

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

# Option Pricing using Hidden Markov Models



Michael Anderson

**MSc** Thesis Submitted on **14 July 2006**

Presented for the Degree of **Master of Science in Mathematics of Finance** in the Department of Statistical Sciences, University of Cape Town, South Africa.

Supervisor: Professor Renkuan Guo

## Declarations

### Copyright:

I hereby grant the University of Cape Town permission to copy and disseminate this work, or any part thereof, using any means, for the purposes of study and research.

### Plagiarism:

1. This dissertation is my own work.
2. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work
3. I have used the American Psychological Association convention for citation and referencing. Each significant contribution to, and quotation in, this dissertation from the work of other people has been cited and referenced.

Michael Anderson (ANDMIC011)

Date: 14 July 2006

## Abstract

The Black-Scholes model for option pricing allows for preference-free, and thus unique, option prices to be calculated. Due to this feature, it represented a major step forward in the theory of option pricing. However, it also features several flaws, chief of which is the fact that it assumes that the parameters required for option pricing remain constant over time. This is unlikely to be the case, however, since these parameters have been observed to vary over time in the financial markets. This work will present an option pricing model that accommodates parameters that vary over time, whilst still retaining a closed-form expression for option prices: the Hidden Markov Option Pricing Model. This is possible due to the macro-structure of this model and provides the added advantage of ensuring efficient computation of option prices. This model turns out to be a very natural extension to the Black-Scholes model, allowing for time-varying input parameters. The model is also very flexible and extensible, as will be illustrated by a sample of the possible extensions that can be made to the 'basic form' of the model. The model will also be applied to develop a new approach for valuing Bermudan-style options, which have several possible exercise dates. This approach has the advantage, over traditional methods for valuing such options such as finite difference methods, of accommodating input parameters that vary over time. This may allow for more accurate pricing of these options. The models presented will all be fitted to data from the UK and South African markets in order to provide an illustration of the use of the models, and an indication of their performance.

## Acknowledgements

I would like to thank my supervisor, Professor R. K. Guo, for all his assistance and guidance in the preparation of this work. His suggestions and insight have made this a much better work than it otherwise would have been. His patience in fielding my barrages of questions also deserves recognition, and for this I am especially grateful.

University of Cape Town

## Table of Contents

<i>Declaration</i>	i
<i>Abstract</i>	ii
<i>Acknowledgements</i>	iii
<i>Table of Contents</i>	iv
<i>List of Figures</i>	viii
<i>List of Tables</i>	xii
<i>List of Notation</i>	xiv
<b>Chapter 1. Introduction</b>	<b>1</b>
1.1 Aims of this work	2
1.2 Methodology	3
1.3 Data	4
<b>Chapter 2. The Hidden Markov Model Framework</b>	<b>6</b>
<b>Chapter 3. Options and the Black-Scholes Option Pricing Formula</b>	<b>9</b>
3.1 Basic Options Theory	9
3.2 Pricing Options: The Black-Scholes Formula	10
3.2.1 Pricing via the Black-Scholes PDE	12
3.2.2 Pricing via Risk Neutral Methods	14
3.2.3 Key Features of Option Pricing	17
<b>Chapter 4. Why use Hidden Markov Models to Price Options</b>	<b>19</b>
<b>Chapter 5. Pricing, Arbitrage and Market Completeness when Hidden Markov Models are used to Model Asset Prices</b>	<b>25</b>
5.1 Change of State Securities	25
5.2 The Local Risk Neutral Valuation Relationship	26
5.3 The Regime-switching Esscher Transform	27
<b>Chapter 6. The Hidden Markov Option Pricing Model</b>	<b>31</b>
6.1 The Discrete-time Version of the Model	31
6.2 The Continuous-time Version of the Model	38

<b>Chapter 7. Extensions to the Basic Hidden Markov Option Pricing Model</b>	<b>44</b>
7.1 Time-varying Transition Probabilities	44
7.2 GARCH effects in the Variance of the Observed Process	48
7.3 Feedback effects between the State-dependent Variables and the Observed Values	49
7.4 Modelling Information Effects in the Market	50
<b>Chapter 8. Pricing Bermudan-style Options using the Hidden Markov Option Pricing Model</b>	<b>53</b>
<b>Chapter 9. A Lattice Pricing Method for the HMM Framework</b>	<b>56</b>
<b>Chapter 10. Fitting the HMM to the Variable Underlying the Option Contract</b>	<b>62</b>
10.1 A Basic Summary Analysis of the Data	62
10.1.1 JSE All Share Index Data	63
10.1.2 FTSE 100 Index Data	65
10.2 Estimating the Parameters of the HMM of the Log return Data	66
10.2.1 The Update Equations used for Estimating the Parameters and their Standard Errors	67
10.2.2 Selecting the number of States and the Start-up Parameters	72
10.2.3 Results of Fitting the HMM's to the Log return Data	74
<b>Chapter 11. Pricing Options on the JSE All Share and FTSE 100 Indices using the fitted HMM Parameters</b>	<b>85</b>
11.1 Option Prices Computed using the 2-state Model of the JSE Data	85
11.2 Option Prices Computed using the 3-state Model of the JSE Data	90
11.3 Option Prices Computed using the 2-state Model of the FTSE 100 Data	93

Option Pricing using Hidden Markov Models	vi
11.4 Option Prices Computed using the 3-state Model of the FTSE 100 Data	96
11.5 Brief General Conclusions	99
<b>Chapter 12. Calculating Prices for Bermudan-style Options using the Hidden Markov Option Pricing Model</b>	<b>101</b>
12.1 Basic Summary Analysis of the Data	102
12.2 Estimated Values for the 2-state Model of the Data	104
12.3 Estimated Values for the 3-state Model of the Data	106
12.4 Pricing Bermudan-style Payer Swaptions	108
12.4.1 Swaption Prices Computed using the 2-state Model of the Data	109
12.4.2 Swaption Prices Computed using the 3-state Model of the Data	113
<b>Chapter 13. Pricing Options on the JSE All Share and FTSE 100 Indices using an HMM with Time-varying Transition Probabilities</b>	<b>117</b>
13.1 Estimating the Model Parameters and the Occupation Time Probabilities	117
13.2 Results for the JSE Data	118
13.2.1 Parameter Estimates for the 2-state Model of the JSE Data	118
13.2.2 Option Prices Computed using the 2-state Model of the JSE Data	119
13.2.3 Parameter Estimates for the 3-state Model of the JSE Data	123
13.2.4 Option Prices Computed using the 2-state Model of the JSE Data	124
13.3 Results for the FTSE 100 Data	126
13.3.1 Parameter Estimates for the 2-state Model of the FTSE 100 Data	126

13.3.2 Option Prices Computed using the 2-state Model of the FTSE 100 Data	127
13.3.3 Parameter Estimates for the 3-state Model of the FTSE 100 Data	130
13.3.4 Option Prices Computed using the 2-state Model of the FTSE 100 Data	131
13.4 Some Conclusions	133
<b>Chapter 14. Conclusions</b>	<b>134</b>
14.1 Advantages to using the Hidden Markov Option Pricing Model	134
14.2 Problems with the Model	135
14.3 Performance of the Model when used to Evaluate European-style Index Option Prices	136
14.4 Performance of the Model when used to Evaluate Bermudan-style Option Prices	138
14.5 Performance of the Model with Time-varying Transition Probabilities	139
14.6 Further Areas of Work on this Model	139
<i>Glossary</i>	141
<i>Bibliography</i>	144

## List of Figures

Figure 1	A graphical representation of the HMM dependence structure	7
Figure 2	Payoff functions of a call and a put option at exercise	10
Figure 3	Mixtures of Normal Distributions	21
Figure 4	The pentanomial lattice structure of Bollen (1998)	58
Figure 5	A comparison of the recombining properties of the quadrinomial and pentanomial lattices	58
Figure 6	Histogram of the log returns of the JSE All Share Index	64
Figure 7	Histogram of the log returns of the FTSE 100 Index	66
Figure 8	Smoothed Probability of being in State 1 in the 2-state Model of the JSE Data	76
Figure 9	Smoothed Probability of being in State 2 in the 2-state Model of the JSE Data	76
Figure 10	Smoothed Probability of being in State 1 in the 3-state Model of the JSE Data	78
Figure 11	Smoothed Probability of being in State 2 in the 3-state Model of the JSE Data	78
Figure 12	Smoothed Probability of being in State 3 in the 3-state Model of the JSE Data	79
Figure 13	Smoothed Probability of being in State 1 in the 2-state Model of the FTSE 100 Data	80
Figure 14	Smoothed Probability of being in State 2 in the 2-state Model of the FTSE 100 Data	81
Figure 15	Smoothed Probability of being in State 1 in the 3-state Model of the FTSE 100 Data	83
Figure 16	Smoothed Probability of being in State 2 in the 3-state Model of the FTSE 100 Data	83

Figure 17	Smoothed Probability of being in State 3 in the 3-state Model of the FTSE 100 Data	84
Figure 18	Implied Volatility for Call options under the 2-state Model of the JSE Data	87
Figure 19	Implied Volatility for 30 and 60 day Call options under the 2-state Model of the JSE Data	88
Figure 20	Implied Volatility for Put options under the 2-state Model of the JSE Data	89
Figure 21	Implied Volatility for Call options under the 3-state Model of the JSE Data	91
Figure 22	Implied Volatility for Call options (2) under the 3-state Model of the JSE Data	91
Figure 23	Implied Volatility for Put options under the 3-state Model of the JSE Data	92
Figure 24	Implied Volatility for Call options under the 2-state Model of the FTSE 100 Data	94
Figure 25	Implied Volatility for Put options under the 2-state Model of the FTSE 100 Data	94
Figure 26	Implied Volatility for Put options (2) under the 2-state Model of the FTSE 100 Data	95
Figure 27	Implied Volatility for Call options under the 3-state Model of the FTSE 100 Data	97
Figure 28	Implied Volatility for Call options (2) under the 3-state Model of the FTSE 100 Data	98
Figure 29	Implied Volatility for Put options under the 3-state Model of the FTSE 100 Data	98
Figure 30	Histogram of the log returns of the 5-year UK Swap Rate	103
Figure 31	Smoothed Probability of being in State 1 in the 2-state Model of the Data	105
Figure 32	Smoothed Probability of being in State 2 in the 2-state Model of the Data	105

Option Pricing using Hidden Markov Models		x
Figure 33	Smoothed Probability of being in State 1 in the 3-state Model of the Data	107
Figure 34	Smoothed Probability of being in State 2 in the 3-state Model of the Data	107
Figure 35	Smoothed Probability of being in State 3 in the 3-state Model of the Data	108
Figure 36	Swaption prices for 91-day swaptions under the 2-state model of the swap data, with 6 possible exercise dates	112
Figure 37	Swaption prices for 182-day swaptions, under the 2-state model of the swap data, with 6 possible exercise dates	113
Figure 38	Swaption prices for 60-day swaptions, under the 3-state model of the swap data, with 2 possible exercise dates	116
Figure 39	Swaption prices for 91-day swaptions, under the 3-state model of the swap data, with 6 possible exercise dates	116
Figure 40	Implied Volatility for Call options under the 2-state model, with time-varying transition probabilities, of the JSE Data	120
Figure 41	Implied Volatility for Put options under the 2-state model, with time-varying transition probabilities, of the JSE Data	121
Figure 42	Implied Volatility for Put Options under the JSE 2-state model, positive solution	122
Figure 43	Implied Volatility for Call options under the 3-state model, with time-varying transition probabilities, of the JSE Data	125
Figure 44	Implied Volatility for Put options under the 3-state model, with time-varying transition probabilities, of the JSE Data	125
Figure 45	Implied Volatility for Call options under the 2-state model, with time-varying transition probabilities, of the FTSE 100 Data	128
Figure 46	Implied Volatility for Put options under the 2-state model, with time-varying transition probabilities, of the FTSE 100 Data	129

Option Pricing using Hidden Markov Models		xi
Figure 47	Implied Volatility for Call options under the 3-state model, with time-varying transition probabilities, of the FTSE 100 Data	132
Figure 48	Implied Volatility for Put options under the 3-state model, with time-varying transition probabilities, of the FTSE 100 Data	132

University of Cape Town

## List of Tables

Table 1	Call option prices computed using the 2-state Model of the JSE Data	85
Table 2	Put option prices computed using the 2-state Model of the JSE Data	86
Table 3	Call option prices computed using the 3-state Model of the JSE Data	90
Table 4	Put option prices computed using the 3-state Model of the JSE Data	90
Table 5	Call option prices computed using the 2-state Model of the FTSE 100 Data	93
Table 6	Put option prices computed using the 2-state Model of the FTSE 100 Data	93
Table 7	Call option prices computed using the 3-state Model of the FTSE 100 Data	96
Table 8	Put option prices computed using the 3-state Model of the FTSE 100 Data	96
Table 9	Bermudan swaption prices computed using the 2-state model of the swap data	109
Table 10	Bermudan swaption prices computed using the lattice method of Bollen (1998)	110
Table 11	Bermudan swaption prices computed using a trinomial tree method	110
Table 12	Bermudan swaption prices computed using the 3-state model of the swap data	113
Table 13	Bermudan swaption prices computed using a trinomial tree method (2)	114
Table 14	Bermudan swaption prices computed using a finite difference method	114

Table 15	Call option prices computed using the 2-state Model of the JSE Data with time-varying Transition Probabilities	119
Table 16	Put option prices computed using the 2-state Model of the JSE Data with Time-varying Transition Probabilities	119
Table 17	Call option prices computed using the 3-state Model of the JSE Data with Time-varying Transition Probabilities	124
Table 18	Put option prices computed using the 3-state Model of the JSE Data with Time-varying Transition Probabilities	124
Table 19	Call option prices computed using the 2-state Model of the FTSE 100 Data with Time-varying Transition Probabilities	127
Table 20	Put option prices computed using the 2-state Model of the FTSE 100 Data with Time-varying Transition Probabilities	127
Table 21	Call option prices computed using the 3-state Model of the FTSE 100 Data with Time-varying Transition Probabilities	131
Table 22	Put option prices computed using the 3-state Model of the FTSE 100 Data with Time-varying Transition Probabilities	131

## List of Notation

$(\cdot)^+$ :	Operator that returns the maximum of the argument and 0
$\langle \cdot, \cdot \rangle$ :	Inner product operator
$C()$ :	Call price, for a given set of arguments
$\Delta t$ :	Size of time step in a lattice, as a fraction of the overall time to expiry of the option being priced
$E[A   B]$ :	Conditional expectation of A, conditional on B
$f(\cdot)$ :	Probability density function
$f_{ij}(t)$ :	Probability of first passage from hidden state i to hidden state j at time t
$\mathcal{F}_t$ :	Filtration at time t. Natural filtration unless otherwise specified
$I_A[\cdot]$ :	Indicator function of event A, evaluates to 1 if the argument is A and 0 otherwise
$K$ :	Strike price of a given option
$N$ :	Number of hidden states in a given Hidden Markov Model
$\lambda_t$ :	Market price of risk at time t, where $\lambda = \frac{\mu - r}{\sigma}$ , the excess mean rate of return per unit of volatility, or 'risk'
$\mu$ :	Mean rate of return, per annum
$\mu_t$ :	Annualised mean rate of return at time t, if time dependent
$O_i^t$ :	Occupation time in hidden state i between time 0 and time t
$\Pr(A)$ :	Probability of event A occurring
$\Pr(A   B)$ :	Conditional probability of event A occurring, conditional on B occurring
$\pi(j)$ :	Probability that the hidden state in force at time 0 is State j
$p_{ij}$ :	Probability of moving from hidden state i to hidden state j in one time step

$p_{ij}(t)$ :	Probability of moving from hidden state $i$ to hidden state $j$ after $t$ time steps
$p_{hh}$ :	Probability of remaining in the high-volatility state in the lattice model of Bollen (1998)
$p_{ll}$ :	Probability of remaining in the low-volatility state in the lattice model of Bollen (1998)
<b>P</b> :	Matrix of transition probabilities $p_{ij}$
$P$ :	Objective, or 'real world' probability measure
$Q$ :	Risk neutral probability measure
$\frac{dQ}{dP}$ :	Radon-Nikodým derivative of $Q$ with respect to $P$
$\theta$ :	Vector of parameters for a given model
$r$ :	Risk-free rate of return, per annum
$r_t$ :	Annualised risk-free rate of return at time $t$ , if time dependent
$\sigma(X_u; 0 \leq u \leq t)$ :	Smallest sigma-algebra generated by the hidden state process from time 0 to time $t$
$\sigma$ :	Annualised volatility value
$\sigma_t$ :	Annualised volatility value at time $t$ , if time dependent
$S_t$ :	Value of the asset that the option contract is written on, at time $t$
$T$ :	Date of final value in the sample used to fit the model, measured from time 0 and thus also the number of time points in the sample
$T_i$ :	Date of $i$ -th possible exercise date in the case of a Bermudan-style option, measured from time 0
$\tau$ :	Expiry date of the option in question, measured from time 0
$\zeta_i$ :	Value that the occupation time in State $i$ can take, being one of a set of fixed, deterministic values
$W_t$ :	Value of a Brownian motion, with respect to $P$ , at time $t$
$\hat{W}_t$ :	Value of a Brownian motion, with respect to $Q$ , at time $t$
$\{X_t\}$ :	Hidden state process

$X_t$ :	Hidden state value at time $t$
$\{Y_t\}$ :	Observed value process, for a given HMM
$Y_t$ :	Value of observed series, in an HMM, at time $t$

Clarification: Although the variable that an option is written on will be referred to as 'the asset' in the text below, it can in practice be a variable such as an interest rate, which is not strictly an 'asset' in the conventional sense

## Chapter 1. Introduction

The well-known Black-Scholes option pricing formula, developed by Black and Scholes (1973), overcame a major obstacle in the theory and practice of pricing options. This was the dependence of previous option pricing methods on the individual risk preferences of the investor pricing the option. This arose from the need to model future values of the asset that the option contract was based on, to determine the value of the option to the holder in the future. These projected future values could then be used to estimate the present value of the option to the holder. Most of the models of the future value of the asset in question included a term representing the mean expected rate of return on the asset. The return that a particular investor will demand (expect) from a given asset will depend on the level of risk associated with those returns and the preferences of the investor with regard to risk, i.e. risk averse, risk seeking, etc. Thus, any option pricing formula that depends on a model of future asset values which includes the expected mean rate of return on the asset depends on the individual risk preferences of the investor concerned. This, in turn, resulted in an option price that depended on the individual performing the pricing. This meant that it was impossible to find a unique, preference-free, option price applicable to all investors.

