information is fully reflected in the present price and that markets respond immediately to any new information about an asset. Thus the price of an asset is effectively affected by the arrival of new information.

Suppose the price of an asset at any time t is S. Consider during a small time interval dt in which the asset price changes from S to $S + dS$. Wilmott etal (1995) models the corresponding return on the asset, $\frac{dS}{S}$, by decomposing the return into two components. The first component is a measure of the average rate of growth, μ, of the asset price known as *drift*. Over the time interval dt, this makes a contribution μdt to the return $\frac{dS}{S}$. The drift is often a constant in simple models but in more complicated models, such as for exchange rates, μ can be a function of S and t.

The second component measures the standard deviation of the returns. This models the random change in the asset price in response to external effects such as unexpected news and is represented by a random sample taken from a normal distribution with mean zero. The contribution of this to $\frac{dS}{S}$ is σdX, where σ is a number called *volatility* and the quantity dX is the sample from a normal distribution. Putting these components together leads to an equation similar to equation 2.3.3 called **stochastic differential equation**

$$\frac{dS}{S} = \sigma dX + \mu dt. \tag{2.3.22}$$

The stochastic differential equation is a particular example of a random walk model that is often used to describe the price process of many assets. When the volatility is zero, equation 2.3.22 reduces to an ordinary differential equation

$$\frac{dS}{S} = \mu dt \tag{2.3.23}$$

which is solved to give an exponential growth in the value of the asset

$$S = S_0 e^{\mu(t-t_0)}, \tag{2.3.24}$$

where S_0 is the value of the asset at $t = t_0$. This implies that if $\sigma = 0$ the future price of an asset can be predicted with certainty. However, in reality volatility is never zero. Consequently the component σdX is certainly a feature of the asset price process. The term dX is known as a **Wiener process** and has the following properties: dX is a random variable drawn from a normal distribution; the mean of dX is zero and the variance of dX is dt.

There are other versions of the random walk model. For example, if P_t is the price of stock at time t and let its residual be ϵ_t, Granger and Morgenstern (1970) propose a random walk model written, in its simplest form, as

$$P_t = P_{t-1} + \epsilon_t \tag{2.3.25}$$

where $E[\epsilon_t] = 0$, $cov[\epsilon_t, \epsilon_{t-s}] = 0$, for all $s \neq 0$. The expression $cov[\epsilon_t, \epsilon_{t-s}] = 0$ implies that the residuals whose mean is zero are uncorrelated with all previous residuals. In this form, the implications of the random walk model is that the best predictor of the following day's price is the current price. More generally the best predictor of any future price is the most recently available price. It is shown formally through reapplication of equation (2.3.25) as follows. Using equation (2.3.25), it follows that

$$P_{t+1} = P_t + \epsilon_{t+1}$$

implies that

$$P_{t+2} = P_{t+1} + \epsilon_{t+2},\ P_{t+3} = P_{t+2} + \epsilon_{t+3}$$

and finally

$$P_{t+n} = P_t + \Sigma_{j=1}^{n} \epsilon_{t+j}.$$

(see Granger and Morgenstern (1970)).

Since $E[\epsilon_t] = 0$ and $cov[\epsilon_t, \epsilon_{t-s}] = 0$, then $E[\Sigma_{j=1}^{n} \epsilon_{t+j}] = 0$ and the result follows.

The model given by equation 2.3.22 fits real time data series very well especially equities and indices (Wilmott etal (1995)). Although real data exhibit higher probability of large rises or falls than the model predicts, the random walk model turns out to be the basis for more sophisticated models such as the Mean Reverting Process and the Ornstein-Uhlenbeck Process (see Neftci (1996)). Blasco et al (1997), in their study of the random walk hypothesis in the Spanish stock market, conclude that while stock returns are not independent and identically distributed, stock prices appeared to follow the random walk. This view is also shared by many financial economists as well as statisticians (Malkiel (2003)).

While there is very strong empirical evidence in favour of the random walk model, it is not an absolutely perfect fit for all price series or over all time intervals. It appears valid for markets that have the characteristic of the stock market (Granger and Morgenstern (1970)). For instance, in previous studies cotton prices did not appear to be following the random walk hypothesis (Mandelbrot (1963)). Lo and MacKinlay (1988) reject the random walk model for weekly returns for the entire sample period ($1962-1982$). They provide evidence that stock prices do not follow random walks by using a simple specification test based on variance estimators. Darrat and Zhong (2000) tested the random walk hypothesis on daily stock price data of China's two official stock markets (Shanghai and Shenzhen). The results obtained did not support the random walk hypothesis. Another study that does not support the random walk hypothesis was carried out by Niederhoffer (1965).

There are various other models that have been proposed which build on the independent increments of returns. For instance, Praetz (1972) presents a scaled t-distribution model which appears to be a good fit to weekly share price indices from the Sydney Stock Exchange for the period $1958-1966$ (see Rydberg (2000)).

Another model of the discrete time type that has been proposed in recent times is the Autoregressive Conditional Duration (ACD) model developed by Engle and Russell (see Rydberg (2000) and Engle and Russell (1998)). In this model, the arrival of transactions are described as a counting process and the duration between events follows a process of the type

$$Y_i = \omega + \alpha x_{i-1} + \beta Y_{t-1} \tag{2.3.26}$$

for $\alpha \geq 0$, $\beta \geq 0$, $\omega > 0$ for all i, $i = 1, 2, \cdots, n$ where x_i denotes the duration between events at time t_{i-1} and t_i. A thorough coverage of other models can be found in Rydberg (2000)).

CHAPTER 3

The Double Gamma Model, Data and Methodology

The Double Gamma model proposed in this thesis studies the distribution of differences between closing stock prices on successive days. Until recently, stock price differences have been modeled as being either normal or log-normal (Brada et al (1965) and Mandelbrot (1963)). However, normal quantile-quantile ($Q-Q$) plots for indices and stocks studied reveal the contrary. This is shown in Appendix A.

3.1. The Data

The raw data for analysis consists of 25,786 daily observations of three major indices and 36,377 values of stock prices for six firms. The data comprises daily closing stock values for major firms trading on the London Stock Exchange, New York Stock Exchange and Tokyo Stock Exchange. Indices consist of Dow Jones Industrial Average, Japan's Nikkei 225 and Financial Times 100 Index which are reviewed in the following subsection. Table 3.1 below summarises the nature of the data used in this thesis.

Table 3.1. **Summary of sampled data**

Index/Stock	Sample Period	Sample Size	Country/Region
FTSE	02/04/1984 - 19/01/2007	5,761	U.K. (Europe)
Nikkei 225	04/01/1984 - 19/01/2007	5,672	Japan (Asia)
S&P 500	03/01/1950 - 19/01/2007	14,353	USA (America)
Sony Corporation	06/04/1983 - 19/01/2007	5,997	Japan (Asia)
Toyota Corporation	13/04/1993 - 19/01/2007	3,464	Japan (Asia)
Microsoft Corporation	13/03/1986 - 19/01/2007	5,261	USA (America)
General Motors	02/01/1962 - 19/01/2007	11,340	USA (America)
GlaxoSmithKline plc	09/07/1986 - 19/01/2007	5,179	UK (Europe)
Barclays plc	10/09/1986 - 19/01/2007	5,136	UK (Europe)

The data set was downloaded from the finance subdirectory of the website 'Yahoo.com'. The sample period covered ranges from 1950 to January 2007 with varying commencement dates. Data from developing countries has not been considered owing to the low level of activity on stock exchange markets and stock data inaccessibility. Research has shown that low volume and thinly traded markets are inappropriate for efficiency since they lack liquidity and do not provide smooth transfer of information. Further, price indices in small markets tend to exhibit inflated volatility thereby complicating statistical inference (see Darrat and Zhong (2000)).

3.1.1. Overview of the selected indices and stocks.

- **Financial Times 100 Index**: The Financial Times 100 Index (FTSE 100) is a share index that is commonly used as a benchmark for the performance of stocks traded on the London Stock Exchange. The FTSE index consists of the 100 largest companies traded on the London Stock Exchange (based on market capitalization). The companies in the list include BP, British Airways, Barclays Bank, GlaxoSmithKline, Unilever and Vodafone just to mention a few. The stocks represent about 80 percent of the value of all issues traded on the exchange. The FTSE index is used as a benchmark for success of the British economy.

- **Nikkei 225 Index**: The Nikkei 225 Index is a stock market index for the Tokyo Stock Exchange (TSE). It is composed of 225 leading stocks traded on the Tokyo Stock Exchange. Major companies in this index include Sony Corporation, Sharp Corporation, Toyota Motor Corporation and Japan Airlines Corporation.

- **Standard and Poors 500**: The Standard and Poors 500 (S&P 500) Index is a market index based on a portfolio of 500 different stocks that are traded on the New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and the Nasdaq National Market System. The 500 stocks is composed of 400 industrials, 40 utilities, 20 transportation companies, and 40 financial institutions. Compared to the DJIA Index, the S&P 500 index is viewed as a better representation of the US market as it incorporates more firms from a wide range of fields. These firms include Chevron Corporation, Apple Computer, Cisco Systems, General Motors and PepsiCo Inc.

- **Stocks**: The stocks for Barclays plc and GlaxoSmithKline plc traded on the London Stock Exchange are used. US stocks used are those for Microsoft Corporation and General Motors while from the Tokyo Stock Exchange we use stocks for Toyota Motor Corporation and Sony Corporation.

3.2. Model Specification

Under the Double Gamma model, we model the differences between the closing prices of stock on day t and day $t-1$, that is, $P_t - P_{t-1}$. The notation $P(t)$ and $P(t-1)$ is used interchangeably to mean P_t and P_{t-1} respectively. Accordingly we define the following:

Let $\Delta P_t = P_t - P_{t-1}$ be the change in closing price from day $t-1$ to day t.

In particular, denote $$\Delta P_t = \left[\begin{array}{c} \mathrm{X}_t = P_t - P_{t-1} > 0 \text{ with probability } P \\ \mathrm{P}_t - P_{t-1} = 0 \text{ with probability } R \\ \mathrm{Y}_t = P_t - P_{t-1} < 0 \text{ with probability } Q \end{array} \right.$$
where $P + R + Q = 1$.

3.2.1. Model Assumptions. The proposed model is based on the following assumptions:

(i) The process followed by P_t is a discrete time process.
(ii) ΔP_t are independent and identically distributed.
(iii) ΔP_t is independent of time t.
(iv) $X_t \sim Ga(\alpha, \beta)$
(v) $Y_t = -X_t$
(vi) Variations in asset price are random.
(vii) The present price P_t possesses a Markov property.
(viii) On two successive trading days, the asset price can either increase or decrease, not stay constant.

Owing to the nature of the data used, the closing prices can only be quoted at the end of each trading day. Thus it is reasonable to assume that the process followed by P_t is a discrete time process. Further, since all known information is used optimally by market participants, variations in asset prices are random and that the present price P_t is independent of all past prices (see section 1.3).

Charts produced depicting the stock price changes $P(t) - P(t-1)$ shown in Appendix A are revealing. A closer look at the histograms produced, it is tempting to conclude that actual differences, $P(t)-P(t-1)$ (positive and negative), suggest a normal distribution for each of the indices and stocks. On the other hand, histograms of the absolute differences, $|P(t) - P(t-1)|$, for each index and stock suggest a distribution of the gamma type. (See figure below and Appendix A, parts (b), (c) and (d)).

A common approach to testing the normal fit is the use of quantile-quantile $(Q - Q)$ plots. Parts (d) of the charts in Appendix A show normal $Q-Q$ plots for the differences, $P(t)-P(t-1)$, for each of the indices and stocks. From Figure 3.1, showing the histogram of absolute price differences and the $Q-Q$ plot for the S&P 500 index, the deviations from the straight line and thus from normality are obvious (see also parts (d) of each figure in Appendix A).

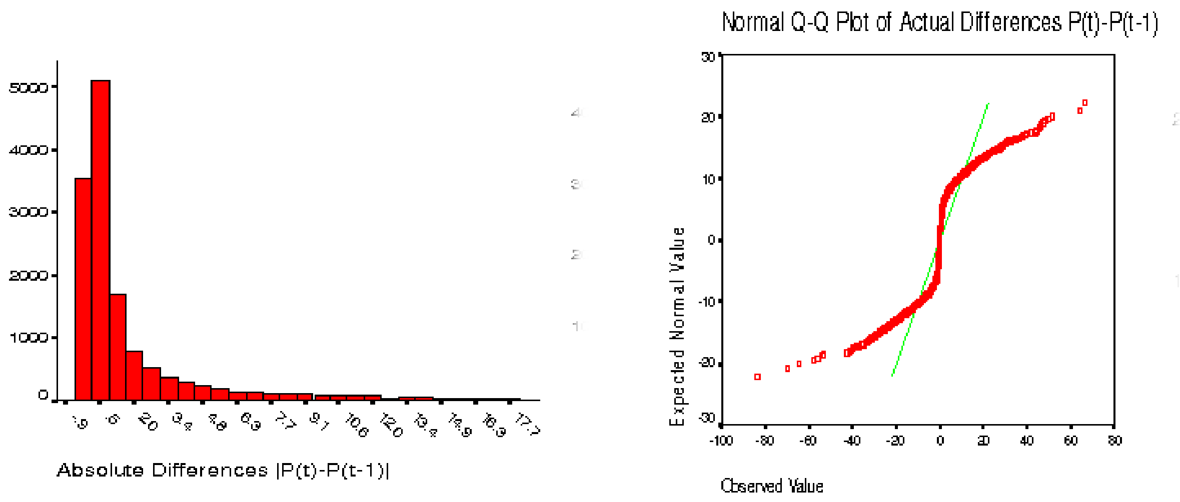

FIGURE 3.1. Histogram of Absolute Price Differences and Q-Q Plot for S&P 500

It can also be seen from Figure 3.1 that the curved pattern with slopes increasing from left to right in the histogram (on the left) for the absolute differences suggests that the data distribution is skewed to the right. This is a feature that is synonymous with the gamma distribution.

The last assumption is not completely true in the real world. However, it can be seen from Appendix B that stock prices or values of indices are rarely constant. Over 5,761 trading days the FTSE index changed approximately 0.30% of the time while the S&P

500 index remained constant 124 times over a period of 14,353 trading days representing 0.86% of the time. Overall, out of a total 62,163 daily closing values, stock prices and index values remained constant 2,612 times which represents 4.2% of the time. It is, therefore, reasonable to assume that stock prices either increase or decrease on any two succesive trading days. Thus the assumption holds at least 95.8% of the time.

3.3. Methodology

Two models are considered for simulating stock prices. These are the lognormal model used in the Black-Scholes option pricing model and the *double gamma* model that is proposed in this thesis. Simulation is done using the Statistical Package for the Social Sciences (SPSS) and Microsoft Excel. Charts are drawn using SPSS while random uniform numbers in the interval $[0, 1]$ are generated by the random number generator in Excel.

The methodology used depends on the model being simulated and the approach followed. The modeling is carried out using three approaches. Two approaches have been used in the simulation under the *double gamma* model while the third approach is modeled along the lognormal distribution. These are outlined below.

(1) **Approach 1**: ΔP_t is modeled as following a *modified double gamma* distribution with probability p of an "up" movement (i.e. $P_t - P_{t-1} > 0$) and probability $1-p$ of a "down" movement (i.e. $P_t - P_{t-1} < 0$).

(2) **Approach 2**: ΔP_t is modeled as following the plain *double gamma*.

(3) **Approach 3**: $ln\dfrac{P_t}{P_{t-1}}$ is modeled as following the *normal* distribution.

In each of the three approaches sample means of the differences between successive observed closing prices are obtained. From the law of large numbers, the sample mean of the observed ΔP_t is used as an estimator of

$$E[\Delta P_t] = E[P_t - P_{t-1}] = \alpha\beta$$

of the modeled $Gamma(\alpha, \dfrac{1}{\beta}) \equiv \Gamma(\alpha, \beta)$ for approaches 1 and 2 above.

3.3.1. Parameter estimation for the models. We use the gamma maximum likelihood estimators that were highlighted in subsection (2.2.1) for approaches 1 and 2.

The scale and location parameters of the gamma distribution are estimated using

$$\widehat{\alpha} = \frac{\overline{\Delta P_t}^2}{s^2} \quad and \quad \widehat{\beta} = \frac{s^2}{\overline{\Delta P_t}},$$

where s^2 is sample variance of price differences while $\overline{\Delta P_t}$ is sample mean of price differences.

Random numbers are generated by the computer using the random number generator command in Microsoft Excel. Using the parameters that have been estimated using historical data, we use the Excel function 'GAMMAINV' to obtain a value for the gamma probabilistic input.

