Empirical Evidences of Coherent Market Hypothesis

by

Peter, Ta-Chao Kao

B.Bus.Sc.(Hon.), University of Cape Town (2000)

Submitted to the Department of Mathematics and Applied Mathematics
in partial fulfillment of the requirements for the degree of

Masters of Science in Mathematics of Finance

at the

UNIVERSITY OF CAPE TOWN

December 2002

Signature of Author . . . .
Department of Mathematics and Applied Mathematics
11 November 2002

Certified by . . . .
Dr RenKuan Guo
Associate Professor, Department of Statistical Sciences, University of Cape Town
Thesis Supervisor

Accepted by . . . .
Professor B.D. Reddy
Dean, Faculty of Science, University of Cape Town

Empirical Evidences of Coherent Market Hypothesis

by

Peter, Ta-Chao Kao

Submitted to the Department of Mathematics and Applied Mathematics
on 11 November 2002, in partial fulfillment of the
requirements for the degree of
Masters of Science in Mathematics of Finance

Abstract

Financial markets are comprised of many interacting agents, and thus are better modelled as complex systems. Over the last fifty years, the dominant academic theory of financial market has been the Efficient Market Hypothesis (EMH). However, no matter how proponents of EMH have attempted to defend the EMH, the analyses of observed data from the financial market have continuously produced more evidences against EMH (or the random walk model).

Based on these empirical evidence, it is clear that financial markets are non-linear in nature. Vaga (1990) proposed the Coherent Market Hypothesis (CMH), which is an alternative non-linear model of financial market returns and includes EMH as a special case.

In this dissertation, empirical explorations of basic properties of the CMH-based returns distribution will be conducted on the Johannesburg Stock Exchange. This is followed by a theoretical exploration of the stochastic differential equations that governs the underlying market dynamics.

Thesis Supervisor: Dr RenKuan Guo
Title: Associate Professor, Department of Statistical Sciences, University of Cape Town

Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

My sincerest thanks to Dr Renkuan Guo, my dissertation supervisor, for his confidence in me, encouragement, insights, dedication and patience. I am most grateful to Prof. Casparus G. Troskie for the provision of JSE-OVER data, encouragement, experience, stimulating discussions and insights. My thanks to the Department of Statistical Sciences and their staff for providing the support and friendly working environment.

I acknowledge John P. Nolan from Math/Stat Department at the American University, Washington, U.S.A. for authoring and making "STABLE 3.04" available. A big thanks to Fernando Durrell for his assistance and patience with crucial parts of the programming.

My thanks to the Department of Mathematics and Applied Mathematics for hosting the MSc in Mathematics of Finance programme and for accommodating me. Thanks to Dr Peter W. Ouwehand for coordinating the MSc programme and for providing me insights into the mathematical foundation underlying the linear financial market theory.

I thank Melina for always being there for me even in my most desperate moments. Finally, I thank my parents deeply for all their sacrifices in raising me and for offering me all the opportunities in the world.

Chapter 1

INTRODUCTION

The goal of this dissertation is to empirically and theoretically explore the Vaga (1990) Coherent Market Hypothesis returns distribution on the Johannesburg Stock Exchange (JSE.).

1.1 Problem Statements For The Dissertation

The financial market models are statistical in nature, thus must be data-driven. It follows that the merit of the financial models should be judged on whether the observed market dynamics have been reasonably captured. The academic motivation for working with non-linear models of financial market dynamics is based on the following stylised findings of financial market price movements as summarized by J. Voit (2001):

1. The empirical data from all financial markets have fat-tailed (leptokurtic) probability distributions comparing to the assumed Gaussian market theory. In Gaussian market, big price changes are very likely the consequence of many small changes. In real markets, they are very likely the consequence of very few big price changes. The actual stock price is much less continuous than a random walk. Statistical methods based on Gaussian distribution will become questionable.

2. The Black-Scholes analysis of option pricing becomes problematic. Geometric Brownian motion is a necessary condition. Risk-free portfolios can no longer be constructed in theory - not to mention the problems encountered in Black-Scholes continuous adjustment

of positions when the stochastic process followed by the underlying security is discontinuous. The risk associated with an investment is strongly underestimated by Gaussian distributions or geometric Brownian motion. The standard arguments for risk control by diversification may no longer work.

3. The variance of an infinite data sample will be infinite to the extent that a stable Lévy distribution with index $1 \leq \mu \leq 2$ can describe the data. For finite data samples, the variance of course is finite, but it will not converge smoothly to a limit when the sample size is increased.

4. According to economic wisdom, stock prices reflect both the present situation as well as future expectations. While the actual situation most likely evolves continuously, future expectations may suffer discontinuous changes, because they depend on factors such as information flow and human psychology.

5. The trading activity is very non-stationary. There are quiescent periods changing with hectic activity, and sometimes trading is stopped altogether.

Non-linear models are developed in mainly in physics, but also in other scientific fields such as chemistry, biology, earth sciences, social sciences, medical sciences to name a few. Financial non-linear models are often "borrowed" from other fields, particularly from physics in recent years. Based on empirical evidence from financial markets, Vaga (1990)'s Coherent Market Hypothesis (CMH) may be considered as the best starting point for non-linear financial market modelling explorations. However, it has been almost twelve years since the CMH model was first proposed. Yet,there have been no serious empirical data evidence in the academic literature to support CMH model. The exception is a paper by Steiner and Wittkemper (1997).

1.2 Objectives of the Dissertation

The primary goals for this Masters dissertation is to explore, both empirically and theoretically, the following:

1. Basic properties of the CMH-based returns distribution, and

2. The stochastic differential equations (SDE) governing the market dynamics.

In order to achieve the above two goals, the seven objectives for this dissertation are defined as follows:

Objective 1: Using the ideas of technical analysis to split the time series of log returns on JSE Overall Index (from 27 March 1985 to 21 June 2002) into five distinct market states (namely: Random Walk, Coherent Bull, Coherent Bear, Transition, Chaotic).

Objective 2: Calculate the mean, variance, skewness coefficient and excess kurtosis of the log returns in each state.

Objective 3: Graphically explore the relationships between the first four moments of the log returns in each state and the parameters of the CMH return distribution.

Objective 4: Derive the related partial derivatives of the log-likelihood function, assuming the CMH daily return distributions, to prepare for the maximum likelihood estimation procedure.

Objective 5: Estimating CMH return distribution parameters for the Random Walk and Coherent Bull (Bear), Chaotic as well as the transitional states tentatively,

Objective 6: Estimating the five market states by the stable distribution family, and finally

Objective 7: Exploring the CMH returns distributions based on the related diffusion processes.

1.3 Structure of the Dissertation

Excluding this introduction, this dissertation consists of five chapters. A brief overview for each chapter are provided as follows:

Chapter 2 Presents a critical evaluation of the Efficient Market Hypothesis (EMH) upon which the linear financial modelling methodologies are based. The major successes of linear financial models are briefly discussed. It will be shown that researches have continuously produced convincing evidence against EMH and the linear financial models. There is a strong motivation for non-linear financial models that are more able to incorporate essential features from empirical data. A survey of the main non-linear models include Fractal Market Hypothesis (FMH), Lévy processes, and behavioural financial models.

Chapter 3 Presents a detailed discussion of Vaga (1990)'s Coherent Market Hypothesis (CMH) model of annualised returns.

Chapter 4 Presents an exploration of the empirical properties of JSE by means of extensive statistical tests of the JSE overall index. Based on the idea of technical analysis, the observed daily returns where partitioned into five market states as described by Vaga (1990). Finally, stable distribution is fitted to the partitioned daily returns in each market state.

Chapter 5 The CMH model is fitted to the partitioned daily return in each market state by maximum likelihood estimation of the CMH control parameters. The derivation of the procedure is provided and the results from the fitting discussed.

Chapter 6 Presents a theoretical exploration into the underlying stochastic differential equation that governs the CMH returns distribution.

Chapter 7 Provides concluding remarks on the CMH model based evidences provided from chapters 4 to 6. The bibliography lists the references used for the completion of this dissertation.

The appendices contains the program codes used to calculate the various statistics required for the analysis..

Chapter 2

LITERATURE REVIEW AND DISCUSSION OF RELEVANT THEORY

2.1 Review of the Efficient Market Hypothesis

Fama (1970) provided a landmark survey of the theoretical and empirical literature on the efficient market models. In this paper, an "efficient" market is one in which the prices "fully reflect" available information. Fama (1970) went on to define three subsets of information to which security prices adjusts, and hence three forms of market efficiency which he termed: weak form, semi-strong form, and strong-form. LeRoy (1989) provided a comprehensive critique of Fama's EMH both on empirical evidences and from the economic theoretic view. Most importantly, LeRoy pointed out that empirical tests of EMH are usually tests of martingale model, but the association between the intuition of market efficiency and the martingale model is far from direct. However, there are many problems with the statements of EMH and its implications for empirical research. The most serious of which is the fact that market efficiency must be tested with an equilibrium-pricing model. Even Fama (1991) had to admit that this joint hypothesis problem effectively renders tests of market efficiency impossible. Fama (1991) appealed to researchers to focus instead on how EMH can assist in describing time series and cross-sectional behaviour of security prices.

The most stringent requirement of implied by EMH is that the observations had to be independent or had to have a short-term memory; the current change in prices could not be inferred from previous changes. This could occur only if price changes were a random walk and if the best estimate of the future price was the current price. The process would then be a "martingale" or a fair game. The random walk model said that future price changes could not be inferred from past price changes. It did not refer to exogenous information, i.e.) economic or fundamental information. Hence, random walk theory was primarily an attack on technical analysis. Furthermore, the EMH, in its "semi-strong" form, argued that current prices reflected all public information, all past prices, published reports and economic news because of fundamental analysis. All investors has equal access to this information and, being 'rational', they would value the security accordingly. Thus, investors, in aggregate, could not profit from the market because it 'efficiently' valued securities at a price that reflected all known information.

One of the immediate implication of EMH is that technical and fundamental analysis cannot make abnormal profits and that buy-sell decision should be based solely on the market's long-term uptrend. The rest is noise accounted for by the standard deviation of annual returns. This implies that a buy-and-hold strategy would be as effective as or more effective than the average investment timing strategy.

The EMH attempts to explain the statistical structure of the markets, but the theory emerged after the imposition of a statistical structure. Bachelier (1900) first proposed with little proof that markets followed a random walk and can be modelled by standard probability calculus. Eventually, a number of mathematicians realised that stock market were a time series, and as long as the market fulfilled certain restrictive requirements, they can be modelled by probability calculus. The advantage of this approach is that it offered a large body of tools for research but there was a division in the mathematical community about whether statistics could be applied to a time series.

Normal scientific reasoning requires that a certain behaviour and structure are first observed in a system or process, that a theory is then developed to fit the known facts and that the theory is modified or revised as new facts become known. The theory of EMH was developed to justify the use of statistical tools that require independence or very short term memory. The theory was often in disagreement with observed behaviour. There are too many large up-and-down changes at all frequencies for the normal curve to be fitted to these distributions. However, the large changes were labeled special events or :'anomalies' and were left out of the frequency distribution. Leaving out the large changes and renormalise results in the normal distribution. Price changes were labelled 'approximately normal'. Alternatives like the stable Paretian distribution were rejected even though they fit the observed values without modification because standard statistical analysis could not be applied using those distributions. Hence, there is a need to develop a market hypothesis that fits the observed facts and takes into account why markets exist to begin with.

2.1.1 EMH and Mathematical Finance

One of the most famous success story of linear financial modelling is the Black-Scholes-Merton approach to pricing options, warrants and, more generally, derivative securities. Its fame is derived from the speed at which academics and investment professionals have accepted the idea. Its success is derived from the creation of a wide range of new investment possibilities and a complete revolution in risk management philosophies. The fundamental insight of the option pricing models of Black and Scholes (1973) and Merton (1973) is that under the **no-arbitrage** (or **Law of One Price**) condition, an option's payoff can be exactly replicated by a **self-financing** investment strategy constructed from the underlying stock and the riskless debt. Furthermore, this no-arbitrage condition also allows for options that does not exist to be replicated synthetically. This method of no-arbitrage pricing is one of the great achievements in modern financial economics, and has identified an area in finance where pure mathematicians can apply their considerable expertise. It has since developed into a legitimate scientific framework for derivative security pricing that gave rise to a whole new discipline called mathematical finance.

However, Campbell, Lo, MacKinlay (1997) pointed out that the crucial weakness of the mathematical finance framework is that the prices are derived completely independent of empirical data. Under the no-arbitrage paradigm, prices are determined exactly. Hence, there is no statistical error in these prices that needs to be minimised using traditional statistical inference.

Campbell, Lo, MacKinlay (1997) have identified the following two stages in the implementation of derivative security pricing models that requires statistical inference:

1. Parameter estimation of continuous-time price processes, which are inputs for the parametric option pricing formulas, and

2. Statistical considerations in the monte Carlo simulation of path-dependent security prices.

However, this does not detract from the fact that derivative security pricing models have in no ways reflect the empirical reality of the financial market. This is evidenced in the peculiar observation that the only distributions that mathematical finance employs to describe ANY financial variable is either the normal (associated with arithmetic Brownian motion), or the log-normal (associated with geometric Brownian motion) distribution.

2.1.2 "Stylised" Facts of Financial Price Movements

On account of the growing empirical evidence against the EMH, Voit (2001) summarised the stylised findings on financial market price movements:

1. All empirical data have fat-tailed (leptokurtic) probability distributions.

2. The variance of an infinite data sample will be infinite to the extent that these data are described by an α-stable Lévy distribution, where $\alpha \in [1, 2]$. For finite data samples, the variance will be finite, but it will not converge smoothly to a limit as the sample size increases.

3. Quantities derived from the probability distribution (such as mean, variance, or other higher order moments) will be extremely sample-dependent.

4. Statistical methods based on Gaussian distribution will become questionable.

5. The central limit theorem predicts a convergence to a Gaussian, which empirically does not take place.

6. Apparently, special time scale are eliminated by arbitrage.

7. The actual stock price is much less continuous than a random walk.

8. In Gaussian market, big price change are very likely the consequence of many small changes. In real markets, they are very likely the consequence of very few big price changes.

9. The trading activity is very nonstationary. There are quiescent periods changing with hectic activity, and sometimes trading is stopped altogether.

10. According to economic wisdom, stock prices reflect both the present situation as well as future expectations. While the actual situation most likely evolves continuously, future expectations may suffer discontinuous changes, because they depend on factors such as information flow and human psychology.

11 One consequence, namely that filters cannot work, has been discussed in Sect. 5.3.1. A necessary condition is that stock price flows a continuous stochastic process. On the contrary, the process giving rise to Lévy distribution must be rather discontinuous.

12. The assumption of a complete market is not always realistic. With discontinuous price changes, there will be no buyer and no seller at certain prices.

13. Stop-loss orders are not suitable as a protection against big losses. They require a continuous stochastic process to be efficient.

14. The risk associated with an investment is strongly underestimated by Gaussian distributions or geometric Brownian motion.

15. The standard arguments for risk control by diversification may no longer work.

16. The Black-Scholes analysis of option pricing becomes problematic. Geometric Brownian motion is a necessary condition. Risk-free portfolios can no longer be constructed in theory - not to mention the problems encountered in Black-Scholes continuous adjustment of positions when the stochastic process followed by the underlying security is discontinuous.

Thus, it is important to approach financial theory with a sceptical mind in order to seek a better approximation to the market realities.

2.2 Fractal Market Hypothesis and R/S Analysis

Despite the fact that Capital markets are not well described by the normal distribution and random walk theory, the Efficient Market Hypothesis still remains the dominant paradigm for how the market works. Peters (1994) offers an alternative theory of market structure. The Fractal Market Hypothesis provides an economic and mathematical structure to fractal market analysis. Through the Fractal Market Hypothesis it is possible to understand why self-similar statistical structures exist as well as how risk is shared and distributed among investors.

Before a detailed explanation of this alternative is given, we need to first review briefly the

EMH and answer the questions: why do markets exist and what do participants expect and require from markets ?

2.2.1 Distinction Between Stable Markets and Efficient Markets

Market liquidity is one of the major concerns to all investors. Investors need sufficient liquidity to:

1. Ensure that the price at which securities are traded is close to what the market considers fair, and
2. Avoid "panics" in the market, due to imbalances in supply and demand for securities, by allowing investors with different investment horizons to invest indifferently and anonymously.

Technological advances have made trading of large volumes of stock easier by matching up buyers and sellers with different investment horizons quickly and efficiently. However, the technology can only ensure that it is easier for a seller to find a buyer (and vice versa). There is no agreed-on mechanism for the determination of the "fair price" for a security. The capitalist system of supply and demand is strictly adhered to.

However, market liquidity is not the same as trading volume as market crashes have known to occur when there is high trading volume in the market, but low liquidity. Thus, the level of liquidity in the market should more correctly be interpreted as the degree of imbalances in trading volume.

The EMH does not account for liquidity as it simply says that prices are always fair irrespective of whether sufficient market liquidity exists in the market. However, when liquidity vanishes and a trade needs to take place, then obtaining a "fair" price may not be as important as completing the trade at all cost. Thus, the EMH framework cannot explain market crashes and stampedes. On the other hand, a stable market is a liquid market, which is quite different from an "efficient" market. If the market is liquid, then the price can be considered close to "fair". At times when there is a lack of liquidity on the market, the participating investors are willing to take any price then can , whether fair or not.

2.2.2 The Source of Liquidity

An important characteristic of investors is that they are not homogenous. Thus, the importance of information is largely dependent on the investment horizon of the investor. A six standard deviation technical drop would not change the outlook of the weekly trader, who looks at either longer technical or fundamental information. Therefore the day trader's six sigma event is a 0.15-sigma event to the weekly trader. The latter steps in, buys and creates liquidity, which in turn, stabilises the market.

All investors trading in the market simultaneously have different investment horizons and the information that is important at each investment horizon is different. The source of liquidity is investors with different investment horizons, different information sets and consequently different concepts of "fair price".

2.2.3 Information Sets and Investment Horizons

For a day trader, virtually all the tools are technical. For a portfolio manager both technical and fundamental information are used: the buy-and-sell decision will depend on fundamental information, although technical analysis may be used in the course of trading.

Short term investors primarily follow technical analysis. Longer-term investors are more likely to follow fundamentals. Thus, the "fair" value of a stock is identical in two ways:

1. For day traders, it is high for the day if they are selling, or low for the day if they are buying.

2. For long term investors, the actual buying or selling price becomes less important, relative to the high or low of the day.

2.2.4 The Fractal Market Hypothesis

Peter (1994) proposed an alternative theory of financial market, which he called a Fractal Market Hypothesis (FMH):

Proposition 1 (Fractal Market Hypothesis) *The Fractal Market Hypothesis proposes that:*

1. *The market is stable when it consists of investors covering a large number of investment horizons. This ensures that there is ample liquidity for traders.*

2. *The information set is more related to market sentiment and technical factors than in the long term. As investment horizons increase, longer-term fundamental information dominates. Price changes may reflect information important only to that investment horizon.*

3. *If an event makes the validity of fundamental information questionable, long-term investors either stop participating in the market or begin trading based on the short-term information set. When the overall investment horizon of the market shrinks to a uniform level, the market becomes unstable. There are no long term investors to stabilise the market by offering liquidity to short term investors.*

4. *Prices reflect a combination of short term technical trading and long term fundamental valuation. The underlying trend in the market is reflective of changes in expected earnings. based on the changing economic environment. Short term trends are more likely the result of crowd behaviour.*

5. *If a security has no tie to the economic cycle, there will not be a long term trend.*

The key idea of FMH is that financial markets will remain stable when many investors participate and have many different investment horizons. The important features of FMH are:

1. It emphasises the impact of liquidity and investment horizons on the behaviour of investors.

2. It places no statistical requirements on the process.

3. It provides a model of investor behaviour and market price movements that better explains the empirical observations on the markets.

Peters (1994) applies the FMH to explain the impact of the following three historical events on the U.S. financial market:

Example 2 (Kennedy assassination on 22 November 1963) *When President Kennedy was assassinated on November 22, 1963, long-term investors either did not participate that day or they panicked and became short term investors. Once fundamental information lost its value, these long-term investors shortened their investment horizon and began trading on overwhelmingly negative technical dynamics. The market was closed and by the time it reopened, investors were better able to judge the impact of the President's death on the economy, long-term assessment returned, and the market stabilised.*

Example 3 (The market crash on 19 October 1987) *On 19 October 1987, long-term investors had began selling their equity holdings. The crash was due to the market being dominated entirely by traders with extremely short investment horizons. Short-term information (or technical information) dominated in the event. As a result, the market reached new heights of instability and did not stabilise until long-term investors stepped in to buy during the following days.*

Example 4 (Gulf War 1990) *On 19 January 1990, the market roller coaster that took place was caused by the Gulf War. The uncertainty created by the pending war caused investors to concentrate on the short term and disregard the fundamental information. When two statesmen from the U.S. and Iraq met for longer than expected, the Dow Jones Industrials soared 40 points on expectation that a negotiated solution was at hand. When the meeting ended with no progress made, the market immediately plummeted 39 points. There was no fundamental reason for such a wide swing in the market. Investors had become short-term oriented or else the long-term ones did not participate. In either case, the market lost liquidity and became unstable.*

The Fractal statistical structure exists because it is a stable structure. As long as investors with different investment horizons are participating, a panic at one horizon can be absorbed by the other investment horizons as a buying (or selling) opportunity. When the investment horizon becomes uniform, lack of liquidity results and discontinuities appear in the pricing sequence. If the information received by the market is important both to the short and long-term horizons, then liquidity can also be affected.

The FMH owes much to the Coherent Market Hypothesis (CMH) of Vaga (1990) and the K-Z model of Larrain (1991). Like the CMH, the FMH is based on the premise that the market assumes different states and can shift between stable and unstable regimes. Like the K-Z model, the FMH finds that the chaotic regime occurs when investors lose faith in long-term fundamental information. FMH argues that the market is stable when it has no characteristic time scale or investment horizon. Instability occurs when the market loses its Fractal structure and assumes that a fairly uniform investment horizon.

Most standard market analysis assumes that the market processes is, essentially, stochastic. For testing the Efficient Market Hypothesis (EMH), this assumption causes few problems. For the FMH, many of the standard tests does not have enough power. Thus,new methodologies are needed to take advantage of the FMH market structure. One of them that has already been developed is the R/S analysis, which is a robust form of time series analysis.

2.2.5 Long-Range Dependence and the R/S Analysis

Financial time series are non-i.i.d. and non-linear stochastic systems consisting of both randomness and determinism. In order to accommodate this non-linearity, non-parametric statistical analysis is required, where no prior assumption about the probability distribution of the data studied is required.

The Development of the R/S Analysis

The Re-scaled Range or R/S statistic was originally developed by the English hydrologist Harold Edwin Hurst. Hurst (1951) applied R/S analysis to the annual discharges of the Nile River to assist with the modelling of reservoir design. Later on, he applied this very robust non-parametric statistical technique to many natural systems in order to:

1. Distinguish between random and non-random systems,
2. Identify the persistence of trends, and
3. Identify the duration of cycles.

In the analysis of financial time series, R/S analysis can be used to distinguish between random time series and Fractal time series.

Time Series with Long-Range Dependence

If the time series being measured increases by greater than

$$R = \sqrt{T},$$

where R is the distance covered, and T is the time index used., then this provides evidence that the event measured is not random. The phenomenon of long-range dependence is one departure from the random walk hypothesis that is beyond the statistical framework that have been developed so far. There are two definitions of long-range dependence:

Definition 5 (Long Range Dependent in Autocovariance) *Long-range dependent processes may be defined as stochastic processes $\{p_t; t \geq 0\}$ with autocovariance functions $\gamma_p(k)$ such that as $k \to \infty$*

$$\gamma_p(k) \sim \left\{ \begin{array}{ll} k^{\nu} f_1(k), & for\ \nu \in (-1, 0) \\ -k^{\nu} f_1(k), & for\ \nu \in (-2, -1) \end{array} \right\},$$

where $f_1(k)$ is any slowly varying function at infinity (i.e., $\lim_{x \to \infty} \frac{f(tx)}{f(x)} = 1$ for all $t \in [a, \infty)$).

Definition 6 (Long Range Dependence in Spectral Density) *Long-range dependence may also be defined as stochastic processes with spectral density functions $s(\lambda)$ such that as $\lambda \to 0$:*

$$s(\lambda) \sim \lambda^{-\alpha} f_2(k) \ for\ \alpha \in (-1, 1),$$

where $f_2(k)$ is a slowly varying function.

Long-range-dependent time series exhibit an unusually high degree of persistence. Thus, observations in the recent past are non-trivially correlated with observations in the distant future, even as the time span between the two observations increases. Nature's predilection towards long-range dependence has been well-documented in the natural sciences such as hydrology, meteorology, and geophysics. Some have argued that economic time series are also long-range dependent. In the frequency domain, such time series exhibit power at the lowest frequencies.

Tests for long-range dependence: The Hurst-Mandelbrot Rescaled Range statistic and R/S analysis

The importance of long-range dependence in asset markets was first studied by Mandelbrot (1971), who proposed using the range over standard deviation, or R/S, statistic, also called the *re-scaled range*, to detect long-range dependence in economic time series.

Definition 7 (Classical R/S Statistics) *Consider a sample of continuously compounded asset returns:*

$$\{r_1, r_2, \cdots, r_n\}$$

Let $\bar{r}_n$ denote the sample mean, this is:

$$\bar{r}_n = \frac{1}{n}\sum_j r_j$$

Then, the classical re-scaled-range statistic, denoted Q_n is defined as:

$$Q_n = \frac{1}{s_n}\left[\max_{1\leq k\leq n}\sum_{j=1}^{k}(r_j - \bar{r}_n) - \min_{1\leq k\leq n}\sum_{j=1}^{k}(r_j - \bar{r}_n)\right], \quad (2.1)$$

where s_n is the maximum likelihood estimator of standard deviation:

$$s_n = \left[\frac{1}{n}\sum_j (r_j - \bar{r}_n)^2\right]^{\frac{1}{2}}.$$

Remark 8 *The first term in brackets of equation (2.1) is the maximum (over k) of the partial sums of the first k deviations of r_j from the sample mean. Since the sum of all n deviations of $r_j^{\prime}s$ from their mean is zero, this maximum is always nonnegative. The second term in (2.1) is the minimum (over k) of this same sequence of partial sums, and hence it is always non-positive. The difference of the two quantities called the adjusted range, is always nonnegative and hence $Q_n \geq 0$.*

Data requirements for R/S analysis

Time series of data to be analysed using R/S analysis should have the following three properties:

1. Consists of large number of observations,

2. Observations cover a long time interval, and

3. Observation should be sampled less frequently.

Advantages of R/S analysis

The main advantages of the R/S analysis are:

1. **It can detect long-range dependence** in highly non-Gaussian time series with large skewness and/or kurtosis, which can be shown by Monte Carlo simulations. According to Mandelbrot (1972, 1975) there is an almost-sure convergence of the R/S statistic for stochastic processes with infinite variances which is a distinct advantage over autocorrelations and variance ratios.

2. **It can detect both periodic and non-periodic cycles**, as argued by Mandelbrot (1972). Spectral analysis can only detects periodic cycles,

3. **It is robust to noise in the time series.**

Even though these claims can be contested, it is a well-established fact that long-range dependence can indeed be detected by the "classical" R/S statistic.

Disadvantage of R/S analysis

The most important shortcoming of the R/S statistics is its **sensitivity to short-range dependence**. This implies that any incompatibility between the data and the predicted behaviour of the R/S statistic under the null hypothesis need not come from the long-range dependence, but may merely be a symptom of short-term memory.

Example 9 *Lo (1991) showed that under RW1, the asymptotic distribution of:*

$$\left(\frac{1}{\sqrt{N}}\right) \cdot Q_n$$

is given by the random variable ν, the range of a Brownian bridge. But, under a stationary AR(1) specification with Autoreggressive coefficient ϕ, the normalised R/S statistic converges to $\xi\nu$, where

$$\xi \equiv \sqrt{\frac{(1+\phi)}{(1-\phi)}}.$$

To account for the effects of short-range dependence, Lo (1991) developed a modification on the R/S statistic by:

1. Deriving an asymptotic sampling theory under several null and alternative hypothesis, and

2. Demonstrated under Monte Carlo simulations and empirical examples that the modified R/S statistic is considerably more accurate, often yielding inferences that contradict those of its classical counterpart.

2.3 Review of Non-Linear Financial Models

The Gaussian family, including Brownian motion as its special member, is the most popular stochastic process used in the financial market modelling. Logically, wealth in the market is an aggregate outcome of thousands of small financial gains and losses up to date. Statistically, such financial information flow indexed by continuous time can be treated as the sum of large number of *i.i.d.* random variates and according to the Central Limit Theorem, after an appropriate *shifting* and *scaling* the asymptotic distribution, if it exists, should follow *Lévy* distribution. Thus, the wealth processes are in nature of Lévy processes.

2.3.1 Lévy processes

Mandelbrot (1963) was the first to propose geometric Lévy processes in financial market modelling triggered by his observation that the logarithm of relative price changes in market demonstrated long-tail behavior and he performed statistical analysis using financial market data. Following steps of Mandelbrot, empirical studies on market returns, interest rate, foreign exchange rate changes, and commodities price movements were carried out, for example Aase (2000).

Lévy distributions possess scaling or fractal properties and could provide a better statistical representation of financial market data. If we define::

$$\begin{aligned}\kappa &= \text{kurtosis of a distribution}\\ &= \frac{E\left[(X-\mu_X)^4\right]}{\left\{E\left[(X-\mu_X)^2\right]\right\}^2} - 3,\end{aligned}$$

then $\kappa = 0$ is for Gaussian distribution. However, analysis of typical financial market data produces kurtosis $\kappa \gg 0$. For example, Cont (1997) provided the following results:

Data	Frequency	Approx. κ
US dollar/German DM exchange rate futures	5 minutes	74
US dollar/Swiss Franc exchange rate futures	5 minutes	74
S&P500 index futures	5 minutes	16

The empirical evidences provided above suggests that the Gaussian distribution-based option pricing and hedging theory failed to capture the fundamental feature of most price change distribution that they have fat tails (i.e., large price changes appears more frequently than that assumed by the Gaussian theory).

In his Ph.D. thesis, Raible (2000) systematically investigated the fundamental aspects of the financial market: option pricing and hedging, term structure, bond price, and stochastic volatility under the framework of Lévy processes theory, paralleling in certain sense to the standard mathematical finance developments initiated by Black, Scholes (1973).

Definition 10 (Lévy Processes) *Let* $\mathbb{P}$ *be a probability measure on* $(\Omega, \mathcal{F})$ *with* $\mathbb{P}(\zeta = \infty) = 1$. *An adapted process* $\mathbf{L} = \{L_t, t \in \mathbb{R}^+\}$ *with* $X_0 = 0$ a.s. *is a Lévy process, if:*

(i) $\mathbf{L}$ *has independent increments, i.e.,* $L_t - L_s$ *is independent of* $\mathcal{F}_s$, $0 \le s < t < \infty$;

(ii) $\mathbf{L}$ *has stationary increments, i.e.,* $L_t - L_s$ *has the same distribution as* L_{t-s}, $0 \le s < t < \infty$;

(iii) L_t *is continuous at* t *in probability, i.e.*

$$\Pr\left[\lim_{t \to s} L_t = L_s\right] = 1.$$

Remark 11 *From the above definition, Lévy processes have stationary (or homogeneous) independent increments. They can be thought of as analogues of random walks in continuous time.*

Consider an arbitrary Lévy process $(X, \mathbb{P})$. Using the decomposition

$$\begin{aligned} X_1 &= X_{1/n} + (X_{2/n} - X_{1/n}) + \ldots + (X_{n/n} - X_{n-1/n}) \\ &= \sum_{i=1}^{n} \left(X_{i/n} - X_{i-1/n}\right), \end{aligned}$$

we observe the distribution $\mathbb{P}(X_1 \in)$ is infinitely divisible, and we denote its characteristic exponent by Ψ,

$$\mathbb{E}(\exp\{i\langle\lambda, X_1\rangle\}) = \exp\{-\Psi(\lambda)\}, \text{ for } \lambda \in \mathbb{R}^d.$$

By a similar argument, we see that for any rational number $t \geq 0$, $P(X_1 \in)$ is infinitely divisible as well, and its characteristic function is given by

$$\mathbb{E}(\exp\{i\langle\lambda, X_1\rangle\}) = \exp\{-t\Psi(\lambda)\}, \text{ where } \lambda \in \mathbb{R}^d. \tag{2.2}$$

Because X is right-continuous a.s., the mapping $t \to \mathbb{E}(\exp\{i\lambda, X_t\})$ is right-continuous and (2.2) holds for all $t \geq 0$. The function $\Psi : \mathbb{R}^d \to \mathbb{C}$ is called the *characteristic exponent* of the Lévy process X. It characterises the law $\mathbb{P}$ in the sense that two Lévy processes with the same characteristic exponent have the same law. More precisely, they have the same one-dimensional distributions, thus by the homogeneity and independence of the increments, they have the same finite-dimensional distributions, and finite-dimensional distributions determine laws on Ω. One of our main concerns will be to relate analytic properties of the characteristic exponent with the probabilistic behaviour of the Lévy process.

The two well known Lévy processes are:

1. Poisson process, with characteristic exponent is given by:

$$\Psi(\lambda) = c(1 - e^{i\lambda}),$$

and

2. Brownian motion with characteristic exponent given by:

$$\Psi(\lambda) = \frac{1}{2} \mid \lambda \mid^2$$

Now, a third significant prototype will be introduced. Let $\xi_1, ..., \xi_n, ...$ be independent random variables all having the same distribution v on $\mathbb{R}^d - \{0\}$, and $S(n) = \xi_1 + ... + \xi_n$ be the corresponding random walk. Introduce $N = (N_t t \geq 0)$, a Poisson process with parameter $c > 0$, independent of the ξ_n's. Then, it is easy to check using the lack of memory of the exponential law that the process in continuous time,

$$e(t) = \begin{cases} \xi_n & , \text{if } N_{t-} < n = N_t \\ 0 & , \text{otherwise} \end{cases}$$

is a Poisson point process with characteristic measure cv, where 0 serves as isolated point.

Definition 12 (Compound Poisson Process) *The time-changed random walk*

$$S \circ N_t = \sum_{i-1}^{N_t} \xi_i = \sum_{0 \leq s \leq t} e(s), \text{ for } t \geq 0.$$

is a Lévy process called a compound Poisson process with Levy measure cv.

Finally, it follows from the exponential formula that

$$\mathbb{E}\left[\exp\left(i \langle \lambda, S \circ N_t \rangle\right)\right] = \exp\left[-t\psi(\lambda)\right], \ \lambda \in \mathbb{R}^d$$

where

$$\psi(\lambda) = c \int_{\mathbb{R}^d} (1 - e^{i(\lambda, x)}) v(dx),$$

a quantity that can also be written as:

$$-ic \int_{\mathbb{R}^d} (\lambda, x) 1_{\{|x|<1\}} v(dx) + \int_{\mathbb{R}^d} (1 - e^{i(\lambda,x)} + i(\lambda, x) 1_{\{|x|<1\}}) ev(dx), \tag{2.3}$$

Equation (2.3) is a special case of the Lévy-Khintchine formula. The Levy measure is cv, and it appears as the intensity of jumps of $S \circ N$. This also suggests the possibility of approaching a given infinitely divisible law using a sequence of compound Poisson processes.

Another important sub-family of Lévy processes is the follow:

Definition 13 (α−Stable Lévy Process) *For every $\alpha \in (0, 2]$, a Lévy process with characteristic exponent Ψ is called a stable process with index α if $\Psi(k, \lambda) = k^{\alpha}\Psi(\lambda)$ for every $k > 0$ and $\lambda \in \mathbb{R}^d$.*

The scaling property may be shown as follows. For every $k < 0$, the rescaled process $(k^{-1/\alpha}X_{k,t}, t \geq 0)$ has the same law as X. For $\alpha \neq 2$, the Lévy measure of a stable process of index α can be expressed in polar co-ordinates $(r, \vartheta) \in [, \infty) \times S_{d-1}$ (where S_{d-1} is the unit sphere of the Euclidean space $\mathbb{R}^d$) in the form

$$\Pi(dr, d\vartheta) = r^{-\alpha-d}dr v(d\vartheta),$$

where v is some finite measure on S_{d-1}. Because Π is a Lévy measure, one must have $\int_0^{\infty}(1 \wedge r^2)r^{-\alpha-1} < \infty$, which explains the restriction on the range of the index α.

The main result of this sub-section is that any infinitely divisible probability measure μ on $\mathbb{R}^d$ can be viewed as the distribution of a Lévy process evaluated at time 1 (the converse is obvious). The proof provides an explicit construction of this Lévy process, and sheds a probabilistic light on the Lévy-Khintchine formula.

Theorem 14 (Lévy-Khintchine Formula) *Consider $a \in \mathbb{R}^d$, a positive semi-definite quadratic form Q on $\mathbb{R}^d$ and a measure Π on $\mathbb{R}^d - \{0\}$ such that $\int(1 \wedge |x|^2)\Pi(dx) < \infty$. Put for every $\lambda = \mathbb{R}^d$*

$$\Psi(\lambda) = 1(a, \lambda) + \frac{1}{2}Q(\lambda) + \int_{\mathbb{R}^d}(1 - e^{i(\lambda, x)} + i(\lambda, x)1_{\{|x|<1\}})\Pi(dx).$$

Then there exists a unique probability measure $\mathbb{P}$ on Ω under which X is a Lévy process with characteristic exponent Ψ. Moreover, the jump process of X, namely $\triangle X = (\triangle X_t, t \geq 0)$, is a Poisson point process with characteristic measure Π.

Proof. Consider $B = (B_t, t \geq 0)$, a Brownian motion in $\mathbb{R}^d$, and $\triangle = (\triangle_t, t \geq 0)$, an independent Poisson point process with characteristic measure Π. Let $\sqrt{Q}$ be any matrix such that $\langle\sqrt{Q}\lambda, \sqrt{Q}\lambda\rangle = Q(\lambda)$, and put $X_t^{(1)} = \sqrt{Q}B_t - at(t \geq 0)$. Using the Gaussian property of B, it is immediate that X^1 is a Levy process with characteristic exponent

$$\Psi^{(1)}(\lambda) = i\langle a, \lambda\rangle \frac{1}{2}Q(\lambda).$$

Let us next dwell on large values of $\triangle$ and introduce

$$\triangle_t^{(2)} = \begin{cases} \triangle_t & \text{if} \quad |\triangle_t| \geq 1, \\ 0 & \text{otherwise} \end{cases}$$

Observe that $\triangle^{(2)}$ is a Poisson point process with characteristic measure $\Pi^{(2)}(dx) = 1_{\{|x|<1\}})\Pi(dx)$. The total mass of $\Pi^{(2)}$ is finite and $\triangle^{(2)}$ is discrete. We consider the partial sum, $X_t^{(2)} = \sum_{s\leq t} \triangle_s^{(2)}(t \geq 0)$. This process has stationary independent increments and its sample paths are right-continuous with left limits. Hence $X^{(2)}$ is a Lévy process (actually, a compound Poisson process since the process that counts the jumps of $X^{(2)}$ is a non-degenerate Poisson process), and by the exponential formula of section 0.5, its characteristic exponent is

$$\Psi^{(2)}(\lambda) = \int_{\mathbb{R}^d} (1 - e^{i\langle\lambda,x\rangle} 1_{\{|x|<1\}})\Pi(dx).$$

Then we need to deal with the small values of $\triangle$ and introduce

$$\triangle_t^{(3)} = \begin{cases} \triangle_t & \text{if} \quad |\triangle_t| \geq 1, \\ 0 & \text{otherwise} \end{cases}$$

Note that $\triangle^{(3)}$ is a Poisson point distribution with characteristic measure $\Pi^{(3)}(dx) = 1_{\{|x|<1\}})\Pi(dx)$, and is independent of $\triangle^{(2)}$ (because $\triangle^{(2)}$ and $\triangle^{(3)}$ are two Poisson point processes in the same filtration which obviously never jump simultaneously). Consider for every $\varepsilon > 0$ the process of compensated partial sums,

$$X_t^{(\varepsilon,3)} = \sum_{s\leq 1} 1\{\varepsilon < |\triangle_s| < t\} \triangle_s - t\int_{\mathbb{R}^d} x 1_{\{\varepsilon<|x|<1\}}\Pi(dx) \qquad (t \geq 0).$$

On the one hand, by the same argument as above, $X^{(\varepsilon,3)}$is a Levy process with characteristic exponent

$$\Psi^{(\varepsilon,3)}(\lambda) = \int_{\mathbb{R}^d} (1 - e^{i(\lambda,x)} + i(\lambda,x)1_{\{\varepsilon<|x|<1\}})\Pi(dx).$$

On the other hand, the maximal inequality for compensated sums of section 0.5 yields for every $t > 0$ and $\eta \in (0, \varepsilon)$,

$$\mathbb{E}(\sup \left|X_S^{(\varepsilon,3)}\right|^2) \leq 4t \int_{\mathbb{R}^d} |x|^2 1_{\{\varepsilon<|x|<1\}})\Pi(dx).$$

Because the integral $\int(1 \wedge |x|^2 \Pi(dx)$ converges, this last quantity tends to 0 as ε goes to 0 |. Hence $(X^{(\varepsilon,3)}, \varepsilon > 0)$ is a Cauchy family for the norm

$$\|Y\| = \mathbb{E}(\sup\{|Y_S|^2, 0 \leq s \leq t\})^{1/2}$$

Its limit as ε goes to 0, denoted by $X^{(3)}$, has stationary independent increments and its sample path is right-continuous with left limits. It is a Levy process with characteristic exponent

$$\Psi^{(3)}(\lambda) = \int_{\mathbb{R}^d} (1 - e^{i(\lambda,x)} + i\langle\lambda, x\rangle)1_{\{|x|<1\}}\Pi(dx).$$

Moreover $X^{(3)}$ is measurable in the sigma-field generated by $\Delta^{(3)}$ and this is independent of $X^{(2)}$.

Finally, $X^{(1)} + X^{(2)} + X^{(3)}$ is a Lévy process with characteristic exponent $\Psi = \Psi^{(1)} + \Psi^{(2)} + \Psi^{(3)}$ (because $X^{(1)}, X^{(2)}$ and $X^{(3)}$ are independent). By construction, its jump process is Δ. Its law $\mathbb{P}$ on Ω fulfils the requirements of the theorem. ■

Remark 15 *The proof of Lévy-Khintchine formula actually says a bit more than was stated. Recall that a Lévy process can be represented as the sum of three independent Lévy processes* $X^{(1)}, X^{(2)}$ *and* $X^{(3)}$, *where* $X^{(1)}$ *is a linear transform of a Brownian motion with drift (and in particular is continuous),* $X^{(2)}$ *is a compound Poisson process having only jumps of size at least 1, and finally* $X^{(3)}$ *is a pure-jump martingale having only jumps of size less than 1. This decomposition is clearly unique. It is interesting to note that the compound Poisson component* $X^{(2)}$ *does not contribute to the initial sample path behaviour of a Levy process. This allows us to reduce some general studies to the case when the Lévy measure has compact support; in a* *picturesque style, one says that one throws away the big jumps.*

2.3.2 Truncated Lévy processes

Although, modelling price increments by Lévy processes does account for the fat tails and self-similarity behaviours. However, as Cont, Potters, Bouchaud (1997) pointed out, empirical evidences have shown that the scale invariance does not hold at various time-scale and the self-similarity holds only for short time scale less than a week ($\widehat{\alpha} \in [1.7, 2]$, typically) and finite variances are shown. Further empirical evidences have shown tails from Lévy law is too fat to fit the real price movements, i.e., the actual price distribution tails is fatter than Gaussian but thinner than Lévy such that the tail behaviors demonstrated an exponential tail pattern.

Matacz (1998) suggested that the truncated Lévy processes can better capture the price changes in that the central part of the price change distributions coincide with stable law, but the tails tend to zero faster than power laws. The characteristic function for $\alpha \neq 1$,

$$\phi(\theta) = \exp\left\{ -\frac{c^{\alpha}}{\cos\left(\frac{\pi\delta}{2}\right)} \left[\left(\kappa^2 + \theta^2\right)^{\frac{\delta}{2}} \cos\left(\delta \arctan\left(\frac{|\theta|}{\kappa} \right) - \kappa^{\delta} \right) \right] \right\},$$

where κ is the threshold parameter. The truncated Lévy distribution are specified in such a way that it possesses all the finite moments. The probability density function (p.d.f.) $f(x)$ is

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi(\theta) e^{i\theta x} d\theta,$$

The k^{th} moments of the distribution

$$E\left[X_t^k\right] = \left[i^k \frac{\partial^{(k)}}{\partial \theta^k} \phi(\theta, t) \right]_{\theta=0}$$

The tails has a form

$$F_{Tail}(x) \to \frac{c^{\alpha}\Gamma(1+\alpha)\sin\left(\frac{\pi\alpha}{2}\right) e^{-\kappa|x|}}{\pi |x|^{\alpha+1}}, \text{ as } |x| \to \infty$$

Thus, the variance is

$$\sigma_{TL}^2 = \frac{\alpha(1-\alpha)c^{\alpha}\kappa^{\alpha-2}}{\cos\left(\frac{\pi\alpha}{2}\right)},$$

and the kurtosis is

$$\kappa_{TL} = \frac{\cos\left(\frac{\pi\alpha}{2}\right)(\alpha-2)(\alpha-3)}{\alpha(1-\alpha)c^{\alpha}\kappa^{\alpha}}.$$

As the first order approximation,

$$E\left[e^{nX}\right] \simeq e^{\frac{n^2 E[X^2]}{2}}, \text{ for } \kappa^2 \gg n^2.$$

2.3.3 Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs) are non-linear, non-parametric statistical models that depends on the entire dataset to determine its structure and parameters. Thus, these models allows the data to be fully utilised and requires no restrictive assumptions about parametric models. The structural and methodological developments of ANNs have received contributions from researchers in cognitive science, neuroscience, psychology, biology, computer science, mathematics, physics, and statistics. ANNs are widely used in signal processing, medical imaging and economic modeling to name but a few. Over the past decade, due to the rapid advances in computational power, the general high quality of financial data and the scarcity of suitable financial models to explain these data, ANN's have received a lot of attention from financial researchers. Qi (1996) provides a general review of the origins, structure, implementation and interpretation of ANNs.

As the name suggests, ANN were originally developed to simulate the thought processes that takes place in a human brain. Thus, ANNs have powerful pattern recognition or "learning" properties that may potentially outperform many of the current modelling techniques. Since ANNs are defined by the way it "learns" from the data, thus they are commonly identified by its learning rules. In essence, a learning rule is an estimation method for the parameters and model. Currently, the most common ANNs for statistical use are:

1. Multilayer feed-forward back propagation models

2. Recurrent and statistical networks

3. Associative memory networks

4. Self-organisation networks

Of those listed above, the multi-layer feed-forward back propagation network is the most popular ANN for financial application.

Implementation and Interpretation of ANN

Although, the usefulness of ANNs are largely due to its independence of parametric assumptions. However, there are several limitations that may limit its use:.

1. There is no formal methodology for determining optimal network structure. The appropriate number of layers and middle layer units must be determined by experimentation.

2. There is no optimal algorithm for determining the global minimum for error.

3. Statistical inference on ANNs is impossible due to its lack of statistical properties.

4. It is difficult to interpret a trained ANN model

5. ANN's are often described as "black boxes", because they produce some interesting and useful results but no confidence bounds may be calculated on the output.

These issues must be addressed before ANNs can be accepted as a mainstream statistical modelling tool. Cheng and Titterington (1994) believe this calls for further study in the following three areas:

1. Mathematical modeling of real cognitive processes

2. Theoretical investigations into networks and neurocomputing

3. Development of useful tools for practical prediction and pattern recognition

Financial Applications of ANNs

ANNs have been applied in the following areas of finance:

1. **Option Pricing**. Hutchinson, Lo, Poggio (1994) have shown that ANNs is successful when the parametric Black-Scholes model fails and can outperform the Black-Scholes model. Qi and Maddala (1995) have shown that an ANN trained with two years worth of daily data outperformed the Black-Scholes model out of sample.

2. **Bankruptcy Prediction**. Using discriminant analysis(DA) and the logit model, one can model ruin using ANNs. A model developed by Tam and Kiang (1992) uses jackknife methodology for determining type I (misclassifying a failed bank as a non-failed bank) and type II (misclassifying a non-failed bank as a failed bank) errors. Decision makes can then choose a trade-off between these two errors. Salchenberger, Cinar, Lash (1992) have shown that multi-layer ANNs outperform alternatives even with a reduced data set.

3. **Exchange Rate Forecasting.** Exchange rates are highly unpredictable and the use of linear models to predict their movements has severe limitations. ANNs have been used to model exchange rates with mixed success. Models by Kuan and Liu (1995) had mediocre performance while Abu-Mostafa (1995) and Hsu, Hsu, Tenoria (1995) document more successful models. However, no studies exist comparing the ANN performance to the benchmark models.

4. **Stock Market Prediction**. Existing methods such as CAPM and APT have not been successful in predicting stock returns. Research indicates that non-linear ANN's are more successful at predicting stock returns than linear models and neural networks have been shown to be superior to linear regression in predicting returns. However, no results exist at present to provide evidence for the forecastability of ANN's in this regard.

Summary

ANNs can thus be seen as potentially useful tools. But, despite its 60 years of existence, little work has been done on the statistical properties of ANNs, which have placed severe constraints on its usefulness. Nevertheless, ANNs are very efficient in option pricing and bankruptcy prediction, but less so in the areas of forecasting stock returns and exchange rates. Research is required into optimal network structure, minimisation techniques, statistical properties and interpretation of ANNs. In addition, for ANN"s to become mainstream statistical tools, areas in which it clearly outperform current modelling techniques must be identified and developed further.

Chapter 3

THE COHERENT MARKET HYPOTHESIS AND MARKET STATE TRANSITION

The aims of this chapter are as follows:

1. Motivate and develop Vaga (1990)'s Coherent Market Hypothesis (CMH), and
2. Carry out the second and third objectives for this dissertation, namely:

Objective 2 :Calculate the mean, variance, skewness coefficient and excess kurtosis of the log returns in each state.

Objective 3 :Graphically explore the relationships between the first four moments of the log returns in each state and the parameters of the CMH return distribution.

Although, in principle, it is obvious that the four parameters $\alpha-$stable processes should fit the distribution of price changes data better than the two parameter normal distribution. However, one of the major drawbacks of $\alpha-$stable processes that have prevented it from being the obvious practical model of price changes is that the parameters does not have direct financial interpretations. In addition, it is computationally difficult to estimate its parameters from the data as the probability density form often does not have a closed form. This motivates the search for an alternative stochastic process whose parameter have direct financial interpretation and, if possible, have a closed for solution for its probability density function.

3.1 A Stochastic Model for the Formation of Public Opinion

In this section, a stochastic model for the formation of public opinion will be formulated. Like the modelling of many observed processes in economics, finance and sociology, this will be achieved by making analogies between the process with an established mathematical model and the process required to be modelled. The modelling procedure is as follows:

1. Identify the process that is required to be modelled.
2. Select a mathematical model with well-established theory such that the process modelled by this chosen model is similar to the process that is required to be modelled.
3. Carefully identify and justify the analogies between the two processes.
4. If the analogies can be reasonably justified, then one should be able to re-interpret the parameters of the established model to describe the key features of the process being modelled.
5. Finally, use the established model to model the process that is seemingly unrelated to the process described by the established model.

The aim of this section is to apply the procedure described above to the modelling of indicators for financial market fluctuations.

3.1.1 Introduction

The behaviour of ferromagnets is a carefully studied problem in statistical physics. In effect, there are independent domains of activity within a ferromagnetic material, both at the microscopic and the macroscopic level:

1. At the microscopic level, each individual molecule act as a tiny magnet, interacts with neighbouring magnets while being subjected to external magnetic and random thermal forces. It is these molecular interactions that determine whether useful state transitions in a particular material will take place.

2. At the macroscopic level, all the microscopic molecular interactions averages out to give a total magnetisation that is either in the "spin up" phase or the "spin down" phase.

The problem at the macroscopic level is that random fluctuations in external thermal forces will cause unexpected spontaneous phase transition (from one spin direction to the opposite) in the ferromagnetic material's total magnetisation, which leads to a bit error in memory. Although, such problem can be eliminated by employing some error detection and correction methods, but this leads to a loss of efficiency due to the additional bits of information required to carry out the function.

3.1.2 Ferromagnetism and the Ising Model of Ferromagnetic Phase Transition

In order to make analogies between ferromagnetism and behaviour of social groups, it is important to first understand the key dynamics and properties of ferromagnetic materials. Then, it is important to understand the assumptions that underlies the development of the Ising model, which models the phase transition of ferromagnetic materials. Neat overviews of principles in statistical mechanics, ferromagnetism and Ising model are provided by Giordano (1997) and Girifalco (2000).

For a macroscopic system like the ferromagnetic materials, it is impossible to describe, or investigate the exact microscopic behaviour of each individual particle in such a system. Thus, analyses on these systems is limited to its average properties as measured by thermodynamic quantities like the temperature or pressure and correlation functions. Statistical mechanics provides the connection between the microscopic and macroscopic level by supplying methods to calculate the macroscopic properties from the microscopic information, like the interaction energy between the particles. The essential tool in statistical mechanics is the probability distribution, i.e. the collection of occupancies of different configurations (microscopic states).

In statistical mechanics, magnetism is explained in terms of:

1. The electron spin of the molecules making up the solid state material, and
2. The magnetic moment associated with the spins.

Since spin is a quantum mechanical phenomena, thus magnetism may be regarded as an inherently quantum phenomena. Ferromagnetic materials are viewed as systems of many interacting electron spins that exhibits the phase transition phenomena under changes in external environment, particularly that of temperature. The phenomena of ferromagnetism arises when the interaction of a collection of electron spins causes an overall alignment of these spins in the same direction. This alignment leads to a total moment that has a macroscopic value. The essential features of ferromagnetic phase transitions that needs to be understood are:

1. The mechanism of the interaction between electron spins on a microscopic level that gives rise to the overall alignment on the macroscopic level, and
2. The dynamics of the observed temperature dependence of the ferromagnet's magnetic properties.

The spin model of a ferromagnet consists of a collection of magnetic moments. The magnetic moments can be thought of as atoms with:

$$Spin_{atom} = \frac{1}{2}\,(magnetic\ moments)\,.$$

Then, the following assumptions are made in the model building process:

1. **Condition 16**

   i) Electron spins are situated on a regular lattice.

   ii) Each electron spin can only point vertically upwards $(+z)$, or vertically downwards $(-z)$. No other orientations is possible.

   iii) Microscopic interaction only exists between spins that are nearest neighbours (i.e. there is no long-range correlation).

iv) The exchange constant $J \geqslant 0$.

v) Spin system is in equilibrium with external environment (called heat bath *in physics) at temperature T.*

Definition 17 (Ising Spin) *Suppose that the ferromagnet consists of N electron spins and that condition i) and ii) holds. Then, each electron spin is called an* Ising spin. *For $i = 1, \cdots, N$, define:*

$$\begin{aligned} s_i &= \text{value of the } i^{th} \text{ Ising spin} \\ &= \begin{cases} +1, & \text{if } i^{th} \text{ spin is pointing vertically up} \\ -1, & \text{if } i^{th} \text{ spin is pointing vertically down} \end{cases}. \end{aligned}$$

In any spin system, each Ising spin interacts with every other spins in the regular lattice. In order to determine the energy of the spin system, it is necessary to first determine the distance at which there are significant forces of interaction between spins. It is noted that for a permanent magnetic material, the force of interaction between spins is:

$$\begin{cases} \text{Largest} & \text{, for spins that are nearest neighbours} \\ \text{Rapidly vanishing} & \text{, for increasing distance between two spins} \end{cases}.$$

The above observation motivates condition iii).Thus, the simple Ising model is defined as follows:

Proposition 18 (Simple Ising Model) *Suppose conditions i) to iv) holds. Then, the energy function for the Ising spin system defined by:*

$$E = -J \sum_{\langle i,j \rangle} s_i \cdot s_j, \tag{3.1}$$

where the sum is over all pairs of nearest neighbour Ising spins $\langle i, j \rangle$, and $J \geqslant 0$ is called the simple Ising model.

From the energy function (3.1) defined above, two neighbouring spins will produce an energy of interaction

$$E = \begin{cases} -J & \text{, if both spins point in same direction (parallel)} \\ +J & \text{, if spins point in opposite direction (antiparallel)} \end{cases}$$

Since $J \geqslant 0$ by condition iv), thus the interaction between Ising spins in the system will favour the parallel alignment of neighbouring spins (called *coupling*). But, when each spin is parallel to its neighbour, then every spin will be parallel to every other spin in the regular lattice. Thus, there is an alignment of all of the magnetic moments, which leads to a non-zero magnetic moment for the spin system as a whole. This produces the ferromagnet.

Definition 19 (Spontaneous Magnetisation) *A system have a spontaneous magnetisation if it has a non-negative magnetic moment in the absence of an external magnetic field.*

Now, the disordering effect of temperature must be accounted for. Assume condition v) holds. Then, the macroscopic behaviour of the material is described by the canonical ensemble.

Definition 20 (Canonical Ensemble) *A canonical ensemble is a collection of closed systems at constant temperature.*

Remark 21 *Closed systems have fixed number of particles and volume, but can exchange energy with their surroundings. Thus, energy of a closed system is not constant.*

Given that the macroscopic behaviour may be described by the canonical ensemble, thus a fundamental result of statistical mechanics for a system in equilibrium with its environment follows:

Proposition 22 *For a system in equilibrium with its surrounding environment, the probability of finding the system in any particular state is proportional to the Boltzmann factor:*

$$\begin{aligned} P_\alpha &\equiv \Pr(\textit{system in state } \alpha) \\ &\propto \exp\left(-\frac{E_\alpha}{\kappa_B T}\right) \\ &= \frac{\exp\left(-\frac{E_\alpha}{\kappa_B T}\right)}{\sum_\alpha \exp\left(-\frac{E_\alpha}{\kappa_B T}\right)}, \end{aligned}$$

where

$$E_\alpha \equiv \text{energy of state } \alpha \text{ calculated from (3.1), and}$$
$$\kappa_B \equiv \text{Boltzmann's constant.}$$

Definition 23 (Microstate of System) *A microstate of the system is defined as the particular configuration of spins associated with each of the state α.*

On the microscopic level, it is the interaction of the spin system with a surrounding environment (hence gaining or losing energy) that causes the system to undergo transitions from one microstate to another. On the macroscopic level, measures of macroscopic property effectively averages over the many microstates that the system assumes during the course of the measurement. Thus, to calculate the macroscopic behaviour, the probability that the system will be in microstate α is required

$$P_\alpha = \Pr\left[\text{system in microstate } \alpha\right].$$

An important measure of macroscopic property of ferromagnetic material is given by the total magnetic moment, which will be referred to as the measured magnetisation:

Definition 24 (Measured Magnetisation)

$$\begin{aligned} M &\equiv \text{Measured magnetisation of the system} \\ &= \sum_\alpha M_\alpha \cdot P_\alpha, \end{aligned}$$

where:

$$\begin{aligned} M_\alpha &\equiv \text{Total magnetic moment of microstate } \alpha \\ &= \sum_j s_{j,\alpha}, \end{aligned}$$

and

$$s_{j,\alpha} = \text{spin value corresponding to spin directions in microstate } \alpha.$$

Remark 25 *From the above definition, the measured magnetisation can thus take stricity positive (average spin up) or strictly negative (average spin down) values each representing one of the two possible macroscopic state. For a zero measured magnetisation, there is a net cancel out in spin orientation, thus no magnetisation.*

For a lattice containing N Ising spins, each spin can be in one of two states. Thus, there will be a total of 2^N possible distinct microstates as a whole. Furthermore, for the sake of generality, the systems of interest are ones with $N \to \infty$, which means that the number of microstates α will increase exponentially. When required to calculate M where there are a large number of terms in the sum, it is sometimes possible to ignore many terms due to its negligible value. However, near a phase transition such simplification is not possible, because all the terms in the sum will have significantly large value in that case.

Remark 26 *Although, various generalisations of the simple Ising model is possible, but researches have shown that the simple Ising model captures many of the essential features of the phase transition to the ferromagnetic state.*

One of the key property of ferromagnetic materials is that they undergo phase transition due to changes in external surrounding environment, particularly external temperature. At temperatures above some critical temperature T_c say, the ferromagnet's spins point in random directions. This results in a zero macroscopic magnetisation, because the spins' magnetic moments cancel each other out when summed. However, at temperatures below T_c, the ferromagnet's spins become aligned (lined up in an orderly fashion), which leads to a macroscopic magnetisation when the spins' magnetic moments are summed. Thus, the order/alignment on the spins at the microscopic level have led to the emergence of a new macroscopic feature in the material. Hence, the following definitions from statistical physics:

Definition 27 (Phase) *The macroscopic property of a system resulting from aggregated microscopic states (in the case of ferromagnets, the spins' magnetic moments) is called the* phase *that the system is in.*

Definition 28 (Phase Transition) *The change from one phase (due to changes in aggregated microscopic state) to another is called* phase transition.

In statistical physics, the order parameter serves to identify the phase that a system is in. Hence,

Definition 29 (Order Parameter) *An* order parameter *is a thermodynamic function that takes on distinct values in each equilibrium phase.*

It follows that the macroscopic order parameter for a ferromagnet may be taken to be the system's measured magnetisation M, which has the following equivalent meaning:

$$\begin{aligned} M &= \sum_{\alpha} M_{\alpha} \cdot P_{\alpha} \\ &= (\text{Number of spins pointing one direction}) - (\text{Number of spins pointing the opposite direction}) \end{aligned}$$

where M_{α} and P_{α} are as defined in the definition of M. Thus, if the number of spins pointing in one direction is nearly equal to the number of spins pointing in the other direction, then the ferromagnet's order parameter will take on a small number (and 0, if exactly the same number point in each direction). Under certain external conditions imposed on the ferromagnet, the direction that the spins are pointing may become highly polarised (i.e., nearly all pointing the same direction), which leads to a high degree of macroscopic order when aggregating over these microscopic states.

In the ferromagnet, there are essentially two types of phase transition that may take place:

1. **First order transition** At the critical (Curie) temperature, the two possible macroscopic states coexists and there is a discontinuous change in measured magnetisation and hence the phase of the material. The correlation length between spins in the system are finite.

2. **Second order transition**. At the critical (Curie) temperature, there is a continuous change in the measured magnetisation and hence phase of the material. The correlation length between spins in the system becomes infinite. Empirically, magnetic phase transition is usually observed to be second order.

To appreciate the Ising model, consider a bar of iron in which individual molecular magnetic spins point either up or down:

1. When the bar's surrounding temperature is high. The thermal energy supplied will cause each molecule within the system to be randomly bombarded by its neighbouring molecules. The force of the random bombardment will exceed the weak magnetic interaction between adjacent molecules, which results in molecular orientation becoming disordered. Although, on the microscopic level, more molecules may point up than down, or vice versa from time to time. However, the macroscopic magnetic field will fluctuate randomly around zero, in a state of disorder.