Black and Scholes (1973) solved this problem by constructing a portfolio using the option in question and the asset that it was based on. By further assuming that arbitrage was not possible in the market, Black and Scholes (1973) were able to derive a formula for option prices that was independent of the risk preferences of the individual investor. Assuming that arbitrage is not possible implies that it is impossible to generate risk-free profits by buying and selling any combination of assets that trade in the market, with no initial net investment. This, in turn, effectively implies that all securities whose value depends on the value of another security, such as options, are consistently priced relative to those securities. The new pricing formula was a major step forward in the theory and practice of option pricing and formed the basis of all further work in the field. However, the Black-Scholes formula has various limitations associated with it. Chief of these is the fact that it assumes that the input parameters that are required to compute option prices are constant over the life of the option being priced. These parameters include the risk-free rate of return and the volatility of the price of the asset that the option is based upon. It has been observed that these variables do not remain constant over time. Furthermore, the level of volatility implied by market option prices is not constant for options with differing strike prices, but with the same maturity. This is known as the 'volatility smile' because of the shape of the curve of implied volatility versus the strike price of the options in question.

Various attempts have been made to accommodate these time-varying parameters into the option pricing procedure. The main focus has been on the

volatility parameter, since this has the largest impact on the price generated by the option pricing formula. One approach is that of the 'stochastic volatility' models such as those developed by Hull and White (1987) and Wiggins (1987). These assume a set of dynamics for the volatility of the asset that the option is based on, independently of the dynamics of the asset price. Typically, they do not admit closed-form expressions for option prices and have to be evaluated by Monte Carlo simulation. This is a disadvantage since such simulations tend to not be very efficient, which may be undesirable if a price is required on short notice. Another problem with these models is the specification of the volatility dynamics, which may have to be made on an ad-hoc basis. Another approach that allows volatility to vary over time is that of the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) option pricing models, such as those of Duan (1996) and Heston and Nandi (2000). These assume that the volatility process can be described by a GARCH model, as developed by Bollerslev (1986). Again, these models do not always admit for a closed-form option pricing formula, which could be a disadvantage. There are also problems associated with assuming a single unconditional level of volatility, as is the case in the GARCH model. This can lead to the model overestimating the persistence of volatility values, as is discussed by Lamoureux and Lastrapes (1990). This could lead to inaccurate option prices in models using the GARCH values for volatility as input.

What is required, then, is a model for option pricing that allows for the input parameters, especially volatility, to vary over time but avoids the problems associated with the models described above. It would also be desirable to allow the risk-free rate of return to vary over time, since interest rates change as economic circumstances change over time. Furthermore, it would be an advantage if the model could still admit a closed-form option pricing formula. This would allow for efficient computation of option prices and a clearly understandable pricing procedure. One such model is the Hidden Markov Option Pricing Model, as is described by Ishijima and Kihara (2005) and Elliott et al. (2005). It is this model that will be studied in this work.

## 1.1 Aims of this work

This work aims to describe and explain the Hidden Markov Option Pricing Model, as is presented by Ishijima and Kihara (2005) and Elliott et al. (2005). In order to fully explain the model, a brief digression into the theory of Hidden Markov Models (HMM's) and classical option pricing theory will be necessary. Once these have been discussed, it will be possible to study the option pricing model in detail. Apart from meeting the requirements outlined above, it turns out that this form of option pricing model is very flexible and easily extensible. It can be modified to explicitly take account of effects such as the 'leverage effect', identified by Black (1976), being a relationship between the mean rate of return on equities and the associated volatility of returns, as well as the influence of insider trading on market prices, among others. This further recommends it for

study and use. A discussion of these, and other advantages of the model, will be undertaken to motivate its use for pricing options.

Despite all the apparent advantages of this model, it remains to be seen whether the model performs well when fitted to data. One of the key factors in determining whether this is the case will be whether the model can incorporate the volatility smile effect. This would be the case if option prices produced by the model, for options with a range of different strike prices but the same maturity, followed the pattern found in market option prices that produces the implied volatility smile. Another means of testing the performance of the model would be a comparison between prices generated by the model and those produced by the Black-Scholes formula. If the prices generated by the model were not significantly different to the Black-Scholes equivalent values, then there would be little value in using the model. The Black-Scholes formula, which is well known and understood, might as well be used instead. Furthermore, if the two sets of prices are significantly different then the manner in which they differ may provide an indication as to what extra information the model prices include, that the Black-Scholes values did not, that was resulting in this difference. This could be used to determine the usefulness of the model, by comparing these extra effects that it incorporates to market conditions. Thus, another aim of this work is to test the model by fitting it to data and using approaches such as those described above to determine whether it is a good model for pricing options, or not.

A further aim of this work will be to use the Hidden Markov Option Pricing Model to develop a new approach for valuing Bermudan-style options. These options allow the holder to exercise the option at any of a discrete set of time points prior to expiry of the option. The form of the model in discrete-time, as described by Ishijima and Kihara (2005), lends itself particularly well to this application. This is due to the fact that it can implicitly take account of the discrete nature of the exercise date structure of the Bermudan options. An advantage in using this approach to value these options is that it results in a pricing formula that is close to being closed-form, whilst still allowing the input parameters of the formula to vary stochastically. Such options are found in interest rate markets and can be used to approximate the value of American-style options, which are more widely encountered. It would thus be useful to develop this new approach for use in these applications.

## 1.2 Methodology

Once the model has been described and explained, it will be fitted to data using the Baum-Welch algorithm, with scaling to avoid numerical underflow as in Rabiner (1989). It turns out that this is a form of the Expectation-Maximisation (EM) algorithm described by Dempster *et al.* (1977), which is used for fitting models to data with missing entries. This model fitting procedure will be carried out using the statistical freeware package R, version 2.1.1. R may be downloaded from <http://cran.r-project.org>, where more recent versions of the

software are available. All the code that was used to fit the model and price options can be found on the CD-ROM accompanying this work, in the "R Code" folder. Once the model has been fitted, and the parameters estimated, the results will be used to price a range of options. It will be assumed that these options are based on equity (stock) market indices. Option prices for a range of maturities, as well as a range of strike prices for each maturity, will be computed. This will determine whether the model performs differently for different maturities and whether it incorporates the volatility smile effect.

Black-Scholes prices will also be computed using parameters estimated using the model fitting procedure, but with the assumption that these then remain constant over the life of the option being priced. These values will be used as a 'base case' for comparison with the prices generated by the model. This will allow for an evaluation as to whether the extra information that the model prices include, in the form of the time-varying input parameters, results in prices that are significantly different from the Black-Scholes values. If they are not, then there would be little value in going to the trouble of computing option prices based on time-varying parameters. If the model prices are significantly different, then this may indicate that they are picking up effects that the Black-Scholes prices do not, which may result in the model prices being more accurate than the Black-Scholes equivalent values.

In order to illustrate the new scheme for pricing Bermudan-style options, this will be applied to swap rate data. Options that are based on swaps are known as 'swaptions' and are often Bermudan-style options. This is thus likely to be a useful practical setting for the use of this new scheme, which motivates its use here to test this method. As for the options on the indices mentioned above, prices will be computed for a range of maturities and strike prices. This will provide an indication as to whether the method behaves differently for these different scenarios, and whether the deviations in behaviour are as would be expected. In order to provide 'base case' values for comparison against the prices generated using the Hidden Markov approach, prices based on constant parameters will be computed using trinomial trees and finite difference methods. A lattice option pricing method, described by Bollen (1998), which takes account of hidden state effects will be also be used to compute prices for comparison here. This method is somewhat limited, however, since it only includes two possible hidden states.

### 1.3 Data

The data that will be used to fit the models, as described above, will be daily data from the period 15/03/1996 – 15/03/2006. Weekend days are not included in the dataset. Public holidays are assigned the value of the series in question for the previous day. The series in question is therefore assumed to have a zero rate of return over a public holiday. A large dataset allows for better estimation of the hidden state effects in the model, which is why such a dataset is used here. The

indices that will be used for pricing options will be the Johannesburg Securities Exchange (JSE) All Share Index, from the South African market and the Financial Times Stock Exchange (FTSE) 100 Index, from the UK market. These two different indices will be used since one is from a market in a developed economy, whilst the other is based in an emerging market. The aim will be to determine if, and how, the model performs differently in these two settings. All data is sourced from Thomson Datastream, using the "Market value" setting. These may be based on closing prices, since all other alternatives, such as opening prices, intraday highs, etc, were explicitly accounted for.

The swap data that will be used for illustrating the Bermudan-style options will run over the same period as the index data. The 5-year swap rate, being the swap rate on a swap with a maturity of 5 years, from the UK market will be used. Swap rate data for the South African market was not available, which is why UK data was used here. The maturity of the swap used was somewhat arbitrarily chosen, with no significant reasons existing for this length of swap being used. The data was again classed as "Market values" by Datastream, and the conventions for public holidays and weekends is the same as above.

All data, together with the summary statistics for that data, can be found on the CD-ROM accompanying this work, in the "Source Data" folder.

## Chapter 2. The Hidden Markov Model Framework

A Hidden Markov Model (HMM) consists of a sequence of unobservable, or hidden, states  $\{X_t\}_{t \geq 0}$ , each of which is associated with a set of parameters of a model or probability distribution, which generates the observed data  $\{Y_t\}_{t \geq 0}$  that is being modelled. Bickel *et al.* (1998) suggest that the  $X_t$  can be thought of as 'selecting' the parameters used to generate  $Y_t$  (Bickel *et al.* (1998) p1614). The unobserved state changes over time according to a set of transition probabilities that satisfy the Markov property:

$$\Pr(X_t = x_t | X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \Lambda, X_0 = x_0) = \Pr(X_t = x_t | X_{t-1} = x_{t-1}) \quad (1)$$

Thus the unobserved state sequence can be modelled by a Markov Chain. The time index can be discrete or continuous, although the former is more common. Each  $Y_t$  belongs to a given parametric family, such as that of the Normal distribution, the Poisson distribution, etc, or a given family of models. The latter case is often referred to as a 'Markov switching' model. The model family might be an autoregressive (AR) model of order 4, as studied by Hamilton (1989). In this case, the autoregressive coefficients change when the unobserved state,  $X_t$ , changes. In the case of the parametric family, the parameters of the distribution, such as the mean and variance for the Normal distribution, change when the unobserved state changes. In this way,  $X_t$  can be said to 'govern' the behaviour of  $Y_t$ , but does not generate the value that is actually observed:  $y_t$ . This is done by the given model or parametric family, which is sometimes referred to as the 'data generating process'.

The above formulation may, at first, appear complex. However, it allows for the modelling of time series that undergo structural changes, perhaps many times over the period of interest. A structural change is accommodated by a change in the value of the unobserved state, which results in a change of the parameters of the data generating process. This, in turn, then results in a change in the behaviour of the observed values of the series. Time series with structural changes often have a complicated set of dynamics, with the complication arising largely from the structural changes. However, assuming a Hidden Markov framework to deal with these structural changes provides a relatively simple solution to the problem. The simplicity arises from the assumption that the states possess the Markov property. This allows the state process to be modelled by a Markov Chain. The advantage of this model is its simplicity and tractability. The further assumption that, conditional on the unobserved state  $X_t$ , the  $Y_t$  are

independent of each other simplifies the joint density of the observed values and the unobserved states. This allows for a relatively simple estimation procedure. The advantages of this framework have led to its use in a wide variety of applications, such as speech processing, investigated by (among others) Juang and Rabiner (1991), face recognition as in Nefian and Hayes (1998) and biological sequencing such as in Krogh *et al.* (1994). A graphical representation of the Hidden Markov dependence structure is provided below.

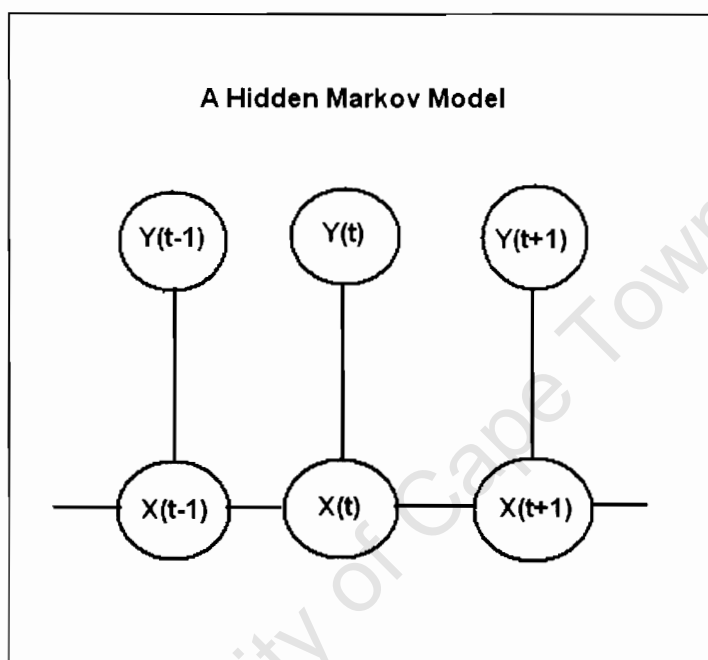


Figure 1: A graphical representation of the HMM dependence structure.

Note: This is based on a similar graphic by MacDonald and Zucchini (1997) p 67

Note that the  $Y_t$  are not connected, except through the  $X_t$ , which represents the fact that, conditional on  $X_t$ , the  $Y_t$  are independent. In the more general case of the Markov switching model, where the observed values  $Y_t$  are generated by, for example, an autoregressive process; in this case, the value  $Y_t$  will depend on all the previous values of the observed series, for each value of  $t$ .

The model is fully specified by the set of parameters:  $\{\Theta, P, \pi, B\}$   $\mathbf{P}$  is the transition probability matrix of the unobserved states:  $\mathbf{P} = \{p_{ij}\}$   $i, j \in \{1, 2, \dots, K, N\}$  where  $p_{ij} = \Pr(X_t = j | X_{t-1} = i)$ , which in the case of a continuous-time model will be the generator for the Markov Chain.

In such a case  $\mathbf{P} = \mathbf{P}(t) = \{p_{ij}(t)\}$  and  $X_t = X_0 + \int_0^t P(s)X_s ds + M_t$ , where  $\{M_t\}$  is a

martingale increment process. Notice that  $\sum_{j=1}^N p_{ij} = 1$ , for the  $p_{ij}$  to satisfy the definition of probabilities.  $\Theta$  is the vector of parameters associated with the given model, or parametric family that generates the values of  $Y_t$ .  $\pi$  is the vector containing the initial distribution of the unobserved states:

$\pi(1) = \Pr(X_0 = 1), \pi(2) = \Pr(X_0 = 2), \dots, \pi(N) = \Pr(X_0 = N)$ .  $\mathbf{B}$  is the vector of so-called 'emission probabilities':  $B(i) = f(Y_t | X_t = i; \Theta)$ , where  $f()$  is the conditional probability density of  $Y_t$ , conditional on  $\Theta$  and the unobserved state value  $X_t$ .