Under approach 3, we use the procedure in Hull (1989), to estimate the volatility empirically. Since $ln\frac{S_T}{S_t}$ is normally distributed with mean $(\mu - \frac{\sigma^2}{2})(T-t)$ and variance $\sigma^2(T-t)$ (see Hull (1989)), two parameters are estimated: μ and σ.

Assuming no intermediate cash flows such as dividends, let

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

where S_i is the closing price of the asset at the end of the ith time-interval. As $S_i = S_{i-1}e^{u_i}$ is the continuously compounded return in the ith interval, an unbiased estimator, s of the standard deviation of the $u_i's$ is given by

$$s = \sqrt{\frac{1}{n-1}\Sigma_{i=1}^{n}(u_i - \overline{u})^2},$$

where $\overline{u}$ is the mean of the $u_i's$.

With the parameters estimated and random probabilities generated, we use the Excel function 'NORMINV' to obtain a value for a probabilistic input that is normally distributed. Then the estimated stock price under the lognormal model is obtained as

$$\widehat{P_t} = P_{t-1}\widehat{z_t}$$

where $\widehat{P_t}$ is the simulated stock price at time t, P_{t-1} is the actual closing stock price at time $t-1$ while $\widehat{z_t}$ is the simulated probabilistic input under the normal distribution obtained using the Excel command function 'NORMINV'.

3.3.2. Model Properties. Using the assumption that the asset price can either increase or decrease, then $\Delta P_t = \left[\begin{array}{l} \mathrm{X}_t = P_t - P_{t-1} > 0 \text{ with probability } P \\ \mathrm{Y}_t = P_t - P_{t-1} < 0 \text{ with probability } Q \end{array}\right.$

since $R = 0$.

Thus the probability of an "up" movement and "down" movement would, respectively, be redefined as

$$p = \frac{P}{P+Q} and q = 1 - p = \frac{Q}{P+Q}$$

Further as $0 < X_t < \infty$ and $-\infty < Y_t < 0$, then under the gamma distribution assumption, $X_t \sim Ga(\alpha_1, \beta_1)$ and

$$f(x) = \frac{\beta_1(\beta_1 x)^{\alpha_1 - 1} e^{-\beta_1 x}}{\Gamma \alpha_1}, \qquad 0 < x < \infty$$

while $Y_t \sim Ga(\alpha_2, \beta_2)$ and

$$g(y) = \frac{\beta_2(-\beta_2 y)^{\alpha_2 - 1} e^{\beta_2 y}}{\Gamma \alpha_2}, \qquad -\infty < y < 0,$$

where $x \equiv X_t$ and $y \equiv Y_t$.

Let M be the number "up" movements (that is, when $P(t) > P(t-1)$), and N be the number of "down" movements in time interval $[0, T]$.

Let μ be the rate at which prices change such that the rate at which the "up" movements occur is μp while the rate at which the "down" movements occur is $\mu(1-p)$. Then we define

$$P_t = P_{t-1} + \Sigma_{k=1}^{t-1} \Delta P_k = P_{t-1} + \Sigma_{k=1}^{M} \Delta X_k + \Sigma_{k=1}^{N} \Delta Y_k, \tag{3.3.1}$$

where M and N are random.

Let M and N be independent. Then by Theorem 2.1.18,

$$M \sim Poiss((\mu p)T)$$

and

$$N \sim Poiss((\mu(1-p))T).$$

Further since M and N are independent and using properties of the Poisson distribution it follows that for period T,

$$M + N \sim Poiss(\mu T).$$

Let M and X_k be independent and N and Y_k be independent. Then

$$\begin{aligned}
E[P_t] &= P_{t-1} + E[M]E[X_k] + E[N]E[Y_k] \\
&= P_{t-1} + E[M]E[X_k] - E[N]E[X_k] \\
&= P_{t-1} + E[X_k]\{E[M] - E[N]\} \\
&= P_{t-1} + \frac{\alpha}{\beta}[\mu pT - (\mu(1-p))T] \\
&= P_{t-1} + \frac{\alpha}{\beta}[\mu pT - \mu T + \mu pT] \\
&= P_{t-1} + \frac{\alpha}{\beta}(2\mu pT - \mu T) \\
&= P_{t-1} + \frac{\alpha\mu T}{\beta}(2p-1),
\end{aligned}$$

where

(i) the parameters for $Ga(\alpha, \beta)$ are obtainable from the historical data using maximum likelihood estimators.

(ii) $\mu = \frac{M+N}{T}$, T is the time period.

(iii) $p = \dfrac{P}{P+Q}$.

Since $\mu = \frac{M+N}{T}$, then the model simplifies to

$$E[P_t] = P_{t-1} + \frac{\alpha(M+N)}{\beta}(2p-1),$$

Further since under the last assumption the price does not remain constant, then over period T the total number of 'up' and 'down' movements must be equal to the number of trading days, that is, $M + N = T$. Thus

$$E[P_t] = P_{t-1} + \frac{\alpha T}{\beta}(2p-1). \tag{3.3.2}$$

This result is used to estimate future stock prices and index values. Hence

$$\widehat{P_t} = P_{t-1} + \frac{\alpha T}{\beta}(2p-1). \tag{3.3.3}$$

Let the residual error at time t be denoted ε_t where

$$\varepsilon_t = \widehat{P_t} - P_t.$$

The approximation is improved by introducing a shift factor. The shift factor is found by taking the mean of absolute residual errors, $\overline{\varepsilon_{t-1}}$, from the first trading day to day $t-1$.

This is used since it represents the average absolute deviations from the true value. The shift factor is applied as follows:

- if on day $t-1$, $\varepsilon_{t-1} < 0$, then add $\overline{\varepsilon_{t-1}}$ and thus $\widehat{P_t} = P_{t-1} + \frac{\alpha T}{\beta}(2p-1) + \overline{\varepsilon_{t-1}}$.
- if on day $t-1$, $\varepsilon_{t-1} > 0$, then subtract $\overline{\varepsilon_{t-1}}$ and thus $\widehat{P_t} = P_{t-1} + \frac{\alpha T}{\beta}(2p-1) - \overline{\varepsilon_{t-1}}$.

Hence

$$\widehat{P_t} = P_{t-1} + \frac{\alpha T}{\beta}(2p-1) \pm \overline{\varepsilon_{t-1}}. \tag{3.3.4}$$

In addition, with $\mu = \frac{M+N}{T}$ specified, the probability density function of $\Delta P_t = P_t - P_{t-1}$ would be estimated by the double gamma distribution given by

$$f(\Delta P_t) = \frac{1}{2}\frac{|\Delta P_t - \mu|^{\alpha-1} e^{\frac{|\Delta P_t - \mu|}{\beta}}}{\beta^{\alpha}\Gamma\alpha}, \tag{3.3.5}$$

where μ is the location parameter while α and β are shape and scale parameters respectively.

3.3.3. Measuring Model Performance. To see how well a model performs, we look at the relative errors generated by the models. Suppose $\widehat{P_t}$ is the stock price at time t generated by the model and P_t is the actual or observed price on the market. The relative error (RE) is calculated by

$$RE = \frac{|\widehat{P_t} - P_t|}{P_t} \tag{3.3.6}$$

If the relative error (expressed in percentage) is small, it means the model gives a good approximation to the stock price reflected on the market. Conversely, if the relative error is big, then the model is considered to be a poor approximation to the market price. However, apart from checking how close the model price is to the market price, it is also important to check whether the model produces over- or underestimates to the market. In other words, it is important to see when a model underprices and overprices a specific stock. To achieve this, we modify the relative error formula to

$$RE = \frac{\widehat{P_t} - P_t}{P_t} \tag{3.3.7}$$

A negative relative error then means that the model underprices the specific stock whereas a positive relative error means that the model overprices the specific stock.

The other approach is to look at the absolute value of the errors. The absolute values of the errors are used to assess the magnitude of the error. The best model is the one that results in least values of the absolute errors. The absolute value of the errors are given as

$$E = |P_t - \widehat{P}_t| \tag{3.3.8}$$

We use both approaches in the analysis of results obtained.

In order to assess the performance of the proposed model, simulation is carried for each stock/index thirty (30) times. The estimate $\widehat{P}_t$ for each simulation is recorded and the absolute errors $\varepsilon = |P_t - \widehat{P}_t|$ are calculated. The mean values of the errors and the estimates $\widehat{P}_t$ are also calculated. The approach that results in the least value of the mean errors is considered to be the best approach for modeling prices of stocks and indices that have been studied.

CHAPTER 4

Results, Analysis and Conclusion

4.1. Results and Analysis

In this chapter, results that were simulated using the proposed double gamma model and lognormal model as described in section 3.3 are presented and compared. The simulation results can be found in Appendix C. In Appendix D, relative errors for all the stocks and indices are presented. The relative errors given are for the last 34 days to present time t, that is, 19th January 2007. Percentage relative errors may be obtained by multiplying by 100.

Table 4.1 shows partial simulated results obtained for Microsoft Corporation in comparison with the actual closing stock price of 31.11 as recorded on 19 January 2007.

TABLE 4.1. **Simulated Results for Microsoft Corporation**

Simulation	$\widehat{P}_t$ App. 1	Error App.1	$\widehat{P}_t$ App. 2	Error App. 2	$\widehat{P}_t$ App. 3	Error App. 3
1	31.8	0.7	34.0	2.9	33.3	2.19
2	32.3	1.2	31.2	0.1	32.1	0.99
3	32.7	1.8	32.0	0.9	34.9	3.79
4	29.8	1.5	29.9	1.2	33.3	2.19
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮
28	34.0	2.9	33.4	2.3	33.0	1.89
29	36.4	5.3	40.1	9.0	33.4	2.29
30	33.7	2.6	33.1	2.0	32.2	1.09
Mean Values	33.0	3.5	33.5	2.9	33.6	2.5

Considering the size of absolute errors for each approach, it can be seen that the lognormal model outperforms the other approaches using the proposed double gamma modeling. As simulated results for Microsoft Corporation indicate, after 30 simulations approaches 1 and 2 gave average values of 33.0 and 33.5 respectively against the value of 33.6 obtained using approach 3. However, the average values using approaches 1 and 2 are closer to the actual recorded value than what approach 3 generates.

Approach 1 turned out to be the best in modeling the FTSE and Nikkei indices outperforming the traditional lognormal model. While the actual FTSE index value on day t (19th January 2007) was $P_t = 6237.2$, after 30 simulations approach 1 gave an average value of $\widehat{P_t} = 6237.8$ which is a very good approximation. On the other hand approach 2 outperformed the rest in modeling the values of the S&P 500 index. The predicted value at time t given by the average was found to be $\widehat{P_t} = 1427.7$ compared to the actual value of 1430.5. However, both approaches 1 and 2 turned out to be extremely poor at modeling some stock prices. The worst results are obtained particularly for General Motors, GlaxoSmithkline, Sony and Toyota Corporation. However, the simulation results show that approach 3, lognormal model, is appropriate for modeling prices of all stocks that were studied. Table 4.2 summarises the ranking of each approach at modeling stocks/indices.

TABLE 4.2. **Model Performance Ranking**

Index/Stock	**Approach 1**	**Approach 2**	**Approach 3**
FTSE	1	2	3
SP 500	2	1	3
GlaxoSmithKline	3	2	1
Microsoft Corp.	3	2	1
General Motors	3	2	1
Sony	3	2	1
Toyota Corp.	3	2	1
Barclays plc	3	2	1
Nikkei	1	3	2

While approaches 1 and 2 faired poorly at modeling prices of nearly all stocks, from the results it can be observed that the poor results were extreme in cases where $N > M$ (that is, where the number of simulated 'down' movements were greater than the number of 'up' movements. Simulation figures for General Motors and Sony Corporation in Appendix B reflect this.

From the simulation results in Appendix B, it can be observed that the double gamma approaches are more accurate in modeling high valued stocks and indices. This is evident in the prediction of the FTSE and Nikkei indices by approach 1 and the prediction of the S&P 500 index using approach 2. This is where the lognormal model performs very poorly. One common feature of the stocks for which the lognormal model produced good estimates is that all of them are lowly valued.

While the absolute error approach for assessing model performance may suggest some weaknesses in approaches 1 and 2, the use of relative errors suggests otherwise. Results of relative errors shown in Appendix D show that while the lognormal model is a better predictor of all stocks, the other two approaches are still competitive. For instance approach 2 is not extremely bad in predicting stock prices for Barclays plc since the relative errors are on average less than 5%. To illustrate this, partial relative errors for Barclays plc are reproduced in table 4.3.

Table 4.3. **Relative Errors for Barclays plc**

P_t Actual	$\widehat{P}_t$ App. 1	$\widehat{P}_t$ App. 2	$\widehat{P}_t$ App. 3	RE App. 1	RE App. 2	RE App. 3
59.4	67.688885	61.25519	56.988994	0.1395435	0.03123	0.040589
59.14	67.97722	61.54478	58.459572	0.1494288	0.04066	0.011505
59.43	88.275556	61.84436	58.92972	0.1488399	0.04063	0.008418
59.73	68.533891	62.10395	58.773339	0.1473948	0.03974	0.016016
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮
53.89	62.557285	56.16242	56.28486	0.1608329	0.04217	0.04444
54.06	62.535621	56.14201	52.747384	0.1567817	0.03851	0.024281
54.04	61.643956	55.2561	56.746861	0.1407098	0.02242	0.05009

As Appendix D shows, relative errors relating to approaches 1 and 2 are large: at least 10% for Sony Corporation stocks, more than 40% for both Toyota Corporation and GlaxoSmithkline plc stocks and at least 20% for General Motors stocks. This is in agreement with our earlier observation. On the other hand the relative errors of about 1% for Microsoft Corporation under approach 2 and Nikkei under approach 1 suggest that the double gamma performs very well. The relative errors of less than 1% for both S&P 500 under approaches 1 and 2 and FTSE under approach 1 confirm that the double gamma model has effective prediction capability.

When $p = 0.5$,

$$E[P_t] \;=\; P_{t-1} + \frac{\alpha\mu T}{\beta}(2p-1)$$

reduces to

$$E[P_t] \;=\; P_{t-1}$$

which confirms the notion that the best estimate of stock price at time t is its value at time $t-1$. This, however, did not arise in the study of the stocks and indices that were used.

4.2. Conclusion

In mathematical modeling, assumptions are formulated to simplify the number of variables in the problem. While this makes the analysis simpler but it, however, makes the model less accurate. The double gamma model developed is supposed to be valid while the assumptions that were made hold. In the real market place some of the assumptions may not hold while others may hold. On the other hand, certain factors such as payment of dividends that were not incorporated into the model may negatively impact on the results. The model developed in this thesis is no exception.

Before developing the model, it was necessary to review the behaviour of the stock prices, the financial market system and the justification for studying stock prices and indices. This was done in Chapter 1. The unpredictable nature of stock prices provoked the need to review processes that are random in nature called stochastic processes as well as methods of estimating some important statistics representing such processes and simulation. Efforts by other researchers in modeling stock prices and financial time series are also reviewed. One of the models, the lognormal model of stock prices, is used as a comparator to the model developed in this thesis. The concepts and models have been covered in Chapter 2. The concepts covered are used in the development of the model in this thesis.

In Chapter 3, the data is reviewed and the model is developed. The methodology used to obtain the results is also explained. The results obtained are presented and analysed in Chapter 4. The results show that the model developed has strong prediction capabilities of indices and stock prices through Markov chain modeling. It was seen that the model developed in this thesis outperformed the traditional lognormal model in modeling some stocks or indices. However, due to the random nature of stock prices, deviations from actual stock and index values are expected.

While the double gamma model has shown strong prediction capabilities of indices and stocks, further improvements to the model can be made. One possible way would be to increase the number of stocks and indices that are modeled. A few more stocks and indices taken from different stock exchanges could be examined. The results of such studies would assist in making modifications to the model.

Further since the model performed very well for two of the three indices, this may suggest that it would be more applicable to indices than stocks. Thus more indices could be studied to check whether the model would still perform very well. In addition, a few more simulations could also help improve the proposed model compared to the thirty simulations that were undertaken in this study.

The limitation of this study lies in two aspects. Firstly only those stock and index values that are nonconstant can be modeled. Thus caution must be exercised for generalisation of the results to other data. Secondly, there is need to have a large set of data values which may somewhat be difficult to find. Without a reasonable amount of data available, the model may not yield the expected results. Other assumptions, such as the price differences being identically distributed, may not be valid due to the fact that data taken over a long time range may have been exposed to different economic environments that may impact on their distributions.

The gamma distribution provides a considerable flexibility as the distribution can take a number of shapes. Thus what is required is the estimation of shape and location parameters, α and β respectively. These parameters can be estimated from historical prices of stocks and index values.