2. When the temperature of the iron bar is lowered below some critical point. In this case, the magnetic interaction between adjacent molecules becomes stronger than the force of random bombardment caused by external thermal energy. Thus, if there is an external magnetic force aligning some clusters of molecules in one direction, then the neighbouring molecules will tend to follow and most clusters will align in the same direction. Although, some clusters will align in the other direction due to random bombardment caused by external thermal energy. But, on averarge, fluctuations in macroscopic magnetic field will be stable around some large net value. In this situation, the iron has a high degree of order. However, the orientation of the macroscopic magnetic field may still "flip" as a result of changes in external forces or even by chance.

3.1.3 Derivation of the Ising Model By Diffusion Approximation

The distinguishing feature of the **Ising model of ferromagnetic phase transition** is that it is one of the rare interacting many-particle system in statistical physics that allowed for a full and accessible description of its microscopic molecular behaviour. Although, microscopically, on a ferromagnetic's atomic lattice sites, the development of magnetic moments and the subsequent ordering of the moments is an extremely complex process involving the cooperative phenomena of many interacting electrons. Yet, different macroscopic ferromagnetic systems displays the same thermodynamic properties, particularly the temperature dependence of specific heat. This is due to the fact that the averaged variables, appropriate for collective long-wavelength and time scales at the microscopic level, are slowly varying continuous fields describing the collective motion of a macroscopic set of particles.

Derivation of the Fokker-Planck Equation

First, to quantify the Ising model of ferromagnetic crystals, define the following:

Definition 30 *Assume that there is a population of n spins in the ferromagnetic crystal. Then,*

$$N_+ \equiv \textit{Number of spins out of } n \textit{ with spins pointing up,}$$

and

$$N_- \equiv \textit{Number of spins out of } n \textit{ with spins pointing down.}$$

The probability of finding N_+ molecules pointing up, and N_- molecules pointing down at any given time t is given by the following transition distribution function:

$$f(N_+, N_-, t) = \Pr[N_+ \text{ spins up } \textbf{AND } N_- \text{ spins down at time } t]. \tag{3.2}$$

To evaluate transition distribution function (3.2) above, it is necessary to first evaluate the transition probabilities. The transition probability is defined as follows:

Definition 31 *The likelihood of a single molecule flipping its orientation from a spin pointing down to spin pointing up and vice versa is defined by the transition probabilities:*

$$p_{-+}(N_+, N_-) \equiv \textit{probability of transition from } (N_+, N_-) \textit{ to } (N_+ + 1, N_- - 1), \tag{3.3}$$

and

$$p_{+-}(N_+, N_-) \equiv \textit{probability of transition from } (N_+, N_-) \textit{ to } (N_+ - 1, N_- + 1). \tag{3.4}$$

From principles of physics, the rate of change in the population must satisfy the continuity equation:

$$\frac{df(N_+, N_-, t)}{dt} = (\text{Gains into } (N_+, N_-)) - (\text{Decreases from } (N_+, N_-))$$

Now, Gains into the state (N_+, N_-) may come from neighbouring states:

1. A state with one more molecule pointing up and one less pointing down has a certain probability for a spin to flip from up to down and, therefore, go into the state (N_+, N_-).

2. A state with one less molecules pointing up and one more pointing down has a probability of a molecule flipping from down to up and going into the state (N_+, N_-).

Decreases from the state (N_+, N_-) takes place at the rate spins flip from down to up and go into the $(N_+ + 1, N_- - 1)$ state and vice versa.

Based on the above analysis, it can be shown that the "master equation" is given by:

$$\begin{aligned}
&\frac{df(N_+, N_-, t)}{dt} \\
&= (\text{Increases in population}) - (\text{Decreases in population}) \\
&= [(N_+ + 1) \cdot p_{+-}(N_+ + 1, N_- - 1)] \cdot f(N_+ + 1, N_- - 1; t) \\
&\quad + [(N_- + 1) \cdot p_{-+}(N_+ - 1, N_- + 1)] \cdot f(N_+ - 1, N_- + 1; t) \\
&\quad - [N_+ \cdot p_{+-}(N_+, N_-) + N_- \cdot p_{-+}(N_+, N_-)] \cdot f(N_+, N_-; t) . \qquad (3.5)
\end{aligned}$$

Determination of Transition Probability

In order to derive the Fokker-Planck equation, the transition probabilities in the master equation (3.5) needs to be explicitly determined. First, transform the discrete-valued variables (N_+, N_-) into another discrete valued variable r, which will be chosen as the order parameter:

Definition 32 *The order parameter describing the macroscopic magnetic field is:*

$$\begin{aligned}
r &\equiv \textit{measure of the net polarisation of Ising spin} \\
&= \frac{N_+ - N_-}{2(N_+ + N_-)} = \frac{N_+ - N_-}{2n},
\end{aligned}$$

where:

$$\begin{aligned}
n &\equiv \textit{Total number of spins in the Ising model} \\
&= N_+ + N_- .
\end{aligned}$$

Then, approximate the discrete process r by a diffusion process, thus the following assumption is required:

Assumption 3: Population of spins is sufficiently large to treat r as a continuous parameter.

Since the individual interactions between the spins are too complex, thus Haken (1983) suggested to write down the transition probabilities p by plausibility arguments. Now, consider the competing forces that will influence the transition probabilities of molecular spin in ferromagnetic materials:

1. Internal coupling of magnetic forces between individual and neighbouring molecules,
2. External magnetic field influencing individual molecule's orientation, and
3. External temperature surrounding ferromagnetic material.

Hence, the transition probabilities may be *quantitatively* expressed as follows:

$$\begin{aligned} p_{+-}(N_+, N_-) &\equiv p_{+-}(r) \\ &= \nu \cdot \exp\left[\frac{-(Ir+H)}{\kappa_B T}\right] \\ &= \nu \cdot \exp\left[\left(\frac{-I}{\kappa_B T}\right) r + \frac{-H}{\kappa_B T}\right] \\ &= \nu \cdot \exp\left[-(kr+\rho)\right], \end{aligned} \tag{3.6}$$

and

$$\begin{aligned} p_{-+}(N_+, N_-) &\equiv p_{-+}(r) \\ &= \nu \cdot \exp\left[\frac{Ir+H}{\kappa_B T}\right] \\ &= \nu \cdot \exp\left[\left(\frac{I}{\kappa_B T}\right) r + \frac{H}{\kappa_B T}\right] \\ &= \nu \cdot \exp\left[kr+\rho\right], \end{aligned} \tag{3.7}$$

where:

$$\begin{aligned}
\nu &= \text{Rate of molecular spin flips,} \\
I &= \text{Magnetic coupling between the molecular spins,} \\
H &= \text{External magnetic field,} \\
\kappa_B &= \text{Boltzmann's constant,} \\
T &= \text{External temperature,} \\
\kappa &= \frac{I}{\kappa_B T}, \text{ and} \\
\rho &= \frac{H}{\kappa_B T}.
\end{aligned}$$

Now, set the transition probabilities as:

$$\begin{aligned}
w_{+-}(r) &\equiv N_+ p_{+-}(N_+, N_-) \\
&= N\left(\frac{1}{2}+r\right) p_{+-}(r),
\end{aligned}$$

and

$$\begin{aligned}
w_{-+}(r) &\equiv N_- p_{-+}(N_+, N_-) \\
&= N\left(\frac{1}{2}-r\right) p_{-+}(r). \qquad (3.8)
\end{aligned}$$

Then, the master equation (3.5) may be re-written as:

$$\begin{aligned}
&\frac{\partial f(r,t)}{\partial t} \\
= &\left\{\left[w_{+-}(r) + w_{-+}(r)\right] \cdot f(r,t) + \left[w_{+-}(r+\Delta r)\right] \cdot f(r+\Delta r, t)\right\} \\
&-\left[w_{-+}(r-\Delta r)\right] \cdot f(r-\Delta r, t), \qquad (3.9)
\end{aligned}$$

where:

$$\Delta r = \frac{1}{n} = \frac{1}{N_+ + N_-}.$$

A Taylor expansion of the finite difference equation (3.9) leads immediately to the following Fokker-Planck equation governing the probability distribution for finding the system in any particular state r:

$$\frac{\partial f(r,t)}{\partial t} = -\frac{\partial}{\partial r}\left[K(r)\cdot f(r,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial r^2}\left[Q(r)\cdot f(r,t)\right], \tag{3.10}$$

where the drift and diffusion coefficients are given respectively by:

$$\begin{aligned} K(r) &= \Delta r\cdot\left[w_{-+}(r) - w_{+-}(r)\right], \text{ and} \\ Q(r) &= \Delta r\cdot 2\left[w_{+-}(r) + w_{-+}(r)\right]. \end{aligned}$$

Time-Stationary Solution to the Homogeneous Fokker-Planck Equation

For a special stationary case, a closed form solution of Fokker-Planck equation (3.5) may be derived. Define the time-homogeneous transition density function by:

$$f_{stat}(r) = \lim_{t\to\infty} f(r,t),$$

and assume that the Fokker-Planck equation is homogeneous. Then, (3.10) is transformed into the following partial differential equation:

$$0 = -\frac{\partial}{\partial r}\left[K(r)\cdot f_{stat}(r)\right] + \frac{1}{2}\frac{\partial^2}{\partial r^2}\left[Q(r)\cdot f_{stat}(r)\right]. \tag{3.11}$$

Furthermore, note that there is a reflecting barrier at $r = -\frac{1}{2}$. Thus, the boundary condition to the above partial differential equation is given by:

$$\frac{1}{2}\frac{\partial^2}{\partial r^2}\left[Q(r)\cdot f_{stat}(r)\right]\Big|_{r=-\frac{1}{2}} = K\left(-\frac{1}{2}\right)\cdot f_{stat}\left(-\frac{1}{2}\right) \tag{3.12}$$

The general solution to PDE (3.11) subject to boundary condition (3.12) may be found by quadrature, and is called the Ising model of ferromagnetism:

Proposition 33 (The Ising Model of Ferromagnetism) *The Ising model of the order pa-*

rameter r is given by:

$$f_{stat}(r) = C \cdot Q^{-1}(r) \cdot \exp\left[2 \int\limits_{u=-\frac{1}{2}}^{r} \frac{K(u)}{Q(u)} du\right], \tag{3.13}$$

where the normalising constant is defined as follows:

$$C^{-1} = \int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} Q^{-1}(r) \exp\left[2 \int\limits_{r=-0.5}^{r} \frac{K(u)}{Q(u)} du\right] dr. \tag{3.14}$$

where:

$$\begin{aligned} K(r) &\equiv \textit{Drift coefficient} \\ &= \nu \cdot [\sinh(\kappa r + \rho) - 2r \cosh(\kappa r + \rho)], \end{aligned} \tag{3.15}$$

$$\begin{aligned} Q(r) &\equiv \textit{Diffusion coefficient} \\ &= \frac{\nu}{n} \cdot [\cosh(\kappa r + \rho) - 2r \sinh(\kappa r + \rho)], \end{aligned} \tag{3.16}$$

$$\begin{aligned} \kappa &= \frac{I}{\kappa_B T}, \\ \rho &= \frac{H}{\kappa_B T}, \\ n &= \textit{Degrees of freedom (Total number of spins)}, \\ \nu &= \textit{Rate of molecular spin flips}, \\ I &= \textit{Magnetic coupling between molecular spins}, \\ H &= \textit{External magnetic field}, \\ T &= \textit{External temperature, and} \\ \kappa_B &= \textit{Boltzmann's constant.} \end{aligned}$$

3.1.4 Modelling Social Group Behaviour: A Theory of Social Imitation

Weidlich (1971) was the first to develop a detailed statistical model of the polarisation phenomena in society to describe intense polarisation of opinions in social groups, where he analysed the French student revolution of the 1960s as follows:

1. Start with the Ising model of ferromagnetic phase transition, then
2. Re-interpret the order and control parameters to describe the polarisation of opinions and other cooperative behaviour of individuals in social groups.

Callen and Shapero (1974) generalised Weidlich's idea further to propose "A Theory of Social Imitation" in which they suggested that in fact there exists a wide variety of biological and social groups displays the characteristic transitions between disordered/structureless states and more ordered/structured states. This lead them to make the comment that "Fish aligned in schools, fireflies flashing in unison, and even humans following the dictates of fashion are examples of ordered systems to which we can apply ferromagnet theory." Callen (1974) suggested that the Ising model could be used to describe the behaviour of a wide variety of social groups, particularly the phenomenon of polarization of opinion. Their Theory of Social Imitation is based on the assumption that, on a macroscopic level, individuals in a group behave in a manner similar to the molecules in a bar of iron. Under some conditions, the individual will act independently of each other. Under the right conditions, however, the same individuals' thinking may become polarised, i.e, the individuals will act as a crowd and individual rational thinking will be replaced by a collective "group think."

According to Haken (1975), social groups can be thought of as large dynamical systems, which are typically made up of a large number of sub-systems (or *degrees of freedom*, as termed by the physicists). Subsequently, Haken (1983) suggests that there exists at least two description levels for systems that are made up of many sub-systems, namely:

1. Analysing the individual system and its interaction with its surrounding, and
2. Describing the statistical (aggregate) behaviour using macroscopic variables.

It is on the second level that interacting social groups becomes mathematically quantifiable. The macroscopic behaviour of these dynamical systems is described by its *order parameter*.

In general, transitions from disorder to order tend to share the same macroscopic characteristics, whether the sub-systems are from physics, chemistry, biology or sociology. Even though the magnet and social groups have vastly different sub-systems, their macroscopic behaviours shares the same properties of all transitions from disorder to order. Thus,if equal number of individuals are for as well as against a particular opinion, then there is no net preference (or order) for one opinion over the other. If individuals become highly polarised over a particular opinion, then there will be uniformity or order in the macroscopic behaviour of that social group.

Although, it is intuitively obvious that the phenomena of public opinion formation (and hence the polarisation of opinion) is of a cooperative nature. However, it is extremely difficult (if not impossible) to model public opinion formation on a rigorous mathematical basis due to the numerous (often unknown) factors that determines the actions of individuals.

The analogy between the physical system of ferromagnetic materials and the social groups can be motivated as follows:

1. (**Ordered States**) In ferromagnetic materials, spins point in a preferred direction and generate a large magnetic field that maintains the particular alignment long after the original aligning force is gone. Thus,

   $$\text{Ordered state} \equiv \text{Orientation of Ising spin}$$

   In social groups, public opinion within a social group may tend to persist for a long time when the individuals within the group become strongly polarised for or against a particular opinion. Thus,

   $$\text{Ordered state} \equiv \text{Agreement on a particular opinion.}$$

2. (**Critical Temperature**) In ferromagnetic materials, there is a critical temperature (called Curie temperature in statistical physics) at which the ferromagnet may abruptly change its macroscopic behaviour which leads to a phase transition from disorder to order, or vice versa. When the external temperature falls below this critical level, the material

becomes spontaneously magnetised. Thus,

Temperature ≡ Intensity of random external forces influencing total alignment/order of spins.

In social groups, when a group becomes susceptible to crowd behaviour, even slight external biasing forces can have a large, long-lasting impact on the polarisation of opinions within the group. The analogy with the Ising Model suggests that there will be a critical "social temperature" that results in crowd behaviour.

Having shown that the analogy between ferromagnet systems and social groups is plausible, it follows that it is reasonable to model social groups with the much researched Ising model.

Based on the theory of social imitation, the following assumptions are made:

Condition 34

(i) The rate of change of an individual's opinion is:

$$\begin{cases} \textit{Enhanced by the group of individuals with an opposite opinion, and} \\ \textit{Diminished by the group of individuals with the same opinion.} \end{cases}$$

(ii) There exists certain overall social climate which either promotes or discourage the change of opinion.

(iii) There exists external influences on each individual.

Now, translate the above three assumptions into a mathematical form as follows:

Firstly, re-interpret the direction of spin (spin up or spin down) as the direction of opinion (opinion for + or opinion against −). In order to avoid the debate of how best to measure public opinion, the following simplifying assumption is made:

Assumption 1: Public opinion can only have two directions: for (+) or against (−).

Then, the microscopic state of the social groups may be quantified by:

$$(N_+, N_-),$$

where:

$$N_+ \equiv \text{the number of individuals with } + \text{ opinion, and}$$
$$N_- \equiv \text{the number of individuals with } - \text{ opinion.}$$

Secondly, applying the theory of social imitation to public opinion formation, the formation of most individual's opinion in a social group is often influenced by the presence of groups of individuals with the same or opposite opinion. Thus, the formation of public opinion is a cooperative effect. Thus, there is uncertainty about what the pair (N_+, N_-) will be from one instance in time to another. In order to incorporate this uncertainty into the model, the following assumption is required:

Assumption 2:∃ probability for changes in (N_+, N_-) per unit time

Thirdly, transform the discrete-valued variables (N_+, N_-) into another discrete valued variable r, which will be chosen as the order parameter for the social group:

Definition 35 *The order parameter describing the macroscopic behaviour of social groups is:*

$$\begin{aligned} r &\equiv \textit{measure of the net polarisation of opinion in the group} \\ &= \frac{N_+ - N_-}{2(N_+ + N_-)} = \frac{N_+ - N_-}{2n}, \end{aligned}$$

where:

$$\begin{aligned} n &\equiv \textit{Total number of individuals in the social group} \\ &= N_+ + N_-. \end{aligned}$$

Fourthly, in order to approximate a discrete process r by a diffusion process, the following assumption is required:

Assumption 3: Social groups are sufficiently large to treat r as a continuous parameter.

where:

$$\Delta r = \frac{1}{n} = \frac{1}{N_+ + N_-}.$$

A Taylor expansion of the finite difference equation (3.18) leads immediately to the following re-interpreted Fokker-Planck equation governing the probability distribution for finding the social group in any particular state described by r:

$$\frac{\partial f(r,t)}{\partial t} = -\frac{\partial}{\partial r}\left[K(r) \cdot f(r,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial r^2}\left[Q(r) \cdot f(r,t)\right], \tag{3.19}$$

where the drift and diffusion coefficients are given respectively by:

$$\begin{aligned} K(r) &= \Delta r \cdot \left[w_{-+}(r) - w_{+-}(r)\right], \text{ and} \\ Q(r) &= \Delta r \cdot 2\left[w_{+-}(r) + w_{-+}(r)\right]. \end{aligned}$$

For a closed form solution of the re-interpreted Fokker-Planck equation (3.19) may be derived. Define the time-homogeneous transition density function by:

$$f_{stat}(r) = \lim_{t \to \infty} f(r,t),$$

and assume that the Fokker-Planck equation is homogeneous. Then, (3.19) is transformed into the following partial differential equation:

$$0 = -\frac{\partial}{\partial r}\left[K(r) \cdot f_{stat}(r)\right] + \frac{1}{2}\frac{\partial^2}{\partial r^2}\left[Q(r) \cdot f_{stat}(r)\right]. \tag{3.20}$$

Furthermore, note that there is a reflecting barrier at $r = -\frac{1}{2}$. Thus, the boundary condition to the above partial differential equation is given by:

$$\frac{1}{2}\frac{\partial^2}{\partial r^2}\left[Q(r) \cdot f_{stat}(r)\right]\bigg|_{r=-\frac{1}{2}} = K\left(-\frac{1}{2}\right) \cdot f_{stat}\left(-\frac{1}{2}\right) \tag{3.21}$$

The general solution to PDE (3.20) subject to boundary condition (3.21) may be found by quadrature and is gives the Ising model reinterpreted as a model of polarisation in social groups.

Proposition 36 (Model of Public Opinion Formation) *Based on the Theory of Social*

Imitation, a stochastic model for public opinion formation is given by:

$$f_{stat}(r) = C \cdot Q^{-1}(r) \cdot \exp\left[2 \int_{u=-\frac{1}{2}}^{r} \frac{K(u)}{Q(u)} du\right], \tag{3.22}$$

where the normalising constant is defined as follows:

$$C^{-1} = \int_{r=-\frac{1}{2}}^{\frac{1}{2}} Q^{-1}(r) \exp\left[2 \int_{r=-0.5}^{r} \frac{K(u)}{Q(u)} du\right] dr. \tag{3.23}$$

where:

$$\begin{aligned} K(r) &\equiv \textit{Drift coefficient} \\ &= \nu \cdot [\sinh(\kappa r + \rho) - 2r \cosh(\kappa r + \rho)], \end{aligned} \tag{3.24}$$

and

$$\begin{aligned} Q(r) &\equiv \textit{Diffusion coefficient} \\ &= \frac{\nu}{n} \cdot [\cosh(\kappa r + \rho) - 2r \sinh(\kappa r + \rho)]. \end{aligned} \tag{3.25}$$

$k \equiv$ *Coupling of opinions among individuals within a social group,*

$\rho \equiv$ *External preference/bias for one opinion over the other,*

$n =$ *Total number of individuals in the social group,*

$\nu \equiv$ *Average Rate of change of opinion,*

$I \equiv$ *Strength of adaptation to neighbouring individuals,*

$H \equiv$ *Preference parameter ($H > 0$ means opinion + is preferred to −),*

$T =$ *Social temperature*

$\kappa_B =$ *Boltzmann constant, and*

$\kappa_B T =$ *Social climate parameter.*

Remark 37 *It is important to note that there are fundamental differences between molecules in a bar of iron and people in a social group. One obvious difference is that the law of conservation of energy applies to a magnet, but does not apply to opinions in a social group.*

When the degree of coupling of opinions k exceeds some critical level, the theory predicts that there will be a transition from a state of net macroscopic disorder (r fluctuating around zero) to a state of order (r fluctuating around either two large values or some stable points far from zero).

In the absence of the external parameter, two typical shapes for $f_{stat}(r)$ are expected from a direct knowledge of the Ising model:

1. High-temperature limit. Relatively frequent changes in opinion will result in a centered distribution of opinions.

2. If the social climate factor $\kappa_B T$ is lowered, or the coupling strength between individuals is increased, then two distinct groups of opinions occurs. This can be taken as a description for the "polarisation phenomenon" in a society.

The aim is to eventually motivate that the annualised returns on an index for a financial market comprising of numerous industrial groups is a reasonable order parameter for that financial market.

3.1.5 Stability of Continuous Markov Processes

Vaga (1994) examined the conditions under which a deterministic process would be stable or become unstable. The same issue is applicable to random processes.

In this case, the characterising functions $K(r,t)$ and $Q(r,t)$ determines whether or not the process will be stable. It is convenient to define a *potential function* by:

$$\varphi(r) = -\int_{u=2}^{r} \frac{K(u)}{Q(u)} du, \tag{3.26}$$

where the characteristic functions $K(u)$ and $Q(u)$ do not explicitly depend on time.

A stationary Markov state density function, $f_{stat}(r)$, can then be expressed as:

$$f_{stat}(r) = \frac{2C}{Q(r)} \exp[-\phi(r)]. \tag{3.27}$$

The potential function is analogous to the potential energy function of classical mechanics. The potential function slopes upwards or downwards at a point, r, depending on whether $K(r)$ is positive or negative. If a process is in state, r, at time t, it will be more likely to move:

$$\begin{cases} \text{Right} & \text{, if } K(r) > 0 \text{ and } \frac{\partial}{\partial r}\phi(r) < 0 \\ \text{Left} & \text{, if } K(r) < 0 \text{ and } \frac{\partial}{\partial r}\phi(r) > 0. \end{cases}$$

The larger the value of $\frac{\partial}{\partial r}\phi(r)$, the greater the probability bias. Hence, a continuous Markov process has a tendency to move towards the nearest local minimum of the potential function, and these are the stable states or the most probable regions in which the Markov process will be found.

The stable states are separated by a *barrier function* defined as follows:

$$B(r) = \xi^{-1} \exp[\phi(r)]. \tag{3.28}$$

The barrier function is of particular importance in computing the average time required for a state transition to occur between the stable states. The problem of transition times between stable states is known as the *first exit time problem*. The mean stable state transition times for a bistable process are given by Gillespie (1992) as:

$$T(r_1 \to r_2) = \int_{r_1}^{r_2} B(r)\, dr \int_{-\infty}^{r} f_{stat}(u)\, du, \tag{3.29}$$

and

$$T(r_2 \to r_1) = \int_{r_1}^{r_2} B(r)\, dr \int_{r}^{-\infty} f_{stat}(u)\, du. \tag{3.30}$$

This can be approximated as:

$$T(r_i \to r_f) \approx \left[\int_{r_b} B(r)\, dr\right] \cdot \left[\int_{r_i} f_{stat}(u)\, du\right]. \tag{3.31}$$

Hence, the average time for a spontaneous transition is approximately equal to the product of the area under the barrier function $B(r)$ and the area under the initial stable state probability distribution.

If the slope of the distribution, $f(r)$, near a peak is approximately Gaussian and the diffusion coefficient is a constant, Q, then the mean transition time from stable state,r_i, to a stable state, r_{stat}, is approximately:

$$T(r_i \rightarrow r_f) \approx 2\pi \cdot \left(\frac{1}{\sqrt{K'(r_i) \cdot K'(r_b)}} \right) \cdot \exp \left[2 \cdot Q^{-1} \left(\left| \int_{r_i}^{r_b} K(r)\, dr \right| \right) \right]. \tag{3.32}$$

Note in (3.32) the following:

1. *size of the fluctuations* is of particular importance in transition times: if Q approaches zero, the average transition time approaches infinity at an exponential rate, and
2. The standard deviation of the first exit time estimates is typically as large as the norm. Hence, the uncertainty of the approximation is small relative to the uncertainty of the process itself.

3.2 The CMH Returns Distribution

Vaga (1990) made the following observation of the US market:

"A coherent bull market is characterised by an "inversion" in the historical risk-reward ratio. Over the period 1930 – 1990, the U.S. stock market has provided an average of 10% total annual return, with a standard deviation of 20%. During periods of coherent bull markets, the return from the stock market averages 25% with a standard deviation of only around 10%. Hence, if a portfolio manager does not trade on these low-risk opportunities, then his/her portfolio may underperform. This is due to that the rest of the time the market will present more risk than reward during periods of chaotic markets and periods of true random walk."

Motivated by the above observation, Vaga (1990) describe the stock market fluctuations in terms of "A Theory of Social Imitation" developed by Callen and Shapero (1974). Under Callen and Shapero's Theory of Social Imitation, stock price path is modelled by a non-linear statistical model, which has the features that it:

1. May be regarded as a statistical version of chaos theory, and

2. Includes the random walk as a special case.

Non-linear statistical models forecasts the conditions under which transitions in market state may be expected. by defining the ***probability distribution*** governing stock market fluctuations, rather than the specific future path of stock prices.

3.2.1 Stock Market State Transitions

Coherent behaviour in financial market is a state of macroscopic order in a complex physical or social system, which is made up of a large number of subsystems. Each sub-system is free to act independently. Periods of coherent behaviour in the stock market occur when the annualised return from a stock market index is greater than the annualised volatility of the index, reflecting a strong bias (or orderly trend) in price fluctuations).

The stock market is an "open" system that requires a continuous flow of capital to maintain a transition from a disordered to a more ordered state. Vaga (1990,1994) made the following assumptions:

Condition 38

1. *The industry groups comprising the stock market are analogous to the molecules in a ferromagnetic material*

2. *Let r be stock market return. Then,*

$$r \propto N_+ - N_-$$

where:

$$\begin{aligned} N_+ &\equiv \text{Number of industry groups trending higher } (+), \text{ and} \\ N_- &\equiv \text{Number of industry groups trending lower } (-). \end{aligned}$$

Remark 39 *Market returns may either fluctuate randomly around zero, or exhibit a high degree* *of polarisation leading to a large net difference between gainers and losers and corresponding large market moves.*

Then, the three input parameters for the stochastic model of public opinion are re-interpreted the following:

1. **Sentiment** (κ): is a measure of whether the level of "group think" is above or below a critical transition threshold.

2. **Fundamental Bias** (ρ) is a measure of external preference toward bullish or bearish sentiment.

3. **Degrees of Freedom** (n): is interpreted by Vaga (1990) as the number of industry groups on the stock exchange under study. (It is assumed to be 186 in sample calculations.)

Remark 40 *Vaga (1990,1994) asserts that under different market state, κ and ρ may vary widely, but n should remain relatively constant.*

Based on the above re-interpreted parameters, Vaga (1990) proposed the following model of annualised returns, which he called the Coherent Market Hypothesis:

Proposition 41 (Coherent Market Hypothesis) *Based on the* theory of social Imitation. *Vaga (1990) proposed the following model of annualised market return:*

$$\begin{aligned} f_R(r|n,\kappa,\rho) &\equiv \text{p.d.f. of annualised market return } R \text{, given the parameter values } n,\kappa,\rho \\ &= c(\kappa,\rho,n)\cdot\Psi^{-1}(\mathbf{r};\kappa,\rho,n)\cdot\exp\left[2\cdot\int_{u=-\frac{1}{2}}^{\mathbf{r}}\frac{\Pi(u;\kappa,\rho)}{\Psi(u;\kappa,\rho,n)}du\right], \end{aligned} \tag{3.33}$$

where the two control parameters of the p.d.f. is re-interpreted as follows:

$$\begin{aligned} \kappa &= \text{degree of market sentiment,} \\ \rho &= \text{degree of fundamental bias,} \end{aligned}$$

and the number of sub-systems making up the market (or degrees of freedom) is re-interpreted as:

$$n = \text{Number of industrial groups making up the financial market.}$$

Now, make the following assumption about the average rate of change of buying and selling:

$$\nu = 1.$$

Then, the Drift coefficient function and Drift coefficient function is respectively given by:

$$\begin{aligned} \Pi(r;\kappa,\rho) &= \sinh(\kappa r+\rho) - 2r\cosh(\kappa r+\rho)\,, \textit{ and} \\ \Psi(r;\kappa,\rho,n) &= \frac{1}{n}\left[\cosh(\kappa r+\rho) - 2r\sinh(\kappa r+\rho)\right]. \end{aligned}$$

Finally, the normalising constant for the p.d.f. is given by:

$$c^{-1}(\kappa,\rho,n) = \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \left(\Psi^{-1}(r;\kappa,\rho,n)\cdot\exp\left[2\int_{u=-\frac{1}{2}}^{r} \frac{\Pi(u;\kappa,\rho)}{\Psi(u;\kappa,\rho,n)}du\right]\right) dr.$$

3.2.2 States of Market Dynamics

The two most important factors that controls the state of the market are:

1. Investor sentiment, and
2. Prevailing bias in economic fundamentals.

The most prevalent market states

When investor sentiment is not conducive to "group-think" or crowd behaviour, the market is likely to be in a **random walk** (efficient market) state. In such relatively quiet periods, the stock market is least likely to outperform fixed currency and fixed income alternatives.

The combination of strong positive fundamentals and investor sentiment conducive to crowd behaviour leads to the safest and most rewarding state, **a coherent bull** market. In these market periods, a fully invested position is necessary to avoid the risk of underperforming the market averages.

When the fundamental bias is neither strongly bullish nor bearish during a period of crowd behaviour, the result is a **chaotic** market. This is a dangerous and "quasi-efficient" market state to invest in as random news is discounted quickly (but with a bias) and sentiment may switch abruptly from bullish to bearish.

Finally, the coherent bear market state is a theoretically possible yet historically rare market state, where there is highly negative fundamentals with crowd behaviour. The bearish fundamentals usually dampen investor enthusiasm and prices tend to drift lower, in a random walk, over an extended bear market.

Characteristics associated with each market state

When crowd behaviour prevails, market fluctuations will tend to follow a bimodal distribution. Under these conditions, market action can best be described as a biased random walk, or "quasi-efficient." On a short-term basis, new developments will be discounted quickly, but investors will tend to respond to good news while ignoring bad news (or vice versa). Under these conditions, sentiment may also, by chance, switch abruptly from one state to its mirror image, producing anomalous short-term volatility. During periods of crowd behaviour, the market is highly sensitive to any underlying bias in economic fundamentals.

Peters (1989) presented evidence of biased random walks, which led him to conclude that "Pure random walk theory does not apply to capital markets. The capital markets instead follows a biased random walk". However, there are periods when random walk theory received good empirical support. Thus, instead of rejecting the random walk theory outright, Callen and Shapero (1974)'s theory of social imitation suggests that, at times, the market is in a true random walk state. If these periods of true random walk are eliminated, then the remaining periods of crowd behaviour may be expected to exhibit even greater persistence (i.e. trendings in returns) than that indicated by Peter's data. This conclusion would have important implications for market timing and asset allocation. Techniques such as tactical asset allocation may not be futile in periods of coherent markets. However, when true random walk markets prevail, trend following strategies and relative-strength approaches will not be effective.

Table 3.1 below presents theoretical expected returns and standard deviations for a range of combinations of sentiment and fundamentals as suggested by Vaga (1990):

Market State	Sentiment (k)	Fundamental (ρ)	Expected ann. return r (%)	Volatility of ann. return (σ)
Random Walk	1.8	+0.02	+8%	10%
		0	0	10
		−0.02	−8	10
Transition	2.0	+0.02	14	12
		0	0	16
		−0.02	−14	12
Coherent Bull	2.2	+0.03	+27	8
		+0.02	+25	11
Chaotic	2.2	+0.01	+16	18
		+0.005	+10	21
		0	0	23
		−0.005	−10	21
		−0.01	−16	18
Coherent Bear	2.2	−0.02	−25	11
		−0.03	−27	8

Table 3.1: Vaga (1990) Recommended CMH Parameter Values

From Table 3.1 above, Vaga (1990) made the following interpretations:

1. $k = 2$ is identified as the critical state transition threshold. Below the critical state transition threshold, the random walk market state will prevail. Above the transition threshold, coherent bull markets will occur if fundamentals are strongly positive, while chaotic markets are likely if fundamentals don't provide a clear direction for investors.

2. Expected return increases non-linearly with respect to changes in the fundamental bias, ρ.

3. In the random-walk market state, the impact of a change in fundamentals is considerably less than it is during period of crowd behaviour.

4. In the coherent market state, the magnitude of the expected return is more than twice its standard deviation. This suggests a potential quantitative definition of coherent behaviour in capital markets.

Market State 1: Random Walk

The CMH returns distribution, as defined by (3.33), can be considerably simplified for the special case when sentiment is not conducive to crowd behaviour, i.e $k < 2$. If we further assume that fundamentals are neutral ($h = 0$), the probability distribution of returns, $f(q)$ can be expressed as follows:

$$f(q) = \left(\frac{1}{g\sqrt{\pi}}\right) \cdot \exp\left(-\frac{q^2}{g^2}\right), \tag{3.34}$$

where the variance is:

$$g^2 = \frac{1}{(2-k)\,n}. \tag{3.35}$$

This is the "normal" distribution corresponding to a snapshot in time of a random-walk process.

When non-linear effects are significant, transition probabilities of steps up or down are no longer equal. In effect, steps in one direction may be larger and more likely than steps in the other direction. This amounts to a biased random walk, where the bias may behave either in coherent or chaotic fashion depending on prevailing sentiment or fundamentals. While the non-linear model reduces to the usual random walk as a special (linear) case, the general non-linear, time-dependent situation is quite complex.

Figure 3-1 below shows the CMH distributions in random walk market state with bearish, neutral and bullish fundamental bias:

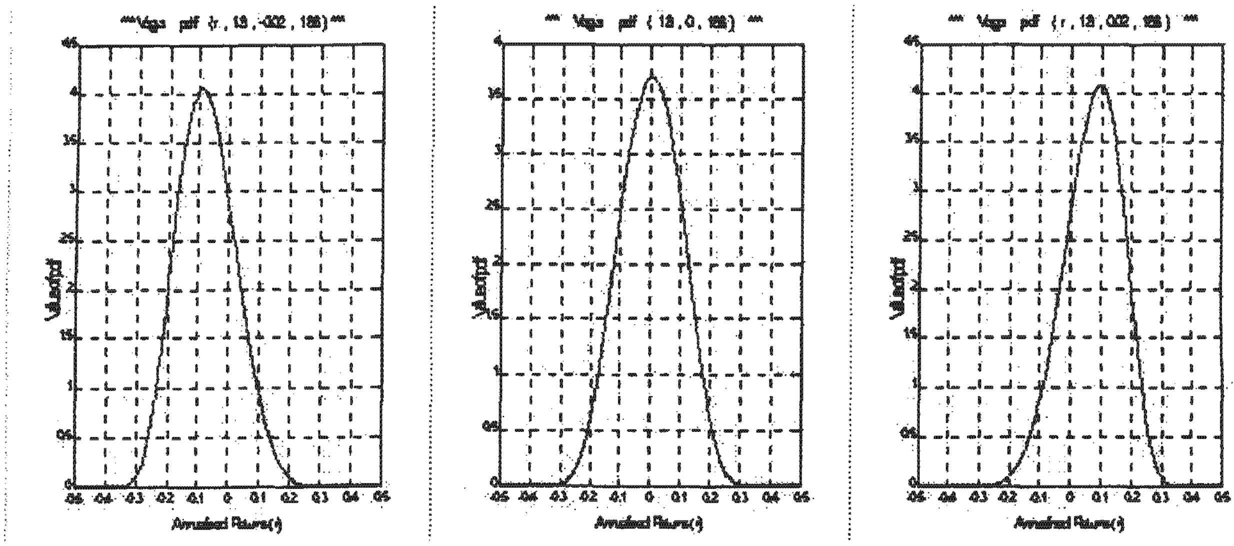

Figure 3-1: CMH Distribution in Random Walk State

Theoretically, random-walk markets could provide either a small, stable bullish return or a small, stable bearish return, depending on fundamentals. Historically, however, random-walk markets have provided a stable negative return and are most frequently associated with bear markets.

Market State 2: Transition

The variance in the normal probability distribution in equation (3.34) becomes very large as k approaches two, the critical transition threshold. In this situation, the normal distribution no longer applies. The random-walk model is no longer valid during the transition to crowd behaviour. At transition, an instability occurs, which implies a highly inefficient market in which large, long-lasting sentiment swings must be expected.

Figure 3-2 below shows the CMH distributions in transition market state:

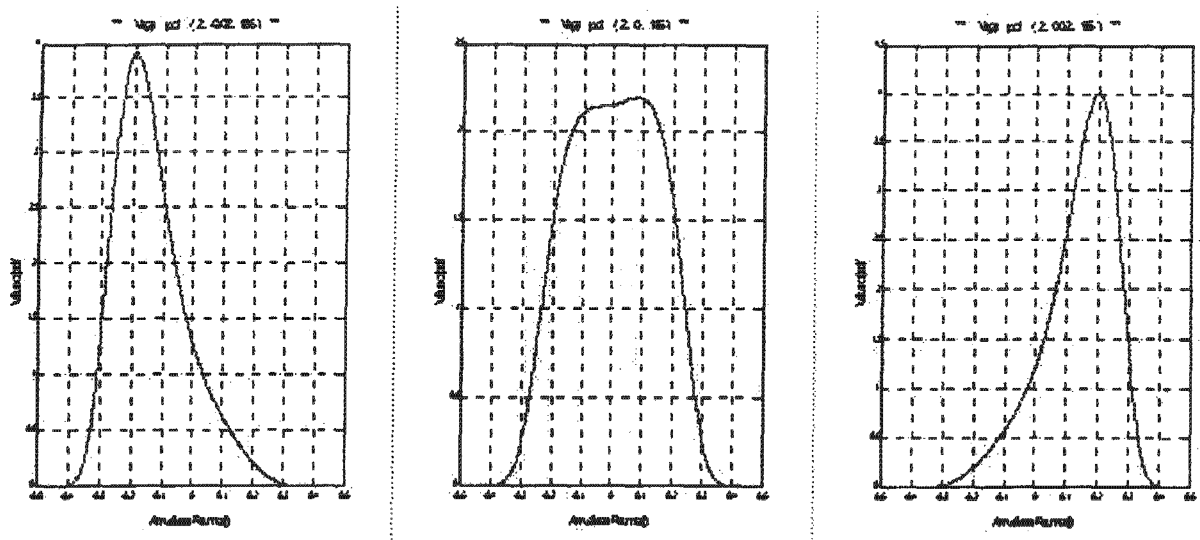

Figure 3-2: CMH Distribution in Transition State

The potential well for the transition to crowd behaviour nearly flat over a wide range of expected returns. Nearly anything can happen in a period of instability. The case shown is for neutral fundamentals; even a slight fundamental bias would tend to skew the distribution strongly in the direction of the bias. **Uniform Distribution**: A specific type of probability distribution, which best represents stock market behaviour during the unstable transition from a true random walk market to a coherent (or chaotic) market. Unlike the bell-shaped normal distribution governing the price fluctuations in a true random walk market, the uniform probability distribution is much wider and flatter, which implies that large, long-lasting price fluctuations are more likely during the period of instability.

Market State 3: Chaotic

As κ increases above the critical threshold of 2, a high degree of polarisation exists among investors. But, without a strong fundamental bias, there is no clear indication as to whether the crowd will stabilise in a bullish or bearish state. Furthermore, there is a possibility of abrupt sentiment shifts (chaotic fluctuations) from bullish to bearish, or vice versa.

In a chaotic market, a "bimodal" probability distribution may occur, where extremes are more likely than the centre of the distribution function. The probability of a large sentiment shift is greatest when prevailing investor sentiment runs counter to a small external bias in fundamentals. Figure 3-3 below shows the CMH distributions in chaotic market state with bearish fundamental bias:

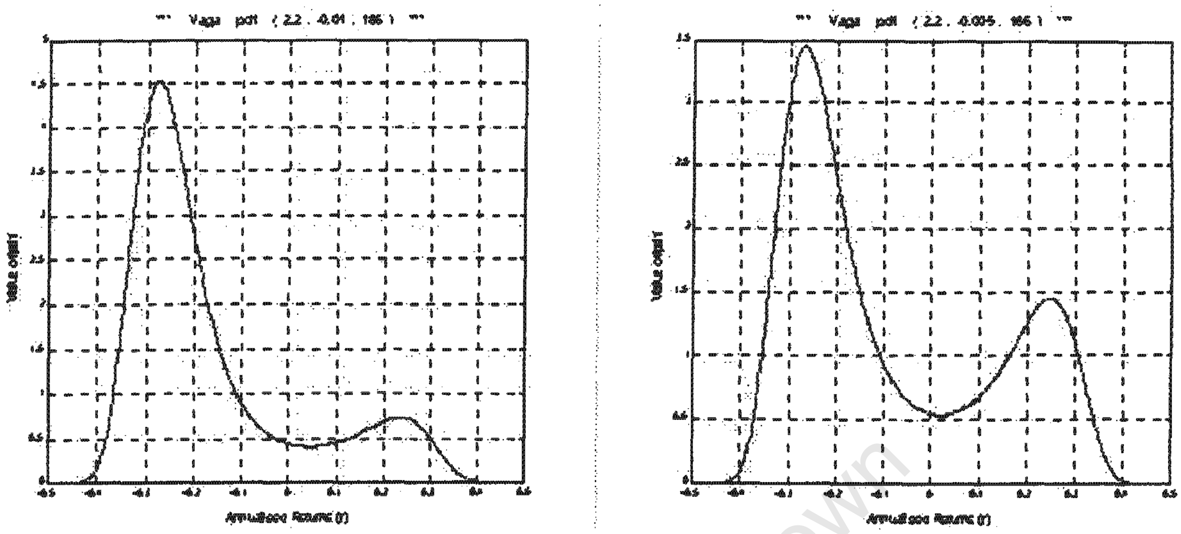

Figure 3-3: CMH Distribution in Bear Chaotic State

The climate prior to the crash was conducive to crowd behaviour; this was indicated by the extremes in market volume and breadth. Fundamentals were somewhat bearish, as a result of rising interest rates throughout 1987 as well as unusually high valuations by historical norms.

Although a bullish state is quite possible when crowd behaviour is combined with bearish fundamentals and may be interpreted as a period of market "mania," but it is a unstable and potentially dangerous situation. The specific news prior to a market crash was less important as a causal factor than the prevailing combination of sentiment and fundamentals.

Figure 3-4 below shows the CMH distributions in true chaotic market state:

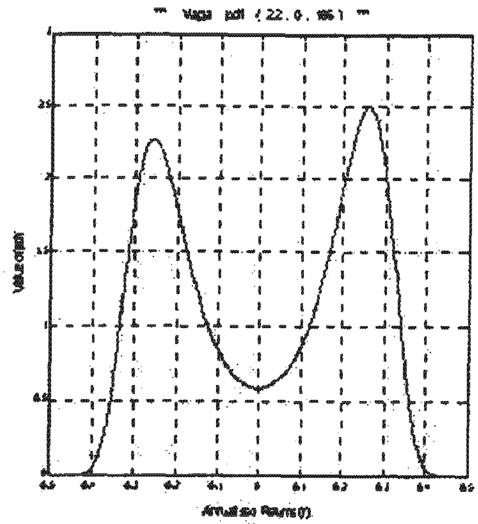

Figure 3-4: CMH Distribution in True Chaotic State

Figure 3-5 below shows the CMH distributions in chaotic market state with bullish fundamental bias:

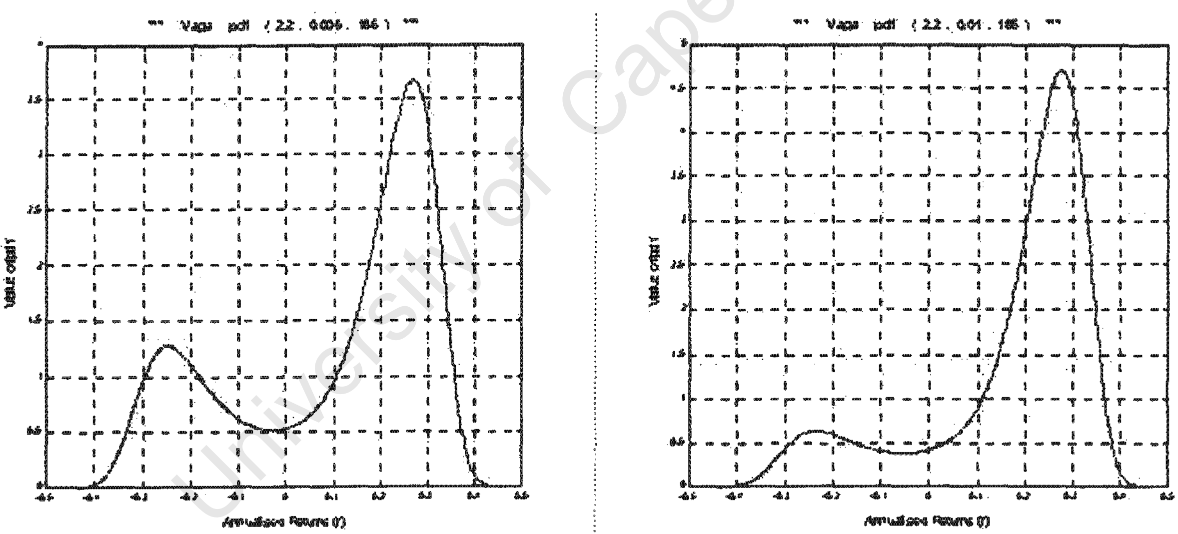

Figure 3-5: CMH Distribution in Bull Chaotic State

A chaotic market may be described as quasi-efficient. As long as sentiment remains stable, new developments will be discounted quickly, but with a bias reflecting prevailing sentiment.

Market State 4: Coherent Bull

Takes place when fundamentals are strongly positive during a period of crowd behaviour.

A coherent bull market can be thought of as a chaotic market in which the bearish side of the potential well is high, and the associated lobe of the probability distribution becomes small. Figure 3-6 below shows the CMH distributions in coherent bull market state:

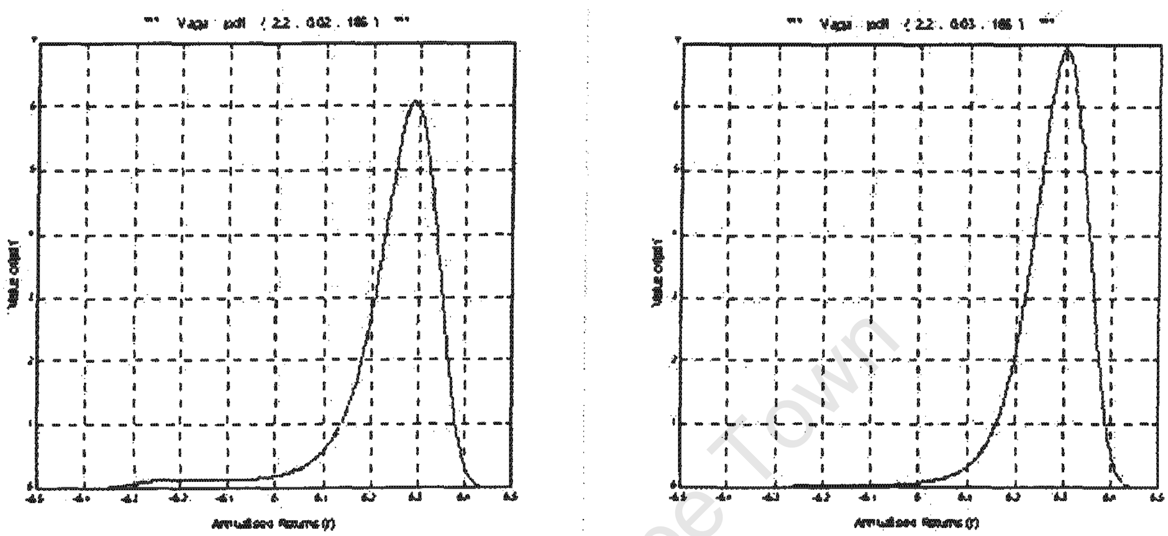

Figure 3-6: CMH Distribution in Bullish State

The market's expected return and its standard deviation show the inversion of the historical risk-reward ratio. The peak of the distribution is even higher than the expected return. However, the distribution has a long tail that goes deep into negative territory. The non-linear model predicts that, even under extremely bullish conditions, there is a small but non-zero probability that the market will provide a negative return.

Coherent bull markets typically occur when capital are unusually high. A continuous supply of available funds is needed to maintain the positive uptrend in stock prices. Maintaining excessive cash reserves during a coherent bull market simply locks in underperformance. Most of the market's long-term gains can be attributed to coherent markets. After a coherent market ends, and a chaotic or random-walk state begins, it is too late to invest with the hope of making up the forgone returns.

Market State 5: Coherent Bear

Coherent bear markets are the mirror image of coherent bull markets. Crowd behaviour with strong bearish fundamentals. The standard deviation is the same as for the coherent bull market, but expected return is negative. Figure 3-7 below shows the CMH distributions in coherent bear market state:

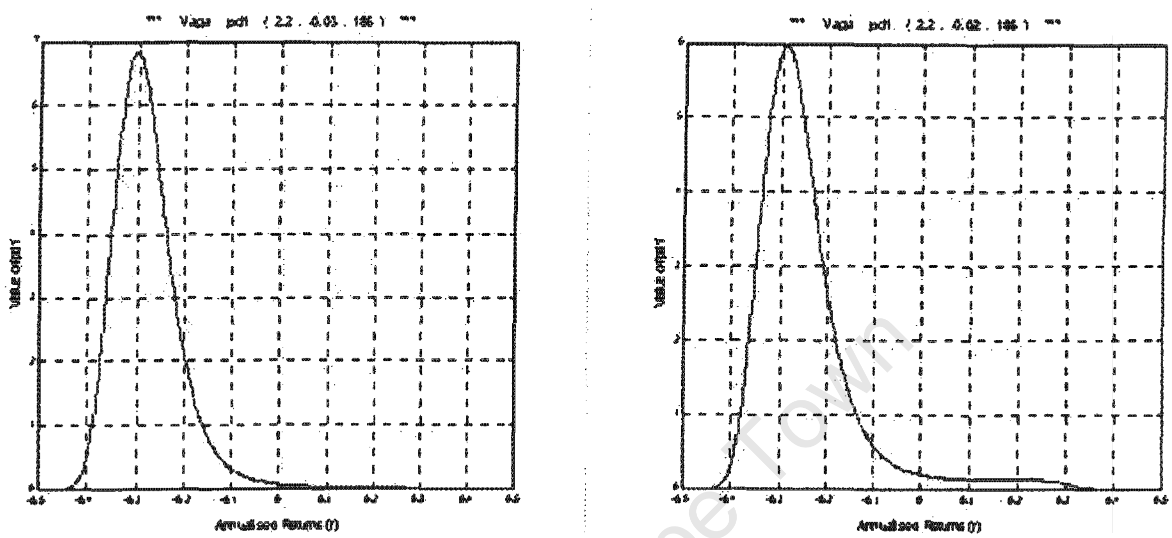

Figure 3-7: CMH Distribution in Bearish State

3.2.3 Market State Indicators

The remainder of this section is devoted to describing the theoretical basis and practical indicators of coherent markets. the results suggests that both technical and fundamental analysis can add real value to the investment decision-making process.

In the Ising model, the probability distribution governing the stock market's risk-reward outlook is completely determined by the following three parameters.

1. Sentiment (k) reflects the degree of coupling of opinion among investors in different market sectors (industry groups); it ranges from independent ($k = 1.8$) through unstable transition ($k = 2.0$) to "group think" or crowd behaviour ($k = 2.2$).

2. Fundamentals (h) reflect prevailing Federal Reserve policy, ranging from typically from tightening ($h = -0.02$) to neutral ($h = 0$) to stimulating growth ($h = 0.02$).

3. Degrees of freedom (n) is assumed to be the number of industry groups making up the market ($n = 186$ for the illustrations).

The first two parameters determine the market state, given a fixed n. They may be rated subjectively or based on quantitative indicators.

Sentiment

Risk-reward outlook may be assessed by determining whether the climate is conducive to crowd behaviour. For financial markets, bullish crowd behaviour can provide the best investment opportunities, provided economic fundamentals are not bearish. Investors who does not take advantage of periods of bullish crowd behaviour (coherent bull markets) will underperform in the long run. Crowd behaviour may be identified by extremes in various sentiment indicators. Long-term gains occur during coherent bull markets. The rest of the time the market offers risk without reward.

Fundamentals

During crowd behaviour, economic fundamentals are critical. The underlying bias can be reduced to whether monetary conditions are supporting or curtailing economic growth. The two fundamental indicators suggested by Vaga (1990) are as follows:

1. Interest rate policy. A reduction in interest rate signals a strong bullish bias in fundamentals. Changing fundamentals during crowd behaviour can shift the risk-reward outlook from the safest to the most dangerous time to be in the stock market. Highly bullish fundamentals coupled with strong market action (crowd behaviour) accompany coherent bull markets.

2. Market's price-to-earnings (P/E) ratio. When stocks were cheap ($P/E < 10$), the annualised return is high and the standard deviation low.

The CMH Model and Chaos

Table 3.2 below summarises the classes of mathematical models that can be used to quantify the fluctuations of an order parameter:

Quantitative Models	Linear	Non-Linear
Statistical	Random Walk (damped harmonic oscillator + random forces)	Statistical Chaos (A Theory of Social Imitation)
Deterministic	Simple Pendulum (damped harmonic oscillator)	Deterministic Chaos (damped anharmonic oscillator)

Table 3.2: Mathematical Models of Order Parameter

Table 3.2 above proposes the following two types of non-linear model for the macroscopic behaviour of the system:

1. **Deterministic Chaos.** The simplest model of deterministic chaos is the non-linear pendulum (heavily damped anharmonic oscillator) whose motion is governed by the *simple cusp potential* function:

$$V(r) = \frac{1}{2}\left(ar^2\right) + \frac{1}{4}\left(br^4\right), \tag{3.36}$$

where:

r is the order parameter, and

a, b are the control parameters.

The simple cusp potential function will bifurcate from a single stable point (linear system) to a bistable configuration (non-linear system) according to the values of the control parameters. The bistable configuration exhibits chaotic behaviour (i.e. any slight changes in the initial conditions will lead to wide variations in the future state of the system).

2. **Statistical Chaos.** If random forces are significant, then a statistical analysis is required. In this case, the probability distribution function governing the fluctuations of the order parameter f is:

$$\begin{aligned} f(r,t) &= \exp\left[-\left(\frac{2}{C}\right)\cdot\left(\frac{1}{2}ar^2 + \frac{1}{4}br^4\right)\right] \\ &= \exp\left[-2\left(\frac{V(r)}{C}\right)\right], \end{aligned} \tag{3.37}$$

where:

$V(r)$ is the potential well function,

a, b are the control parameters, and

C is the correlation coefficient of random forces.

Then, the control parameters will determine the shape of the probability distribution as follows:

i) If non-linearities are insignificant ($b = 0$), then (3.37) reduces to the normal distribution, and

ii) If non-linearities are significant ($b > 0, a < 0$), then the normal distribution bifurcates and (3.37) splits into a symmetrical bimodal distribution.

Note in equation (3.37) above that both V and C becomes part of the solution for the probability distribution, where C is analogous to the temperature in physical systems. Thus, equation (3.37) may be regarded as a statistical version of chaos theory.

The macroscopic behaviour of the system, as described by its order parameter, undergoes transition from one characteristic state to another depending on the two competing external forces on its sub-systems as follows:

1. In the absence of random thermal forces and minimal presence of coupling/feedback among subsystems, non-linear effects are negligible. Thus, the simple pendulum (a damped harmonic oscillator) would adequately characterise the linear behaviour of the system.

2. As the forces of coupling/feedback increases above the random forces, non-linear effects of the system becomes more dominant. Thus, the dynamics of the system can exhibit chaotic behaviour.

3. As the random forces increases above the forces of coupling/feedback, a statistical description of the system becomes necessary.

The Theory of Social Imitation is a non-linear statistical model based on the diffusion approximation of the Ising model, which includes both a statistical version of catastrophe theory and the random walk as special cases. Thus, it avoids the limitations of deterministic chaos theory models and has greater flexibility to describe stock market transitions from periods of true random walk to coherent price trends, and the chaotic fluctuations associated with panics and crashes. This theory includes random walk as a special case and may be thought of as a statistical form of chaos.

3.2.4 Theoretical Explorations of the CMH Returns Distribution

An investigation of the relationship between the first four moments of the CMH model with changes in the control parameters (κ, ρ) are best carried out by contour plots of the first four moments. Contour plots of the CMH distribution are computed using Matlab for $\kappa \in [-190, 10]$, $\rho \in [-10, 10]$, and $n = 20, 80, 150$. The interval of κ and ρ values were chosen for the following reasons:

1. They are the most reasonable range of values based on the obtained maximum likelihood estimates, and
2. It is the widest interval of values possible for which the computing time is still reasonable. The larger the number of (κ, ρ) values for which Matlab requires to calculate the first moments, the longer it will take to produce the result.

Figure 3-8 below shows the contour plots for expected return, standard deviation of return, coefficient of skewness, and excess kurtosis of the CMH distribution, for each $n = 20, 80, 150$:

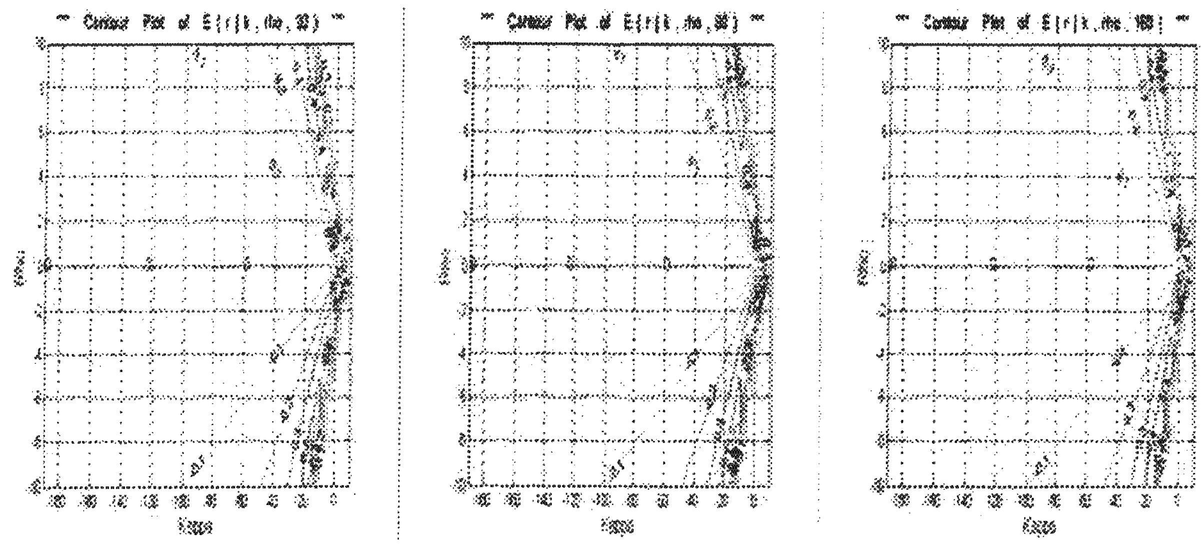

Figure 3-8: Relationship Between Expected Return and CMH control parameters (κ, ρ)

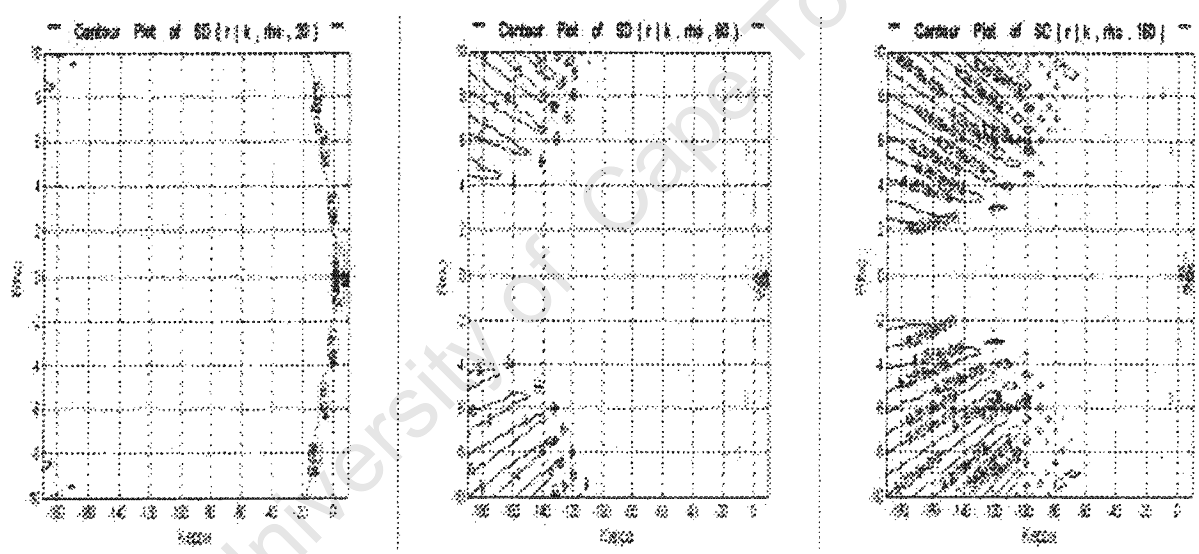

Figure 3-9: Relationship Between Std Dev of Return and CMH control parameters (κ, ρ)

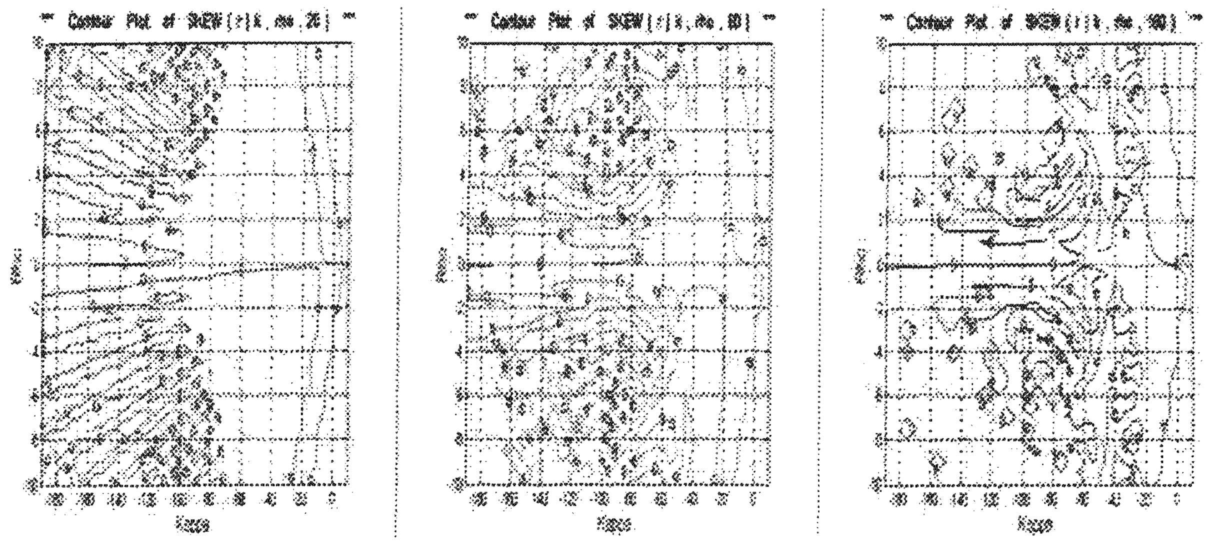

Figure 3-10: Relationship Between Skewness Coefficient and CMH control parameters (κ, ρ)

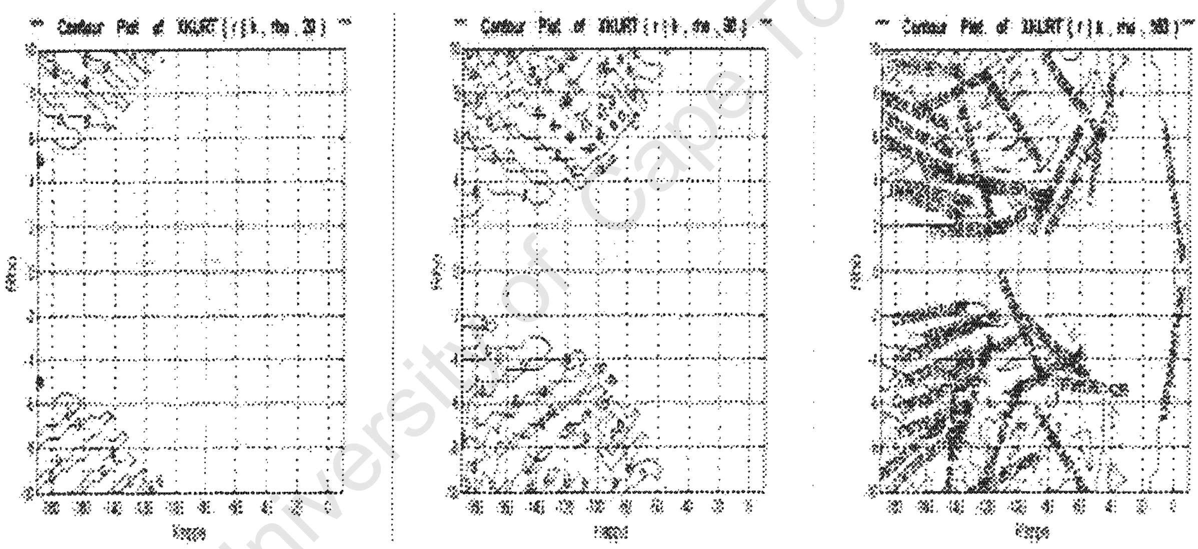

Figure 3-11: Relationship Between Excess Kurtosis and CMH control parameters (κ, ρ)

From Figures 3-8 to 3-11 above, the following interpretations may be made $\kappa \in [-190, 10]$, $\rho \in [-10, 10]$, and $n = 20, 80, 150$:

1. When $\rho = 0$, the mean of the CMH distribution is a constant with respect to κ. The relationship of the mean with (κ, ρ) is not very sensitive to changes in value of n.