Using the properties of the HMM outlined above, it is now possible to find an expression for the unconditional distribution of the observed values in the series,  $Y_t$ , as well as the joint distribution of the observed data and the unobserved states:

$$\begin{aligned}
 & \Pr(Y_T = y_T, Y_{T-1} = y_{T-1}, \mathbf{K}, Y_0 = y_0, X_T = x_T, X_{T-1} = x_{T-1}, \mathbf{K}, X_0 = x_0 | \Theta) \\
 &= \Pr(Y_T = y_T, Y_{T-1} = y_{T-1}, \mathbf{K}, Y_0 = y_0 | X_T = x_T, X_{T-1} = x_{T-1}, \mathbf{K}, X_0 = x_0; \Theta) \times \\
 & \Pr(X_T = x_T, X_{T-1} = x_{T-1}, \mathbf{K}, X_0 = x_0) \tag{2} \\
 &= \Pr(X_T = x_T, X_{T-1} = x_{T-1}, \mathbf{K}, X_0 = x_0) \prod_{t=0}^T f(Y_t = y_t | X_t = x_t; \Theta) \\
 &= \Pr(X_0 = x_0) \prod_{t=1}^T \Pr(X_t = x_t | X_{t-1} = x_{t-1}) \prod_{t=0}^T f(Y_t = y_t | X_t = x_t; \Theta)
 \end{aligned}$$

Notice the simplifications possible in lines 4 and 5 above, due to the assumption that, conditional on  $X_t$ , the  $Y_t$  are independent and that  $X_t$  possesses the Markov property, respectively. To obtain the unconditional distribution of the  $Y_t$ , the uncertainty relating to the value of  $X_t$  must be 'integrated out'. Thus:

$$\begin{aligned}
 & \Pr(Y_T = y_T, Y_{T-1} = y_{T-1}, \mathbf{K}, Y_0 = y_0) = \\
 & \sum_{x_T=1}^N \sum_{x_{T-1}=1}^N \Lambda \sum_{x_0=1}^N \Pr(X_T = x_T, X_{T-1} = x_{T-1}, \mathbf{K}, X_0 = x_0) \prod_{t=0}^T f(Y_t = y_t | X_t = x_t; \Theta) \tag{3}
 \end{aligned}$$

These formulations will be used in the estimation procedures that will be implemented in later chapters.

## Chapter 3. Option Pricing and the Black-Scholes Option Pricing Formula

### 3.1 Basic Options Theory

Options are derivative instruments that grant the holder the right, but not the obligation, to enter into some transaction in the future, which will be specified in the option contract. Typically, the transaction will be the purchase or sale of a specified asset, or basket of assets, for a specified price, which is known as the strike price, at a specified time in the future. Derivative instruments are so called because their value depends on, or is derived from, other simpler variables. In the case of an option based on a particular asset, the underlying variable is then that asset. An option that grants the right to buy a specified asset is known as a call option. An option granting the right to sell the specified asset is known as a put option. Options can also be written on interest rates, financial indices, or even on prevailing weather conditions. The specification of the contract is thus completely flexible. The only limit on whether a specific type of option can be traded is whether a willing counterparty to the trade can be found.

There are two main 'styles' of options: American-style and European-style. An American-style option allows the holder to enter into the contracted transaction at any time up until some specified expiry date. A European-style option allows the transaction to take place only on the expiry date. Of course, the holder of the option will only enter into the transaction if it is profitable for them to do so. If not, then the option will merely be allowed to expire worthless. It is this feature of optionality that the holder of the option enjoys, of not having to enter into the contracted transaction if it is not profitable to do so, that gives an option its value. This is why options have a price, called a premium, which must be paid by the holder to the seller, or writer, of the option. It is this premium which all option pricing models seek to evaluate. Due to the dependence of the value of the option on the underlying variable, it follows that the method of pricing must result in a price for the option that is consistent with the value of the underlying variable.

If the holder of the option chooses to enter the contracted transaction, the option is said to have been exercised. They will do so if, for a European option, the payoff is greater than 0. For American options, prior to expiry, exercise will only occur if the payoff is greater than the value of the option, at that particular point in time. At exercise, the payoff that the holder of the option will receive will depend on the difference between the strike price and the price of the specified asset that the option is written on. For simple call options, the payoff will be the amount by which the price of the asset exceeds the strike price. This is because the holder of the option profits when they can buy the asset from the writer for a price that is below that prevailing in the market. They can exercise the option, buy the asset

for the strike price and then sell it for the (higher) market price. Alternatively, the option contract may specify that the writer will pay the holder the difference between the market price of the asset and the strike price, if this difference is greater than 0. This may be preferable, given that buying and selling the asset will incur transaction costs. If this is the case, then the option is 'settled for cash'. A similar argument is true for a put option, but for this type of option the payoff is the amount by which the strike price exceeds the market price. This is because the contracted transaction which can be entered is the sale of the asset, rather than the purchase.

Below are two graphs, which represent the payoff of an option at the time of exercise, as a function of the market price of the asset that the option is written on at that time. The call option payoff is on the left, whilst the put option payoff is on the right.  $K$  is the strike price of the option. The payoff is represented in red.

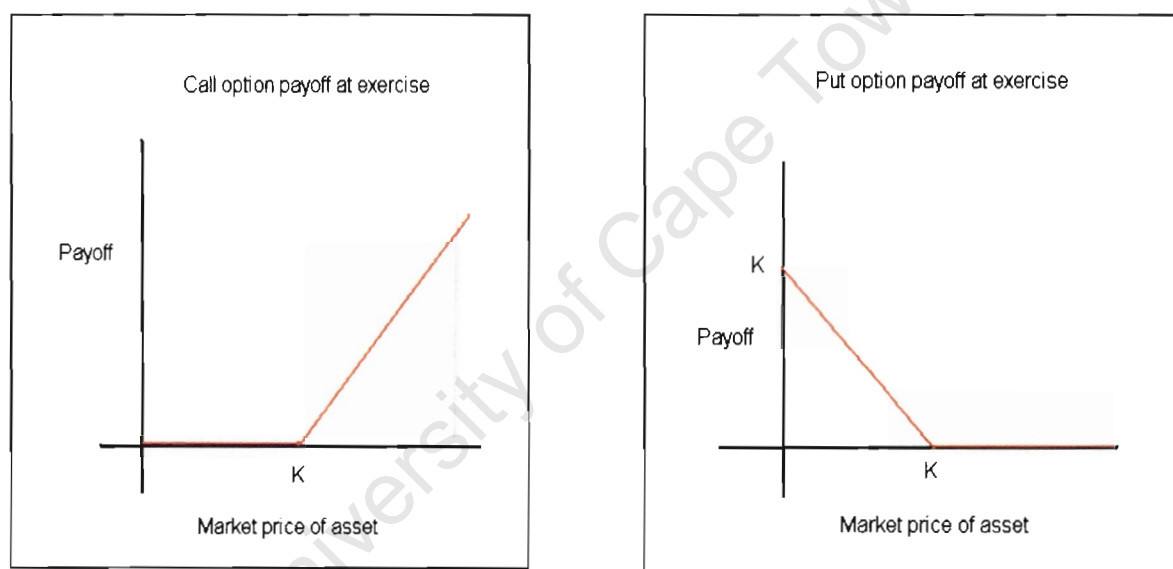


Figure 2: Payoff functions of a call and a put option at exercise

## 3.2 Pricing Options: the Black-Scholes Formula

Having outlined the basic features of options, a method of pricing them will now be introduced. This is the celebrated Black-Scholes option pricing formula, developed by Fischer Black and Myron Scholes in Black and Scholes (1973). It is not the intention of this work to develop the results from probability theory and stochastic calculus that will be used in the discussion below. The reader can find a full explanation of these in Björk (2004), Musiela and Rutkowski (1998), Karatzas and Shreve (2000) or Meyer (2001). The two equivalent approaches to

deriving the formula, the Black-Scholes partial differential equation (PDE) and risk neutral valuation, will both be examined.

The Black-Scholes model makes several assumptions about the nature of the market for assets and the options that are to be priced. These are detailed below and can be found in Black and Scholes (1973) as well as Hull (2003), *inter alia*. The market is assumed to consist of one risky asset, typically an equity security, and one risk-free asset, typically a bank account or short maturity government bond. The option is written on the risky asset.

#### Further Assumptions:

1. The options to be priced are European-style options. To further simplify the discussion, it will be assumed that the option to be priced is a call option.
2. The market for securities is frictionless: there are no taxes or transaction costs, information is freely available and all securities are perfectly divisible.
3. Short sales are permitted, without penalty. Thus, a short seller has access to the entire proceeds of a short sale, which can then be invested at the risk-free rate of return.
4. The asset that the option is written on pays no dividends over the life of the option.
5. The market trades continuously in time. There are no discrete jumps in the price of the asset.
6. The risk free rate of return is constant over the life of the option and for all maturities.
7. The instantaneous rate of return,  $\mu$ , and volatility,  $\sigma$ , of the asset are assumed to be constant over the life of the option.
8. There are no arbitrage opportunities in the market. This implies that the Law of One Price holds: any two assets, or portfolios of assets, with the same future payoffs must have the same current value.

(Black and Scholes (1973) p640, Hull (2003) p242)

The above assumptions essentially ensure that the market is 'ideal', in that there are no complicating factors to the valuation of the option such as transaction costs. The lack of arbitrage opportunities essentially ensures the consistency between the value of the underlying variable and the price of the option. Having established the nature of the market in which the formula will be developed, the dynamics of the market securities need to be specified. As a European option is being considered, the decision of whether or not to exercise the option will depend only on the value of the asset at the expiry date of the option. However, the value of the asset is risky (meaning that it changes randomly) and thus this value at the future expiry date is unknown when the option is written. In fact, the value of the asset is a stochastic process. A model for this asset price process

must thus be chosen. The Black-Scholes model assumes that the asset price can be described by a Geometric Brownian Motion, and so that the value of the asset (at a given instant in time) is lognormally distributed. The dynamics can thus be represented as below:

$$\frac{dS}{S} = \mu dt + \sigma dW_t \quad (4)$$

$$S_0 = s$$

where  $W_t$  is a standard Brownian motion.

And for the risk free security:

$$dB_t = rB_t dt \quad (5)$$

$$B_0 = 1$$

Since the instantaneous mean rate of return,  $\mu$ , is that of a risky security it will include a risk premium. This will be the amount by which  $\mu$  exceeds  $r$ . It is the compensation that an investor in  $S$  will demand for bearing the risks associated with  $S$ . If there were no risk premium, i.e.  $\mu = r$ , then the investor would prefer to invest in  $B$  and earn a return of  $r$  while facing no risks. The relevance of this factor to the current discussion is that the value that an investor places on the risk premium contained in  $\mu$  depends on their degree of risk aversion. This in turn depends on each individual investor's utility function. Thus, any option pricing formula that is based solely on the dynamics (4) and (5) above will result in prices that depend on the preferences of the individual investor. Clearly this is a problem, since these preferences will be unknown in general. This was the state of affairs prior to Black and Scholes (1973). However, it is possible to escape this difficulty.

There are two main methods of doing so; the first is presented by Black and Scholes (1973) and relies on the construction of a portfolio that is instantaneously riskless. In other words, at every instant of time ( $dt$ ), the portfolio does not contain any risk. The second method relies on results from probability theory, particularly the Girsanov theorem for changes of measure, to develop a pricing scheme in a risk neutral setting. The first argument was the first to be developed and so will be covered first here.

### 3.2.1 Pricing via the Black-Scholes PDE

Consider a portfolio consisting of one option and  $J$  units of the risky asset. By a careful choice of  $J$ , the risk contained in the portfolio can be eliminated. Note that this risk arises both from the asset and the option, since the value of the option depends on the value of the asset, which changes unpredictably and is thus

risky. Now, let  $V(t)$  be the value of the portfolio at time  $t$  and  $g(t)$  be the value of the option at time  $t$ . Then, by Itô's lemma, the dynamics of the portfolio are:

$$dV = \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial S} dS + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 g}{\partial S^2} dt + JdS \quad (6)$$

The only terms in this expression which are not deterministic, and are thus risky, are those terms containing  $dS$ . This is because  $dS$  represents the instantaneous change in the value of the risky asset. If these terms can be eliminated, then the portfolio will be instantaneously riskless. Now, it is clear that if  $J$  is chosen to be

$-\frac{\partial g}{\partial S}$  then the terms containing  $dS$  will cancel each other out in (6). However, the

value of  $\frac{\partial g}{\partial S}$  changes as the value of  $S$  changes at each instant of time. This

means that the amount  $J$  of the risky asset that is held will need to be altered continually at every instant of time for the portfolio to remain instantaneously riskless. Reducing the amount of risk in a portfolio is known as 'hedging'. Thus, for the given portfolio to remain risk free a 'dynamic hedging' scheme is required. This means that the hedging procedure must be updated over time to maintain the required level of risk reduction.

Assuming that such an instantaneous dynamic hedge is possible, then the return on the portfolio is deterministic and thus predictable and risk free. By assumption 8, this means that the rate of return that the portfolio should earn is the risk free rate of return  $r$ . It should not be possible to earn more than the risk free rate without taking on any risk, since this would lead to arbitrage opportunities. Thus:

$dV = r(g - \frac{\partial g}{\partial S} S)dt$ . Equating this to (6) leads to, after some simplification:

$$\frac{\partial g}{\partial t} + rS \frac{\partial g}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 g}{\partial S^2} - rg = 0 \quad (7)$$

This is the famous Black-Scholes PDE. All options in the Black-Scholes market satisfy this PDE. Before a solution can be found, a boundary condition must be specified. Since the option to be priced is a European option, this will be the payoff at the expiry date of the option:  $\tau$ . By the argument presented in section 3.1 it can be seen that the payoff of a call option can be represented as  $(S_\tau - K)^+$ , where  $(\cdot)^+$  is the function which returns the maximum of  $(S_\tau - K)$  and 0. The PDE is thus a final value problem, since the boundary condition is specified for the future expiry date of the option. It can be solved using standard techniques, assuming that the current time is time 0, to give:

$$g(0) = S_0 N(d_1) - Ke^{-r\tau} N(d_2)$$

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}, \quad (8)$$

$$d_2 = d_1 - \sigma\sqrt{\tau}$$

Note that  $N(\cdot)$  is the cumulative standard normal distribution function. This is the Black-Scholes option pricing formula for a European call option on an asset that pays no dividends over the life of the option.

### 3.2.2 Pricing via Risk Neutral Methods

An alternative, equivalent, method relies on so-called 'risk neutral' pricing methods. This method is a probabilistic approach to the valuation problem. The equivalence of the result derived in this manner to (8) is a consequence of the Feynman-Kac theorem. Consider again the asset price dynamics given in (4) and (5). These dynamics are specified under the 'objective', or real world, probability measure  $P$ . However, using standard results from probability theory, it is possible to change the probability measure that the asset dynamics are specified under. Specifically, it is possible to change from  $P$  to an equivalent probability measure  $Q$  that assigns probabilities in a risk neutral setting. The rationale for performing such a change is that all assets have a return equal to the risk free rate of return in a risk neutral world. This is because risk neutral investors require no risk premium for bearing risk; they have no preferences with regard to the level of risk that they bear. This means that, in the risk neutral world,  $\mu$  need no longer be considered. The instantaneous mean rate of return on  $S$  will simply be  $r$  and the problems of individual preferences relating to  $\mu$  will no longer apply.

The price of the option can then be calculated as the discounted expected payoff of the option at expiry. This is due to the fact that the equivalent measure  $Q$  is chosen so that discounted asset prices are martingales under  $Q$ . The discount rate that is used is the risk free rate of return, since  $Q$  is a risk neutral probability measure. A martingale is, essentially, a process whose expected value at a future date, conditional on the information available at the present, is its current value. Thus, under  $Q$ :  $e^{-r(T-t)} E^Q[S_T | \mathcal{F}_t] = S_t$ , where  $\mathcal{F}_t$  is the natural filtration on  $S$ . Thus  $Q$  is referred to as an 'Equivalent Martingale Measure' (EMM). The reason that the option can be priced in the risk neutral setting, and that the price that is calculated there is the same as that in the real world, is that  $Q$  is equivalent to  $P$ . This means that it assigns positive probabilities to the same events that  $P$  does, although these will not have the same value as those assigned by  $P$ , since  $Q$  exists in a risk neutral setting. Since this is 'corrected' for

by using the risk-free rate of return as the instantaneous rate of return, the price will be the same as that calculated using the 'objective' measure and the instantaneous mean rate of return including the risk premium:  $\mu$ .