This study has in a way achieved the main objective and has also managed to set the platform for further research in modeling stock prices and indices by exploring the distribution of the differences $P_t - P_{t-1}$. The degree of accuracy with which the model is able to estimate index values provides the strongest hint that the double gamma distribution can play a major role in modeling the stock market, in general, and indices, in particular. Although this research focussed on stock prices and indices, it could be extended to include the possibility of pricing options and incorporating other factors such as payment of dividends into the model.

Index of Definitions

Bibliography

[1] Anderson, D.R., Sweeney, D.J. and Williams, T.A., *Quantitative Methods for Business*, 8th Edition, South-Western College Publishing, 2001.

[2] Bhat, U.N., *Elements of Applied Stochastic Processes*, Second Edition, John Wiley & Sons, Inc, 1984.

[3] Bibby, B.M. and Sorensen, M., *A hyperbolic diffusion model for stock prices*, Finance Stochast 1, 25-41, 1997

[4] Bibby, B.M. and Sorensen, M., *Simplified Estimating Functions for Diffusion Models with a High-dimensional Parameter*, Scandinavian Journal of Statistics, Volume 28, Issue 1, 99 112, 2001.

[5] Bibby, B.M. and Sorensen, M., *Martingale Estimation Functions for Discretely Observed Diffusion Processes*, Bernoulli, Vol. 1, No. 1/2, pp 17 - 39, $March - June$1995.

[6] Bierman, H. Jnr. and Hass, J., *Normative Stock Price Models,* Journal of Financial and Quantitative Analysis, 1971.

[7] Bierman, H. Jnr., Downes, D.H. and Hass, J.E, *Closed-form Stock Price Models*, Journal of Financial and Quantitative Analysis, 1972.

[8] Black, F. and Scholes M., *The Pricing of Options and Corporate Liabilities,*, Journal of Political Economy, 81, 637-659, 1973.

[9] Blasco, N., Del Rio, C. and Santamar, R., *The Random Walk Hypothesis in the Spanish Stock Market: 1980-1992*, Journal of Business Finance and Accounting, 1997.

[10] Bollerslev, T., *Generalized Autoregressive Conditional Heteroskedasticity*, Journal of Econometrics 31, 307-327, 1986).

[11] Boswijk, H.P., Hommes, C.H. and Manzan, S., *Behavioral Heterogeneity in Stock Prices*, A paper presented at the IFAC symposium Computational Economics, Finance and Engineering, Seattle, June 28-30, 2003, 2003

[12] Brada, J., Ernst, H. and Tassel, J., *The Distribution of Stock Price Differences:Not Gaussian After All?*,

[13] Brealey, R.A., *An Introduction to Risk and Return from Common Stocks*, Second Edition, The MIT Press, 1983.

[14] Chou, R.Y., *Modeling the Asymmetry of Stock Movements Using Price Ranges*, Institute of Economics, Academia Sinica, Paper presented at the 2002 NTU International Conference of Finance on Financial Crisis and Financial Distress, 2001.

[15] Cockraine, J.H., *Asset Pricing,* Princeton University Press, 2001.

[16] Contreras, J., Espinola, R., Nogales, F.J. and Conejo, A.J., *ARIMA models to predict next-day electricity prices*, Power Systems, IEEE Transactions on, Volume 18, Issue 3, 1014- 1020, 2003.

[17] Cox, J.C., Ross, S.A. and Rubinstein, M., *Option Pricing: A Simplified Approach, Journal of Financial Economics*, 7:229-263, 1979.

[18] Darrat, A.F. and Zhong, M., *On Testing the Random-Walk Hypothesis: A Model-Comparison Approach*, The Financial Review, Volume 35, Issue 3, 105-124, 2000.

[19] Eberlein, E. and Keller, U., *Hyperbolic Distributions in finance*, Bernoulli 1 (3), 281-299, 1995.

[20] Engle, R., *The Use of ARCH/GARCH Models in Applied Econometrics*, Journal of Economic Perspectives, Volume 15, No. 4, 157168, 2001.

[21] Engle, R.F. and Russell, J.R., *Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data*, Econometrica, Vol. 66, No. 5, pp. 1127-1162, 1998.

[22] Fama, E.F., *Efficient Capital Markets: A Review of Theory and Empirical Work*, Journal of Finance, 25, 383-417, 1970.

[23] Fama, E.F., *Efficient Capital Markets: II*, Journal of Finance . VOL. XLVI, NO. 5, 1991.

[24] Fama, E.F., *The Behaviour of Stock-Market Prices,* The Journal of Business, VOL. 38, Issue 1, 34-105, 1965.

[25] Follmer, H., *A Microeconomic Approach to Diffusion Models for Stock Prices,* Mathematical Finance 3, 1-23, 1993.

[26] Granger, C.W.J. and Morgenstern, O., *Predictability of Stock Market Prices,* D.C. Heath and Company, 1970.

[27] Geman, H., *Pure Jump Lvy Processes for Asset Price Modelling,* Journal of Banking and Finance, 2002.

[28] Hogg, R.V, and Craigg, A.T., *Introduction to Mathematical Statistics*, Fourth Edition, Macmillan Publishing Co., Inc, 1978.

[29] Hull, J., *Options, Futures, and Other Derivative Securities,* Prentice-Hall, Inc, 1989.

[30] Joshi, M.S., *The Concepts and Practice of Mathematical Finance,* Cambridge University Press, 2004.

[31] Kao, E.P.C., *An Introduction to Stochastic Processes,* Wadsworth Publishing Company, 1997.

[32] Kessler, M. *Simple and explicit estimating functions for a discretely observed diffusion process*, Scandinavian Journal of Statistics, Volume 27, Issue 1, 65 82, 2000.

[33] Lo, A.W. and MacKinlay, A.C., *Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test,* The Review of Financial Studies Volume 1, Number 1, pp 41-66, 1988.

[34] Lunde, A. and Timmermann, A., *Completion Time Models for Stock Prices,* Journal of Economic Literature **232** (2003).

[35] Lutkepohl, H., *Introduction to Multiple Time Series Analysis,* Second Edition, Springer-Verlag, 1993.

[36] Malkiel, B.G., *The Efficient Market Hypothesis and Its Critics*, Journal of Economic PerspectivesVolume 17, (2003) : 5982.

[37] Mandelbrot, B., *The Variation of Certain Speculative Prices*, J. Business 36, 307-332, 1963

[38] Mandelbrot, B., *The Variation of Some Other Speculative Prices*, J. Business 36, 394-419, 1963

[39] Medhi, J., *Stochastic Processes,* Wiley Eastern Limited, 1982.

[40] Moran, P.A.P, *An Introduction to Probability Theory*, Oxford University Press, 1968.

[41] Myung, I.J., *Tutorial on Maximum Likelihood Estimation*, Journal of Mathematical Psychology **232** $(2003) : 90 - 100$.

[42] Neftci, S.N., *An Introduction to the Mathematics of Financial Derivatives,* Academic Press, 1996.

[43] Niederhoffer, V., *Clustering of Stock Prices,* Operations Research, Vol. 13, No. 2., pp. 258-265, 1965..

[44] Norvaisa, R., *Modelling of stock price changes: A real analysis approach*, Journal of Finance and Stochastics, **232** $(2000) : 343 - 369$.

[45] Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, Third Edition, McGraw-Hill, Inc, 1991.

[46] Praetz, P.D., *The Distribution of Share Price Changes*, The Journal of Business, Vol. 45, No. 1, 49-55, 1972.

[47] Predota, M, *Asymptotic Estimators for Parameters of Hyperbolic Densities in Finance*, Quality Technology & Quantitative Management Vol. 3, No. 1, pp. 1-16, 2006.

[48] Rice, J.A., *Mathematical Statistics and Data Analysis*, Wadsworth, Inc, 1988.

[49] Ross, S.M., *A First Course in Probability*, Second Edition, Macmillan, 1984.

[50] Ross, S.M., *An Introduction to Mathematical Finance: Options and Other Topics*, Cambridge University Press, 1999.

[51] Ross, S.M., *Stochastic Processes*, Second Edition, John Wiley & Sons, Inc, 1996.

[52] Rydberg, T.H. and Shephard, N., *A modelling framework for the prices and times of trades made on the New York stock exchange*, Paper presented at the Workshop on Mathematical Finance, University of Bremen, Germany, February 1998. 1998

[53] Rydberg, T.H., *Realistic Statistical Modelling of Financial Data*, International Statistical Review, Vol. 68, No. 3, pp. 233 - 258, 2000.

[54] Samuelson, P.A., *Mathematics of Speculative Prices*, SIAM Review, Vol. 15 No. 1, 1973

[55] Shepp, L., *A Model for Stock Price Fluctuations Based on Information*, DIMACS Technical Report, Rurgers University, 2000.

[56] Straja, S.R., *Stochastic Modeling of Stock Prices*, Montgomery Investment Technology, Inc, 2006

[57] Willmott, P., Howison, S. and Dewynne, J., *The Mathematics of Financial Derivatives: A Student Introduction,* Cambridge University Press, 1995.

APPENDICES

APPENDIX A

Charts

FIGURE 1: Charts for Observed FTSE Values

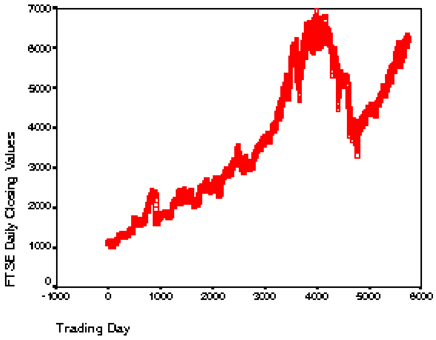

Fig. 1 (a): Scatter Plot for FTSE 100

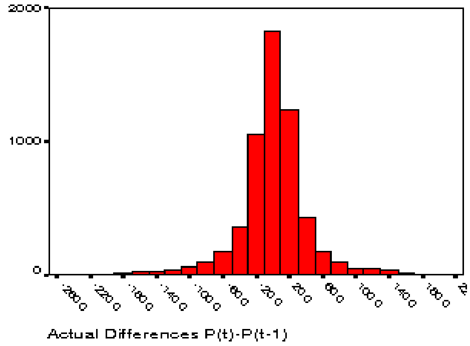

Fig. 1 (b): Histogram for FTSE 100

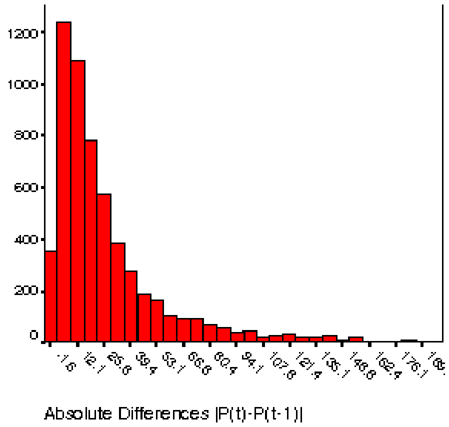

Fig. 1 (c): Histogram for FTSE 100

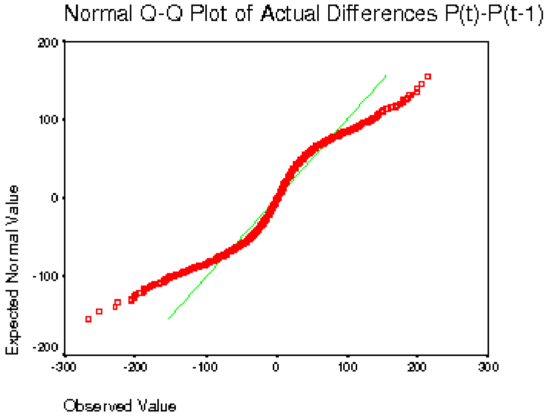

Fig. 1 (d): Q-Q Plots for FTSE 100

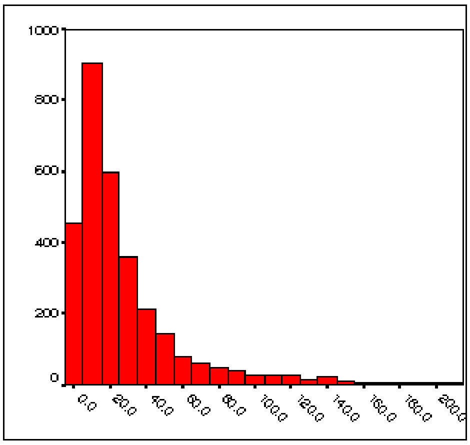

Fig. 1 (e): Histogram: P(t)-P(t-1)>0

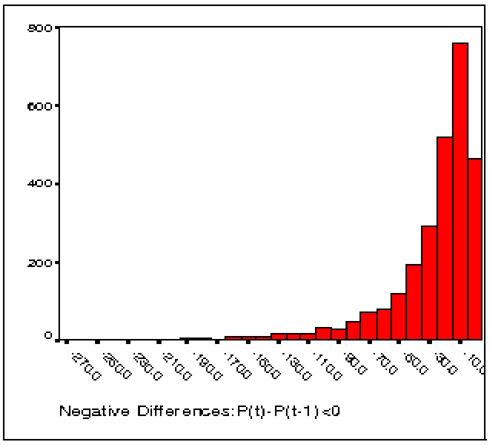

Fig. 1 (e): Histogram: P(t)-P(t-1)<0

FIGURE 2: Charts for Observed S&P 500 Values

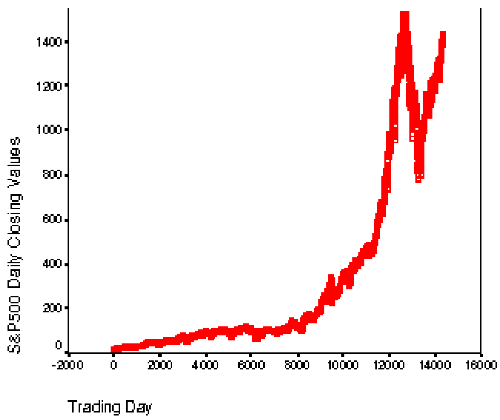

Fig. 2 (a): Scatter Plot for S&P 500 (Observed Values)

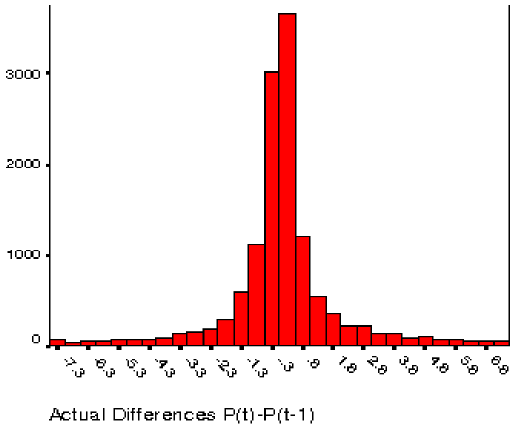

Fig. 2 (b): Histogram for S&P 500

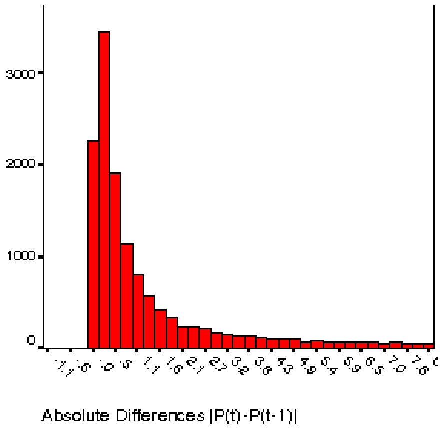

Fig. 2 (c): Histogram for S&P 500

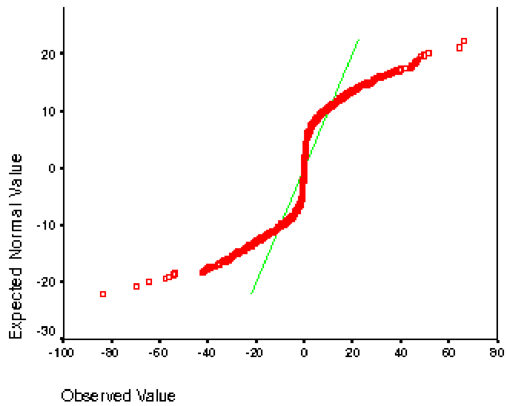

Fig. 2 (d): Q-Q Plots for S&P 500

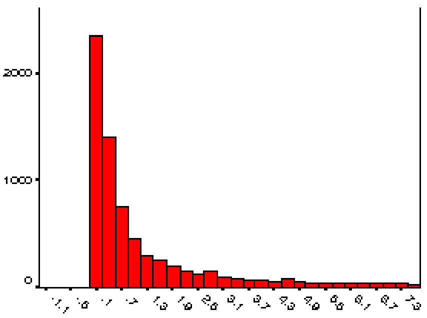

Fig. 2 (e): Histogram: P(t)-P(t-1)>0

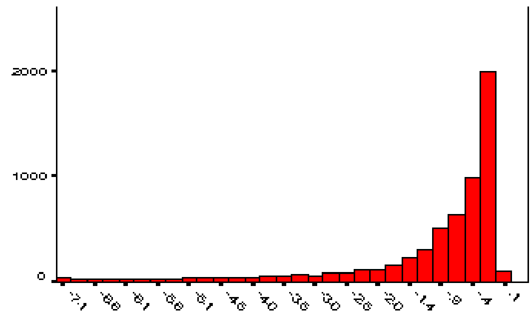

Fig. 2 (e): Histogram: P(t)-P(t-1)<0

FIGURE 3: Charts for Observed Nikkei Values

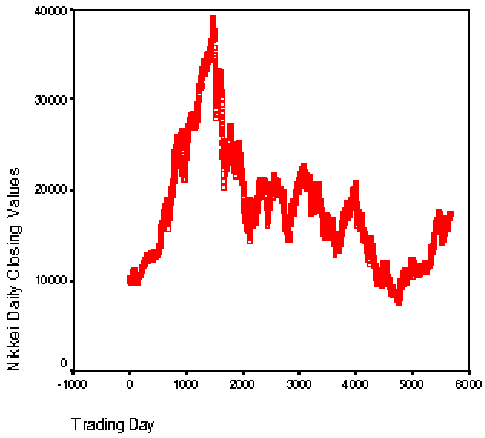

Fig. 3 (a): Scatter Plot for Nikkei (Observed Values)