2. The standard deviation of the CMH distribution is non-linearly related to (κ, ρ). There seems to be a discontinuity at roughly $\rho = 0, \kappa = 5$. This may represent the occurrence of "critical phenomena" in the corresponding market state. The relationship of standard deviation with (κ, ρ) is very sensitive to changes in value of n.

3. The coefficient of skewness of the CMH distribution is non-linearly related to (κ, ρ). There seems to be a discontinuity at roughly $\rho = -4, \kappa = -165$ for $n = 80$ and 150. This may represent the occurrence of "critical phenomena" in the corresponding market state. The relationship of coefficient of skewness with (κ, ρ) is quite sensitive to changes in value of n.

4. The excess kurtosis of the CMH distribution is non-linearly related to (κ, ρ). There seems to be a discontinuity at roughly $\rho = -4, \kappa = -165$ for $n = 80$ and 150 again. This may represent the occurrence of "critical phenomena" in the corresponding market state. Additionally, for $n = 150$, the values of the excess kurtosis seemed to have "blown up" or become very large for all (κ, ρ). This may be due to errors in the code or discontinuous behaviour of the CMH distribution. The relationship of excess kurtosis with (κ, ρ) is very sensitive to changes in value of n.

5. The degree of non-linearity with (κ, ρ) is more marked for standard deviation and excess kurtosis.

Chapter 4

EMPIRICAL EXPLORATIONS OF CMH ON THE JSE SECURITIES EXCHANGE

The aim of this chapter is to carry out the first and the sixth objective for this dissertation, namely:

Objective 1: Using the ideas of technical analysis to split the time series of log returns on JSE Overall Index (from 27 March 1985 to 21 June 2002) into five distinct states (namely: Random Walk, Coherent Bull, Coherent Bear, Transition, Chaotic).

Objective 6: Estimating the five market states by the stable distribution family.

4.1 Background on the JSE Securities Exchange South Africa

This section contains a brief summary of the background on the Johannesburg Stock Exchange Securities Exchange South Africa (JSE). The main source of information for this summary is the official JSE world wide web site (http://www.jse.co.za).

In 1886, the Witwatersrand gold fields, the richest of its kind in the world at that time, was discovered in Johannesburg. Little more than a year later, the Johannesburg Stock Exchange Securities Exchange South Africa (JSE) was founded in 1887 to enable new mines and its financiers to raise funds for the development of the booming mining industry. Today, majority of the listed companies on the JSE are non-mining organisations.

4.1.1 Features of the JSE

The JSE is the only stock exchange for equities and other securities in South Africa. Although, the JSE has been classified as an "emerging market" when compared to the global stock market. But, the South African economy is characterised by both established first world fundamentals as well as third world features. From Monday 10 June 1996, all trades on the JSE is being conducted through the automated trading system, *JET* (JSE Equities Trading).

The JSE is a self-regulatory organisation that is governed by a set of rules drawn up by the elected JSE Committee. The JSE is a member of the Fédération Internationale des Bourses de Valeurs (FIBV) and was granted designation status by the Japanese Securities Dealers Association effective on 14 December 1994. On 1 March 1995 the JSE was included in the Morgan Stanley index for emerging markets and on 7 April 1995 was included in the IFC Emerging Markets Global and Investable Indices.

4.1.2 Roles of the JSE in South Africa

The main functions of the JSE are:

1. To provide an orderly primary and secondary markets for trades in equities and other securities in order to create new investment opportunities in South Africa.

2. To create liquidity, thus ensuring the primary market fulfils its function of raising new investment capital.

3. To re-channel cash resources into productive economic activities, thus allowing for the raising of primary capital required to develop the country's economy.

4. To make the services it provides accessible to the entire nation.

5. To ensure that the nation is suitably informed in the advantages and risks of share ownership.

4.1.3 Re-structuring of the JSE

In anticipation of the fundamental changes to the South African political and economic environment that is to follow the first democratic election in 1994, the JSE applied to the South African Parliament for a re-structuring. In September 1995, the Stock Exchanges Control Amendment Act was approved by the Parliament, which allowed for the re-structuring of JSE. The restructuring plan, as approved by the JSE Committee, aimed to ensure:

1. The stock exchange de-regulated efficiently and successfully,
2. The stock exchange contributes towards the needs of the new political and economic regime, and
3. The JSE's attractiveness to local and foreign investors will be further enhanced.

The JSE is re-structured without bias towards any particular business or social community. The restructuring has impacted on:

1. Membership in the JSE,
2. Trading principles and systems,
3. Clearing and settlement,
4. Transfer and registration,
5. Capital requirements of member firms, and
6. Financial structure of the JSE.

The overall benefits to the economy from this re-structuring were suggested as follows:

1. A listing on the JSE enables substantial amounts of capital to be raised for the financing of new businesses, expansion of existing businesses, and the creation of new employment opportunities.
2. To allow for speculative buying and selling of shares for individuals that are knowledgeable about the performance of selected shares.

4.1.4 Trading Hours on the JSE

The trading days on the JSE are from Monday to Friday, but excluding the public holidays as declared by the South African Parliament. The continuous trading hours on a trading day are from 09H00 to 17H00 (South African time). The JSE SETS, the new automated Equities Trading Systems introduced by JSE, operates every trading day from 08H25 until 18H00, with pre-opening sessions applicable from 08H25 to 09H00, and the runoff session applicable from 17H00 to 18H00.

4.2 The JSE Overall Index

This section contains a brief summary of the financial index measuring the overall performance of the JSE over a trading day. The main source of information for this summary is the official FTSE/JSE Africa Index Series world wide web site (http://ftse.jse.co.za).

In financial markets, financial indices are published by the regulators to summarise the price movements of a group of listed shares on the market. The primary functions of these financial indices are:

1. To describe the market at a point in time in terms of price levels, dividend yield and earnings yields.

2. To enable investors and traders to compare the performance of a particular listed share with the performance of a market index.

3. It is an important economic indicators as movements in JSE-OVER index values may be interpreted to provide an indication of the investors' confidence in and market's expectations for the different sectors within the market as well as the overall market.

On the JSE, the Johannesburg Stock Exchange Overall (**JSE-OVER**) Index measures the performance of the overall equity market. According to JSE's "old" (Prior to 24 June 2002) sector classification system, the JSE is comprised of 5 sectors and 53 sub-sectors. The JSE-OVER index is a *weighted arithmetic* index, where the weights are the market capitalisation (or market cap) of each constituent security.

4.2.1 Market capitalisation of JSE-OVER index's constituent securities

The determination of capitalisation for each constituent company is the most complex task in the index calculation process. The calculation of JSE-OVER index is based on the *full market capitalisation* of constituent securities, which is defined as:

$$\text{Full Market Cap} \equiv (\text{Entire listed share issue of a security}) \times (\text{Current price of security}).$$

Remark 42 *From the above definition of full market capitalisation, note the following:*

1. *Listed share issue includes shares which are issued partly, or nil paid where the call dates are already determined and known.*

2. *Convertible preference shares and loan stocks are excluded from the above definition until such time that they are converted.*

3. *In order to prevent a large number of insignificant weighting changes, the entire share issue is subject to the one percent rule. The one percent rule states that the number of shares in issue for each company is amended only when the total shares in issue changes by more than 1% on a cumulative basis.*

From the above definition of market cap, it follows that the price movement of constituent companies with a larger market cap will have a larger impact on the index than constituent companies with a smaller market cap.

4.2.2 The daily JSE-OVER index

The daily JSE-OVER index measures the relative performance of the overall equity market over a particular trading day. Let $t = 1, 2, \cdots$ be the number of trading days since the starting date of JSE-OVER. Then, the daily JSE-OVER index value is calculated according to the following formula:

$$\begin{aligned} JSE - OVER(t) &= \frac{\text{Total market value of all constituents at end of day } t}{\text{Index Divisor for day } t}, \\ &= \sum_{i=1}^{n} \frac{[p_i(t) \cdot e_i(t)] \cdot s_i(t)}{d(t)}, \end{aligned}$$

where:

$$
\begin{aligned}
n &= \text{Number of constituent securities comprising the Index at any time.} \\
p_i(t) &= \text{Trade price of } i^{th} \text{ constituent security at end of day } t \\
e_i(t) &= \text{Exchange rate, required to convert the } i^{th} \text{ constituent security's} \\
&\quad \text{home currency into index's base currency, at end of day } t. \\
s_i(t) &= \text{Number of shares in issue for the } i^{th} \text{ security.} \\
d(t) &= \text{Index divisor for day } t \\
&= \sum_{i=1}^{n} \frac{[p_i(t-1) \cdot e_i(t-1)] \cdot s_i(t-1)}{JSE-OVER(t-1)}, \text{ and} \\
JSE-OVER(0) &= \text{an arbitrary round number (e.g.10,100,1000) chosen at} \\
&\quad \text{the starting time of the index to fix the index's starting value.}
\end{aligned}
$$

The value of the daily JSE_OVER index is a single number representing the total market value of all constituent securities comprising the index at the end of trading day t relative to the total market value of all constituent securities comprising the index at the end of trading day $t-1$. Changes to the classification of a company or the composition of the sectors are accommodated in the indices, e.g. if the sectoral classification of a constituent changes, the security is deleted from the old sector and is tested for eligibility in the new sector.

4.2.3 The launch of FTSE/JSE Africa Index Series

On 24 June 2002, the new FTSE/JSE Africa Index Series was formally implemented. The most important changes to the "old" daily JSE-OVER index calculation are:

1. The move from indices based on full market capitalisation to free float adjusted indices.
2. Classifies securities into sectors according to the FTSE Global Classification System (comprised of 10 economic groups, 39 sectors, and 116 sub-sectors).

Free float is the amount of shares freely available to investors, which excludes shares where shareholding is restricted to specific individuals or groups of individuals. Free Float Market

Capitalisation is the adjusted market capitalisation used to determine the weighting of securities in different indices in the new FTSE/JSE Africa Index Series. It is calculated by applying the free float banding percentage to the full market capitalisation. The use of free float weightings in index calculation is aimed to provide a more representative view of the overall market performance. This should allow investors to track an index more closely and invest accordingly.

As a result, the FTSE/JSE ALL SHARE index have replaced the old JSE-OVER index. The JSE owns and disseminate the index information, but FTSE Group is responsible for the calculation and global marketing of the indices. Further information about the FTSE Group is obtainable from its official world wide web site (http://www.ftse.com).

4.3 Description of Financial Dataset Under Study

The aims of this section are to describe the dataset that will be used to achieve the objectives of this dissertation. The data includes the JSE overall index. The preliminary results of the data will be presented. The purpose of the computations is to familiarise with the structure of the data.

4.3.1 Source of Data

The financial indicators that summarises the price movements on the JSE are the JSE/Actuaries Indices and the Derivative Indices, which are:

1. Administered and maintained by the JSE's indices department, who is responsible for the daily loading and verifying of capital structure events.

2. Sold by the JSE's new business-data services department to subscribing data vendors.

The new business-data services department, via the Equity Data Dissemination system, provides the following information:

1. Equity price related information on daily trades and share movements, and

2. Capital structure changes, Sector, Indices statistics and Warrants information.

The data described above is processed in batch mode and transmitted to all subscribing users at the end of each trading day. The subscribing users includes information vendors, Members, and Financial Institutions. The Equity Data Dissemination vendors subscribes to the JSE to obtain the JSE equity data.

The daily JSE-OVER index dataset (as described above) used in this dissertation to empirically explore the CMH distribution for the JSE, is an extract accessed from the JSE indices database maintained by Professor C.G. Troskie from the Department of Statistical Sciences at the University of Cape Town, South Africa. The database maintained by Professor Troskie downloads the JSE data from InvestorData (Ntobi) cc who is one of the South African Equity Data Dissemination vendor identified by the JSE. InvestorData (Ntobi) cc checks and corrects the error in the dataset they receive from the JSE before releasing to their clients. The dataset is stored in the ASCII file format under the following column headings:

1. STOCK NAME (JSE-OVER),
2. DATE (yyyymmdd),
3. LAST (Last trade price)
4. HIGH (Highest price of the day)
5. LOW (Lowest price of the day)
6. VOL (Trading volume for the day)
7. BIDPRICE (The last floor price of the day)

4.3.2 The financial dataset under study

Recall that the CMH provides a distribution for return on the overall financial market movement. Thus, the financial indicator chosen for the exploration of CMH model on the JSE is the Johannesburg Stock Exchange Overall Index (**JSE-OVER**).

The JSE-OVER index dataset to be used for the exploration of CMH model in this dissertation is a sample of daily JSE-OVER Index values (prices) recorded sequentially at the end of each day that the JSE opens for trade starting from 27 March 1985 (inclusive) to 21 June 2002 (inclusive). Additional information available on the sample are: highest price of the trading day (HIGH), lowest price of the trading day (LOW), trading volume for the day (VOL), and the last floor price of the trading day (BIDPRICE). A summary of the important characteristics of the daily JSE-OVER index sample are given as follows:

1. The sample consists of $4,303$ raw daily JSE-OVER index values covering a time period of $6,296$ calendar days (or approximately 18 calendar years).

2. Daily JSE-OVER index values prior to 27 March 1985 is not available from the database maintained by Professor Troskie.

3. The last index value in the sample is recorded on 21 June 2002, the last date that the JSE-OVER index is calculated by the JSE. Thus, there is no additional problem with the adjustments required to be made to the new FTSE/JSE All Shares Index in order to make the index values comparable.

4. Daily JSE-OVER index values is not available on any day that the JSE does not trade. The non-trading days include weekends, public holidays as declared by the South African Parliament, and days when the JSE do not trade for special reasons. Thus, the sample of index values are irregularly spaced in calendar time.

Since the JSE-OVER index is a financial variable and the daily JSE-OVER index values are recorded sequentially in calendar time, thus the sample is an observed (univariate) financial time series. In order to avoid unnecessary complications in the analysis, the calendar time at which the daily JSE-OVER index values in the sample is observed will be re-labelled to obtain a new financial time series defined as follows:

Definition 43 (Observed Daily JSE-OVER Index Time Series) *The sample of daily JSE-OVER index values used for this dissertation is an observed (univariate) time series of daily JSE-OVER index defined by:*

$$(p_t)_{t=1}^{4303} = (p_1, p_2, \cdots, p_{4303}),$$

where:

t ≡ *Number of observed index values in the sample since trade began on 27 March 1985 (Date at which first index value in sample is recorded).*

p_t ≡ *Value of daily JSE-OVER index recorded sequentially at the end of the calendar date associated with t.*

Remark 44 *Based on the above definition, the following comments may be made:*

1. *The round bracket around P_t is used to stress that the P_t's are ordered sequentially in t. Thus,*

2. *The definition of t is carefully worded in an attempt to distinguish calendar/physical time (time at which value of index is recorded) from the transaction time (counter of number of observations since the start). The procedure of re-labelling a time series in order to obtain a regularly spaced time series is called time deformation.*

3. *For the ease of interpretation, t shall be called "Trading Day."*

4.3.3 Motivation for the Chosen Time Scale

According to Tsay (2002), financial data that are observed daily or at an even smaller time scales are regarded in financial research as high frequency data. These data have become available due mainly to the rapid advances made in data acquisition and processing techniques over the last ten years. There are two main reasons why studying high frequency financial data is most desirable in financial research:

1. Large sample size. Higher frequency data means that over the same given time period, more observations are made on the financial variable, which results in a larger sample size. According to Taylor (1986), large sample sizes is most desirable in financial researches. This is because a large sample sizes will lower the variance of parameter estimates and increase the power of statistical tests. Large sample size is a particularly important requirement for non-parametric statistical tests like R/S analysis.

2. It allows for the empirical study of the market microstructure. Jarrow, Maksimovic, Ziemba (1995) describes market microstructure as the study of the interaction of institutional trading rules with investors' information and trading preferences in order to determine the price process. For a critical review of the market microstructure literature, see Easley and O'Hara (1995).

At the time that the sample is obtained, the JSE-OVER index is calculated on four different time scales, namely: Daily, Weekly, Monthly, and Annually. Thus, the daily JSE-OVER index are the highest frequency data available on the JSE. The daily JSE-OVER index dataset is chosen for the exploration of CMH distribution on the JSE for the following reasons:

1. It provides the largest possible sample size out of all the JSE datasets available on Professor Troskie's database.

2. It provides the largest possible amount of detail on the dynamics of the JSE.

4.4 Preliminary Analysis of Data

4.4.1 Introduction to Statistical Time Series analysis

Recall that a stochastic process is a family of random variables indexed by set of time, which is defined formally as follows:

Definition 45 (Stochastic Process) *Let*

$$(\Omega, \Sigma, P) \text{ be a probability space,}$$

where:

Ω *is the sample space.*

Σ *is the sigma-algebra defined on subsets of* Ω

$\mathbf{P}$ *is the probability measure defined on* Σ.

Let T be a time index set. Define a real-valued random function $X(\cdot,\cdot)$ such that

$$X : (T \times \Omega) \to \mathbb{R}.$$

Then,

(i) *$X = \{X(t,\omega) : t \in T, \omega \in \Omega\}$ is called a stochastic process.*

(ii) *For a fixed $\omega \in \Omega$, $X(t,\omega)$ is a real-valued function of t, which is called a **realisation** (or **sample function**) of the stochastic process X.*

(iii) *The collection of all possible realisations (or sample functions) is called the **ensemble** of realisations (or sample functions).*

The key insight that made parametric statistical theory of time series analysis possible is the realisation that any given time series $(x_t)_{t\geqslant 0}$ can be regarded as an observation made on a family of underlying random variables $\{X_t : t \in T\}$, where T is some time index set. It follows that any observed time series can be regarded as a realisation from the underlying stochastic process. The utility of this conceptualisation is that if an observed time series $(x_t)_{t\geqslant 0}$ is assumed to be a realisation from its underlying stochastic process $\{X_t : t \in T\}$, then it is possible to make statistical inference on the probability law governing the underlying stochastic process given the observed time series. For a more detailed discussion of this idea see Parzen (1962), Box, Jenkins, Reinsel (1994), Fuller (1996), Chatfield (2001).

Applying the idea from the above discussion to the analysis of the observed daily JSE-OVER time series, the following assumption is made:

Condition 46 *It is reasonable to assume that $(p_t)_{t=1}^{4303}$, the observed daily JSE-OVER index time series can be regarded as a finite realisation of the underlying stochastic process $\{P_t : t = 1, 2, \cdots, 4303\}$.*

Thus, from now onwards, characteristics of an observed time series will be regarded as characteristic of the underlying stochastic process.

A stochastic process is characterised by its finite dimensional distribution function, which is defined formally as follows:

Definition 47 (Finite-Dimensional Distribution Function) *Let $\{X_{t_1}, X_{t_2}, \cdots, X_{t_n}\}$ be a finite subset of random variables from the stochastic process $X = \{X(t,\omega) : t \in T, \omega \in \Omega\}$. Then, the joint distribution function of $X_{t_1}, X_{t_2}, \cdots, X_{t_n}$ is defined by:*

$$\begin{aligned} & F_{X_{t_1}, X_{t_2}, \cdots, X_{t_n}}\left(x_{t_1}, x_{t_2}, \cdots, x_{t_n}\right) \\ = & \Pr\left[\left\{\omega : X(t_1,\omega) \leqslant x_{t_1}, X(t_2,\omega) \leqslant x_{t_2}, \cdots, X(t_n,\omega) \leqslant x_{t_n}\right\}\right] \end{aligned}$$

For practical implementation of models for stochastic processes, the estimation of parameters of the stochastic process must be mathematically tractable. This is not possible if the parameters of the stochastic process changes with time, which is characterised by the dependence of the finite-dimensional distribution of the stochastic process, $F_{X_{t_1}, X_{t_2}, \cdots, X_{t_n}}\left(x_{t_1}, x_{t_2}, \cdots, x_{t_n}\right)$ on the choice of time $t_1, t_2, \cdots, t_n$. Thus, it is desirable if the stochastic process under study has the following property:

Definition 48 (Strictly Stationary Stochastic Process) *Let $X = \{X(t,\omega) : t \in T\}$ be a stochastic process. Let $t_1, t_2, \cdots, t_n$ be any finite subset of T. Then, X is a strictly stationary stochastic process, if:*

(i) *The multivariate distribution of the random variables $X_{t_1+h}, X_{t_2+h}, \cdots, X_{t_n+h}$ is independent of h, i.e.)*

$$F_{X_{t_1}, X_{t_2}, \cdots, X_{t_n}}\left(x_{t_1}, x_{t_2}, \cdots, x_{t_n}\right) = F_{X_{t_1+h}, X_{t_2+h}, \cdots, X_{t_n+h}}\left(x_{t_1}, x_{t_2}, \cdots, x_{t_n}\right),$$

and

(ii) *h is chosen such that the translated index set is still a subset of T.*

Thus, strictly stationary stochastic processes requires that the finite-dimensional distributions characterising the stochastic process be invariant under a time shift. This is a very strong condition and requires one to empirically verify a large number of conditions. Thus, the first assumption in standard statistical analysis of time series is the follow:

Definition 49 (Weakly Stationary Stochastic Processes) *Let $X = \{X(t,\omega) : t \in T\}$ be* *a stochastic process. Let $t_1, t_2, \cdots, t_n$ be any finite subset of T. Then, X is a weakly stationary stochastic process, if:*

(i) $E\left[|X_t|^2\right] < \infty$, *for* $t \in T$

(ii) $E[X_t] = m$, $\forall t \in T$ *and* $m \in \mathbb{R}$

(iii) $\gamma_X(r,s) = \gamma_X(r+t, s+t)$, $\forall r,s,t \in T$, *where:*

$$\gamma_X(r,s) = E[(X_r - E[X_r]) \cdot (X_s - E[X_s])].$$

Remark 50 *Weakly stationary stochastic processes are also called wide sense (Doob (1953)), covariance stationary (Hamilton (1994)), or second-order (Gouriéroux (1997)) stationary stochastic processes.*

Remark 51 *Weak stationarity is theoretically less restrictive and implies that the underlying stochastic process can be characterised by its first two moments.*

Remark 52 (Relationship Between Weak and Strict Stationarity) *From the definitions of stationarity, a strictly stationary stochastic process with finite first and second moment is also a weakly stationary stochastic process. However, weakly stationary stochastic process does not imply that a stochastic process is strictly stationary. For an example, see Brockwell and Davis (1991).*

In order to make statistical inferences on the probability laws that governs the underlying stochastic process from a single realisation, the stochastic process needs to satisfy conditions of the Ergodic theorem. The ergodic theorem says that time averages for a single realisation converge to ensemble averages. This allows consistent estimates of the properties of a stationary process to be obtained from a single finite realisation. Doob (1953) shows that strictly and weakly stationary stochastic processes satisfies the ergodic properties, where the convergence to ensemble averages is stated as the Law of Large Numbers. For a good and concise explanation of the ergodicity properties for stochastic processes, see Parzen (1962). For discussion of the ergodicity assumption for time series in particular, see Hamilton (1994).

Remark 53 *Chatfield (2001) explains the ergodic property as that properties of the random variable at time t can be estimated with observations made at other times.*

Thus, for any statistical analysis of time series to be possible, one must ensure that the series under study is stationary. For a comprehensive account of non-linear time series analysis, see Tong (1996).

4.4.2 Time Plots and Tests of Stationarity

Figure 4-1 below shows the time plot of the observed daily JSE-OVER index time series computed by EVIEWS::

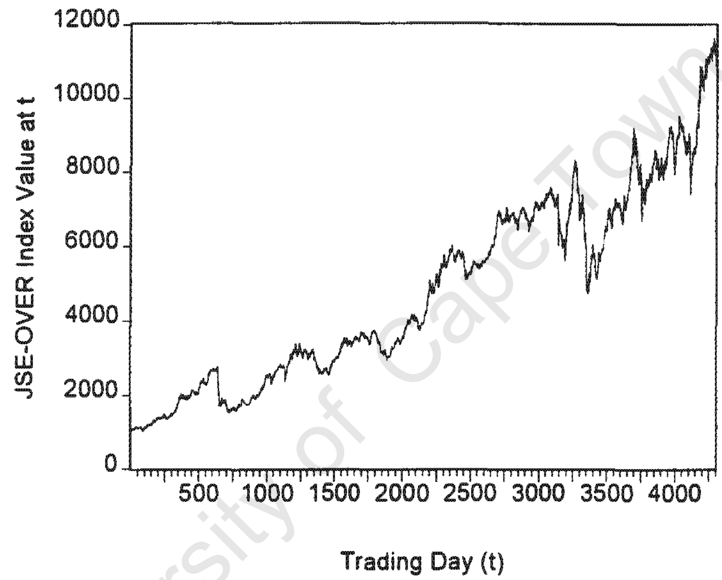

Figure 4-1: Time Plot of Observed Daily JSE-OVER Index Time Series

From Figure 4-1 above, the following characteristics are noted:

1. There is an exponentially increasing trend with variance appearing to increase with the trend. Thus, the observed daily JSE-OVER time series is not (weakly) stationary.
2. The increasing trend also implies that consecutive index values will be highly positively correlated, which was confirmed by an examination of the associated correlogram.

Log-Difference Transformation of the Observed Daily JSE-OVER Time Series

Empirical studies in financial economics are carried out (almost always) on *returns* for which Campbell, Lo and MacKinlay (1997) provided the following two reasons:

1. Return is a complete and scale-free summary of investment opportunities on the financial markets.

2. Returns have more desirable statistical properties than prices, particularly (weak) stationarity and hence ergodicity.

Campbell, Lo and MacKinlay (1997) then went on to define, in some detail, the various types of return that may be modelled, and motivated why the continuously compounded (or log) return should be the standard definition of return used for modelling by most current financial researches.

Definition 54 (Observed Daily Returns Time Series) *For $t = 2, \cdots, 4303$, the (continuously compounded) log-return on daily JSE-OVER index over the time interval $[t-1, t]$ is defined as:*

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right).$$

Thus, the observed daily return time series is defined by:

$$(r_t)_{t=2}^{4302} = (r_2, r_3, \cdots, r_{4303})$$

As was the case with the observed JSE-OVER index time series, the following assumption needs to be made in order to apply statistical theory on the observed daily returns time series:

Condition 55 *It is reasonable to assume that $(r_t)_{t=2}^{4303}$, the observed daily return time series can be regarded as a finite realisation of the underlying stochastic process $\{R_t : t = 2, \cdots, 4303\}$.*

Time Plot and Tests for Stationarity of Observed Daily Returns Time Series

Figure 4-2 below shows the time plot of observed daily returns time series computed by EVIEWS:

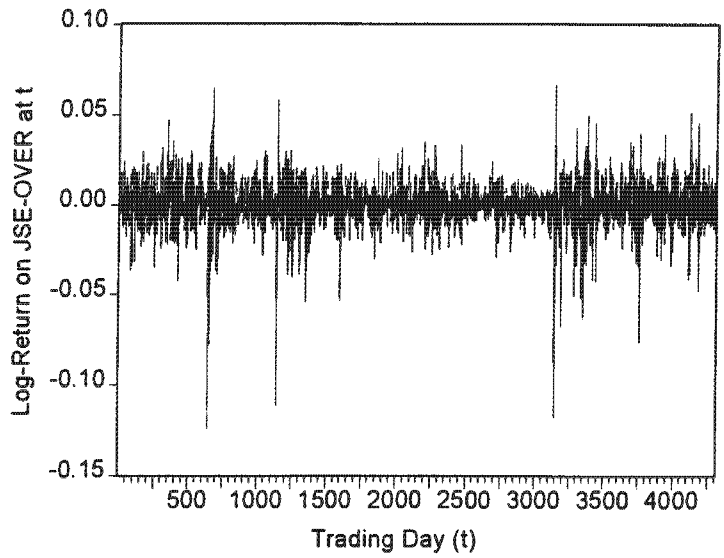

Figure 4-2: Time Plot of the Observed Daily Returns Time Series

From Figure 4-2 above, the following characteristics are noted:

1. The observed daily-returns time series looks relatively stationary with a constant mean daily return of roughly 0.
2. For r_t values that are close to or equal to zero, p_t and p_{t-1} are not very different.
3. Relatively large positive r_t value means that p_t is much greater than p_{t-1}, which indicates an increase in the daily index from $t-1$ to t. This suggest an upward movement on the JSE.
4. Relatively r_t values that are much smaller than 0, p_{t-1} is far much greater than p_t, which indicates a decrease in the daily index from $t-1$ to t. This suggests a downward movement on the JSE.
5. The very large positive returns are less than half the magnitude of the very negative returns. This is explained by the observation that prices takes longer to rise (due to holding of and excess demand for stock) than to fall (due to large quantities of stocks put up for sells at very low price).

6. The period prior to the 3100^{th} trading day have shown that extreme positive or negative returns were rare. Over this period, extreme returns occurred during two periods:

Trading day interval	Calendar Period	Major Market Event
$(600, 750)$	Sep 1987 to Apr 1988	Oct 1987 crash
$(1100, 1250)$	Sep 1989 to Apr 1990	Jan 1990 Gulf War

   The less major extreme returns occurred between $(1300, 1750)$.

7. The JSE is the least volatile over the period from 1600^{th} trading day to 3100^{th} trading day. This corresponded to the market boom during that period.

8. The extreme positive or negative returns have appeared more regularly with increasing magnitudes over the period starting from the 3100^{th} trading day until the end of the sample period. Loosely speaking, this observation may be described as the JSE have become more volatile over the last 5 years (roughly from June 1997 to June 2002).

Although, the time plot of the observed daily returns time series looks relatively stationary, but an unit root test on the observed daily returns time series is required to ensure that the series is indeed stationarity. Table 4.1 below shows the results of the Augmented Dicky-Fuller (ADF) test on the observed daily returns time series:

ADF Test Statistics =	-28.08216
MacKinnon critical values for rejection of H_0 :	1% Critical Value = -3.4350 5% Critical Value = -2.8627 10% Critical Value = -2.5674

Table 4.1: Results From Augmented Dicky-Fuller Test

From Table 4.1 above, note that:

$$\text{ADF Test Statistics} = -28.08216 \ll -2.8627 = 5\% \text{ Critical Value.}$$

Thus, the null hypothesis of a unit root (or non-stationary) time series is rejected at the 5% level. This implies that there is strong evidence to suggest that the observed daily returns time series is stationary.

4.4.3 Investigation of Distribution for Observed Daily Returns Time Series

Since the observed daily returns time series is weakly stationary, thus it is reasonable to assume that this time series is ergodic.

Histogram and Descriptive Statistics

According to Chatfield (2001), histogram are only useful if the time series is untrended and deseasonalised. Figure 4-3 below shows the histogram and the important descriptive statistics for the observed daily returns time series computed by EVIEWS:

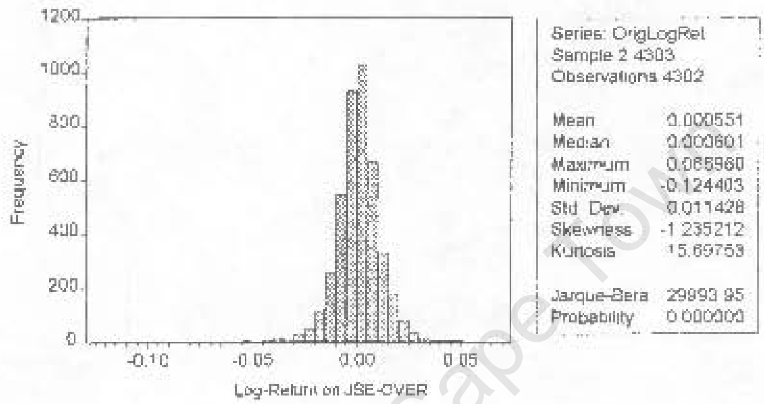

Figure 4-3: Histogram and Descriptive Statistics for Observed Daily Returns

From Figure 4-3 above, the following characteristics are noted:

1. The observed daily returns series has a small, but positive mean daily return of 0.000551 (or 0.055%).

2. The median daily return of 0.000601 (or 0.060%) is greater than the mean daily return. This means that there are more positive daily returns, but the magnitude of the negative daily returns are larger than the magnitude of the positive daily returns. Thus, the distribution for observed daily returns time series should have a longer left tail than right tail.

3. The standard deviation of observed daily returns of 0.011428 (or 1.143%) means that the

largest negative (smallest) return is 10.838 standard deviations away from the mean daily return while the largest positive return is 5.811 standard deviations away from its sample mean. Now, for a normally distributed log returns series, the probability of occurrence of a value that is 4 standard deviations away from its sample mean is 0 almost surely. Yet, an 11 standard deviation observation has occurred within a sampling period of 18 calendar years leads one to suspect that the observed daily returns time series may not be normally distributed.

4. The range of the observed daily returns time series is:

$$\begin{aligned} Range(\{r_t\}) &= \max\{r_t\} - \min\{r_t\} \\ &= 0.066960 - (-0.124403) \\ &= 0.19136 \text{ (or } 19.136\%) \end{aligned}$$

This means that the observed daily returns have a wide range of values. Thus, the standard deviation of observed daily returns is relatively small when compared to the range. This is due to a large number of observed daily returns with values close to the mean daily return, which offsetted the extreme positive and negative daily returns. This suggests that the distribution of observed daily returns time series is highly peaked at the mean and is heavy-tailed.

5. The coefficient of skewness of $-1.235212 \ll 0$ indicates that the distribution for the observed daily returns time series is asymmetric about its mean, and is significantly more negatively skewed than the corresponding normal distribution (which has a coefficient of skeweness of exactly 0). This means that the distribution for observed daily returns have tails that extends further towards the left than to the right due to the occurrence of more large negative returns than large positive returns. This confirms characteristic 2. noted above.

6. The coefficient of kurtosis of $15.69753 \gg 3$ indicates that the distribution of the observed daily returns time series is significantly leptokurtic (i.e. much higher-peaked at its mean and heavy-tailed than the corresponding normal distribution, which has a coefficient of

kurtosis of exactly 3. This confirms characteristic 4. noted above.

7. The observed Jarque-Bera test statistics value (calculated from the coefficients of skewness and kurtosis) of 29993.95 is far larger than the critical chi-squared value ($\chi^2_{0.95,2}$) of 5.99, which is confirmed by a reported p-value of 0.000000. This implies that mill hypothesis of normality is rejected at the 5% (and even at the 1%) significance level. Thus, there is very strong evidence to support the claim that the distribution of observed daily returns time series does not have the shape of a normal distribution.

Sample Auto and Partial Autocorrelation

Since the observed daily returns time series is weakly stationary, thus it is meaningful to estimate the sample auto and partial autocorrelation function. Figure 4-4 below shows the sample autocorrelation function (acf) and partial autocorrelation function (pacf) of the observed daily returns time series for lags up to 15 trading days as computed by EVIEWS:

Sample: 1 4303
Included observations: 4302

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.147	0.147	92.919	0.000
		2	0.060	0.039	108.29	0.000
		3	0.035	0.021	113.54	0.000
		4	0.004	-0.006	113.62	0.000
		5	-0.014	-0.017	114.47	0.000
		6	-0.014	-0.011	115.38	0.000
		7	0.035	0.042	120.79	0.000
		8	0.063	0.056	138.09	0.000
		9	0.029	0.010	141.69	0.000
		10	0.005	-0.008	141.81	0.000
		11	0.017	0.012	143.10	0.000
		12	0.029	0.026	146.69	0.000
		13	-0.015	-0.022	147.72	0.000
		14	-0.035	-0.033	153.14	0.000
		15	0.020	0.027	154.94	0.000

Figure 4-4: Correlogram for Observed Daily Returns Time Series

From Figure 4-4 above, the following interpretations may be made:

1. The lag 1 acf and pacf estimate of 0.147 is far larger than the 95% standard error bounds (as indicated by the dotted line). This means that the lag 1 acf and pacf estimate is significantly different from zero at the (approximate) 5% significance level. However, an acf and pacf value of only 0.147 means that it only explains $0.147^2 = 0.021609$ or 2.16% of the variation. Thus, the evidence is inconclusive to suggest that the observed daily returns time series has a significant first order acf and pacf.

2. The approximate 95% standard error for both the acf and pacf estimates is $\frac{2}{\sqrt{T}}$, where T is the number of observations in the observed time series. The observed daily returns time series has a sample size of 4302 observations, which is extremely large. As a result, the standard error bound of the estimates will be extremely small ($\pm\frac{2}{\sqrt{4302}} = \pm 0.030493$). Thus, the fact that an order k acf or pacf estimate have a value exceeding this error bound cannot be taken as conclusive evidence for a significance.

3. The lag 1 autocorrelation has a Q-Statistic value of 92.919, which is extremely large. However, note that the reported probabilities (p-values) for the Q-statistics are 0.000 for all lags. This means that the null hypothesis of no autocorrelation up to order k is rejected at all $k = 1, 2, \cdots, 4301$, and every acf is significant at the 5% level. Again, the Q-statistics is dependent on the sample size, and the larger the sample size the larger the value of the Q statistics, which explains why the reported probabilities at all lags are zero. Thus, a significant Q-statistics cannot be taken as conclusive evidence for a significant order k acf.

4. Interpretations 1. to 3. are typical feature of high frequency data, where there is always a large sample size. Thus, classical time series statistical tests becomes meaningless and no conclusions may be drawn based on it.

5. There is no exponential decay in the acf as lags increases. This implies that the observed daily returns time series cannot be explained by a pure autoregressive process of order k. Furthermore, the relatively large autocorrelation at higher lags suggests that there may exists long memory in the time series.

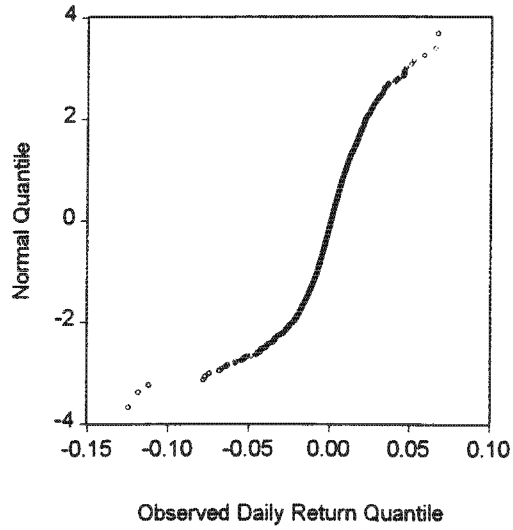

Figure 4-5: The Normal Quantile-Quantile Plot of the Observed Daily Returns Time Series

Normal Quantile-Quantile Plots

The normal quantile-quantile (Q-Q) plot is a plot of the observed daily returns quantile against the corresponding quantile of the normal distribution. It provides visual evidence on the observed daily returns' departure from normality. If the quantiles from the observed daily return exactly matches the quantiles from the corresponding normal distribution, then the normal Q-Q plot will lie exactly on a straight 45° line. The manner at which the normal Q-Q plot departs from this 45° line will provide an indication of how and where the observed daily return's departure from normality. Figure 4-5 below shows the normal Q-Q plot for the observed daily returns time series computed by EVIEWS:From Figure 4-5 above, the following interpretations may be made:

1. The normal Q-Q plot does not lie on the straight line given by Normal Quantile=Observed Daily Returns Quantile. Thus, the distribution for observed daily returns time series is clearly not normal.

2. The steep slope of the normal Q-Q plot over the observed daily return (empirical) quantile interval $(-0.025, 0.04)$ indicates that there is a high concentration of observed daily returns

with values in that interval. Thus, the distribution for observed daily returns is much higher-peaked than the corresponding normal distribution over some part of that empirical quantile interval. Since the steep slope occurs over a wider positive empirical quantile interval than the negative empirical quantile interval, thus there is more positive observed daily returns than negative, which means that it is reasonable to expect the mean daily return to be a positive value.

3. The relatively abrupt increase in the slope of the normal Q-Q plot from over empirical quantile interval $(-0.10, -0.025)$ (less steep) to empirical quantile interval $(-0.025, 0.04)$ (more steep) indicates that the distribution for observed daily return is asymmetric about its mean and is skewed to the left.

4. The three extreme negative quantiles appearing over the empirical quantile interval $(-0.15, -0.10)$ has larger magnitude than the four extreme positive quantiles appearing over empirical quantile interval $(0.05, 0.10)$. This indicates that the distribution for observed daily returns will have a longer left tail than the right tail. Since coefficient of skewness is very sensitive to outlying observations (hence not robust), thus it is reasonable to expect that the distribution for observed daily returns will have a negative coefficient of skewness.

4.4.4 Tests for Long-Term Memory: Results From R/S Analysis

The R/S analysis is performed on the observed daily return time series from 27 March 1985 to 26 July 2000, which consists of 3826 observations covering a time length of 5601 calendar days (or approximately 15 calendar years). The none availability of data on days when JSE does not trade will not affect R/S analysis, because Hurst exponent is a rough measure of the memory in the market. Since FMH emphasises the identification of global determinism, thus a few days of discontinuity is unlikely to affect the overall pattern exhibited by the financial market.

Figure 4-6 below shows the log-log plot of the sample and the expected R/S values computed based on the VBA programs written (see Appendix D):

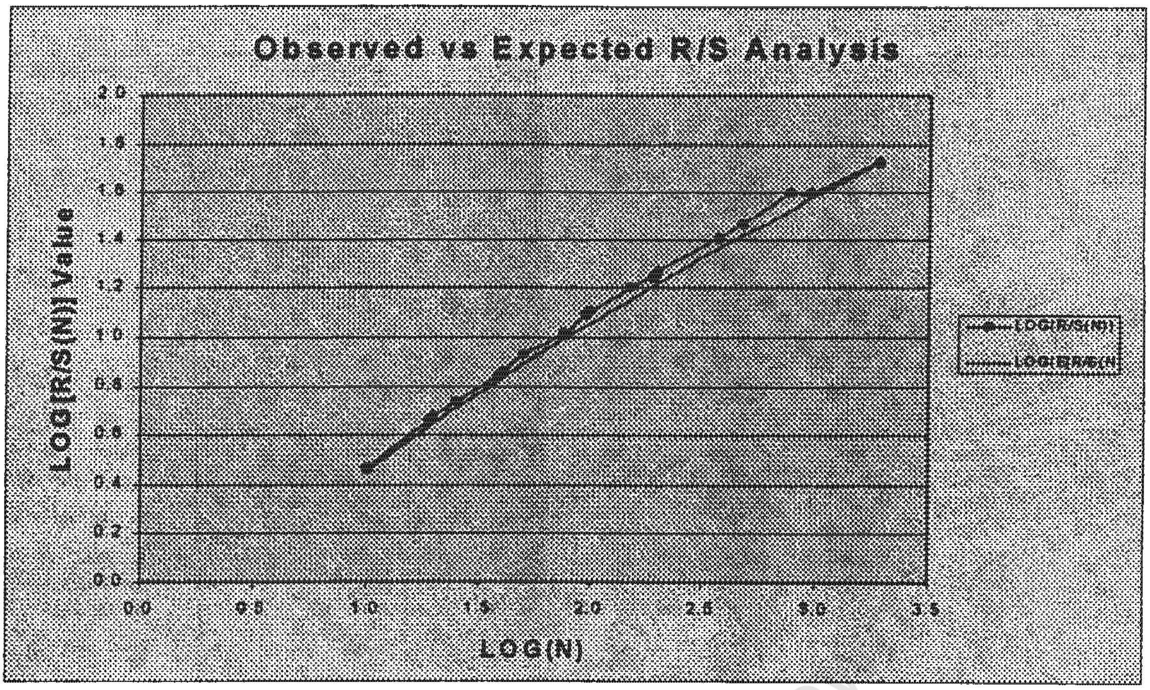

Figure 4-6: Sample versus Expected R/S Values

From Figure 4-6 above, the following characteristics were noted:

1. There is a persistent upward bias of the sample R/S values from the expected R/S values under the null hypothesis.

2. There seems to be a discontinuity in the sample log-log plot at $log(n)$ of roughly 2 and 3 respectively. These discontinuities indicates the presence of cycles that changes direction at these points.

Table 4.2 below shows the regression output for the estimated sample Hurst exponent for **all values of** n:

From Table 4.2 above, the estimated sample Hurst exponent is 0.560, which implies that the log returns series exhibits persistence. Table 4.3 below shows the expected Hurst exponent for all values of n, which is required to test for the significance of the estimated sample Hurst exponent value:

(10 < n < 1900) Hurst Exponent (β)	LOG(Constant) (α)
0.560	-0.040
0.010	0.022
0.995	0.027
3105.695	16.000
2.233	0.012

Table 4.2: Estimated Sample Hurst Exponent over all values of N

(10 < n < 1900) E(H) (β)	LOG(Constant) (α)
0.552	-0.059
0.005	0.011
0.999	0.014
11646.672	16.000
2.171	0.003

Table 4.3: expected Hurst exponent for all values of N

and the standard deviation of the expected Hurst exponent for all values of n is:

$$s.d.\left[E\left(H\right)\right] = \frac{1}{\sqrt{T}} = \frac{1}{\sqrt{3800}} = 0.016$$

From table 4.3 above, the expected Hurst exponent is 0.552. Thus, at the 5% significance level, the sample Hurst exponent over all values of n is insignificant. Thus, one concludes that there is insufficient evidence to reject the null hypothesis at the 5% level and that the log returns are independent of each other and normally distributed.

However, it was noted previously that the sample log-log plot have shown that there is discontinuities at roughly $log(n) = 2$ and 3 respectively. Thus, the V-statistics method was implemented to identify the time of the discontinuity more precisely and hence determine the cycle exhibited by this log returns series. Figure 4-7 below shows the plot of sample and expected V-statistics against $log(n)$:

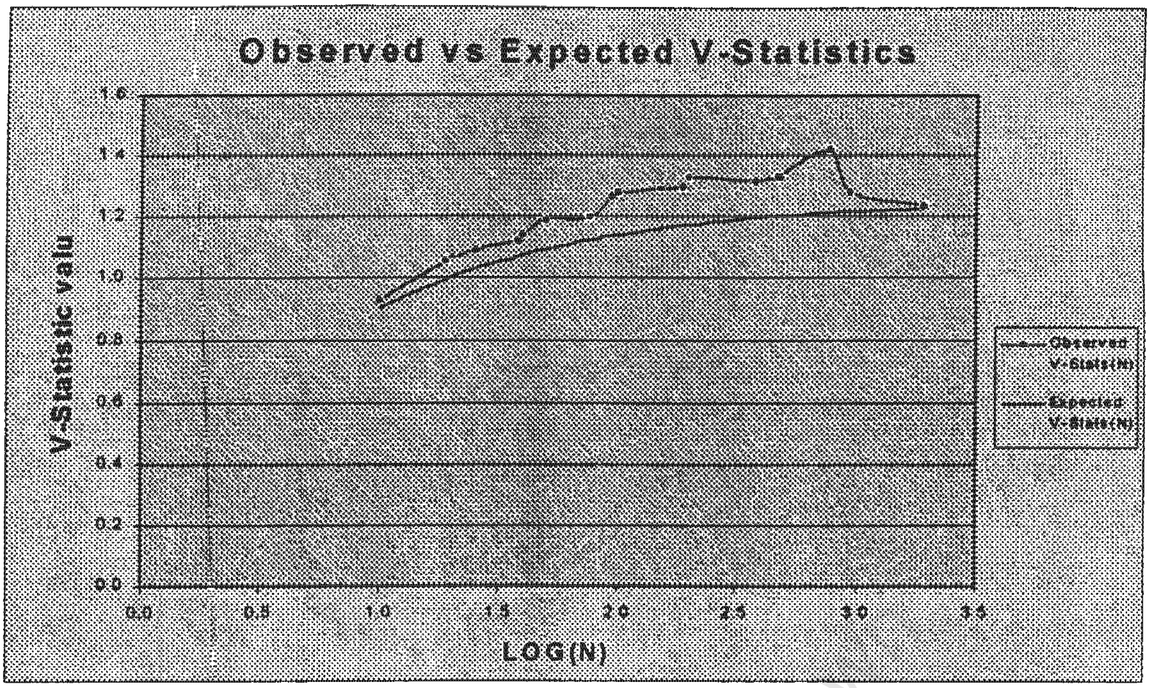

Figure 4-7: Sample versus Expected V-Statistic Values

From Figure 4-7 above and by examining the V_n values for the various $n's$, the following interpretations may be made:

1. The sample V-statistics is consistently greater than the expected V-statistics for all values of n.

2. The sample plot levelled off at $log(n) = 1.699$, which implies that the log returns process has changed from slightly persistent to independent. This indicates that a minor cycle is changing direction.

3. The sample plot had a sharp peak at $log(n) = 2.881$, which implies that there is an abrupt change in the characteristic of the log returns process from strongly persistent to strongly anti-persistent. This indicates a major/primary cycle that is changing its direction from one extreme event to another.

Based on the above observations, one can conclude that the discontinuity occurs at $log(n) =$

1.699 and 2.881, which translates to 50 and 760 days or roughly 2 months and 2 years respectively.

In order to assess the statistical significance of the memory and the cycles identified above, one needs to estimate the sample and the expected Hurst exponent over the following intervals of n:

Interval 1: $10 \leq n \leq 50$

Interval 2: $50 \leq n \leq 760$

Interval 3: $760 \leq n \leq 1900$

by running the regression over these intervals. Table 4.4 below shows the estimated regression coefficients (representing the estimated sample Hurst exponent):

	H	$E[H]$	**P-value**
$10 \leq n \leq 50$	0.641	0.615	$P < 0.10$
$50 \leq n \leq 760$	0.556	0.536	$P < 0.20$
$760 \leq n \leq 1900$	0.372	0.514	$P < 0.0001$

Table 4.4: Estimated Sample Hurst Exponent Over Different intervals of n

From the Table 4.4 above, the significance (at the 5% level) of the sample Hurst exponent is as follows:

1. For $10 \leq n \leq 50$, the sample Hurst exponent indicates a nearly significant persistence.

2. For $50 \leq n \leq 760$, the sample Hurst exponent indicates a possibly significant persistence.

3. For $760 \leq n \leq 1900$, the sample Hurst exponent indicates a very highly significant anti-persistence.

The above results, confirmed the observations made from Figure 4-7

In order to detect and measure the effects of the bias on the above R/S analysis, one needs to compare plot of log first difference of the log returns with the AR(1) residuals in order to determine whether a significant serial correlation exists in the log returns. Figure 4-8 below shows the V-statistics plot for the log returns and Expected R/S values against the AR(1) residuals:

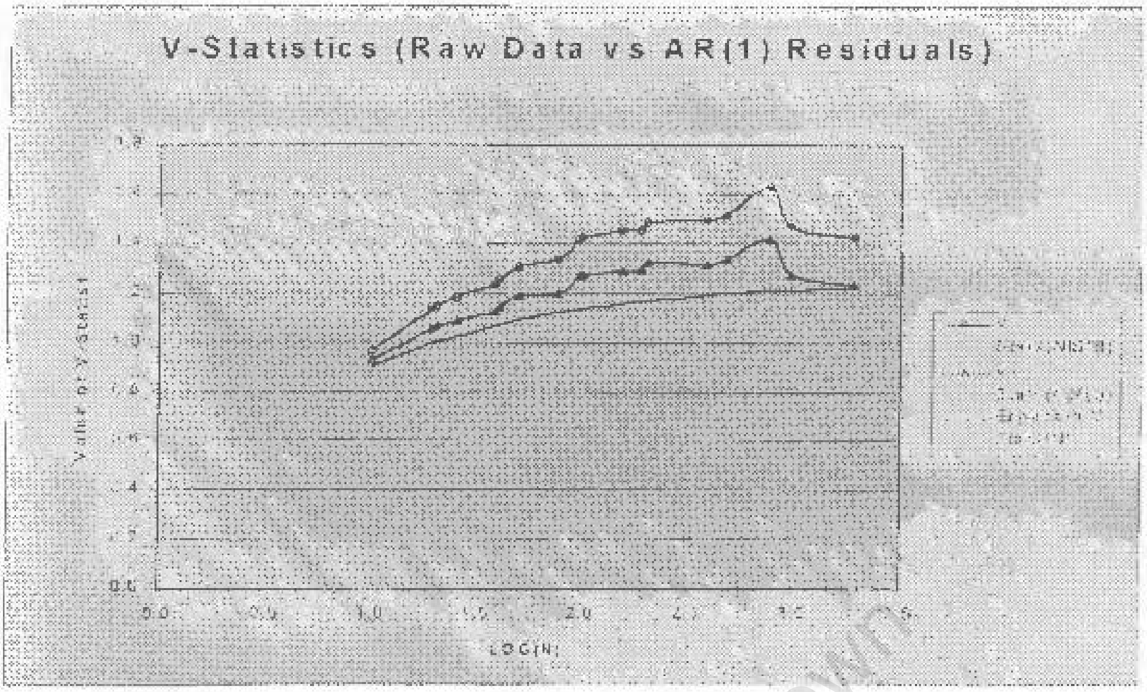

Figure 4-8: V-Statistic Values for Log-Transformed JSE OVERALL Index against its AR(1) Residual

From Figure 4-8 above, the following interpretations may be made:

1. A small AR(1) bias in the log returns causes the expected R/S values to be a little higher than when using AR(1) residuals.

2. The Hurst exponent estimate is also slightly biased.

Explanations for the above interpretations are as follows:

1. Since all R/S values are biased upwards, thus the scaling feature, as measured by the Hurst exponent, is slightly affected by the bias, although the bias is definitely present.

2. This may be caused by the impact of higher frequency (daily) of data, which seems to worsen the impact of short-term memory process on R/S analysis.

4.4.5 Conclusions From Preliminary Analysis of Data

Based on the findings above, the following conclusions may be drawn:

1. The distribution for the observed daily returns time series is weakly stationary.

2. The distribution for the observed daily return time series is asymmetric about its mean and it is skewed to the left.

3. The distribution for the observed daily returns time series have a relatively higher peak at the mean and a much heavier tail than the corresponding normal distribution.

4. Due to the large sample size, no conclusions may be drawn on the significance of sample acf and pacf at lag k based on the classical hypothesis tests.

5. Results from the R/S analysis have provided evidence to suggest that the JSE exhibits long-term dependence, and hence have a self-similar structure. This means that the distribution for observed daily returns time series will not be normally distributed.

6. R/S analysis have further identified a major cycle of length 2 years and a minor cycle within it of length 2 months on the JSE. Based on these empirical evidence, one may make the interpretation that, on average, the JSE has turned from an extreme positive market to an extreme negative market (or vice versa) every 2 years, while there was a less significant turn every 2 months.

7. The major problem with R/S analysis is that it does not have enough statistical power to reject the null hypothesis. Thus, a large number of observations covering a long time interval is required to obtain a statistically significant sample Hurst exponent. The analysis performed in this dissertation was not able to find conclusive evidence that there is long-term memory in the observed daily returns time series. However, this is not because that long-term memory does not exist on the JSE at all, but because there is insufficient amount of data available to support the claim.

8. If one ignores the extreme observations in a log returns series, then the returns will actually be normally distributed. Although, one can justify such action by arguing that these extreme observations are comparatively rare and thus will be averaged out in a large sample. However, such argument would be flawed for financial returns series, because extreme observations impacts on the participants of the market the most. Large amounts

of profits and losses may be caused by these extreme observations. In fact, one can go further and argue that actually one should only concentrate on these observations as one's fortune is depended upon it.

4.5 Empirical Market State Partition of Observed Log-Returns Series

In this section, characteristics for every observed daily return is quantified using technical analysis indicators. Then, each observed daily return will be classified as being in one of the five market states based on the characteristics that it exhibits.

According to Vaga (1990), technical analysis is an effective tool in identifying upward trends in the market, if the market dynamic is in the coherent bull state. The general goal of technical analysis is summarised by Lo, Mamaysky, Wang (2000) as the identification of regularities in financial time series by extracting non-linear patterns from the noise in the data. The tools used by technical analysis is geometry and pattern recognition. The technical analysis indicators used for the partitioning of the observed daily returns time series is calculated by the software **VisuaData**, which is copyrighted to NewWave Intelligent Business systems, Inc. Table 4.5 below summarises the technical indicators used:

Indicator	Name	Inputs	Lookback Period
rwih (short)	Random walk index on high price	CLOSE	2
roc (mo)	Rate of change (momentum)	CLOSE	10
dm_dx	Net directional movement	HIGH,LOW	14
volatilityc	Chaikin's volatility	HIGH,LOW	10

Table 4.5: Technical Indicators for Daily Return Partitioning

From Table 4.5 above, the technical analysis indicators are interpreted (according to **VisuaData** Users' Manual) as follows:

1. The random walk index (RWI) was developed by Poulos (1992) to distinguish trending prices from random prices:

$$rwi \begin{cases} > 1 & \text{trending} \\ \leq 1 & \text{random walk expected} \end{cases}$$

2. The rate of change or momentum index (roc /mo) to identify when a signal is rising (gaining momentum) or falling (losing momentum):

$$roc \begin{cases} > 0 & \text{time series is rising} \\ = 0 & \\ < 0 & \text{time series is falling} \end{cases}$$

3. The net directional movement index (dm_dx) is a measure of the total strength of the trend:

$$dm\_dx \begin{cases} > 0 & \text{Bullish, buy signal} \\ = 0 & \\ < 0 & \text{Bearish, sell signal} \end{cases}$$

Based on the above technical analysis indicators, the observed daily returns series will be partitioned into five market states.

4.5.1 Step 1: Classify returns as Random Walk and Non-Random Walk

The initial criterion is that if a index has $rwih(short) < 1$, then that index can be interpreted as showing no trending (over the last two trading days) and hence we expect it to be a random walk. Based on the stated criterion, we got the RW random walk log-return series. However, the RW's Jarque-Bera value is highly significant at the 5% level, hence it does not have a normal distribution as expected. In order to obtain a log-return series from RW that has a normal distribution, we have to further remove all log-return values that is beyond 3 standard deviations away from its sample mean of -0.0021, i.e.

$$\begin{aligned} RW1 &\in (-0.0474, 0.0432) \\ &= \text{Mean}_{RW} \pm (3 \times \text{Std. Dev.}_{RW}) \end{aligned}$$

However, $RW1$'s Jarque-Bera is still highly significant at the 5% level, hence still does not have a normal distribution. Thus, we need to further remove log-returns in $RW1$ that has value beyond 3 standard deviations from its sample mean of −0.0016, i.e.

$$
\begin{aligned}
RW2 &\in (-0.0373, 0.0341) \\
&= \text{Mean}_{RW1} \pm (3 * \text{Std. Dev.}_{RW1})
\end{aligned}
$$

Unfortunately, $RW2$'s Jarque-Bera is still significant at the 5% level, hence still does not have a normal distribution. Thus, we need to further remove log-returns in $RW2$ that has value beyond 3 standard deviations from its mean of −0.0014, i.e.

$$
\begin{aligned}
RW3 &\in (-0.0347, 0.0319) \\
&= \text{Mean}_{RW2} \pm (3 * \text{Std. Dev.}_{RW2})
\end{aligned}
$$

Finally, $RW3$'s Jarque-Bera is not significant at the 5% level, hence it has a normal distribution. A summary of the process is given below:

	RW	RW1	RW2	**RW3**
No. of Obs.	910	894	883	**868**
Mean	−0.0021	−0.0016	−0.0014	**−0.0011**
Std. Dev.	0.0151	0.0119	0.0111	**0.0103**
Skewness	−1.5132	−0.2923	−0.1852	**−0.0732**
Kurtosis	14.1512	4.3794	3.7212	**3.2590**
Jarque-Bera	5062.19	83.61	24.19	**3.2026**
Prob.	0.0000	0.0000	0.0000	**0.2016**

Thus, we take $RW3$ as our Random walk log-return series. Then, the original log-returns series less the random walk series can be regarded as our "non-random walk" series (NonRW), with the following characteristics:

	NonRW
No. of Obs.	3417
Mean	0.0010
Std. Dev.	0.0117
Skewness	−1.4708
Kurtosis	17.8808
Jarque-Bera	32759.02
Prob.	0.0000

4.5.2 Step 2: Classify Non-RW returns into Chaotic/Transition and Coherent Bull/Bear

According to the criterion:

Market State	dm_dx	$roc(mo)$	$\frac{r}{\sigma}$(σ of returns)	$rwih(short)$
Coherent Bull	> 0	> 0	> 1.0	> 1
Coherent Bear	< 0	< 0	< -1.0	> 1
Chaotic/Transition	0?	0?	$[-1, 1]$	> 1

Then, the NonRW log returns series is classified into the following three series with the following characteristics:

	Bearb4	ChaosTrans3	Bullb4
No. of Obs.	305	2347	765
Mean	−0.0167	−0.0003	0.0119
Std. Dev.	0.0171	0.0078	0.0073
Max.	−0.0014	0.0670	0.0654
Min.	−0.1244	0.0485	0.0006
Skewness	−3.096	0.7411	1.7897
Kurtosis	15.5885	12.5471	9.2614
Jarque-Bera	2501.28	9128.18	1658.05
Prob.	0.0000	0.0000	0.0000

It is interesting to note that the Bullb4 log-returns series does not include the max. log-return value in original series. It is instead included in ChaosTrans3. Is there something wrong with the indicators or the way we are choosing?