In order to be able to apply this risk neutral pricing method, there are further conditions that must be satisfied by the market in which pricing is to occur. Firstly, the market must be complete. A market is complete if every contingent claim is attainable. A contingent claim is a security whose value is a random variable. In this case, the contingent claim is the option to be priced. The value of the option, at a given instant in time, is a random variable because the asset price on which it depends is a random variable. If the claim is attainable, then it is possible to construct a self-financing portfolio from the risky and the risk-free assets which has the same payoff, under all conditions, as the contingent claim at expiry. Such a portfolio is known as the replicating portfolio, since it replicates the future value of the contingent claim. Since the replicating portfolio has the same future values as the contingent claim, in this case the option, then by assumption 8 it has the same current value. The self-financing condition requires that no further funds be withdrawn from, or added to, the portfolio once it has been created. All changes in value of the portfolio are due to changes in the value of its components only. It is possible to transfer funds between the risky and risk free assets, after a change in value has occurred, however. The usefulness of such a portfolio is that it consists of assets for which known pricing methods exist, such as bank accounts and equities. Thus, if the composition of the portfolio can be determined, then so can its value. Since the Law of One Price holds, this value will be equal to the price of the contingent claim, which was being sought.

If, as is often the case, the composition of the replicating portfolio is difficult to find then the method of risk neutral pricing can be applied. Under conditions of no arbitrage, guaranteed by assumption 8, implies that there exists an EMM. This has been shown by Harrison and Pliska (1981) and Harrison and Kreps (1979). The fact that the replicating portfolio exists is a result of the Martingale Representation Theorem. This is an existence theorem only and gives no indication as to the composition of the portfolio. However, existence of the portfolio is the only condition required. The further condition of market completeness, as discussed above, implies that the EMM is unique. This means that the contingent claim can be uniquely priced by changing to the EMM  $Q$ . This change is effected by an application of the Girsanov theorem, from probability theory.

Begin with the dynamics in (4):

$$dS = \mu S dt + \sigma S dW_t$$

Then, the Girsanov transformation is:

$$dS = \mu S dt + \sigma S \left( d\hat{W}_t - \left( \frac{\mu - r}{\sigma} \right) dt \right) \quad (9)$$

$$\Rightarrow dS = rS dt + \sigma S d\hat{W}_t$$

Where  $\hat{W}_t$  is a standard Brownian motion under  $Q$ . The dynamics are now expressed in a risk neutral setting, which can be seen from the fact that the instantaneous mean rate of return has changed from  $\mu$  to  $r$ .

Now, an application of Itô's lemma gives:

$$\ln S_t \sim \phi \left( \ln S_0 + \left( r - \frac{1}{2} \sigma^2 \right) t, \sigma^2 t \right)$$

$$\Rightarrow S_t = S_0 \exp \left( \left( r - \frac{1}{2} \sigma^2 \right) t + \sigma \hat{W}_t \right) \quad (10)$$

where,  $\phi(\cdot)$  is the standard normal density function. Knowing the distributional properties of  $S$  at any time  $t$ , the price of the option written on  $S$  can now be found.

By the principle of risk neutral valuation, using the fact that  $Q$  is a martingale measure:

$$g(0) = e^{-r\tau} E^Q \left[ (S_\tau - K)^+ \mid \mathfrak{F}_t \right]$$

$$= e^{-r\tau} \int_{-x}^x (e^{\ln S_\tau} - K)^+ \frac{1}{\sqrt{2\pi\sigma^2\tau}} \exp \left( \frac{-(\ln S_\tau - (\ln S_0 + (r - \frac{1}{2}\sigma^2)\tau))^2}{2\sigma^2\tau} \right) d \ln S_\tau$$

$$= e^{-r\tau} \int_{\ln K}^x (e^u - K) \frac{1}{\sqrt{2\pi\sigma^2\tau}} \exp \left( \frac{-(u - (\ln S_0 + r - \frac{1}{2}\sigma^2)\tau)^2}{2\sigma^2\tau} \right) du$$

Where  $u = \ln S_\tau$

This simplifies, eventually, to:

$$g(0) = S_0 N(d_1) - Ke^{-r\tau} N(d_2)$$

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}$$

$$d_2 = d_1 - \sigma\sqrt{\tau}$$

The above is, again, the Black-Scholes formula for the price of a European call option on an asset that pays no dividends over the life of the option. It is exactly the same as the result in (8), as a consequence of the Feynman-Kac theorem mentioned previously.

A similar development to both the above cases, either the PDE approach or the risk neutral pricing method, can be used to determine the value of a European put on a risky asset that pays no dividends over the life of the option. The result can also be found via put-call parity.

The price of the put option is:

$$\begin{aligned} g(0) &= Ke^{-r\tau} N(-d_2) - S_0 N(-d_1) \\ &= Ke^{-r\tau} (1 - N(d_2)) - S_0 (1 - N(d_1)) \end{aligned} \tag{11}$$

with  $d_1$  and  $d_2$  as before.

### 3.2.3 Key Features of Option Pricing

Having outlined the two main methods of pricing options in the Black-Scholes framework, the key theoretical results and requirements will now be re-examined. These results and requirements will feature in later chapters and so need to be made clear here. Firstly, using a continuous dynamic hedging strategy it is possible to construct a portfolio that is instantaneously riskless. From the dynamics of this portfolio, which must be the same as those of the risk free asset in (5) under conditions of no arbitrage, a PDE can be found. This PDE is satisfied by all options that are traded in the Black-Scholes market, prior to exercise or expiry. In the case of American options, these two dates need not be the same. The PDE can then be solved once the boundary condition has been specified, in this case the payoff function of the option, to give the price of the option. Although the assumption that continuous dynamic hedging is possible is somewhat unrealistic, the PDE method is important in the pricing of American

options. It becomes a free-boundary problem, which complicates the valuation of the option.

The alternative to the PDE approach is that based on risk neutral pricing. Here, the probability measure under which the asset dynamics are specified is changed from the 'objective' measure  $P$  to the EMM  $Q$ .  $Q$  exists due to the lack of arbitrage, which is assumed under the model. The fact that a unique EMM can be found is a result of the completeness of the market. Market completeness is, essentially, the condition that all risks in the market, in this case those associated with the risky asset, can be hedged. This means that a replicating portfolio for all contingent claims can be found, with the replicating portfolio always having the same future values as the contingent claim under all conditions. If a risk factor existed that could not be hedged in the market, then the replicating portfolio (for a given contingent claim) cannot be uniquely determined. In this case there are many EMMs, which all can be used to derive an arbitrage free price for the contingent claim. The problem in the incomplete market setting is then which EMM, and thus which price, to choose. This issue will be revisited in Chapter 5.

Once the measure has been changed and the distributional properties of the risky asset found, under the new measure, then the option price can be found by evaluating the discounted expected future payoff of the option at expiry. This is the risk neutral pricing step of the valuation. The pricing formula that is derived in this way is, under the Black-Scholes assumptions, the same as that derived using the PDE approach. This follows from the Feynman-Kac theorem. Although this model represented a great breakthrough in option pricing, by eliminating individual risk preferences from the pricing formula, it has several limitations. It can only be used to price European-style options, whereas options in the marketplace often have a more complex exercise date structure. Furthermore, it assumes that the input parameters of the formula remain constant over time. This, too, is unrealistic. The Hidden Markov approach will specifically deal with this problem, providing a framework where these parameters are allowed to vary over time.

## Chapter 4. Why use Hidden Markov Models to Price Options?

The previous chapter outlined, at some length, a means for establishing the price of European options. Any further efforts aimed at finding a different method to price options thus require some justification, given that prices are readily found by applying the Black-Scholes formula developed there. The first reason for seeking an alternative is perhaps the most obvious: the assumptions of the Black-Scholes model are not satisfied in practice. This problem is identified by, among others, Tzavalis and Chourdakis (1999) and Driffill *et al.* (2002). These authors point out that the conditional means and volatilities of asset returns, equities in particular, vary over time (Tzavalis and Chourdakis (1999) p2). The Black-Scholes model assumes that these values are constant. Of the two quantities the volatility is the most important, since, as was shown in the previous chapter, the mean rate of return that is used to price options is the risk-free rate of return.

Volatility varies not only over time, but also for options with different strike prices written on the same asset. This effect, observed since the stock market crash of October 1987, is the well documented 'volatility smile'. See Campa *et al.* (1997), Buraschi and Jackwerth (2001) and Derman and Kani (1994b) *inter alia*. For options written on equities, those with lower strike prices have a higher implied volatility than options with higher strike prices. The effect is particularly pronounced for out-of-the-money put options, where the strike price of the option is well below the prevailing market price of the asset that the option is written on. In an attempt to explain this phenomenon the suggestion has been made that market participants value the 'market crash insurance' feature of these put options, which arises by the placing of a floor on the price that the holder of the option can sell their securities for. This valued 'insurance feature' leads to higher demand for these options, which then drives their price, and thus implied volatility up. Option prices generated when an HMM is used to model the asset price process can replicate this 'smile' effect, as demonstrated by Ishijima and Kihara (2005) pp. 14-17.

Extensions to the Black-Scholes model to take account of variable volatility include the stochastic volatility models such as those of Hull and White (1987) and Wiggins (1987). Such models postulate a separate set of dynamics for the volatility process, along with the asset price dynamics. Pricing is typically implemented using Monte Carlo simulation techniques, as the models do not admit a closed-form solution for option prices. There are also Generalised Autoregressive Conditional Heteroscedastic (GARCH) option pricing models, such as those of Heston (1993), Duan (1996), Heston and Nandi (1997) and Ritchken and Trevor (1999). These models use a GARCH process to account for

the variation in volatility. The formulae that they derive are then a form of extended Black-Scholes formula where the assumption of constant volatility has been relaxed. In some cases, a closed-form solution for the option price is possible in these models. Other possibilities include the so-called 'local volatility' models such as that developed by Derman and Kani (1994a), which use a lattice framework to compute market implied volatilities for options with various expiry dates and strike prices.

In the face of all these alternatives, the question of justification for yet another method, such as the HMM method, becomes even more pressing. However, there are several reasons for using such a method. Firstly, the above models only allow for variable volatility. What of a time varying risk-free rate of return? Furthermore, the risk-free rate of return (as measured by the yield on government bonds in the local currency) is very seldom the same for all maturities. In other words, the yield curve is not very often flat. This point is particularly important when the option to be priced has a 'long' time left to expiry, say three to five years. In the spirit of the stochastic volatility models another set of dynamics, this time for the risk-free rate of return, could be added to the Black-Scholes model specification. However, this would be a very ad-hoc response. Further, the short term risk-free rate of return will often be co-integrated with the benchmark rate of interest set by the central bank. The benchmark rate is adjusted in a discrete manner, when the monetary policy committee of the central bank meets. Thus a model that is based on discrete changes in its parameters, such as an HMM, could be expected to be a more reasonable model of the risk-free rate of return, especially in the short term.

In terms of volatility modelling, an HMM may be preferable to the alternatives identified above because it allows the unconditional level of volatility to vary. GARCH models are a type of Autoregressive Integrated Moving Average (ARIMA) process and thus have only one unconditional level of volatility. This is essentially an assumption under these models that the time series being modelled does not undergo any structural changes. Although GARCH models perform well in modelling the persistence observed in volatility time series, this single unconditional volatility level can be problematic. Lamoureux and Lastrapes (1990) have shown that assuming only one unconditional level of volatility in a GARCH model can cause the level of persistence in the values of the volatility series to be overestimated by the GARCH model. This is an undesirable state of affairs. An HMM can also estimate persistence in volatility, by assigning a high probability to the event of remaining in a given volatility state. By allowing the unconditional level of volatility to switch to a different level, however, the problem of overestimating volatility persistence can be reduced, as discussed by Tzavalis and Chourdakis (1999) p2. The model also admits closed-form expressions for option prices, which the stochastic volatility models do not and the GARCH option pricing models can only achieve under certain specific conditions.

Using an HMM for the asset price process can also account for the observed negative skewness and leptokurtosis of the distribution of asset price returns. This is because the conditional distribution of the returns, in an HMM at time  $t$ , is a dependent mixture of normal distributions. A mixture distribution is one that is the result of a weighted sum of component distributions, with the weights being the 'mixing probabilities'. These mixing probabilities are the probability that a given observation from the mixture distribution was from the associated component distribution, for the mixing probability in question. A mixture of normal distributions, where the component distributions have differing variances will tend to have a higher kurtosis, overall, than a single normal distribution, whilst a mixture of normal distributions where the component distributions have differing means will tend to be skew.

The HMM has an advantage over independent mixture models in that the mixing probabilities of the mixture distribution are the unconditional probabilities of the hidden Markov Chain. This induces a Markov dependence structure in the asset price returns, as is found in the diffusion models for asset returns such as (4) in the previous chapter. 'Standard' mixture models assume that the component distributions are independent of each other. Below are two sketches of mixture distributions. The overall mixture distribution is represented in blue, whilst those of the component distributions are represented in red and black, respectively. Notice how the mixture distribution can be skew, whilst neither of the component distributions are in the case on the right, whilst allowing the variance of the component distributions to be different results in a mixture distribution with a larger kurtosis, overall, than either of the components. The interested reader is referred to McLachlan and Peel (2000) for a full technical treatment of independent mixture distributions, where the features below are explained in further detail.

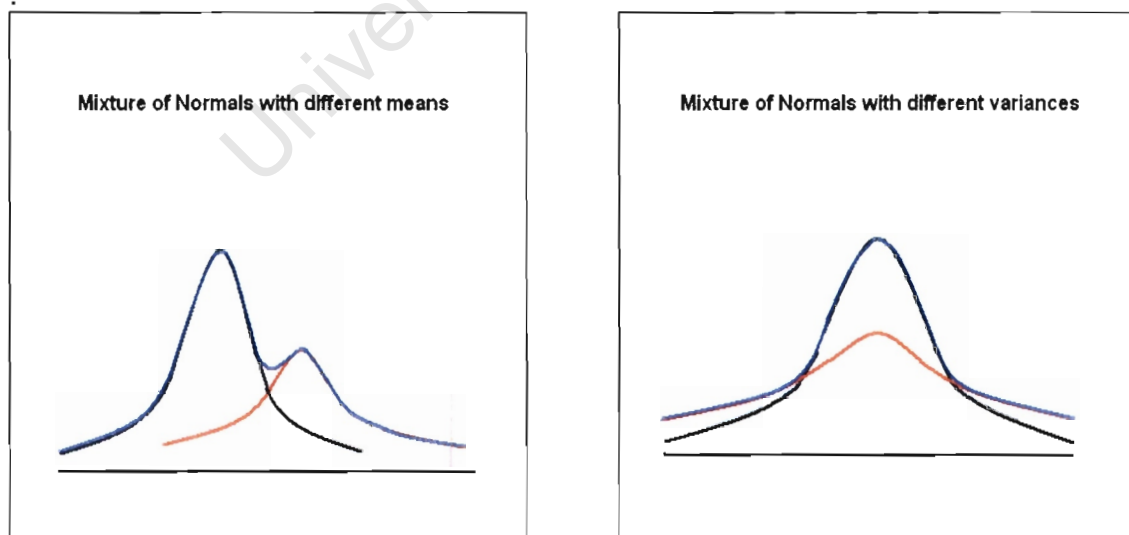


Figure 3: Mixtures of Normal Distributions

A further advantage of using an HMM framework for option pricing, identified by Duan *et al.* (2002) as well as Rossi and Gallo (2002), is that such a model can reproduce the asymmetric response of volatility to the rate of return (Duan *et al.* (2002) pp. 117-118). Low returns are associated with high levels of volatility and vice versa. This is sometimes known as the 'leverage' effect and was first studied by Black (1976). It is especially pronounced when asset returns fall. This fall is usually assumed to be due to new information that is unfavourable, which will influence the asset price, entering the market. By relating the pair of a given conditional mean and volatility of the asset price process to a given single state in the hidden Markov Chain, this association between the volatility and mean return can be included in the Hidden Markov framework.

From a philosophical point of view, using an HMM for asset prices is also appealing in that many changes in the economy are of a discrete nature. Government policy and regulations change in a discrete fashion. New rules and policies will be sufficiently different from the old that the two sets of policies can comfortably be modelled as belonging to two different 'states'. For the modelling of equity securities, corporate events that have an impact on the share price occur in a discrete fashion. Mergers and acquisitions, changes in the nature of the competition that a firm faces, changes in the composition of the management structure and team, etc, all are discrete events. News events other than those identified that affect the price of a given asset also arrive in a discrete fashion. Any changes in the nature of the asset price process are thus also likely to be discrete.