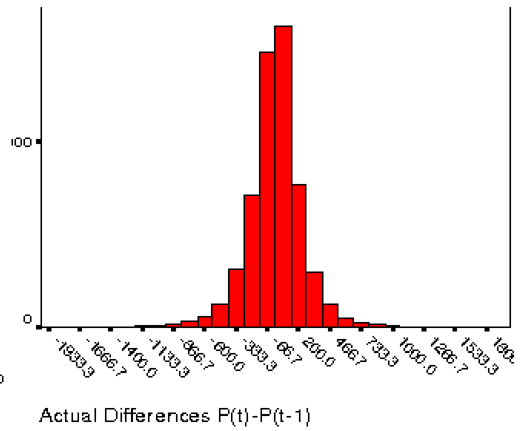

Fig. 3 (b): Histogram for Nikkei

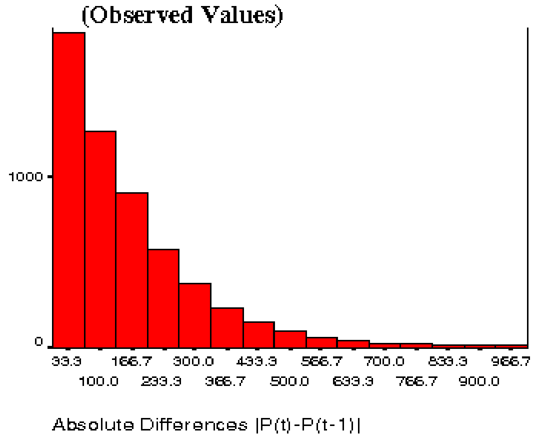

Fig. 3 (c): Histogram for Nikkei

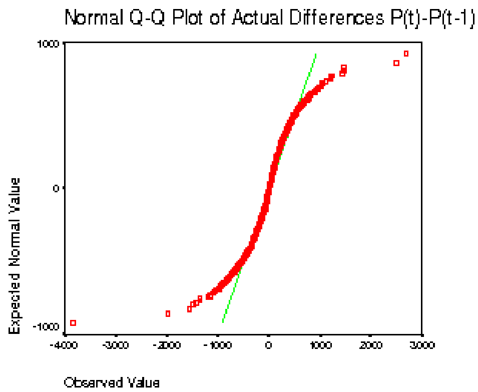

Fig. 3 (d): Q-Q Plots for Nikkei

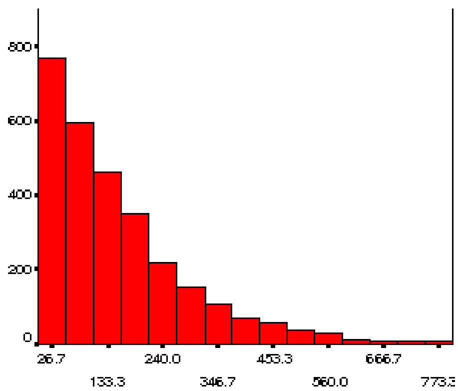

Fig. 3 (e): Histogram: P(t)-P(t-1)>0

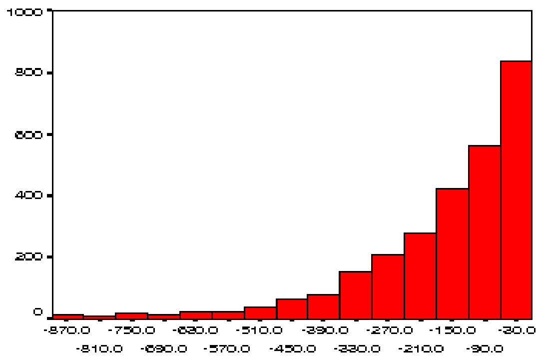

Fig. 3 (f): Histogram: P(t)-P(t-1)<0

FIGURE 4: Charts for Observed Glaxosmithkline Values

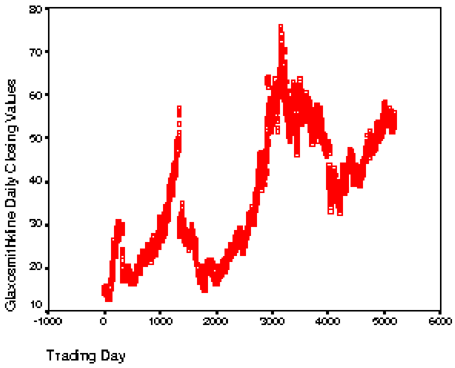

Fig. 4 (a): Scatter Plot for Glaxo. (Observed Values)

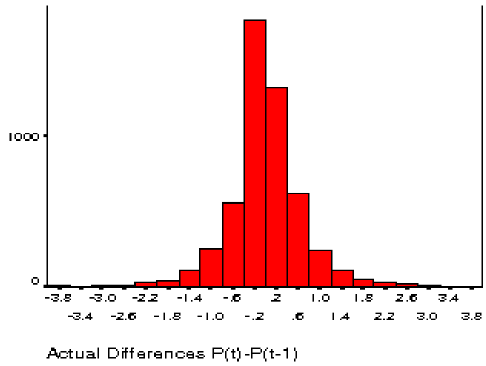

Fig. 4 (b): Histogram for Glaxo.

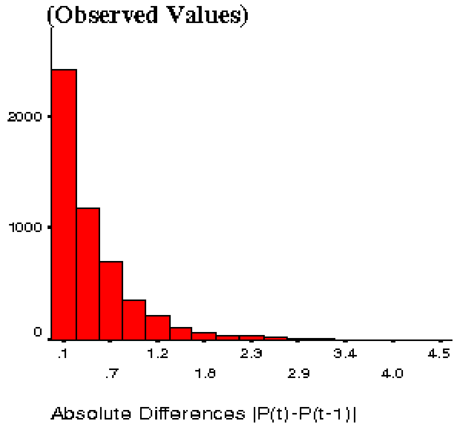

Fig. 4 (c): Histogram for Glaxo.

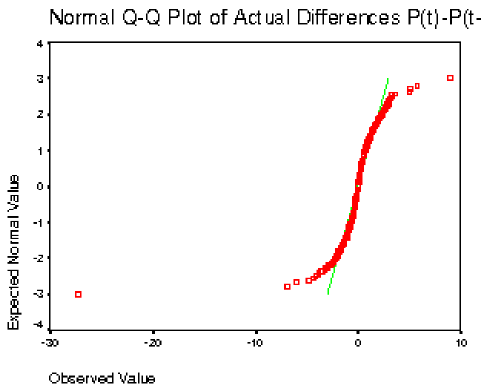

Fig. 4 (d): Q-Q Plots for Glaxo.

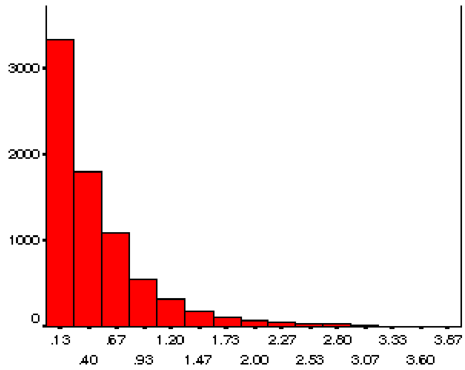

Fig. 4 (e): Histogram: P(t)-P(t-1)>0

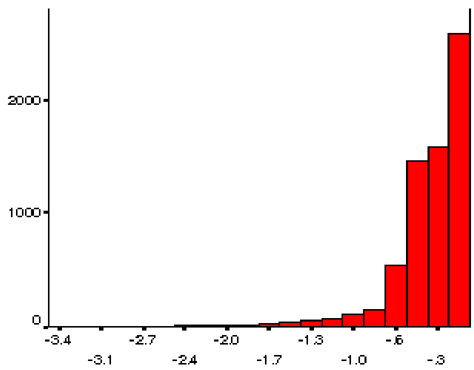

Fig. 4 (f): Histogram: P(t)-P(t-1)<0

FIGURE 5: Charts for Observed Microsoft Corporation Values

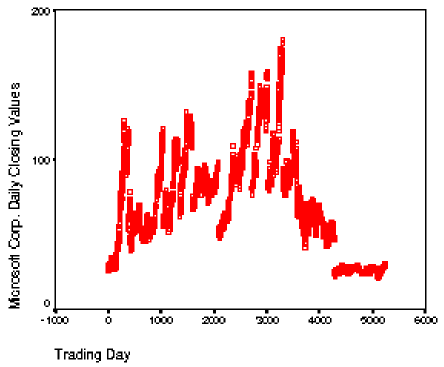

Fig. 5 (a): Scatter Plot for Microsoft Corp.

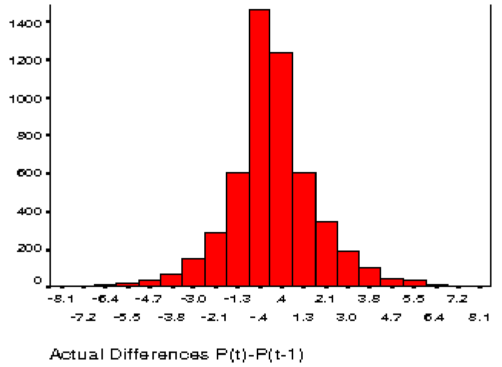

Fig. 5 (b): Histogram for Microsoft Corp.

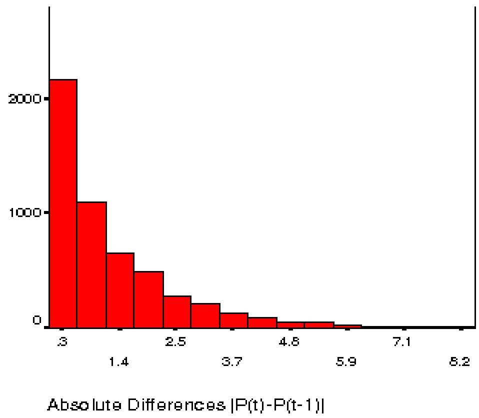

Fig. 5 (c): Histogram for Microsoft Corp.

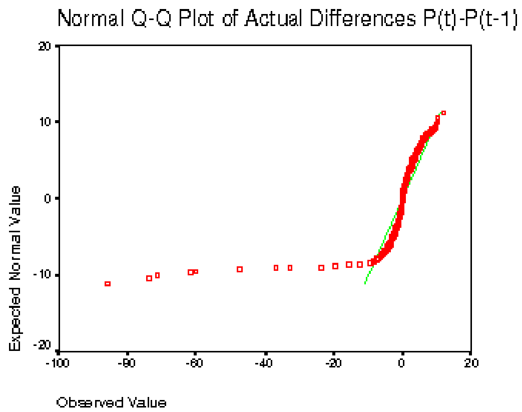

Fig. 5 (d): Q-Q Plots for Microsoft Corp.

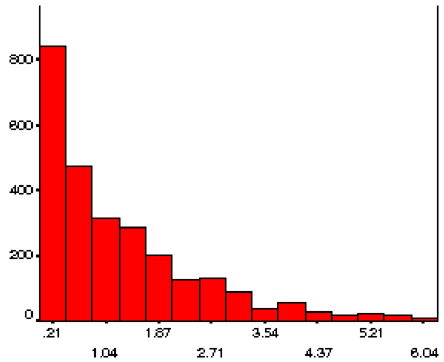

Fig. 5 (e): Histogram: P(t)-P(t-1)>0

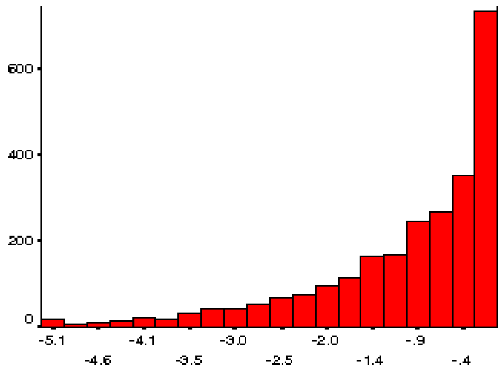

Fig. 5 (f): Histogram: P(t)-P(t-1)<0

FIGURE 6: Charts for Observed General Motors Values

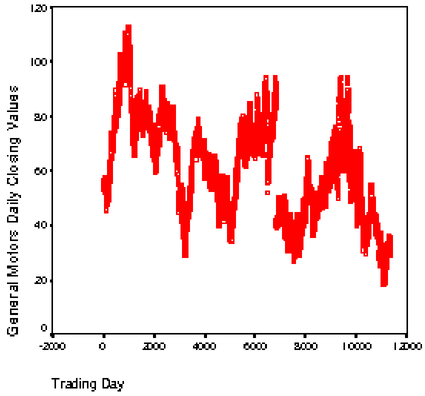

Fig. 6 (a): Scatter Plot for General Motors (Observed Values)

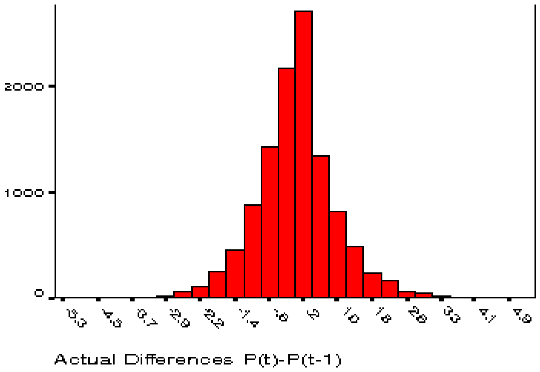

Fig. 6 (b): Histogram for Gen. Motors

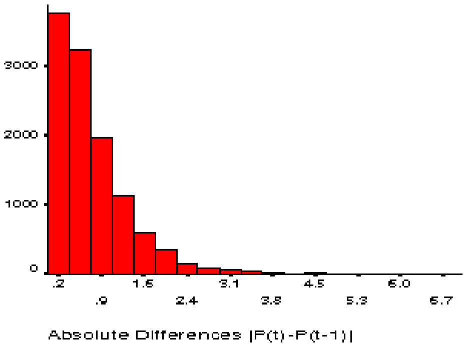

Fig. 6 (c): Histogram for Gen. Motors

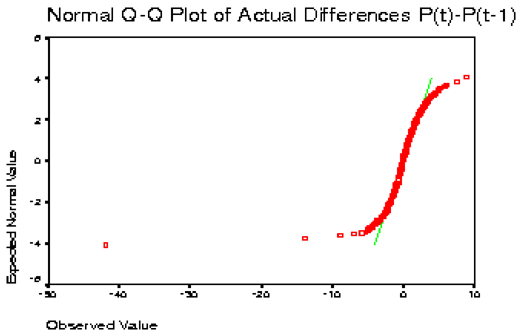

Fig. 6 (d): Q-Q Plots for Gen. Motors

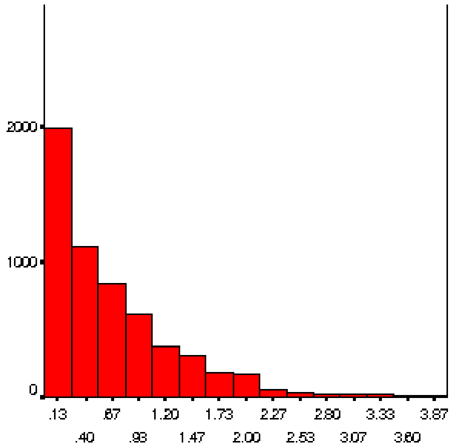

Fig. 6 (e): Histogram: P(t)-P(t-1)>0

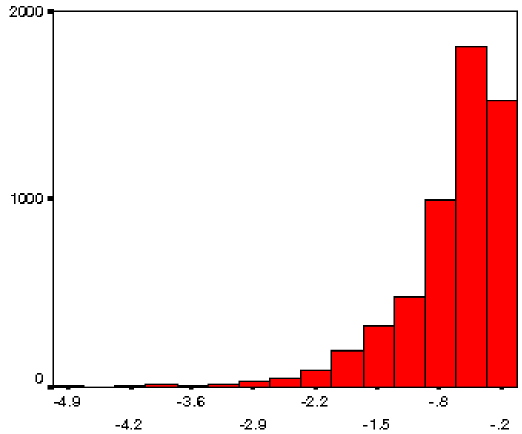

Fig.6 (f): Histogram: P(t)-P(t-1)<0

FIGURE 7: Charts for Observed Sony Corporation Values

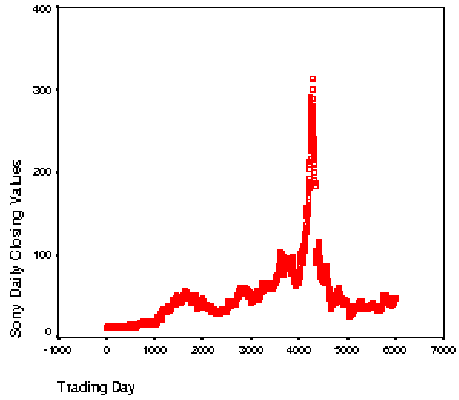

Fig. 7 (a): Scatter Plot for Sony Corp (Observed Values)