Figure (4-9) below summarises the results for log-returns in bullish market state:

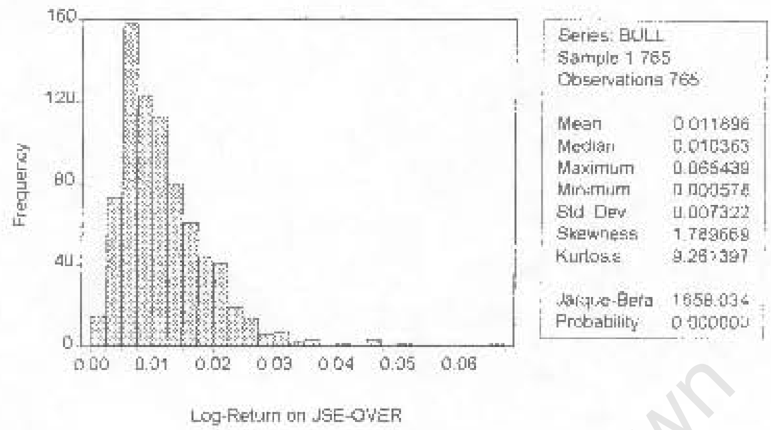

Figure 4-9: Histogram and Descriptive Statistics for BULL log-returns

From Figure 4-9 above, the following characteristics are noted:

1. The distribution is skewed to the right (i.e. it has a long right-tail). This is confirmed by a coefficient of skewness of $1.79 \gg 0$.

2. The distribution has a positive mean daily return of 1.19% and a standard deviation of 0.73% over the sampled period.

Figure 4-10 below summarises the results for log-returns in bearish market state:

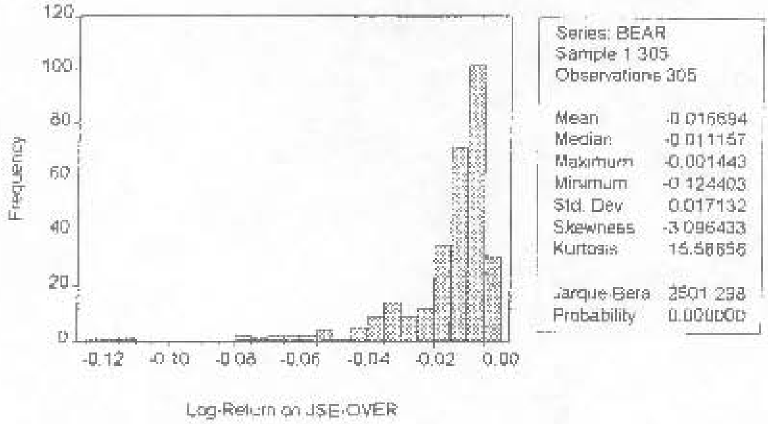

Figure 4-10: Histogram and Descriptive Statistics for BEAR log-returns

From Figure 4-10 above, the following characteristics are noted:

1. The distribution is skewed to the left (i.e. there is a long left tail). This is confirmed by a coefficient of skewness of $-3.10 \ll 0$.

2. The distribution has a negative mean daily return of -1.67% and a standard deviation of 1.71% over the sampled period.

4.5.3 Step 3: Classify Chaotic/Transition returns into Chaotic and Transition

In order to identify daily returns that exhibits characteristics of Chaotic and Transition market state, the volatility associated with each daily return is required to be computed. Thus, before the filtering criterion may be applied, an exploratory investigation on the distributional properties of volatility of original log-return series is required. Figure 4-11 below shows the log-return volatility series calculated using the technical analysis indicator *volatilityc*:

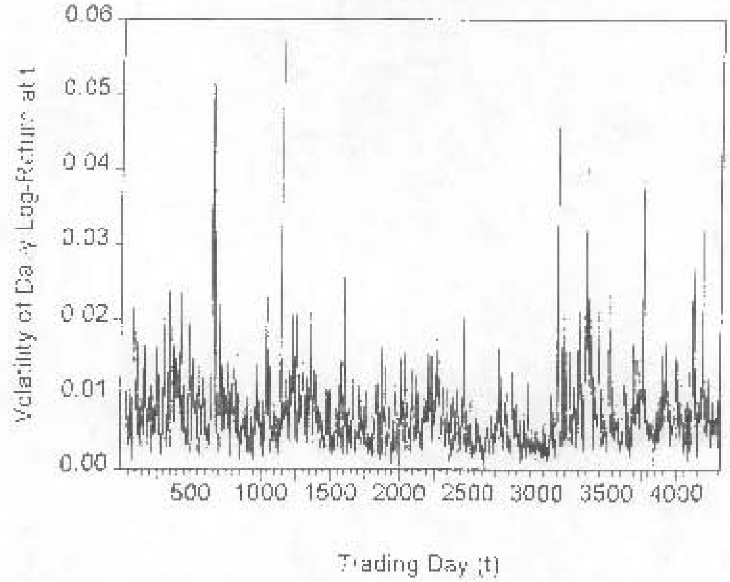

Figure 4-11: Volatility Series Associated with Original Daily Returns

From Figure 4-11 above, note that there is a clumping of large volatilities from trading day 3000 until the end of the sample period.

Figure 4-12 below shows the histogram and standard descriptive statistics for the daily volatility series:

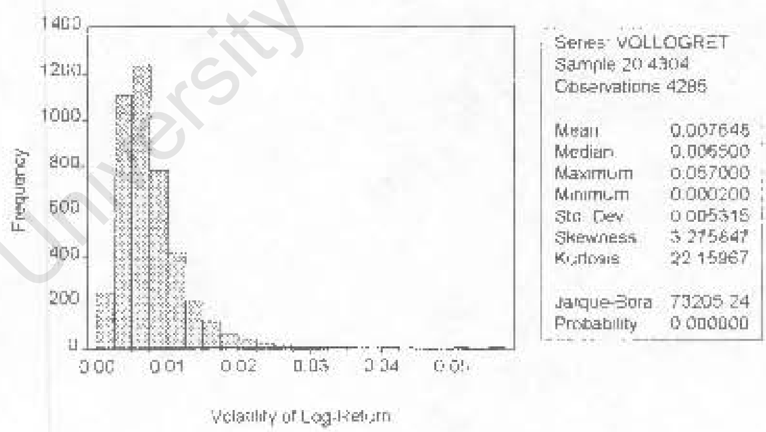

Figure 4-12: Histogram and Descriptive Statistics for Daily Volatility

From Figure 4-12 above, note that the coefficient of skewness of 3.28 is considerably larger than that for the normal distribution. It is skewed to the right.

Based on the computed daily volatility series, the following filtering criterion will be applied in order to identify daily returns that exhibits characteristics of chaotic and transition market state:

	$VolLogret$	dm_dx	$roc(mo)$	$\frac{r}{\sigma}(\sigma$ of returns)	$rwih(short)$
Chaotic	$> mean + S.D.$	0?	0?	$[-1,1]$	> 1
Transition	$otherwise$	0?	0?	$[-1,1]$	> 1

Then, the ChaosTrans log-return series is classified into the following two series, with the following characteristics:

	ChaoticSDa	TransSDa
No. of Obs.	237	2110
Mean	0.0018	−0.0005
Std. Dev.	0.0165	0.0061
Skewness	0.5402	−0.1539
Kurtosis	4.7521	7.6142
Jarque-Bera	41.84	1880.14
prob.	0.0000	0.0000

We note that ChaoticSDa have more or less of a bimodal histogram.

Figure 4-13 below summarises the results for log-returns in the chaotic market state:

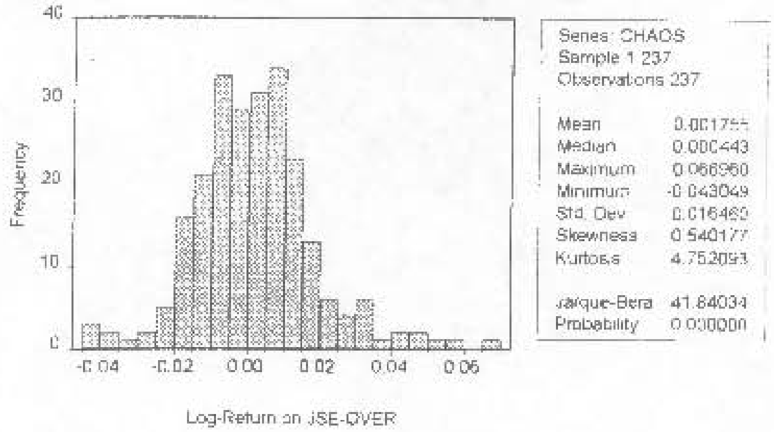

Figure 4-13: Histogram and Descriptive Statistics for CHAOS log-return

From Figure 4-13 above, the following characteristics are noticed:

1. There is a slight bimodality in the distribution with modes at ± 0.01.
2. The mean return is 0.18% with a standard deviation of 1.65% over the sampled period.
3. The distribution is slightly skewed to the right.
4. The distribution have very heavy tails, with the left tail heavier than the right. This is reflected by a comparatively low excess kurtosis of 1.75, since the extreme negative returns cannot be completely balanced by extreme positive returns.

4.5.4 Step 4: Classify Transition returns into Random Walk and True Transition

Under Vaga (1990)'s CMH, the original log-returns series less the following series:

1. Random Walk ($RW3$),
2. Coherent Bull ($Bullb4$),
3. Coherent Bear ($Bearb4$), and
4. Chaotic ($Chaoticb4$)

should leave us with the log-return series that is in a transition state, which should have an uniform distribution. However, our *TransSDa* series still has a normal distribution looking histogram. Since sum of a large number of uniform distributions will give you a normal distribution (unclear! verify this!), we next attempt to use volatility of random walk log-returns to separate out all the normally distributed log-returns out of our *TransSDa* series. It is hoped that by recombining various component normal distributions we will get the desired uniform distribution. But, it is not yet determined what the volatility of a random walk return should be. Thus, we take the standard deviation of our *RW3* (Std. Dev_{RW3}=0.0103) series to be representative of what the random walk volatility should be. Then, we classify the log-returns of our *TransSDa* series according to which $S.D._{RW3}$ interval does the volatility of the returns fall into. Initially, we partition SD starting from 0 and increasing at increments of $0.5 * SD_{RW3}$ and got the following results:

	NesRW1	NesRW2	**NesRW3**
Volatility	[0, 0.0052)	[0.0052, 0.0103)	[0.0103, 0.0155)
No. of Obs.	767	1117	**226**
Mean	−0.0006	−0.0005	**−0.0007**
Std. Dev.	0.0038	0.0068	**0.0086**
Max.	0.0175	0.0432	**0.0209**
Min.	−0.0187	-0.0485	**−0.0273**
Skewness	−0.1887	−0.0503	**−0.3443**
Kurtosis	4.4521	7.6062	**3.2231**
Jarque-Bera	71.9438	987.9324	**4.9331**
Prob.	0.0000	0.0000	**0.0849**

From the above table, we note the following:

1. There are no log-returns in *TransSDa* series that has a volatility greater than $1.5 * S.D._{RW3}$.

2. NesRW3 has a normal distribution, since at the 5% significance level its Jarque-Bera statistics value of 4.9331 is not significant.

3. Neither NesRW1 nor NesRw2 series have a normal distribution. Both have a much higher kurtosis than 3, thus are more higher peaked than normal distribution. However, NesRW1 is more normally distributed than NesRW2. But, NesRW2 have more than half of the observations of *TransSDa* series.

Based on the above observations, we shall further classify NesRW1 and NesRW2 by partitioning $[0, SD_{RW3})$ into four equal sub-intervals. We obtain the following results:

	NesRW11	NesRW12	NesRW21	NesRW22
Volatility	$[0, 0.0026]$	$(0.0026, 0.0052]$	$(0.0052, 0.0078]$	$(0.0078, 0.0103]$
No. of Obs.	124	643	688	429
Mean	-0.0009	-0.0005	-0.0002	-0.0010
Std. Dev.	0.0028	0.0040	0.0061	0.0078
Max.	0.0063	0.0175	0.0432	0.0334
Min.	-0.0109	-0.0187	-0.0209	-0.0485
Skewness	-0.8336	-0.4655	-0.6901	-0.5327
Kurtosis	4.4445	4.2704	7.7695	6.6062
Jarque-Bera	25.1421	46.1769	706.7311	252.7404
Prob.	0.0000	0.0000	0.0000	0.0000

From the above table we note the following:

1. NesRW12 returns has the higher volatility and its max. and min. value is also the max. and min. value of NesRW1. Thus, this is consistent with our intuition that the higher the volatility, the more extreme the log-return will be. Both NesRW11 and Nes RW22 are nearly normally distributed. It is not due to the new extreme observations (i.e. the max. and min. of *TransSDa*.

2. However, the intuition that higher volatility should produce more extreme returns did not materialise for NesRW21 (lower volatility and contains max. of *TransSDa*) and NesRW22 (higher volatility and contains min. of *TransSDa*) series. These two series also has the most observations (688 and 429 respectively), thus there is further investigation needed for this part of the original log-returns series.

4.5.5 Step 5: Combining NesRW3 and RW3

Figure 4-14 below summarises the results from combining the random walk part of log-returns in transition state ($NesRW3$) and random walk log-returns identified in the beginning ($RW3$):

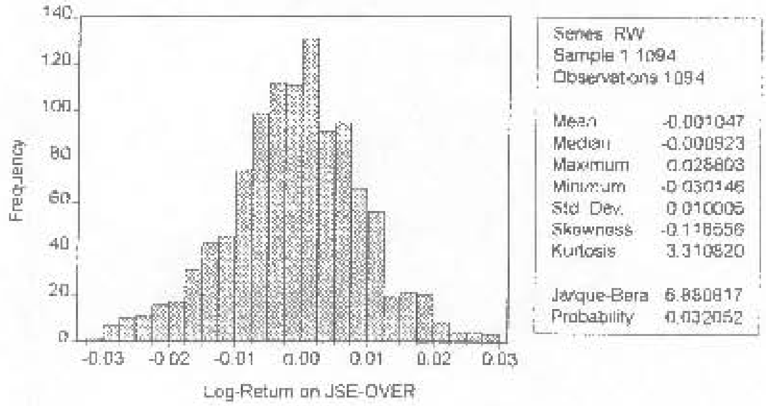

Figure 4-14: Histogram and Descriptive Statistics for RW log-Returns

From Figure 4-14 above, the following characteristics are noted:

1. The mean return is small but negative at -0.10% with a standard deviation of 1% over the sampled period.

2. The distribution is slightly skewed to the left and the excess kurtosis is almost zero. However, the Jarque-Bera statistic is still slightly significant at the 5% level. But, this distribution can be taken as sufficiently normal.

Figure 4-15 below summarises the results for log-returns in transition market state:

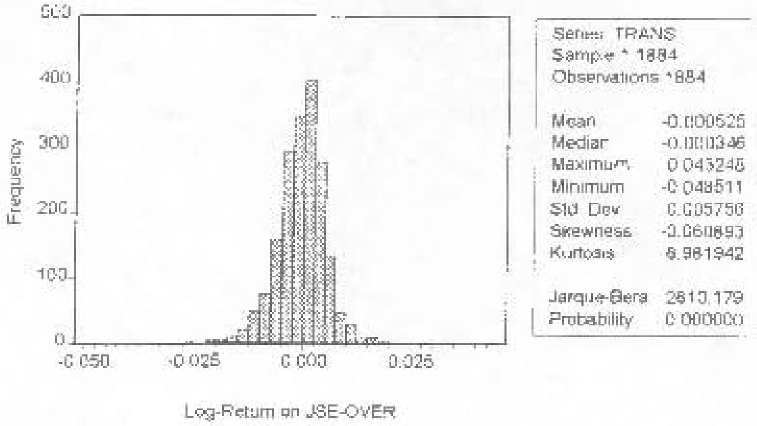

Figure 4-15: Histogram and Descriptive Statistics for TRANS log-returns

From Figure 4-15 above, the following characteristics are noticed:

1. The mean is very small but negative at -0.05% with a standard deviation of 0.56%.
2. The distribution is slightly skewed to the left.
3. The distribution is highly peaked at its mode, but the excess kurtosis is only 5.98. This is because those very few extreme returns are mostly balanced by the large number of small returns.

4.6 Fitting α Stable Distribution to Daily Return in Each Market State

4.6.1 A Review of Stable distributions

Definition 56 (Infinite Divisibility) *The characteristic function of a random variable ϕ is infinitely divisible if and only if there exists a characteristic function ϕ_n such that*

$$\phi = (\phi_n)^n, \ \forall \ integer \ n \geq 1.$$

Let $X_1, X_2, \ldots, X_n$ be n *i.i.d.* random variables each with characteristic function ϕ_X. If

$$Y_n = \sum_{j=1}^{n} X_j,$$

then

$$\phi_{Y_n} = (\phi_X)^n .$$

Thus, it is convenient to represent an infinitely divisible random variable in terms of equal-spaced increment:

$$X_t - X_s = \sum_{j=0}^{n-1} \left(X_{s+(j+1)\Delta_n} - X_{s+j\Delta_n} \right)$$

where

$$\Delta_n = \frac{t-s}{n}, \text{ for } n \geq 1, \, s < t.$$

Remark 57 *The well-known infinitely divisible classes are Gaussian, Poisson, Compound Poisson and stable distributions.*

Theorem 58 (Zolotarev) *The infinitely divisible distribution class is defined by the characteristic exponent of form:*

$$\psi(\theta) = i\theta\alpha - \beta\theta^2 + \int_{\{x \neq 0\}} \left[e^{i\theta x} - 1 - ix\sin(x) \right] d\Psi(x),$$

where $\alpha, \beta \geq 0$ *and* $\Psi(x)$ *is called the spectral function of the infinitely divisible distribution and well defined on real-line except at* $x = 0$. *Furthermore,* $\Psi(x)$ *has the following properties:*

1. *Non-decreasing on the semi-axis* $x > 0$ *and* $x < 0$,
2. $\Psi(x) \to 0$ *as* $|x| \to \infty$, *and*
3. $\int_{\{|x|<1\}} x^2 d\Psi(x) < \infty$.

Definition 59 (Stable Random Variable) *Let* X *be a random variable with cumulative distribution function* F_X. *If* $\forall n \in \mathbf{k}^+$ *and* $\forall$ *sets of i.i.d. random variables* $\{X_1, X_2, \cdots, X_n\}$ *each with the same distribution as* X, $\exists$ *constants* $a_n > 0, b_n$ *such that*

$$F_{X_1+X_2+\cdots+X_n} = F_{a_n X + b_n},$$

then X is stable (or have a stable distribution).

A stable distribution may also be defined in terms of its characteristic function, which Lévy (1924,1925) and Gnedenko and Kolmogorov (1954) have shown to be given by the following :

Definition 60 (Characteristic Function of Stable Distribution) *A random variable X have a stable distribution if its characteristic function has the following form:*

$$\begin{aligned}\phi_X(\theta) &= E\left[\exp(i\theta X)\right] \\ &= \begin{cases} exp\left\{i\delta\theta - |\gamma\theta|^{\alpha}\left[1 - i\beta sign(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right]\right\} & , \text{if } \alpha \neq 1 \\ exp\left\{i\delta\theta - |\gamma\theta|\left[1 + i\beta\frac{2}{\pi} sign(\theta)\ln|\theta|\right]\right\} & , \text{if } \alpha = 1 \end{cases}, \end{aligned} \quad (4.1)$$

where:

$$\begin{aligned} \alpha &\in (0,2] \text{ is the index of stability,} \\ \beta &\in [-1,1] \text{ is the skewness parameter,} \\ \gamma &\in (0,\infty) \text{ is the scale parameter,} \\ \delta &\in (-\infty,\infty) \text{ is the location parameter} \\ sign(\theta) &= \begin{cases} 1 & , \text{if } \theta > 0 \\ 0 & , \text{if } \theta = 0 \\ -1 & , \text{if } \theta < 0 \end{cases}, \text{ and} \\ i &= \sqrt{-1}. \end{aligned}$$

From the above definition of stable distribution, since (4.1) is characterised by the four parameters α, β, γ, and δ, thus Samorodnitsky and Taqqu (1994) introduced the following notation:

Notation 61 *If X is a stable distribution, then it is denoted by:*

$$X \sim S(x; \alpha, \beta, \gamma, \delta),$$

where:

$$\begin{aligned} \alpha &\in (0,2] \text{ is the index of stability,} \\ \beta &\in [-1,1] \text{ is the skewness parameter,} \\ \gamma &\in (0,\infty) \text{ is the scale parameter,} \\ \delta &\in (-\infty,\infty) \text{ is the location parameter} \end{aligned}$$

as for the definition.

Remark 62 *It follows that standardising X does not change its stable distribution, i.e.)*

$$S(x;\alpha,\beta,\gamma,\delta) = S\left(\frac{x-\delta}{\gamma};\alpha,\beta,1,0\right).$$

Remark 63 *The* standard *stable distribution is defined as:*

$$S_{\alpha\beta}(x) = S(x;\alpha,\beta,1,0).$$

As shown by Zolotarev (1986), the probability densities function of stable distribution always exist and are continuous, but they do not have closed form. Thus, implementing stable distribution for parametric statistical inference will always be difficult. However, the FORTRAN program, STABLE 3.04, written by John P. Nolan from the Math/Stat Department of American University in Washington, DC calculates the density (p.d.f.), cumulative distribution function (c.d.f.), and quantiles for a general stable distribution. For parametric statistical inference, the program also performs maximum likelihood estimation of stable parameters and some exploratory data analysis techniques for assessing the fit of a data set. The aim is to introduce the characteristic function form used by Nolan in the program.

However,.for the maximum likelihood estimations, Nolan (2001) suggests that a more useful parameterisation is a variation of the Zolotarev parameterisation:

Definition 64 *If* $X \sim S(0;\alpha,\beta,\gamma,\delta_0)$, *then the characteristic function of the stable distribu-*

tion is given by:

$$\phi_X(\theta) = E[\exp(i\theta X)]$$
$$= \begin{cases} exp\left\{i\delta_0\theta - |\gamma\theta|^\alpha\left[1 + i\beta sign(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\left((\gamma|\theta|)^{1-\alpha} - 1\right)\right]\right\} & , if\ \alpha \neq 1 \\ exp\left\{i\delta_0\theta - |\gamma\theta|\left[1 + i\beta\frac{2}{\pi}sign(\theta)(\ln|\theta| + \ln(\gamma))\right]\right\} & , if\ \alpha = 1 \end{cases}, \quad (4.2)$$

where:

$$\begin{aligned} \alpha &\in (0,2] \text{ is the index of stability,} \\ \beta &\in [-1,1] \text{ is the skewness parameter,} \\ \gamma &\in (0,\infty) \text{ is the scale parameter,} \\ \delta_0 &\in (-\infty,\infty) \text{ is the location parameter} \end{aligned}$$

According to Nolan (2001), the advantages of the $S(0;\alpha,\beta,\gamma,\delta_0)$ parameterisation are as follows:

1. The characteristic functions (hence the corresponding densities and distribution function) are jointly continuous in all four parameters.

2. α and β have a much clearer meaning as measures of the heaviness of the tails and skewness parameters. The parameters α, β and γ have the same interpretation for both parameterisations. But, the location parameters of the two parameterisation are related by:

$$\delta_1 = \begin{cases} \delta_0 - \beta\left(\tan\frac{\pi\alpha}{2}\right)\gamma & \text{, if } \alpha \neq 1 \\ \delta_0 - \beta\frac{2}{\pi}\gamma\ln\gamma & \text{, if } \alpha = 1 \end{cases}.$$

3. $S(0;\alpha,\beta,\gamma,\delta_0)$ is a location and scale family. If

$$Y \sim S(0;\alpha,\beta,\gamma,\delta_0),$$

then for any $a \neq 0$, b,

$$aY + b \sim S\left(0; \alpha, \left(sign\left(a\right)\right)\beta, \left|a\right|\gamma, a\delta_0 + b\right)$$

4. Reduces the correlation between the parameter estimates especially when α is near 1 (where there is a discontinuity).

4.6.2 Fitting α−Stable Distribution

The four parameters $\alpha, \beta, \gamma, \delta$ of the $S\left(0; \alpha, \beta, \gamma, \delta_0\right)$ is estimated using the STABLE 3.04 FORTRAN program . Table 4.6 below presents the estimated parameters of the α-stable distribution for partitioned daily return in each given market state:

State	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\gamma}$	$\widehat{\delta}$
RW	2.0000	−0.4885	0.707107×10^{-2}	-0.104704×10^{-2}
BULL	1.4000	1.0000	0.350323×10^{-2}	0.910219×10^{-2}
BEAR	1.1000	−1.0000	0.400129×10^{-2}	-0.907659×10^{-2}
CHAOS	1.6898	0.4001	0.963188×10^{-2}	0.945423×10^{-4}
TRANS	1.7935	−0.2474	0.345013×10^{-2}	-0.326662×10^{-3}

Table 4.6: Results of Statewise Fitting of Stable Distribution to Daily Returns

From Table 4.6 above, the following interpretations may be made:

1. The distribution of partitioned daily return in RW market state follows a normal distribution ($\widehat{\alpha} = 2$), but is positively skewed with the median greater than the mode ($\widehat{\beta} < 0$). It is shifted slightly to the left of zero ($\widehat{\delta} < 0$) and has a standard deviation $\sigma = \sqrt{2} \cdot \widehat{\gamma} = 0.01$.

2. The distribution of partitioned daily return in BULL market state has infinite variance ($\widehat{\alpha} < 2$) and is negatively skewed with the median less than the mode ($\widehat{\beta} > 0$). It is shifted slightly to the right of zero ($\widehat{\delta} > 0$).

3. The distribution of partitioned daily return in BEAR market state has infinite variance ($\widehat{\alpha} < 2$) and can almost be described by the Cauchy distribution ($\alpha = 1$). It is positively skewed with the median greater than the mode ($\widehat{\beta} < 0$), and is shifted slightly to the left of zero ($\widehat{\delta} < 0$).

4. The distribution of partitioned daily return in CHAOS market state has infinite variance ($\widehat{\alpha} < 2$) and is negatively skewed with the median less than the mode ($\widehat{\beta} > 0$). It is shifted slightly to the right of zero ($\widehat{\delta} > 0$).

5. The distribution of partitioned daily return in TRANS market state has infinite variance ($\widehat{\alpha} < 2$), but is positively skewed with the median greater than the mode ($\widehat{\beta} < 0$). It is shifted slightly to the left of zero ($\widehat{\delta} < 0$).

6. Stable distribution fitted to the BEAR daily returns have the thickest tail, because it has the smallest α.

4.6.3 Diagnostics for Assessing Stability Using the Stabilized Probability Plot

Now, to assess the goodness-of-fit of the four parameter stable distribution to the various sets of daily returns, Standard p-p plots tend to over-emphasise the behaviour of the distribution near the mode (where there are more variation) and under emphasis the tails. Michael (1983) defined a "stabilised" probability plots that eliminates this problem. The major advantage of the stabilised probability plot is that the variances of the plotted points are approximately equal, which enhances the interpretability of the plot. The feature of approximate equal variance lead to the definition of the goodness-of-fit statistics D_{SP} as the maximum deviation of the plotted points from its theoretical values. Acceptance regions for $Q-Q$, $P-P$ and other probability plots can be constructed using either the standard Kolmogorov-Smirnov Statistics D or D_{SP}, which helps to remove much of the subjectivity from the interpretation of these probability plots.

Let

$$r_{(1)} \leq \cdots \leq r_{(n)}$$

be the realisation of an ordered random sample of size n from the distribution $F(\bullet)$. A *quantile-quantile* (Q-Q) plot for a continuous hypothesised location-scale distribution

$$F_0\left[\frac{r-\mu}{\sigma}\right]$$

is constructed by plotting each sample quantile r_i versus a corresponding theoretical standard quantile

$$x_i = F_0^{-1}(t_i),$$

where t_i is an appropriate cumulative proportion, which is chosen to be:

$$t_i = \frac{i - \frac{1}{2}}{n}.$$

Similarly, a *percent-percent* (P-P) plot is constructed by plotting each transformed value

$$u_i = F_0\left(\frac{r_i - \mu}{\sigma}\right)$$

versus the uniform quantile t_i. If μ and σ are unknown, then they are replaced by their maximum likelihood estimates, as discussed by Wilk (1968).

In many situations, the problem with Q-Q plots is that certain points determined by $F(\bullet)$ are much more variable than other points. For the P-P plots, the opposite is true when $F_0 = F$, regardless of the form of $F(\bullet)$. To overcome this problem of unequal variability, a transformation to stabilise the variances of the plotted points is described in the following sub-section.

The stabilised probability plot

Definition 65 (Realisation of Uniform Order Statistics) *If the following parameters are known:*

1. $F(\bullet) = F_0(\bullet)$,
2. μ, *and*
3. σ,

then u_i can be regarded as the realisation of a uniform order statistics.

Definition 66 (Asymptotic Realisation of Uniform Order Statistics) *If the parameters are:*

1. *Unknown, but*

2. *Efficiently estimated,*

then u_i can be asymptotically regarded as the realisation of a uniform order statistics.

If u_i can be asymptotically regarded as the realisation of a uniform order statistics, then an *arcsine transformation* can be applied to stabilise the variance of an uniform order statistics as follows.

Suppose the random variable S is defined as a arcsine transformation of the uniform random variable by:

$$S = \frac{2}{\pi} \arcsin\left(\sqrt{U}\right),$$

where:

$$U \sim Uniform\,(0, 1).$$

Then, the probability density function of S is:

$$f_S(s) = \begin{cases} \frac{\pi}{2} \sin(\pi s) & ,\text{for } s \in [0,1] \\ 0 & ,\text{Otherwise.} \end{cases} \quad (4.3)$$

The distribution corresponding to (4.3) can be termed the *sine distribution*, since (4.3) is directly proportional to a half-cycle of a sine wave.

The sine distribution has the important property that its order statistics have the same asymptotic variance as stated in the following proposition:

Proposition 67 *If $S_{(1)} \leq \cdots \leq S_{(n)}$ is an ordered random sample from the sine distribution:*

$$f_S(s) = \begin{cases} \frac{\pi}{2} \sin(\pi s) & ,\textit{for } s \in [0,1] \\ 0 & ,\textit{Otherwise.} \end{cases},$$

then as $n \to \infty$ and $\frac{i}{n} \to p$, the asymptotic variance of $nS_{(i)}$ is:

$$Var\left[nS_{(i)}\right] = \frac{1}{\pi^2},$$

which is independent of p.

Based on the above proposition, the stabilised probability plot is then defined as follows:

Definition 68 (Stabilised Probability Plot) *For each $i = 1, \cdots, n$, the stabilised probability plot (SP) is defined as the plot of :*

$$s_i = \frac{2}{\pi} \arcsin\left(\sqrt{u_i}\right) \quad \textit{(Ordinate)}$$

versus

$$r_i = \frac{2}{\pi} \arcsin\left(\sqrt{t_i}\right) \quad \textit{(Abscissa).}$$

Remark 69 *If $F_0(\bullet) = F(\bullet), \mu, \sigma$ are known, then r_i is the mode of $S_{(i)}$.*

Table 4.7 below, summarises the formulae for the three distribution plots:

Plot Type	Abscissa	Ordinate
Q-Q	$x_i = F_0^{-1}\left(\frac{i-\frac{1}{2}}{n}\right)$	y_i
P-P	$t_i = \frac{i-\frac{1}{2}}{n}$	$u_i = F_0\left(\frac{y_i-\mu}{\sigma}\right)$
SP	$r_i = \frac{2}{\pi}\arcsin\left(\sqrt{\frac{i-\frac{1}{2}}{n}}\right)$	$\frac{2}{\pi}\arcsin\left(\sqrt{F_0\left(\frac{y_i-\mu}{\sigma}\right)}\right)$

Table 4.7: Formulae for abscissa and ordinate of different distribution plots

Goodness-of-Fit statistics for stabilised p-p plot

The standard Kolmogorov-Smirnov statistics D can be expressed as follows:

$$D = \max\left\{|t_i - u_i|\right\} + \frac{1}{2n}.$$

Analogous to D, we now define the goodness-of-fit statistics for stabilised p-p plot D_{SP} as follows:

$$D_{SP} = \max\left\{|r_i - s_i|\right\},$$

which can be used to test the null hypothesis:

$$H_0 : F(\bullet) = F_0(\bullet).$$

The exact critical points for D_{SP} for testing a simple hypothesis were computed using a recursive algorithm described by Noe (1972). However, due to time constraints, critical points was not able to be determined for sample size up to 1900 observations. Thus, no hypothesis tests on goodness-of-fit were performed.

4.6.4 Diagnostic for goodness-of-fit of stable distribution

Based on Michael (1983)'s idea, the STABLE 3.04 program produced p-p and stabilised p-p plots for the estimated parameters. Figure 4-16 below shows the pp-plots for the entire daily returns time series:

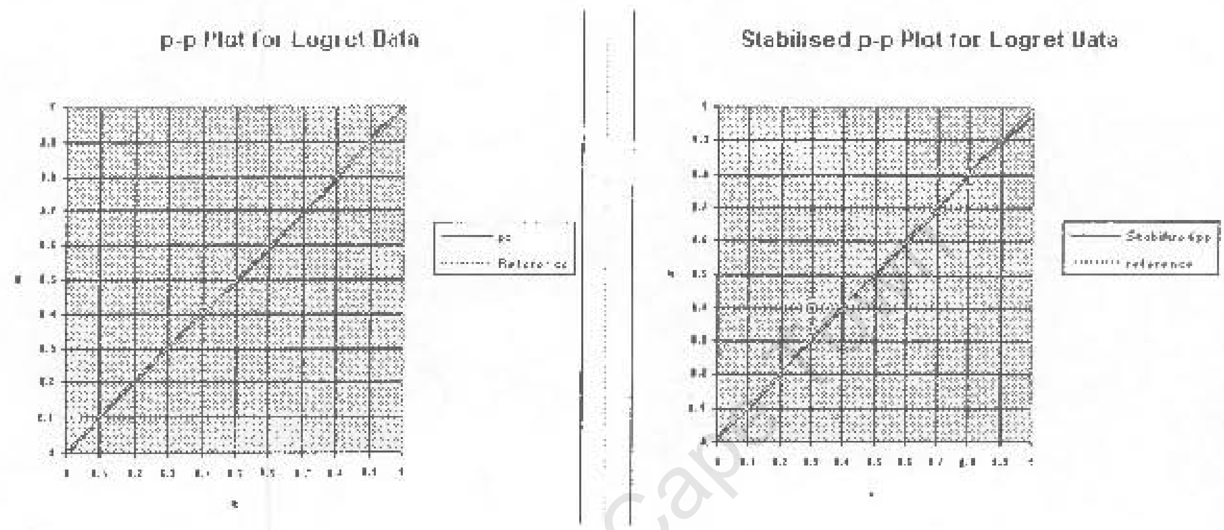

Figure 4-16: Probability Plots for the Entire Daily Returns Time Series

From Figure 4-16 above, it appears that the stable distribution have fitted all parts of the observed daily return distribution satisfactorily.

Figure 4-17 below shows the pp-plots for daily returns in the random walk (RW) market state:

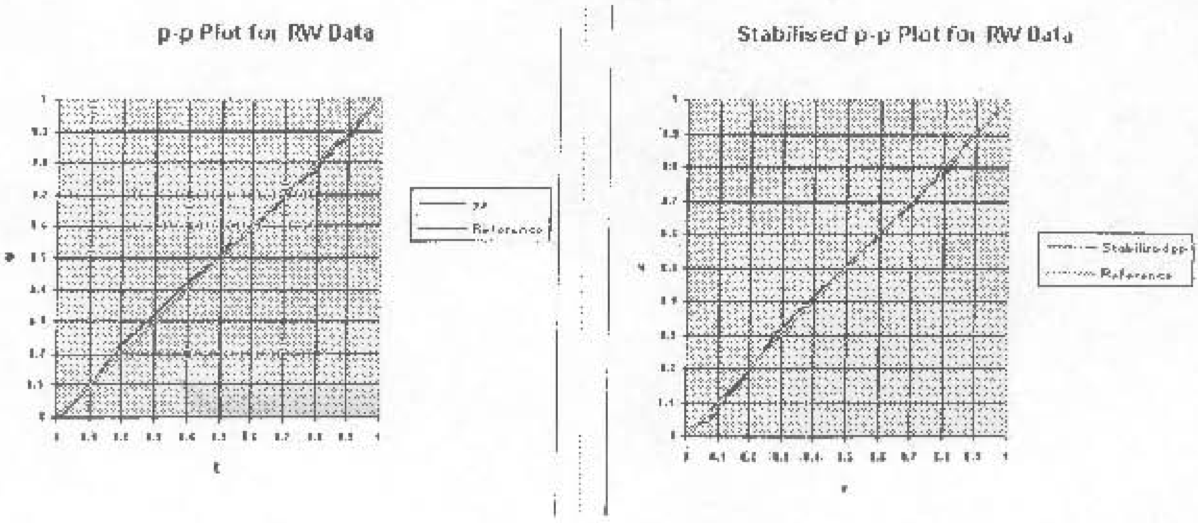

Figure 4-17: Probability Plots for Daily Returns in the Random Walk State

From Figure 4-17 above, there appears to be some oscillation of the p-p plots around the reference line. This means that he stable distribution did not fit the RW daily return distribution as well as it fitted the observed daily return distribution.

Figure 4-18 below shows the pp-plots for daily returns in the coherent bull (BULL) market state:

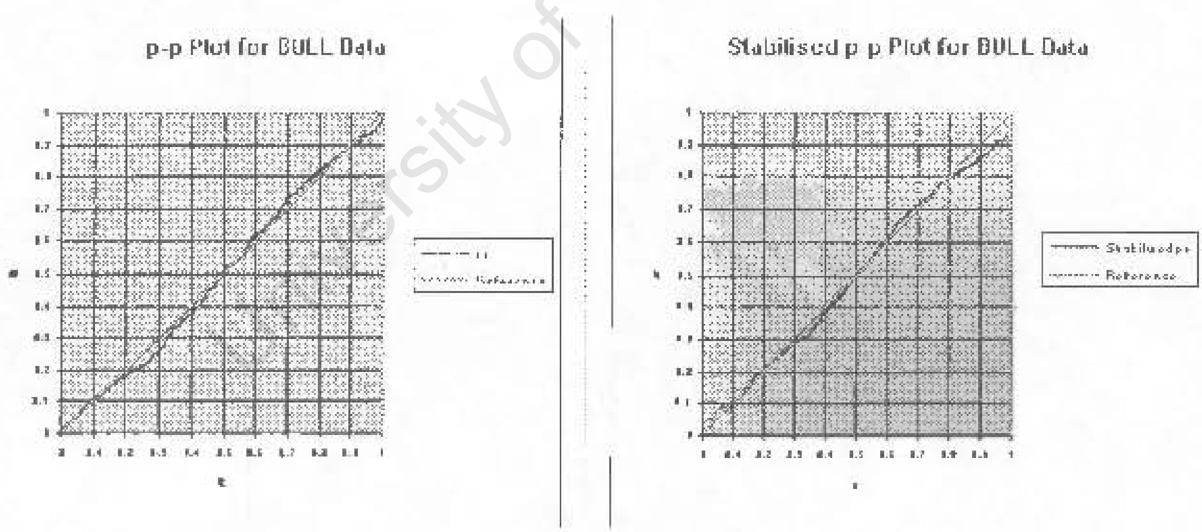

Figure 4-18: Probability Plots for Daily Returns in the Coherent Bull State

From Figure 4-18 above, there appears to be a clear systematic departure of the stabilised p-p plot from the reference line at the right end. This means that the right tail of the BULL daily return is not well fitted by the stable distribution. This is expected as distribution for BULL daily return is skewed to the right and have a very heavy right-tail.

Figure 4-19 below shows the pp-plots for daily returns in the coherent bear (BEAR) market state:

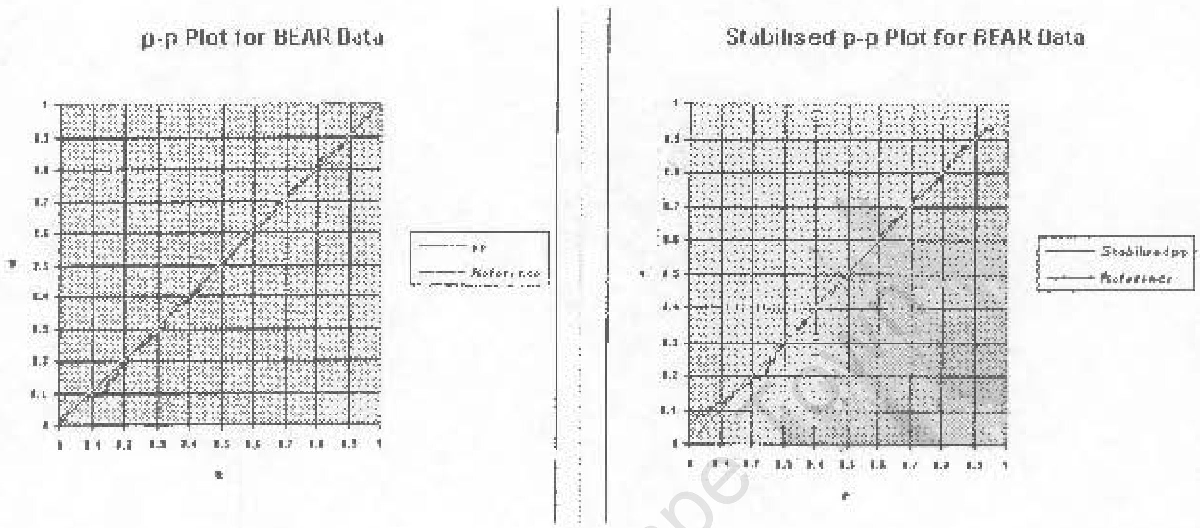

Figure 4-19: Probability Plots for Daily Returns in the Coherent Bear State

From Figure 4-19 above, there appears to be a clear systematic departure of the stabilised p-p plot from the reference line at the left end. This means that the left-tail of the BEAR daily return is not well fitted by the stable distribution. This is expected as distribution for BEAR daily return is skewed to the left and have a very heavy left-tail. This is exactly opposite to the BULL daily return as expected.

Figure 4-20 below shows the pp-plots for daily returns in the chaotic (CHAOS) market state:

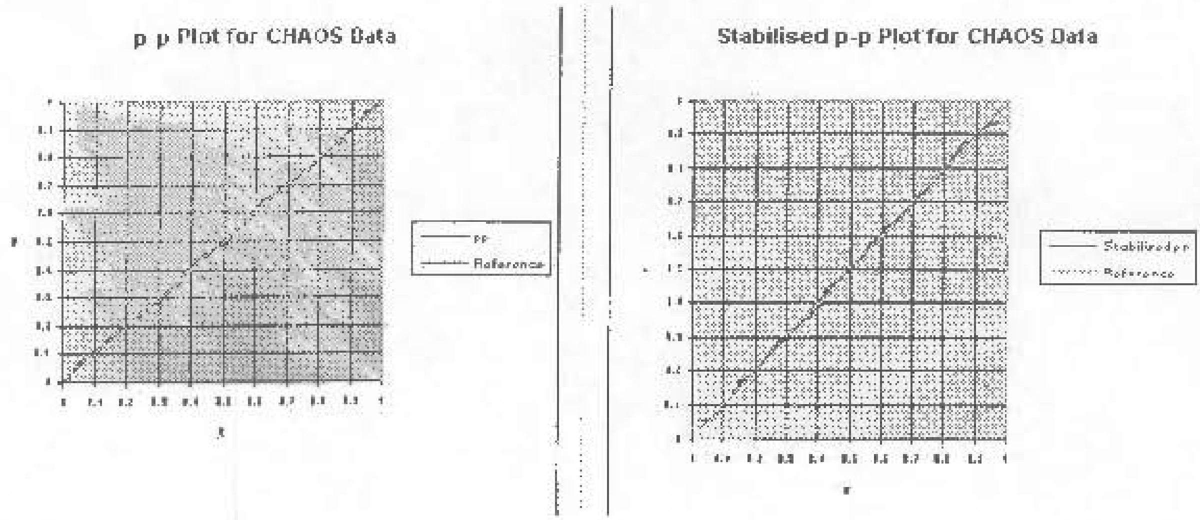

Figure 4-20: Probability Plots for Daily Returns in Chaotic State

From Figure 4-20 above, there appears to be a sufficiently significant departure of the stabilised p-p plot from its reference line. This means that both tails of the CHAOS daily return distribution is not well fitted by the stable distribution.

Figure 4-21 below shows the pp-plots for daily returns in the transition (TRANS) market state:

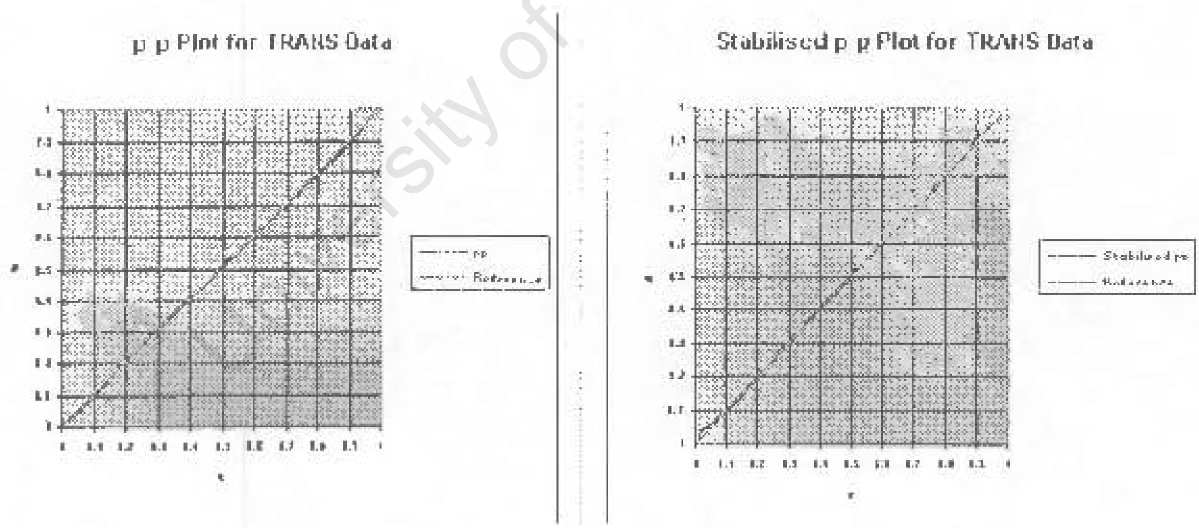

Figure 4-21: Probability Plots for Daily Returns in Transition State

From Figure 4-21 above, the stable distribution appears to have fitted the TRANS daily return distribution quite satisfactorily.

Chapter 5

PARAMETRIC STATISTICAL INFERENCE ON CMH DISTRIBUTION

The aim of this chapter is to carry out the fourth and fifth objective for this dissertation, namely:

Objective 4: Derive the related partial derivatives of the log-likelihood function, assuming the CMH daily return distributions, to prepare for the maximum likelihood estimation procedure.

Objective 5: Estimating CMH return distribution parameters for the Random Walk and Coherent Bull (Bear), Chaotic as well as the transitional states tentatively.

5.1 Derivation of Maximum Likelihood Estimation Procedure for the CMH Parameters

Recall that based on the *theory of social Imitation*, Vaga (1990)'s Coherent Market Hypothesis (CMH) models the annualised market return by a random variable, R, with p.d.f. (3.33) given below:

$$
\begin{aligned}
& f_R(r; n, \kappa, \rho) \\
= \quad & c(\kappa, \rho, n) \cdot \Psi^{-1}(r; \kappa, \rho, n) \cdot \exp\left[2 \int_{u=-\frac{1}{2}}^{r} \frac{\Pi(u; \kappa, \rho)}{\Psi(u; \kappa, \rho, n)} du\right],
\end{aligned}
$$

where:

1. The two control parameters are defined as:

$$
\begin{aligned}
\kappa &= \text{degree of market sentiment, and} \\
\rho &= \text{degree of fundamental bias.}
\end{aligned}
$$

2. The number of sub-systems making up the market (or degrees of freedom) is re-interpreted as:

   n = Number of industrial groups making up the financial market.

3. The drift coefficient function is defined as:

$$
\Pi(u; \kappa, \rho) = \sinh(\kappa u + \rho) - 2u \cosh(\kappa u + \rho).
$$

4. The diffusion coefficient function is defined as:

$$
\Psi(r; \kappa, \rho, n) = \frac{1}{n} \left[\cosh(\kappa u + \rho) - 2u \sinh(\kappa u + \rho)\right].
$$

5. The normalisation constant is defined as:

$$
c^{-1}(\kappa, \rho, n) = \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \Psi^{-1}(r; \kappa, \rho) \exp\left[2 \int_{u=-\frac{1}{2}}^{r} \frac{\Pi(u; \kappa, \rho)}{\Psi(u; \kappa, \rho, n)} du\right] dr.
$$

Appendix C contains the derivation, by first principle, of the maximum likelihood estimation procedure for the parameters of the distribution based on (3.33). It is noticed that the form of the CMH returns distribution, as given by (3.33), have lead to some unnecessary complications in the derivation and the maximum likelihood estimation procedure so derived cannot be coded. Hence, it is necessary to re-state the CMH p.d.f. in a mathematically more convenient form.

Firstly, notice that there exists some intrinsic relations between the diffusion $\Psi(u;\kappa,\rho)$ and drift $\Pi(u;\kappa,\rho)$ coefficient functions , which may greatly simplify the derivation of the log-likelihood function and its associated partial derivatives. Thus, define the following new function:

$$\Phi(u;\kappa,\rho) \equiv \cosh(\kappa u+\rho) - 2u\sinh(\kappa u+\rho). \tag{5.1}$$

Then, the diffusion coefficient function, $\Psi(r;\kappa,\rho,n)$, can be re-stated as:

$$\Psi(r;\kappa,\rho,n) \equiv \frac{1}{n}\Phi(u;\kappa,\rho). \tag{5.2}$$

It follows that Vaga's CMH distribution can be re-defined in terms of $\Pi(u;\kappa,\rho)$ and $\Phi(u;\kappa,\rho)$ functions as follows:

$$\begin{aligned}
& f_R(r;n,\kappa,\rho) \\
= \; & c(\kappa,\rho,n)\cdot\Psi^{-1}(r;\kappa,\rho,n)\cdot\exp\left[2n\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(u;\kappa,\rho)}{\Phi(u;\kappa,\rho)}du\right] \\
= \; & n\cdot C^{-1}(\kappa,\rho,n)\cdot\Phi^{-1}(r;\kappa,\rho)\cdot\exp\left[2nB(r;\kappa,\rho)\right],
\end{aligned} \tag{5.3}$$

where

$$\begin{aligned}
\Pi(u;\kappa,\rho) &= \sinh(\kappa u+\rho) - 2u\cosh(\kappa u+\rho) \\
\Phi(u;\kappa,\rho) &= \cosh(\kappa u+\rho) - 2u\sinh(\kappa u+\rho) \\
A(u;\kappa,\rho) &= \frac{\Pi(u;\kappa,\rho)}{\Phi(u;\kappa,\rho)} \\
B(r;\kappa,\rho) &= \int\limits_{-\frac{1}{2}}^{r}\frac{\Pi(u;\kappa,\rho)}{\Phi(u;\kappa,\rho)}du = \int\limits_{-\frac{1}{2}}^{r}A(u;\kappa,\rho)\,du \\
C(\kappa,\rho,n) &= c^{-1}(\kappa,\rho,n) = n\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\Phi^{-1}(r;\kappa,\rho)e^{2nB(r;\kappa,\rho)}dr
\end{aligned}$$

Now, derive the first and second order partial derivatives of $\Pi(u;\kappa,\rho)$ and $\Phi(u;\kappa,\rho)$ with respect to κ and ρ to prepare for the derivation of log-likelihood function for (5.3). The first order partial derivatives of $\Pi(u;\kappa,\rho)$ are given by:

$$\begin{aligned}
\frac{\partial}{\partial\kappa}\Pi(u;\kappa,\rho) &= \frac{\partial}{\partial\kappa}\sinh(\kappa u+\rho) - 2u\frac{\partial}{\partial\kappa}\cosh(\kappa u+\rho) \\
&= u\cosh(\kappa u+\rho) - 2u^2\sinh(\kappa u+\rho) \\
&= u\cdot\Phi(u;\kappa,\rho) \\
\frac{\partial}{\partial\rho}\Pi(u;\kappa,\rho) &= \frac{\partial}{\partial\rho}\sinh(\kappa u+\rho) - 2u\frac{\partial}{\partial\rho}\cosh(\kappa u+\rho) \\
&= \cosh(\kappa u+\rho) - 2u\sinh(\kappa u+\rho) \\
&= \Phi(u;\kappa,\rho)
\end{aligned}$$

and the first order partial derivatives of $\Phi(u;\kappa,\rho)$ are given by:

$$\begin{aligned}
\frac{\partial}{\partial\kappa}\Phi(u;\kappa,\rho) &= \frac{\partial}{\partial\kappa}\cosh(\kappa u+\rho) - 2u\frac{\partial}{\partial\kappa}\sinh(\kappa u+\rho) \\
&= u\left(\sinh(\kappa u+\rho) - 2u\cosh(\kappa u+\rho)\right) \\
&= u\cdot\Pi(u;\kappa,\rho) \\
\frac{\partial}{\partial\rho}\Phi(u;\kappa,\rho) &= \frac{\partial}{\partial\rho}\cosh(\kappa u+\rho) - 2u\frac{\partial}{\partial\rho}\sinh(\kappa u+\rho) \\
&= \sinh(\kappa u+\rho) - 2u\cosh(\kappa u+\rho) \\
&= \Pi(u;\kappa,\rho)
\end{aligned}$$

Based on the above first order partial derivatives, the second order partial derivatives of $\Pi(u;\kappa,\rho)$ and $\Phi(u;\kappa,\rho)$ with respect to κ and ρ can be derived. The second order partial derivatives of $\Pi(u;\kappa,\rho)$ are given by:

$$\begin{aligned}
\frac{\partial^2}{\partial\kappa^2}\Pi(u;\kappa,\rho) &= u\frac{\partial}{\partial\kappa}\Phi(u;\kappa,\rho) = u^2\cdot\Pi(u;\kappa,\rho) \\
\frac{\partial^2}{\partial\kappa\partial\rho}\Pi(u;\kappa,\rho) &= \frac{\partial}{\partial\rho}\left[u\Phi(u;\kappa,\rho)\right] = u\cdot\Pi(u;\kappa,\rho) \\
\frac{\partial^2}{\partial\rho^2}\Pi(u;\kappa,\rho) &= \frac{\partial}{\partial\rho}\Phi(u;\kappa,\rho) = \Pi(u;\kappa,\rho)
\end{aligned}$$

and the second order partial derivatives of $\Phi(u;\kappa,\rho)$ is given by:

$$\begin{aligned}
\frac{\partial^2}{\partial\kappa^2}\Phi(u;\kappa,\rho) &= u\frac{\partial}{\partial\kappa}\Pi(u;\kappa,\rho) = u^2\cdot\Phi(u;\kappa,\rho) \\
\frac{\partial^2}{\partial\kappa\partial\rho}\Phi(u;\kappa,\rho) &= u\frac{\partial}{\partial\rho}\Pi(u;\kappa,\rho) = u\cdot\Phi(u;\kappa,\rho) \\
\frac{\partial^2}{\partial\rho^2}\Phi(u;\kappa,\rho) &= \frac{\partial}{\partial\rho}\Pi(u;\kappa,\rho) = \Phi(u;\kappa,\rho)
\end{aligned}$$

Given the first and second order partial derivatives of $\Pi(u;\kappa,\rho)$ and $\Phi(u;\kappa,\rho)$ derived above, the derivation of the log-likelihood function and its associated partial derivatives may now commence. From the maximum likelihood large samples theory, the likelihood function for the re-stated CMH distribution (5.3) is given by:

$$\begin{aligned}
L(\kappa,\rho|\,\mathbf{r}) &= \prod_{i=1}^{m}[f_R(r_i;n,\kappa,\rho)] \\
&= \prod_{i=1}^{m}\left[nC^{-1}(\kappa,\rho,n)\,\Phi^{-1}(r_i;\kappa,\rho)\exp\left[2nB(r_i;\kappa,\rho)\right]\right] \\
&= n^m C^{-m}(\kappa,\rho,n)\prod_{i=1}^{m}\left[\Phi^{-1}(r_i;\kappa,\rho)\exp\left[2nB(r_i;\kappa,\rho)\right]\right]
\end{aligned}$$

However, it is easier to maximise the natural logarithm of the likelihood function. Hence, taking the natural logarithm of the likelihood function, the log-likelihood function for the re-stated CMH distribution (5.3) is given by:

$$l(\kappa,\rho|\,\mathbf{r}) = m\ln n - m\ln C(\kappa,\rho,n) - \sum_{i=1}^{m}\ln\Phi(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m}B(r_i;\kappa,\rho). \tag{5.4}$$

Now, before attempting the derivations of the first and second order partial derivatives of (5.4) required for the maximum likelihood estimation procedure, notice that first and second order partial derivatives of $A(u;\kappa,\rho)$, $B(r_i;\kappa,\rho)$, and $C(\kappa,\rho,n)$ will be required in several parts of the derivation. These needs to be first worked out to keep the derivation of the partial derivatives of $l(\kappa,\rho|\,\mathbf{r})$ manageable and efficient.