If, on the other hand, the asset being modelled is a debt instrument, such as a bond, then a strong case for changes in the nature of the price process being discrete can also be made. As previously noted, the benchmark rate of interest that the central bank sets, which has an effect on the value of debt of all maturities, changes in a discrete fashion. Further, events such as a credit rating upgrade or downgrade, a default on the obligation that the debt represents or inclusion or exclusion of the debt in a reference index are all discrete events that have an impact on the value, and thus the price, of debt instruments. Again, for such reasons, a process that models changes in the nature of the price process as being discrete in nature may be reasonable.

On a more general level, the economy that all financial instruments exist in can be thought of as transiting between various discrete states. Indeed, some of the earliest work involving HMM's in econometrics, by Hamilton (1989) and (1990), was to model the business cycle where the economy grows, falls into recession, recovers and grows again. Thus the underlying environment in which assets exist, and which must have an impact on their value at a very fundamental level, appears to change in a discrete manner. This reinforces the case for using HMM's to model asset prices.

Finally, from a model specification perspective, the HMM framework is a generally appealing one. Whilst remaining a relatively simple and parsimonious

structure, HMM's allow for a great flexibility in specification. The fact that the unobserved state process is modelled by a Markov Chain affords the overall model a good macro-structure, which admits a closed-form expression for option prices. This results in the model being a very natural extension to the Black-Scholes model, but being able to accommodate time-varying parameters. In terms of the possible specifications available for the model, the asset parameters such as mean rate of return, volatility, etc, can be specified as discrete values or processes of their own. In this case the parameters of the process, as opposed to the values that the variable can take, will change when the unobserved state changes. This allows, for example, for a 'switching GARCH' model to be specified for the volatility process, which will remove the problems of overestimating the persistence of volatility values associated with the standard GARCH model. Other possibilities for the specification of a rich structure include allowing for a feedback from the changes in the return value to the volatility value, such as that presented by Duan *et al.* (2002). This is an alternative method of accounting for leverage effects. These, along with other possible generalisations, should give an indication of the richness of structure that is possible within the relatively simple Hidden Markov framework.

Of course, no model or model structure is without its defects. A key question when using the Hidden Markov framework is that of determining the number of hidden states that the model should contain. Often, as in Hamilton (1994) or Tzavalis and Chourdakis (1999), the simplest case of only two states is chosen in the interests of simplicity and parsimony. Sometimes, as in Ishijima and Kihara (2005) and Rossi and Gallo (2002), a variety of models with differing numbers of states are specified and estimated and then compared using some criterion such as Akaike's Information Criterion. Others have proposed nonparametric Bayesian methods for estimating the number of states, either separately, as in Otranto and Gallo (2001), or jointly with the model parameters as in Robert *et al.* (2000), Chopin and Pelgrin and Casarin (2003). Yet another alternative is to try and use hypothesis testing to determine the correct number of states, as in Krolzig (1997) and Hamilton (1994). As this brief review shows, there is no agreement in the literature on how the number of states should be determined. Furthermore, none of these methods are without their own difficulties and complications.

Problems arise when the number of states chosen is large, say five or more. Not only in estimation, in the face of a large number of parameters, but also from an economic and philosophical point of view. If the model of an equity security price process has seven states, then what do these states represent economically? Do these contain different volatility levels, ranked from lowest to highest, or different levels of mean return? If so, on what basis are all seven levels required to explain the asset price returns? The problem of model overfit is clearly a real one if the recommendation of some 'empirical' method for choosing the number of states is accepted blindly. In later chapters a balance between model fit and economic rationale will be attempted when choosing the number of possible hidden states to use in the model.

The above considerations notwithstanding, the Hidden Markov approach to option pricing presents many advantages in overcoming the shortcomings of the Black-Scholes model. It allows for parameters that vary over time, while avoiding problems associated with other methods for achieving this, such as the overestimation of volatility persistence by the GARCH models. It also allows for the incorporation of the leverage and volatility smile effects. Furthermore, the unconditional distribution of asset price returns that the model implies can potentially incorporate the skewness and leptokurtosis often found in market price returns. Finally, despite including so many features and effects, the model remains relatively parsimonious and tractable. It also admits a closed-form solution for option prices, as will be seen in Chapter 6. These advantages all recommend the Hidden Markov approach for use in pricing options.

## Chapter 5. Pricing, Arbitrage and Market Completeness when Hidden Markov Models are used to Model Asset Prices

When Hidden Markov Models are used to model the asset price process of the asset that the option in question is to be written on, an unobserved state process,  $\{X_t\}_{t \geq 0}$ , is assumed to exist. There are thus fewer securities in the market than there are sources of randomness, since there is no security, other than the risky asset, whose value depends on  $X_t$ . The HMM market is thus not complete, in the sense of Harrison and Pliska (1981) and Harrison and Kreps (1979). This fact is highlighted by Elliott and Buffington (2002), Elliott *et al.* (2005) as well as Guo (2001). This means that there is no longer a unique equivalent martingale measure (EMM). This means that the pricing methods outlined in Chapter 3 can no longer be used to compute a unique price for the option that is being valued. The price of the option now depends on the EMM that is used in the valuation procedure. The existence of at least one EMM is guaranteed by the continuing assumption of a lack of arbitrage opportunities in the market, which will be carried throughout the remainder of this work. Before a Hidden Markov option pricing scheme can be developed, then, this issue must be resolved. This is because the aim is to find a unique option price, for which a unique choice of EMM is required.

### 5.1 Change of State Securities

There are several approaches to resolving this problem in the literature. Guo (2001) proposes assuming the existence of Arrow-Debreu-type securities for the process  $\{X_t\}$ . Since the framework presented by Guo (2001) is in continuous time, the securities are 'Change of State' (COS) securities which pay 1 at the stopping time  $\tau(t)$  when the unobserved state changes value. Once the state does change and payment is made, a new COS security comes into being and lasts until the next change of state. In this case, a further assumption about the distribution of arrival times of the state changes must be made. For example, Guo (2001) assumes that the state changes can be modelled by a Poisson process and that the times between state changes are then exponentially distributed. In the discrete time case, these securities would pay 1 if the state in force at time  $t$  were a given value, and 0 otherwise. Inference about the state value could then be made from the Markov Chain being used to model the state process  $\{X_t\}_{t \geq 0}$ .

These hypothetical securities then complete the market, since the 'state risk' can now be hedged. If the Poisson model of Guo (2001) is assumed, then the price of the COS is:  $V(0) = E^P \left[ \exp(-[r+k](\tau(t)-t)) \middle| \mathfrak{F}_t^X \right]$ . Here  $k$  is a risk premium factor, for the risk arising from the uncertainty surrounding the value of  $X_t$ , and  $\mathfrak{F}_t^X = \sigma(X_u; 0 \leq u \leq t)$  is the natural filtration on  $X_t$ . Since the market is now complete, a unique EMM exists and so a unique price for the option can now be found. Under the unique risk neutral measure  $Q$ , the prices of the state securities are now:

$$V(0) = E^Q \left[ \exp(-r(\tau(t)-t)) \middle| \mathfrak{F}_t^X \right] = \frac{\lambda^Q(i)}{r + \lambda^Q(i)} \quad \text{where } \lambda^Q(i) \text{ are the point process}$$

intensities under  $Q$ . The point process models the random arrival of information events. It is now possible to apply the usual risk neutral pricing theory, since the COS securities complete the market.

There is a certain air of the ad-hoc about this method of resolving the market incompleteness problem that may discourage some from using it. It does provide a method of resolving the problem that is relatively straight forward, however, and it does not require any further delving into the theory of probability measures to be understood and implemented. It simply fulfils a need thrown up by the use of an HMM for the asset price: there is no security with which to hedge the 'state risk', ergo one will be assumed into being.

## 5.2 The Local Risk Neutral Valuation Relationship

An alternative to this approach is provided by Duan (1995) and Duan (1996). This is the use of a Local Risk Neutral Valuation Relationship (LRNVR). Duan accepts that the HMM market is incomplete and that an EMM will have to be chosen from those that exist. Duan (1996) proposes to use an 'equilibrium measure'  $Q$ , defined by:  $\frac{dQ}{dP} = e^{(r-\beta)T} \frac{U'(C_T)}{U'(C_t)}$  where  $U(C_T)$  is the utility function of

consumption, evaluated at time  $T$ ,  $\beta$  is the subjective intertemporal discount rate and  $\frac{dQ}{dP}$  is the Radon-Nikodým derivative of  $Q$ , with respect to  $P$ . This

'equilibrium measure' satisfies the LRNVR if the following conditions hold:

1.  $P$  is equivalent to  $Q$  ( $P \sim Q$ )

2.  $\ln \left( \frac{S_t}{S_{t-1}} \right) \middle| \mathfrak{F}_{t-1}^{S,X}$  is normally distributed under  $Q$  with  $E^Q \left( \frac{S_t}{S_{t-1}} \middle| \mathfrak{F}_{t-1}^{S,X} \right) = e^r$

$$\text{and } Var^Q \left( \frac{S_t}{S_{t-1}} \middle| \mathfrak{F}_{t-1}^{S,X} \right) = Var^Q \left( \frac{S_t}{S_{t-1}} \middle| \mathfrak{F}_{t-1}^n \right) \text{ P-almost surely.}$$

The EMM chosen in this fashion is not entirely independent of preferences, as Duan (1996) notes, it depends on preferences indirectly through the unit risk

premium. This may be considered as undesirable, since the aim in option pricing is to find a preference-free price. However, the above change of measure is specifically formulated in a discrete model. From this perspective, it is an appealing solution to the market incompleteness problem for the case of the discrete-time HMM option pricing model that is developed in the next chapter. Indeed, it is the method that will be used to select an EMM there.

### 5.3 The Regime-switching Esscher Transform

Finally, a further alternative, proposed by Elliott *et al.* (2005), is to use the Esscher transform to change the measure to an EMM. Since the framework of interest is that of HMM's, a slightly modified version of the Esscher transform is required. Elliott *et al.* (2005) refer to this modified form as the 'regime switching' Esscher transform. The motivation for using this transformation of measure method is that the EMM it changes the measure to is the minimum entropy EMM.

Entropy is defined as  $I_p(Q) = E^P \left[ \ln \left( \frac{dQ}{dP} \right) \mathfrak{F}_t^X \right]$  and the minimum entropy EMM is

that EMM which satisfies:  $I_p(Q^*) \leq I_p(Q) \quad \forall Q \in \mathcal{S}(P)$ , where  $\mathcal{S}(P)$  is the set of all martingale measures that are equivalent to  $P$ . Minimum entropy EMM's are a tool that is used in option pricing in incomplete markets, when the incompleteness is not due to model specification, as is the case in the HMM market, such as the pricing of weather derivatives. Delbaen *et al.* (2002) have established that the choice of the measure derived from the Esscher transform will maximise terminal investor utility (at time  $\tau$  in the case of an option expiring  $\tau$  time units from now) if investors have an exponential utility function. This can be deemed a further possible justification for using this method. As will also be seen below, the appealing feature of this method is that it is very similar to the change of measure used in the previous chapter, for risk neutral valuation. It is, furthermore, underpinned by the solid theoretical basis of the Esscher transform.

The regime-switching Esscher transform can be defined in the following manner, as described by Elliott *et al.* (2005) pp.426-428. Define  $\theta_t = \theta(t, X_t)$  to be the regime switching Esscher parameter and  $Q_\theta$  as the EMM defined by the regime switching Esscher transform measure.

Then:

$$\frac{dQ_\theta}{dP} \Big|_{\mathfrak{F}_t^{Y,X}} = \frac{\exp\left(\int_0^t \theta_s dY_s\right)}{E^P\left[\exp\left(\int_0^t \theta_s dY_s\right) \Big| \mathfrak{F}_t^X\right]}$$

where

$$Q_\theta \sim P$$

$$Y_t = \int_0^t \left(\mu_s - \frac{1}{2}\sigma_s^2\right) ds + \int_0^t \sigma_s dW_s$$

and  $W_t$  is a standard Brownian motion under  $P$ .

Now, under  $P$ :  $\int_0^t \theta_s dY_s \Big| \mathfrak{F}_t^X \sim \phi\left(\int_0^t \theta_s (\mu_s - \frac{1}{2}\sigma_s^2) ds, \int_0^t \theta_s^2 \sigma_s^2 ds\right)$  so, it follows that:

$$E^P\left[\exp\left(\int_0^t \theta_s dY_s\right) \Big| \mathfrak{F}_t^X\right] = \exp\left(\int_0^t \theta_s (\mu_s - \frac{1}{2}\sigma_s^2) ds + \frac{1}{2}\int_0^t \theta_s^2 \sigma_s^2 ds\right)$$

so

$$\frac{dQ_\theta}{dP} \Big|_{\mathfrak{F}_t^X} = \exp\left(\int_0^t \theta_s \sigma_s dW_s - \frac{1}{2}\int_0^t \theta_s^2 \sigma_s^2 ds\right)$$

If  $Q_\theta$ , defined as above is to satisfy the condition for being an EMM, then the following must hold, substituting the asset price  $S_t$  for  $Y_t$ :

$$S_0 = \exp\left(-\int_0^t r_s ds\right) E^{Q_\theta}[S_t \Big| \mathfrak{F}_t^X]$$

Although this condition seems to rely on the information about  $X_t$ ,  $\mathfrak{F}_t^X$ , being available, the so-called 'tower law' for conditional expectations implies that if the result holds in the above conditional case, it also holds when  $\mathfrak{F}_t^X$  is not known, as Elliott *et al.* (2005) indicate. (Elliott *et al.* (2005) p427)

The next step of the derivation requires Bayes' Rule for Conditional Expectations, which states that:  $E^{Q_\theta}[\bullet] = \frac{E^P[\Lambda \bullet]}{E^P[\Lambda]}$  where  $\Lambda$  is the Radon-Nikodým derivative:

$\frac{dQ_\theta}{dP}$ . Using this definition, together with the martingale condition and the value of  $\frac{dQ_\theta}{dP}$  above:

$$\exp\left(-\int_0^t r_s ds\right) E^{Q_\theta} [S_t | \mathfrak{F}_t^X] = \frac{E^P \left[ \exp\left(-\int_0^t r_s ds\right) S_t \exp\left(\int_0^t \theta_s \sigma_s dW_s - \frac{1}{2} \int_0^t \theta_s^2 \sigma_s^2 ds\right) \middle| \mathfrak{F}_t^X \right]}{E^P \left[ \exp\left(\int_0^t \theta_s \sigma_s dW_s - \frac{1}{2} \int_0^t \theta_s^2 \sigma_s^2 ds\right) \right]}$$

and

$$S_t = S_0 \exp\left(\int_0^t \left(\mu_s - \frac{1}{2} \sigma_s^2\right) ds + \int_0^t \sigma_s dW_s\right)$$

so

$$\exp\left(-\int_0^t r_s ds\right) E^{Q_\theta} [S_t | \mathfrak{F}_t^X] = S_0 E^P \left[ \exp\left(\int_0^t \left(\mu_s - \frac{1}{2} \sigma_s^2\right) - r_s - \frac{1}{2} \theta_s^2 \sigma_s^2 ds + \int_0^t (\theta_s + 1) \sigma_s dW_s\right) \middle| \mathfrak{F}_t^X \right]$$

noting

$$\begin{aligned} \int_0^t \theta_s \sigma_s dW_s &\sim \phi\left(0, \int_0^t \theta_s^2 \sigma_s^2 ds\right) \Rightarrow E^P \left[ \exp\left(\int_0^t \theta_s \sigma_s dW_s - \frac{1}{2} \int_0^t \theta_s^2 \sigma_s^2 ds\right) \right] \\ &= E^P \left[ \exp\left(\frac{1}{2} \int_0^t \theta_s^2 \sigma_s^2 ds - \frac{1}{2} \int_0^t \theta_s^2 \sigma_s^2 ds\right) \right] = E^P [\exp(0)] = 1 \end{aligned}$$

Now

$$\begin{aligned} &S_0 E^P \left[ \exp\left(\int_0^t \left(\mu_s - \frac{1}{2} \sigma_s^2\right) - r_s - \frac{1}{2} \theta_s^2 \sigma_s^2 ds + \int_0^t (\theta_s + 1) \sigma_s dW_s\right) \middle| \mathfrak{F}_t^X \right] \\ &= S_0 \exp\left(\int_0^t \left(\mu_s - \frac{1}{2} \sigma_s^2\right) - r_s - \frac{1}{2} \theta_s^2 \sigma_s^2 ds + \frac{1}{2} \int_0^t (\theta_s^2 + 2\theta_s + 1) \sigma_s^2 ds\right) \end{aligned}$$

Now define  $\tilde{\theta}_s$  to be the Esscher transform parameter that ensures that the measure  $Q_\theta$  is an EMM: i.e.  $\tilde{\theta}_s$  is chosen so that the martingale condition is satisfied by S.