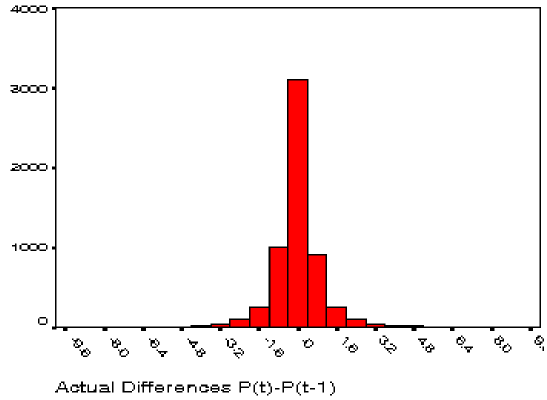

Fig. 7 (b): Histogram for Sony Corp.

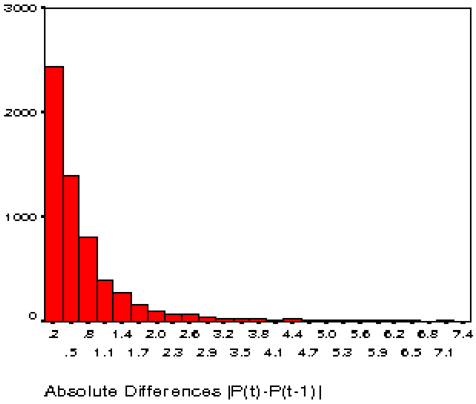

Fig. 7 (c): Histogram for Sony Corp.

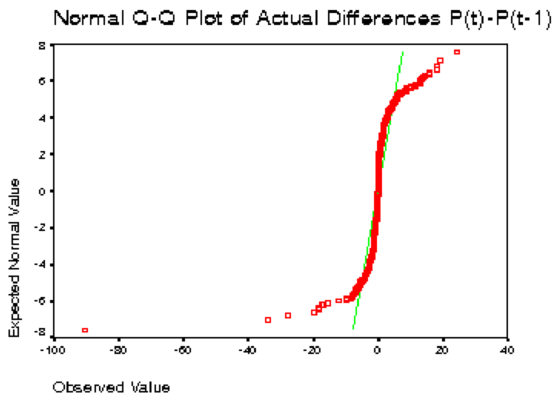

Fig. 7 (d): Q-Q Plots for Sony Corp.

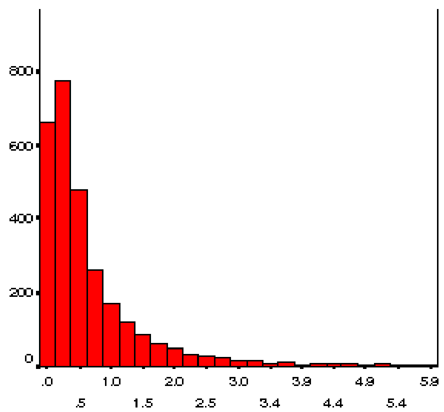

Fig. 7 (e): Histogram: P(t)-P(t-1)>0

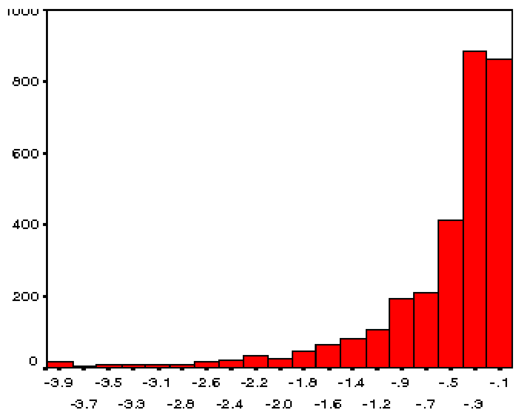

Fig. 7 (f): Histogram: P(t)-P(t-1)<0

FIGURE 8: Charts for Observed Toyota Corporation Values

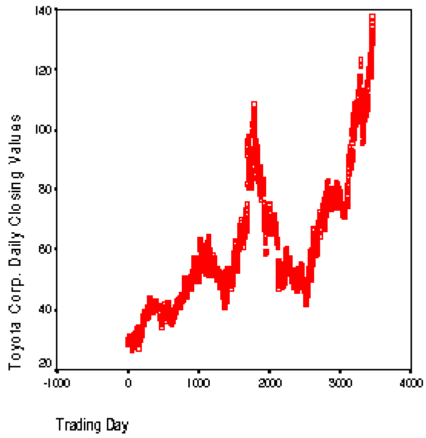

Fig. 8 (a): Scatter Plot for Toyota Corp

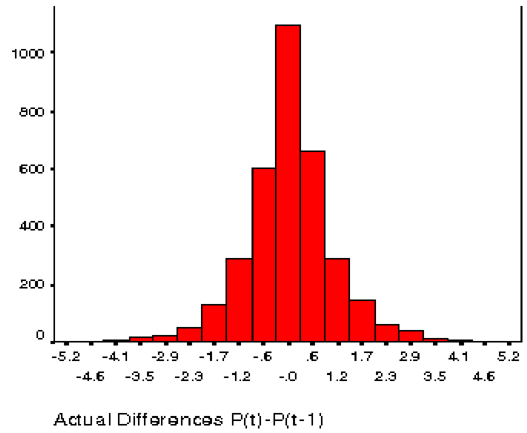

Fig. 8 (b): Histogram for Toyota Corp.

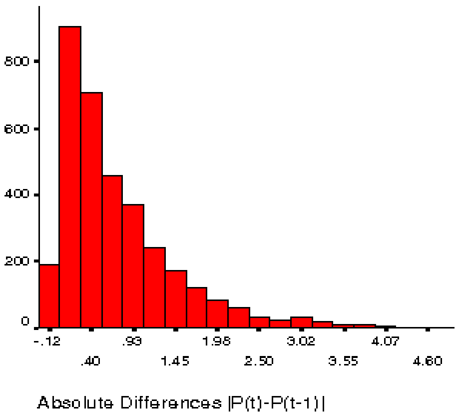

Fig. 8 (c): Histogram for Toyota Corp.

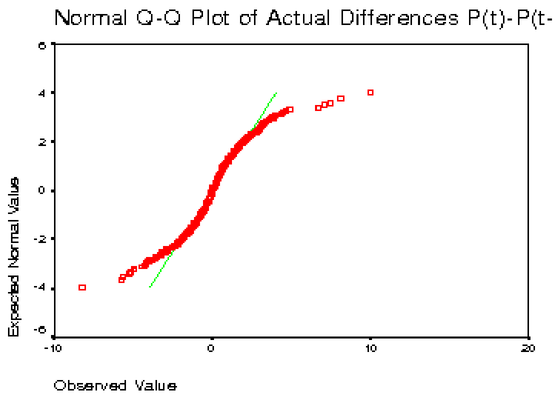

Fig. 8 (d): Q-Q Plots for Toyota Corp.

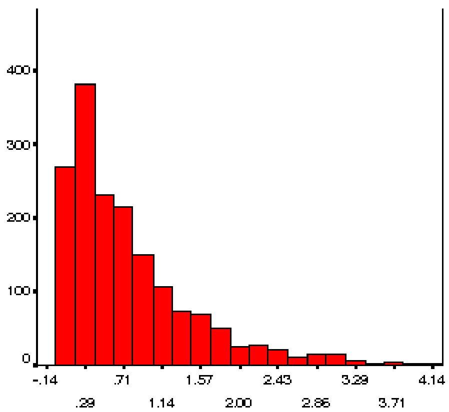

Fig. 8 (e): Histogram: P(t)-P(t-1)>0

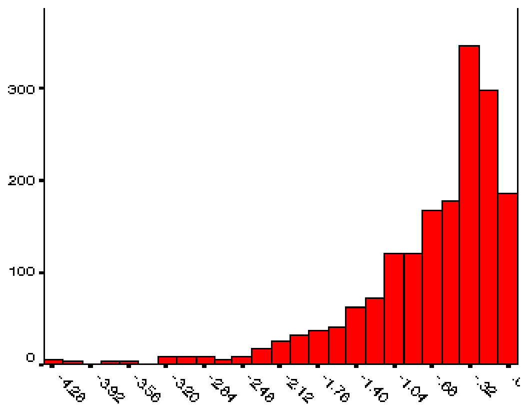

Fig. 8 (f): Histogram: P(t)-P(t-1)<0

FIGURE 9: Charts for Observed Barclays plc Values

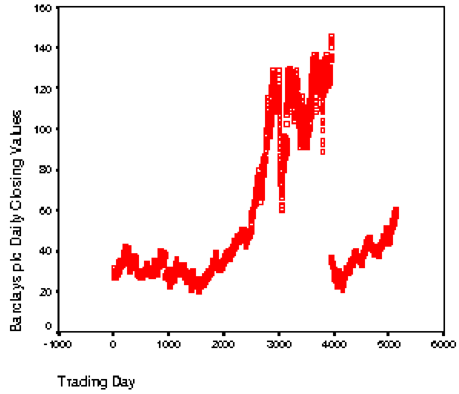

Fig. 9 (a): Scatter Plot for Barclays plc

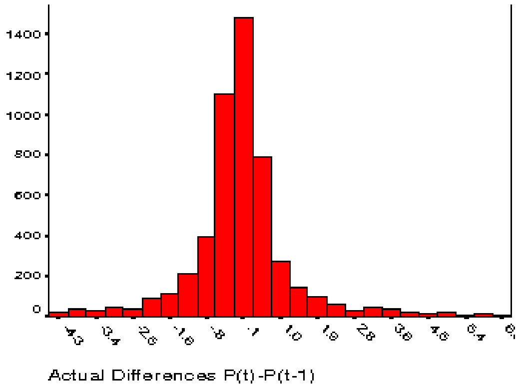

Fig. 9 (b): Histogram for Barclays plc.

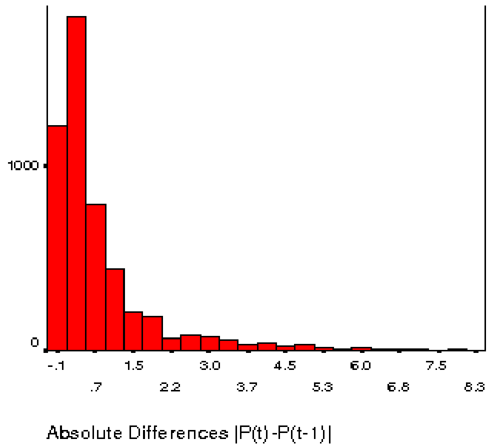

Fig. 9 (c): Histogram for Barclays plc.

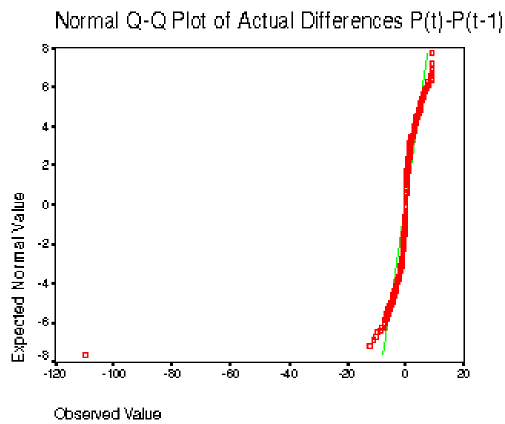

Fig. 9 (d): Q-Q Plots for Barclays plc.

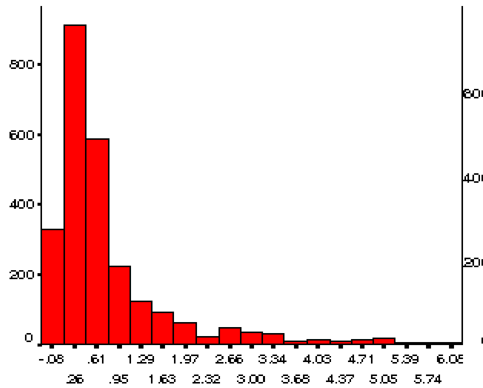

Fig. 9 (e): Histogram: P(t)-P(t-1)>0

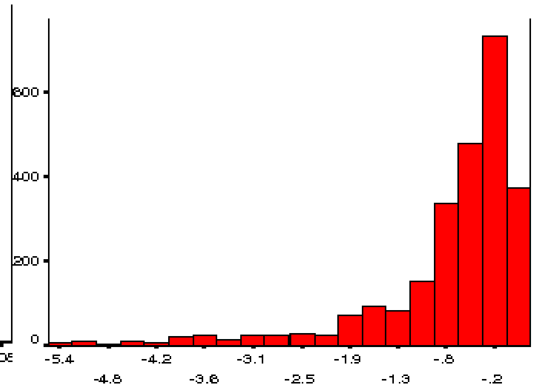

Fig. 9 (f): Histogram: P(t)-P(t-1)<0

APPENDIX B

Frequency of price changes

Index/Stock	Number of trading days	Frequency of 'up' movements	Frequency of 'down' movements	Frequency of 'no' movement
FTSE	5,761	3027	2717	17
Nikkei 225	5,672	2,910	2,749	13
SP 500	14,353	7,564	6,665	124
Sony Corp.	5,997	2,713	2,877	407
Toyota Corp.	3,464	1,667	1,618	179
Microsoft Corp.	5,261	2,597	2,491	173
General Motors	11,340	5,212	5,426	702
GlaxoSmithKline plc	5,179	2,394	2,374	411
Barclays plc	5,136	2,348	2,202	586