The partial derivatives of $A(u;\kappa,\rho)$ needs to be first derived. The first order partial derivatives of $A(u;\kappa,\rho)$ are given by:

$$
\begin{aligned}
\frac{\partial}{\partial\kappa}A(u;\kappa,\rho) &= \frac{\partial}{\partial\kappa}\left[\frac{\Pi(u;\kappa,\rho)}{\Phi(u;\kappa,\rho)}\right] \\
&= \frac{\Phi(u;\kappa,\rho)\frac{\partial}{\partial\kappa}\Pi(u;\kappa,\rho) - \Pi(u;\kappa,\rho)\frac{\partial}{\partial\kappa}\Phi(u;\kappa,\rho)}{\Phi^2(u;\kappa,\rho)} \\
&= \frac{\Phi(u;\kappa,\rho)u\Phi(u;\kappa,\rho) - \Pi(u;\kappa,\rho)u\Pi(u;\kappa,\rho)}{\Phi^2(u;\kappa,\rho)} \\
&= u\frac{\Phi^2(u;\kappa,\rho) - \Pi^2(u;\kappa,\rho)}{\Phi^2(u;\kappa,\rho)} \\
&= \frac{u(1-4u^2)}{\Phi^2(u;\kappa,\rho)},
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial}{\partial\rho}A(u;\kappa,\rho) &= \frac{\partial}{\partial\rho}\left[\frac{\Pi(u;\kappa,\rho)}{\Phi(u;\kappa,\rho)}\right] \\
&= \frac{\Phi(u;\kappa,\rho)\frac{\partial}{\partial\rho}\Pi(u;\kappa,\rho) - \Pi(u;\kappa,\rho)\frac{\partial}{\partial\rho}\Phi(u;\kappa,\rho)}{\Phi^2(u;\kappa,\rho)} \\
&= \frac{\Phi(u;\kappa,\rho)\Phi(u;\kappa,\rho) - \Pi(u;\kappa,\rho)\Pi(u;\kappa,\rho)}{\Phi^2(u;\kappa,\rho)} \\
&= \frac{\Phi^2(u;\kappa,\rho) - \Pi^2(u;\kappa,\rho)}{\Phi^2(u;\kappa,\rho)} \\
&= \frac{1-4u^2}{\Phi^2(u;\kappa,\rho)}
\end{aligned}
$$

Definition 70 *Denote the two first order partial derivatives of $A(u;\kappa,\rho)$ as*

$$
\begin{aligned}
A1(u;\kappa,\rho) &= \frac{\partial}{\partial\kappa}A(u;\kappa,\rho) = \frac{u(1-4u^2)}{\Phi^2(u;\kappa,\rho)} \\
A2(u;\kappa,\rho) &= \frac{\partial}{\partial\rho}A(u;\kappa,\rho) = \frac{1-4u^2}{\Phi^2(u;\kappa,\rho)}
\end{aligned}
$$

Remark 71 *From the above definition, notice that a simple relationship between $A1(u;\kappa,\rho)$ and $A2(u;\kappa,\rho)$ is given below:*

$$A1(u;\kappa,\rho) = uA2(u;\kappa,\rho).$$

For the derivations of the second order partial derivatives of $A(u;\kappa,\rho)$, the function defined below is required in a few parts of the derivation:

Definition 72

$$\begin{aligned} D(u;\kappa,\rho) &= A(u;\kappa,\rho)A2(u;\kappa,\rho) \\ &= \frac{\Pi(u;\kappa,\rho)}{\Phi(u;\kappa,\rho)} \times \frac{1-4u^2}{\Phi^2(u;\kappa,\rho)} \\ &= \frac{\left(1-4u^2\right)\Pi(u;\kappa,\rho)}{\Phi^3(u;\kappa,\rho)} \end{aligned}$$

Then, the second order partial derivatives of $A(u;\kappa,\rho)$ with respect to to κ and ρ are given as follows:

$$\begin{aligned} \frac{\partial^2}{\partial\kappa^2}A(u;\kappa,\rho) &= \frac{\partial}{\partial\kappa}\left[\frac{u(1-4u^2)}{\Phi^2(u;\kappa,\rho)}\right] \\ &= \frac{-2u(1-4u^2)\frac{\partial}{\partial\kappa}\Phi(u;\kappa,\rho)}{\Phi^3(u;\kappa,\rho)} \\ &= \frac{-2u(1-4u^2)u\Pi(u;\kappa,\rho)}{\Phi^3(u;\kappa,\rho)} \\ &= \frac{-2u^2(1-4u^2)A(u;\kappa,\rho)}{\Phi^2(u;\kappa,\rho)} \\ &= -2u^2D(u;\kappa,\rho) \end{aligned}$$

$$\begin{aligned} \frac{\partial^2}{\partial\rho^2}A(u;\kappa,\rho) &= \frac{\partial}{\partial\rho}\left[\frac{1-4u^2}{\Phi^2(u;\kappa,\rho)}\right] \\ &= \frac{-2(1-4u^2)\frac{\partial}{\partial\rho}\Phi(u;\kappa,\rho)}{\Phi^3(u;\kappa,\rho)} \\ &= \frac{-2(1-4u^2)\Pi(u;\kappa,\rho)}{\Phi^3(u;\kappa,\rho)} \\ &= \frac{-2(1-4u^2)A(u;\kappa,\rho)}{\Phi^2(u;\kappa,\rho)} \\ &= -2D(u;\kappa,\rho) \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial\kappa\partial\rho}A(u;\kappa,\rho) &= \frac{\partial}{\partial\rho}\left[\frac{u(1-4u^2)}{\Phi^2(u;\kappa,\rho)}\right] \\
&= \frac{-2u(1-4u^2)\frac{\partial}{\partial\rho}\Phi(u;\kappa,\rho)}{\Phi^3(u;\kappa,\rho)} \\
&= \frac{-2u(1-4u^2)A(u;\kappa,\rho)}{\Phi^2(u;\kappa,\rho)} \\
&= -2uD(u;\kappa,\rho)
\end{aligned}$$

Definition 73 *Denote the three second order partial derivatives of $A(u;\kappa,\rho)$ as:*

$$\begin{aligned}
A11\,(u;\kappa,\rho) &= \frac{\partial^2}{\partial\kappa^2}A(u;\kappa,\rho) = -2u^2D(u;\kappa,\rho) \\
A12\,(u;\kappa,\rho) &= \frac{\partial^2}{\partial\kappa\partial\rho}A\,(u;\kappa,\rho) = -2uD(u;\kappa,\rho) \\
A22\,(u;\kappa,\rho) &= \frac{\partial^2}{\partial\rho^2}A(u;\kappa,\rho) = -2D(u;\kappa,\rho)
\end{aligned}$$

Next, the partial derivatives of $B(r_i;\kappa,\rho)$ needs to be worked out Recall:

$$B(r_i;\kappa,\rho) = \int_{u=-\frac{1}{2}}^{r_i} A\,(u;\kappa,\rho)\,du$$

Definition 74 *Denote the partial derivatives of $B(r_i;\kappa,\rho)$ as follows:*

$$\begin{aligned}
B1(r_i;\kappa,\rho) &= \frac{\partial}{\partial\kappa}B(r_i;\kappa,\rho) \\
B2(r_i;\kappa,\rho) &= \frac{\partial}{\partial\rho}B(r_i;\kappa,\rho) \\
B11(r_i;\kappa,\rho) &= \frac{\partial^2}{\partial\kappa^2}B(r_i;\kappa,\rho) \\
B12(r_i;\kappa,\rho) &= \frac{\partial^2}{\partial\kappa\partial\rho}B(r_i;\kappa,\rho) \\
B22(r_i;\kappa,\rho) &= \frac{\partial^2}{\partial\rho^2}B(r_i;\kappa,\rho)
\end{aligned}$$

Then, the first order partial derivatives of $B(r_i;\kappa,\rho)$ with respect to to κ and ρ are given by:

$$\begin{aligned}
\frac{\partial}{\partial\kappa}B(r_i;\kappa,\rho) &= \int_{-\frac{1}{2}}^{r_i}\frac{\partial}{\partial\kappa}A\left(u;\kappa,\rho\right)du = \int_{-\frac{1}{2}}^{r_i} uA2(u;\kappa,\rho)du \\
\frac{\partial}{\partial\rho}B(r_i;\kappa,\rho) &= \int_{-\frac{1}{2}}^{r_i}\frac{\partial}{\partial\rho}A\left(u;\kappa,\rho\right)du = \int_{-\frac{1}{2}}^{r_i} A2(u;\kappa,\rho)du
\end{aligned}$$

and the second order partial derivatives of $B(r_i;\kappa,\rho)$ with respect to to κ and ρ are given by:

$$\begin{aligned}
\frac{\partial^2}{\partial\kappa^2}B(r_i;\kappa,\rho) &= \int_{-\frac{1}{2}}^{r_i}\frac{\partial^2}{\partial\kappa^2}A\left(u;\kappa,\rho\right)du \\
&= \int_{-\frac{1}{2}}^{r_i} A11\left(u;\kappa,\rho\right)du \\
&= -2\int_{-\frac{1}{2}}^{r_i} u^2 D(u;\kappa,\rho)du
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial\kappa\partial\rho}B(r_i;\kappa,\rho) &= \int_{-\frac{1}{2}}^{r_i}\frac{\partial^2}{\partial\kappa\partial\rho}A\left(u;\kappa,\rho\right)du \\
&= \int_{-\frac{1}{2}}^{r_i} A12\left(u;\kappa,\rho\right)du \\
&= -2\int_{-\frac{1}{2}}^{r_i} uD(u;\kappa,\rho)du
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2}{\partial\rho^2}B(r_i;\kappa,\rho) &= \int_{-\frac{1}{2}}^{r_i} \frac{\partial^2}{\partial\rho^2}A(u;\kappa,\rho)\,du \\
&= \int_{-\frac{1}{2}}^{r_i} A22(u;\kappa,\rho)\,du \\
&= -2\int_{-\frac{1}{2}}^{r_i} D(u;\kappa,\rho)du
\end{aligned}
$$

Finally, the derivations of the partial derivatives of $C(\kappa,\rho,n)$ with respect κ and ρ requires some care and organisation. Recall:

$$
C(\kappa,\rho,n) = n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \Phi^{-1}(r;\kappa,\rho)e^{2nB(r;\kappa,\rho)}dr
$$

Definition 75 *Denote the partial derivatives of $C(\kappa,\rho,n)$ as follows:*

$$
\begin{aligned}
C1(\kappa,\rho,n) &= \frac{\partial}{\partial\kappa}C(\kappa,\rho,n) \\
C2(\kappa,\rho,n) &= \frac{\partial}{\partial\rho}C(\kappa,\rho,n) \\
C11(\kappa,\rho,n) &= \frac{\partial^2}{\partial\kappa^2}C(\kappa,\rho,n) \\
C12(\kappa,\rho,n) &= \frac{\partial^2}{\partial\kappa\partial\rho}C(\kappa,\rho,n) \\
C22(\kappa,\rho,n) &= \frac{\partial^2}{\partial\rho^2}C(\kappa,\rho,n)
\end{aligned}
$$

Then, first order partial derivatives of $C(\kappa, \rho, n)$ with respect κ and ρ are given by:

$$\begin{aligned}
\frac{\partial}{\partial\kappa}C(\kappa,\rho,n) &= n\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial}{\partial\kappa}\left[\Phi^{-1}(r;\kappa,\rho)e^{2nB(r;\kappa,\rho)}\right]dr \\
&= n\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left[\frac{\partial}{\partial\kappa}\left[\Phi^{-1}(r;\kappa,\rho)\right]e^{2nB(r;\kappa,\rho)} + \Phi^{-1}(r;\kappa,\rho)\frac{\partial}{\partial\kappa}e^{2nB(r;\kappa,\rho)}\right]dr \\
&= n\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left(-\Phi^{-2}(r;\kappa,\rho\right)\frac{\partial}{\partial\kappa}\left[\Phi(r;\kappa,\rho)\right]e^{2nB(r;\kappa,\rho)}dr \\
&\quad + n\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \Phi^{-1}(r;\kappa,\rho)e^{2nB(r;\kappa,\rho)}\frac{\partial}{\partial\kappa}2nB\left(r;\kappa,\rho\right)dr \\
&= n\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[\frac{-\frac{\partial}{\partial\kappa}\left[\Phi(r;\kappa,\rho)\right]}{\Phi(r;\kappa,\rho)} + 2n\frac{\partial}{\partial\kappa}B\left(r;\kappa,\rho\right)\right]dr \\
&= n\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[\frac{-r\Pi(r;\kappa,\rho)}{\Phi(r;\kappa,\rho)} + 2n\frac{\partial}{\partial\kappa}B\left(r;\kappa,\rho\right)\right]dr \\
&= n\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB1\left(r;\kappa,\rho\right) - rA\left(r;\kappa,\rho\right)\right]dr
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial}{\partial\rho}C(\kappa,\rho,n) &= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial}{\partial\rho}\left[\Phi^{-1}(r;\kappa,\rho)e^{2nB(r;\kappa,\rho)}\right]dr \\
&= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \left[\frac{\partial}{\partial\rho}\left[\Phi^{-1}(r;\kappa,\rho)\right]e^{2nB(r;\kappa,\rho)} + \Phi^{-1}(r;\kappa,\rho)\frac{\partial}{\partial\rho}e^{2nB(r;\kappa,\rho)}\right]dr \\
&= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} -\Phi^{-2}(r;\kappa,\rho)\frac{\partial}{\partial\rho}\Phi(r;\kappa,\rho)e^{2nB(r;\kappa,\rho)}dr \\
&\quad +n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \Phi^{-1}(r;\kappa,\rho)e^{2nB(r;\kappa,\rho)}\frac{\partial}{\partial\rho}2nB\left(r;\kappa,\rho\right)dr \\
&= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[\frac{-\frac{\partial}{\partial\rho}\left[\Phi(r;\kappa,\rho)\right]}{\Phi(r;\kappa,\rho)} + 2n\frac{\partial}{\partial\rho}B\left(r;\kappa,\rho\right)\right]dr \\
&= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[\frac{-\Pi(r;\kappa,\rho)}{\Phi(r;\kappa,\rho)} + 2n\frac{\partial}{\partial\rho}B\left(r;\kappa,\rho\right)\right]dr \\
&= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB2\left(r;\kappa,\rho\right) - A\left(r;\kappa,\rho\right)\right]dr
\end{aligned}
$$

and the second order partial derivatives of $C(\kappa, \rho, n)$ with respect κ and ρ are given by:

$$\begin{aligned}
\frac{\partial^2}{\partial \kappa^2} C(\kappa, \rho, n) &= n \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial}{\partial \kappa} \left\{ \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)} \left[2nB1\left(r;\kappa,\rho\right) - rA\left(r;\kappa,\rho\right)\right] \right\} dr \\
&= n \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial}{\partial \kappa} \left\{ \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)} \right\} \left[2nB1\left(r;\kappa,\rho\right) - rA\left(r;\kappa,\rho\right)\right] dr \\
&\quad + n \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)} \frac{\partial}{\partial \kappa} \left\{ \left[2nB1\left(r;\kappa,\rho\right) - rA\left(r;\kappa,\rho\right)\right] \right\} dr \\
&= n \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)} \left[2nB1\left(r;\kappa,\rho\right) - rA\left(r;\kappa,\rho\right)\right]^2 dr \\
&\quad + n \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)} \left[2nB11\left(r;\kappa,\rho\right) - rA1\left(r;\kappa,\rho\right)\right] dr
\end{aligned}$$

$$
\begin{aligned}
& \frac{\partial^2}{\partial\kappa\partial\rho}C(\kappa,\rho,n) \\
= \; & n \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial}{\partial\rho}\left\{\frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB1\left(r;\kappa,\rho\right)-rA\left(r;\kappa,\rho\right)\right]\right\}dr \\
= \; & n \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial}{\partial\rho}\left\{\frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\right\}\left[2nB1\left(r;\kappa,\rho\right)-rA\left(r;\kappa,\rho\right)\right]dr \\
& +n \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\frac{\partial}{\partial\rho}\left\{\left[2nB1\left(r;\kappa,\rho\right)-rA\left(r;\kappa,\rho\right)\right]\right\}dr \\
= \; & n \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB2\left(r;\kappa,\rho\right)-A\left(r;\kappa,\rho\right)\right]\left[2nB1\left(r;\kappa,\rho\right)-rA\left(r;\kappa,\rho\right)\right]dr \\
& +n \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB12\left(r;\kappa,\rho\right)-rA2\left(r;\kappa,\rho\right)\right]dr
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial^2}{\partial\rho^2}C(\kappa,\rho,n) &= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial}{\partial\rho}\left\{\frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB2\left(r;\kappa,\rho\right)-A\left(r;\kappa,\rho\right)\right]\right\}dr \\
&= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial}{\partial\rho}\left\{\frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\right\}\left[2nB2\left(r;\kappa,\rho\right)-A\left(r;\kappa,\rho\right)\right]dr \\
&\quad +n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\frac{\partial}{\partial\rho}\left\{\left[2nB2\left(r;\kappa,\rho\right)-A\left(r;\kappa,\rho\right)\right]\right\}dr \\
&= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB2\left(r;\kappa,\rho\right)-A\left(r;\kappa,\rho\right)\right]^2dr \\
&\quad +n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB22\left(r;\kappa,\rho\right)-A2\left(r;\kappa,\rho\right)\right]dr
\end{aligned}
$$

Now, derivation for partial derivatives of the log-likelihood function with respect to κ and ρ may commence. Recall that the log-likelihood function for the re-stated CMH distribution (5.3) is given by (5.4) as:

$$
l(\kappa,\rho|\,\mathbf{r}) = m\ln n - m\ln C(\kappa,\rho,n) - \sum_{i=1}^{m}\ln\Phi(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m}B(r_i;\kappa,\rho).
$$

Definition 76 *Denote the partial derivatives of $l(\kappa,\rho|\,\mathbf{r})$ as follows:*

$$
\begin{aligned}
l1(\kappa,\rho) &= \frac{\partial}{\partial\kappa}l(\kappa,\rho|\,\mathbf{r}) \\
l2(\kappa,\rho) &= \frac{\partial}{\partial\rho}l(\kappa,\rho|\,\mathbf{r}) \\
l11(\kappa,\rho) &= \frac{\partial^2}{\partial\kappa^2}l(\kappa,\rho|\,\mathbf{r}) \\
l12(\kappa,\rho) &= \frac{\partial^2}{\partial\kappa\partial\rho}l(\kappa,\rho|\,\mathbf{r}) \\
l22(\kappa,\rho) &= \frac{\partial^2}{\partial\rho^2}l(\kappa,\rho|\,\mathbf{r})
\end{aligned}
$$

Then, the first order partial derivatives of $l(\kappa,\rho)$ with respect κ and ρ are given by:

$$
\begin{aligned}
l1(\kappa,\rho) &= \frac{\partial}{\partial\kappa} l(\kappa,\rho|\,\mathbf{r}) \\
&= \frac{\partial}{\partial\kappa}\left\{-m\ln C(\kappa,\rho,n) - \sum_{i=1}^{m}\ln\Phi(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m}B(r_i;\kappa,\rho)\right\} \\
&= \frac{-mC1(\kappa,\rho,n)}{C(\kappa,\rho,n)} - \sum_{i=1}^{m}\frac{r_i\Pi(r_i;\kappa,\rho)}{\Phi(r_i;\kappa,\rho)} + 2n\sum_{i=1}^{m}B1(r_i;\kappa,\rho) \\
&= \frac{-mC1(\kappa,\rho,n)}{C(\kappa,\rho,n)} - \sum_{i=1}^{m}r_iA(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m}B1(r_i;\kappa,\rho) \\
l2(\kappa,\rho) &= \frac{\partial}{\partial\rho} l(\kappa,\rho|\,\mathbf{r}) \\
&= \frac{\partial}{\partial\rho}\left\{-m\ln C(\kappa,\rho,n) - \sum_{i=1}^{m}\ln\Phi(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m}B(r_i;\kappa,\rho)\right\} \\
&= \frac{-mC2(\kappa,\rho,n)}{C(\kappa,\rho,n)} - \sum_{i=1}^{m}\frac{\Pi(r_i;\kappa,\rho)}{\Phi(r_i;\kappa,\rho)} + 2n\sum_{i=1}^{m}B2(r_i;\kappa,\rho) \\
&= \frac{-mC2(\kappa,\rho,n)}{C(\kappa,\rho,n)} - \sum_{i=1}^{m}A(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m}B2(r_i;\kappa,\rho)
\end{aligned}
$$

and the second order partial derivatives of $l(\kappa,\rho|\,\mathbf{r})$ with respect κ and ρ are given by:

$$
\begin{aligned}
l11(\kappa,\rho) &= \frac{\partial^2}{\partial\kappa^2} l(\kappa,\rho|\,\mathbf{r}) \\
&= \frac{\partial}{\partial\kappa}\left[\frac{\partial}{\partial\kappa} l(\kappa,\rho|\,\mathbf{r})\right] \\
&= \frac{\partial}{\partial\kappa}\left[\frac{-mC1(\kappa,\rho,n)}{C(\kappa,\rho,n)} - \sum_{i=1}^{m}r_iA(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m}B1(r_i;\kappa,\rho)\right] \\
&= \frac{-m\left[C11(\kappa,\rho,n)C(\kappa,\rho,n) - C1^2(\kappa,\rho,n)\right]}{C^2(\kappa,\rho,n)} \\
&\quad - \sum_{i=1}^{m}r_iA1(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m}B11(r_i;\kappa,\rho)
\end{aligned}
$$

$$
\begin{aligned}
l12(\kappa,\rho) &= \frac{\partial^2}{\partial\kappa\partial\rho} l(\kappa,\rho|\,\mathbf{r}) \\
&= \frac{\partial}{\partial\rho}\left[\frac{\partial}{\partial\kappa} l(\kappa,\rho|\,\mathbf{r})\right] \\
&= \frac{\partial}{\partial\rho}\left[\frac{-mC1(\kappa,\rho,n)}{C(\kappa,\rho,n)} - \sum_{i=1}^{m} r_i A(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m} B1(r_i;\kappa,\rho)\right] \\
&= \frac{-m\left[C12(\kappa,\rho,n)C(\kappa,\rho,n) - C1(\kappa,\rho,n)C2(\kappa,\rho,n)\right]}{C^2(\kappa,\rho,n)} \\
&\quad - \sum_{i=1}^{m} r_i A2(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m} B12(r_i;\kappa,\rho)
\end{aligned}
$$

$$
\begin{aligned}
l22(\kappa,\rho) &= \frac{\partial^2}{\partial\rho^2} l(\kappa,\rho|\,\mathbf{r}) \\
&= \frac{\partial}{\partial\rho}\left[\frac{\partial}{\partial\rho} l(\kappa,\rho|\,\mathbf{r})\right] \\
&= \frac{\partial}{\partial\rho}\left[\frac{-mC2(\kappa,\rho,n)}{C(\kappa,\rho,n)} - \sum_{i=1}^{m} A(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m} B2(r_i;\kappa,\rho)\right] \\
&= \frac{-m\left[C22(\kappa,\rho,n)C(\kappa,\rho,n) - C2^2(\kappa,\rho,n)\right]}{C^2(\kappa,\rho,n)} \\
&\quad - \sum_{i=1}^{m} A2(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m} B22(r_i;\kappa,\rho)
\end{aligned}
$$

5.2 Programming Scheme for Matlab Coding

Based on the partial derivatives of (5.4) derived in the previous section, this section presents the maximum likelihood estimation procedure for implementation by Matlab.

5.2.1 The Re-Formulated CMH Returns Distribution

Vaga's CMH returns distribution, re-defined in terms of $\Pi(u;\kappa,\rho)$ and $\Phi(u;\kappa,\rho)$, is given by (5.3) as:

$$
\begin{aligned}
& f(r;n,\kappa,\rho) \\
= \quad & c(\kappa,\rho,n)\cdot\Psi^{-1}(r;\kappa,\rho,n)\cdot\exp\left[2n\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(u;\kappa,\rho)}{\Phi(u;\kappa,\rho)}du\right] \\
= \quad & nC^{-1}(\kappa,\rho,n)\,\Phi^{-1}(r;\kappa,\rho)\exp\left[2nB(r;\kappa,\rho)\right],
\end{aligned}
$$

where:

$$
\begin{aligned}
\Pi(u;\kappa,\rho) &= \sinh(\kappa u+\rho) - 2u\cosh(\kappa u+\rho) \\
\Phi(u;\kappa,\rho) &= \cosh(\kappa u+\rho) - 2u\sinh(\kappa u+\rho) \\
A(u;\kappa,\rho) &= \frac{\Pi(u;\kappa,\rho)}{\Phi(u;\kappa,\rho)} \\
B(r;\kappa,\rho) &= \int\limits_{-\frac{1}{2}}^{r}\frac{\Pi(u;\kappa,\rho)}{\Phi(u;\kappa,\rho)}du = \int\limits_{-\frac{1}{2}}^{r}A(u;\kappa,\rho)\,du \\
C(\kappa,\rho,n) &= c^{-1}(\kappa,\rho,n) = n\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\Phi^{-1}(r;\kappa,\rho)e^{2nB(r;\kappa,\rho)}dr
\end{aligned}
$$

5.2.2 List of Pre-defined Functions

The list of pre-defined functions required for the maximum likelihood estimation procedure are as follows:

$$\Pi(u;\kappa,\rho) = \sinh(\kappa u+\rho) - 2u\cosh(\kappa u+\rho) \tag{5.5}$$

$$\Phi(u;\kappa,\rho) = \cosh(\kappa u+\rho) - 2u\sinh(\kappa u+\rho) \tag{5.6}$$

$$A(u;\kappa,\rho) = \frac{\Pi(u;\kappa,\rho)}{\Phi(u;\kappa,\rho)} \tag{5.7}$$

$$B(r;\kappa,\rho) = \int_{-\frac{1}{2}}^{r} \frac{\Pi(u;\kappa,\rho)}{\Phi(u;\kappa,\rho)} du = \int_{-\frac{1}{2}}^{r} A(u;\kappa,\rho)\, du \tag{5.8}$$

$$C(\kappa,\rho,n) = c^{-1}(\kappa,\rho,n) = n \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \Phi^{-1}(r;\kappa,\rho) e^{2nB(r;\kappa,\rho)} dr \tag{5.9}$$

$$A1(u;\kappa,\rho) = \frac{\partial}{\partial\kappa} A(u;\kappa,\rho) = \frac{u(1-4u^2)}{\Phi^2(u;\kappa,\rho)} \tag{5.10}$$

$$A2(u;\kappa,\rho) = \frac{\partial}{\partial\rho} A(u;\kappa,\rho) = \frac{1-4u^2}{\Phi^2(u;\kappa,\rho)} \tag{5.11}$$

$$D(u;\kappa,\rho) = \frac{(1-4u^2)\,\Pi(u;\kappa,\rho)}{\Phi^3(u;\kappa,\rho)} = A(u;\kappa,\rho)A2(u;\kappa,\rho) \tag{5.12}$$

$$A11(u;\kappa,\rho) = \frac{\partial^2}{\partial\kappa^2} A(u;\kappa,\rho) = -2u^2 D(u;\kappa,\rho) \tag{5.13}$$

$$A12(u;\kappa,\rho) = \frac{\partial^2}{\partial\kappa\partial\rho} A(u;\kappa,\rho) = -2uD(u;\kappa,\rho) \tag{5.14}$$

$$A22(u;\kappa,\rho) = \frac{\partial^2}{\partial\rho^2} A(u;\kappa,\rho) = -2D(u;\kappa,\rho) \tag{5.15}$$

$$B1(r_i;\kappa,\rho) = \frac{\partial}{\partial\kappa}B(r_i;\kappa,\rho) = \int_{-\frac{1}{2}}^{r_i} uA2(u;\kappa,\rho)du \tag{5.16}$$

$$B2(r_i;\kappa,\rho) = \frac{\partial}{\partial\rho}B(r_i;\kappa,\rho) = \int_{-\frac{1}{2}}^{r_i} A2(u;\kappa,\rho)du \tag{5.17}$$

$$B11(r_i;\kappa,\rho) = \frac{\partial^2}{\partial\kappa^2}B(r_i;\kappa,\rho) = -2\int_{-\frac{1}{2}}^{r_i} u^2D(u;\kappa,\rho)du \tag{5.18}$$

$$B12(r_i;\kappa,\rho) = \frac{\partial^2}{\partial\kappa\partial\rho}B(r_i;\kappa,\rho) = -2\int_{-\frac{1}{2}}^{r_i} uD(u;\kappa,\rho)du \tag{5.19}$$

$$B22(r_i;\kappa,\rho) = \frac{\partial^2}{\partial\rho^2}B(r_i;\kappa,\rho) = -2\int_{-\frac{1}{2}}^{r_i} D(u;\kappa,\rho)du \tag{5.20}$$

$$C(\kappa,\rho,n) = n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \Phi^{-1}(r;\kappa,\rho)e^{2nB(r;\kappa,\rho)}dr \tag{5.21}$$

$$
\begin{aligned}
C1(\kappa,\rho,n) &= \frac{\partial}{\partial\kappa}C(\kappa,\rho,n) \\
&= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB1\left(r;\kappa,\rho\right)-rA\left(r;\kappa,\rho\right)\right]dr
\end{aligned} \tag{5.22}
$$

$$
\begin{aligned}
C2(\kappa,\rho,n) &= \frac{\partial}{\partial\rho}C(\kappa,\rho,n) \\
&= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB2\left(r;\kappa,\rho\right)-A\left(r;\kappa,\rho\right)\right]dr
\end{aligned} \tag{5.23}
$$

$$
\begin{aligned}
C11(\kappa,\rho,n) &= \frac{\partial^2}{\partial\kappa^2}C(\kappa,\rho,n) \\
&= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB1\left(r;\kappa,\rho\right)-rA\left(r;\kappa,\rho\right)\right]^2 dr \\
&\quad +n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB11\left(r;\kappa,\rho\right)-rA1\left(r;\kappa,\rho\right)\right]dr
\end{aligned} \tag{5.24}
$$

$$
\begin{aligned}
C12(\kappa,\rho,n) &= \frac{\partial^2}{\partial\kappa\partial\rho}C(\kappa,\rho,n) \\
&= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB2\left(r;\kappa,\rho\right)-A\left(r;\kappa,\rho\right)\right]\left[2nB1\left(r;\kappa,\rho\right)-rA\left(r;\kappa,\rho\right)\right]dr \\
&\quad +n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB12\left(r;\kappa,\rho\right)-rA2\left(r;\kappa,\rho\right)\right]dr
\end{aligned} \tag{5.25}
$$

$$
\begin{aligned}
C22(\kappa,\rho,n) &= \frac{\partial^2}{\partial\rho^2}C(\kappa,\rho,n) \\
&= n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB2\left(r;\kappa,\rho\right)-A\left(r;\kappa,\rho\right)\right]^2 dr \\
&\quad +n\int_{r=-\frac{1}{2}}^{\frac{1}{2}} \frac{e^{2nB(r;\kappa,\rho)}}{\Phi(r;\kappa,\rho)}\left[2nB22\left(r;\kappa,\rho\right)-A2\left(r;\kappa,\rho\right)\right]dr
\end{aligned} \tag{5.26}
$$

5.2.3 The Maximum Likelihood Estimation Procedure

The maximum likelihood estimation procedure for the control parameters of CMH returns distribution is as follows:

1. Maximise the log-likelihood function of (5.3), given by (5.4) as:

$$l(\kappa, \rho|\, \mathbf{r}, n) = m \ln n - m \ln C(\kappa, \rho, n) - \sum_{i=1}^{m} \ln \Phi(r_i; \kappa, \rho) + 2n \sum_{i=1}^{m} B(r_i; \kappa, \rho)$$

   subject to $C(\kappa, \rho, n) > 0$. This is a constrained non-linear optimisation problem.

2. Check that $(\hat{\kappa}, \hat{\rho})$ obtained from implementation of 1. is a satisfactory local maximum based on the following criteria:

$$|l1\,(\hat{\kappa}, \hat{\rho})| + |l2\,(\hat{\kappa}, \hat{\rho})| \approx 0,$$

   where the first-order partial derivatives of (5.4) are:

$$\begin{aligned} l1(\kappa, \rho) &= \frac{\partial}{\partial \kappa} l(\kappa, \rho) \\ &= \frac{-mC1(\kappa, \rho, n)}{C(\kappa, \rho, n)} - \sum_{i=1}^{m} r_i A(r_i; \kappa, \rho) + 2n \sum_{i=1}^{m} B1(r_i; \kappa, \rho) \qquad (5.27) \\ l2(\kappa, \rho) &= \frac{\partial}{\partial \rho} l(\kappa, \rho) \\ &= \frac{-mC2(\kappa, \rho, n)}{C(\kappa, \rho, n)} - \sum_{i=1}^{m} A(r_i; \kappa, \rho) + 2n \sum_{i=1}^{m} B2(r_i; \kappa, \rho) \qquad (5.28) \end{aligned}$$

3. Calculate the variance-covariance matrix for $(\hat{\kappa}, \hat{\rho})$ given by:

$$\left. \begin{pmatrix} \frac{\partial^2 l(\kappa,\rho)}{\partial \kappa^2} & \frac{\partial^2 l(\kappa,\rho)}{\partial \kappa \partial \rho} \\ \frac{\partial^2 l(\kappa,\rho)}{\partial \rho \partial \kappa} & \frac{\partial^2 l(\kappa,\rho)}{\partial \rho^2} \end{pmatrix}^{-1} \right|_{(\kappa=\hat{\kappa}, \rho=\hat{\rho})}$$

$$= \begin{pmatrix} l11\,(\hat{\kappa}, \hat{\rho}) & l12\,(\hat{\kappa}, \hat{\rho}) \\ l12\,(\hat{\kappa}, \hat{\rho}) & l22\,(\hat{\kappa}, \hat{\rho}) \end{pmatrix}^{-1},$$

where the second-order partial derivatives of (5.4) are:

$$\begin{aligned} l11(\kappa,\rho) &= \frac{\partial^2}{\partial\kappa^2} l(\kappa,\rho) \\ &= \frac{-m\left[C11(\kappa,\rho,n)C(\kappa,\rho,n) - C1^2(\kappa,\rho,n)\right]}{C^2(\kappa,\rho,n)} \\ &\quad -\sum_{i=1}^{m} r_i A1(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m} B11(r_i;\kappa,\rho), \end{aligned} \tag{5.29}$$

$$\begin{aligned} l12(\kappa,\rho) &= \frac{\partial^2}{\partial\kappa\partial\rho} l(\kappa,\rho) \\ &= \frac{-m\left[C12(\kappa,\rho,n)C(\kappa,\rho,n) - C1(\kappa,\rho,n)C2(\kappa,\rho,n)\right]}{C^2(\kappa,\rho,n)} \\ &\quad -\sum_{i=1}^{m} r_i A2(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m} B12(r_i;\kappa,\rho), \end{aligned} \tag{5.30}$$

and

$$\begin{aligned} l22(\kappa,\rho) &= \frac{\partial^2}{\partial\rho^2} l(\kappa,\rho) \\ &= \frac{-m\left[C22(\kappa,\rho,n)C(\kappa,\rho,n) - C2^2(\kappa,\rho,n)\right]}{C^2(\kappa,\rho,n)} \\ &\quad -\sum_{i=1}^{m} A2(r_i;\kappa,\rho) + 2n\sum_{i=1}^{m} B22(r_i;\kappa,\rho). \end{aligned} \tag{5.31}$$

5.3 Fitting CMH Return Distribution to Partitioned Daily Return in Each Market State

The maximum likelihood estimation procedure derived in the previous section is coded and implemented in Matlab. According to Vaga (1990), the CMH returns distribution as given by (3.33) is specified only by its control parameters (κ, ρ). The third parameter n (the number of industrial groups in the market under study) is taken as fixed, which Vaga (1990) have assumed to be 186. However, the CMH distribution specified only by (κ, ρ) does not have sufficient flexibility to fit the data, because the optimisation procedure does not produce a convergent result for partitioned daily returns in any of the five market states. Further investigations reveals the effect of each of the three parameters κ, ρ, n on the shape of the CMH distribution:

1. ρ determines the mode and skewness of the CMH distribution, and
2. κ and n **jointly** determines the shape and kurtosis of the CMH distribution.

Based on the above insight, it is necessary to include n as an additional parameter specifying the CMH distribution.

5.3.1 Maximum Likelihood Estimates of CMH Control Parameters in Each Market State

Table (5.1) below summarises the results from the maximum likelihood estimation procedure given partitioned daily returns from each of the five market states:

State	M.L.E.	$L1 = \frac{\partial l(\kappa,\rho)}{\partial \kappa}$	$L2 = \frac{\partial l(\kappa,\rho)}{\partial \rho}$	$\|L1\| + \|L2\|$
RW (1)	$\widehat{\kappa} = -58.1413$ $\widehat{\rho} = -0.0475$ $\widehat{n} = 79.2857$	-1.1276×10^{-4}	0.1312	0.1313
BULL (2)	$\widehat{\kappa} = -110.3242$ $\widehat{\rho} = 1.2913$ $\widehat{n} = 77.4958$	-4.3998×10^{-4}	-0.0367	0.0371
BEAR (3)	N/A	N/A	N/A	N/A
CHAOS (4)	$\widehat{\kappa} = -65.3561$ $\widehat{\rho} = 0.0778$ $\widehat{n} = 19.9067$	-2.5659×10^{-4}	-0.0092	0.0095
TRANS (5)	$\widehat{\kappa} = -86.3889$ $\widehat{\rho} = -0.0367$ $\widehat{n} = 161.0610$	-1.5605×10^{-4}	0.0883	0.0885

Table 5.1: mle of CMH Parameters for Each Market State

From Table (5.1) above, the following are noted

1. The maximum likelihood procedure did not produce a convergent result given daily returns

in BEAR market state. Thus, parameter values are marked "N/A". Possible reasons are:

(i) There are errors in the codes, and

(ii) The optimisation procedure given the form of the log-likelihood function is incapable of producing a convergent result in the presence of large negative daily returns value.

Due to time constraints, further investigations into the cause for the non-convergent result was not possible.

2. The estimated parameter values $\hat{\kappa}$, $\hat{\rho}$,and $\hat{n}$ for each market state are far different to the values in Table 3.1 suggested by Vaga (1990). The most surprising results are that $\hat{\kappa}'s$ have relatively large negative value in all four of the market states for which an estimate is obtained. An obvious explanation would be that the CMH distribution has been fitted to daily returns and not annualised returns. This confirms that the CMH distribution is indeed a model of the annualised returns. However, the financial interpretation of κ needs to be re-considered for daily returns.

3. The $\hat{\rho}$ have values relatively close to zero for RW, CHAOS and TRANSITION market state. However, $\hat{\rho}$ is a relatively large positive value for BULL market state, which is as expected.

4. For different market state, $\hat{\kappa}$ values have varied significantly while the $\hat{\rho}$ values have varied little. This supports the intuition that market sentiment should vary more wildly than economic fundamentals.

5. The values of $\hat{n}$ also varied considerably from one market state to another. This is in contradiction to Vaga's assertion that n remains relatively constant with changes in κ and ρ.

Table 5.2 below presents the variance-covariance matrices for each market state that produced maximum likelihood estimates:

From Table 5.2 above, the following are noted in the four market states that produced maximum likelihood estimates:

State	Sample Size	Var/Cov	$s.d.(\widehat{\kappa})$	$s.d.(\widehat{\rho})$	$Corr\ (\widehat{\kappa},\widehat{\rho})$
RW (1)	1094	$\begin{pmatrix} 6.8010 & 0.0051 \\ 0.0051 & 0.0003 \end{pmatrix}$	2.6079	0.0173	0.1129
BULL (2)	765	$\begin{pmatrix} 23.7742 & -0.2795 \\ -0.2795 & 0.0041 \end{pmatrix}$	4.8759	0.0640	−0.89523
BEAR (3)	305	N/A	N/A	N/A	N/A
CHAOS (4)	237	$\begin{pmatrix} 26.4869 & -0.0209 \\ -0.0209 & 0.0043 \end{pmatrix}$	5.1465	0.0655	−0.06193
TRANS (5)	1884	$\begin{pmatrix} 6.2801 & 0.0058 \\ 0.0058 & 0.0002 \end{pmatrix}$	2.506	0.0141	0.16366

Table 5.2: Variance-Covariance Matrices for mle of CMH Parameters

1. The standard deviations for the $\hat{\kappa}'s$ are relatively small compared to the value of the estimates.

2. The standard deviations for the $\hat{\rho}'s$ are relatively larger when compared to the value of the estimates.

3. The standard deviation of $\hat{\kappa}$ and $\hat{\rho}$ are the smallest in market state for which there is a large sample size and vice versa.

4. The correlation between $\hat{\kappa}$ and $\hat{\rho}$ in the RW, CHAOS and TRANS market states are all very close to zero and may be regarded as uncorrelated. However, there is a peculiarly large negative correlation between $\hat{\kappa}$ and $\hat{\rho}$ in the BULL market state of −0.90. This may be interpreted to mean that in the BULL market state, an increase in the value of κ will lead to a decrease in the value of ρ, and vice versa. Of course, the changes cannot be so large that CMH distribution changes to another market state. Not excluding the possibility of an error in computation, but an investigation into the cause of this relationship may provide certain insights into the market.

5.3.2 Properties of CMH Distribution Given the Maximum Likelihood Estimates

For each of the market state that produced a m.l.e. triplet $(\hat{\kappa}, \hat{\rho}, \hat{n})$, the first four moments of the CMH distribution is computed by Matlab to describe the properties of the distribution in that market state. Table 5.3 below presents the first four moments of the CMH distribution under each triplets of parameter estimates:

State	Mean	Std Dev	$\frac{Mean}{StdDev}$	Skewness	Excess Kurtosis
RW	−0.0010	0.0100	−0.10	0.0005	0.6589
BULL	0.0113	0.0072	1.57	0.0056	0.9587
BEAR	N/A	N/A	N/A	N/A	N/A
CHAOS	0.0011	0.0160	0.07	0.0006	1.4650
TRANS	−0.0006	0.0058	−0.10	−0.0038	0.5523

Table 5.3: First Four Moments for CMH Distribution given mle parameters

From Table 5.3 above, the following interpretations may be made for the four market states that produced the estimates:

1. The mean daily returns are all very small.

2. The BULL market state has the largest mean daily return as it should be.

3. The TRANS market state has the smallest standard deviation. The BULL market has a slightly larger standard deviation.

4. The BULL market state produced the by far the largest mean daily return per unit of standard deviation. The CHAOS market state also produced a positive mean daily return per unit standard deviation.

5. The coefficient of skewness are all relatively close to zero. The RW market state has a coefficient of skewness that is closest to zero and hence has the most symmetrical distribution about its mean. The magnitude of coefficient of skewness in BULL market state is the largest and hence the most (positively) asymmetric distribution.

6. The CHAOS market state has the largest excess kurtosis, while the TRANS market state the smallest.

For each of the market state that produced a m.l.e. triplet $(\hat{\kappa}, \hat{\rho}, \hat{n})$, the CMH distribution is plotted by Matlab. Figure 5-1 below shows the CMH distribution fitted to daily returns in Random Walk market state:

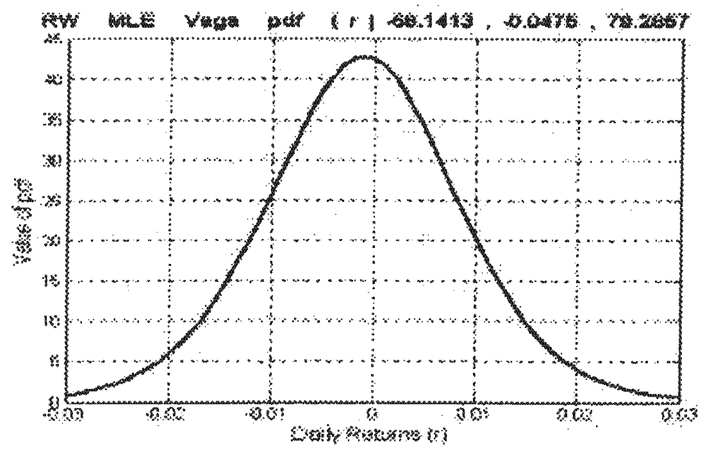

Figure 5-1: CMH Distribution Fitted to Returns in RW Market State

From Figure 5-1 above, the following characteristics are noted:

1. The distribution has the characteristic "bell-shape" for the normal distribution and appears to fit the distribution of partitioned daily returns in RW market state (see Figure 4-14) satisfactorily.

2. The mode of the distribution is slightly less than zero.

3. Recall that the standard deviation in the RW state is exactly 0.01. Then, note that the value of p.d.f. is quite close to zero for daily returns at ±3 standard deviations from its mean. this is in agreement with the properties of normal distribution.

4. Based on the above three observations, the CMH distribution appears relatively normal in the RW market state.

Figure 5-2 below shows the CMH distribution fitted to daily returns in Coherent Bull market state:

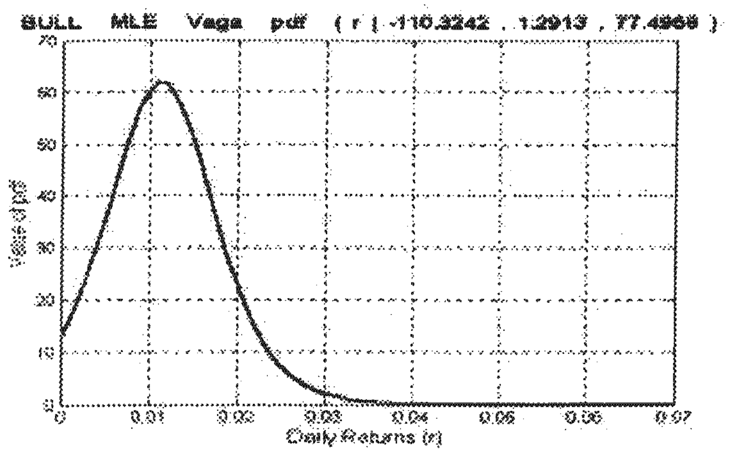

Figure 5-2: CMH Distribution Fitted to Returns in BULL Market State

From Figure 5-2 above, the following characteristics are noted:

1. The mode of the distribution is slightly greater than 0.01.

2. The distribution appears to have a long positive tail and hence positively skewed.

Figure 5-3 below shows the CMH distribution fitted to daily returns in Chaotic market state:

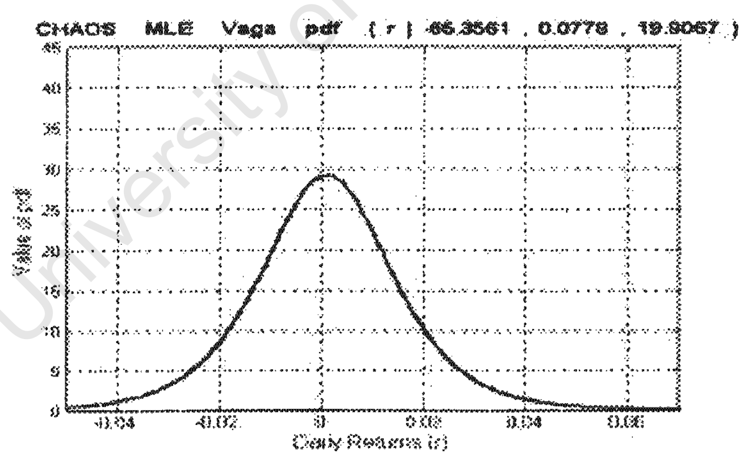

Figure 5-3: CMH Distribution Fitted to Returns in CHAOS Market State

From Figure 5-3 above, the following characteristics are noted:

1. The mode of the distribution is slightly greater than zero.

2. The distribution appears relatively symmetric about its mean.

3. The distribution appears relatively flat-peaked with a slightly longer right tail than the left.

Figure 5-4 below shows the CMH distribution fitted to daily returns in Transition market state:

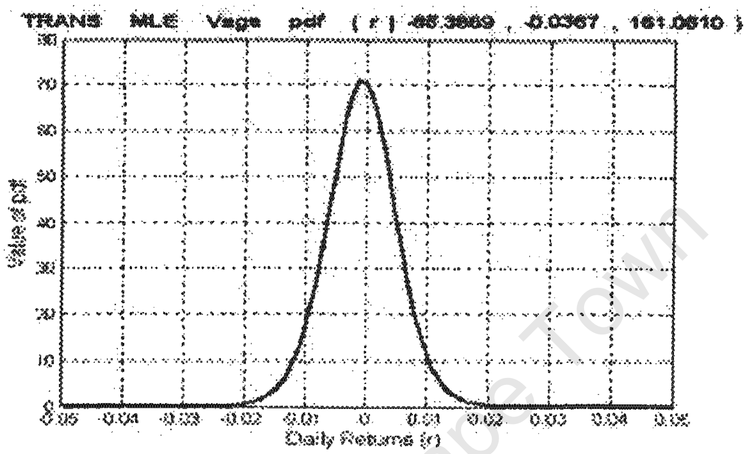

Figure 5-4: CMH Distribution Fitted to Returns in TRANS Market State

From Figure 5-4 above, the following characteristics are noted:

1. The mode of the distribution is slightly less than zero.

2. The distribution appears relatively symmetric about its mean.

3. The distribution has a very distinct sharp peaked with no visibly longer tail on either side of the mode.

Chapter 6

DIFFUSION PROCESSES ASSOCIATED WITH CMH

The aim of this chapter is to carry out the seventh objective for this dissertation:

Objective 7: Exploring the CMH returns distribution based on the related diffusion processes.

6.1 The Kolmogorov Equations for Diffusion Processes

Let $(X_t)_{t\geq 0}$ be some given stochastic process. Suppose $X(t) = x$, i.e. the stochastic process is in state x at time t. We want to analyse the incremental change in the process over the time interval $[t, t+\Delta t]$ given by:

$$X(t+\Delta t) - X(t).$$

In order to satisfy the condition that the stochastic process changes continuously over time, we require the change $X(t+\Delta t) - X(t)$ to be small whenever changes in time Δt is small. Then, it follows that for any $\delta > 0$, we have:

$$\lim_{\Delta t \to 0} \left\{ \Pr\left[\left| X(t+\Delta t) - X(t) \right| > \delta \left| X(t) = x \right. \right] \right\} = 0. \tag{6.1}$$

Now, define the transition distribution function of $(X_t)_{t\geq 0}$ by:

$$F(s, x; t, y) = \Pr\left[X(t) < y \left| X(s) = x \right. \right].$$

Then, the continuity of stochastic process (6.1) can be re-written as:

$$\lim_{\Delta t\to 0}\left\{\int\limits_{|y-x|>\delta} d_y F(t,x;t+\Delta t,y)\right\}=0. \tag{6.2}$$

Now, some usual conditions on $(X_t)_{t\geq 0}$ required in the definition of a diffusion process will be stated as follows:

Condition 77 *Let $(X_t)_{t\geq 0}$ be some given stochastic process with transition distribution function $F(s,x;t,y)$.*

1. ***(Continuity Condition)*** *To ensure the continuity of $(X_t)_{t\geq 0}$, the following rather strong condition is usually required:*

$$\lim_{\Delta t\to 0}\left\{\frac{1}{\Delta t}\int\limits_{|y-x|>\delta} d_y F(t,x;t+\Delta t,y)\right\}=0, \tag{6.3}$$

or the even stronger condition:

$$\lim_{\Delta t\to 0}\left\{\frac{1}{\Delta t}\int\limits_{|y-x|>\delta} (y-x)^2\, d_y F(t,x;t+\Delta t,y)\right\}=0. \tag{6.4}$$

2. ***(Time Proportionality of Increment)*** *If the increment $X(t+\Delta t)-X(t)$ is directly proportional to Δt, then we can define a function $a(t,x)$ by:*

$$\lim_{\Delta t\to 0}\left\{\frac{1}{\Delta t}\int\limits_{-\infty}^{\infty} (y-x)\, d_y F(t,x;t+\Delta t,y)\right\}=a(t,x), \tag{6.5}$$

and

3. ***(Time Proportionality of Variability in Increment)*** *If the variance of the increment $Var[X(t+\Delta t)-X(t)]$ is directly proportional to Δt, then we can define a function*

$b(t,x)$ by:

$$\lim_{\Delta t \to 0}\left\{\frac{1}{\Delta t}\int_{-\infty}^{\infty}(y-x)^2\,d_yF(t,x;t+\Delta t,y)\right\}=b(t,x)>0. \tag{6.6}$$

Remark 78 *The above limits are all assumed to be uniform with respect to x.*

Proposition 79 *If condition (6.4) holds, then conditions (6.5) and (6.6) can be re-written respectively as:*

$$a(t,x)=\lim_{\Delta t\to 0}\left\{\frac{1}{\Delta t}\int_{|y-x|\le\delta}(y-x)\,d_yF(t,x;t+\Delta t,y)\right\}, \tag{6.7}$$

and

$$b(t,x)=\lim_{\Delta t\to 0}\left\{\frac{1}{\Delta t}\int_{|y-x|\le\delta}(y-x)^2\,d_yF(t,x;t+\Delta t,y)\right\}>0. \tag{6.8}$$

Based on the conditions stated above, we shall now present two mathematical definitions of diffusion processes. In general, a diffusion process is defined as follows:

Definition 80 (Diffusion Processes-General) *If a Markov process $\{X_t, t\in\mathbb{R}^+\}$ satisfies conditions (6.3), (6.7) and (6.8), then it is said to be a diffusion process.*

Sometimes, a diffusion process is also defined as follows:

Definition 81 (Diffusion Processes-Alternative) *If a Markov process $\{X_t, t\in\mathbb{R}^+\}$ satisfies conditions (6.2), (6.4), and (6.6), then it is said to be a diffusion process.*

Now that we have defined what is a diffusion process, we can define its parameter functions:

Definition 82 *If $\{X_t, t\in\mathbb{R}^+\}$ is a diffusion process, then:*

1. *The function*

$$a(t,x)=\begin{cases}\lim_{\Delta t\to 0}\left\{\frac{1}{\Delta t}\int_{|y-x|\le\delta}(y-x)\,d_yF(t,x;t+\Delta t,y)\right\} & \text{, if (6.4) holds}\\ \lim_{\Delta t\to 0}\left\{\frac{1}{\Delta t}\int_{-\infty}^{\infty}(y-x)\,d_yF(t,x;t+\Delta t,y)\right\} & \text{, otherwise}\end{cases}$$

is called the drift coefficient *of the diffusion process, and*

2. *The function*

$$b(t,x) = \begin{cases} \lim_{\Delta t \to 0} \left\{ \frac{1}{\Delta t} \int\limits_{|y-x| \le \delta} (y-x)^2 \, d_y F(t,x;t+\Delta t,y) \right\} > 0 & \text{, if (6.4) holds} \\ \lim_{\Delta t \to 0} \left\{ \frac{1}{\Delta t} \int\limits_{-\infty}^{\infty} (y-x)^2 \, d_y F(t,x;t+\Delta t,y) \right\} > 0 & \text{, otherwise} \end{cases}$$

is called the diffusion coefficient *of the diffusion process.*

Having defined the diffusion processes and its drift and diffusion coefficients, now the Kolmogorov equations that governs the transition probability density function of diffusion processes can be stated. First, Kolmogorov's forward equation is given by:

Theorem 83 (Kolmogorov Forward Equation) *Let*

$$\{X(t) : t \in T = [a,b]\} \text{ be a stochastic process,}$$

and

$$p(s,x;t,y) \equiv \text{the transition probability density function of } X(t).$$

If $p(s,x;t,y)$ *and the following partial derivatives of exists:*

$$\frac{\partial}{\partial t} p(s,x;t,y)$$
$$\frac{\partial}{\partial y} [a(t,y) \cdot p(s,x;t,y)]$$
$$\frac{\partial^2}{\partial y^2} [b(t,y) \cdot p(s,x;t,y)].$$

Then, $p(s,x;t,y)$ *satisfies the following partial differential equation:*

$$\frac{\partial}{\partial t} p(s,x;t,y) = -\frac{\partial}{\partial y} [a(t,y) \cdot p(s,x;t,y)] + \frac{1}{2} \frac{\partial^2}{\partial y^2} [b(t,y) \cdot p(s,x;t,y)].$$

Proof. Let $Q(y)$ be an arbitrarily chosen, twice differentiable function, $Q(y) = 0$ for $y \notin [a,b]$. Then, from the continuity of $Q(y), Q'(y), Q''(y)$, we can see that $Q(y), Q'(y), Q''(y)$

is equal to zero at $y = a$ and $y = b$. Now, let us investigate the following integral:

$$
\begin{aligned}
I &= \int_{y=-\infty}^{\infty} Q(y) \frac{\partial p}{\partial t} dy \\
&= \int_{y=a}^{b} Q(y) \frac{\partial}{\partial t} p(s, x; t, y)\, dy \\
&= \frac{\partial}{\partial t} \int_{y=a}^{b} Q(y)\, p(s, x; t, y)\, dy \\
&= \lim_{\Delta t \to 0} \left[\int_{y=a}^{b} Q(y) \frac{p(s, x; t + \Delta t, y) - p(s, x; t, y)}{\Delta t} dy \right] \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left\{ \int_{y=a}^{b} Q(y) \left[\int_{u=-\infty}^{\infty} p(s, x; t, u) \cdot p(t, u; t + \Delta t, y)\, du \right] dy - \int_{y=a}^{b} Q(y) \cdot p(s, x; t, y)\, dy \right\} \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left\{ \int_{u=-\infty}^{\infty} p(s, x; t, u) \left[\int_{y=a}^{b} p(t, u; t + \Delta t, y) \cdot Q(y)\, dy - Q(u) \right] du \right\}.
\end{aligned}
$$

Since $Q(y)$ is bounded and follows condition (6.3), we have:

$$
\begin{aligned}
& \int_{|y-u|>\delta} d_y F(t, u; t + \Delta t, y) \\
= & \int_{|y-u|>\delta} p(t, u; t + \Delta t, y)\, dy \\
= & \ o(\Delta t)
\end{aligned}
$$

and

$$
\begin{aligned}
& \int_{|y-u|>\delta} p(t, u; t + \Delta t, y) \cdot Q(y)\, dy \\
= & \ o(\Delta t).
\end{aligned}
$$

Thus, we have:

$$\begin{aligned} &\frac{1}{\Delta t}\left[\int_{y=a}^{b} p\left(t,u;t+\Delta t,y\right)\cdot Q\left(y\right)dy - Q\left(u\right)\right] \\ = \; &\frac{1}{\Delta t}\left[\int_{|y-u|\leq\delta} p\left(t,u;t+\Delta t,y\right)\cdot Q\left(y\right)dy - Q\left(u\right)\right] + o\left(1\right). \end{aligned} \tag{6.9}$$

Now, in the neighbourhood of u, if we Taylor expand $Q\left(y\right)$, we get:

$$Q\left(y\right) = Q\left(u\right) + Q'\left(u\right)\left(y-u\right) + \frac{1}{2}Q''\left(u\right)\left(y-u\right)^2 + o\left[\left(y-u\right)^2\right]. \tag{6.10}$$

Substituting (6.10) into (6.9), we get:

$$\begin{aligned} &\frac{1}{\Delta t}\left[\int_{|y-u|\leq\delta} p\left(t,u;t+\Delta t,y\right)\cdot Q\left(y\right)dy - Q\left(u\right)\right] \\ = \; &\frac{1}{\Delta t}\left\{\int_{|y-u|\leq\delta} p\left(t,u;t+\Delta t,y\right)\left[Q\left(u\right) + Q'\left(u\right)\left(y-u\right) + \frac{Q''\left(u\right)}{2}\left(y-u\right)^2 + o\left(\left(y-u\right)^2\right)\right]dy - Q\left(u\right)\right. \\ = \; &Q\left(u\right)\frac{1}{\Delta t}\int_{|y-u|>\delta} p\left(t,u;t+\Delta t,y\right)dy \\ &+Q'\left(u\right)\frac{1}{\Delta t}\int_{|y-u|\leq\delta}\left(y-u\right)p\left(t,u;t+\Delta t,y\right)dy \\ &+\frac{1}{2}Q''\left(u\right)\left[\frac{1}{\Delta t}\int_{|y-u|\leq\delta}\left(y-u\right)^2\cdot p\left(t,u;t+\Delta t,y\right)dy\right] \\ &+o\left(1\right). \end{aligned}$$

Applying conditions (6.3),(6.7),(6.8), we can see that as $t \to 0$, the above tends to:

$$Q'\left(u\right)a\left(t,u\right) + \frac{1}{2}Q''\left(u\right)b\left(t,u\right).$$

It follows that:

$$\begin{aligned} I &= \int_{u=-\infty}^{\infty} p(s,x;t,u)\left[Q'(u)\,a(t,u)+\frac{1}{2}Q''(u)\,b(t,u)\right]du \\ &= \int_{u=-\infty}^{\infty} pQ'a\,du+\frac{1}{2}\int_{u=-\infty}^{\infty} pQ''b\,du, \end{aligned}$$

and using integration by parts, we get:

$$\int_{u=-\infty}^{\infty} apQ'du = apQ|_{-\infty}^{\infty} - \int_{u=-\infty}^{\infty} Q\frac{\partial}{\partial u}(ap)\,du.$$

Since $Q(u)=0$ for $u \notin [a,b]$, thus the first term of above is equal to zero. Hence, we get:

$$\int_{u=-\infty}^{\infty} apQ'du = -\int_{u=-\infty}^{\infty}\left[\frac{\partial}{\partial u}(ap)\right]\cdot Q\,du,$$

and similarly we get:

$$\int_{u=-\infty}^{\infty} bpQ''du = \int_{u=-\infty}^{\infty}\left[\frac{\partial^2}{\partial u^2}(bp)\right]Q\,du.$$

Thus, we get:

$$\begin{aligned} I &= \int_{u=-\infty}^{\infty} Q\frac{\partial p}{\partial t}du \\ &= -\int_{u=-\infty}^{\infty}\left[\frac{\partial}{\partial u}(ap)\right]Q\,du+\frac{1}{2}\int_{u=-\infty}^{\infty}\left[\frac{\partial^2}{\partial u^2}(bp)\right]Q\,du \\ &= \int_{u=-\infty}^{\infty} Q\left[-\frac{\partial}{\partial u}(ap)+\frac{1}{2}\frac{\partial^2}{\partial u^2}(bp)\right]du. \end{aligned}$$

Given the arbitrariness of Q, and the continuity property of $\frac{\partial p}{\partial t}, \frac{\partial}{\partial u}(ap), \frac{\partial^2}{\partial u^2}(bp)$, we immediately get:

$$\frac{\partial}{\partial t}p(s,x;t,u) = -\frac{\partial}{\partial u}\left[a(t,u)\cdot p(s,x;t,u)\right]+\frac{1}{2}\frac{\partial^2}{\partial u^2}\left[b(t,u)\cdot p(s,x;t,u)\right],$$

as required. ■

Remark 84 *In physics, the Kolmogorov Forward equation is called the Fokker-Planck Equation.*

Now, we will state Kolmogorov's backward equation:

Theorem 85 (Kolmogorov's Backward Equation) *Let* $\{X(t) : t \in T = [a, b]\}$ *be a stochastic process. If the transition distribution function of* $X(t)$ *given by:*

$$F(s, x; t, y)$$

exists together with the following partial derivatives:

$$\frac{\partial}{\partial x} F(s, x; t, y),$$
$$\frac{\partial^2}{\partial x^2} F(s, x; t, y),$$

then $F(s, x; t, y)$ *satisfies the following partial differential equation:*

$$\frac{\partial}{\partial s} F(s, x; t, y) = -a(s, x) \frac{\partial}{\partial x} F(s, x; t, y) - \frac{1}{2} b(s, x) \frac{\partial^2}{\partial x^2} F(s, x; t, y).$$

Proof. Firstly, we have the following:

$$\begin{aligned}
& \frac{F(s - \Delta s, x; t, y) - F(s, x; t, y)}{\Delta s} \\
= & \frac{1}{\Delta s} \left\{ \left[\int_{u=-\infty}^{\infty} d_u F(s - \Delta s, x; s, u) \cdot F(s, u; t, y) \right] - F(s, x; t, y) \right\} \\
= & \frac{1}{\Delta s} \int_{u=-\infty}^{\infty} d_u F(s - \Delta s, x; s, u) \cdot [F(s, u; t, y) - F(s, x; t, y)] \\
= & I_1 + I_2, \qquad (6.11)
\end{aligned}$$

where:

$$
\begin{aligned}
|I_1| &= \left| \frac{1}{\Delta s} \int\limits_{|u-x|>\delta} d_u F(s-\Delta s, x; s, u) \cdot [F(s, u; t, y) - F(s, x; t, y)] \right| \\
&\leq \frac{2}{\Delta s} \int\limits_{|u-x|>\delta} d_u F(s-\Delta s, x; s, u) \to 0, \text{ as } \Delta s \to 0
\end{aligned}
$$

and

$$
\begin{aligned}
I_2 &= \frac{1}{\Delta s} \int\limits_{|u-x|\leq\delta} d_u F(s-\Delta s, x; s, u) \cdot [F(s, u; t, y) - F(s, x; t, y)] \\
&= \frac{1}{\Delta s} \int\limits_{|u-x|\leq\delta} d_u F(s-\Delta s, x; s, u) \left[\frac{\partial F(s, x; t, y)}{\partial x} (u-x) + \frac{1}{2} \frac{\partial^2 F(s, x; t, y)}{\partial x^2} (u-x)^2 + o\left[(u-x)^2\right.\right. \\
&= \left[\frac{\partial F(s, x; t, y)}{\partial x} \cdot \frac{1}{\Delta s} \int\limits_{|u-x|\leq\delta} (u-x)\, d_u F(s-\Delta s, x; s, u) \right] \\
&\quad + \left[\left(\frac{1}{2} \cdot \frac{\partial^2 F(s, x; t, y)}{\partial x^2} + o(1) \right) \cdot \frac{1}{\Delta s} \int\limits_{|u-x|\leq\delta} (u-x)^2\, d_u F(s-\Delta s, x; s, u) \right] \\
&\to a(s, x) \frac{\partial F(s, x; t, y)}{\partial x} + \frac{1}{2} b(s, x) \frac{\partial^2 F(s, x; t, y)}{\partial x^2}, \text{ as } \Delta s \to 0.
\end{aligned}
$$

As $\Delta s \to 0$, we also have:

$$
\frac{F(s-\Delta s, x; t, y) - F(s, x; t, y)}{\Delta s} \to -\frac{\partial}{\partial s} F(s, x; t, y). \tag{6.12}
$$

Combining (6.11) and (6.12), we have:

$$
\begin{aligned}
& -\frac{\partial}{\partial s} F(s, x; t, y) \\
&= a(s, x) \frac{\partial}{\partial x} F(s, x; t, y) + \frac{1}{2} b(s, x) \frac{\partial^2}{\partial x^2} F(s, x; t, y),
\end{aligned}
$$

and hence

$$\begin{aligned}
&\frac{\partial}{\partial s} F(s, x; t, y) \\
= \quad & -a(s, x) \frac{\partial}{\partial x} F(s, x; t, y) - \frac{1}{2} b(s, x) \frac{\partial^2}{\partial x^2} F(s, x; t, y),
\end{aligned}$$

as required. ■

6.2 Removing the Drift of a Diffusion Process

Recall that the one-dimensional Stochastic Differential Equation (SDE) with a drift term has the following general form:

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t, \tag{6.13}$$

where:

$$\begin{aligned}
b &: \mathbb{R} \to \mathbb{R}, \text{ and} \\
\sigma &: \mathbb{R} \to \mathbb{R}
\end{aligned}$$

are Borel-measurable coefficients.