Thus:

$$\begin{aligned}
 S_0 \exp\left(\int_0^t (\mu_s - \frac{1}{2}\sigma_s^2) - r_s - \frac{1}{2}\tilde{\theta}_s^2 \sigma_s^2 ds + \frac{1}{2}\int_0^t (\tilde{\theta}_s^2 + 2\tilde{\theta}_s + 1)\sigma_s^2 ds\right) &= S_0 \\
 \Rightarrow \exp\left(\int_0^t (\mu_s - \frac{1}{2}\sigma_s^2) - r_s + \left(\frac{2\tilde{\theta}_s + 1}{2}\right)\sigma_s^2 ds\right) &= 1 \\
 \Rightarrow \int_0^t (\mu_s - \frac{1}{2}\sigma_s^2) - r_s + \left(\frac{2\tilde{\theta}_s + 1}{2}\right)\sigma_s^2 ds &= 0 \\
 \Rightarrow (\mu_t - \frac{1}{2}\sigma_t^2) - r_t + \left(\frac{2\tilde{\theta}_t + 1}{2}\right)\sigma_t^2 &= 0 \\
 \Rightarrow \left(\frac{2\tilde{\theta}_t + 1}{2}\right)\sigma_t^2 &= \frac{1}{2}\sigma_t^2 + r_t - \mu_t \\
 \Rightarrow \left(\tilde{\theta}_t + \frac{1}{2}\right)\sigma_t^2 &= \frac{1}{2}\sigma_t^2 + r_t - \mu_t \\
 \Rightarrow \tilde{\theta}_t &= \frac{r_t - \mu_t}{\sigma_t^2} \\
 &= -\frac{\lambda_t}{\sigma_t}
 \end{aligned}$$

where  $\lambda_t$  is the market price of risk, being the return per unit of risk demanded by investors for taking on that risk, at time  $t$ .

Finally then, the risk neutral regime switching Esscher EMM can be defined by:

$$\frac{dQ_{\tilde{\theta}_s}}{dP} = \exp\left(\int_0^t \left(\frac{r_s - \mu_s}{\sigma_s}\right) dW_s - \frac{1}{2}\int_0^t \left(\frac{r_s - \mu_s}{\sigma_s}\right)^2 ds\right)$$

It follows that if the dynamics of

$S_t$  under  $P$  are  $dS_t = \mu_s S_t ds + \sigma_s S_t dW_s$ , then changing the measure to  $Q_{\tilde{\theta}_s}$  changes the dynamics to:

$$\begin{aligned}
 dS_t &= \mu_s S_t ds + \sigma_s S_t \left( \hat{W}_t - \int_0^t \left(\frac{r_s - \mu_s}{\sigma_s}\right) ds \right) \\
 &= r_s S_t ds + \sigma_s S_t d\hat{W}_t
 \end{aligned}$$

The last line above now reflects the dynamics that would be expected under a risk neutral measure. Once this change of measure has been effected, then option pricing can proceed using the methods outlined in Chapter 3.

## Chapter 6. The Hidden Markov Option Pricing Model

Once a decision has been made on the approach to the problem of market incompleteness that is to be taken, the explicit form of the option pricing model that is to be used can be specified. There are two major alternatives available: a discrete-time version of the model and a continuous-time version. Whilst the continuous-time model may be more theoretically appealing, due to the fact that the asset price process is modelled by a diffusion-type process (as in (4)), the discrete-time model is far easier to estimate. Furthermore, the data that will be used to estimate the parameters of the model is of a discrete nature: daily, weekly monthly, etc. Even so-called 'high frequency' data that is recorded every few minutes, or even seconds, is still discrete in nature. It thus follows that an estimation method for a discrete-time model will be much easier to implement, given that the data that the model parameters are to be estimated from is in a discrete-time form. It also seems more reasonable, given the nature of the data.

Furthermore, it will become apparent below that a key value of interest, to be determined, is the probability that amount of time that the unobserved state process  $\{X_t\}_{t \geq 0}$  spends in each state, over a given time interval, was a certain value. Now, since these amounts are discrete values and the probabilities are to be estimated from the data, a case can be made that the continuous-time model is overly complicated for option pricing. If the amount of time spent in each state is a continuous variable, then it turns out that finding the probability density function of this variable is not an easy exercise. Further, once this density has been determined, it is a complicated function that is not easy to manipulate. Given that the probabilities that are to be determined are for discrete events, this appears somewhat unnecessary. However, for the sake of completeness, both versions of the model are presented here.

### 6.1 The Discrete-time Version of the Model

The model that is discussed and presented here is, with minor modifications, essentially that studied by Ishijima and Kihara (2005). What follows is thus very similar to Ishijima and Kihara (2005) pp. 3-8. Given the HMM framework described in Chapter 2, the following model of the asset price can be specified.

As in the case of the Black-Scholes model, it will be assumed that the option to be priced is a European call option with strike price  $K$ . Assume that the overall economy can occupy any one of  $N$  unobserved states at time  $t$ . The state that the economy is in will determine the value of the risk-free rate, as well as the volatility and mean rate of return of the asset that the option to be priced is written on. Let this asset be denoted  $S$ . Thus, the value of  $S$  at time  $t$  is  $S_t$ . The model is thus parameterised according to the scheme below.

Let  $\{X_t\}_{t=1}^T$  be the unobserved state process such that  $X_t \in \{e_1, e_2, \dots, e_N\}$  where the  $e_i$  are unit vectors in  $\mathbb{R}^N$ , with unity in the  $i$ -th position and 0's elsewhere:  $e_i = (0, 0, \dots, 0, 1, 0, \dots, 0)'$  where ' denotes the transpose operator. Thus, if  $x_t = e_i$  then the state in force at time  $t$  is the  $i$ -th state. Let  $\mathbf{P}$  be the homogenous transition probability matrix of the Markov Chain that is used to model  $X_t$ . Let  $\pi$  be the probability distribution of the initial state value,  $X_0$ , as in Chapter 2.  $\mathbf{r}$  is the vector of possible risk free interest rates,  $\sigma$  is the vector of possible volatilities and  $\mu$  is the vector of possible mean rates of return. Thus:  $\mathbf{r} = (r_1, r_2, \dots, r_N)'$ ,  $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)'$  and  $\mu = (\mu_1, \mu_2, \dots, \mu_N)'$ . Under this formulation, the risk-free rate in force at time  $t$ ,  $r(X_t) = \langle \mathbf{r}, X_t \rangle$ , where the operator  $\langle \cdot, \cdot \rangle$  denotes the inner product. This is due to the fact that  $X_t$  is a unit vector. Similarly,  $\sigma(X_t) = \langle \sigma, X_t \rangle$  and  $\mu(X_t) = \langle \mu, X_t \rangle$ .

Now denote the filtration on  $X_t$  as  $\mathfrak{F}_t^X := \sigma(X_0, X_1, \dots, X_t)$ , i.e. the smallest sigma-algebra generated by the values of  $X_t$  up to and including time  $t$ . Define the filtration on the observed values,  $S_t$  in a similar manner:  $\mathfrak{F}_t^S := \sigma(S_0, S_1, \dots, S_t)$ . The joint filtration on both the observed and unobserved values can then be defined as:  $\mathfrak{F}_t^{X,S} := \{\mathfrak{F}_t^X, \mathfrak{F}_t^S\}$ . Define the dynamics of the model, under the 'objective' or real world measure  $P$  as follows:

$$X_t = \mathbf{P} X_{t-1} + M_t \quad (12)$$

where  $\{M_t\}$  is a  $\mathfrak{F}_t^X$ -martingale increment process

$$B_T = \exp\left(\sum_{t=0}^T r(X_t)\right) \quad (13)$$

$$B_0 = 1$$

$$\ln\left(\frac{S_t}{S_{t-1}}\right) X_t = \mu(X_t) - \frac{1}{2} \sigma^2(X_t) + \sigma(X_t) \varepsilon_t \quad (14)$$

$$S_0 = s$$

where  $\varepsilon_t \sim \phi(0,1)$  under  $P$ .

Equation (14) above is really more of an observation equation, conditional on the value of  $X_t$ . In this formulation of the model, the asset price returns do not have any dynamics. In a Markov switching model, on the other hand, they could well follow some dynamic model such as an autoregressive model. The simpler case, without such dynamics is considered here.

As in Chapter 3, the above equations involve a term that includes a risk premium:  $\mu(X_t)$ , thus including the risk preferences of the individual investor. Again, this term needs to be eliminated to allow for preference-free pricing. However, unlike in Chapter 3, the market is no longer complete. As discussed in the previous chapter, the uncertainty relating to  $X_t$  is not priced in the market, which leads to market incompleteness, as was discussed in the previous chapter. In this case, the Locally Risk-Neutral Valuation Relationship of Duan (1995) and Duan (1996) will be used to resolve this problem. Under this relationship, it follows that (14) becomes:

$$\ln\left(\frac{S_t}{S_{t-1}}\right) \Big| X_t = r_t - \frac{1}{2}\sigma^2(X_t) + \sigma(X_t)\varepsilon_t^* \quad (15)$$

where  $\varepsilon_t^* \sim \phi(0,1)$  under the equilibrium measure  $Q$ .

Now, under risk neutral conditions, a Black-Scholes-type price can be developed. From (15) above, it follows that:

$$\ln S_t \sim \phi\left(\ln S_0 + \sum_{t=0}^{\tau} r(X_t) - \frac{1}{2} \sum_{t=0}^{\tau} \sigma^2(X_t), \sum_{t=0}^{\tau} \sigma^2(X_t)\right)$$

Thus, using a similar method to the risk neutral pricing scheme outlined in Chapter 3, and assuming that  $r_t$  and  $\sigma_t$  were deterministic (and thus known at  $t$ ), the price of the call option is:

$$\begin{aligned} g(0, S_0, \mathfrak{F}_t^X) &= E^Q \left[ \exp\left(-\sum_{t=0}^{\tau} r(X_t)\right) (S_{\tau} - K)^+ \right] \\ &= S_0 N(d_1) - Ke^{-\sum_{t=0}^{\tau} r(X_t)} N(d_2) \end{aligned} \quad (16)$$

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \sum_{t=0}^{\tau} \left(r(X_t) + \frac{1}{2}\sigma^2(X_t)\right)}{\sqrt{\sum_{t=0}^{\tau} \sigma^2(X_t)}}$$

$$d_2 = d_1 - \sqrt{\sum_{t=0}^{\tau} \sigma^2(X_t)}$$

Again, as in Chapter 3, the time to expiry of the option is  $\tau$  and the current time is 0. As can be seen, this is a slightly modified form of the Black-Scholes formula that was the result in (8). This is partly because the standard Black-Scholes model is a continuous-time model, with the asset dynamics represented as

diffusions, whilst the above model exists in discrete time. It is also a result of the fact that the Black-Scholes model assumes that all the inputs are constant, whereas here, we have allowed them to vary over time. Of course, the above model is not correct, since it was assumed that the state variable,  $X_t$ , was unobservable. This means that the value of  $X_t$  is unknown, and thus  $r(X_t)$  and  $\sigma(X_t)$  are also unknown.

This apparent impasse can be overcome by defining a new variable,  $O_\tau^i$  which represents the amount of time that  $X_t$  spends in state  $i$  from time 0 to time  $\tau$ . This is known as the 'occupation time' of  $X$  in state  $i$  up to time  $\tau$ . It follows that  $O_\tau^i = \sum_{t=0}^{\tau} \langle X_t, e_i \rangle$ , since  $\langle X_t, e_i \rangle$  acts as a sort of indicator function, as it is equal to 1 if  $X_t = e_i$  and 0 otherwise. Now, by the properties of inner products:

$$\sum_{t=0}^{\tau} r(X_t) = \sum_{t=0}^{\tau} \langle r, X_t \rangle = \sum_{i=1}^N r(i) O_\tau^i \quad (17)$$

Similarly:

$$\sum_{t=0}^{\tau} \sigma(X_t) = \sum_{i=1}^N \sigma(i) O_\tau^i \quad (18)$$

By noting that  $\sum_{i=1}^N O_\tau^i = \tau$ , and that thus once  $(N-1)$  of the  $O_\tau^i$  have been determined, the  $N$ -th is fixed, it is possible to rewrite (17) and (18) above as follows:

$$\begin{aligned} \sum_{t=0}^{\tau} r(X_t) &= \sum_{i=1}^{N-1} r(i) O_\tau^i + r_N \left( \tau - \sum_{i=1}^{N-1} O_\tau^i \right) \\ \sum_{t=0}^{\tau} \sigma(X_t) &= \sum_{i=1}^{N-1} \sigma(i) O_\tau^i + \sigma_N \left( \tau - \sum_{i=1}^{N-1} O_\tau^i \right) \end{aligned}$$

Having determined these quantities, it is now possible to reformulate (16) to take account of the fact that  $X_t$  is unobservable. This will require probabilistic inferences about the value of each  $O_\tau^i$  to be made. The final formula can then be derived by taking an expectation over the random variables  $\{O_\tau^i\}_{i=(1, \dots, N-1)}$ :

$$g(0, \tau, S_0) = \sum_{\zeta_1=0}^r \Lambda \sum_{\zeta_{N-1}=0}^r \Pr(O_\tau^1 = \zeta_1, K, O_\tau^{N-1} = \zeta_{N-1}) \left[ S_0 N(d_1) - K \exp\left(-\sum_{i=1}^{N-1} r_i \zeta_i - r_N \left(\tau - \sum_{i=1}^{N-1} \zeta_i\right)\right) N(d_2) \right]$$

$$= \sum_{j=1}^N \pi(j) g(0, \tau, S_0, X_0 = e_j)$$

(19)

where  $\pi$  has been defined and  $g(0, \tau, S_0, X_0 = e_j)$  is the call price, conditional on the value of the initial state  $X_0$ . This formulation arises due to the fact that a Markov Chain is only fully specified by its transition probability matrix as well as the distribution of the value of its initial state. Since  $\{X_t\}$  is being modelled by a Markov Chain, the result follows.

Further:

$$g(0, \tau, S_0, X_0 = e_j) := \sum_{\zeta_1=0}^r \Lambda \sum_{\zeta_{N-1}=0}^r \Pr(O_\tau^1 = \zeta_1, K, O_\tau^{N-1} = \zeta_{N-1} | X_0 = e_j) \times$$

$$\left[ S_0 N(d_1) - K \exp\left(-\sum_{i=1}^{N-1} r_i \zeta_i - r_N \left(\tau - \sum_{i=1}^{N-1} \zeta_i\right)\right) N(d_2) \right]$$

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \sum_{i=1}^{N-1} \left(r_i + \frac{1}{2} \sigma_i^2\right) \zeta_i + \left(r_N + \frac{1}{2} \sigma_N^2\right) \left(\tau - \sum_{i=1}^{N-1} \zeta_i\right)}{\sqrt{\sum_{i=1}^{N-1} \sigma_i^2 \zeta_i + \sigma_N^2 \left(\tau - \sum_{i=1}^{N-1} \zeta_i\right)}} \quad (20)$$

$$d_2 = d_1 - \sqrt{\sum_{i=1}^{N-1} \sigma_i^2 \zeta_i + \sigma_N^2 \left(\tau - \sum_{i=1}^{N-1} \zeta_i\right)}$$

It can be seen that if  $N = 1$ , then this formula reduces to the Black-Scholes formula as in (8), albeit in a discrete-time setting as opposed to the continuous-time version found there. Notice that this is a weighted average of call prices, one for each state evolution path, given the initial state value. The weights are then the probabilities that the actual evolution path was that used in determining the given price. Summing over all the possible paths that the state process may take then 'averages out' the uncertainty relating to the hidden state process. This pricing formula is thus a very natural extension to the Black-Scholes formula, taking into account time-varying parameters.

The last term which now remains to be determined is the value:  $\Pr(O_\tau^1 = \zeta_1, O_\tau^2 = \zeta_2, K, O_\tau^{N-1} = \zeta_{N-1} | X_0 = e_j)$ . The Markov dependence structure of  $\{X_t\}$  can be used to calculate this value in a relatively simple manner.