APPENDIX C

Simulation Results

Table S1: FTSE Simulation Results

FTSE: Actual P_t on 19/01/07 = 6237.2						
Simulation	$\hat{P}_t$ App. 1	Error App. 1	$\hat{P}_t$ App. 2	Error App. 2	$\hat{P}_t$ App. 3	Error App. 3
1.0	6236.9	0.3	6231.0	6.2	6297.51	60.31
2.0	6237.2	0.0	6229.7	7.5	6189.2	48
3.0	6235.9	1.3	6235.4	1.8	6343.2	106
4.0	6237.5	0.3	6231.7	5.5	6228.4	8.8
5.0	6238.0	0.8	6232.8	4.4	6279.2	42
6.0	6239.6	2.4	6232.7	4.5	6426.8	189.6
7.0	6236.7	0.5	6230.4	6.8	6214.7	22.5
8.0	6238.2	1.0	6232.6	4.6	6300.3	63.1
9.0	6237.7	0.5	6231.7	5.5	6242.3	5.1
10.0	6236.9	0.3	6231.4	5.8	6263.6	26.4
11.0	6239.7	2.5	6232.5	4.7	6243.8	6.6
12.0	6239.8	2.6	6231.1	6.1	6296.4	59.19
13.0	6237.8	0.6	6231.3	5.9	6242.6	5.4
14.0	6235.8	1.4	6234.2	3.0	6303.4	66.2
15.0	6238.4	1.2	6231.4	5.8	6299.3	62.1
16.0	6238.8	1.6	6232.6	4.6	6275.6	38.4
17.0	6239.4	2.2	6232.6	4.6	6292.6	55.4
18.0	6235.3	1.9	6229.6	7.6	6074.0	163.2
19.0	6236.3	0.9	6230.6	6.6	6232.8	4.4
20.0	6237.7	0.5	6231.9	5.3	6390.6	153.4
21.0	6238.2	1.0	6230.9	6.3	6284.8	47.6
22.0	6238.3	1.1	6230.7	6.5	6112.0	125.2
23.0	6238.1	0.9	6231.8	5.4	6312.0	74.8
24.0	6237.8	0.6	6232.6	4.6	6236.7	0.5
25.0	6238.3	1.1	6233.4	3.8	6236.7	0.5
26.0	6238.6	1.4	6231.0	6.2	6243.8	6.56
27.0	6240.8	3.6	6233.4	3.8	6133.9	103.3
28.0	6236.0	1.2	6228.9	8.3	6163.0	74.2
29.0	6239.2	2.0	6232.8	4.4	6245.6	8.4
30.0	6236.2	1.0	6230.2	7.0	6297.1	59.9
Mean Values	**6237.8**	**1.2**	**6231.8**	**5.4**	**6256.7**	**56.2**

Table S2: S&P 500 Simulation Results

S&P 500: Actual P_t on 19/01/07 = 1430.5						
Simulation	$\hat{P}_t$ **App. 1**	**Error App. 1**	$\hat{P}_t$ **App. 2**	**Error App. 2**	$\hat{P}_t$ **App. 3**	**Error App. 3**
1.0	1439.0	8.5	1427.0	3.5	1430.9	0.4
2.0	1440.5	10.0	1429.4	1.1	1403.5	27
3.0	1438.9	8.4	1426.2	4.3	1434.5	4
4.0	1439.2	8.7	1427.2	3.3	1419.6	10.9
5.0	1438.9	8.4	1427.5	3.0	1424.5	6
6.0	1440.2	9.7	1427.7	2.8	1409.4	21.1
7.0	1440.6	10.1	1429.7	0.8	1420.7	9.8
8.0	1440.8	10.3	1428.8	1.7	1455.1	24.6
9.0	1437.8	7.3	1423.8	6.7	1413.6	16.9
10.0	1438.9	8.4	1426.2	4.3	1418.6	11.9
11.0	1438.7	8.2	1424.3	6.2	1446.0	15.5
12.0	1439.6	9.1	1426.8	3.7	1413.2	17.3
13.0	1438.6	8.1	1427.4	3.1	1443.3	12.8
14.0	1439.3	8.8	1427.3	3.2	1395.9	34.6
15.0	1440.5	10.0	1429.4	1.1	1411.7	18.8
16.0	1440.2	9.7	1429.4	1.1	1422.5	8
17.0	1441.1	10.6	1429.1	1.4	1420.0	10.5
18.0	1439.2	8.7	1427.3	3.2	1423.6	6.9
19.0	1440.5	10.0	1430.0	0.5	1420.8	9.7
20.0	1438.0	7.5	1424.3	6.2	1430.1	0.4
21.0	1440.1	9.6	1428.9	1.6	1447.2	16.7
22.0	1440.9	10.4	1430.1	0.4	1419.8	10.7
23.0	1437.8	7.3	1424.1	6.4	1441.8	11.3
24.0	1440.2	9.7	1429.5	1.0	1410.6	19.9
25.0	1440.3	9.8	1427.3	3.2	1426.5	4
26.0	1439.7	9.2	1428.0	2.5	1425.6	4.9
27.0	1437.7	7.2	1424.6	5.9	1433.3	2.8
28.0	1441.6	11.1	1429.9	0.6	1415.6	14.9
29.0	1440.9	10.4	1428.7	1.8	1443.4	12.9
30.0	1441.6	11.1	1430.9	0.4	1438.1	7.6
Mean Values	**1439.7**	**9.2**	**1427.7**	**2.8**	**1425.3**	**12.4**

Table S3: GlaxoSmithkline plc Simulation Results

GlaxoSmithkline plc: Actual P_t on 19/01/07 = 55.95						
Simulation	$\hat{P}_t$ App. 1	Error App. 1	$\hat{P}_t$ App. 2	Error App. 2	$\hat{P}_t$ App. 3	Error App. 3
1.0	81.1	25.2	56.4	0.4	56.5	0.55
2.0	32.6	23.4	27.5	28.5	57.7	1.75
3.0	47.9	8.1	43.4	12.6	55.1	0.85
4.0	161.6	105.7	136.7	80.8	57.3	1.35
5.0	92.6	36.7	77.3	21.4	55.0	0.95
6.0	86.8	30.9	55.2	0.8	55.0	0.95
7.0	79.2	23.3	56.4	0.4	54.6	1.35
8.0	47.3	8.7	42.8	13.2	57.0	1.05
9.0	51.8	4.2	46.0	10.0	55.9	0.05
10.0	98.3	42.4	81.1	25.2	55.3	0.65
11.0	44.7	11.3	35.1	20.9	55.1	0.85
12.0	56.4	35.1	49.8	6.2	55.4	0.55
13.0	39.0	17.0	30.0	26.0	57.2	1.25
14.0	148.2	92.3	134.8	78.9	56.8	0.85
15.0	242.1	186.2	221.0	165.1	55.8	0.15
16.0	84.9	29.0	65.8	9.8	56.1	0.15
17.0	119.4	63.5	100.3	44.4	55.3	0.65
18.0	96.4	40.5	73.4	17.5	57.2	1.25
19.0	136.7	80.8	125.2	69.3	55.2	0.75
20.0	19.8	36.2	14.7	41.3	58.1	2.15
21.0	27.5	28.5	23.0	33.0	54.5	1.45
22.0	44.1	11.9	40.9	15.1	55.0	0.95
23.0	130.9	75.0	102.2	46.3	55.9	0.05
24.0	30.0	26.0	24.3	31.7	55.9	0.05
25.0	48.6	7.4	44.1	11.9	55.4	0.55
26.0	69.6	13.7	53.0	3.0	57.5	1.55
27.0	88.8	32.9	63.9	8.0	55.9	0.05
28.0	86.8	30.9	65.8	9.8	54.5	1.45
29.0	184.6	128.7	169.3	113.4	55.6	0.35
30.0	205.7	149.8	178.8	122.9	55.6	0.35
Mean Values	**89.4**	**46.8**	**74.6**	**35.6**	**55.9**	**0.8**

Table S4: **Microsoft Corporation** Simulation Results

Microsoft Corporation: Actual P_t on 19/01/07 = 31.11						
Simulation	$\hat{P}_t$ App. 1	Error App. 1	$\hat{P}_t$ App. 2	Error App. 2	$\hat{P}_t$ App. 3	Error App. 3
1.0	31.8	0.7	34.0	2.9	33.3	2.19
2.0	32.3	1.2	31.2	0.1	32.1	0.99
3.0	32.7	1.6	32.0	0.9	34.9	3.79
4.0	29.6	1.5	29.9	1.2	33.3	2.19
5.0	32.1	1.0	30.1	1.0	35.7	4.59
6.0	35.5	4.4	37.3	6.2	34.9	3.79
7.0	32.8	1.7	32.4	1.3	35.4	4.29
8.0	32.6	1.5	32.3	1.2	35.1	3.99
9.0	34.2	3.1	34.8	3.7	33.5	2.39
10.0	35.0	3.9	37.4	6.3	34.4	3.29
11.0	32.9	1.8	32.2	1.1	33.0	1.89
12.0	33.2	35.1	32.9	1.8	32.4	1.29
13.0	31.8	0.7	30.8	0.3	32.8	1.69
14.0	36.4	5.3	41.5	10.4	35.6	4.49
15.0	33.7	2.6	33.9	2.8	32.6	1.49
16.0	33.0	1.9	33.0	1.9	33.7	2.59
17.0	30.3	0.8	30.4	0.7	34.4	3.29
18.0	33.6	2.5	33.1	2.0	32.3	1.19
19.0	29.1	2.0	28.9	2.2	33.8	2.69
20.0	34.2	3.1	34.8	3.7	34.2	3.09
21.0	31.3	0.2	30.0	1.1	34.6	3.49
22.0	34.6	3.5	38.0	6.9	34.1	2.99
23.0	34.9	3.8	36.7	5.6	33.8	2.69
24.0	33.7	2.6	32.6	1.5	32.4	1.29
25.0	32.9	1.8	32.7	1.6	32.8	1.69
26.0	28.6	2.5	28.8	2.3	33.0	1.89
27.0	34.4	3.3	35.6	4.5	32.6	1.49
28.0	34.0	2.9	33.4	2.3	33.0	1.89
29.0	36.4	5.3	40.1	9.0	33.4	2.29
30.0	33.7	2.6	33.1	2.0	32.2	1.09
Mean Values	**33.0**	**3.5**	**33.5**	**2.9**	**33.6**	**2.5**

Table S5: **General Motors Simulation Results**

General Motors: Actual P_t on 19/01/07 = 31.55						
Simulation	$\hat{P}_t$ **App. 1**	**Error App. 1**	$\hat{P}_t$ **App. 2**	**Error App. 2**	$\hat{P}_t$ **App. 3**	**Error App. 3**
1.0	-82.1	113.7	31.6	0.1	31.8	0.25
2.0	-72.8	104.4	50.4	18.9	31.3	0.25
3.0	-59.9	91.5	42.0	10.5	30.4	1.15
4.0	-134.0	165.6	-24.7	56.3	32.3	0.75
5.0	-95.1	126.7	8.7	22.9	31.9	0.35
6.0	-15.4	47.0	180.9	149.4	32.5	0.95
7.0	-76.6	108.2	29.1	2.5	33.0	1.45
8.0	-41.3	72.9	164.3	132.8	31.8	0.25
9.0	26.3	5.3	367.1	335.6	31.9	0.35
10.0	-55.3	86.9	36.5	5.0	33.5	1.95
11.0	-109.0	140.6	11.4	20.2	31.5	0.05
12.0	-99.7	35.1	-8.0	39.6	32.1	0.55
13.0	-125.6	157.2	-44.1	75.7	31.3	0.25
14.0	-58.0	89.6	75.3	43.8	32.4	0.85
15.0	-60.8	92.4	86.5	55.0	31.6	0.05
16.0	-60.8	92.4	89.2	57.7	31.4	0.15
17.0	-78.4	110.0	11.4	20.2	32.6	1.05
18.0	-56.2	87.8	72.6	41.1	32.7	1.15
19.0	-123.8	155.4	-23.8	55.4	31.3	0.25
20.0	-94.1	125.7	17.9	13.7	31.6	0.05
21.0	-114.5	146.1	-12.7	44.3	32.1	0.55
22.0	-72.8	104.4	31.6	0.1	32.5	0.95
23.0	-182.1	213.7	-82.1	113.7	32.3	0.75
24.0	-110.8	142.4	-9.9	41.5	32.4	0.85
25.0	-149.7	181.3	-72.8	104.4	32.3	0.75
26.0	-73.8	105.4	39.3	7.8	31.8	0.25
27.0	-82.1	113.7	28.2	3.4	32.1	0.55
28.0	-84.9	116.5	27.2	4.4	32.2	0.65
29.0	-83.0	114.6	2.2	29.4	30.7	0.85
30.0	-38.6	70.2	144.8	113.3	31.8	0.25
Mean Values	**-82.2**	**110.5**	**42.3**	**53.9**	**32.0**	**0.6**

Table S6: Sony Corporation Simulation Results

Sony Corporation: Actual P_t on 19/01/07 = 47.06						
Simulation	$\hat{P}_t$ App. 1	Error App. 1	$\hat{P}_t$ App. 2	Error App. 2	$\hat{P}_t$ App. 3	Error App. 3
1.0	15.3	31.8	49.7	2.6	47.9	0.84
2.0	14.8	32.3	-41.9	89.0	50.1	3.04
3.0	14.5	32.6	-41.9	89.0	48.2	1.14
4.0	16.1	31.0	-38.6	85.7	47.8	0.74
5.0	14.8	32.3	-39.1	86.2	48.9	1.84
6.0	15.4	31.7	-39.2	86.3	47.4	0.34
7.0	15.4	31.7	-39.3	86.4	47.8	0.74
8.0	17.3	29.8	-34.3	81.4	47.0	0.06
9.0	14.9	32.2	-39.8	86.9	46.6	0.46
10.0	17.0	30.1	-33.2	80.3	47.9	0.84
11.0	16.1	31.0	-36.8	83.9	48.9	1.84
12.0	14.5	32.6	-42.1	89.2	47.8	0.74
13.0	13.8	33.3	-43.4	90.5	48.1	1.04
14.0	16.2	30.9	45.3	1.8	47.3	0.24
15.0	15.1	32.0	-37.7	84.8	49.2	2.14
16.0	16.5	30.6	-37.0	84.1	48.9	1.84
17.0	15.2	31.9	-41.3	88.4	49.0	1.94
18.0	45.5	1.6	48.2	1.1	48.4	1.34
19.0	15.5	31.6	-32.2	79.3	47.7	0.64
20.0	15.1	32.0	-39.5	86.6	48.0	0.94
21.0	15.3	31.8	-40.2	87.3	47.6	0.54
22.0	13.7	33.4	-44.3	91.4	46.7	0.36
23.0	15.3	31.8	-40.2	87.3	48.4	1.34
24.0	16.0	31.1	-36.5	83.6	47.4	0.34
25.0	14.6	32.5	-41.4	88.5	47.6	0.54
26.0	15.4	31.7	-38.4	85.5	46.7	0.36
27.0	14.9	32.2	-40.9	88.0	49.0	1.94
28.0	13.3	33.8	-44.5	91.6	49.4	2.34
29.0	15.5	31.6	-38.9	86.0	47.8	0.74
30.0	16.0	31.1	-37.7	84.8	49.1	2.04
Mean Values	**16.3**	**30.8**	**-30.6**	**77.9**	**48.1**	**1.1**

Table S7: **Toyota Motor Corporation** Simulation Results

Toyota Motor Corporation: Actual P_t on 19/01/07 = 132.33						
Simulation	$\hat{P}_t$ **App. 1**	**Error App. 1**	$\hat{P}_t$ **App. 2**	**Error App. 2**	$\hat{P}_t$ **App. 3**	**Error App. 3**
1.0	247.5	115.2	155.0	22.7	131	1.33
2.0	256.5	124.2	172.9	40.6	131.0	1.33
3.0	193.8	61.5	125.2	7.1	135.5	3.17
4.0	211.7	79.4	134.3	2.0	133.6	1.27
5.0	-224.9	357.2	-260.7	393.0	131.5	0.83
6.0	-259.7	392.0	-279.6	411.9	132.9	0.57
7.0	-236.8	369.1	-258.7	391.0	130.4	1.93
8.0	-267.6	399.9	-294.0	426.3	132.4	0.07
9.0	-261.7	394.0	-282.6	414.9	125.0	7.33
10.0	-303.4	435.7	331.3	199.0	132.5	0.17
11.0	229.6	97.3	146.0	13.7	132.6	0.27
12.0	129.2	35.1	110.2	22.1	131.5	0.83
13.0	105.3	27.0	82.4	49.9	130.4	1.93
14.0	211.7	79.4	137.1	4.8	130.8	1.53
15.0	399.8	267.5	322.1	189.8	130.5	1.83
16.0	157.9	25.6	112.2	20.1	130.3	2.03
17.0	328.1	195.8	211.7	79.4	131.5	0.83
18.0	223.6	91.3	137.1	4.8	134.1	1.77
19.0	163.9	31.6	120.2	12.1	130.3	2.03
20.0	126.2	6.1	104.3	28.0	133.9	1.57
21.0	178.8	46.5	118.2	14.1	133.8	1.47
22.0	187.8	55.5	126.2	6.1	130.2	2.13
23.0	301.3	169.0	220.6	88.3	131.5	0.83
24.0	134.3	2.0	112.2	20.1	132.6	0.27
25.0	172.9	40.6	105.3	27.0	134.1	1.77
26.0	271.4	139.1	223.6	91.3	130.7	1.63
27.0	256.5	124.2	187.8	55.5	132.5	0.17
28.0	190.8	58.5	126.2	6.1	131.8	0.53
29.0	181.8	49.5	129.2	3.1	131.1	1.23
30.0	118.2	14.1	103.3	29.0	131.9	0.43
Mean Values	**114.2**	**142.8**	**82.6**	**102.5**	**131.7**	**1.4**