Engelbert and Schmidt (1984) shows that the solution to a zero drift SDE cannot explode. Thus, it is desirable to transform SDE (6.13) such that the drift is removed. This reduction requires the following two assumptions:

1. *Non-degeneracy* (ND) condition:

$$\sigma^2(x) > 0, \text{ for all } x \in \mathbb{R}, \tag{ND}$$

   and

2. *Local integrability* (LI) condition, which states that $\forall x \in \mathbb{R}, \exists \epsilon > 0$ such that:

$$\int_{x-\epsilon}^{x+\epsilon} \frac{|b(y)|}{\sigma^2(y)} dy < \infty. \tag{LI}$$

Definition 86 *Under the assumptions of (ND) and (LI), the* scale function *is defined as follows:*

$$p(x) \equiv \int_{\xi=c}^{x} \exp\left[-2\int_{u=c}^{\xi} \frac{b(u)}{\sigma^2(u)}du\right] d\xi, \tag{6.14}$$

where:

$$\begin{aligned} c &\in \mathbb{R} \text{ is a fixed number, and} \\ x &\in \mathbb{R}. \end{aligned}$$

The function p defined above has the following two properties:

1. p has a continuous, strictly positive derivative, and
2. p'' exists almost everywhere and satisfies:

$$p''(x) = -\frac{2b(x)}{\sigma^2(x)} \cdot p'(x). \tag{6.15}$$

From now on, p'' will be taken to mean the locally integrable function defined by (6.15) on the entire real number line $\mathbb{R}$. This definition is possible due to the assumption of (ND).

The function p maps $\mathbb{R}$ onto $(p(-\infty), p(\infty))$, and has a continuously differentiable inverse:

$$q : (p(-\infty), p(\infty)) \to \mathbb{R}.$$

This inverse function q has the following two properties:

1. The first derivative is absolutely continuous and is given by:

$$q'(y) = \frac{1}{p'[q(y)]},$$

and

2. The second derivative is given by:

$$\begin{aligned} q''(y) &= -\frac{p''[q(y)]}{p'[q(y)]} \cdot \left[q'(y)\right]^2 \\ &= \left(\frac{2b[q(y)]}{\sigma^2[q(y)]}\right) \cdot \left(\frac{1}{[p'(q(y))]^2}\right), \text{ a.e. in } (p(-\infty), p(\infty)). \end{aligned}$$

Furthermore, extend p to $[-\infty, \infty]$ and q to $[p(-\infty), p(\infty)]$ so that the resulting functions are continuous in topology on the extended real number system.

Proposition 87 (SDE Solution) *Assume (ND) and (LI). A process*

$$X = \{X_t, F_t : t \in [0, \infty)\}$$

is a weak (or strong) solution of SDE (6.13) if and only if the process

$$Y = \{Y_t \equiv p(X_t), F_t : t \in [0, \infty)\}$$

is a weak (or strong solution) of:

$$Y_t = Y_0 + \int_{s=0}^{t} \tilde{\sigma}(Y_s)\, dW_s, \text{ for } t \in [0, \infty), \tag{6.16}$$

where:

$$Y_0 \in (p(-\infty), p(\infty)), \text{ a.s.} \tag{6.17}$$

and

$$\tilde{\sigma}(y) = \begin{cases} p'[q(y)] \cdot \sigma[q(y)] & , \text{ for } y \in (p(-\infty), p(\infty)) \\ 0 & , \text{ otherwise.} \end{cases} \tag{6.18}$$

The process X may explode in finite time, but the process Y does not.

Proof. (Karatzas and Shreve (1991), p.340) ■

Remark 88 *The above proposition raises the interesting issue of determining necessary and sufficient conditions for explosion of the solution X to SDE (6.13). Since Y given by (6.16)* *does not explode, and $Y_t = p(X_t)$, it is clear that the condition*

$$p(\pm\infty) = \pm\infty$$

guarantees that X is also non-exploding. However, this sufficient condition is unfortunately not necessary (see remark below). Instead, the necessary and sufficient condition used is Feller's test for explosion *(Feller (1952)).*

Remark 89 *Consider the case of*

$$\begin{aligned} b(x) &= sgn(x), \\ \sigma(x) &\equiv \sigma > 0. \end{aligned}$$

The scale function p, defined in (6.14), is bounded. According to Proposition (Strong Solution), the SDE (6.13) has non-exploding, unique strong solution for any initial distribution.

Theorem 90 (Engelbert & Schmidt (1984)) *For every initial distribution μ, the stochastic differential equation*

$$dX_t = \sigma(X_t) \cdot dW_t$$

has a solution which is unique in the sense probability law if and only if

$$I(\sigma) = Z(\sigma).$$

Proof. (Karatzas and Shreve (1991), p.335) ■

Theorem 91 (Weak Solution) *Assume that σ^{-2} is locally integrable at every point in $\mathbb{R}$, and conditions (ND) and (LI) holds. Then, for every initial distribution μ, the equation (6.13) has a weak solution up to an explosion time, and this solution is unique in the sense of probability law.*

Proof. Let $\widetilde{\sigma}$ be as defined by (6.18). According to theorem (Engelbert & Schmidt (1984)) and proposition (SDE solution), it suffices to prove that

$$I(\widetilde{\sigma}) = Z(\widetilde{\sigma}).$$

Now,

$$Z(\widetilde{\sigma}) = (p(-\infty), p(\infty))^c,$$

and $I(\widetilde{\sigma})$ contains this set. We must show that $\widetilde{\sigma}^{-2}$ is locally integrable at every point $y_0 \in (p(-\infty), p(\infty))$. At such a point y_0, choose $\epsilon > 0$ so that

$$p(-\infty) < y_0 - \epsilon < y_0 + \epsilon < p(\infty),$$

and write:

$$\int_{y=y_0-\epsilon}^{y_0+\epsilon} \frac{1}{\widetilde{\sigma}^2(y)} dy = \int_{x=q(y_0-\epsilon)}^{q(y_0+\epsilon)} \frac{1}{p'(x)\,\sigma^2(x)} dx.$$

The second integral is finite, because p' is bounded away from zero on finite intervals and σ^{-2} is locally integrable. ∎

Corollary 92 *Assume that σ^{-2} is locally integrable at every point in $\mathbb{R}$, and that conditions (ND), (LI), and*

$$\left.\begin{array}{l} |b(x) - b(y)| \leq K|x-y|, \text{ where } (x,y) \in \mathbb{R}^2 \\ |\sigma(x) - \sigma(y)| \leq h|x-y|, \text{ where } (x,y) \in \mathbb{R}^2 \end{array}\right\} \tag{6.19}$$

holds, where

$$K \text{ is a positive constant,}$$

and

$$h : [0,\infty) \to [0,\infty)$$

is a strictly increasing function for which $h(0) = 0$, and

$$\int_{(0,\epsilon)} h^{-2}(u)\,du = \infty, \ \forall \epsilon > 0$$

holds. Then, for every initial condition ξ independent of the driving Brownian motion $W = \{W_t, F_t : t \in [0,\infty)\}$, the one-dimensional SDE (6.13)

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t$$

has a unique strong solution (possibly up to an explosion time).

Proof. (Karatzas and Shreve (1991), p.341). ■

Proposition 93 (Strong Solution) *Assume that*

$$b : \mathbb{R} \to \mathbb{R} \text{ is bounded,}$$

and

$$\sigma : \mathbb{R} \to \mathbb{R} \text{ is Lipschitz continuous,}$$

with σ^2 bounded away from zero on every compact subset of $\mathbb{R}$. Then, for every initial condition ξ independent of the driving Brownian motion $W = \{W_t, F_t : t \in [0, \infty)\}$, the one-dimensional SDE (6.13)

$$dX_t = b(X_t)\, dt + \sigma(X_t)\, dW_t$$

has a non-exploding, unique strong solution.

Proof: Karatzas and Shreve (1991). First, the *boundedness of b prevents the explosion of any solution X* to SDE (6.13) needs to be shown. Lets fix $t \in [0, \infty)$.Let $n \to \infty$ in

$$\Pr\left[X_{t \wedge S_n} = X_0 + \int_{s=0}^{t} b(X_s)\, I_{\{s \leq S_n\}} ds + \int_{s=0}^{t} \sigma(X_s)\, I_{\{s \leq S_n\}} dW_s : t \in [0, \infty)\right] = 1. \tag{6.20}$$

Then, the Lebesgue integral

$$\int_{s=0}^{t} b(X_s)\, I_{\{s \leq S_n\}} ds \to \int_{s=0}^{t \wedge S} b(X_s)\, ds,$$

which is a finite expression as b is bounded. On the other hand, the stochastic integral

$$\int_{s=0}^{t} \sigma(X_s)\, I_{\{s \leq S_n\}} dW_s \to \int_{s=0}^{t \wedge S} \sigma(X_s)\, dW_s$$

on the event

$$A \equiv \left\{ \int_{s=0}^{t\wedge S} \sigma^2 (X_s)\, ds < \infty \right\},$$

and does not have a limit on A^c. It develops that the limit of the right-hand side of (6.20) exists, is finite $a.s.$ on A, and does not exist on A^c. On the left-hand side of (6.20), we have:

$$\lim_{n\to\infty} X_{t\wedge S_n} = X_{t\wedge S},\ a.s.$$

and is equal to $\pm\infty$ on $\{S < t\}$. It follows that

$$\Pr[S < t],\ \forall t > 0,$$

thus

$$\Pr[S = \infty] = 1.$$

Secondly, the *existence of a weak solution* for SDE (6.13) needs to be shown. The assumptions on b and σ imply (ND), (LI), and the local integrability of σ^{-2}. Thus, weak existence follows from theorem (Weak solution). According to proposition (SDE solution), the pathwise uniqueness of SDE (6.13) is equivalent to that of (6.16). Since

$$p'' \equiv -\left(\frac{2b}{\sigma^2}\right) p'$$

is locally bounded, p' is locally Lipschitz. It follows that $\widetilde{\sigma}(y)$ defined by (6.18) is locally Lipschitz at every point $y \in (p(-\infty), p(\infty))$.

Finally, the *pathwise uniqueness* of the solution needs to be shown. It has already been shown that any solution X to SDE (6.13) does not explode, thus for any solution Y to (6.16), (6.17) must remain in the interval $(p(-\infty), p(\infty))$. Under the above conditions, pathwise uniqueness holds for (6.16). We appeal to the corollary to Proposition of Yamada & Watanabe (1971) to conclude the argument. ■

Remark 94 *The linear growth condition*

$$|b(x)| + |\sigma(x)| \le K(1 + |x|),\ \forall x \in \mathbb{R}$$

is sufficient for $\Pr[S = \infty] = 1$.

6.3 Linearisation of the Solution to SDE

The aim of this section is to remove the drift from the SDE derived from Theory of Social Imitation and solve the resultant SDE.

First, define the following:

$$\begin{aligned}
a &= v(\sinh(kx+r) - 2x\cosh(kx+r)) \\
b &= \frac{v}{n}(\cosh(kx+r) - 2x\sinh(kx+r)) \\
c &= \frac{kv}{n}a - \frac{2v}{n}\sinh(r+kx) \\
d &= \frac{k^2v^2}{n}b - \frac{2kv(v+1)}{n}\cosh(r+kx) \\
Z &= \frac{b}{a} - \frac{1}{2}c \\
E &= \frac{d}{dx}Z \\
F &= \frac{d^2}{dx^2}Z \\
G &= \frac{d^3}{dx^3}Z
\end{aligned}$$

It is difficult to decide whether the following equation holds:

$$\frac{d}{dx}\left(\frac{\frac{d}{dx}(bE)}{E}\right) = 0.$$

6.3.1 Removing the Drift of the SDE

Based on the Theory of Social Imitation, the SDE derived has the form:

$$dr_t = \mu(r_t)dt + \sigma(r_t)dB_t, \tag{6.21}$$

where the drift coefficient function is:

$$\mu(r_t) = v\left[\sinh\left(\kappa r_t + \rho\right) - 2r_t \cosh\left(\kappa r_t + \rho\right)\right]$$

and the diffusion coefficient function is:

$$\sigma(r_t) = \frac{v}{n}\left(\cosh\left(\kappa r_t + \rho\right) - 2r_t \sinh\left(\kappa r_t + \rho\right)\right).$$

Recall that the domain of observed daily return is $r_t \in [-0.5, 0.5]$. First, the ND condition needs to be checked, i.e.

$$\sigma^2(r) > 0, \text{ for } r \in [-0.5, 0.5].$$

By definition, $\Phi(x)$ and $\Pi(x)$ is given by:

$$\Phi(x) \triangleq \cosh\left(\kappa x + \rho\right) - 2x \sinh\left(\kappa x + \rho\right),$$

and

$$\Pi(x) = \sinh\left(\kappa x + \rho\right) - 2x \cosh\left(\kappa x + \rho\right).$$

Thus, the ND condition of $\sigma(x)$ depends only on $\Phi(x)$.

Now, the behaviour of $\Phi(x)$ for a given pair (κ, ρ) needs to be investigated. Lets start by taking $(\kappa, \rho) = (1.8, 0.02)$, then:

$$\Phi(x|1.8, 0.02) = \cosh\left(1.8x + 0.02\right) - 2x \sinh\left(1.8x + 0.02\right),$$

and its graph is shown in Figure 6-1 below:

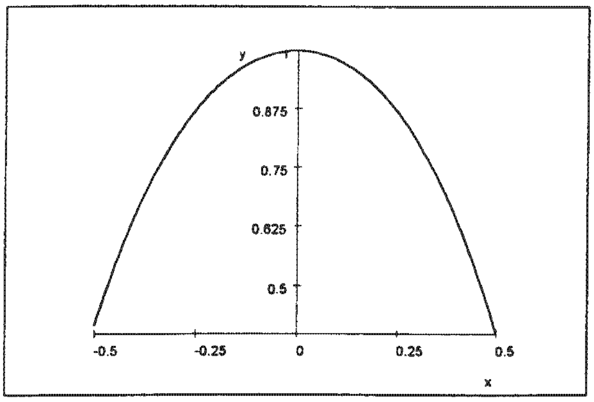

Figure 6-1: Plot of $\phi\left(x|\kappa=1.8,\rho=0.02\right)$

Similarly, take $(\kappa,\rho)=(-1.8,0.02)$, then:

$$\Phi(x)=\cosh\left(-1.8x+0.02\right)-2x\sinh\left(-1.8x+0.02\right),$$

and its graph is shown in Figure 6-2 below:

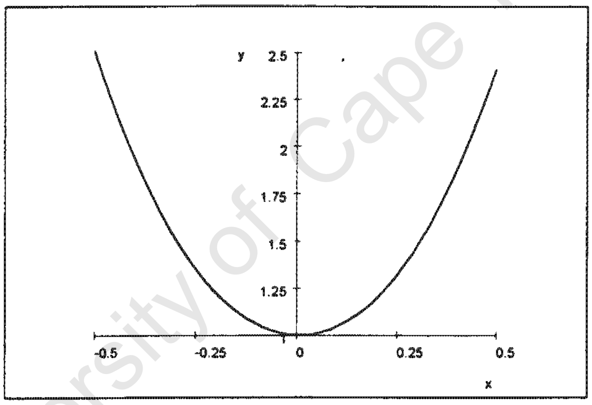

Figure 6-2: Plot of $\phi\left(x|-1.8,0.02\right)$

Figures 6-1 and 6-2 above suggests that

$$\sigma^2(x)>0,\text{ for }x\in[-0.5,0.5]$$

holds. Thus, ND condition of $\sigma(x)$ is justified.

Applying the triangle inequality to (6.22) above, the following is obtained:

$$|\mu(x)| \leq \frac{v}{\kappa}\left|\Phi'(x)\right| + \frac{2v}{\kappa}\left|\sinh\left(\kappa x + \rho\right)\right|. \tag{6.23}$$

Dividing both sides of (6.23) by $\sigma^2(y)$, then taking integral with respect to y from $x-\epsilon$ to $x+\epsilon$, get the following:

$$\begin{aligned}
\int_{y=x-\epsilon}^{x+\epsilon} \frac{|\mu(y)|}{\sigma^2(y)}dy &= \int_{y=x-\epsilon}^{x+\epsilon} \frac{\frac{v}{\kappa}\left|\Phi'(y) + 2\sinh\left(\kappa y + \rho\right)\right|}{\sigma^2(y)}dy \\
&\leq \frac{n^2}{\kappa v}\int_{x-\epsilon}^{x+\epsilon} \frac{|\Phi'(y)| + |2\sinh\left(\kappa y + \rho\right)|}{|\Phi(y)|^2}dy \\
&= \frac{n^2}{\kappa v}\left(\underbrace{\int_{x-\epsilon}^{x+\epsilon} \frac{|\Phi'(y)|}{\Phi(y)^2}dy}_{(1)} + \underbrace{\int_{x-\epsilon}^{x+\epsilon} \frac{|2\sinh\left(\kappa y + \rho\right)|}{|\Phi(y)|^2}dy}_{(2)}\right)
\end{aligned} \tag{6.24}$$

Now, term (1) is:

$$\begin{aligned}
\int_{x-\epsilon}^{x+\epsilon} \frac{|\Phi'(y)|}{\Phi^2(y)}dy &= \int_{x-\epsilon}^{x+\epsilon} \frac{\pm\Phi'(y)}{\Phi^2(y)}dy = \int_{x-\epsilon}^{x+\epsilon} \frac{\pm d\Phi(y)}{\Phi^2(y)} \\
&= \mp\int_{x-\epsilon}^{x+\epsilon} d\left(\frac{1}{\Phi(y)}\right) < \infty,
\end{aligned}$$

and term (2) is:

$$\int_{x-\epsilon}^{x+\epsilon} \frac{|2\sinh\left(\kappa y + \rho\right)|}{|\Phi(y)|^2}dy < \infty,$$

since

$$|\Phi(y)|^2 \geq |2\sinh\left(\kappa y + \rho\right)|.$$

Thus, it has been shown that

$$\int_{y=x-\epsilon}^{x+\epsilon} \frac{|\mu(y)|}{\sigma^2(y)}dy < \infty$$

Hence, the LI condition is satisfied. Thus, the drift in SDE (6.21) may be removed.

6.3.2 Deriving An Expression for the Solution to SDE

Define the *scale function*

$$p(x) = \int_{s=c}^{x} \exp\left(-2\int_{u=c}^{s} \frac{\mu(u)}{\sigma^2(u)}du\right) ds.$$

Then, the first ordinary derivative of $p(x)$ is given by:

$$p'(x) = \exp\left(-2\int_{c}^{x} \frac{\mu(s)}{\sigma^2(s)}ds\right),$$

and the second ordinary derivative of $p(x)$ is given by:

$$p''(x) = -\left[\frac{2\mu(s)}{\sigma^2(s)}\right] \cdot \exp\left(-2\int_{c}^{x} \frac{\mu(s)}{\sigma^2(s)}ds\right).$$

Now, define the continuous differential inverse function of $p(\cdot)$ by:

$$q(y) = \int_{0}^{y} \frac{1}{p'\left[q(u)\right]}du = \inf\left\{u : p(u) > y\right\}.$$

Then, the first ordinary derivative of $q(y)$ is:

$$q'(y) = \frac{1}{p'\left[q(y)\right]},$$

and the second ordinary derivative of $q(y)$ is:

$$q''(y) = \frac{2\mu\left[q(y)\right]}{\sigma^2\left[q(y)\right]} \exp\left(4\int_{c}^{q(y)} \frac{\mu(s)}{\sigma^2(s)}ds\right), a.e.$$

The solution U_t to the SDE:

$$dU_t = \tilde{\sigma}(U_t)dB_t$$

is easily shown to be:

$$U_t = U_0 + \int_0^t \widetilde{\sigma}(U_s)dB_s,$$

where:

$$p(-0.5) \leq U_0 \leq p(0.5), \; a.s.$$

and

$$\widetilde{\sigma}(y) = \begin{cases} p'(q(y))\sigma(q(y)) & , \text{for } p(-.5) \leq y \leq p(.5) \\ 0 & , \text{otherwise} \end{cases}.$$

Therefore, the solution to the SDE

$$dr_t = \mu(r_t)dt + \sigma(r_t)dB_t$$

is:

$$r_t = q\left(p(r_0) + \int_0^t \widetilde{\sigma}\left[p(r_s)\right] dB_s\right).$$

Thus, $\{r_t, t \in \mathbb{R}^+\}$ is a solution to SDE (6.21).

6.3.3 Condition for Solution of SDE to be a Brownian Motion

Based on the following theorem:

Theorem 95 *An Itô process*

$$\begin{aligned} dY_t &= vdB_t; \\ Y_0 &= 0, \; v(t,\omega) \in V_{\mathcal{H}}^{n\times m} \end{aligned}$$

coincide (in law) with n-dimensional Brownian motion if and only if

$$vv^T = \mathbf{I}_n \text{ for } a.a.(t,\omega) \text{ w.r.t. } dt \times dP,$$

where $\mathbf{I}_n$ is the n-dimensional identity matrix.

It follows that for $\{r_t, t \in \mathbb{R}^+\}$ to be a Brownian motion, the following equality must hold:

$$[\widetilde{\sigma}(p(x))]^2 = 1.$$

Now, the aim is to identify the parameter combination (κ, ρ, n, v) such that the above equality holds, which is a problem in seeking solutions of (κ, ρ, n, v) to the following equation:

$$\widetilde{\sigma}\left[p(x; \kappa, \rho, n, v)\right] = 1.$$

The above equation may be empirically solved using Matlab. Then, compute contour plots to identify the relationship between the (κ, ρ, n, v) combinations and each of the five market states.

6.3.4 Conditions for An Itô Process To Be A Diffusion Process

The Itô formula states that if one apply a C^2 function:

$$\phi : U \subset \mathbb{R}^n \to \mathbb{R}^n$$

to an Itô process X_t, then the resultant functional:

$$\phi\left(X_t\right)$$

is another Itô process. The aim of this section is to investigate if X_t is an Itô diffusion process, then under what conditions will $\phi\left(X_t\right)$ be an Itô diffusion process too.

One way of seeing that the process Y_t coincides in law with one-dimensional Brownian motion $\widetilde{B_t}$ is to apply the following results from Øksendal (1998):

Theorem 96 (8.4.2) *Let* $\nu\left(t, \omega\right) \in V_H^{n \times m}$. *An Itô process:*

$$\begin{cases} dY_t = \nu dB_t \\ Y_0 = 0 \end{cases}$$

coincides (in law) with n-dimensional Brownian motion if and only if

$$\nu \cdot \nu^{tr}(t,\omega) = I_n, \text{ for a.a. } (t,\omega) \text{ w.r.t. } dt \times dP, \tag{6.25}$$

where I_n is the n-dimensional identity matrix.

Theorem 97 (8.4.3) *Let X_t be an Itô diffusion process given by:*

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dB_t,$$

where:$b \in \mathbb{R}^n, \sigma \in \mathbb{R}^{n\times m}, X_0 = x$, and let Y_t be an Itô process given by:

$$dY_t = u(t,\omega)\,dt + \nu(t,\omega)\,dB_t,$$

where:$u \in \mathbb{R}^n, \nu \in \mathbb{R}^{n\times m}, Y_0 = x$. Then, $X_t \simeq Y_t$ if and only if

$$E^x[u(t,\cdot)|\, N_t] = b(Y_t^x),$$

and

$$\nu \cdot \nu^{tr}(t,\omega) = \sigma \cdot \sigma^{tr}(Y_t^x) \tag{6.26}$$

for a.a. (t,ω) w.r.t. $dt \times dP$, where N_t is the σ-algebra generated by Y_s $(s \leq t)$.

Lemma 98 (8.4.4) *Let*

$$\begin{aligned} dY_t &= u(t,\omega)\,dt + \nu(t,\omega)\,dB_t, \\ Y_0 &= x \end{aligned}$$

be as in Theorem 6.26. Then, there exists an N_t-adapted process $W(t,\omega)$ such that

$$\nu \cdot \nu^{tr}(t,\omega) = W(t,\omega), \text{ for a.a. } (t,\omega).$$

Corollary 99 (How to recognise a Brownian motion) *Let*

$$dY_t = u(t,\omega)\,dt + \nu(t,\omega)\,dB_t$$

be an Itô process in $\mathbb{R}^n$. *Then,* Y_t *is a Brownian motion if and only if*

$$E^x[u(t,\cdot)|N_t] = 0,$$

and

$$\nu \cdot \nu^{tr}(t,\omega) = I_n, \tag{6.27}$$

for a.a. (t,ω).

Chapter 7

CONCLUDING DISCUSSIONS

CMH model is based on applying the Social Imitation Theory (an idea which stemmed from statistical physics) to market dynamics research. The CMH model has two distinct characteristics:

1. The model includes one parameter quantifying the market sentiment (investor psychology), and the other parameter quantifying the market's economic fundamentals, and

2. The model is able to identify the different states (Vaga (1990) have identified five market states) in the market dynamic, which is described by distinct combinations of values for control parameters. The Gaussian market is merely one of the five identified market states. The market dynamic makes transitions from one state to another according to the market sentiments and economic fundamentals.

Hence, the CMH have explicitly modelled both the investors' behaviour and the economic fundamentals in the financial market. In this respect, the CMH distribution have also decomposed financial returns into economic fundamentals and market sentiment. Other advantages of the CMH distribution are as follows:

1. Its probability density function has a closed form, which makes parametric statistical inference on data easier.

2. Its control parameters have financial interpretation, which directly describes the state of the market dynamic.

3. It allows one to explain, at least in a qualitative manner, financial market movements, which is essentially an aggregation of human behaviour. However, such statistical descriptions certainly do not allow unique predictions due to the stochastic nature of the process described.

4. It can assist in the understanding of general features of cooperative behaviour of investors, even though the behaviour of each individual investor may be extremely complicated and not accessible to a mathematical description.

5. It justifies the use of both technical analysis and fundamental analysis of financial market. The effectiveness of these analyses in identifying trends and values are dependent on which state the market dynamic is in.

6. It provides a very general description of the financial market. It incorporates almost all of the important properties of financial market, namely: existence of short-term trends (technical analysis), importance of the fundamental value (fundamental analysis), investor sentiment (behavioural financial modelling), long-term memory and fractal structure (FMH), non-linearity, heavy-tailed return distribution, scaling, self-similarity (stable processes), and random walk and martingale properties (Gaussian processes). Finally, CMH is able to describe the transition between these important properties by borrowing the idea of phase transition and critical phenomena from statistical physics. This makes CMH the most flexible of all parametric non-linear financial models.

This dissertation have carried out all the necessary statistical analyses required to achieve objectives 1 to 7 of this dissertation as specified in the introduction. Based on the results of the statistical analyses, the following remarks on the properties of the CMH distribution may be made:

1. The CMH distribution of Vaga (1990) is indeed a model of the annualised returns. This evidenced by the difficulties encountered in obtaining a convergent result in the optimisation procedure. Furthermore, the fundamental and sentiment parameter values for CMH distribution fitted to partitioned daily returns in each state are quite different from those suggested by Vaga (1990). Since annualised returns have much larger values and are far

less volatile than daily returns, thus the substantially different parameter values obtained is expected. However, the large negative values obtained for the sentiment parameter κ forces one to re-think its financial interpretation.

2. Degrees of freedom n have varied wildly with changes in (κ, ρ). Considering this together with Vaga (1990)'s interpretation of number of industrial groups leads to the suggestion that n is describing an important factor that determines market state.

3. The first moments of the CMH distribution is non-linearly related to its control parameter. Thus, the CMH distribution is better able to capture observed non-linearities in market dynamics.

4. Fitting stable distribution to observed daily returns have supported the properties of the fitted CMH distribution. This also highlights the advantage of the CMH distribution in that its control parameters have financial interpretation that directly describes particular market states. In contrast, parameters of stable distribution does not directly describe the market state, but merely provide an idea of the return distribution shape.

5. The distribution form proposed by Vaga (1990) is a limiting (steady state) distribution. At best, only the time-homogeneous distribution may be derived from the diffusion process with state-dependent drift and diffusion coefficients. Since daily returns are not time-homogeneous and is non-stationary in nature, thus it is difficult to obtain parameter estimates using the classical maximum likelihood method.

6. It is more descriptive to regard physical temperature in the Ising model as "social atmosphere" for the financial market. Temperature is a very important factor that determines phase transition in the Ising model. Since the CMH distribution is derived from Theory of Social Imitation based on analogies made between physical system and financial market, thus the analogies should be made in greater detail.

Bibliography

[1] Aase, K.K. (2000), "An Equilibrium Asset Pricing Model Based on Lévy Processs: Relations to Stochastic Volatility, and Survival Hypothesis," *Insurance: Mathematics and Economics*, Amsterdam: Elsevier B.V., Vol. 27, pp. 345-363.

[2] Abu-Mostafa, Y.S. (1995), "Financial Market Applications of Learning From Hints," in A.-P. Refeness (ed.), *Neural Networks in the Capital Markets*, Chichester: John Wiley & Sons, pp. 221-232.

[3] Bachelier, L. (1900), "Théorie de la Spéculation", *Annales de l'Ecole Normale Superieure*, Series 3, Vol. 17, pp. 21-86.

[4] Barnett, W.A., Serletis, A. (2000), "Martingales, Nonlinearity, and Chaos," *Journal of Economic Dynamics & Control*, Elsevier Science B.V., Vol. 24, pp. 703-724.

[5] Black, F., Scholes, M. (1973), "The Pricing of Options and Corporate Liabilities," *The Journal of Political Economy*, Chicago: The University of Chicago Press, Vol. 81 (3), pp. 637-654.

[6] Bochner, S. (1955), *Harmonic Analysis and the Theory of Probability*, Berkeley: University of California Press.

[7] Box, G.E.P., Jenkins, G.M., Reinsel, G.C. (1994), *Time Series Analysis: Forecasting and Control*, Third Edition, Englewood Cliffs, NJ: Prentice-Hall, Inc..

[8] Breiman, L. (1968), *Probability*, Reading,Mass.:Addison-Wesley.

[9] Brockwell, P.J., Davis, R.A. (1991), *Time Series: Theory and Methods*, Second Edition, New York: Springer-Verlag New York, Inc..

[10] Callen, E., Shapero, D. (1974), "A Theory of Social Imitation", *Physics Today*, New York: American Institute of Physics, Inc., Vol. 27, No. 7, pp. 23-28.

[11] Campbell, J.Y., Lo, A.W., MacKinlay, A.C. (1997), *The Econometrics of Financial Markets*, Princeton, NJ: Princeton University Press.

[12] Chatfield, C. (2001), *Time-Series Forecasting*, Boca Raton: Chapman & Hall/CRC.

[13] Cheng, B., Titterington, D. (1994), "Neural Networks: A Review From Statistical Perspective," *Statistical Science*, Beachwood, OH: Institute of Mathematical Statistics, Vol. 9, pp. 2-54.

[14] Cont, R., Potters, M., Bouchaud, J.-P. (1997), "Scale Invariance and Beyond," *Proceedings of the CNRS Workshop on Scale Invariance*, EDP-Springer.

[15] Doob, J.L., (1953), *Stochastic Processes*, New York: John Wiley & Sons, Inc..

[16] De Bondt, W.F.M., Thaler, R.H. (1995), "Financial Decision-Making in Markets and Firms: A Behavioural Perspective," in Jarrow, R.A., Maksimovic, V., Ziemba, W.T. (ed.), *Handbooks in Operations Research and Management Science, Volume 9: Finance*, Amsterdam: Elsevier Science B.V., pp.385-410.

[17] Easley, D., O'Hara, M. (1995), "Market Microstructure," in Jarrow, R.A., Maksimovic, V., Ziemba, W.T. (ed.), *Handbooks in Operations Research and Management Science, Volume 9: Finance*, Amsterdam: Elsevier Science B.V., pp.357-383.

[18] Engelbert, H.J., Schmidt, W. (1984), "On One-dimensional Stochastic Differential Equations with Generalized Drift," *Lecture Notes in Control and Information Sciences*, Berlin: Springer-Verlag, Vol. 69, pp. 143-155.

[19] Fama, E.F. (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, MA: Blackwell Publishing, Vol. 25 (2), pp. 383-417.

[20] Fama, E.F. (1991), "Efficient Capital Markets: II," *Journal of Finance*, MA: Blackwell Publishing, Vol. 46 (5), pp. 1575-1617.

[21] Feller, W. (1952), "The Parabolic Differential Equations and the Associated Semi-Groups of Transformations," *Annals of Mathematics*, NJ: Princeton University and Institute for Advanced Study, Vol. 55, pp. 468-519.

[22] Fuller, W.A. (1996), *Introduction to Statistical Time Series*, Second Edition, New York: John Wiley & Sons, Inc..

[23] Gillespie, D.T. (1992), *Markov Processes: An Introduction for Physical Scientists*, San Diego, California: Academic Press, Inc..

[24] Giordano, N.J. (1997), *Computational Physics*, New Jersey: Prentice-Hall, Inc..

[25] Girifalco, L.A. (2000), *Statistical Mechanics of Solids*, New York: Oxford University Press, Inc..

[26] Gnedenko, B.V., Kolmogorov, A.N. (1954), *Limit Distribution for Sums of Independent Variables*, Translated by K.L. Chung, Cambridge, Massachusetts: Addison-Wesley Press.

[27] Gouriéroux, C. (1997), *ARCH Models and Financial Applications*, New York: Springer-Verlag, Inc..

[28] Haken, H. (1975), "Cooperative Phenomena in Systems Far From Thermal Equilibrium and in Non Physical Systems", *Review of Modern Physics*, Vol. 47, No. 1.

[29] Haken, H. (1983), *Synergetics-An Introduction: Non-Equilibrium Phase Transitions and Self-Organization in Physics, Chemistry and Biology*, Third Revised and Enlarged Edition, Berlin Heidelberg: Springer-Verlag.

[30] Hamilton, J.D. (1994), *Time Series Analysis*, New Jersey: Princeton University Press.

[31] Hsu, W., Hsu, L.S., Tenoria, M.F. (1995), "A Neural Network Procedure for Selecting Predictive Indicators in Currency Trading," in A.-P. Refenes (ed.), *Neural Networks in the Capital Markets*, Chichester: John Wiley & Sons, pp. 245-257.

[32] Hurst, H.E. (1951), "Long Term Storage Capacity of Reservoirs", *Transactions of the American Society of Civil Engineers*, 116, pp.770-799.

[33] Hutchinson, J.A., Lo, A., Poggio, T. (1994), "A Nonparametric Approach to Pricing and Hedging Derivative Securities Via Learning Networks," *Journal of Finance*, MA: Blackwell Publishing, Vol. 99, pp. 851-889.

[34] Jarrow, R.A., Maksimovic, V., Ziemba, W.T. (1995), *Handbooks in Operations Research and Management Science, Volume 9: Finance*, Amsterdam: Elsevier Science B.V..

[35] Karatzas,I., Shreve, S.E. (1991), *Brownian Motion and Stochastic Calculus*, Second Edition, New York: Springer-Verlag,Inc..

[36] Kuan, C., Liu, T. (1995), "Forecasting Exchange rates Using Feedforward and Recurrent Neural Networks," *Journal of Applied Econometrics*, New York: John Wiley & Sons, Vol. 10, pp. 347-364.

[37] Larrain, M. (1991), "Empirical Tests of Chaotic Behaviour in a Nonlinear Interest Rate Model," *Financial Analyst Journal*, Vol. 47.

[38] LeRoy, S.F. (1989), "Efficient Capital Markets and Martingales," *Journal of Economic Literature*, Nashville, TN: American Economic Association, Vol. 27 (4), pp. 1583-1612.

[39] Lévy, P. (1924), "Théorie des erreurs la loi de Gauss et les lois exceptionelles," *Bulletin de la Société Mathématique de France*, Vol. 52, pp.49-85.

[40] Lévy, P. (1925), *Calcul des Probabilities*, Paris: Gauthier-Villars.

[41] Lo, A. (1991), "Long Term Memory in Stock Market Prices," *Econometrica*, Vol. 59, pp. 1279-1313.

[42] Lo, A.W., Mamaysky, H., Wang, J. (2000), "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation," *Journal of Finance*, MA: Blackwell Publishing, Vol. 55 (4), pp. 1705-1765.

[43] Mandelbrot, B. (1963), "The variation of certain speculative prices," *Journal of Business*, Chicago: The University of Chicago Press, Vol. 36, pp. 394-419.

[44] Matacz, A. (1998), "Financial Modelling and Option Pricing Theory with the Truncated Lévy Processes," *International Journal of Theoretical and Applied Finance*, World Scientific Publishing Company, Vol. 3 (1), pp. 143-160.

[45] Merton, R.C. (1973), "Theory of Rational Option Pricing," *Bell Journal of Economics & Management Science*, Santa Monica, CA: RAND Journal of Economics, Vol. 4 (1), pp. 141-183.

[46] Michael, J.R. (1983), "The Stabilized Probability Plot", *Biometrika*, Great Britain: Oxford Univerity Press, Vol. 70 (1), pp. 11-17.

[47] Noe, M. (1972), "The Calculation of Distributions of Two-Sided Kolmogorov-Smirnov Type statistics," *Annals of Mathematical Statistics*, OH: Institute of Mathematical Statistics, Vol. 43, pp. 58-64.

[48] Nolan, J.P. (2001), "Maximum Likelihood Estimation of Stable Parameters," in Barndorff-Nielsen, O.E., Mikosch, T. Resnick, S.I. (ed.), *Lévy Processes: Theory and Applications*, Boston: Birkhäuser

[49] Parzen, E. (1962), *Stochastic Processes*, San Francisco: Holden-Day, Inc..

[50] Peters, E.E. (1989), "Fractal Structure in the Capital Market," *Financial Analysts Journal*, July/August.

[51] Peters, E.E. (1994), *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics*, New York: John Wiley & Sons, Inc.

[52] Poulos, E. M. (1992), "Are There Persistent Cycles?", *Technical Analysis of Stocks & Commodities*, Technical Analysis, Inc. , Vol. 10, Chap. 9 (September), pp. 385-389.

[53] Øksendal, B. (1998), *Stochastic Differential Equations: An Introduction with Applications*, Fifth Edition, Berlin Heidelberg: Springer-Verlag.

[54] Qi, M. (1996), "Financial Applications of Artificial Neural Networks," in Maddala, G.S., Rao, C.R. (ed.), *Handbook of Statistics, Volume 14: Statistical Methods in Finance*, Amsterdam: Elsevier Science B.V., pp.529-552.

[55] Qi, M., Maddala, G.S. (1995), "Option Pricing Using ANN: The Case of S&P 500 Index Call Options," *Neural Networks in Financial Engineering: Proceedings of 3rd International Conference on Neural Networks in the Capital Markets*, London, pp. 78-91.

[56] Raible, S. (2000), *Lévy Processes in Finance: Theory, Numerics, and Empirical Facts*, PhD Thesis. Mathematics Faculty, Freiburg University, Freiburg.

[57] Risken, H. (1989), *The Fokker-Planck Equation: Methods of Solution and Applications*, Second Edition, Berlin Heidelberg: Springer Verlag.

[58] Salchenberger, L., Cinar, E., Lash, N. (1992), "Neural Networks: A New Tool for Predicting Bank Failures," *Decision Sciences*, Vol. 23, pp. 899-916.

[59] Samorodnitsky, G. and Taqqu, M.S. (1994), *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, New York: Chapman & Hall.

[60] Steiner, M., Wittkemper, H.-G. (1997), "Portfolio Optimization with a Neural Network Implementation of the Coherent Market Hypothesis", *European Journal of Operational Research*, Elsevier Science B.V., 100, pp.27-40.

[61] Tam, K.Y., Kiang, Y.M. (1992), "Managerial Application of Neural Networks: The Case of Bank Failure Predictions," *Management Science*, Institute for Operations Research and the Management Sciences, Vol. 38, pp. 926-947.

[62] Taylor, S.J. (1986), *Modelling Financial Time Series*, New York: John Wiley and Sons Ltd.

[63] Tong, H. (1990), *Non-Linear Time Series: A Dynamical Systems Approach*, New York: Oxford University Press, Inc..

[64] Tsay, R.S. (2002), *Analysis of Financial Time Series*, New York: John Wiley & Sons, Inc..

[65] Weidlich, W. (1971), "The Statistical Description of Polarization Phenomena in Society," *British Journal of Mathematical and Statistical Psychology*, Vol. 24, pp. 251-266.

[66] Wilk, M.B., Gnanadesikan, R. (1968), "Probability Plotting Methods for the Analysis of Data," *Biometrika*, Great Britain: Oxford University Press, Vol. 55, pp. 1-17.

[67] Vaga, T. (1990), "The Coherent Market Hypothesis", *Financial Analysts Journal*, Nov/Dec, pp. 36~49.

[68] Vaga, T. (1994), *Profiting From Chaos: Using Chaos Theory for Market Timing, Stock Selection, and Option Valuation*, New York: McGraw-Hill, Inc..

[69] Voit, J. (2001), *The Statistical Mechanics of Financial Markets*, Berlin Heidelberg: Springer -verlag.

[70] Yamada, T., Watanabe, S. (1971), "On the Uniqueness of Solutions of Stochastic Differential Equations," *Journal of Mathematics Kyoto University*, Kyoto: Kyoto University, Vol. 11, pp.155-167.

Appendix A

MATLAB CODES LISTINGS

This section of the appendix contains a listing of the Matlab codes written to perform the following:

1. Optimisation procedure for maximum likelihood estimation.
2. Calculation of the first four moments associated with CMH distribution in a particular state.
3. CMH distributional plots and Contour plots.

A.1 Maximum Likelihood Estimation

A.1.1 Main Optimisation Routine

```
%The crucial step in the maximum likelihood estimation procedure is constrained/unconstrained
    %nonlinear optimisation. But, nonlinear optimisation is very sensitive to
    %the quality of one's initial quess (unlike Newton-Raphson method).
    %Hence, to make the numerical scheme work, I first
    %look at the histogram of the data, and fiddle around with the parameters unitil the plot of
    %Vaga's pdf more or less resembles the shape of histogram. The key things to look out for are:
    %1. location of mode of histogram (indicated by rho parameter),and
    %2. the spread of return and 'height' of pdf (controlled by n and kappa parameters together)
```

```
%---------------------------------------------------------------
%incr=0.01;
%r=-1:incr:1;
%m = length(r);
%hist(r)
%---------------------------------------------------------------
%input the logreturn vector
logret = importdata('logret.txt');
state = importdata('state.txt');
%identifying which state the logreturn belongs to.
s=input('Please enter market state: ');
mkt = (state==s); %state=1 is RW
                  %state=2 is BULL
                  %State=3 is BEAR
                  %State=4 is CHAOS
                  %State=5 is TRANS
%Returns data dependent inputs for objective function
r = logret(mkt)';
m = length(r);
%Specifying initial guesses for parameter
k = input('Initial guess for kappa (sentiment) parameter: ');
rho = input('Initial guess for rho (fundamentals) parameter: ');
n = input('Initial guess for degrees of freedom for the market under study: ');
coeffs = [k,rho,n];
[coeffs,fval,exitflag,output] = fmincon(@objfunc,coeffs,[],[],[],[],[],[],@nonlcon,[], m,r)
```

```
function y = objfunc(coeffs,m,r)
k = coeffs(1);
rho = coeffs(2);
n = coeffs(3);
```

```
%first_two = m*log(n) - m*log(C(k,rho,n));
const_one = m*log(n);
const_two = -m*log(C(k,rho,n)); %Had warning:log of 0 for BEAR,CHAOS,TRANS
                                 %the smallest no. that Matlab can handle is
                                 %10^-323=9.8813e-324
sum = 0;
for i = 1:m
    if phi(r(i),k,rho) ~= 0 %phi(r(i),k,rho) > 0
       sum = sum - log(abs(phi(r(i),k,rho))) + 2*n*B(r(i),k,rho);
    else
      sum = sum - log(abs(phi(r(i),k,rho))+eps) + 2*n*B(r(i),k,rho);
    end
end
y = -(const_one + const_two + sum);
```

```
function [c, ceq] = nonlcon(coeffs,m,r)
k = coeffs(1);
rho = coeffs(2);
n = coeffs(3);
ceq = [];
c = -C(k,rho,n) + 2*eps;
```

```
function y = phi(r,k,rho)
%Calculates the value of the phi(r,k,rho) function
%one of the two fundamental function for Vaga pdf. The other being A(r,k,rho)
%used for: Plotting of Vaga pdf for annualised return
% : nonlinear optimisation procedure ('fmincon') to find mle for k,rho
y = cosh((k*r)+rho)-(2*r.*sinh((k*r)+rho));
```

```
function y = B(r,k,rho)
```

```
%input parameters are:r,k,rho
incr=(r+0.5)/1000; %Set integration precision at 1,000 steps.
%(10,000 steps was too time consuming:took 1 day!)
%Computing speed and accuracy are two important but competing objectives.
%Since B is in the power of exponential,thus C and objfunc will be
%quite sensitive to the accuracy of B.
%However, value of B does not change up to four decimal places
%for 10,000 or 100,000 steps.
%In this instance, computing speed should take preference.
%Hence, the 10,000 step integration precision.
u=-0.5:incr:r;
sum = 0;
for i = 1:length(u)
    sum = sum + (A(u(i),k,rho)*incr) ;
end
y = sum;
```

```
function y = C(k,rho,n)
%used to evalaute the normalising constant for Vaga pdf for annulaised return
%the parameters are: k, rho, n
%the integration runs over r=-0.5:0.001:0.5
%integral approximated simply as area under the curve, with step size=0.001
%excludes terms that phi(r,k,rho) ~= 0
incr = 0.01; %set integration precision at (0.5-(-0.5))/100 = 0.01
r = -0.5:incr:0.5;
sum = 0;
for i = 1:length(r)
   if phi(r(i),k,rho) ~= 0 %phi(r(i),k,rho) > 0
     sum = sum + (exp(2*n*B(r(i),k,rho))./phi(r(i),k,rho)*incr);
  else
```

```
        sum = sum + (exp(2*n*B(r(i),k,rho))./(phi(r(i),k,rho)+eps)*incr);
    end
end
y = n*sum;
```

A.1.2 Fundamental Functions

```
function y = A(u,k,rho)
%Calculates the value of the A(r,k,rho) function
%Accepts a ROW vector of inputs u=[]
%one of the two fundamental function for Vaga pdf. The other being phi(r,k,rho)
%used for: Plotting of Vaga pdf for annualised return
%           : nonlinear optimisation procedure ('fmincon') to find mle for k,rho
%Warning: due to the function defined as a fraction of pifunc over phi, need to beware
% of combinations of (u,k,rho) such that phi will take on very small value or zero
% (see zeros of phi)
%Attempted to overcome the zero of phi problem by the Matlab command:
%                          'phi=phi+(phi==0)*eps'
%            : will affect whether B(r,k,rho) is defined or not
pifunc=sinh((k*u)+rho)-(2*u.*cosh((k*u)+rho));
phiu=cosh((k*u)+rho)-(2*u.*sinh((k*u)+rho));
phiu=phiu+(phiu==0)*eps; %attempt to avoid division by zero
y = pifunc./phiu;
```

```
function y = A2(u,k,rho)
%warning: only works for u a constant. Does not work for u a vector, because phiu is a matrix!!
phiu = cosh((k*u)+rho)-(2*u.*sinh((k*u)+rho));
phiu = phiu+(phiu==0)*eps; %attempt to avoid division by zero
y = (1-(4*u.^2))./(phiu.^2);
function y = D(u,k,rho)
y = A(u,k,rho).*A2(u,k,rho);
```

```
function y = D(u,k,rho)
y = A(u,k,rho).*A2(u,k,rho);
```

A.1.3 First Partial Derivatives

Codes for $L1\left(\kappa,\rho,n\right)$

```
function y = L1(k,rho,n)
    %input the logreturn vector
    logret=importdata('logret.txt');
    state=importdata('state.txt');
    %identifying which state the logreturn belongs to.
    s=input('PLease enter the market state: ');
    mkt=(state==s);
    r=logret(mkt)';
    %need to specify k and rho according to which state is being fitted (see Vaga)
    k=input('Please enter value of kappa estimate: ');
    rho=input('Please enter value of rho estimate: ');
    n=input('Please enter the degrees of freedom: '); %n=degrees of freedom
    m=length(r); %m=no. of elements in the particular state vector
    const = -m*C1(k,rho,n)/C(k,rho,n);
    sum = 0;
    for i = 1:m

        sum = sum - (r(i)*A(r(i),k,rho)) + (2*n*B1(r(i),k,rho));

    end
    y = const + sum;
```

Codes for $L2\left(\kappa,\rho,n\right)$

```
function y = L2(k,rho,n)
```

```
%input the logreturn vector
logret=importdata('logret.txt');
state=importdata('state.txt');
%identifying which state the logreturn belongs to.
s=input('PLease enter the market state: ');
mkt=(state==s);
r=logret(mkt)';
%need to specify k and rho according to which state is being fitted (see Vaga)
k=input('Please enter value of kappa estimate: ');
rho=input('Please enter value of rho estimate: ');
n=input('Please enter the degrees of freedom: '); %n=degrees of freedom
m=length(r); %m=no. of elements in the particular state vector
const = -m*C2(k,rho,n)/C(k,rho,n);
sum = 0;
for i = 1:m
    sum = sum - (A(r(i),k,rho)) + (2*n*B2(r(i),k,rho));
end
y = const + sum;
```

Codes for potential function $B(r, \kappa, \rho)$

```
function y = B1(r,k,rho)
    incr=(r+0.5)/1000; %Set integration precision at 1,000 steps.
                       %(10,000 steps was too time consuming:took 1 day!)
                       %Computing speed and accuracy are two important but competing objectives.
                       %Since B is in the power of exponential,thus C and objfunc will be
                       %quite sensitive to the accuracy of B.
                       %However, value of B does not change up to four decimal places
                       %for 10,000 or 100,000 steps.
                       %In this instance, computing speed should take preference.
                       %Hence, the 10,000 step integration precision.
```

```
u=-0.5:incr:r;
sum = 0;
for i = 1:length(u)
   sum = sum + (u(i)*A2(u(i),k,rho)*incr) ;
end
y = sum;
```

```
function y = B2(r,k,rho)
incr = (r+0.5)/1000; %Set integration precision at 1,000 steps.
                     %(10,000 steps was too time consuming:took 1 day!)
                     %Computing speed and accuracy are two important but competing objectives.
                     %Since B is in the power of exponential,thus C and objfunc will be
                     %quite sensitive to the accuracy of B.
                     %However, value of B does not change up to four decimal places
                     %for 10,000 or 100,000 steps.
                     %In this instance, computing speed should take preference.
                     %Hence, the 10,000 step integration precision.
u = -0.5:incr:r;
sum = 0;
for i = 1:length(u)
   sum = sum + (A2(u(i),k,rho)*incr) ;
end
y = sum;
```

Codes for normalising constant $C(\kappa, \rho, n)$

```
function y = C1(k,rho,n)
   incr = 0.01; %set integration precision at (0.5-(-0.5))/100 = 0.01
   r = -0.5:incr:0.5;
   sum = 0;
   for i = 1:length(r)
```

```
  if phi(r(i),k,rho) ~= 0
    sum = sum + ((exp(2*n*B(r(i),k,rho))./phi(r(i),k,rho)).*diffC1(r(i),k,rho,n))*incr;
  else
    sum = sum + ((exp(2*n*B(r(i),k,rho))./(phi(r(i),k,rho)+eps)).*diffC1(r(i),k,rho,n))*incr;
  end
end
y = n*sum;
```

```
function y = C2(k,rho,n)
incr = 0.01; %set integration precision at (0.5-(-0.5))/100 = 0.01
r = -0.5:incr:0.5;
sum = 0;
for i = 1:length(r)
  if phi(r(i),k,rho) ~= 0
    sum = sum + ((exp(2*n*B(r(i),k,rho))./phi(r(i),k,rho)).*diffC2(r(i),k,rho,n))*incr;
  else
    sum = sum + ((exp(2*n*B(r(i),k,rho))./(phi(r(i),k,rho)+eps)).*diffC2(r(i),k,rho,n))*incr;
  end
end
y = n*sum;
```

```
function y=diffC1(r,k,rho,n)
pir=sinh((k*r)+rho)-(2*r.*cosh((k*r)+rho));
phir=cosh((k*r)+rho)-(2*r.*sinh((k*r)+rho));
phir=phir+(phir==0)*eps;
Ar=pir./phir;
y=(2*n*B1(r,k,rho))-(r.*Ar);
```

```
function y = diffC2(r,k,rho,n)
pir=sinh((k*r)+rho)-(2*r.*cosh((k*r)+rho));
```

```
phir=cosh((k*r)+rho)-(2*r.*sinh((k*r)+rho));
phir=phir+(phir==0)*eps;
Ar=pir./phir;
y = (2*n*B2(r,k,rho))-Ar;
```

A.1.4 Second Partial Derivatives

Codes for $L11\,(\kappa, \rho, n)$

```
function y = L11(k,rho,n)
    %input the logreturn vector
    logret=importdata('logret.txt');
    state=importdata('state.txt');
    %identifying which state the logreturn belongs to.
    s=input('PLease enter the market state: ');
    mkt=(state==s);
    r=logret(mkt)';
    %need to specify k and rho according to which state is being fitted (see Vaga)
    k=input('Please enter value of kappa estimate: ');
    rho=input('Please enter value of rho estimate: ');
    n=input('Please enter the degrees of freedom: '); %n=degrees of freedom
    m=length(r); %m=no. of elements in the particular state vector
    const = -m*(C11(k,rho,n)*C(k,rho,n)-(C1(k,rho,n)^2))/(C(k,rho,n)^2);
    sum = 0;
    for i = 1:m

        sum = sum - (r(i)^2*A2(r(i),k,rho)) + (2*n*B11(r(i),k,rho));

    end
    y = const + sum;
```

Codes for $L12(\kappa, \rho, n)$

```
function y = L12(k,rho,n)
    %input the logreturn vector
    logret=importdata('logret.txt');
    state=importdata('state.txt');
    %identifying which state the logreturn belongs to.
    s=input('PLease enter the market state: ');
    mkt=(state==s);
    r=logret(mkt)';
    %need to specify k and rho according to which state is being fitted (see Vaga)
    k=input('Please enter value of kappa estimate: ');
    rho=input('Please enter value of rho estimate: ');
    n=input('Please enter the degrees of freedom: '); %n=degrees of freedom
    m=length(r); %m=no. of elements in the particular state vector
    const = -m*(C12(k,rho,n)*C(k,rho,n)-(C1(k,rho,n)*C2(k,rho,n)))/(C(k,rho,n)^2);
    sum = 0;
    for i = 1:m

        sum = sum - (r(i)*A2(r(i),k,rho)) + (2*n*B12(r(i),k,rho));

    end
    y = const + sum;
```

Codes for $L22(\kappa, \rho, n)$

```
function y = L22(k,rho,n)
    %input the logreturn vector
    logret=importdata('logret.txt');
    state=importdata('state.txt');
    %identifying which state the logreturn belongs to.
    s=input('PLease enter the market state: ');
```

```
mkt=(state==s);
r=logret(mkt)';
%need to specify k and rho according to which state is being fitted (see Vaga)
k=input('Please enter value of kappa estimate: ');
rho=input('Please enter value of rho estimate: ');
n=input('Please enter the degrees of freedom: '); %n=degrees of freedom
m=length(r); %m=no. of elements in the particular state vector
const = -m*(C22(k,rho,n)*C(k,rho,n)-(C2(k,rho,n)^2))/(C(k,rho,n)^2);
sum = 0;
for i = 1:m

    sum = sum - (A2(r(i),k,rho)) + (2*n*B22(r(i),k,rho));

end
y = const + sum;
```

Codes for potential function $B(r, \kappa, \rho)$

```
function y = B11(r,k,rho)
incr = (r+0.5)/1000; %Set integration precision at 1,000 steps.
%(10,000 steps was too time consuming:took 1 day!)
%Computing speed and accuracy are two important but competing objectives.
%Since B is in the power of exponential,thus C and objfunc will be
%quite sensitive to the accuracy of B.
%However, value of B does not change up to four decimal places
%for 10,000 or 100,000 steps.
%In this instance, computing speed should take preference.
%Hence, the 10,000 step integration precision.
u = -0.5:incr:r;
sum = 0;
for i = 1:length(u)
```

```
   sum = sum + (u(i)^2*D(u(i),k,rho)*incr) ;
end
y = -2*sum;
```

```
function y = B12(r,k,rho)
incr = (r+0.5)/1000; %Set integration precision at 1,000 steps.
%(10,000 steps was too time consuming:took 1 day!)
%Computing speed and accuracy are two important but competing objectives.
%Since B is in the power of exponential,thus C and objfunc will be
%quite sensitive to the accuracy of B.
%However, value of B does not change up to four decimal places
%for 10,000 or 100,000 steps.
%In this instance, computing speed should take preference.
%Hence, the 10,000 step integration precision.
u = -0.5:incr:r;
sum = 0;
for i = 1:length(u)
   sum = sum + (u(i)*D(u(i),k,rho)*incr) ;
end
y = -2*sum;
```

```
function y = B22(r,k,rho)
incr = (r+0.5)/1000; %Set integration precision at 1,000 steps.
%(10,000 steps was too time consuming:took 1 day!)
%Computing speed and accuracy are two important but competing objectives.
%Since B is in the power of exponential,thus C and objfunc will be
%quite sensitive to the accuracy of B.
%However, value of B does not change up to four decimal places
%for 10,000 or 100,000 steps.
%In this instance, computing speed should take preference.
```

```
%Hence, the 10,000 step integration precision.
u = -0.5:incr:r;
sum = 0;
for i = 1:length(u)
  sum = sum + (D(u(i),k,rho)*incr) ;
end
y = -2*sum;
```

Codes for normalising constant $C(\kappa, \rho, n)$

```
function y = C11(k,rho,n)
  incr = 0.01; %set integration precision at (0.5-(-0.5))/100 = 0.01
  r = -0.5:incr:0.5;
  sum = 0;
  for i = 1:length(r)
    if phi(r(i),k,rho) ~= 0
      sum = sum + (exp(2*n*B(r(i),k,rho))./phi(r(i),k,rho))*(diffC1(r(i),k,rho,n)^2+diffC11(r(i),k,rho,n))*incr;
    else
      sum = sum + (exp(2*n*B(r(i),k,rho))./(phi(r(i),k,rho)+eps))*(diffC1(r(i),k,rho,n)^2+diffC11(r(i),k,rho,n))*incr;
    end
  end
  y = n*sum;
```

```
function y = C12(k,rho,n)
incr = 0.01; %set integration precision at (0.5-(-0.5))/100 = 0.01
r = -0.5:incr:0.5;
sum = 0;
for i = 1:length(r)
  if phi(r(i),k,rho) ~= 0
    sum = sum + (exp(2*n*B(r(i),k,rho))./phi(r(i),k,rho))*(diffC1(r(i),k,rho,n)*diffC2(r(i),k,rho,n)+diffC12(r(i),k,rho,n))*incr;
  else
```

```
    sum = sum + (exp(2*n*B(r(i),k,rho))./(phi(r(i),k,rho)+eps))*(diffC1(r(i),k,rho,n)*diffC2(r(i),k,rho,n)+diffC12(r(i),k,rho,n))*incr;
  end
end
y = n*sum;
```

```
function y = C22(k,rho,n)
incr = 0.01; %set integration precision at (0.5-(-0.5))/100 = 0.01
r = -0.5:incr:0.5;
sum = 0;
for i = 1:length(r)
  if phi(r(i),k,rho) ~= 0
    sum = sum + (exp(2*n*B(r(i),k,rho))./phi(r(i),k,rho))*(diffC2(r(i),k,rho,n)^2+diffC22(r(i),k,rho,n))*incr;
  else
    sum = sum + (exp(2*n*B(r(i),k,rho))./(phi(r(i),k,rho)+eps))*(diffC2(r(i),k,rho,n)^2+diffC22(r(i),k,rho,n))*incr;
  end
end
y = n*sum;
```

```
function y = diffC11(r,k,rho,n)
%warning: only works for r a constant. Does not work for r a vector, because A2r is a matrix!!
phir = cosh((k*r)+rho)-(2*r.*sinh((k*r)+rho));
phir = phir+(phir==0)*eps; %attempt to avoid division by zero
A2r=(1-(4*r.^2))./(phir^2);
y = (2*n*B11(r,k,rho))-(r^2*A2r);
```

```
function y = diffC12(r,k,rho,n)
%warning: only works for r a constant. Does not work for r a vector, because A2r is a matrix!!
phir = cosh((k*r)+rho)-(2*r.*sinh((k*r)+rho));
phir = phir+(phir==0)*eps; %attempt to avoid division by zero
A2r=(1-(4*r.^2))./(phir^2);
```

```
y = (2*n*B12(r,k,rho))-(r*A2r);
```

```
function y = diffC22(r,k,rho,n)
%warning: only works for r a constant. Does not work for r a vector, because A2r is a matrix!!
phir = cosh((k*r)+rho)-(2*r.*sinh((k*r)+rho));
phir = phir+(phir==0)*eps; %attempt to avoid division by zero
A2r=(1-(4*r.^2))./(phir^2);
y = (2*n*B22(r,k,rho))-A2r;
```

A.2 First Four Moments for CMH Returns Distribution

```
function y=moments(k,rho,n)
    %Calculates the first four measures of spread for Vaga's pdf
    %for annualised returns r=-0.5:0.001:0.5
    %three parameters needs to be specified: k,rho,n
    %outputs a row vector: [mean stddev skewcoeff kurtosiscoeff]
    incr=0.001;
    r=-0.5:incr:0.5; %size(r)=1001. But, r vector will have both f(-0.5) and f(0.5),
                        %which means that this area under the curve approx. is neither
                        %a LH sum or a RH sum. However, both these f values are zero,
                        %thus won't be a problem.
    r2=r.^2;
    r3=r.^3;
    r4=r.^4;
    const = n/C(k,rho,n);
    f = const*(exp(2*n*plotB(r,k,rho))./phi(r,k,rho));
    z=f';
    expect = r*z*incr;
    m2=r2*z*incr;
    spread=m2-(expect^2);
```

```
stdev=sqrt(spread);
m3=r3*z*incr;
skew=(m3-(3*m2*expect)+(2*expect^3))/stdev^3;
m4=r4*z*incr;
kurtosis=(m4-(4*m3*expect)+(6*m2*expect^2)-(3*expect^4))/stdev^4;
excess=kurtosis-3;
y=[expect stdev skew excess];
```

A.3 Graphical Plots

A.3.1 2D-Plots for CMH Returns Distribution $f_R(r|\kappa,\rho,n)$

```
function y=plotyrpdf(k,rho,n)
  %Plots Vaga's pdf for annualised returns over r = -0.5:0.001:0.5
  %The controlling parameters for the pdf are:
  %k = value of "sentiment" parameter,
  %rho = value of "fundamental" parameter ,
  %n = degrees of freedom for market unde study (Vaga interpret it as no. of industrial groups
  % in sample calculation and assumed to be 186 for US. SA is more like 50)
  %the component functions required is A,plotB,C
  r = -0.5:0.001:0.5;
  const = n/C(k,rho,n);
  y = const*(exp(2*n*plotB(r,k,rho))./phi(r,k,rho)); %used for plotting Vaga's pdf
  plot(r,y),grid;
  title('*** Vaga pdf for annualised returns***')
  ylabel('value of pdf')
  xlabel('annualised returns (r)')
```

```
function y = plotB(r,k,rho)
%plots the integral of A(u,k,rho) from -0.5 to r.
%Mainly used for the plotting of Vaga pdf for annualised return
```

```
%accepts a ROW vector of returns inputs r=[]
%other input parameters are: k, rho
intg=[];
for j=1:length(r)
     incr(j)=(r(j)+0.5)/1000;
     u=-0.5:incr(j):r(j);
    sum = 0;
    for i = 1:length(u)
        sum = sum + (A(u(i),k,rho)*incr(j)) ;
    end
    intg=[intg;sum];
end
y=intg'; %y must be a row vector, otherwise there will be a warning for 'matrix dimension must agree'
```