Firstly:

$$\Pr(O_\tau^1 = \varsigma_1, \mathbf{K}, O_\tau^{N-1} = \varsigma_{N-1} | X_0 = e_j) = \Pr(O_\tau^1 = \varsigma_1 | X_0 = e_j) \cdot \Pr(O_\tau^2 = \varsigma_2 | O_\tau^1 = \varsigma_1, X_0 = e_j) \mathbf{K} \\ \mathbf{K} \Pr(O_\tau^{N-1} = \varsigma_{N-1} | O_\tau^1 = \varsigma_1, O_\tau^2 = \varsigma_2, \mathbf{K}, O_\tau^{N-2}, X_0 = e_j)$$

using the definition of conditional probability. The above conditional probabilities can then be rewritten in a recursive form, as below. This alternative expression will allow for the development of a recursive algorithm to calculate

$$\Pr(O_\tau^1 = \varsigma_1, O_\tau^2 = \varsigma_2, \mathbf{K}, O_\tau^{N-1} = \varsigma_{N-1} | X_0 = e_j).$$

Thus:

$$\Pr(O_\tau^2 = \varsigma_2 | X_0 = e_j) = \Pr(O_{\tau-\varsigma_1}^2 = \varsigma_2 | X_0 = e_j) \quad \text{if } 0 \leq \varsigma_2 \leq \tau - \varsigma_1 \\ = 0 \quad \text{otherwise}$$

$$\Pr(O_\tau^{N-1} = \varsigma_{N-1} | O_\tau^1 = \varsigma_1, O_\tau^2 = \varsigma_2, \mathbf{K}, O_\tau^{N-1} = \varsigma_{N-1}, X_0 = e_j) \\ = \Pr(O_{\tau - \sum_{i=1}^{N-2} \varsigma_i}^{N-1} = \varsigma_{N-1} | X_0 = e_j) \quad \text{if } 0 \leq \varsigma_{N-1} \leq \tau - \sum_{i=1}^{N-1} \varsigma_i \\ = 0 \quad \text{otherwise}$$

Before  $\Pr(O_\tau^1 = \varsigma_1, O_\tau^2 = \varsigma_2, \mathbf{K}, O_\tau^{N-1} = \varsigma_{N-1} | X_0 = e_j)$  can be computed, using the above, two further definitions are required. Firstly, let  $\mathbf{P}_u = (p_{ij}(u))_{1 \leq i, j \leq N} = (\Pr(X_{t+u} = e_j | X_t = e_i))_{1 \leq i, j \leq N}$  ( $u = 1, \mathbf{K}, \tau$ ) i.e.  $\mathbf{P}_u$  is the transition probability matrix for transitions of more than one step. Using the Chapman-Kolmogorov equation for time-homogenous Markov Chains, it follows that:

$$\mathbf{P}_u = \mathbf{P}^u \quad (21)$$

( $u = 1, \dots, t$ ).

Secondly, let  $f_{ji}(t)$  be the 'first-passage' probability. This is the probability that the first time that  $X_t$  switches from state  $i$  to state  $j$  after time 0 is time  $t$ :

$$f_{ij}(t) = \Pr(X_t = e_j, X_{t-1} \neq e_j, \mathbf{K} X_1 \neq e_j | X_0 = e_i) \quad t = (2, \mathbf{K}, \tau)$$

Using (21), it is possible to set up a recursive scheme to evaluate the above probabilities.

Hence:

$$f_{ij}(t) = p_{ij}(t) - \sum_{u=1}^{t-1} p_{jj}(t-u)f_{ij}(u) \quad (22)$$

$$f_{ij}(1) = \Pr(X_1 = e_j | X_0 = e_i) = p_{ij}$$

$$t = (2, K, \tau)$$

Using these definitions, it is possible to compute the occupation time probabilities as follows:

$$\begin{aligned} \Pr(O_t^i = 1 | X_0 = e_j) &= \Pr(X_1 = e_i | X_0 = e_j) = p_{ji} \\ \Pr(O_t^i = 0 | X_0 = e_j) &= \Pr(X_t \neq e_i, K, X_1 \neq e_i | X_0 = e_j) \\ &= 1 - \sum_{u=1}^t f_{ji}(u) \end{aligned} \quad (23)$$

$$t = (1, K, \tau)$$

$$\begin{aligned} \Pr(O_t^i = 1 | X_0 = e_j) &= \Pr(\{X_u = e_i\} \mathbf{I} \{X_v \neq e_i\} | X_0 = e_j) \\ &= \Pr\left(\bigcap_{u=1}^{t-1} \{X_t \neq e_i, K, X_{u+1} \neq e_i, X_u = e_i, X_{u-1} \neq e_i, K, X_1 \neq e_i\} | X_0 = e_j\right) + \\ &\Pr(X_t = e_i, X_{t-1} \neq e_i, K, X_1 \neq e_i | X_0 = e_j) \\ &= \Pr\left(\left(\bigcap_{u=1}^{t-1} \{X_u = e_i, X_{u-1} \neq e_i, K, X_1 \neq e_i\}\right) \mathbf{I} \left(\bigcap_{v=u+1}^{t-1} \{X_t \neq e_i, K, X_{v+1} \neq e_i, X_v \neq e_i\}\right) | X_0 = e_j\right) + \\ &\Pr(X_t = e_i, X_{t-1} \neq e_i, K, X_1 \neq e_i | X_0 = e_j) \\ &= \Pr\left(\left(\bigcap_{u=1}^{t-1} \{X_u = e_i, X_{u-1} \neq e_i, K, X_1 \neq e_i\}\right) \mathbf{I} \left(\bigcap_{v=u+1}^{t-1} \{O_{v \rightarrow (t-1)}^i = 0\}\right) | X_0 = e_j\right) + \\ &\Pr(X_t = e_i, X_{t-1} \neq e_i, K, X_1 \neq e_i | X_0 = e_j) \\ &= \sum_{u=1}^{t-1} \Pr(X_u = e_i, X_{u-1} \neq e_i, K, X_1 \neq e_i | X_0 = e_j) \times \Pr\left(\sum_{s=u+1}^t \langle X_s, e_i \rangle = 0 | X_0 = e_j\right) \\ &+ \Pr(X_t = e_i, X_{t-1} \neq e_i, K, X_1 \neq e_i | X_0 = e_j) \\ &= \sum_{u=1}^{t-1} f_{ji}(u) \Pr(O_{t-u}^i = 0 | X_0 = e_j) + f_{ji}(t) \end{aligned}$$

$$(24)$$

where  $t = (2, K, \tau)$  and  $O_{v \rightarrow (t-1)}^i$  is the occupation time in state  $i$  between times  $v$  and  $(t-1)$

And finally:

$$\begin{aligned}
& \Pr(O_t^i = \zeta_i | X_0) \\
&= \Pr\left(\left(\prod_{u=1}^{t-\zeta_i+1} \{O_{(t-\zeta_i+1)}^i = 1\}\right) \mathbf{I}\left(\prod_{u+1}^t \{O_{(u+1) \rightarrow t}^i = \zeta_i - 1\}\right) \middle| X_0 = e_j\right) \\
&= \Pr\left(\left(\prod_{u=1}^{t-\zeta_i+1} \{O_{(t-\zeta_i+1)}^i = 1\}\right) \middle| X_0 = e_j\right) \times \Pr\left(\prod_{u+1}^t \{O_{(u+1) \rightarrow t}^i = \zeta_i - 1\} \middle| X_0 = e_j\right) \\
&= \sum_{u=1}^{t-\zeta_i+1} \Pr(X_u = e_i, X_{u-1} \neq e_i, K, X_1 \neq e_i | X_0 = e_j) \times \Pr\left(\prod_{u+1}^t \{O_{(u+1) \rightarrow t}^i = \zeta_i - 1\} \middle| X_0 = e_j\right) \\
&= \sum_{u=1}^{t-\zeta_i+1} \Pr(X_u = e_i, X_{u-1} \neq e_i, K, X_1 \neq e_i | X_0 = e_j) \times \Pr\left(\sum_{v=u+1}^t \langle X_v, e_i \rangle = \zeta_i - 1 \middle| X_0 = e_j\right) \\
&= \sum_{u=1}^{t-\zeta_i+1} f_{ji}(u) \Pr(O_{t-u}^i = \zeta_i - 1 | X_0 = e_j)
\end{aligned}$$

$$(t = 2, K, \tau; \zeta_i = 2, K, t)$$

(25)

These results, although appearing unwieldy and complicated, are in effect relatively simple. They follow from careful application of the definition of the various events, as well as use of the values and relations defined in (21) and (22). A further crucial assumption is that of the time homogeneity of the transition probability matrix of the Markov Chain. This allows the definition of (21) to be used. This is especially important when calculating the  $f_{ij}(t)$ . A further advantage is that these formulae are linear, which means that although the computations involving them might be lengthy and tedious, they will remain reasonably simple. This is a result of the simple structure of the Markov Chain that is being used to model the dynamics of  $\{X_t\}$ . Armed with these occupation time probabilities, the price of a call option can now be evaluated using (19) and (20).

## 6.2 The Continuous-time Version of the Model

In the continuous-time version of the model, many of the quantities defined in the previous section remain the same, albeit in a continuous, rather than a discrete, setting. The summation terms of the previous section become integrals and the observed process now follows a set of dynamics, as opposed to an observation equation as previously. Furthermore, the occupation times that were necessary to evaluate the option price are also required here, although they now have a

probability density function (*pdf*) since time is a continuous variable. This *pdf* needs to be derived and, as will become clear below, it turns out to be a complex function. This leads to the argument that this form of the model is not especially useful in a practical setting, as it overcomplicates the evaluation of the occupation times and occupation time probabilities.

However, leaving such considerations to one side, the continuous time model can be specified as follows. This specification is essentially that of Elliott and Buffington (2002) pp. 498-500 and so, as in the previous section, the description that follows will closely resemble the source material.

As before,  $\{X_t\}_{t=1}^T$  is the unobserved state variable, with state space  $\{e_1, e_2, \dots, e_N\}$ .  $\mathbf{P}(t)$  is now the generator for  $\{X_t\}$ , with  $\mathbf{P}(t)$  defined as in Chapter 2.

$\mu_t = \langle \mu, X_t \rangle$ ,  $\sigma_t = \langle \sigma, X_t \rangle$  and  $r_t = \langle r, X_t \rangle$ , in a similar manner to that previously, bearing in mind that  $t$  is now a continuous variable and so that these definitions are for every instant of time and not time points, as before.  $\mathfrak{F}_t^X, \mathfrak{F}_t^S$  and  $\mathfrak{F}_t^{X,S}$  are all defined as before. Now define the dynamics of the model under  $P$  as:

$$dX_t = PX_t dt + dM_t \quad (26)$$

where  $\{M_t\}$  is a  $\mathfrak{F}_t^X$ -martingale increment process

$$\begin{aligned} dS_t &= \mu_t S_t dt + \sigma_t S_t dW_t \\ S_0 &= s \end{aligned} \quad (27)$$

where  $W_t$  is a standard Brownian motion under  $P$

$$\begin{aligned} dB_t &= r_t B_t dt \\ B_0 &= 1 \end{aligned} \quad (28)$$

The above dynamics contain terms that are preference dependent, specifically  $\mu_t$ , and are thus not suitable for the derivation of an option pricing formula. A change of measure is required, to an EMM, so that the method of risk neutral pricing can be applied. However, as was noted in the previous chapter, the process  $\{X_t\}_{t=1}^T$  introduces uncertainty into the market that is not priced. The market is thus incomplete, in the sense of Harrison and Pliska (1981). There are thus many EMM's and one must be selected for pricing. Elliot and Buffington (2002) avoid this issue by assuming that the dynamics are specified in a risk neutral setting, without considering how that setting was attained. It is proposed that the method of the Esscher transform proposed by Elliott *et al.* (2005) be the means by which a change of measure is effected.

Thus:

$$dS_t = r_t S_t dt + \sigma_t S_t d\hat{W}_t \quad \text{where } \hat{W}_t \text{ is a standard Brownian motion under } Q$$

Now, by an application of Itô's lemma:

$$\ln S_t \sim \phi(\ln S_0 + \int_0^t r_u du - \frac{1}{2} \int_0^t \sigma_u^2 du, \int_0^t \sigma_u^2 du)$$

If  $r_t$  and  $\sigma_t$  were deterministic, and thus predictable, it would be possible to derive a formula for the price of a European call option using the methods of Chapter 3 to yield:

$$\begin{aligned} g(0, \tau, S_0, X_0) &= E^Q \left[ \exp \left( - \int_0^\tau r_u du \right) (S_\tau - K)^+ \mid \mathfrak{F}_0^X \right] \\ &= S_0 N(d_1(0)) - K \exp \left( - \int_0^\tau r_u du \right) N(d_2(0)) \end{aligned} \quad (29)$$

$$d_1(0) = \frac{\ln \left( \frac{S}{K} \right) + \left( \int_0^\tau r_u du + \frac{1}{2} \int_0^\tau \sigma_u^2 du \right)}{\sqrt{\int_0^\tau \sigma_u^2 du}}$$

$$d_2(0) = d_1(0) - \sqrt{\int_0^\tau \sigma_u^2 du}$$

However,  $r_t$  and  $\sigma_t$  are not deterministic and so the above formula contains quantities that are unknown. As in the previous section, this difficulty is overcome by defining the occupation times:  $O_\tau^i$ , which denote the amount of time that  $X_t$  spends in state  $i$  up until time  $t$ . Under a continuous time setting these are now

calculated as:  $O_\tau^r = \int_0^\tau \langle X_u, e_i \rangle du$ . Notice that  $\sum_{i=1}^N O_\tau^i = \tau$  as before, and so the set of

$\{O_\tau^i\}$  is completely determined by  $N-1$  of its elements. Taking an expectation over these  $N-1$  values will then determine the call price, as in the previous section:

$$g(0, \tau, S_0, X_0) = \int_0^\tau K \int_0^\tau g(0, \tau, S_0, r_\tau, \sigma_\tau) \varphi(O_\tau^1, K, O_\tau^{N-1}) d_{\xi_1} K d_{\xi_{N-1}} \quad (30)$$

where  $\varphi(O_\tau^1, K, O_\tau^{N-1})$  is the *pdf* of the first  $N-1$  occupation times

Notice that  $r_t = \int_0^t \langle r, X_u \rangle du = \sum_{i=1}^N r_i O_\tau^i$ , as in the previous section, and similarly  $\sigma_t =$

$\sum_{i=1}^N \sigma_i O_\tau^i$ . Since these are exactly the same quantities as in (17) and (18), with the occupation time now being a continuous variable, it appears that the continuous time specification is somewhat unnecessary. The events of interest are, ultimately, discrete and the continuous nature of the time index does not seem to add any value to the model. Indeed, in light of the nature of the *pdf* of the occupation times, derived below, this feature seems merely to complicate matters.

Finally, as in the previous section, the unconditional call price is:

$$g(0, \tau, S_0) = \sum_{j=1}^N \pi(j) g(0, \tau, S_0, X_0)$$

It now remains to derive the form of the *pdf* of the occupation times used in the evaluation of (30). This derivation can be found in appendix A1 of Elliott and Buffington (2002).

First, begin by deriving the characteristic function of  $\varphi$ :  $\Psi_\varphi(\theta) = E(\exp(i\theta O_\tau))$  where  $i^2 = -1$ ,  $\theta$  is a vector of parameters and  $O_\tau$  is the  $N$ -vector containing the occupation times of the first  $N-1$  states, up to time  $\tau$ , and 0.

Then:

$$\begin{aligned}\Psi_\varphi(\theta) &= E \left[ \exp \left( i \sum_{j=1}^{N-1} \theta_j O_\tau^j \right) \right] \\ &= E \left[ \exp \left( i \int_0^\tau \langle \theta, X_u \rangle du \right) \right]\end{aligned}$$

because

$$\begin{aligned}\sum_{j=1}^{N-1} \theta_j O_\tau^j &= \sum_{j=1}^{N-1} \theta_j \int_0^\tau \langle X_u, e_j \rangle du \\ &= \int_0^\tau \sum_{j=1}^{N-1} \theta_j \langle X_u, e_j \rangle du = \int_0^\tau \sum_{j=1}^{N-1} \langle \theta_j X_u, e_j \rangle du \\ &= \int_0^\tau \langle \theta, X_u \rangle du\end{aligned}$$

define

$$Z_t := \exp \left( i \int_0^t \langle \theta, X_u \rangle du \right) X_t$$

then

$$\begin{aligned}dZ_t &= \exp \left( i \int_0^t \langle \theta, X_u \rangle du \right) dX_t + X_t \exp \left( i \int_0^t \langle \theta, X_u \rangle du \right) (i \langle \theta, X_t \rangle) dt \\ &= \exp \left( i \int_0^t \langle \theta, X_u \rangle du \right) (\mathbf{P} X_t dt + dM_t) + X_t \exp \left( i \int_0^t \langle \theta, X_u \rangle du \right) (i \langle \theta, X_t \rangle) dt \\ &= \mathbf{P} \exp \left( i \int_0^t \langle \theta, X_u \rangle du \right) X_t dt + X_t \exp \left( i \int_0^t \langle \theta, X_u \rangle du \right) (i \langle \theta, X_t \rangle) dt + \exp \left( i \int_0^t \langle \theta, X_u \rangle du \right) dM_t \\ &\Rightarrow Z_t = X_0 + \int_0^t (\mathbf{P} + i \times \text{diag}(\theta)) Z_u du + \int_0^t \exp \left( i \int_0^u \langle \theta, X_u \rangle du \right) dM_u \\ &\Rightarrow E[Z_t] = X_0 + \int_0^t (\mathbf{P} + i \times \text{diag}(\theta)) E[Z_u] du \\ &\Rightarrow E[Z_t] = X_0 \exp(\mathbf{P} + i \times \text{diag}(\theta)) t\end{aligned}$$

Now,  $Z_t$  is a vector-valued process, of dimension  $N$ . The elements of  $Z$  must therefore be summed to find the characteristic function value:

$$\begin{aligned}
& E \left[ \exp \left( i \int_0^{\tau} \langle \theta, X_u \rangle du \right) \right] \\
&= E \left[ \left\langle \exp \left( i \int_0^{\tau} \langle \theta, X_u \rangle du \right) X_{\tau, \mathbf{1}} \right\rangle \right] \\
&= \langle \exp((\mathbf{P} + i \times \text{diag}(\theta))), \mathbf{1} \rangle
\end{aligned} \tag{31}$$

where  $\text{diag}(\theta)$  is the  $(N-1) \times (N-1)$  matrix with the elements of  $\theta$  on its diagonal and 0 elsewhere.  $\mathbf{1}$  is the  $(N-1)$  vector containing unity in each element.