Table S8: **Barclays plc Simulation Results**

Barclays plc: Actual P_t on 19/01/07 = 59.4						
Simulation	$\hat{P}_t$ **App. 1**	**Error App. 1**	$\hat{P}_t$ **App. 2**	**Error App. 2**	$\hat{P}_t$ **App. 3**	**Error App. 3**
1.0	63.8	4.4	63.3	3.9	59.2	0.2
2.0	60.1	0.7	57.3	2.1	57.2	2.2
3.0	60.9	1.5	56.8	2.6	59.5	0.1
4.0	62.8	3.4	60.5	1.1	59.5	0.1
5.0	61.8	2.4	58.3	1.1	56.4	3
6.0	64.6	5.2	64.4	5.0	58.9	0.5
7.0	61.6	2.2	58.2	1.2	57.4	2
8.0	58.6	0.8	56.0	3.4	55.7	3.7
9.0	63.8	4.4	63.2	3.8	55.9	3.5
10.0	65.1	5.7	65.9	6.5	59.7	0.3
11.0	68.7	9.3	76.4	17.0	57.5	1.9
12.0	63.4	35.1	60.0	0.6	56.3	3.1
13.0	63.4	4.0	61.9	2.5	57.7	1.7
14.0	61.5	2.1	58.7	0.7	56.2	3.2
15.0	62.6	3.2	60.5	1.1	57.1	2.3
16.0	66.6	7.2	69.1	9.7	60.8	1.4
17.0	61.2	1.8	57.8	1.6	59.8	0.4
18.0	62.4	3.0	59.3	0.1	57.9	1.5
19.0	63.7	4.3	59.7	0.3	59.6	0.2
20.0	61.7	2.3	58.6	0.8	60.3	0.9
21.0	64.2	4.8	63.7	4.3	57.8	1.6
22.0	59.8	0.4	57.3	2.1	56.2	3.2
23.0	63.6	4.2	62.0	2.6	54.6	4.8
24.0	60.3	0.9	57.0	2.4	59.2	0.2
25.0	63.0	3.6	59.9	0.5	58.8	0.6
26.0	63.0	3.6	59.9	0.5	58.8	0.6
27.0	63.4	4.0	60.6	1.2	57.6	1.8
28.0	61.2	1.8	58.1	1.3	59.0	0.4
29.0	63.8	4.4	62.7	3.3	55.4	4
30.0	61.3	1.9	57.5	1.9	57.5	1.9
Mean Values	**62.7**	**4.4**	**60.8**	**2.8**	**57.9**	**1.7**

Table S9: **Nikkei Simulation Results**

Nikkei: Actual P_t on 19/01/07 = 17310.44						
Simulation	$\hat{P}_t$ App. 1	Error App. 1	$\hat{P}_t$ App. 2	Error App. 2	$\hat{P}_t$ App. 3	Error App. 3
1.0	17430.5	120.1	17430.0	119.6	17815.2	504.76
2.0	17430.8	120.4	17430.2	119.8	17576.6	266.16
3.0	17430.2	119.8	1429.5	15880.9	17661.0	350.56
4.0	17430.1	119.7	17429.5	119.1	17521.6	211.16
5.0	17430.9	120.5	17430.3	119.9	17581.9	271.46
6.0	17429.7	119.3	17429.2	118.8	17732.2	421.76
7.0	17430.2	119.8	17429.4	119.0	17882.6	572.16
8.0	17430.6	120.2	17429.8	119.4	17693.1	382.66
9.0	17430.2	119.8	17429.4	119.0	17378.8	68.36
10.0	17430.1	119.7	17429.5	119.1	17992.5	682.06
11.0	17429.7	119.3	17429.1	118.7	17602.1	291.66
12.0	17430.2	35.1	17429.5	119.1	17895.8	585.34
13.0	17430.3	119.9	17429.7	119.3	17599.9	289.41
14.0	17429.7	119.3	17429.1	118.7	17185.1	125.34
15.0	17429.9	119.5	17429.4	119.0	17454.7	144.26
16.0	17429.9	119.5	17429.2	118.8	17519.5	209.06
17.0	17430.6	120.2	17429.9	119.5	17319.3	8.86
18.0	17430.2	119.8	17429.5	119.1	17722.9	412.46
19.0	17430.3	119.9	17429.6	119.2	17434.5	124.04
20.0	17430.4	120.0	17429.7	119.3	17795.4	484.99
21.0	17429.9	119.5	17429.3	118.9	17771.9	461.46
22.0	17424.2	113.8	17423.5	113.1	17557.0	246.56
23.0	17424.6	114.2	17423.9	113.5	17546.1	235.66
24.0	17424.5	114.1	17423.9	113.5	17521.4	210.96
25.0	17424.6	114.2	17424.0	113.6	17605.5	295.06
26.0	17424.8	114.4	17424.2	113.8	17625.2	314.76
27.0	17424.9	114.5	17424.4	114.0	17877.3	566.86
28.0	17424.5	114.1	17424.0	113.6	17616.1	305.66
29.0	17424.0	113.6	17423.3	112.9	17402.0	91.56
30.0	17425.4	115.0	17424.7	114.3	17368.6	58.16
Mean Values	**17428.5**	**115.3**	**16894.6**	**642.8**	**17608.5**	**306.4**

APPENDIX D

Relative Errors

Table R1: FTSE Relative Errors

P_t Actual	$\hat{P}_t$ App.1	$\hat{P}_t$ App. 2	$\hat{P}_t$ App. 3	RE App 1	RE App2	RE App3
6237.2	6216.9	6210.5	6344.3294	0.0032625	0.00428	0.017176
6210.3	6211.0	6204.7	6282.1187	0.0001208	0.0009	0.011564
6204.5	6222.2	6215.9	6224.914	0.0028606	0.00184	0.00329
6215.7	6270.0	6263.7	6256.4449	0.0087436	0.00773	0.006555
6263.5	6245.5	6239.2	6332.7401	0.0028664	0.00387	0.011055
6239	6236.6	6230.3	6159.9628	0.0003774	0.00139	0.012668
6230.1	6167.2	6160.9	6247.4782	0.010089	0.0111	0.002789
6160.7	6202.6	6196.3	6296.9702	0.0068082	0.00578	0.022119
6196.1	6200.7	6194.4	6302.364	0.0007492	0.00027	0.01715
6194.2	6226.6	6220.3	6206.5747	0.0052373	0.00422	0.001998
6220.1	6293.5	6287.2	6252.8488	0.0118068	0.01079	0.005265
6287	6325.5	6319.2	6311.1207	0.0061299	0.00513	0.003837
6319	6317.4	6311.1	6249.7564	0.0002473	0.00125	0.010958
6310.9	6227.3	6221.0	6300.8693	0.0132412	0.01424	0.001589
6220.8	6247.4	6241.1	6106.009	0.0042816	0.00327	0.018453
6240.9	6251.7	6245.4	6238.2508	0.001736	0.00073	0.000424
6245.2	6196.5	6190.2	6216.6218	0.0077927	0.0088	0.004576
6190	6190.2	6183.9	6202.6368	3.744E-05	0.00098	0.002041
6183.7	6205.1	6198.8	6185.1363	0.0034657	0.00245	0.000232
6198.6	6210.4	6204.1	6239.0154	0.0019084	0.00089	0.00652
6203.9	6253.9	6247.6	6256.0912	0.008064	0.00705	0.008413
6247.4	6266.5	6260.2	6205.8487	0.0030616	0.00205	0.006651
6260	6234.5	6228.2	6322.7726	0.0040693	0.00508	0.010028
6228	6199.0	6192.7	6165.982	0.0046524	0.00566	0.009958
6192.5	6162.9	6156.6	6127.9623	0.0047761	0.00579	0.010422
6156.4	6166.3	6160.0	6198.8049	0.0016118	0.00059	0.006888
6159.8	6158.9	6152.6	6109.0582	0.0001426	0.00116	0.008238
6152.4	6138.0	6131.7	6142.9817	0.0023372	0.00336	0.001531
6131.5	6096.8	6090.5	6130.6215	0.0056562	0.00668	0.000143
6090.3	6092.9	6086.6	6116.3008	0.0004299	0.0006	0.004269
6086.4	6056.9	6050.6	6085.4962	0.0048441	0.00588	0.000148
6050.4	6028.0	6021.7	6122.254	0.0036996	0.00474	0.011876
6021.5	6055.3	6049.0	6024.6734	0.0056157	0.00457	0.000527
6048.8	6090.9	6084.6	6044.5944	0.0069623	0.00592	0.000695

Table R2: S&P 500 Relative Errors

P_t Actual	$\hat{P}_t$ App.1	$\hat{P}_t$ App. 2	$\hat{P}_t$ App. 3	RE App 1	RE App2	RE App3
1430.5	1447.996	1430.728	1411.977	0.012231	0.00016	0.012948
1426.37	1452.245	1434.978	1431.717	0.01814	0.006035	0.003749
1430.62	1453.523	1436.258	1450.866	0.016009	0.003941	0.014152
1431.9	1452.352	1435.087	1428.868	0.014283	0.002226	0.002117
1430.73	1445.44	1428.177	1402.513	0.010282	0.001784	0.019722
1423.82	1436.469	1419.207	1422.358	0.008884	0.00324	0.001027
1414.85	1433.727	1416.466	1383.976	0.013342	0.001143	0.021822
1412.11	1434.456	1417.196	1387.108	0.015824	0.003602	0.017705
1412.84	1431.324	1414.066	1424.482	0.013083	0.000868	0.00824
1409.71	1439.953	1422.696	1427.625	0.021453	0.009212	0.012709
1418.34	1438.211	1420.955	1423.666	0.01401	0.001844	0.003755
1416.6	1439.91	1422.655	1414.093	0.016455	0.004274	0.00177
1418.3	1446.338	1429.085	1410.612	0.019769	0.007604	0.005421
1424.73	1448.447	1431.194	1436.119	0.016646	0.004537	0.007994
1426.84	1438.505	1421.254	1428.66	0.008176	0.003915	0.001275
1416.9	1432.364	1415.114	1433.359	0.010914	0.001261	0.011616
1410.76	1439.902	1422.653	1420.194	0.020657	0.008431	0.006687
1418.3	1445.131	1427.883	1430.699	0.018918	0.006757	0.008742
1423.53	1447.149	1429.903	1441.246	0.016592	0.004477	0.012445
1425.55	1444.078	1426.833	1423.681	0.012997	0.0009	0.001311
1422.48	1448.686	1431.442	1430.006	0.018423	0.0063	0.005291
1427.09	1447.085	1429.842	1425.57	0.014011	0.001928	0.001065
1425.49	1434.803	1417.562	1422.269	0.006533	0.005562	0.002259
1413.21	1433.152	1415.911	1448.127	0.014111	0.001911	0.024707
1411.56	1434.63	1417.391	1413.54	0.016344	0.004131	0.001403
1413.04	1431.429	1414.191	1403.834	0.013014	0.000814	0.006515
1409.84	1428.877	1411.64	1412.832	0.013503	0.001277	0.002122
1407.29	1434.486	1417.25	1402.969	0.019325	0.007078	0.003071
1412.9	1436.344	1419.11	1399.799	0.016593	0.004395	0.009273
1414.76	1430.703	1413.47	1405.68	0.011269	0.000912	0.006418
1409.12	1418.291	1401.059	1394.09	0.006508	0.00572	0.010666
1396.71	1422.21	1404.979	1390.051	0.018257	0.00592	0.004768
1400.63	1421.058	1403.829	1413.732	0.014585	0.002284	0.009354
1399.48	1408.297	1391.068	1384.975	0.0063	0.006011	0.010365

Table R3: GlaxoSmithkline plc Relative Errors

P_t Actual	$\hat{P}_t$ App.1	$\hat{P}_t$ App.2	$\hat{P}_t$ App. 3	RE App 1	RE App2	RE App3
55.95	29.555896	16.77828	56.588605	0.4717445	0.70012	0.011414
55.75	29.040954	16.26581	55.622399	0.479086	0.70824	0.002289
55.23	28.356012	15.58333	56.926357	0.4865832	0.71785	0.030714
54.54	27.57107	14.80086	52.881731	0.4944798	0.72862	0.030405
53.75	26.896127	14.12838	53.251743	0.4996069	0.73715	0.00927
53.07	26.561185	13.79591	52.63811	0.4995066	0.74004	0.008138
52.73	26.506243	13.74343	51.875498	0.4973214	0.73936	0.016205
52.67	26.501301	13.74096	53.49251	0.4968426	0.73911	0.015616
52.66	27.806358	15.04848	54.566445	0.4719643	0.71423	0.036203
53.96	28.571416	15.81601	55.364708	0.4705075	0.70689	0.026032
54.72	27.666474	14.91353	54.413714	0.4943992	0.72746	0.005597
53.81	26.621532	13.87106	53.578759	0.5052679	0.74222	0.004297
52.76	26.406589	13.65858	54.776284	0.499496	0.74112	0.038216
52.54	26.511647	13.76611	52.939826	0.4954007	0.73799	0.00761
52.64	26.236705	13.49363	52.525688	0.5015824	0.74366	0.002172
52.36	26.271763	13.53116	51.645703	0.4982475	0.74157	0.013642
52.39	26.02682	13.28868	51.428347	0.5032101	0.74635	0.018356
52.14	26.181878	13.44621	52.063562	0.4978543	0.74211	0.001466
52.29	26.316936	13.58373	52.017038	0.4967119	0.74022	0.00522
52.42	26.191994	13.46126	52.273391	0.5003435	0.7432	0.002797
52.29	25.837051	13.10878	52.060689	0.5058892	0.74931	0.004385
51.93	26.672109	13.94631	53.413886	0.4863834	0.73144	0.028575
52.76	26.827167	14.10383	53.719172	0.4915245	0.73268	0.01818
52.91	26.922225	14.20136	50.7198	0.4911694	0.73159	0.041395
53	26.357282	13.63888	51.683563	0.5026928	0.74266	0.024838
52.43	26.45234	13.73641	51.375094	0.4954732	0.738	0.02012
52.52	26.487398	13.77393	52.771856	0.4956703	0.73774	0.004795
52.55	26.352456	13.64146	53.213865	0.4985261	0.74041	0.012633
52.41	26.767513	14.05898	51.943482	0.4892671	0.73175	0.008901
52.82	27.282571	14.57651	52.075323	0.4834803	0.72403	0.014098
53.33	27.687629	14.98403	51.075545	0.4808245	0.71903	0.042274
53.73	27.092687	14.39156	52.331786	0.4957624	0.73215	0.026023
53.13	27.217744	14.51908	51.928236	0.4877142	0.72673	0.022619
53.25	25.832802	13.13661	50.517766	0.514877	0.7533	0.05131

Table R4: Microsoft Corporation Relative Errors

P_t Actual	$\hat{P}_t$ App.1	$\hat{P}_t$ App.2	$\hat{P}_t$ App. 3	RE App 1	RE App2	RE App3
31.11	36.323614	31.29343	30.406448	0.1675864	0.0059	0.022615
31	36.422602	31.39337	32.44348	0.1749227	0.01269	0.046564
31.1	36.481591	31.45332	32.021758	0.1730415	0.01136	0.029639
31.16	36.530579	31.50326	31.866176	0.1723549	0.01102	0.022663
31.21	36.019567	30.9932	28.854808	0.1541034	0.00695	0.075463
30.7	34.978555	29.95315	32.58519	0.1393666	0.02433	0.061407
29.66	35.277543	30.25309	29.522898	0.1893979	0.02	0.004622
29.96	35.246531	30.22304	28.681981	0.176453	0.00878	0.042658
29.93	34.955519	29.93298	29.784498	0.1679091	1E-04	0.004861
29.64	35.124507	30.10293	30.950276	0.1850374	0.01562	0.044206
29.81	35.173495	30.15287	30.773343	0.1799227	0.0115	0.032316
29.86	35.172483	30.15281	30.084067	0.177913	0.00981	0.007504
29.86	35.291472	30.27276	29.223358	0.1818979	0.01382	0.021321
29.98	35.33046	30.3127	29.310814	0.1784676	0.0111	0.022321
30.02	35.299448	30.28265	27.370455	0.1758643	0.00875	0.088259
29.99	34.948436	29.93259	30.622627	0.1653363	0.00191	0.021095
29.64	35.287424	30.27254	29.959415	0.1905339	0.02134	0.010776
29.98	35.396412	30.38248	31.003371	0.1806675	0.01342	0.034135
30.09	35.2954	30.28242	27.558435	0.1729944	0.00639	0.084133
29.99	35.194388	30.18237	29.369022	0.1735375	0.00641	0.020706
29.89	35.493376	30.48231	28.635819	0.1874666	0.01982	0.04196
30.19	35.372364	30.36226	28.064999	0.1716583	0.00571	0.070388
30.07	34.851353	29.8422	29.335848	0.1590074	0.00758	0.024415
29.55	34.730341	29.72214	28.875966	0.1753076	0.00583	0.02281
29.43	34.839329	29.83209	30.195064	0.1838032	0.01366	0.025996
29.54	34.698317	29.69203	28.283693	0.1746214	0.00515	0.042529
29.4	34.147305	29.14198	28.549264	0.161473	0.00878	0.028937
28.85	34.286293	29.28192	28.097932	0.188433	0.01497	0.026068
28.99	34.425281	29.42187	28.056381	0.1874881	0.0149	0.032205
29.13	34.624269	29.62181	28.326224	0.1886121	0.01688	0.027593
29.33	34.413257	29.41175	29.508937	0.1733126	0.00279	0.006101
29.12	34.652245	29.6517	30.172613	0.189981	0.01826	0.036147
29.36	34.861234	29.86164	29.641147	0.1873717	0.01709	0.009576
29.57	34.680222	29.68159	30.306994	0.1728178	0.00377	0.024924