A.3.2 Contour Plots for First Four Moments of CMH Distribution

```
n = input('Please enter degrees of freedom: ');
    incr = 0.001;
    r = -0.5:incr:0.5;
    [k rho] = meshgrid(-190:5:10,-10:0.5:10);
    const = n./C(k,rho,n);
    sum1=0;
    sum2=0;
    sum3=0;
    sum4=0;
    for i=1:length(r)

     if phi(r(i),k,rho)~=0

       sum1 = sum1 + r(i)*incr*const.*(exp(2*n*B(r(i),k,rho))./phi(r(i),k,rho));
```

```
    sum2 = sum2 + (r(i)^2)*incr*const.*(exp(2*n*B(r(i),k,rho))./phi(r(i),k,rho));
    sum3 = sum3 + (r(i)^3)*incr*const.*(exp(2*n*B(r(i),k,rho))./phi(r(i),k,rho));
    sum4 = sum4 + (r(i)^4)*incr*const.*(exp(2*n*B(r(i),k,rho))./phi(r(i),k,rho));

  else

    sum1 = sum1 + r(i)*incr*const.*(exp(2*n*B(r(i),k,rho))./(phi(r(i),k,rho))+eps);
    sum2 = sum2 + (r(i)^2)*incr*const.*(exp(2*n*B(r(i),k,rho))./(phi(r(i),k,rho))+eps);
    sum3 = sum3 + (r(i)^3)*incr*const.*(exp(2*n*B(r(i),k,rho))./(phi(r(i),k,rho))+eps);
    sum4 = sum4 + (r(i)^4)*incr*const.*(exp(2*n*B(r(i),k,rho))./(phi(r(i),k,rho))+eps);

  end

end
m2=sum2;
m3=sum3;
m4=sum4;
expect = sum1;
figure
[C,h]=contour(k,rho,expect),grid
clabel(C,h)
title('*** 3D Contour Plot for Expected Annualised Return given (k,rho,80) ***')
xlabel('kappa')
ylabel('rho')
zlabel('Expected annualised return')
stdev = m2 - (expect.^2);
figure
[C,h]=contour(k,rho,stdev),grid
clabel(C,h)
title('*** 3D Contour Plot for Std Dev of Annualised Return given (k,rho,80) ***')
```

```
xlabel('kappa')
ylabel('rho')
zlabel('Std Dev of annualised return')
skew = (m3-(3*m2.*expect)+(2*expect.^3))./(stdev.^3);
figure
[C,h]=contour(k,rho,skew),grid
clabel(C,h)
title('*** 3D Contour Plot for Skewness of Annualised Return given (k,rho,80) ***')
xlabel('kappa')
ylabel('rho')
zlabel('Skewness Coefficient')
kurtosis = (m4-(4*m3.*expect)+(6*m2.*(expect.^2))-(3*expect.^4))./(stdev.^4);
excess = kurtosis-3;
figure
[C,h]=contour(k,rho,excess),grid
clabel(C,h)
title('*** 3D Contour Plot for Excess Kurtosis of Annualised Return given (k,rho,80) ***')
xlabel('kappa')
ylabel('rho')
zlabel('Excess Kurtosis')
```

A.3.3 3D Surface Plot for Log-Likelihood Function $l(\kappa, \rho, n|\mathbf{r})$

```
%Input variables are:k,rho,n,r,m
s=input('Please enter market state: ');
logret=importdata('logret.txt');
state=importdata('state.txt');
mkt=(state==s);
r=logret(mkt)';
m=length(r);
[k n]=meshgrid(-190:4:10,10:4:190); %[k n]=meshgrid(-190:1:10,10:0.85:180);
```

```
rho=input('Please specify value of rho (location of mode): ');
const_one = m*log(n);
const_two = -m*log(plotC(k,rho,n)); %Had warning:log of 0 for BEAR,CHAOS,TRANS
                                    %the smallest no. that Matlab can handle is
                                    %10^-323=9.8813e-324
sum = 0;
for i = 1:m

    if phi(r(i),k,rho) ~= 0

       sum = sum - log(abs(phi(r(i),k,rho))) + 2*n.*B(r(i),k,rho);

    else

       sum = sum - log(abs(phi(r(i),k,rho))+eps) + 2*n.*B(r(i),k,rho);

    end

end
y = const_one + const_two + sum;
mesh(k,n,y)
title('*** 3D Surface for Log-Likelihood Function l(k,n|rho=0) ***')
xlabel('kappa')
ylabel('n')
zlabel('Value of Log-Likelihood')
```

Appendix B

VBA Codes Listing

B.1 Codes for Computation of R/S Statistics

```
Sub RSanalysis()
    Dim i, j, k, l, m As Integer
    Dim A(20), x(5000), N(20), sumx(5000), mean(5000), SumDep(5000), AccumDep(5000) As Single
    Dim Max, Min, Maxi(5000), Mini(5000), xsq(5000), xsqsum(5000), sumxsq(5000), sumRS(5000) As
Single
    Dim stddev(10000), R(10000), RS(10000) As Single
    Worksheets("(RS) Analysis (ln)").Activate
    'reading in AR(1) residuals of changes in indices
    For i = 1 To 3800
    x(i) = Range("o27").Offset(i - 1, 0).Value
    Next
    'Loop to calculate the R/S Statistics
    For j = 1 To 18
    N(j) = Range("p2").Offset(j - 1, 0).Value
    A(j) = 3800 / N(j)
    'Calculate mean and standard deviation for the k-th subperiod each of length N(j)
    For k = 1 To A(j)
    'Loop for the l-th element within the k-th subperiod of length N()
```

```
sumx(k) = 0
xsqsum(k) = 0
For l = (1 + ((k - 1) * N(j))) To (N(j) * k)
sumx(k) = sumx(k) + x(l)
mean(k) = sumx(k) / N(j)
'Range("x2").Offset(k - 1, j - 1).Value = mean(k)
xsq(l) = x(l) ^2
xsqsum(k) = xsqsum(k) + xsq(l)
sumxsq(k) = sumx(k) ^2
stddev(k) = Sqr(((N(j) * xsqsum(k)) - (sumxsq(k))) / (N(j) * (N(j) - 1)))
'Range("aq2").Offset(k - 1, j - 1).Value = stddev(k)
Next
Next
'Calculate the re-scaled range for k-th subperiod each of length N(j)
sumRS(j) = 0
For k = 1 To A(j)
'Loop for the l-th element within the k-th subperiod of length N()
sumx(k) = 0
For l = (1 + ((k - 1) * N(j))) To (N(j) * k)
'Calculating accumulated departures from the mean
For m = (1 + ((k - 1) * N(j))) To l
SumDep(m) = SumDep(m - 1) + (x(m) - mean(k))
AccumDep(l) = SumDep(m)
If l = 1 + ((k - 1) * N(j)) Then
Max = AccumDep(l)
Min = AccumDep(l)
ElseIf AccumDep(l) > Max Then
Max = AccumDep(l)
ElseIf AccumDep(l) < Min Then
Min = AccumDep(l)
```

```
End If
Maxi(k) = Max
Mini(k) = Min
Next
Next
R(k) = Maxi(k) - Mini(k)
RS(k) = R(k) / stddev(k)
sumRS(j) = sumRS(j) + RS(k)
RS(j) = sumRS(j) / A(j)
Next
Range("q2").Offset(j - 1, 0).Value = RS(j)
Range("s2").Offset(j - 1, 0).Value = Log(RS(j))
Range("r2").Offset(j - 1, 0).Value = Log(N(j))
Next
End Sub
```

B.2 Codes for Computation of Expected R/S Values Under the Null Hypothesis

```
Sub SigTesting()
    Dim i, j As Integer
    Dim N(20), sqrt(10000), ERS(10000), sumsqrt(10000) As Single
    Worksheets("SigTesting (ln)").Activate
    'reading in values of N divisible into 3800
    For i = 1 To 18
    N(i) = Range("a2").Offset(i - 1, 0).Value
    sumsqrt(i) = 0
    pi = 3.141592654
    For j = 1 To (N(i) - 1)
    sqrt(j) = Sqr((N(i) - j) / j)
```

```
sumsqrt(i) = sumsqrt(i) + sqrt(j)
'Range("i2").Offset(i - 1, 0).Value = sumsqrt(i)
ERS(i) = (1 / (Sqr(((N(i) - 0.5) / N(i)) * (N(i) * (pi / 2))))) * sumsqrt(i)
Next
Range("b2").Offset(i - 1, 0).Value = ERS(i)
Range("c2").Offset(i - 1, 0).Value = Log(N(i))
Range("d2").Offset(i - 1, 0).Value = Log(ERS(i))
Next
End Sub
```

B.3 Codes for Computation of R/S Values for Raw (Undetrended) JSE OVERALL Indices

```
Sub RSanalysis()
Dim i, j, k, l, m As Integer
Dim A(20), x(5000), N(20), sumx(5000), mean(5000), SumDep(5000), AccumDep(5000) As Single
Dim Max, Min, Maxi(5000), Mini(5000), xsq(5000), xsqsum(5000), sumxsq(5000), sumRS(5000) As
Single
Dim stddev(10000), R(10000), RS(10000) As Single
Worksheets("Serial correlation").Activate
'reading in log first differences of changes in indices
For i = 1 To 3800
x(i) = Range("l27").Offset(i - 1, 0).Value
Next
'Loop to calculate the R/S Statistics
For j = 1 To 18
N(j) = Range("p2").Offset(j - 1, 0).Value
A(j) = 3800 / N(j)
'Calculate mean and standard deviation for the k-th subperiod each of length N(j)
For k = 1 To A(j)
```

```
'Loop for the l-th element within the k-th subperiod of length N()
sumx(k) = 0
xsqsum(k) = 0
For l = (1 + ((k - 1) * N(j))) To (N(j) * k)
sumx(k) = sumx(k) + x(l)
mean(k) = sumx(k) / N(j)
'Range("x2").Offset(k - 1, j - 1).Value = mean(k)
xsq(l) = x(l) ^2
xsqsum(k) = xsqsum(k) + xsq(l)
sumxsq(k) = sumx(k) ^2
stddev(k) = Sqr(((N(j) * xsqsum(k)) - (sumxsq(k))) / (N(j) * (N(j) - 1)))
'Range("aq2").Offset(k - 1, j - 1).Value = stddev(k)
Next
Next
'Calculate the re-scaled range for k-th subperiod each of length N(j)
sumRS(j) = 0
For k = 1 To A(j)
'Loop for the l-th element within the k-th subperiod of length N()
sumx(k) = 0
For l = (1 + ((k - 1) * N(j))) To (N(j) * k)
'Calculating accumulated departures from the mean
For m = (1 + ((k - 1) * N(j))) To l
SumDep(m) = SumDep(m - 1) + (x(m) - mean(k))
AccumDep(l) = SumDep(m)
If l = 1 + ((k - 1) * N(j)) Then
Max = AccumDep(l)
Min = AccumDep(l)
ElseIf AccumDep(l) > Max Then
Max = AccumDep(l)
ElseIf AccumDep(l) < Min Then
```

```
Min = AccumDep(l)
End If
Maxi(k) = Max
Mini(k) = Min
Next
Next
R(k) = Maxi(k) - Mini(k)
RS(k) = R(k) / stddev(k)
sumRS(j) = sumRS(j) + RS(k)
RS(j) = sumRS(j) / A(j)
Next
Range("r2").Offset(j - 1, 0).Value = RS(j)
'Range("s2").Offset(j - 1, 0).Value = Log(RS(j))
'Range("r2").Offset(j - 1, 0).Value = Log(N(j))
Next
End Sub
```

Appendix C

MLE OF CMH PARAMETERS BY FIRST PRINCIPLE

C.1 Detailed Derivations of the Partial Derivatives Required for MLE Procedure

Recall that under Vaga (1990)'s Coherent Market Hypothesis (CMH), the annualised market return is a random variable with the following probability density function (p.d.f.):

$$\begin{aligned} f_R(\mathbf{r}; n, \kappa, \rho) &= \text{p.d.f. of annualised market return } R\text{, given the parameter values } n,\kappa,\rho \\ &= c(\kappa, \rho, n) \cdot \Psi^{-1}(\mathbf{r}; \kappa, \rho, n) \cdot \exp\left[2 \cdot \int_{u=-\frac{1}{2}}^{\mathbf{r}} \frac{\Pi(u; \kappa, \rho)}{\Psi(u; \kappa, \rho, n)} du\right], \qquad \text{(C.1)} \end{aligned}$$

where if we first define the parameters to mean:

$$\begin{aligned} n &= \text{number of degrees of freedom} \\ \kappa &= \text{degree of market sentiment} \\ \rho &= \text{degree of fundamental bias,} \end{aligned}$$

then we can define the following:

$$\begin{aligned}
\Pi(u;\kappa,\rho) &= \sinh(\kappa u+\rho) - 2u\cosh(\kappa u+\rho)\,, \\
\Psi(u;\kappa,\rho,n) &= \frac{1}{n}\left[\cosh(\kappa u+\rho) - 2u\sinh(\kappa u+\rho)\right], \text{ and} \\
c^{-1}(\kappa,\rho,n) &= \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \left(\Psi^{-1}(r;\kappa,\rho,n)\cdot\exp\left[2\int_{u=-\frac{1}{2}}^{r} \frac{\Pi(u;\kappa,\rho)}{\Psi(u;\kappa,\rho,n)}du\right]\right) dr.
\end{aligned}$$

For a given realisation of random sample of size m log-return observations, and $n = n^{\star}$ some pre-determined fixed value, the likelihood function of parameters κ and ρ is given by:

$$\begin{aligned}
\mathbf{L}(\kappa,\rho) &= \prod_{i=1}^{m} f(\kappa,\rho;r_i,n^{\star}) = \prod_{i=1}^{m}\left(c(\kappa,\rho;n^{\star})\cdot\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\cdot\exp\left[2\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du\right]\right) \\
&= c(\kappa,\rho;n^{\star})^{m}\cdot\prod_{i=1}^{m}\left(\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\cdot\exp\left[2\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du\right]\right)
\end{aligned}$$

Then, the log-likelihood is a bivariate function of κ and ρ given by:

$$
\begin{aligned}
l(\kappa,\rho) &= \ln[\mathbf{L}(\kappa,\rho)] = \ln\left[\prod_{i=1}^{m} f(\kappa,\rho;r_i,n^\star)\right] \\
&= \ln\left[c(\kappa,\rho;n^\star)^m \cdot \prod_{i=1}^{m}\left(\Psi^{-1}(\kappa,\rho;r_i,n^\star)\cdot \exp\left[2\int_{u=-\frac{1}{2}}^{r_i} \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\right)\right] \\
&= \ln\left[c(\kappa,\rho;n^\star)^m\right] + \ln\left[\prod_{i=1}^{m}\left(\Psi^{-1}(\kappa,\rho;r_i,n^\star)\cdot \exp\left[2\int_{u=-\frac{1}{2}}^{r_i} \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\right)\right] \\
&= m\cdot\ln\left[c(\kappa,\rho;n^\star)\right] + \sum_{i=1}^{m}\ln\left(\Psi^{-1}(\kappa,\rho;r_i,n^\star)\cdot \exp\left[2\int_{u=-\frac{1}{2}}^{r_i} \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\right) \\
&= m\cdot\ln\left[c(\kappa,\rho;n^\star)\right] + \left\{\sum_{i=1}^{m}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)\right]\right\} + 2\sum_{i=1}^{m}\left[\int_{u=-\frac{1}{2}}^{r_i} \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right].
\end{aligned}
\tag{C.2}
$$

Based on the above log-likelihood function (C.2), we want to derive the maximum likelihood estimates (m.l.e.) and its associated variance as follows:

Firstly, we find the maximum likelihood estimate of κ, denoted $\widehat{\kappa}$, by taking the first partial derivative of C.2 with respect to κ as follows:

$$
\begin{aligned}
&\frac{\partial l(\kappa,\rho)}{\partial \kappa} \\
=\;& \frac{\partial}{\partial\kappa}\left(m\cdot\ln\left[c(\kappa,\rho;n^{\star})\right]+\left\{\sum_{i=1}^{m}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\right]\right\}+2\sum_{i=1}^{m}\left[\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du\right]\right) \\
=\;& \frac{\partial}{\partial\kappa}\left(m\cdot\ln\left[c(\kappa,\rho;n^{\star})\right]\right)+\frac{\partial}{\partial\kappa}\left(\sum_{i=1}^{m}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\right]\right) \\
&+\frac{\partial}{\partial\kappa}\left(2\sum_{i=1}^{m}\left[\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du\right]\right) \\
=\;& \underbrace{-m\left(\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^{\star})\right]\right)}_{(B.3.4)}+\sum_{i=1}^{m}\underbrace{\left(\frac{\partial}{\partial\kappa}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\right]\right)}_{(B.3.2)} \\
&+2\underbrace{\sum_{i=1}^{m}\left(\frac{\partial}{\partial\kappa}\left[\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du\right]\right)}_{(B.3.3)},
\end{aligned}
\tag{C.3}
$$

where partial derivative $(B.3.1)$ is given by:

$$
\begin{aligned}
\frac{\partial}{\partial\kappa}\Pi(\kappa,\rho;r_i) &= \frac{\partial}{\partial\kappa}\left[\sinh(\kappa r_i+\rho)-2r_i\cosh(\kappa r_i+\rho)\right] \\
&= r_i\cdot\left[\cosh(\kappa r_i+\rho)-2r_i\sinh(\kappa r_i+\rho)\right] \\
&= r_i n^{\star}\Psi(\kappa,\rho;r_i,n^{\star}),
\end{aligned}
$$

first partial derivative $(B.3.2)$ is given by:

$$
\begin{aligned}
\frac{\partial}{\partial\kappa}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\right] &= \Psi(\kappa,\rho;r_i,n^{\star})\cdot\frac{\partial}{\partial\kappa}\left[\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\right] \\
&= \Psi(\kappa,\rho;r_i,n^{\star})\cdot\left[-\frac{r_i}{n^{\star}}\cdot\frac{\Pi(\kappa,\rho;r_i)}{\Psi^{2}(\kappa,\rho;r_i,n^{\star})}\right] \\
&= -\frac{r_i}{n^{\star}}\cdot\left[\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^{\star})}\right],
\end{aligned}
$$

first partial derivative $(B.3.3)$ is given by:

$$
\begin{aligned}
\frac{\partial}{\partial\kappa}\left[\int_{u=-\frac{1}{2}}^{r_i} \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right] &= \int_{u=-\frac{1}{2}}^{r_i} \frac{\partial}{\partial\kappa}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]du \\
&= \int_{u=-\frac{1}{2}}^{r_i} \left[un^\star - \frac{u}{n^\star}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2\right]du \\
&= \frac{n^\star}{2}\left(r_i^2 - \frac{1}{4}\right) - \frac{1}{n^\star}\int_{u=-\frac{1}{2}}^{r_i} \left[u\cdot\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2\right]du,
\end{aligned}
$$

and first partial derivative $(B.3.4)$ is given by:

$$
\begin{aligned}
\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] &= c(\kappa,\rho;n^\star)\cdot\frac{\partial}{\partial\kappa}\left[c^{-1}(\kappa,\rho;n^\star)\right]\\
&= c(\kappa,\rho;n^\star)\cdot\frac{\partial}{\partial\kappa}\left[\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(\Psi^{-1}(\kappa,\rho;r,n^\star)\cdot\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\right)dr\right]\\
&= c(\kappa,\rho;n^\star)\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\frac{\partial}{\partial\kappa}\left(\Psi^{-1}(\kappa,\rho;r,n^\star)\cdot\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\right)\right]dr\\
&= c(\kappa,\rho;n^\star)\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\frac{\partial}{\partial\kappa}\Psi^{-1}(\kappa,\rho;r,n^\star)\cdot\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]dr\\
&\quad +c(\kappa,\rho;n^\star)\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)\cdot\frac{\partial}{\partial\kappa}\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]dr\\
&= -\frac{c(\kappa,\rho;n^\star)}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[r\left(\frac{\Pi(\kappa,\rho;r_i)}{\Psi^2(\kappa,\rho;r_i,n^\star)}\right)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\right]dr\\
&\quad +c(\kappa,\rho;n^\star)\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\left[2\frac{\partial}{\partial\kappa}\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]dr\\
&= -\frac{c(\kappa,\rho;n^\star)}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[r\left(\frac{\Pi(\kappa,\rho;r_i)}{\Psi^2(\kappa,\rho;r_i,n^\star)}\right)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\right]dr\\
&\quad +c(\kappa,\rho;n^\star)\,n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\left(r^2-\frac{1}{4}\right)dr\\
&\quad -\frac{2c(\kappa,\rho;n^\star)}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\int_{u=-\frac{1}{2}}^{r_i}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right]dr
\end{aligned}
$$

Then, to calculate the variance associated with $\widehat{\kappa}$, we need to derive the second partial derivative of C.2 with respect to κ as follows:

$$
\begin{aligned}
\frac{\partial^2 l(\kappa,\rho)}{\partial \kappa^2} &= \frac{\partial}{\partial \kappa}\left[\frac{\partial l(\kappa,\rho)}{\partial \kappa}\right] \\
&= \frac{\partial}{\partial \kappa}\left[\begin{array}{c} -m\left(\frac{\partial}{\partial \kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\right) + \sum_{i=1}^{m}\left(\frac{\partial}{\partial \kappa}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)\right]\right) \\ +2\sum_{i=1}^{m}\left(\frac{\partial}{\partial \kappa}\left[\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\right)\end{array}\right] \\
&= -m\underbrace{\left(\frac{\partial^2}{\partial \kappa^2}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\right)}_{(B.4.4)} + \sum_{i=1}^{m}\underbrace{\left(\frac{\partial^2}{\partial \kappa^2}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)\right]\right)}_{(B.4.2)} \\
&\quad +2\sum_{i=1}^{m}\underbrace{\left(\frac{\partial^2}{\partial \kappa^2}\left[\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\right)}_{(B.4.3)},
\end{aligned} \tag{C.4}
$$

where second partial derivative $(B.4.1)$ is given by:

$$
\begin{aligned}
\frac{\partial^2}{\partial \kappa^2}\Pi(\kappa,\rho;r_i) &= \frac{\partial}{\partial \kappa}\left[\frac{\partial}{\partial \kappa}\Pi(\kappa,\rho;r_i)\right] \\
&= \frac{\partial}{\partial \kappa}\left[r_i n^\star \Psi(\kappa,\rho;r_i,n^\star)\right] \\
&= r_i n^\star \cdot \left[\frac{\partial}{\partial \kappa}\Psi(\kappa,\rho;r_i,n^\star)\right] \\
&= r_i n^\star \cdot \left[\frac{r_i}{n^\star}\cdot \Pi(\kappa,\rho;r_i)\right] \\
&= r_i^2 \cdot \Pi(\kappa,\rho;r_i),
\end{aligned}
$$

second partial derivative $(B.4.2)$ is given by:

$$\begin{aligned}
\frac{\partial^2}{\partial \kappa^2} \ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)\right] &= \frac{\partial}{\partial \kappa}\left[\frac{\partial}{\partial \kappa}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)\right]\right] \\
&= \frac{\partial}{\partial \kappa}\left(-\frac{r_i}{n^\star}\cdot\left[\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)}\right]\right) \\
&= -\frac{r_i}{n^\star}\cdot\left[\frac{\partial}{\partial \kappa}\left(\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)}\right)\right] \\
&= -\frac{r_i}{n^\star}\cdot\left[r_i n^\star - \frac{r_i}{n^\star}\left(\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)}\right)^2\right] \\
&= -r_i^2 + \left(\frac{r_i}{n^\star}\cdot\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)}\right)^2,
\end{aligned}$$

second partial derivative $(B.4.3)$ is given by:

$$\frac{\partial^2}{\partial\kappa^2}\left[\int\limits_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]=\frac{\partial}{\partial\kappa}\left[\frac{\partial}{\partial\kappa}\left(\int\limits_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\right]$$

$$= \frac{\partial}{\partial\kappa}\left[\frac{n^\star}{2}\left(r_i^2-\frac{1}{4}\right)-\frac{1}{n^\star}\int\limits_{u=-\frac{1}{2}}^{r_i}\left[u\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2\right]du\right]$$

$$= -\frac{1}{n^\star}\cdot\left(\int\limits_{u=-\frac{1}{2}}^{r_i}u\frac{\partial}{\partial\kappa}\left[\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2\right]du\right)$$

$$= -\frac{1}{n^\star}\left(\int\limits_{u=-\frac{1}{2}}^{r_i}u\left[2\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)\frac{\partial}{\partial\kappa}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)\right]du\right)$$

$$= -\frac{1}{n^\star}\int\limits_{u=-\frac{1}{2}}^{r_i}2u\left[\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)\left(un^\star-\frac{u}{n^\star}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2\right)\right]du$$

$$= -\frac{2}{n^\star}\int\limits_{u=-\frac{1}{2}}^{r_i}u\left[\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)un^\star-\frac{u}{n^\star}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^3\right]du$$

$$= -\frac{2}{n^\star}\int\limits_{u=-\frac{1}{2}}^{r_i}\left[u^2n^\star\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)-\frac{u^2}{n^\star}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^3\right]du$$

$$= -2\int\limits_{u=-\frac{1}{2}}^{r_i}u^2\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]du+\frac{2}{(n^\star)^2}\int\limits_{u=-\frac{1}{2}}^{r_i}u^2\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^3du,$$

and second partial derivative $(B.4.4)$ is given by:

$$\frac{\partial^2}{\partial\kappa^2}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] = \frac{\partial}{\partial\kappa}\left[\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\right] = \frac{\partial}{\partial\kappa}\left[c(\kappa,\rho;n^\star)\cdot\frac{\partial}{\partial\kappa}c^{-1}(\kappa,\rho;n^\star)\right]$$

$$= \left[\frac{\partial}{\partial\kappa}c(\kappa,\rho;n^\star)\cdot\frac{\partial}{\partial\kappa}c^{-1}(\kappa,\rho;n^\star)\right] + \left[c(\kappa,\rho;n^\star)\cdot\frac{\partial^2}{\partial\kappa^2}c^{-1}(\kappa,\rho;n^\star)\right]$$

$$= c(\kappa,\rho;n^\star)\left\{\left[\frac{\frac{\partial}{\partial\kappa}c(\kappa,\rho;n^\star)}{c(\kappa,\rho;n^\star)}\cdot\frac{\partial}{\partial\kappa}c^{-1}(\kappa,\rho;n^\star)\right] + \left[\frac{\partial^2}{\partial\kappa^2}c^{-1}(\kappa,\rho;n^\star)\right]\right\}$$

$$= c(\kappa,\rho;n^\star)\left\{\left[-\frac{\partial}{\partial\kappa}\ln c^{-1}(\kappa,\rho;n^\star)\cdot\frac{\partial}{\partial\kappa}c^{-1}(\kappa,\rho;n^\star)\right] + \left[\frac{\partial^2}{\partial\kappa^2}c^{-1}(\kappa,\rho;n^\star)\right]\right\}$$

$$= c(\kappa,\rho;n^\star)\left\{\left[-c(\kappa,\rho;n^\star)\frac{\partial}{\partial\kappa}c^{-1}(\kappa,\rho;n^\star)\cdot\frac{\partial}{\partial\kappa}c^{-1}(\kappa,\rho;n^\star)\right] + \left[\frac{\partial^2}{\partial\kappa^2}c^{-1}(\kappa,\rho;n^\star)\right]\right\}$$

$$= -\left[\underbrace{\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]}_{(B.3.4)}\right]^2 + c(\kappa,\rho;n^\star)\underbrace{\left[\frac{\partial^2}{\partial\kappa^2}c^{-1}(\kappa,\rho;n^\star)\right]}_{(B.4.5)}.$$

Now, deriving the second partial derivative $(B.4.5)$ will take some work:

$$\frac{\partial^2}{\partial\kappa^2}c^{-1}(\kappa,\rho;n^\star) = \frac{\partial}{\partial\kappa}\left[\frac{\partial}{\partial\kappa}c^{-1}(\kappa,\rho;n^\star)\right]$$

$$= \underbrace{\frac{\partial}{\partial\kappa}\left[-\frac{1}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}} r\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)dr\right]}_{(1)}$$

$$\underbrace{+\frac{\partial}{\partial\kappa}\left[n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\left(r^2-\frac{1}{4}\right)dr\right]}_{(2)}$$

$$\underbrace{+\frac{\partial}{\partial\kappa}\left[-\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\int_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right)dr\right.}_{(3)}$$

Lets derive each of (1) $\sim$ (3) individually. For (1), we have:

$$\frac{\partial}{\partial\kappa}\left[-\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} r\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)dr\right]$$

$$= -\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} r\cdot\frac{\partial}{\partial\kappa}\left[\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\right]dr$$

$$= -\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} r\cdot\frac{\partial}{\partial\kappa}\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)dr$$

$$-\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} r\cdot\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\frac{\partial}{\partial\kappa}\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)dr,$$

where:

$$\frac{\partial}{\partial\kappa}\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right) = \frac{\partial}{\partial\kappa}\left[\Pi(\kappa,\rho;r)\cdot\Psi^{-2}(\kappa,\rho;r,n^\star)\right]$$

$$= \left[\frac{\partial}{\partial\kappa}\Pi(\kappa,\rho;r)\cdot\Psi^{-2}(\kappa,\rho;r,n^\star)\right]+\left[\Pi(\kappa,\rho;r)\frac{\partial}{\partial\kappa}\Psi^{-2}(\kappa,\rho;r,n^\star)\right]$$

$$= \left[rn^\star\Psi(\kappa,\rho;r,n^\star)\Psi^{-2}(\kappa,\rho;r,n^\star)\right]+\left[\Pi(\kappa,\rho;r)\frac{\partial}{\partial\kappa}\left(\Psi^{-1}(\kappa,\rho;r,n^\star)\right)^2\right]$$

$$= \frac{rn^\star}{\Psi(\kappa,\rho;r,n^\star)}+\left[\Pi(\kappa,\rho;r)2\Psi^{-1}(\kappa,\rho;r,n^\star)\frac{\partial}{\partial\kappa}\Psi^{-1}(\kappa,\rho;r,n^\star)\right]$$

$$= \frac{rn^\star}{\Psi(\kappa,\rho;r,n^\star)}+2\left[\Pi(\kappa,\rho;r)\Psi^{-1}(\kappa,\rho;r,n^\star)\left(-\frac{r}{n^\star}\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\right]$$

$$= \frac{rn^\star}{\Psi(\kappa,\rho;r,n^\star)}-\frac{2r}{n^\star}\left[\frac{\Pi^2(\kappa,\rho;r)}{\Psi^3(\kappa,\rho;r,n^\star)}\right]$$

$$= rn^\star\Psi^{-1}(\kappa,\rho;r,n^\star)-\frac{2r}{n^\star}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)A^2(\kappa,\rho;r,n^\star)\right]$$

$$= \Psi^{-1}(\kappa,\rho;r,n^\star)\left[rn^\star-\frac{2r}{n^\star}A^2(\kappa,\rho;r,n^\star)\right]$$

$$\frac{\partial}{\partial\kappa}\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)$$

$$= \exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\cdot 2\left[\frac{\partial}{\partial\kappa}\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right]$$

$$= \exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\cdot\left[n^\star\left(r^2-\frac{1}{4}\right)-\frac{2}{n^\star}\int\limits_{u=-\frac{1}{2}}^{r}u\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2du\right]$$

$$= E(\kappa,\rho;r,n^\star)\left[n^\star\left(r^2-\frac{1}{4}\right)-\frac{2}{n^\star}G(\kappa,\rho;r,n^\star)\right]$$

For (2), we have:

$$\frac{\partial}{\partial\kappa}\left[n^\star\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left[2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\left(r^2-\frac{1}{4}\right)dr\right]$$

$$= n^\star\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\frac{\partial}{\partial\kappa}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\right]dr$$

$$= n^\star\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\cdot\frac{\partial}{\partial\kappa}\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)dr$$

$$+n^\star\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\cdot\Psi^{-1}(\kappa,\rho;r,n^\star)\frac{\partial}{\partial\kappa}\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)dr,$$

where:

$$\frac{\partial}{\partial\kappa}\Psi^{-1}(\kappa,\rho;r,n^\star) = -\frac{r}{n^\star}\left[\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right] = -\frac{r}{n^\star}H(\kappa,\rho;r,n^\star)$$

$$\frac{\partial}{\partial \kappa}\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)$$

$$= \exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\cdot\left[n^\star\left(r^2-\frac{1}{4}\right)-\frac{2}{n^\star}\int_{u=-\frac{1}{2}}^{r}u\cdot\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2du\right]$$

$$= E(\kappa,\rho;r,n^\star)\left[n^\star\left(r^2-\frac{1}{4}\right)-\frac{2}{n^\star}G(\kappa,\rho;r,n^\star)\right]$$

For (3), we have:

$$\frac{\partial}{\partial \kappa}\left[-\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\int_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right]^2du\right)dr\right]$$

$$= -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\frac{\partial}{\partial \kappa}\left(\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\int_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right]^2du\right)dr$$

$$= -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(\frac{\partial}{\partial \kappa}\Psi^{-1}(\kappa,\rho;r,n^\star)\cdot\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\int_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right]^2du\right)dr$$

$$-\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(\Psi^{-1}(\kappa,\rho;r,n^\star)\cdot\frac{\partial}{\partial \kappa}\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\int_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right]^2du\right)dr$$

$$-\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(\Psi^{-1}(\kappa,\rho;r,n^\star)\cdot\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\frac{\partial}{\partial \kappa}\int_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right]^2du\right)dr.$$

where:

$$\frac{\partial}{\partial \kappa}\Psi^{-1}(\kappa,\rho;r,n^\star) = -\frac{r}{n^\star}\left[\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right] = -\frac{r}{n^\star}H(\kappa,\rho;r,n^\star)$$

$$
\begin{aligned}
&\frac{\partial}{\partial\kappa}\exp\left[2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\\
=\;&\exp\left[2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\cdot\left[n^\star\left(r^2-\frac{1}{4}\right)-\frac{2}{n^\star}\int\limits_{u=-\frac{1}{2}}^{r}u\cdot\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2du\right]\\
=\;&E(\kappa,\rho;r,n^\star)\left[n^\star\left(r^2-\frac{1}{4}\right)-\frac{2}{n^\star}G(\kappa,\rho;r,n^\star)\right]
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial}{\partial\kappa}\int\limits_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\\
=\;&\int\limits_{u=-\frac{1}{2}}^{r}u\frac{\partial}{\partial\kappa}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du=2\int\limits_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]\frac{\partial}{\partial\kappa}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right]du\\
=\;&2\int\limits_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]\left[un^\star-\frac{u}{n^\star}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2\right]du\\
=\;&2\int\limits_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]un^\star du+2\int\limits_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]\left[-\frac{u}{n^\star}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2\right]du\\
=\;&2n^\star\int\limits_{u=-\frac{1}{2}}^{r}u^2\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]du-\frac{2}{n^\star}\int\limits_{u=-\frac{1}{2}}^{r}u^2\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^3du\\
=\;&2n^\star L(\kappa,\rho;r,n^\star)-\frac{2}{n^\star}M(\kappa,\rho;r,n^\star)
\end{aligned}
$$

Thus, re-write in terms of the pre-defined functions in **Step 0**, we finally get:

$$
\begin{aligned}
&\frac{\partial^2}{\partial\kappa^2}c^{-1}(\kappa,\rho;n^\star) \\
=\;& -\frac{1}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[r\Psi^{-1}(\kappa,\rho;r,n^\star)\left(rn^\star-\frac{2r}{n^\star}A^2(\kappa,\rho;r,n^\star)\right)E(\kappa,\rho;r,n^\star)\right]dr \\
& -\frac{1}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}} rH(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[n^\star\left(r^2-\frac{1}{4}\right)-\frac{2}{n^\star}G(\kappa,\rho;r,n^\star)\right]dr \\
& +n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\left(-\frac{r}{n^\star}H(\kappa,\rho;r,n^\star)\right)E(\kappa,\rho;r,n^\star)dr \\
& +n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\cdot\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[n^\star\left(r^2-\frac{1}{4}\right)-\frac{2}{n^\star}G(\kappa,\rho;r,n^\star)\right]dr \\
& -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(-\frac{r}{n^\star}H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)G(\kappa,\rho;r,n^\star)\right)dr \\
& -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[n^\star\left(r^2-\frac{1}{4}\right)-\frac{2}{n^\star}G(\kappa,\rho;r,n^\star)\right]G(\kappa,\rho;r,n^\star)\right)dr \\
& -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left(2n^\star L(\kappa,\rho;r,n^\star)-\frac{2}{n^\star}M(\kappa,\rho;r,n^\star)\right)\right]dr.
\end{aligned}
$$

Secondly, taking the first partial derivative of C.2 with respect to ρ, we get:

$$
\begin{aligned}
&\frac{\partial l(\kappa,\rho)}{\partial \rho} \\
= & \frac{\partial}{\partial \rho}\left(m \cdot \ln\left[c(\kappa,\rho;n^{\star})\right] + \left\{\sum_{i=1}^{m} \ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\right]\right\} + 2\sum_{i=1}^{m}\left[\int_{u=-\frac{1}{2}}^{r_i} \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du\right]\right) \\
= & \frac{\partial}{\partial \rho}\left(m \cdot \ln\left[c(\kappa,\rho;n^{\star})\right]\right) + \frac{\partial}{\partial \rho}\left(\sum_{i=1}^{m} \ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\right]\right) \\
& + \frac{\partial}{\partial \rho}\left(2\sum_{i=1}^{m}\left[\int_{u=-\frac{1}{2}}^{r_i} \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du\right]\right) \\
= & -m \cdot \underbrace{\left(\frac{\partial}{\partial \rho}\ln\left[c^{-1}(\kappa,\rho;n^{\star})\right]\right)}_{(B.5.4)} + \sum_{i=1}^{m}\underbrace{\left(\frac{\partial}{\partial \rho}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\right]\right)}_{(B.5.2)} \\
& + 2\cdot\sum_{i=1}^{m}\underbrace{\left(\frac{\partial}{\partial \rho}\left[\int_{u=-\frac{1}{2}}^{r_i} \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du\right]\right)}_{(B.5.3)}, \qquad (C.5)
\end{aligned}
$$

where first partial derivative $(B.5.1)$ is given by:

$$
\begin{aligned}
\frac{\partial}{\partial \rho}\Pi(\kappa,\rho;r_i) &= \frac{\partial}{\partial \rho}\left[\sinh(\kappa r_i+\rho) - 2r_i\cosh(\kappa r_i+\rho)\right] \\
&= \cosh(\kappa r_i+\rho) - 2r_i\sinh(\kappa r_i+\rho) \\
&= n^{\star}\cdot\Psi(\kappa,\rho;r_i,n^{\star}),
\end{aligned}
$$

first partial derivative $(B.5.2)$ is given by:

$$
\begin{aligned}
\frac{\partial}{\partial \rho}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\right] &= \Psi(\kappa,\rho;r_i,n^{\star})\cdot\frac{\partial}{\partial \rho}\left[\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\right] \\
&= \Psi(\kappa,\rho;r_i,n^{\star})\cdot\left[\frac{-1}{n^{\star}}\cdot\frac{\Pi(\kappa,\rho;r_i)}{\Psi^2(\kappa,\rho;r_i,n^{\star})}\right] \\
&= \frac{-1}{n^{\star}}\cdot\left[\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^{\star})}\right],
\end{aligned}
$$

first partial derivative $(B.5.4)$ is given by:

$$
\begin{aligned}
&\frac{\partial}{\partial\rho}\ln\left[c^{-1}(\kappa,\rho;n^{\star})\right] = c(\kappa,\rho;n^{\star})\cdot\left[\frac{\partial}{\partial\rho}c^{-1}(\kappa,\rho;n^{\star})\right] \\
= \; & c(\kappa,\rho;n^{\star})\cdot\frac{\partial}{\partial\rho}\left[\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^{\star})\cdot\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du\right)dr\right] \\
= \; & c(\kappa,\rho;n^{\star})\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\frac{\partial}{\partial\rho}\left[\Psi^{-1}(\kappa,\rho;r,n^{\star})\cdot\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du\right)\right]dr \\
= \; & c(\kappa,\rho;n^{\star})\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\frac{\partial}{\partial\rho}\Psi^{-1}(\kappa,\rho;r,n^{\star})\cdot\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du\right)dr \\
& +c(\kappa,\rho;n^{\star})\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^{\star})\cdot\frac{\partial}{\partial\rho}\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du\right)dr \\
= \; & -\frac{c(\kappa,\rho;n^{\star})}{n^{\star}}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^{2}(\kappa,\rho;r,n^{\star})}\right)\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^{\star})}du\right)\right]dr \\
& +c(\kappa,\rho;n^{\star})\,2n^{\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^{\star})\exp\left[2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^{\star})}du\right]\left(r+\frac{1}{2}\right)\right]dr \\
& -\frac{2c(\kappa,\rho;n^{\star})}{n^{\star}}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^{\star})\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^{\star})}du\right)\int\limits_{u=-\frac{1}{2}}^{r}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^{\star})}\right)^{2}du\right]dr
\end{aligned}
$$

Then, to calculate the variance associated with $\widehat{\rho}$, we take the second partial derivative of

C.2 with respect to ρ to get:

$$\begin{aligned}
\frac{\partial^2 l(\kappa,\rho)}{\partial\rho^2} &= \frac{\partial}{\partial\rho}\left[\frac{\partial l(\kappa,\rho)}{\partial\rho}\right] \\
&= \frac{\partial}{\partial\rho}\left[\begin{array}{c} -m\left(\frac{\partial}{\partial\rho}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\right) + \sum_{i=1}^{m}\left(\frac{\partial}{\partial\rho}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)\right]\right) \\ +2\sum_{i=1}^{m}\left(\frac{\partial}{\partial\rho}\left[\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\right)\end{array}\right] \\
&= -m\underbrace{\left(\frac{\partial^2}{\partial\rho^2}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\right)}_{(B.6.4)} + \sum_{i=1}^{m}\underbrace{\left(\frac{\partial^2}{\partial\rho^2}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)\right]\right)}_{(B.6.2)} \\
&\quad +2\sum_{i=1}^{m}\underbrace{\left(\frac{\partial^2}{\partial\rho^2}\left[\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\right)}_{(B.6.3)},
\end{aligned} \tag{C.6}$$

where the second partial derivative $(B.6.1)$ is given by:

$$\begin{aligned}
\frac{\partial^2}{\partial\rho^2}\Pi(\kappa,\rho;r_i) &= \frac{\partial}{\partial\rho}\left[\frac{\partial}{\partial\rho}\Pi(\kappa,\rho;r_i)\right] \\
&= \frac{\partial}{\partial\rho}\left[n^\star\cdot\Psi(\kappa,\rho;r_i,n^\star)\right] \\
&= n^\star\cdot\left[\frac{\partial}{\partial\rho}\Psi(\kappa,\rho;r_i,n^\star)\right] \\
&= n^\star\cdot\left[\frac{1}{n^\star}\cdot\Pi(\kappa,\rho;r_i)\right] \\
&= \Pi(\kappa,\rho;r_i),
\end{aligned}$$

the second partial derivative $(B.6.2)$ is given by:

$$\begin{aligned}
\frac{\partial^2}{\partial\rho^2}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)\right] &= \frac{\partial}{\partial\rho}\left[\frac{\partial}{\partial\rho}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)\right]\right] \\
&= \frac{\partial}{\partial\rho}\left(\frac{-1}{n^\star}\cdot\left[\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)}\right]\right) \\
&= \frac{-1}{n^\star}\cdot\left[\frac{\partial}{\partial\rho}\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)}\right] \\
&= \frac{-1}{n^\star}\cdot\left[n^\star-\frac{1}{n^\star}\left(\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)}\right)^2\right] \\
&= -1+\frac{1}{(n^\star)^2}\left[\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)}\right]^2,
\end{aligned}$$

the second partial derivative $(B.6.3)$ is given by:

$$\begin{aligned}
\frac{\partial^2}{\partial\rho^2}\left[\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right] &= \frac{\partial}{\partial\rho}\left[\frac{\partial}{\partial\rho}\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right] \\
&= \frac{\partial}{\partial\rho}\left(n^\star\cdot\left[r_i+\frac{1}{2}\right]-\frac{1}{n^\star}\int_{u=-\frac{1}{2}}^{r_i}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right) \\
&= -\frac{1}{n^\star}\int_{u=-\frac{1}{2}}^{r_i}\frac{\partial}{\partial\rho}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du \\
&= -\frac{1}{n^\star}\int_{u=-\frac{1}{2}}^{r_i}2\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]\frac{\partial}{\partial\rho}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]du \\
&= -\frac{2}{n^\star}\int_{u=-\frac{1}{2}}^{r_i}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]\left[n^\star-\frac{1}{n^\star}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2\right]du \\
&= -\frac{2}{n^\star}\int_{u=-\frac{1}{2}}^{r_i}\left(\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]n^\star-\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\frac{1}{n^\star}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2\right]\right)du \\
&= -2\int_{u=-\frac{1}{2}}^{r_i}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]du+\frac{2}{(n^\star)^2}\int_{u=-\frac{1}{2}}^{r_i}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^3du
\end{aligned}$$

and the second partial derivative $(B.6.4)$ is given by:

$$\frac{\partial^2}{\partial\rho^2}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] = \frac{\partial}{\partial\rho}\left[\frac{\partial}{\partial\rho}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\right] = \frac{\partial}{\partial\rho}\left(c(\kappa,\rho;n^\star)\cdot\left[\frac{\partial}{\partial\rho}c^{-1}(\kappa,\rho;n^\star)\right]\right)$$

$$= \left[\frac{\partial}{\partial\rho}c(\kappa,\rho;n^\star)\cdot\frac{\partial}{\partial\rho}c^{-1}(\kappa,\rho;n^\star)\right] + \left[c(\kappa,\rho;n^\star)\cdot\frac{\partial^2}{\partial\rho^2}c^{-1}(\kappa,\rho;n^\star)\right]$$

$$= c(\kappa,\rho;n^\star)\left[\left(\frac{\frac{\partial}{\partial\rho}c(\kappa,\rho;n^\star)}{c(\kappa,\rho;n^\star)}\cdot\frac{\partial}{\partial\rho}c^{-1}(\kappa,\rho;n^\star)\right) + \left(\frac{\partial^2}{\partial\rho^2}c^{-1}(\kappa,\rho;n^\star)\right)\right]$$

$$= c(\kappa,\rho;n^\star)\left[\left(-\frac{\partial}{\partial\rho}\ln c^{-1}(\kappa,\rho;n^\star)\cdot\frac{\partial}{\partial\rho}c^{-1}(\kappa,\rho;n^\star)\right) + \frac{\partial^2}{\partial\rho^2}c^{-1}(\kappa,\rho;n^\star)\right]$$

$$= -\left(\underbrace{\frac{\partial}{\partial\rho}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]}_{(B.5.4)}\right)^2 + c(\kappa,\rho;n^\star)\underbrace{\left[\frac{\partial^2}{\partial\rho^2}c^{-1}(\kappa,\rho;n^\star)\right]}_{(B.6.5)}.$$

Again, deriving the second partial derivative $(B.6.5)$ will take some work:

$$\frac{\partial^2}{\partial\rho^2}c^{-1}(\kappa,\rho;n^\star) = \frac{\partial}{\partial\rho}\left(\frac{\partial}{\partial\rho}c^{-1}(\kappa,\rho;n^\star)\right)$$

$$= \underbrace{\frac{\partial}{\partial\rho}\left[-\frac{1}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)dr\right]}_{(4)}$$

$$\underbrace{+\frac{\partial}{\partial\rho}\left[2n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right]\left(r+\frac{1}{2}\right)dr\right]}_{(5)}$$

$$\underbrace{+\frac{\partial}{\partial\rho}\left[-\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)\int_{u=-\frac{1}{2}}^{r}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right)^2du\right]dr\right]}_{(6)}$$

Lets derive each of (4) $\sim$ (6) individually. For (4), we have:

$$\frac{\partial}{\partial\rho}\left[-\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)dr\right]$$

$$= -\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\frac{\partial}{\partial\rho}\left[\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)\right]dr$$

$$= -\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\frac{\partial}{\partial\rho}\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)dr$$

$$-\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\frac{\partial}{\partial\rho}\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)dr,$$

where:

$$\frac{\partial}{\partial\rho}\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right) = \frac{\partial}{\partial\rho}\left[\Pi(\kappa,\rho;r)\cdot\Psi^{-2}(\kappa,\rho;r,n^\star)\right]$$

$$= \left[\frac{\partial}{\partial\rho}\Pi(\kappa,\rho;r)\cdot\Psi^{-2}(\kappa,\rho;r,n^\star)\right] + \left[\Pi(\kappa,\rho;r)\cdot\frac{\partial}{\partial\rho}\Psi^{-2}(\kappa,\rho;r,n^\star)\right]$$

$$= \left[n^\star\Psi(\kappa,\rho;r_i,n^\star)\cdot\Psi^{-2}(\kappa,\rho;r,n^\star)\right] + \left[\Pi(\kappa,\rho;r)\cdot 2\Psi^{-1}(\kappa,\rho;r,n^\star)\frac{\partial}{\partial\rho}\Psi^{-1}(\kappa,\rho;r,n^\star)\right]$$

$$= \frac{n^\star}{\Psi(\kappa,\rho;r_i,n^\star)} + 2\left[\Pi(\kappa,\rho;r)\Psi^{-1}(\kappa,\rho;r,n^\star)\left(\frac{-1}{n^\star}\frac{\Pi(\kappa,\rho;r_i)}{\Psi^2(\kappa,\rho;r_i,n^\star)}\right)\right]$$

$$= \frac{n^\star}{\Psi(\kappa,\rho;r_i,n^\star)} - \frac{2}{n^\star}\left[\frac{\Pi^2(\kappa,\rho;r_i)}{\Psi^3(\kappa,\rho;r_i,n^\star)}\right]$$

$$= n^\star\Psi^{-1}(\kappa,\rho;r_i,n^\star) - \frac{2}{n^\star}\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)A^2(\kappa,\rho;r_i,n^\star)\right]$$

$$= \Psi^{-1}(\kappa,\rho;r_i,n^\star)\left[n^\star - \frac{2}{n^\star}A^2(\kappa,\rho;r_i,n^\star)\right]$$

$$\frac{\partial}{\partial\rho}\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)$$

$$= \exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\cdot 2\frac{\partial}{\partial\rho}\left[\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]$$

$$= \exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\cdot\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}\int_{u=-\frac{1}{2}}^{r}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right]$$

$$= E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]$$

For (5), we have:

$$\frac{\partial}{\partial\rho}\left[2n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right]\left(r+\frac{1}{2}\right)dr\right]$$

$$= 2n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r+\frac{1}{2}\right)\frac{\partial}{\partial\rho}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)\right]dr$$

$$= 2n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r+\frac{1}{2}\right)\cdot\frac{\partial}{\partial\rho}\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)dr$$

$$+2n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r+\frac{1}{2}\right)\cdot\Psi^{-1}(\kappa,\rho;r,n^\star)\frac{\partial}{\partial\rho}\left[\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)\right]dr,$$

where:

$$\frac{\partial}{\partial\rho}\Psi^{-1}(\kappa,\rho;r,n^\star)=\frac{-1}{n^\star}\cdot\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}=\frac{-1}{n^\star}\cdot H(\kappa,\rho;r,n^\star)$$

$$\frac{\partial}{\partial\rho}\left[\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)\right]$$

$$= \exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\cdot\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}\int_{u=-\frac{1}{2}}^{r}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2 du\right]$$

$$= E(\kappa,\rho;r,n^\star)\cdot\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]$$

For (6), we have:

$$\frac{\partial}{\partial\rho}\left[-\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)\int_{u=-\frac{1}{2}}^{r}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right)^2 du\right]dr\right]$$

$$= -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\frac{\partial}{\partial\rho}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)\int_{u=-\frac{1}{2}}^{r}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right)^2 du\right]dr$$

$$\begin{aligned}= & -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\frac{\partial}{\partial\rho}\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)\int_{u=-\frac{1}{2}}^{r}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right)^2 du\right]dr \\
& -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\frac{\partial}{\partial\rho}\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)\int_{u=-\frac{1}{2}}^{r}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right)^2 du\right]dr \\
& -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)\frac{\partial}{\partial\rho}\int_{u=-\frac{1}{2}}^{r}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right)^2 du\right]dr,\end{aligned}$$

where:

$$\frac{\partial}{\partial\rho}\Psi^{-1}(\kappa,\rho;r,n^\star)=\frac{-1}{n^\star}\cdot\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}=\frac{-1}{n^\star}H(\kappa,\rho;r,n^\star)$$

$$
\begin{aligned}
&\frac{\partial}{\partial\rho}\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}du\right)\\
&= \exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\cdot\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}\int\limits_{u=-\frac{1}{2}}^{r}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right]\\
&= E(\kappa,\rho;r,n^\star)\cdot\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial\rho}\int\limits_{u=-\frac{1}{2}}^{r}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right)^2du &= \int\limits_{u=-\frac{1}{2}}^{r}\frac{\partial}{\partial\rho}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right)^2du\\
&= \int\limits_{u=-\frac{1}{2}}^{r}2\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right)\frac{\partial}{\partial\rho}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right)du\\
&= 2\int\limits_{u=-\frac{1}{2}}^{r}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;r,n^\star)}\right)\left[n^\star-\frac{1}{n^\star}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2\right]du\\
&= 2n^\star\int\limits_{u=-\frac{1}{2}}^{r}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)du-\frac{2}{n^\star}\int\limits_{u=-\frac{1}{2}}^{r}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^3du\\
&= 2n^\star B(\kappa,\rho;r,n^\star)-\frac{2}{n^\star}N(\kappa,\rho;r,n^\star)
\end{aligned}
$$

Thus, re-write in terms of the pre-defined functions in **Step 0**, we finally get:

$$
\begin{aligned}
&\frac{\partial^2}{\partial \rho^2} c^{-1}(\kappa,\rho;n^\star) \\
= \; & -\frac{1}{n^\star} \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \Psi^{-1}(\kappa,\rho;r_i,n^\star)\left[n^\star - \frac{2}{n^\star}A^2(\kappa,\rho;r,n^\star)\right] E(\kappa,\rho;r,n^\star)dr \\
& -\frac{1}{n^\star} \int_{r=-\frac{1}{2}}^{\frac{1}{2}} H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right) - \frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right] dr \\
& -2 \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \left(r+\frac{1}{2}\right) H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)dr \\
& +2n^\star \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \left(r+\frac{1}{2}\right)\cdot\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right) - \frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right] dr \\
& +\frac{2}{(n^\star)^2} \int_{r=-\frac{1}{2}}^{\frac{1}{2}} H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)D(\kappa,\rho;r,n^\star)dr \\
& -\frac{2}{n^\star} \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \left[\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right) - \frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right] D(\kappa,\rho;r,n^\star)\right] dr \\
& -\frac{2}{n^\star} \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star B(\kappa,\rho;r,n^\star) - \frac{2}{n^\star}N(\kappa,\rho;r,n^\star)\right] dr
\end{aligned}
$$

Finally, in order to calculate the observed information matrix, we also need to calculate the second cross partial derivative of C.2 first with respect to κ then with respect to ρ as follows:

$$\begin{aligned}
\frac{\partial^2 l(\kappa,\rho)}{\partial\rho\partial\kappa} &= \frac{\partial}{\partial\rho}\left[\frac{\partial l(\kappa,\rho)}{\partial\kappa}\right] \\
&= \frac{\partial}{\partial\rho}\left[\begin{array}{c} -m\cdot\left(\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^*)\right]\right)+\sum_{i=1}^{m}\left(\frac{\partial}{\partial\kappa}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^*)\right]\right) \\ +2\sum_{i=1}^{m}\left(\frac{\partial}{\partial\kappa}\left[\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}du\right]\right)\end{array}\right] \\
&= -m\cdot\underbrace{\left(\frac{\partial^2}{\partial\rho\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^*)\right]\right)}_{(B.7.4)}+\sum_{i=1}^{m}\underbrace{\left(\frac{\partial^2}{\partial\rho\partial\kappa}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^*)\right]\right)}_{(B.7.2)} \\
&\quad +2\sum_{i=1}^{m}\underbrace{\left(\frac{\partial^2}{\partial\rho\partial\kappa}\left[\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}du\right]\right)}_{(B.7.3)},
\end{aligned} \tag{C.7}$$

where the second cross partial derivative $(B.7.1)$ is given by:

$$\begin{aligned}
\frac{\partial^2}{\partial\rho\partial\kappa}\Pi(\kappa,\rho;r_i) &= \frac{\partial}{\partial\rho}\left[\frac{\partial}{\partial\kappa}\Pi(\kappa,\rho;r_i)\right] \\
&= \frac{\partial}{\partial\rho}\left[r_i n^*\Psi(\kappa,\rho;r_i,n^*)\right] \\
&= r_i n^*\cdot\left[\frac{\partial}{\partial\rho}\Psi(\kappa,\rho;r_i,n^*)\right] \\
&= r_i n^*\cdot\left[\frac{1}{n^*}\cdot\Pi(\kappa,\rho;r_i)\right] \\
&= r_i\cdot\Pi(\kappa,\rho;r_i),
\end{aligned}$$

the second cross partial derivative $(B.7.2)$ is given by:

$$\begin{aligned}
\frac{\partial^2}{\partial\rho\partial\kappa}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^*)\right] &= \frac{\partial}{\partial\rho}\left[\frac{\partial}{\partial\kappa}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^*)\right]\right] \\
&= \frac{\partial}{\partial\rho}\left(-\frac{r_i}{n^*}\cdot\left[\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^*)}\right]\right) \\
&= -\frac{r_i}{n^*}\cdot\left[\frac{\partial}{\partial\rho}\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^*)}\right] \\
&= -\frac{r_i}{n^*}\cdot\left[n^*-\frac{1}{n^*}\left(\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^*)}\right)^2\right] \\
&= -r_i+\frac{r_i}{(n^*)^2}\left[\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^*)}\right]^2,
\end{aligned}$$

the second cross partial derivative $(B.7.3)$ is given by:

$$\frac{\partial^2}{\partial\rho\partial\kappa}\left[\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}du\right] = \frac{\partial}{\partial\rho}\left[\frac{\partial}{\partial\kappa}\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}du\right]$$

$$= \frac{\partial}{\partial\rho}\left[\frac{n^*}{2}\left(r_i^2-\frac{1}{4}\right)-\frac{1}{n^*}\int_{u=-\frac{1}{2}}^{r_i}u\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}\right)^2du\right]$$

$$= -\frac{1}{n^*}\int_{u=-\frac{1}{2}}^{r_i}u\left[\frac{\partial}{\partial\rho}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}\right)^2\right]du$$

$$= -\frac{1}{n^*}\int_{u=-\frac{1}{2}}^{r_i}u\left[2\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}\right)\left(\frac{\partial}{\partial\rho}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}\right)\right]du$$

$$= -\frac{2}{n^*}\int_{u=-\frac{1}{2}}^{r_i}u\left[\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}\right)\left(n^*-\frac{1}{n^*}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}\right)^2\right)\right]du$$

$$= -\frac{2}{n^*}\int_{u=-\frac{1}{2}}^{r_i}u\left[\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}\right)n^*-\frac{1}{n^*}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}\right)^3\right]du$$

$$= -2\int_{u=-\frac{1}{2}}^{r_i}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}\right]du+\frac{2}{(n^*)^2}\int_{u=-\frac{1}{2}}^{r_i}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}\right]^3du,$$

and the second cross partial derivative $(B.7.4)$ is given by:

$$\frac{\partial^2}{\partial\rho\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^*)\right]=\frac{\partial}{\partial\rho}\left[\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^*)\right]\right]=\frac{\partial}{\partial\rho}\left(c(\kappa,\rho;n^*)\cdot\frac{\partial}{\partial\kappa}c^{-1}(\kappa,\rho;n^*)\right)$$

$$=\left[\frac{\partial}{\partial\rho}c(\kappa,\rho;n^*)\cdot\frac{\partial}{\partial\kappa}c^{-1}(\kappa,\rho;n^*)\right]+\left[c(\kappa,\rho;n^*)\cdot\frac{\partial^2}{\partial\rho\partial\kappa}c^{-1}(\kappa,\rho;n^*)\right]$$

$$=c(\kappa,\rho;n^*)\left[\frac{\frac{\partial}{\partial\rho}c(\kappa,\rho;n^*)}{c(\kappa,\rho;n^*)}\cdot\frac{\partial}{\partial\kappa}c^{-1}(\kappa,\rho;n^*)+\frac{\partial^2}{\partial\rho\partial\kappa}c^{-1}(\kappa,\rho;n^*)\right]$$

$$=c(\kappa,\rho;n^*)\left[-\frac{\partial}{\partial\rho}\ln c^{-1}(\kappa,\rho;n^*)\cdot\frac{\partial}{\partial\kappa}c^{-1}(\kappa,\rho;n^*)+\frac{\partial^2}{\partial\rho\partial\kappa}c^{-1}(\kappa,\rho;n^*)\right]$$

$$=-\left(\underbrace{\frac{\partial}{\partial\rho}\ln\left[c^{-1}(\kappa,\rho;n^*)\right]}_{(B.5.4)}\cdot\underbrace{\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^*)\right]}_{(B.3.4)}\right)+c(\kappa,\rho;n^*)\underbrace{\left[\frac{\partial^2}{\partial\rho\partial\kappa}c^{-1}(\kappa,\rho;n^*)\right]}_{(B.7.5)}$$

As was the case for the second partial derivatives, the second cross partial derivative $(B.7.5)$ will take some work to derive:

$$\frac{\partial^2}{\partial\rho\partial\kappa}c^{-1}(\kappa,\rho;n^*)=\frac{\partial}{\partial\rho}\left[\frac{\partial}{\partial\kappa}c^{-1}(\kappa,\rho;n^*)\right]$$

$$=\underbrace{\frac{\partial}{\partial\rho}\left[-\frac{1}{n^*}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}r\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^*)}\right)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}du\right)dr\right]}_{(7)}$$

$$+\underbrace{\frac{\partial}{\partial\rho}\left[n^*\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^*)\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}du\right]\left(r^2-\frac{1}{4}\right)dr\right]}_{(8)}$$

$$+\underbrace{\frac{\partial}{\partial\rho}\left[-\frac{2}{n^*}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^*)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}du\right)\int_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^*)}\right]^2du\right]dr\right]}_{(9)}$$

Lets derive each of (7) ~ (9) individually. For (7), we have:

$$\begin{aligned}
&\frac{\partial}{\partial\rho}\left[-\frac{1}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}} r\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)dr\right]\\
&= -\frac{1}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}} r\frac{\partial}{\partial\rho}\left[\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\right]dr\\
&= -\frac{1}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}} r\cdot\frac{\partial}{\partial\rho}\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)dr\\
&\quad -\frac{1}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}} r\cdot\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right)\frac{\partial}{\partial\rho}\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)dr,
\end{aligned}$$

where:

$$\begin{aligned}
\frac{\partial}{\partial\rho}\left(\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}\right) &= \frac{n^\star}{\Psi(\kappa,\rho;r_i,n^\star)}-\frac{2}{n^\star}\left[\frac{\Pi^2(\kappa,\rho;r_i)}{\Psi^3(\kappa,\rho;r_i,n^\star)}\right]\\
&= n^\star\Psi^{-1}(\kappa,\rho;r,n^\star)-\frac{2}{n^\star}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)A^2(\kappa,\rho;r,n^\star)\right]\\
&= \Psi^{-1}(\kappa,\rho;r,n^\star)\left[n^\star-\frac{2}{n^\star}A^2(\kappa,\rho;r,n^\star)\right]
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial}{\partial\rho}\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\\
&= \exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\cdot\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}\int_{u=-\frac{1}{2}}^{r}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right]\\
&= E(\kappa,\rho;r,n^\star)\cdot\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]
\end{aligned}$$

For (8), we have:

$$\frac{\partial}{\partial\rho}\left[n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left[2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]\left(r^2-\frac{1}{4}\right)dr\right]$$

$$= n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\frac{\partial}{\partial\rho}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\right]dr$$

$$= n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\cdot\frac{\partial}{\partial\rho}\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)dr$$

$$+n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\Psi^{-1}(\kappa,\rho;r,n^\star)\frac{\partial}{\partial\rho}\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)dr,$$

where:

$$\frac{\partial}{\partial\rho}\Psi^{-1}(\kappa,\rho;r,n^\star)=\frac{-1}{n^\star}\cdot\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)}=\frac{-1}{n^\star}H(\kappa,\rho;r,n^\star)$$

$$\frac{\partial}{\partial\rho}\exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)$$

$$= \exp\left(2\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\cdot\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}\int_{u=-\frac{1}{2}}^{r}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right]$$

$$= E(\kappa,\rho;r,n^\star)\cdot\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]$$

For (9), we have:

$$\frac{\partial}{\partial\rho}\left[-\frac{2}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\int\limits_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right]dr\right]$$

$$= -\frac{2}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\frac{\partial}{\partial\rho}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\int\limits_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right]dr$$

$$= -\frac{2}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\frac{\partial}{\partial\rho}\Psi^{-1}(\kappa,\rho;r,n^\star)\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\int\limits_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right]dr$$

$$-\frac{2}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\cdot\frac{\partial}{\partial\rho}\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\int\limits_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right]dr$$

$$-\frac{2}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)\cdot\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\frac{\partial}{\partial\rho}\int\limits_{u=-\frac{1}{2}}^{r}u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right]dr.$$

where:

$$\frac{\partial}{\partial\rho}\Psi^{-1}(\kappa,\rho;r,n^\star) = \frac{-1}{n^\star}\cdot\frac{\Pi(\kappa,\rho;r)}{\Psi^2(\kappa,\rho;r,n^\star)} = \frac{-1}{n^\star}H(\kappa,\rho;r,n^\star)$$

$$\frac{\partial}{\partial\rho}\exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)$$

$$= \exp\left(2\int\limits_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right)\cdot\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}\int\limits_{u=-\frac{1}{2}}^{r}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^2du\right]$$

$$= E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]$$

$$
\begin{aligned}
\frac{\partial}{\partial \rho} \int\limits_{u=-\frac{1}{2}}^{r} u \left[\frac{\Pi(\kappa, \rho; u)}{\Psi(\kappa, \rho; u, n^\star)} \right]^2 du &= \int\limits_{u=-\frac{1}{2}}^{r} u \frac{\partial}{\partial \rho} \left[\frac{\Pi(\kappa, \rho; u)}{\Psi(\kappa, \rho; u, n^\star)} \right]^2 du \\
&= \int\limits_{u=-\frac{1}{2}}^{r} u 2 \left[\frac{\Pi(\kappa, \rho; u)}{\Psi(\kappa, \rho; u, n^\star)} \right] \frac{\partial}{\partial \rho} \left[\frac{\Pi(\kappa, \rho; u)}{\Psi(\kappa, \rho; u, n^\star)} \right] du \\
&= 2 \int\limits_{u=-\frac{1}{2}}^{r} u \left[\frac{\Pi(\kappa, \rho; u)}{\Psi(\kappa, \rho; u, n^\star)} \right] \left[n^\star - \frac{1}{n^\star} \left(\frac{\Pi(\kappa, \rho; r_i)}{\Psi(\kappa, \rho; r_i, n^\star)} \right)^2 \right] du \\
&= 2n^\star \int\limits_{u=-\frac{1}{2}}^{r} u \left[\frac{\Pi(\kappa, \rho; u)}{\Psi(\kappa, \rho; u, n^\star)} \right] du - \frac{2}{n^\star} \int\limits_{u=-\frac{1}{2}}^{r} u \left(\frac{\Pi(\kappa, \rho; r_i)}{\Psi(\kappa, \rho; r_i, n^\star)} \right)^3 du \\
&= 2n^\star F(\kappa, \rho; r, n^\star) - \frac{2}{n^\star} O\left(\kappa, \rho; r, n^\star\right)
\end{aligned}
$$

Thus, re-write in terms of the pre-defined functions in **Step 0**, we finally get:

$$
\begin{aligned}
&\frac{\partial^2}{\partial\rho\partial\kappa}c^{-1}(\kappa,\rho;n^\star) \\
= \quad & -\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} r\Psi^{-1}(\kappa,\rho;r,n^\star)\left[n^\star-\frac{2}{n^\star}A^2(\kappa,\rho;r,n^\star)\right]E(\kappa,\rho;r,n^\star)dr \\
& -\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} rH(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]dr \\
& -\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)dr \\
& +n^\star\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]dr \\
& +\frac{2}{(n^\star)^2}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)G(\kappa,\rho;r,n^\star)dr \\
& -\frac{2}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]G(\kappa,\rho;r,n^\star)dr \\
& -\frac{2}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star F(\kappa,\rho;r,n^\star)-\frac{2}{n^\star}O(\kappa,\rho;r,n^\star)\right]dr
\end{aligned}
$$

C.2 The MLE Procedure By First Principle

Recall that based on the *theory of social limitation*, Vaga (1990)'s Coherent Market Hypothesis (CMH) models the annualised market return by a random variable, R, with p.d.f. (3.33) given by:

$$
\begin{aligned}
f_R(\mathbf{r}; n, \kappa, \rho) &= \text{p.d.f. of annualised market return } R\text{, given the parameter values } n, \kappa, \rho \\
&= c(\kappa, \rho, n) \cdot \Psi^{-1}(\mathbf{r}; \kappa, \rho, n) \cdot \exp\left[2 \cdot \int_{u=-\frac{1}{2}}^{\mathbf{r}} \frac{\Pi(u; \kappa, \rho)}{\Psi(u; \kappa, \rho, n)} du\right],
\end{aligned}
$$

where the three parameters are defined as follows:

$$
\begin{aligned}
n &= \text{number of degrees of freedom,} \\
\kappa &= \text{degree of market sentiment,} \\
\rho &= \text{degree of fundamental bias,}
\end{aligned}
$$

the drift coefficient function is defined as:

$$
\Pi(u; \kappa, \rho) = \sinh(\kappa u + \rho) - 2u \cosh(\kappa u + \rho),
$$

the diffusion coefficient function is defined as:

$$
\Psi(u; \kappa, \rho, n) = \frac{1}{n}\left[\cosh(\kappa u + \rho) - 2u \sinh(\kappa u + \rho)\right],
$$

and the normalisation constant is defined as:

$$
c^{-1}(\kappa, \rho, n) = \int_{r=-\frac{1}{2}}^{\frac{1}{2}} \left(\Psi^{-1}(r; \kappa, \rho, n) \cdot \exp\left[2 \int_{u=-\frac{1}{2}}^{r} \frac{\Pi(u; \kappa, \rho)}{\Psi(u; \kappa, \rho, n)} du\right]\right) dr.
$$

Now, assuming that the annualised market return does actually have a distribution of form 3.33, our aim is to estimate this distribution's parameter values κ and ρ based on:

1. A given realisation of random sample of size m log-return observations $(r_i)_{i=1}^{m}$, and

2. A pre-determined degrees of freedom for the p.d.f. of $n = n^{\star}$.

We shall estimate the parameter values κ and ρ using the method of maximum likelihood as follows:

Step 0: Due to the complexity of the form of p.d.f. 3.33, we need to pre-define some functions that are frequently used in the various partial derivatives of the log-likelihood. Firstly, lets define the fundamental function that appears over and over again in the partial derivatives:

$$\begin{aligned} A(\kappa,\rho;u,n^\star) &= \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} = \frac{\sinh(\kappa u+\rho) - 2u\cosh(\kappa u+\rho)}{\frac{1}{n}\left[\cosh(\kappa u+\rho) - 2u\sinh(\kappa u+\rho)\right]} \\ &= n^\star \cdot \left[\frac{\sinh(\kappa u+\rho) - 2u\cosh(\kappa u+\rho)}{\cosh(\kappa u+\rho) - 2u\sinh(\kappa u+\rho)}\right]. \end{aligned}$$

Secondly, lets define those functions that are frequently used in the first partial derivative of log-likelihood:

$$\begin{aligned} B(\kappa,\rho;r,n^\star) &= \int_{u=-\frac{1}{2}}^{r} \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} du = \int_{u=-\frac{1}{2}}^{r} A(\kappa,\rho;u,n^\star)\, du \\ D(\kappa,\rho;r,n^\star) &= \int_{u=-\frac{1}{2}}^{r} \left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2 du = \int_{u=-\frac{1}{2}}^{r} A^2(\kappa,\rho;u,n^\star)\, du \\ E(\kappa,\rho;r,n^\star) &= \exp\left[2 \cdot \int_{u=-\frac{1}{2}}^{r} \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} du\right] = \exp\left[2 \cdot B(\kappa,\rho;r,n^\star)\right] \end{aligned}$$

$$\begin{aligned} F(\kappa,\rho;r,n^\star) &= \int_{u=-\frac{1}{2}}^{r} u \cdot \left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right) du = \int_{u=-\frac{1}{2}}^{r} u \cdot A(\kappa,\rho;u,n^\star)\, du \\ G(\kappa,\rho;r,n^\star) &= \int_{u=-\frac{1}{2}}^{r} u \cdot \left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2 du = \int_{u=-\frac{1}{2}}^{r} u \cdot A^2(\kappa,\rho;u,n^\star)\, du \end{aligned}$$

Thirdly, let us now define those functions that are frequently used in second partial derivative of log-likelihood:

$$H(\kappa,\rho;r,n^\star) = \frac{\Pi(\kappa,\rho;u)}{\Psi^2(\kappa,\rho;u,n^\star)} = \Psi^{-1}(\kappa,\rho;u,n^\star) \cdot A(\kappa,\rho;u,n^\star)$$

$$I(\kappa,\rho;r,n^{\star}) = \int_{u=-\frac{1}{2}}^{r} \left(\frac{1}{(n^{\star})^2} \left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})} \right]^3 - \left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})} \right] \right) du$$

$$= \int_{u=-\frac{1}{2}}^{r} \left[\frac{A^3(\kappa,\rho;u,n^{\star})}{(n^{\star})^2} - A(\kappa,\rho;u,n^{\star}) \right] du$$

$$J(\kappa,\rho;r,n^{\star}) = \int_{u=-\frac{1}{2}}^{r} u^2 \left(\frac{1}{(n^{\star})^2} \left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})} \right]^3 - \left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})} \right] \right) du$$

$$= \int_{u=-\frac{1}{2}}^{r} u^2 \left(\frac{A^3(\kappa,\rho;u,n^{\star})}{(n^{\star})^2} - A(\kappa,\rho;u,n^{\star}) \right) du$$

$$K(\kappa,\rho;r,n^{\star}) = \int_{u=-\frac{1}{2}}^{r} \left(\frac{1}{(n^{\star})^2} \left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})} \right]^3 - u \left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})} \right] \right) du$$

$$= \int_{u=-\frac{1}{2}}^{r} \left[\frac{A^3(\kappa,\rho;u,n^{\star})}{(n^{\star})^2} - u \cdot A(\kappa,\rho;u,n^{\star}) \right] du$$

$$L(\kappa,\rho;r,n^{\star}) = \int_{u=-\frac{1}{2}}^{r} u^2 \left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})} \right) du = \int_{u=-\frac{1}{2}}^{r_i} u^2 A(\kappa,\rho;u,n^{\star})\, du$$

$$M(\kappa,\rho;r,n^{\star}) = \int_{u=-\frac{1}{2}}^{r_i} u^2 \left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})} \right)^3 du = \int_{u=-\frac{1}{2}}^{r_i} u^2 A^3(\kappa,\rho;u,n^{\star})\, du$$

$$N(\kappa,\rho;r,n^{\star}) = \int_{u=-\frac{1}{2}}^{r} \left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})} \right)^3 du = \int_{u=-\frac{1}{2}}^{r_i} A^3(\kappa,\rho;u,n^{\star})\, du$$

$$O(\kappa,\rho;r,n^{\star}) = \int_{u=-\frac{1}{2}}^{r} u \left(\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^{\star})} \right)^3 du = \int_{u=-\frac{1}{2}}^{r} u A^3(\kappa,\rho;u,n^{\star})\, du$$

Step 1: Determine the form of the log-likelihood function. For Vaga (1990)'s p.d.f. 3.33, the log-likelihood is a bivariate function of κ and ρ given by:

$$l(\kappa,\rho) = m\ln\left[c(\kappa,\rho;n^\star)\right]+\left\{\sum_{i=1}^{m}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)\right]\right\}+2\sum_{i=1}^{m}\left[\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right]. \quad \text{(C.8)}$$

Step 2: Derive the first two partial derivatives of the likelihood function with respect to the first parameter. Based on the likelihood function C.8, the first partial derivative with respect to κ is given by:

$$\begin{aligned}
&\frac{\partial l(\kappa,\rho)}{\partial\kappa}\\
=\;& -m\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]+\sum_{i=1}^{m}\frac{\partial}{\partial\kappa}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)\right]+2\sum_{i=1}^{m}\frac{\partial}{\partial\kappa}\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\\
=\;& -m\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]+\sum_{i=1}^{m}-\frac{r_i}{n^\star}\left[\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)}\right]\\
&+2\sum_{i=1}^{m}\left(\frac{n^\star}{2}\left[r_i^2-\frac{1}{4}\right]-\frac{1}{n^\star}\int_{u=-\frac{1}{2}}^{r_i}\left[u\cdot\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2\right]du\right)\\
=\;& -m\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]-m\left(\frac{n^\star}{4}\right)+\sum_{i=1}^{m}\left[n^\star r_i^2-\frac{r_i}{n^\star}\left(\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)}\right)\right]\\
&-\frac{2}{n^\star}\sum_{i=1}^{m}\left[\int_{u=-\frac{1}{2}}^{r_i}u\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2du\right], \qquad \text{(C.9)}
\end{aligned}$$

which can be re-written using the pre-defined functions in **Step 0** as:

$$\begin{aligned}
&\frac{\partial l(\kappa,\rho)}{\partial\kappa}\\
=\;& -m\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]-m\left(\frac{n^\star}{4}\right)+\sum_{i=1}^{m}\left[n^\star r_i^2-\frac{r_i}{n^\star}A(\kappa,\rho;r_i,n^\star)\right]-\frac{2}{n^\star}\sum_{i=1}^{m}G(\kappa,\rho;r_i,n^\star),
\end{aligned}$$

where:

$$
\begin{aligned}
&\frac{\partial}{\partial \kappa} \ln \left[c^{-1}(\kappa, \rho; n^\star)\right] \\
= \; & -\frac{c(\kappa, \rho; n^\star)}{n^\star} \int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left[r \cdot H(\kappa, \rho; r, n^\star) \cdot E(\kappa, \rho; r, n^\star)\right] dr \\
& + c(\kappa, \rho; n^\star)\, n^\star \int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left[\Psi^{-1}(\kappa, \rho; r, n^\star) \cdot E(\kappa, \rho; r, n^\star) \cdot \left(r^2 - \frac{1}{4}\right)\right] dr \\
& - \frac{2c(\kappa, \rho; n^\star)}{n^\star} \int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left[\Psi^{-1}(\kappa, \rho; r, n^\star) E(\kappa, \rho; r, n^\star) G(\kappa, \rho; r, n^\star)\right] dr,
\end{aligned}
$$

and the second partial derivative with respect to κ is given by:

$$
\begin{aligned}
&\frac{\partial^2 l(\kappa, \rho)}{\partial \kappa^2} \\
= \; & -m \frac{\partial^2}{\partial \kappa^2} \ln \left[c^{-1}(\kappa, \rho; n^\star)\right] + \sum_{i=1}^{m} \frac{\partial^2}{\partial \kappa^2} \ln \left[\Psi^{-1}(\kappa, \rho; r_i, n^\star)\right] + 2 \sum_{i=1}^{m} \frac{\partial^2}{\partial \kappa^2} \int\limits_{u=-\frac{1}{2}}^{r_i} \frac{\Pi(\kappa, \rho; u)}{\Psi(\kappa, \rho; u, n^\star)} du \\
= \; & -m \frac{\partial^2}{\partial \kappa^2} \ln \left[c^{-1}(\kappa, \rho; n^\star)\right] + \sum_{i=1}^{m} \left(-r_i^2 + \left[\frac{r_i}{n^\star} \cdot \frac{\Pi(\kappa, \rho; r_i)}{\Psi(\kappa, \rho; r_i, n^\star)}\right]^2\right) \\
& + 2 \sum_{i=1}^{m} \left(-2 \int\limits_{u=-\frac{1}{2}}^{r_i} u^2 \left[\frac{\Pi(\kappa, \rho; u)}{\Psi(\kappa, \rho; u, n^\star)}\right] du + \frac{2}{(n^\star)^2} \int\limits_{u=-\frac{1}{2}}^{r_i} u^2 \left[\frac{\Pi(\kappa, \rho; u)}{\Psi(\kappa, \rho; u, n^\star)}\right]^3 du\right) \\
= \; & -m \frac{\partial^2}{\partial \kappa^2} \ln \left[c^{-1}(\kappa, \rho; n^\star)\right] + \sum_{i=1}^{m} \left(\left[\frac{r_i}{n^\star} \cdot \frac{\Pi(\kappa, \rho; r_i)}{\Psi(\kappa, \rho; r_i, n^\star)}\right]^2 - r_i^2\right) \\
& + 4 \sum_{i=1}^{m} \left[\int\limits_{u=-\frac{1}{2}}^{r_i} u^2 \left(\frac{1}{(n^\star)^2} \left[\frac{\Pi(\kappa, \rho; u)}{\Psi(\kappa, \rho; u, n^\star)}\right]^3 - \left[\frac{\Pi(\kappa, \rho; u)}{\Psi(\kappa, \rho; u, n^\star)}\right]\right) du\right], \qquad \text{(C.10)}
\end{aligned}
$$

which can be re-written using the pre-defined functions in **Step 0** as:

$$
\begin{aligned}
&\frac{\partial^2 l(\kappa,\rho)}{\partial\kappa^2} \\
= \quad & -m\frac{\partial^2}{\partial\kappa^2}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]+\sum_{i=1}^{m}\left[\left(\frac{r_i}{n^\star}\right)^2\cdot A^2(\kappa,\rho;r_i,n^\star)-r_i^2\right]+4\sum_{i=1}^{m}J(\kappa,\rho;r_i,n^\star),
\end{aligned}
$$

where:

$$
\frac{\partial^2}{\partial\kappa^2}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]=-\left(\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\right)^2+c(\kappa,\rho;n^\star)\left[\frac{\partial^2}{\partial\kappa^2}c^{-1}(\kappa,\rho;n^\star)\right]
$$

and

$$
\begin{aligned}
&\frac{\partial^2}{\partial\kappa^2}c^{-1}(\kappa,\rho;n^\star) \\
= \quad & -\frac{1}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[r\Psi^{-1}(\kappa,\rho;r,n^\star)\left(rn^\star-\frac{2r}{n^\star}A^2(\kappa,\rho;r,n^\star)\right)E(\kappa,\rho;r,n^\star)\right]dr \\
& -\frac{1}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}} rH(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[n^\star\left(r^2-\frac{1}{4}\right)-\frac{2}{n^\star}G(\kappa,\rho;r,n^\star)\right]dr \\
& +n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\left(-\frac{r}{n^\star}H(\kappa,\rho;r,n^\star)\right)E(\kappa,\rho;r,n^\star)dr \\
& +n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\cdot\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[n^\star\left(r^2-\frac{1}{4}\right)-\frac{2}{n^\star}G(\kappa,\rho;r,n^\star)\right]dr \\
& -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(-\frac{r}{n^\star}H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)G(\kappa,\rho;r,n^\star)\right)dr \\
& -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[n^\star\left(r^2-\frac{1}{4}\right)-\frac{2}{n^\star}G(\kappa,\rho;r,n^\star)\right]G(\kappa,\rho;r,n^\star)\right)dr \\
& -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left(2n^\star L(\kappa,\rho;r,n^\star)-\frac{2}{n^\star}M(\kappa,\rho;r,n^\star)\right)\right]dr.
\end{aligned}
$$

Step 3: Derive the first two partial derivatives of the likelihood function with respect to the second parameter. Based on the likelihood function C.8, the first partial derivative with respect to ρ is given by:

$$
\begin{aligned}
&\frac{\partial l(\kappa,\rho)}{\partial \rho} \\
= \ & -m\frac{\partial}{\partial \rho}\ln\left[c^{-1}(\kappa,\rho;n^{\star})\right]+\sum_{i=1}^{m}\frac{\partial}{\partial \rho}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\right]+2\sum_{i=1}^{m}\frac{\partial}{\partial \rho}\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du \\
= \ & -m\frac{\partial}{\partial \rho}\ln\left[c^{-1}(\kappa,\rho;n^{\star})\right]+\sum_{i=1}^{m}\left(\frac{-1}{n^{\star}}\cdot\left[\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^{\star})}\right]\right) \\
& +2\sum_{i=1}^{m}\left(n^{\star}\left[r_i+\frac{1}{2}\right]-\frac{1}{n^{\star}}\int_{u=-\frac{1}{2}}^{r_i}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}\right]^2du\right) \\
= \ & -m\frac{\partial}{\partial \rho}\ln\left[c^{-1}(\kappa,\rho;n^{\star})\right]+mn^{\star}+\sum_{i=1}^{m}\left[2n^{\star}r_i-\frac{1}{n^{\star}}\left(\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^{\star})}\right)\right] \\
& -\frac{2}{n^{\star}}\sum_{i=1}^{m}\left[\int_{u=-\frac{1}{2}}^{r_i}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}\right)^2du\right] \qquad \text{(C.11)}
\end{aligned}
$$

which can be re-written using the pre-defined functions in **Step 0** as:

$$
\begin{aligned}
&\frac{\partial l(\kappa,\rho)}{\partial \rho} \\
= \ & -m\frac{\partial}{\partial \rho}\ln\left[c^{-1}(\kappa,\rho;n^{\star})\right]+mn^{\star}+\sum_{i=1}^{m}\left[2n^{\star}r_i-\frac{A(\kappa,\rho;r_i,n^{\star})}{n^{\star}}\right]-\frac{2}{n^{\star}}\sum_{i=1}^{m}D(\kappa,\rho;r_i,n^{\star})
\end{aligned}
$$

where:

$$\frac{\partial}{\partial\rho}\ln\left[c^{-1}\left(\kappa,\rho;n^{\star}\right)\right]$$

$$= -\frac{c\left(\kappa,\rho;n^{\star}\right)}{n^{\star}}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[H(\kappa,\rho;r,n^{\star})\cdot E(\kappa,\rho;r,n^{\star})\right]dr$$

$$+c\left(\kappa,\rho;n^{\star}\right)2n^{\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^{\star})\cdot E(\kappa,\rho;r,n^{\star})\cdot\left(r+\frac{1}{2}\right)\right]dr$$

$$-\frac{2c\left(\kappa,\rho;n^{\star}\right)}{n^{\star}}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^{\star})E(\kappa,\rho;r,n^{\star})D(\kappa,\rho;r,n^{\star})\right]dr.$$

and the second partial derivative with respect to ρ is given by:

$$\frac{\partial^2 l\left(\kappa,\rho\right)}{\partial\rho^2}$$

$$= -m\frac{\partial^2}{\partial\rho^2}\ln\left[c^{-1}\left(\kappa,\rho;n^{\star}\right)\right]+\sum_{i=1}^{m}\frac{\partial^2}{\partial\rho^2}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^{\star})\right]+2\sum_{i=1}^{m}\frac{\partial^2}{\partial\rho^2}\int\limits_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}du$$

$$= -m\frac{\partial^2}{\partial\rho^2}\ln\left[c^{-1}\left(\kappa,\rho;n^{\star}\right)\right]+\sum_{i=1}^{m}\left(-1+\frac{1}{(n^{\star})^2}\left[\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^{\star})}\right]^2\right)$$

$$+2\sum_{i=1}^{m}\left(-2\int\limits_{u=-\frac{1}{2}}^{r_i}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}\right]du+\frac{2}{(n^{\star})^2}\int\limits_{u=-\frac{1}{2}}^{r_i}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}\right)^3du\right)$$

$$= -m\frac{\partial^2}{\partial\rho^2}\ln\left[c^{-1}\left(\kappa,\rho;n^{\star}\right)\right]-m+\frac{1}{(n^{\star})^2}\sum_{i=1}^{m}\left[\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^{\star})}\right]^2$$

$$+4\sum_{i=1}^{m}\int\limits_{u=-\frac{1}{2}}^{r_i}\left[\frac{1}{(n^{\star})^2}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}\right)^3-\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^{\star})}\right)\right]du, \tag{C.12}$$

which can be re-written using the pre-defined functions in **Step 0** as:

$$
\begin{aligned}
&\frac{\partial^2 l(\kappa,\rho)}{\partial\rho^2} \\
&= -m\frac{\partial^2}{\partial\rho^2}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] - m + \frac{1}{(n^\star)^2}\sum_{i=1}^{m}A^2(\kappa,\rho;r_i,n^\star) + 4\sum_{i=1}^{m}I(\kappa,\rho;r_i,n^\star)
\end{aligned}
$$

where:

$$
\frac{\partial^2}{\partial\rho^2}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] = -\left(\frac{\partial}{\partial\rho}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\right)^2 + c(\kappa,\rho;n^\star)\left[\frac{\partial^2}{\partial\rho^2}c^{-1}(\kappa,\rho;n^\star)\right]
$$

and

$$
\begin{aligned}
&\frac{\partial^2}{\partial\rho^2}c^{-1}(\kappa,\rho;n^\star) \\
&= -\frac{1}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r_i,n^\star)\left[n^\star - \frac{2}{n^\star}A^2(\kappa,\rho;r,n^\star)\right]E(\kappa,\rho;r,n^\star)dr \\
&\quad -\frac{1}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right) - \frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]dr \\
&\quad -2\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r+\frac{1}{2}\right)H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)dr \\
&\quad +2n^\star\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r+\frac{1}{2}\right)\cdot\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right) - \frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]dr \\
&\quad +\frac{2}{(n^\star)^2}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)D(\kappa,\rho;r,n^\star)dr \\
&\quad -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right) - \frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]D(\kappa,\rho;r,n^\star)\right]dr \\
&\quad -\frac{2}{n^\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star B(\kappa,\rho;r,n^\star) - \frac{2}{n^\star}N(\kappa,\rho;r,n^\star)\right]dr
\end{aligned}
$$

Step 4: Derive the second cross partial derivatives of the likelihood function first with respect to the one parameter then the other. Based on the likelihood function C.8, the first partial derivative with respect to κ and the ρ is given by:

$$\begin{aligned}
&\frac{\partial^2 l(\kappa,\rho)}{\partial\rho\partial\kappa} \\
= &-m\frac{\partial^2}{\partial\rho\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] + \sum_{i=1}^{m}\frac{\partial^2}{\partial\rho\partial\kappa}\ln\left[\Psi^{-1}(\kappa,\rho;r_i,n^\star)\right] + 2\sum_{i=1}^{m}\frac{\partial^2}{\partial\rho\partial\kappa}\int_{u=-\frac{1}{2}}^{r_i}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du \\
= &-m\frac{\partial^2}{\partial\rho\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] + \sum_{i=1}^{m}\left(-r_i + \frac{r_i}{(n^\star)^2}\left[\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)}\right]^2\right) \\
&+2\sum_{i=1}^{m}\left(-2\int_{u=-\frac{1}{2}}^{r_i} u\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]du + \frac{2}{(n^\star)^2}\int_{u=-\frac{1}{2}}^{r_i}\left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right]^3 du\right) \\
= &-m\frac{\partial^2}{\partial\rho\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] + \sum_{i=1}^{m} r_i\left(\left[\frac{1}{(n^\star)}\cdot\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)}\right]^2 - 1\right) \\
&+4\sum_{i=1}^{m}\int_{u=-\frac{1}{2}}^{r_i}\left[\frac{1}{(n^\star)^2}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^3 - u\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)\right]du, \qquad \text{(C.13)}
\end{aligned}$$

which can be re-written using the pre-defined functions in **Step 0** as:

$$\begin{aligned}
&\frac{\partial^2 l(\kappa,\rho)}{\partial\rho\partial\kappa} \\
= &-m\frac{\partial^2}{\partial\rho\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] + \sum_{i=1}^{m} r_i\left[\left(\frac{\Lambda(\kappa,\rho;r_i,n^\star)}{(n^\star)}\right)^2 - 1\right] + 4\sum_{i=1}^{m}K(\kappa,\rho;r_i,n^\star)
\end{aligned}$$

where:

$$\begin{aligned}
&\frac{\partial^2}{\partial\rho\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] \\
= &-\left(\frac{\partial}{\partial\rho}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\right) + c(\kappa,\rho;n^\star)\left[\frac{\partial^2}{\partial\rho\partial\kappa}c^{-1}(\kappa,\rho;n^\star)\right]
\end{aligned}$$

and

$$
\begin{aligned}
&\frac{\partial^2}{\partial\rho\partial\kappa}c^{-1}(\kappa,\rho;n^*)\\
= &-\frac{1}{n^*}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} r\Psi^{-1}(\kappa,\rho;r,n^*)\left[n^*-\frac{2}{n^*}A^2(\kappa,\rho;r,n^*)\right]E(\kappa,\rho;r,n^*)dr\\
&-\frac{1}{n^*}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} rH(\kappa,\rho;r,n^*)E(\kappa,\rho;r,n^*)\left[2n^*\left(r+\frac{1}{2}\right)-\frac{2}{n^*}D(\kappa,\rho;r,n^*)\right]dr\\
&-\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)H(\kappa,\rho;r,n^*)E(\kappa,\rho;r,n^*)dr\\
&+n^*\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\Psi^{-1}(\kappa,\rho;r,n^*)E(\kappa,\rho;r,n^*)\left[2n^*\left(r+\frac{1}{2}\right)-\frac{2}{n^*}D(\kappa,\rho;r,n^*)\right]dr\\
&+\frac{2}{(n^*)^2}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} H(\kappa,\rho;r,n^*)E(\kappa,\rho;r,n^*)G(\kappa,\rho;r,n^*)dr\\
&-\frac{2}{n^*}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^*)E(\kappa,\rho;r,n^*)\left[2n^*\left(r+\frac{1}{2}\right)-\frac{2}{n^*}D(\kappa,\rho;r,n^*)\right]G(\kappa,\rho;r,n^*)dr\\
&-\frac{2}{n^*}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^*)E(\kappa,\rho;r,n^*)\left[2n^*F(\kappa,\rho;r,n^*)-\frac{2}{n^*}O(\kappa,\rho;r,n^*)\right]dr
\end{aligned}
$$

Step 5: Solve the likelihood equation for $\widehat{\kappa}$ and $\widehat{\rho}$.

Based on the given realisation of the random samples, $(r_i)_{i=1}^{m}$, the maximum likelihood estimates $\widehat{\kappa}$ and $\widehat{\rho}$ is the roots to the following likelihood equations:

$$
\begin{aligned}
0 &= \frac{\partial l(\kappa,\rho)}{\partial\kappa}\\
&= -m\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^*)\right]-m\left(\frac{n^*}{4}\right)+\sum_{i=1}^{m}\left[n^*r_i^2-\frac{r_i}{n^*}A(\kappa,\rho;r_i,n^*)\right]-\frac{2}{n^*}\sum_{i=1}^{m}G(\kappa,\rho;r_i,n^*),
\end{aligned}
\tag{C.14}
$$

and

$$
\begin{aligned}
0 &= \frac{\partial l(\kappa,\rho)}{\partial \rho} \\
&= -m\frac{\partial}{\partial \rho}\ln\left[c^{-1}(\kappa,\rho;n^{\star})\right] + mn^{\star} + \sum_{i=1}^{m}\left[2n^{\star}r_i - \frac{A(\kappa,\rho;r_i,n^{\star})}{n^{\star}}\right] - \frac{2}{n^{\star}}\sum_{i=1}^{m}D(\kappa,\rho;r_i,n^{\star}),
\end{aligned}
\tag{C.15}
$$

where:

$$
\begin{aligned}
&\frac{\partial}{\partial \kappa}\ln\left[c^{-1}(\kappa,\rho;n^{\star})\right] \\
= \quad & -\frac{c(\kappa,\rho;n^{\star})}{n^{\star}}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[r\cdot H(\kappa,\rho;r,n^{\star})\cdot E(\kappa,\rho;r,n^{\star})\right]dr \\
& + c(\kappa,\rho;n^{\star})\,n^{\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^{\star})\cdot E(\kappa,\rho;r,n^{\star})\cdot\left(r^2-\frac{1}{4}\right)\right]dr \\
& - \frac{2c(\kappa,\rho;n^{\star})}{n^{\star}}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^{\star})E(\kappa,\rho;r,n^{\star})G(\kappa,\rho;r,n^{\star})\right]dr,
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial}{\partial \rho}\ln\left[c^{-1}(\kappa,\rho;n^{\star})\right] \\
= \quad & -\frac{c(\kappa,\rho;n^{\star})}{n^{\star}}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[H(\kappa,\rho;r,n^{\star})\cdot E(\kappa,\rho;r,n^{\star})\right]dr \\
& + c(\kappa,\rho;n^{\star})\,2n^{\star}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^{\star})\cdot E(\kappa,\rho;r,n^{\star})\cdot\left(r+\frac{1}{2}\right)\right]dr \\
& - \frac{2c(\kappa,\rho;n^{\star})}{n^{\star}}\int_{r=-\frac{1}{2}}^{\frac{1}{2}}\left[\Psi^{-1}(\kappa,\rho;r,n^{\star})E(\kappa,\rho;r,n^{\star})D(\kappa,\rho;r,n^{\star})\right]dr.
\end{aligned}
$$

$$A(\kappa,\rho;u,n^\star) = \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}$$

$$G(\kappa,\rho;r,n^\star) = \int_{u=-\frac{1}{2}}^{r} u\cdot\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2 du = \int_{u=-\frac{1}{2}}^{r} u\cdot A^2(\kappa,\rho;u,n^\star)\,du$$

$$H(\kappa,\rho;r,n^\star) = \frac{\Pi(\kappa,\rho;u)}{\Psi^2(\kappa,\rho;u,n^\star)} = \Psi^{-1}(\kappa,\rho;u,n^\star)\cdot A(\kappa,\rho;u,n^\star)$$

$$E(\kappa,\rho;r,n^\star) = \exp\left[2\cdot\int_{u=-\frac{1}{2}}^{r}\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}du\right] = \exp\left[2\cdot B(\kappa,\rho;r,n^\star)\right]$$

$$D(\kappa,\rho;r,n^\star) = \int_{u=-\frac{1}{2}}^{r}\left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)}\right)^2 du = \int_{u=-\frac{1}{2}}^{r} A^2(\kappa,\rho;u,n^\star)\,du$$

Step 6: Calculate the variance-covariance matrix associated with the calculated estimates. In our case, a (2×2) observed information matrix defined as follows:

$$\left.\begin{pmatrix} \frac{\partial^2 l(\kappa,\rho)}{\partial\kappa^2} & \frac{\partial^2 l(\kappa,\rho)}{\partial\kappa\partial\rho} \\ \frac{\partial^2 l(\kappa,\rho)}{\partial\rho\partial\kappa} & \frac{\partial^2 l(\kappa,\rho)}{\partial\rho^2}\end{pmatrix}^{-1}\right|_{(\kappa=\hat{\kappa},\rho=\hat{\rho})},$$

where:

$$\frac{\partial^2 l(\kappa,\rho)}{\partial\kappa^2}$$
$$= -m\frac{\partial^2}{\partial\kappa^2}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] + \sum_{i=1}^{m}\left[\left(\frac{r_i}{n^\star}\right)^2\cdot A^2(\kappa,\rho;r_i,n^\star) - r_i^2\right] + 4\sum_{i=1}^{m} J(\kappa,\rho;r_i,n^\star),$$

$$\frac{\partial^2 l(\kappa,\rho)}{\partial\rho^2}$$
$$= -m\frac{\partial^2}{\partial\rho^2}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] - m + \frac{1}{(n^\star)^2}\sum_{i=1}^{m} A^2(\kappa,\rho;r_i,n^\star) + 4\sum_{i=1}^{m} I(\kappa,\rho;r_i,n^\star),$$

and

$$\frac{\partial^2 l(\kappa,\rho)}{\partial\kappa\partial\rho} = \frac{\partial^2 l(\kappa,\rho)}{\partial\rho\partial\kappa}$$

$$= -m\frac{\partial^2}{\partial\rho\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] + \sum_{i=1}^{m} r_i\left[\left(\frac{A(\kappa,\rho;r_i,n^\star)}{(n^\star)}\right)^2 - 1\right] + 4\sum_{i=1}^{m} K(\kappa,\rho;r_i,n^\star)$$

$$\frac{\partial^2}{\partial\kappa^2}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] = -\left(\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\right)^2 + c(\kappa,\rho;n^\star)\left[\frac{\partial^2}{\partial\kappa^2}c^{-1}(\kappa,\rho;n^\star)\right]$$

$$\frac{\partial^2}{\partial\rho^2}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] = -\left(\frac{\partial}{\partial\rho}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\right)^2 + c(\kappa,\rho;n^\star)\left[\frac{\partial^2}{\partial\rho^2}c^{-1}(\kappa,\rho;n^\star)\right]$$

$$\frac{\partial^2}{\partial\rho\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right] = -\left(\frac{\partial}{\partial\rho}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\frac{\partial}{\partial\kappa}\ln\left[c^{-1}(\kappa,\rho;n^\star)\right]\right) + c(\kappa,\rho;n^\star)\left[\frac{\partial^2}{\partial\rho\partial\kappa}c^{-1}(\kappa,\rho;n^\star)\right]$$

$$
\begin{aligned}
&\frac{\partial^2}{\partial \kappa^2} c^{-1}(\kappa,\rho;n^\star) \\
= \; & -\frac{1}{n^\star} \int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left[r\Psi^{-1}(\kappa,\rho;r,n^\star) \left(rn^\star - \frac{2r}{n^\star} A^2(\kappa,\rho;r,n^\star) \right) E(\kappa,\rho;r,n^\star) \right] dr \\
& -\frac{1}{n^\star} \int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} rH(\kappa,\rho;r,n^\star) E(\kappa,\rho;r,n^\star) \left[n^\star \left(r^2 - \frac{1}{4} \right) - \frac{2}{n^\star} G(\kappa,\rho;r,n^\star) \right] dr \\
& +n^\star \int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left(r^2 - \frac{1}{4} \right) \left(-\frac{r}{n^\star} H(\kappa,\rho;r,n^\star) \right) E(\kappa,\rho;r,n^\star) dr \\
& +n^\star \int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left(r^2 - \frac{1}{4} \right) \cdot \Psi^{-1}(\kappa,\rho;r,n^\star) E(\kappa,\rho;r,n^\star) \left[n^\star \left(r^2 - \frac{1}{4} \right) - \frac{2}{n^\star} G(\kappa,\rho;r,n^\star) \right] dr \\
& -\frac{2}{n^\star} \int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left(-\frac{r}{n^\star} H(\kappa,\rho;r,n^\star) E(\kappa,\rho;r,n^\star) G(\kappa,\rho;r,n^\star) \right) dr \\
& -\frac{2}{n^\star} \int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left(\Psi^{-1}(\kappa,\rho;r,n^\star) E(\kappa,\rho;r,n^\star) \left[n^\star \left(r^2 - \frac{1}{4} \right) - \frac{2}{n^\star} G(\kappa,\rho;r,n^\star) \right] G(\kappa,\rho;r,n^\star) \right) dr \\
& -\frac{2}{n^\star} \int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left[\Psi^{-1}(\kappa,\rho;r,n^\star) E(\kappa,\rho;r,n^\star) \left(2n^\star L(\kappa,\rho;r,n^\star) - \frac{2}{n^\star} M(\kappa,\rho;r,n^\star) \right) \right] dr.
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial^2}{\partial \rho^2} c^{-1}(\kappa,\rho;n^\star) \\
=\;& -\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \Psi^{-1}(\kappa,\rho;r_i,n^\star)\left[n^\star-\frac{2}{n^\star}A^2(\kappa,\rho;r,n^\star)\right]E(\kappa,\rho;r,n^\star)dr \\
& -\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]dr \\
& -2\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left(r+\frac{1}{2}\right)H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)dr \\
& +2n^\star\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left(r+\frac{1}{2}\right)\cdot\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]dr \\
& +\frac{2}{(n^\star)^2}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)D(\kappa,\rho;r,n^\star)dr \\
& -\frac{2}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \left[\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]D(\kappa,\rho;r,n^\star)\right]dr \\
& -\frac{2}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} \Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star B(\kappa,\rho;r,n^\star)-\frac{2}{n^\star}N(\kappa,\rho;r,n^\star)\right]dr
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial^2}{\partial\rho\partial\kappa}c^{-1}\left(\kappa,\rho;n^\star\right)\\
= \; & -\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} r\Psi^{-1}(\kappa,\rho;r,n^\star)\left[n^\star-\frac{2}{n^\star}A^2(\kappa,\rho;r,n^\star)\right]E(\kappa,\rho;r,n^\star)dr\\
&-\frac{1}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} rH(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]dr\\
&-\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)dr\\
&+n^\star\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\left(r^2-\frac{1}{4}\right)\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]dr\\
&+\frac{2}{(n^\star)^2}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}} H(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)G(\kappa,\rho;r,n^\star)dr\\
&-\frac{2}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star\left(r+\frac{1}{2}\right)-\frac{2}{n^\star}D(\kappa,\rho;r,n^\star)\right]G(\kappa,\rho;r,n^\star)dr\\
&-\frac{2}{n^\star}\int\limits_{r=-\frac{1}{2}}^{\frac{1}{2}}\Psi^{-1}(\kappa,\rho;r,n^\star)E(\kappa,\rho;r,n^\star)\left[2n^\star F(\kappa,\rho;r,n^\star)-\frac{2}{n^\star}O\left(\kappa,\rho;r,n^\star\right)\right]dr
\end{aligned}
$$

$$
\begin{aligned}
A(\kappa,\rho;u,n^\star) &= \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \\
J(\kappa,\rho;r,n^\star) &= \int\limits_{u=-\frac{1}{2}}^{r} u^2 \left(\frac{1}{(n^\star)^2} \left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \right]^3 - \left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \right] \right) du \\
&= \int\limits_{u=-\frac{1}{2}}^{r} u^2 \left(\frac{A^3(\kappa,\rho;u,n^\star)}{(n^\star)^2} - A(\kappa,\rho;u,n^\star) \right) du \\
E(\kappa,\rho;r,n^\star) &= \exp\left[2 \cdot \int\limits_{u=-\frac{1}{2}}^{r} \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} du \right] = \exp\left[2 \cdot B(\kappa,\rho;r,n^\star)\right] \\
H(\kappa,\rho;r,n^\star) &= \frac{\Pi(\kappa,\rho;u)}{\Psi^2(\kappa,\rho;u,n^\star)} = \Psi^{-1}(\kappa,\rho;u,n^\star) \cdot A(\kappa,\rho;u,n^\star) \\
G(\kappa,\rho;r,n^\star) &= \int\limits_{u=-\frac{1}{2}}^{r} u \cdot \left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \right)^2 du = \int\limits_{u=-\frac{1}{2}}^{r} u \cdot A^2(\kappa,\rho;u,n^\star)\, du \\
L(\kappa,\rho;r,n^\star) &= \int\limits_{u=-\frac{1}{2}}^{r} u^2 \left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \right) du = \int\limits_{u=-\frac{1}{2}}^{r_i} u^2 A(\kappa,\rho;u,n^\star)\, du \\
M(\kappa,\rho;r,n^\star) &= \int\limits_{u=-\frac{1}{2}}^{r_i} u^2 \left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \right)^3 du = \int\limits_{u=-\frac{1}{2}}^{r_i} u^2 A^3(\kappa,\rho;u,n^\star)\, du
\end{aligned}
$$

$$I(\kappa,\rho;r,n^\star) = \int\limits_{u=-\frac{1}{2}}^{r} \left(\frac{1}{(n^\star)^2} \left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \right]^3 - \left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \right] \right) du$$

$$= \int\limits_{u=-\frac{1}{2}}^{r} \left[\frac{A^3(\kappa,\rho;u,n^\star)}{(n^\star)^2} - A(\kappa,\rho;u,n^\star) \right] du$$

$$D(\kappa,\rho;r,n^\star) = \int\limits_{u=-\frac{1}{2}}^{r} \left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \right)^2 du = \int\limits_{u=-\frac{1}{2}}^{r} A^2(\kappa,\rho;u,n^\star)\, du$$

$$B(\kappa,\rho;r,n^\star) = \int\limits_{u=-\frac{1}{2}}^{r} \frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} du = \int\limits_{u=-\frac{1}{2}}^{r} A(\kappa,\rho;u,n^\star)\, du$$

$$N(\kappa,\rho;r,n^\star) = \int\limits_{u=-\frac{1}{2}}^{r} \left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \right)^3 du = \int\limits_{u=-\frac{1}{2}}^{r_i} A^3(\kappa,\rho;u,n^\star)\, du$$

$$K(\kappa,\rho;r,n^\star) = \int\limits_{u=-\frac{1}{2}}^{r} \left(\frac{1}{(n^\star)^2} \left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \right]^3 - u \left[\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \right] \right) du$$

$$= \int\limits_{u=-\frac{1}{2}}^{r} \left[\frac{A^3(\kappa,\rho;u,n^\star)}{(n^\star)^2} - u \cdot A(\kappa,\rho;u,n^\star) \right] du$$

$$F(\kappa,\rho;r,n^\star) = \int\limits_{u=-\frac{1}{2}}^{r} u \cdot \left(\frac{\Pi(\kappa,\rho;u)}{\Psi(\kappa,\rho;u,n^\star)} \right) du = \int\limits_{u=-\frac{1}{2}}^{r} u \cdot A(\kappa,\rho;u,n^\star)\, du$$

$$O(\kappa,\rho;r,n^\star) = \int\limits_{u=-\frac{1}{2}}^{r} u \left(\frac{\Pi(\kappa,\rho;r_i)}{\Psi(\kappa,\rho;r_i,n^\star)} \right)^3 du = \int\limits_{u=-\frac{1}{2}}^{r} uA^3(\kappa,\rho;u,n^\star)\, du$$

Appendix D

DESCRIPTION OF METHODOLOGY USED IN R/S ANALYSIS

D.1 Programming Solution to the Estimation of Sample Hurst Exponent

D.1.1 Detrending the time series of log returns: calculating AR(1) residual of the log returns

In order to detrend the observed daily returns time series, the R/S statistics are calculated from the time series of $AR(1)$ residuals in order to eliminate or minimise serial correlation (or linear dependency) between consecutive price indices. It is important to detrend the log return series, because serial correlation can bias the sample Hurst exponent by causing a significant sample Hurst exponent to be calculated when long-term memory does not exist in the log return series or a Type I error. Although, detrending will not remove all serial correlation in the log return series, but Brock, Dechert and Scheinkman (1987) have shown that detrending is able to minimise serial correlation to insignificant level AR processes of up to order 3.

D.1.2 Selecting length of sub-time interval based on the total number of observations of AR(1) residuals

The selection of sub-time interval $n \in Z^+$ based on the total number of observations in a given price series is crucial in the calculation of R/S statistics. This is because R/S statistics is a function of n and the values of n used will determine the statistical properties of R/S statistics. Sub-time intervals n are selected based on the following requirements:

1. *Sub-time intervals should be non-overlapping and each of length n.*

2. *Use the N (the number of AR(1) residuals to be used in the calculation of R/S statistics) that:*

(i) *Has the largest number of factors.* This ensures that the maximum number of R/S statistics are generated from the series so that the estimated sample Hurst exponent will be reliable. Thus, it is more desirable to have more R/S statistics values than using all the observations of AR(1) residuals.

(ii) *Uses the most number of observations of AR(1) residual.* This will require N to be an even-numbered positive integer less than or equal to the total number of AR(1) residuals observed. This is done to ensure that both the beginning and the end points of the residuals series upon which R/S statistics will be calculated are used in the calculation. Using a N that leaves too many of the observations of AR(1) residuals unused will bias the R/S statistics, because the subjective choice of starting point in the observed AR(1) residuals series will influence the value of R/S statistics for large values of n and cause it to be unstable over time.

3. $A \cdot n = N$, where $A \in Z^+$ *is the number sub-time intervals a given n. Calculate values of n by finding all factors of the N chosen from above that are greater than or equal to 10.* The minimum requirement that n must be greater than or equal to 10 ensures that statistics calculated on each sub-interval is a reasonable summary for that sub-interval. This enables robust R/S statistics to be calculated from these sub-interval statistics.

1. **Remark 100** *A dominant feature of R/S statistics is that its statistical properties are heavily dependent upon the number and reliability of sub-interval statistics.*

Remark 101 *The problem of the bias introduced by the choice of starting point emphasises the importance of the frequency of the observation. Choice of starting point will not significantly influence the R/S statistics for intra-day or daily observations, because it is most unlikely that the underlying characteristic of the market is likely to change over a few months. But for annual observations, the choice of starting point will profoundly influence the R/S statistics as it is well known that business cycles are roughly 4 years in length.*

Based on the above objectives, the following values were chosen based on the given sample of 3824 observations of AR(1) residuals:

2. By trial and error, N was chosen to be 3800. It was decided to discard the first 24 AR(1) residuals in the calculation of R/S statistics.

3. Using EXCEL, the values of n are calculated to be:

$$\{10, 19, 20, 25, 38, 40, 50, 76, 95, 100, 152, 190, 200, 380, 475, 760, 950, 1900\}$$

4. Thus, there are 18 factors in 3800 that is greater than or equal to 10,i.e.) A=18. This implies that the given AR(1) residuals series will produce 18 R/S statistics.

Remark 102 *It is possible to include in the program to calculate R/S statistic, an algorithm that is capable of choosing N and calculate resulting n values. Also, there exists softwares like S-Plus that can easily perform such functions. However, due to the time constraint and the lack of resources to obtain the necessary softwares, these tasks were carried out by trial and error and brute force computation.*

D.1.3 Calculate R/S statistics from the given AR(1) residuals series

In order to calculate the R/S statistics from the AR(1) residuals series, a computer program was written using the **Visual Basic for Applications (VBA)** programming codes in Visual Basic editor supplied with the **Microsoft Excel 97** spreadsheet software. But, to explain the tasks performed by the program, it is necessary to define the following notations:

- Let the given $AR(1)$ series be denoted as follows:

$$\widetilde{X} = (X_3, X_4, \cdots, X_i, \cdots, X_{3826}).$$

- Let the calculated $n's$ be ordered in increasing value and denoted by:

$$\widetilde{n} = (n_1 = 10, n_2 = 19, \cdots, n_{18} = 1900)$$

Then, for $n = n_i (i = 1, 2, \cdots, 18)$, there will be $\frac{3800}{n_i}$ sub-intervals on $\widetilde{X}$, each of length n. Index each sub-interval by a $(a = 1, 2, \cdots, \frac{3800}{n})$.

- Let $X_{i,a}$ denote the AR(1) residual in the i^{th} position of $\widetilde{X}$, which is within the a^{th} sub-interval.

Given the above definitions, the program performs the required tasks according to the following algorithm. Start the algorithm by using $n = 10$, the value of n in the first position $(i = 1)$ of $\widetilde{n}$. Then, the number of sub-intervals is $\frac{3800}{10} = 380$. For $n = n_i$ $(i = 1, 2, \cdots, 18)$, the i^{th} loop of the algorithm calculates the following:

1. **Calculate sample mean of each sub-interval of length $n = n_i$ $(i = 1, 2, \cdots, 18)$.** The sample mean of the a^{th} $(a = 1, 2, \cdots, \frac{3800}{n_i})$ sub-interval, each of length $n = n_i(i = 1, 2, \cdots, 18)$, is calculated by:

$$\overline{X_a} = \frac{1}{n}\left(\sum_{j=1}^{n} X_{j,a}\right),$$

where $X_{j,a}$ is the AR(1) residual in position j $(j = 1, \cdots, n)$ of the a^{th} sub-interval on $\widetilde{X}$.

2. **Calculate the time series of accumulated departures from sample mean for each sub-interval.** For $n = n_i$ $(i = 1, 2, \cdots, 18)$, the j^{th} accumulated departure of the AR(1) residual in position j $(j = 1, \cdots, n)$ of a^{th} sub-interval on $\widetilde{X}$ from the sample mean for this sub-interval is calculated by:

$$\epsilon_{j,a} = \sum_{r=1}^{j} (X_{j,a} - \overline{X_a})$$

3. **Calculate the adjusted range for each sub-interval.** The adjusted range for the a^{th} sub-interval, denoted R_a, is defined as the difference between the maximum and the minimum value of $\epsilon_{j,a}$ for each sub-interval and is calculated by:

$$R_a = \max_{1\leq j\leq n}(\epsilon_{j,a}) - \min_{1\leq j\leq n}(\epsilon_{j,a})$$

4. **Calculate the sample standard deviation for each sub-interval.** The computational form of the sample standard deviation for the a^{th} sub-interval on $\widetilde{X}$, denoted s_a is as follows:

$$s_a = \sqrt{\frac{1}{n-1}\left[\left(\sum_{j=1}^{n} X_{j,a}^2\right) - \frac{1}{n}\left(\sum_{j=1}^{n} X_{j,a}\right)^2\right]},$$

where the $n-1$ is to reflect the fact that sample standard deviation is an estimate the population standard deviation..

5. **Calculate the re-scaled range for each sub-interval.** For the a^{th} sub-interval, calculate the re-scaled range by dividing the adjusted range by the sample standard deviation, i.e.):

$$(R/S)_a = \frac{R_a}{s_a}.$$

6. **Calculate the R/S statistics for $n = n_i$.** Having calculated $\frac{3800}{n_i}$ re-scaled ranges, the R/S statistics for the i^{th} n is the arithmetic average of these re-scaled ranges:

$$(R/S)_n = \left(\frac{3800}{n_i}\right)^{-1} \cdot \sum_{a=1}^{\frac{3800}{n_i}} (R/S)_a \,.$$

In the next section, the set of R/S statistics generated from running the above program:

$$\{(R/S)_n \,; n \in \widetilde{n}\}$$

will then be used to estimate the sample Hurst exponent.

D.1.4 Estimation of sample Hurst exponent

Based on the set of 18 R/S statistics generated, the sample Hurst exponent is estimated as follows:

1. Calculate $Log_{10}[(R/S)_n]$ and $Log_{10}(n)$ over all $n \in \tilde{n}$.

2. Run a simple linear regression of $Log_{10}[(R/S)_n]$ (dependent variable) against $Log_{10}(n)$ independent variable. Estimate the regression constant α and regression coefficient β, using the "LINEST" function in the EXCEL spreadsheet. The outputs from this function is displayed as follows:

β	α
se_β	se_α
R^2	se_y
F-statistics	d.f.
ss_{reg}	ss_{resid}

where the terms used are as follows:

- se: The standard error values, where the subscript indicates the coefficient.
- Se_b : The standard error value for the constant b (se_b = #N/A when const is FALSE).
- R^2 : The coefficient of determination. Compares estimated and actual y-values, and ranges in value from 0 to 1. If it is 1, there is a perfect correlation in the sample $\frac{3}{4}$ there is no difference between the estimated y-value and the actual y-value. At the other extreme, if the coefficient of determination is 0, the regression equation is not helpful in predicting a y-value. For information about how r2 is calculated, see "Remarks" later in this topic.
- se_y : The standard error for the y estimate.
- F :The F statistic, or the F-observed value. Use the F statistic to determine whether the observed relationship between the dependent and independent variables occurs by chance.

- *df* :The degrees of freedom. Use the degrees of freedom to help you find F-critical values in a statistical table. Compare the values you find in the table to the F statistic returned by LINEST to determine a confidence level for the model.

- ss_{reg} : The regression sum of squares.

- ss_{resid} : The residual sum of squares.

The regression coefficient β estimated will be the sample Hurst exponent over all $n \in \widetilde{n}$.

Remark 103 *Note that in all proceeding analysis, the log is taken to the base 10 in accordance with the convention adopted by Peters(1994). However, recall that the Hurst exponent is the slope estimate of the log-log plot, thus it does not matter if the log is to the base of 10 or a natural log as long as the same log base is applied to both $(R/S)_n$ and n. What does change is that natural log scale will be larger than the base 10 scale, which will imply different x-intercept.*

D.1.5 Interpreting the sample Hurst exponent: randomness and persistence

Independent process

If $H = 0.5$, then the events that occurs in the system is i.i.d., implying that they are normally distributed and that it is an independent process.

Persistent time series

For $H \in (0.5, 1]$, the re-scaled range is scaling at a faster rate than the square root of time, which means that the system under study is covering more distance than would a random process. In order for this to happen, the time series of observations has to be serially correlated.

Time series with the above-mentioned H value is termed to be a persistent time series and these time series are characterised by long-term memory effect. This long-term memory occurs regardless of the time scale being used and hence does not have a characteristic time scale, which is one of the key features of Fractal time series as described earlier.

Theoretically, long-memory process implies that the initial value that the system is subjected to will impact the future indefinitely. In chaotic or non-linear dynamics, this amounts to the system being sensitive to its initial conditions. Although, there are Auto-Regressive (AR) processes that may cause this short-term correlation, but it is unlikely that a simple AR(1) or AR(2) process may be the cause. Hurst felt that this process was caused by a biased random walk. In terms of financial time series, this implies that if the market behaves according to a Hurst process, then the time series observed would exhibit trending that will persist until a randomly occurring economic event takes place that will change the bias in both its magnitude and its direction.

Anti-persistent time series

If $H \in [0, 0.5)$, then re-scaled range is scaling at a slower rate than the square root of time. This means that the system under study is covering less distance than would a random process. In order for this to happen, the time series of observations has to be reversing the direction of its trend more frequently than a random process. However, this is not the same as a mean reverting process as mean reverting processes assumes that the system under study has a stable mean and there is no justification to make such assumption in R/S analysis.

Comparing the characteristics of Hurst process with the fractals

The following two properties of the Hurst process is similar to that of the characteristics of fractals:

1. Hurst process scales according to the power law. In the case of time series, this means that as the increments of time increases, the range increases by some power of the time increment. This is synonymous with that of fractals.
2. Re-scaled range analysis is able to describe time series that have no characteristic scale, which is also a characteristic of fractals.

D.2 Programming Solution to the of Expected Hurst Exponent under the Null Hypothesis

Peters(1994) empirically derived the following formula, which may be used to generate the expected values of the R/S statistics and hence estimate the expected Hurst exponent under the null hypothesis:

Proposition 104 *Under the null hypothesis, the values of the R/S statistics generated from an i.i.d. data may be calculated as follows:*

$$E[(R/S)_n] = \left[\left(\frac{n-0.5}{n}\right)\cdot\left(\frac{\pi}{2}\cdot n\right)\right]^{-\frac{1}{2}}\cdot\sum_{r=1}^{n-1}\sqrt{\frac{n-r}{r}}, \tag{D.1}$$

for all subintervals of length $n \in [0, N]$.

Remark 105 *In theory, any range will be appropriate provided the system under study and the $E\left[(R/S)_n\right]$ covers the same set of subintervals, n.*

Remark 106 *For the analysis of financial time series, sub-intervals length should begin with length 10. The final sub-interval length will depend on the system under study.*

Proposition 107 *Given that the R/S statistics is a normally distributed random variable under the null hypothesis, thus one would expect the Hurst exponent estimated as a result will also be normally distributed. Then, the variance of the expected Hurst exponent (under null hypothesis) is given by:*

$$Var\left(E\left[H|T\right]\right) = \frac{1}{T}, \tag{D.2}$$

where T is the total number of observations in the given time series.

Remark 108 *Peters (1994) numerically tested the validity of equation (D.2) by Monte Carlo experiment.*

Remark 109 *Note that the equation (D.2) does not depend on the n used or the value of H. Thus, the variance of the expected Hurst exponent calculated using equation (D.1) only depends on the total number of observations in the given time series. This suggests that there might not be enough power for this hypothesis test to reject H_0.*

Conjecture 110 *Based on equations (D.1) and (D.2), one can conclude that the null distribution for the Hurst exponent is:*

$$E\left(H|T\right) \sim Normal\left(E\left(H\right), \frac{1}{T}\right).$$

Remark 111 *For non-Gaussian independent random processes, Peters (1994) has shown that the mean value of these processes can be predicted by equation (D.1), but its variance differs with the Gaussian independent processes. Since variance for distributions that are not normal differs on an individual basis, thus the confidence interval for $E[H|T]$ can only be constructed under null hypothesis and not for the sample Hurst exponent.*

The expected Hurst exponent under the null hypothesis is estimated as follows:

1. Writing a computer program to calculate the $E[(R/S)_n]$ values, using equation (D.1), for those values of n used in the estimation of sample Hurst exponent.

2. Estimate the expected Hurst exponent from the $E[(R/S)_n]$ values generated by the program written above by Ordinary Least Squares (OLS) estimation.

D.2.1 Hypothesis Testing in R/S Analysis

Having calculated the values of the R/S Statistics and then estimated the Hurst exponent from the given time series, one needs to set up a hypothesis test for statistical significance of the sample Hurst exponent. The null hypothesis for the test is:

H_0: *Underlying process is random (stationary) and hence normally distributed*

First of the main problems with R/S analysis is that for hypothesis testing, it is analytically intractable to derive the distribution for the R/S statistics under the null hypothesis. Thus, in order to set up a hypothesis test for R/S analysis, one has to:

1. Use computer simulation to calculate the expected value of the R/S statistics and estimate the expected Hurst exponent against which the sample Hurst exponent may be compared,

2. Standardise the sample Hurst exponent and test the hypothesis at 5% significance level. The hypothesis test is two-tailed.

3. Apply asymptotic theory in order to assess the statistical significance of the sample Hurst exponent.

D.3 The V-statistics method for Identifying Cycles within the JSE

This is a more precise measure of the cycle length, which works particularly well in the presence of noise.

Definition 112 *The V-statistics is defines as:*

$$V_n = \frac{(R/S)_N}{\sqrt{n}}.$$

Then, to identify and estimate the cycle length in a time series, the V-statistic method are applied in the following way:

1. Calculate the V-statistics V_n and $\log(n)$. Plot V_n against $log(n)$, which can be graphically interpreted as follows:

   Flat if the process was an independent, random process,

   Upward sloping, if the process was persistent,

   Downward sloping, if the process was anti-persistent..

2. Examining the value of V_n for each sub-time interval, n to estimate the cycle length for each frequency.