Table R5: General Motors Relative Errors

P_t Actual	$\hat{P}_t$ App.1	$\hat{P}_t$ App.2	$\hat{P}_t$ App. 3	RE App 1	RE App2	RE App3
31.55	-91.47079	112.3139	31.416398	3.8992325	2.55987	0.004235
30.8	-91.39	112.3767	31.391862	3.9672079	2.64859	0.019216
30.87	-91.39922	112.3495	32.149915	3.9607782	2.63944	0.041461
30.85	-91.48844	112.2423	30.406611	3.9655896	2.63832	0.014372
30.75	-91.36766	112.3451	30.684878	3.9713059	2.6535	0.002118
30.86	-91.60688	112.0879	30.639633	3.9684665	2.63214	0.007141
30.61	-91.43609	112.2407	31.184169	3.9871314	2.6668	0.018758
30.77	-91.61531	112.0435	31.343239	3.9774232	2.64132	0.01863
30.58	-91.94453	111.6964	29.359082	4.0066883	2.65259	0.039925
30.24	-92.53375	111.0892	30.259485	4.0599784	2.67358	0.000644
29.64	-92.71296	110.892	29.207333	4.1279678	2.74129	0.014597
29.45	-91.43218	112.1548	31.256735	4.1046581	2.80831	0.061349
30.72	-91.5214	112.0476	32.023868	3.9792122	2.64738	0.042444
30.62	-91.54062	112.0104	30.929554	3.9895695	2.65808	0.01011
30.59	-92.34984	111.1832	29.076968	4.0189551	2.63463	0.049462
29.77	-92.68905	110.826	28.400872	4.1135053	2.72274	0.04599
29.42	-92.54827	110.9488	29.409839	4.1457604	2.7712	0.000345
29.55	-92.50749	110.9717	29.797662	4.1305411	2.75539	0.008381
29.58	-92.79671	110.6645	29.604752	4.1371436	2.74119	0.000837
29.28	-92.76592	110.6773	29.77867	4.1682351	2.77996	0.017031
29.3	-92.79514	110.6301	29.60879	4.1670697	2.77577	0.010539
29.26	-92.27436	111.1329	28.949847	4.1536008	2.79812	0.0106
29.77	-92.58358	110.8057	29.345401	4.1099623	2.72206	0.014263
29.45	-92.16279	111.2085	29.649987	4.1294667	2.77618	0.006791
29.86	-92.03201	111.3213	29.203044	4.082117	2.72811	0.022001
29.98	-92.42123	110.9142	29.155677	4.0827629	2.6996	0.027496
29.58	-92.94045	110.377	28.625286	4.142003	2.73147	0.032276
29.05	-92.60967	110.6898	28.879647	4.1879403	2.81032	0.005864
29.37	-91.82888	111.4526	31.34246	4.1266218	2.79478	0.067159
30.14	-92.1481	111.1154	30.367734	4.0573358	2.68664	0.007556
29.81	-92.25732	110.9882	28.294592	4.0948447	2.72319	0.050836
29.69	-92.70654	110.521	28.768294	4.1224836	2.7225	0.031044
29.23	-92.42575	110.7838	29.767882	4.1620169	2.79007	0.018402
29.5	-91.94497	111.2466	30.077074	4.1167787	2.77107	0.019562

Table R6: Sony Corporation Relative Errors

P_t Actual	$\hat{P}_t$ App.1	$\hat{P}_t$ App.2	$\hat{P}_t$ App. 3	RE App 1	RE App2	RE App3
47.06	35.3	40.5	45.542482	0.2500903	0.14027	0.032246
46.36	36.5	41.6	48.570039	0.2132745	0.10181	0.047671
47.54	35.9	41.1	47.277591	0.244122	0.13545	0.00552
47	36.6	41.8	48.230572	0.2209301	0.11102	0.026182
47.68	34.6	39.8	44.415831	0.2745778	0.16626	0.06846
45.65	34.7	39.9	46.316309	0.2394309	0.12631	0.014596
45.78	35.3	40.5	47.302524	0.2280073	0.11523	0.033257
46.4	33.8	38.9	44.68796	0.2725502	0.1613	0.036897
44.81	33.7	38.9	44.880315	0.24692	0.13174	0.001569
44.8	32.7	37.9	44.042222	0.2690321	0.15384	0.016915
43.8	31.9	37.0	43.781616	0.2726208	0.15482	0.00042
42.91	31.8	36.9	43.214275	0.2593556	0.13913	0.007091
42.83	32.1	37.3	44.39234	0.2497572	0.12933	0.036478
43.18	32.3	37.5	44.253007	0.2518586	0.13243	0.02485
43.35	31.9	37.0	43.096052	0.2648999	0.14596	0.005858
42.91	31.6	36.8	42.117549	0.2626792	0.14254	0.018468
42.68	31.9	37.1	43.364406	0.2525707	0.1318	0.016036
42.94	32.0	37.1	44.018935	0.2556561	0.13564	0.025127
43	31.8	37.0	44.268549	0.2594425	0.13961	0.029501
42.88	32.1	37.3	42.71944	0.2510303	0.13088	0.003744
43.15	32.4	37.5	44.692862	0.2496485	0.13027	0.035756
43.41	31.8	36.9	43.876438	0.2676915	0.14905	0.010745
42.82	30.7	35.9	43.113577	0.2825465	0.16229	0.006856
41.75	30.0	35.1	40.457128	0.2818395	0.15852	0.030967
41.01	29.4	34.5	40.729068	0.2834663	0.15795	0.00685
40.41	29.1	34.2	37.940068	0.280453	0.15309	0.061122
40.1	28.9	34.0	39.462658	0.2798319	0.15151	0.015894
39.9	28.7	33.9	38.609636	0.2801858	0.15124	0.03234
39.74	28.6	33.7	41.709055	0.2805125	0.15107	0.049548
39.61	28.8	33.9	38.697965	0.2738127	0.14397	0.023025
39.78	28.7	33.8	40.019213	0.2791321	0.14986	0.006013
39.69	28.4	33.5	40.783088	0.2845057	0.15496	0.027541
39.41	28.5	33.6	38.096245	0.276838	0.1464	0.033336
39.51	28.3	33.4	40.043841	0.2839367	0.15385	0.013512

Table R7: Toyota Motor Corporation Relative Errors

P_t Actual	$\hat{P}_t$ App.1	$\hat{P}_t$ App.2	$\hat{P}_t$ App. 3	RE App 1	RE App2	RE App3
47.06	35.3	40.5	45.542482	0.2500903	0.14027	0.032246
46.36	36.5	41.6	48.570039	0.2132745	0.10181	0.047671
47.54	35.9	41.1	47.277591	0.244122	0.13545	0.00552
47	36.6	41.8	48.230572	0.2209301	0.11102	0.026182
47.68	34.6	39.8	44.415831	0.2745778	0.16626	0.06846
45.65	34.7	39.9	46.316309	0.2394309	0.12631	0.014596
45.78	35.3	40.5	47.302524	0.2280073	0.11523	0.033257
46.4	33.8	38.9	44.68796	0.2725502	0.1613	0.036897
44.81	33.7	38.9	44.880315	0.24692	0.13174	0.001569
44.8	32.7	37.9	44.042222	0.2690321	0.15384	0.016915
43.8	31.9	37.0	43.781616	0.2726208	0.15482	0.00042
42.91	31.8	36.9	43.214275	0.2593556	0.13913	0.007091
42.83	32.1	37.3	44.39234	0.2497572	0.12933	0.036478
43.18	32.3	37.5	44.253007	0.2518586	0.13243	0.02485
43.35	31.9	37.0	43.096052	0.2648999	0.14596	0.005858
42.91	31.6	36.8	42.117549	0.2626792	0.14254	0.018468
42.68	31.9	37.1	43.364406	0.2525707	0.1318	0.016036
42.94	32.0	37.1	44.018935	0.2556561	0.13564	0.025127
43	31.8	37.0	44.268549	0.2594425	0.13961	0.029501
42.88	32.1	37.3	42.71944	0.2510303	0.13088	0.003744
43.15	32.4	37.5	44.692862	0.2496485	0.13027	0.035756
43.41	31.8	36.9	43.876438	0.2676915	0.14905	0.010745
42.82	30.7	35.9	43.113577	0.2825465	0.16229	0.006856
41.75	30.0	35.1	40.457128	0.2818395	0.15852	0.030967
41.01	29.4	34.5	40.729068	0.2834663	0.15795	0.00685
40.41	29.1	34.2	37.940068	0.280453	0.15309	0.061122
40.1	28.9	34.0	39.462658	0.2798319	0.15151	0.015894
39.9	28.7	33.9	38.609636	0.2801858	0.15124	0.03234
39.74	28.6	33.7	41.709055	0.2805125	0.15107	0.049548
39.61	28.8	33.9	38.697965	0.2738127	0.14397	0.023025
39.78	28.7	33.8	40.019213	0.2791321	0.14986	0.006013
39.69	28.4	33.5	40.783088	0.2845057	0.15496	0.027541
39.41	28.5	33.6	38.096245	0.276838	0.1464	0.033336
39.51	28.3	33.4	40.043841	0.2839367	0.15385	0.013512

Table R8: Barclays plc Relative Errors

P_t Actual	$\hat{P}_t$ App.1	$\hat{P}_t$ App.2	$\hat{P}_t$ App. 3	RE App 1	RE App2	RE App3
59.4	67.688885	61.25519	56.988994	0.1395435	0.03123	0.040589
59.14	67.97722	61.54478	58.459572	0.1494288	0.04066	0.011505
59.43	68.275556	61.84436	58.92972	0.1488399	0.04063	0.008418
59.73	68.533891	62.10395	58.773339	0.1473948	0.03974	0.016016
59.99	68.152227	61.72354	57.651249	0.1360598	0.0289	0.038986
59.61	67.490562	61.06313	55.529566	0.132202	0.02438	0.068452
58.95	68.358898	61.93272	59.957847	0.1596081	0.0506	0.017097
59.82	67.487233	61.06231	60.681933	0.1281717	0.02077	0.014409
58.95	66.785569	60.36189	54.232268	0.1329189	0.02395	0.080029
58.25	68.043904	61.62148	58.73474	0.1681357	0.05788	0.008322
59.51	67.96224	61.54107	59.759512	0.1420306	0.03413	0.004193
59.43	66.670575	60.25066	59.234389	0.1218337	0.01381	0.003291
58.14	66.908911	60.49025	58.9388	0.1508241	0.04042	0.013739
58.38	66.487246	60.06983	55.507791	0.1388703	0.02895	0.049199
57.96	66.145582	59.72942	57.848809	0.1412281	0.03053	0.001918
57.62	65.823917	59.40901	57.079813	0.1423797	0.03105	0.009375
57.3	65.822253	59.4086	57.003477	0.1487304	0.0368	0.005175
57.3	65.370588	58.95819	56.37655	0.140848	0.02894	0.016116
56.85	65.518924	59.10777	52.616976	0.1524877	0.03971	0.07446
57	65.147259	58.73736	55.652243	0.1429344	0.03048	0.023645
56.63	65.695595	59.28695	58.309366	0.1600847	0.04692	0.029655
57.18	65.74393	59.33654	55.918414	0.1497714	0.03771	0.022063
57.23	65.622266	59.21613	57.731001	0.146641	0.0347	0.008754
57.11	65.120601	58.71572	55.025923	0.1402662	0.02812	0.036492
56.61	65.578937	59.1753	57.202616	0.1584338	0.04532	0.010468
57.07	66.757272	60.35489	59.08003	0.1697437	0.05756	0.03522
58.25	64.295608	57.89448	58.629998	0.1037873	0.0061	0.006524
55.79	62.843943	56.44407	53.995729	0.1264374	0.01172	0.032161
54.34	62.482279	56.08366	52.376309	0.1498395	0.03209	0.036137
53.98	62.420614	56.02324	53.572066	0.1563656	0.03785	0.007557
53.92	62.38895	55.99283	55.027143	0.1570651	0.03844	0.020533
53.89	62.557285	56.16242	56.28486	0.1608329	0.04217	0.04444
54.06	62.535621	56.14201	52.747384	0.1567817	0.03851	0.024281
54.04	61.643956	55.2516	56.746861	0.1407098	0.02242	0.05009

Table R9: Nikkei Relative Errors

P_t Actual	$\hat{P}_t$ App.1	$\hat{P}_t$ App.2	$\hat{P}_t$ App. 3	RE App 1	RE App2	RE App3
17310.44	17371.3	17370.6	17393	0.003515	0.003475	0.004769
17370.93	17261.7	17261.0	16940.62	0.006287	0.006327	0.024772
17261.35	17202.8	17202.1	17313.01	0.003391	0.003431	0.002993
17202.46	17210.3	17209.6	17367.62	0.000455	0.000414	0.009601
17209.92	17057.4	17056.7	16964.9	0.008864	0.008904	0.014237
17057.01	16838.5	16837.8	16893.56	0.012809	0.012849	0.009583
16838.17	16942.8	16942.1	16717.31	0.006212	0.00617	0.007178
16942.4	17238.1	17237.4	17291.49	0.017455	0.017414	0.020604
17237.77	17092.0	17091.3	17144.61	0.008459	0.008499	0.005404
17091.59	17354.0	17353.3	16892.11	0.015355	0.015315	0.011671
17353.67	17226.2	17225.5	17505.57	0.007346	0.007386	0.008753
17225.83	17225.2	17224.5	17596.08	3.82E-05	7.84E-05	0.021494
17224.81	17249.0	17248.3	17515.45	0.001404	0.001364	0.016873
17248.63	17169.6	17168.9	17222.23	0.004585	0.004625	0.001531
17169.19	17093.3	17092.6	17439.95	0.004423	0.004463	0.01577
17092.89	17105.3	17104.6	17574.22	0.000727	0.000687	0.028159
17104.96	17048.2	17047.5	16931.66	0.003319	0.003359	0.010131
17047.83	17011.4	17010.7	16363.69	0.002137	0.002177	0.040131
17011.04	16777.2	16776.5	16794.7	0.013744	0.013785	0.012718
16776.88	16962.5	16961.8	16894.05	0.011062	0.011021	0.006984
16962.11	16914.7	16914.0	16923.93	0.002797	0.002838	0.002251
16914.31	16829.6	16828.9	16586.33	0.00501	0.005051	0.019391
16829.2	16693.3	16692.6	16602.03	0.008076	0.008117	0.013498
16692.93	16638.1	16637.5	16160.81	0.003282	0.003324	0.031877
16637.78	16528.4	16527.7	16589.31	0.006577	0.006619	0.002913
16527.99	16418.2	16417.5	16133.46	0.006644	0.006686	0.023871
16417.82	16473.7	16473.0	16633.84	0.003405	0.003363	0.013158
16473.36	16371.6	16371.0	16375.58	0.006175	0.006217	0.005935
16371.28	16266.1	16265.4	16131.54	0.006423	0.006466	0.014644
16265.76	16304.0	16303.3	16606.72	0.002348	0.002305	0.020962
16303.59	16322.1	16321.5	16265.11	0.001138	0.001095	0.00236
16321.78	16274.7	16274.0	16535.94	0.002885	0.002927	0.013121
16274.33	16076.6	16075.9	16223.38	0.012152	0.012195	0.003131
16076.2	15855.6	15854.9	15927.63	0.013721	0.013764	0.009242