As can be seen, this is a complex function, which leads to a *pdf*, found by taking the inverse Fourier transform of the characteristic function. This complicated function would be difficult to evaluate in practice, which leads to the question of the practicality of this framework. For estimation of parameter values and the option price, it is suggested that the discrete time model would be a more feasible approach.

## Chapter 7. Extensions to the Basic Hidden Markov Option Pricing Model

The form of the Hidden Markov Option Pricing Model described in the previous chapter is not the only possible form of the model. Various extensions and slightly alternative formulations are possible and a selection of these will be discussed below. Some of these alternatives aim to try and increase the accuracy of the model. An example of this would be where the transition probabilities of the Markov Chain that describes  $\{X_t\}$  are allowed to vary through time. Other versions seek to incorporate effects such as a feedback between the asset return innovations and the value of  $X_t$ . This can be used to model the 'leverage' effect identified by Black (1976). This is the phenomenon whereby lower mean rates of return on equities are accompanied by higher levels of volatility of returns. Another effect that can be included, such as that considered by Guo (2001), is the effect of insider information on traded asset prices. This is achieved by considering a state process where the states describe whether insider information is affecting price determination, or not.

The examples of extensions that are presented here are by no means exhaustive. As was noted when considering the advantages of the Hidden Markov Model for option pricing, this is a very extensible, flexible framework which can be modified in many different ways to take account of many particular effects of interest. The intention of this chapter, then, is to identify some of the possible extensions, that have already been presented in the literature, to indicate how further modifications might be undertaken. This will allow the reader to take full advantage of the flexibility of the Hidden Markov framework.

### 7.1 Time-varying Transition Probabilities

The model presented in the previous chapter assumed that the transition probability matrix of the Markov Chain that is used to describe  $\{X_t\}$ ,  $\mathbf{P}$ , remains constant over time. This allowed the use of the identity in (21) for  $\mathbf{P}_u$ , which is the transition probability matrix for  $u$  consecutive transitions. This simple formula then allowed for the straightforward recursion of (22) for the first passage probabilities  $f_{ij}(t)$ . These were then used to evaluate the occupation time probabilities that are used in the option pricing formula. Although this scheme is appealing in its simplicity, there is evidence (for example in Gray (1996)) that allowing the transition probabilities to vary over time may result in a more accurate model of the asset price. This could be due to the additional dynamics in the transition probabilities resulting in the overall model better capturing the overall dynamics of the asset price process. Two approaches to the use of time-varying transition probabilities will be presented here. These can be found in Gray (1996), p37 and Tzavalis and Chourdakis (1999), p12.

Although the exact parameterisation of the two methods differs slightly, the main means of inducing time-varying probabilities is the same in both. A function of an ARMA-type process is specified for the transition probabilities, i.e.  $p_{ij,t} = g(\xi(t))$  where  $g(\cdot)$  is chosen so that its range is the interval  $[0,1]$ . This ensures that the values generated for the  $p_{ij,t}$  satisfy the conditions required to be probabilities.

The condition  $\sum_{j=1}^N p_{ij,t} = 1$  (for each  $t$ ) can be enforced by only generating  $N-1$  of the transition probabilities, for each unobserved state, and then computing the final probability as:  $p_{iN,t} = 1 - \sum_{j=1}^{N-1} p_{ij,t}$ , for each  $i$  and each  $t$ .

Tzavalis and Chourdakis (1999) use a quintic spline function for  $g(\cdot)$  where  $g(a) = 0$  if  $a \leq 0$  and  $g(a) = 1$  if  $a > 1$ .

$$\begin{aligned} g(a) &= 0 \\ g(a) &= a^3(6a^2 - 15a + 10) && \text{if } 0 < a \leq 1 \\ g(a) &= 1 \end{aligned}$$

They motivate the choice of such a function by requiring that  $g(\cdot)$  be a sufficiently smooth map from  $\mathbb{R}$  to  $[0,1]$ . 'Sufficiently' smooth in this case being that the first and second derivatives of  $g(\cdot)$  exist and are continuous, i.e.  $g(\cdot) \in C^2$ . Gray (1996) uses the cumulative standard normal distribution function for  $g(\cdot)$ :  $N(\cdot)$ . This also ensures that the values produced by the equation for  $p_{ij,t}$  lie in  $[0,1]$  and is a smooth function. In fact,  $N(\cdot)$  is differentiable infinitely often and so from the point of view of smoothness may be a preferable specification for  $g(\cdot)$ . The caveat here is that  $N(\cdot)$  is more complicated to compute than the polynomial spline that Tzavalis and Chourdakis (1999) use. Another possibility for  $g(\cdot)$  is a scaled version of the hyperbolic tan function:  $\tanh(\cdot)$ . This is also differentiable infinitely often and is perhaps easier to compute than  $N(\cdot)$ . Thus it might be an acceptable compromise between  $N(\cdot)$  and the polynomial spline, in terms of computability and smoothness criteria. The scaled version of  $\tanh$ , scaled to

have a range of  $[0,1]$  is:  $g(a) = \frac{1}{2}(\tanh(a) + 1)$

The form of  $\xi(t)$  that is chosen can then be fairly general, since  $g(\cdot)$  takes care of the fact that the output of the equation for  $p_{ij,t}$  must lie in the interval  $[0,1]$ . Gray (1996) uses:  $\xi(t) = c_i + d_i a_{t-1}$  where  $a_t$  was the variable being modelled and  $i$  is the unobserved state value. This equation allows the transition probabilities to depend on the level of the observed process being modelled. Gray (1996) shows that this can allow for mean reversion in the observed process (Gray (1996), p37).

Alternatively, Tzavalis and Chourdakis (1999) use  $\xi(t) = a_{ij} + k_{ij} p_{ij,t} + \phi_{ij}(\sigma_{t-1} \varepsilon_{t-1})$ , where  $a$ ,  $k$  and  $\phi$  are constants and  $\varepsilon_{t-1}$  is the asset return innovation from the previous time point. This, of course, assumes that a model is being used for the

asset returns and not just an observation equation, as under the HMM specified in the previous chapter. In this way the current transition probability depends on the previous transition probability (in an autoregressive manner) and the previous return innovation. Incorporating the previous return innovation can, according to Tzavalis and Chourdakis (1999), be thought of as introducing a 'learning' effect into the model. The new information, not previously included in the model, from the most recent return innovation is included in the current transition probabilities, thus updating them by the content of this new information. Clearly, this is by no means an exhaustive list of the possible specifications for  $\xi(t)$ . It does, however, give an indication of the possible specifications that could be used and the effects that these different specifications take into account, such as level effects. Many other possible forms of  $\xi(t)$  could be specified, according to the properties of the model that were required, such as mean reversion (as in Gray (1996)), etc.

Although the framework for specifying time-varying transition probabilities is relatively simple, this extension of the asset price process HMM complicates the evaluation of option prices in the Hidden Markov Option Pricing Model. This is because (21) no longer holds and so the simple recursion for  $f_{ij}(t)$ , (22), breaks down. The evaluation of the occupation time probabilities is now no longer a simple linear recursion, as in Chapter 6. The nonlinear functions for  $p_{ij,t}$ , assuming the parameters of  $\xi(t)$  have been estimated, introduce a considerable amount of complexity into the calculation of the occupation time probabilities. Furthermore, there is now a dependence between successive values of  $p_{ij,t}$ . Tzavalis and Chourdakis (1999) thus propose a Monte Carlo method for evaluating these probabilities, from time  $t$  to time  $\tau$ . It consists of the following steps:

1. Draw values for  $X_t$  and  $\ln\left(\frac{S_{t+h}}{S_{t+h-1}}\right)$  ( $h = 1, 2, \dots, t-\tau$ ), using the parameters that have been estimated from the data,  $\Theta$ , the specified observation equation and given the values  $\Pr(X_t = i | \mathfrak{F}_{t-1})$  ( $i = 1, \dots, N$ ).  $\ln\left(\frac{S_{t+h}}{S_{t+h-1}}\right)$  will be drawn using normal random variables and  $X_t$  will be drawn using an approach similar to that in Lavery *et al.* (2002), using  $\Pr(X_t = i | \mathfrak{F}_{t-1})$ . (Lavery *et al.* (2002), p34)
2. Any further inputs required for the transition probability equation, such as the asset return innovation (being the difference between the actual return and the return value generated by the model), are calculated. The specified transition probability equation is used to update the transition probabilities to  $p_{ij,t+h}$ . These are used, together with Bayes' Rule to update the  $\Pr(X_t = i | \mathfrak{F}_{t-1})$  to  $\Pr(X_{t+h} = i | \mathfrak{F}_{t+h-1})$ . This allows a value for  $X_{t+h}$  to be drawn, again a method similar to the method of Lavery *et al.* (2002), for each  $h$ .

3. Compute the simulated occupation times as  $O_{t \rightarrow \tau}^i = \sum_{h=1}^{\tau-t} I[X_{t+h} = i]$  ( $i = 1, \dots, N$ ).

$I[A]$  is the indicator function for event A.

Repeat steps 1 to 3  $M$  times, where  $M$  is the number of simulations chosen. The simulated occupation time probabilities can then be computed as:

$$\Pr(O_{t \rightarrow \tau}^i = \zeta) = \frac{1}{M} \sum_{n=1}^M I[O_{t \rightarrow \tau}^i(n) = \zeta] \text{ where } I[\cdot] \text{ is again the indicator function.}$$

(Tzavalis and Chourdakis (1999), pp. 23-24)

If the time point  $(t+h)$  lies within the sample, then the equations used to update the  $\Pr(X_{t+h} = i | \mathfrak{F}_{t+h-1})$ , from Gray (1996), are as follows. Let  $\gamma_t = \ln\left(\frac{S_t}{S_{t-1}}\right)$  and  $f(\gamma_t | \mathfrak{F}_{t-1}^S)$  be the conditional density of  $\gamma_t$ , conditional on the information relating to the observed values from previous time points. Then:

$$\begin{aligned} f(\gamma_t | \mathfrak{F}_{t-1}) &= \sum_{i=1}^N f(\gamma_t, X_t = i | \mathfrak{F}_{t-1}) \\ &= \sum_{i=1}^N f(\gamma_t | X_t = i, \mathfrak{F}_{t-1}) \Pr(X_t = i | \mathfrak{F}_{t-1}) \end{aligned}$$

and

$$\Pr(X_t = i | \mathfrak{F}_{t-1}^S) = \sum_{j=1}^N \Pr(X_t = i | X_{t-1} = j, \mathfrak{F}_{t-1}^S) \Pr(X_{t-1} = j | \mathfrak{F}_{t-1}^S) \quad (32)$$

where

$$\begin{aligned} \Pr(X_{t-1} = i | \mathfrak{F}_{t-1}^S) &= \Pr(X_{t-1} = i | \gamma_t, \mathfrak{F}_{t-2}^S) \\ &= \frac{f(\gamma_{t-1} | X_t = i, \mathfrak{F}_{t-2}^S) \Pr(X_{t-1} = i | \mathfrak{F}_{t-2}^S)}{\sum_{i=1}^N f(\gamma_{t-1} | X_{t-1} = i, \mathfrak{F}_{t-2}^S) \Pr(X_{t-1} = i | \mathfrak{F}_{t-2}^S)} \end{aligned}$$

where the last result follows by an application of Bayes' Rule.

Note that the last step in (32) essentially computes what will be referred to as the 'filtered' probability of the unobserved state process taking on a given value in Chapter 10. The second step computes a 'predicted' probability of the unobserved state process taking on a given value, using the transition probabilities of the Markov Chain, at the given time point, and the 'filtered' probability value for the previous time point. This is also similar to a step in the computation of the 'filtered' probabilities in Chapter 10.

If the  $(t+h)$  lies outside the sample used to fit the model, then a predicted value for the updated probability is required. This is computed by multiplying the vector

containing the filtered probabilities of being in each state, at the final time point in the sample, by the successive transition probability matrices generated using the parameter estimates for  $\xi_t$ . The number of multiplications performed is equal to the number of time points between the last time point in the sample and  $(t+h)$ .

## 7.2 GARCH effects in the Variance of the Observed Process

Another extension to the model which may be of interest would be to allow for the variance of the observed process of the HMM to follow a GARCH process. This would allow for the 'volatility clustering' that has been observed in asset price returns to be incorporated into the model. It has been found that periods of high volatility tend to follow periods of high volatility and vice versa. This is another way of saying that there is persistence in the level of the volatility of asset price returns over time. The autoregressive component of the GARCH model could replicate this effect, by directly including previous volatility values in the model of current volatility. If the parameters of the GARCH process are further allowed to depend on the value of the unobserved state process,  $\{X_t\}$ , this could allow for the unconditional level of the volatility implied by the specified GARCH process to change over time. This would thus avoid the problem of overestimating the persistence of the volatility series identified by Lamoureux and Lastrapes (1990), for the case of GARCH models with only a single unconditional level of volatility.

Under such a specification:  $\ln\left(\frac{S_t}{S_{t-1}}\right) \Big| X_t = (\mu_{it} - \frac{1}{2}h_{it}) + \sqrt{h_{it}}Z_t$ , where  $Z_t \sim \phi(0,1)$  and

so  $\ln\left(\frac{S_t}{S_{t-1}}\right)$  is conditionally normally distributed:  $\ln\left(\frac{S_t}{S_{t-1}}\right) \Big| X_t \sim \phi(\mu_{it} - \frac{1}{2}h_{it}, h_{it})$ .  $h_{it}$

is the value of the conditional volatility process, as modelled by the GARCH process. There is a difficulty in this specification as noted by Gray (1996), p34, however. Since the GARCH process has an autoregressive component to it, the Markov dependence structure of the unobserved state process that simplified computations in the 'standard' HMM no longer applies. This is because the autoregressive terms in the GARCH model introduce dependence on all previous volatility values, which depend on previous values of the unobserved state, in a recursive manner. The problem then becomes that at each time point  $t$ , with  $N$  states, there are  $N^t$  components of the likelihood function at  $t$ . This is a result of the uncertainty regarding the previous unobserved state values. This means that any estimation procedures would soon become unfeasible, for large samples.

However, Gray (1996) proposes a simple solution to this problem. The conditional variance at time  $t$ , conditional on available information is  $h_t = E[\gamma_t^2 | \mathfrak{F}_{t-1}] - E[\gamma_t | \mathfrak{F}_{t-1}]^2$ , which can be written, in the HMM case, as:

$$h_t = \sum_{i=1}^N \Pr(X_t = i | \mathfrak{F}_{t-1}) (\mu_{it}^2 + h_{it}) - \left( \sum_{i=1}^N \Pr(X_t = i | \mathfrak{F}_{t-1}) \mu_{it} \right)^2 \quad (33)$$

This value is no longer dependent on the entire previous history of  $\{X_i\}$ , but only on the most recent value of  $\{X_i\}$ . This is now similar to the case in the 'standard' HMM. The model now uses these values,  $h_t$ , as the lagged conditional variance in the GARCH model equation. Thus,  $h_{it} = \omega_i + a_i \varepsilon_{t-1}^2 + b_i h_{t-1}$  ( $i = 1, \dots, N$ ) for a GARCH(1,1) specification, with  $\varepsilon_{t-1}$  being the asset return innovation at time  $t-1$  and  $h_{t-1}$  is calculated as in (29). This allows for the inclusion of persistence effects, by the GARCH model, whilst retaining the tractability of the HMM that is a result of the Markov dependence structure of the unobserved state process.

### 7.3 Feedback effects between the State-dependent Variables and the Observed Process

This form of specification is due to Duan *et al.* (2002) and relies, essentially, on a modified form of the observation equation for the asset price returns to introduce the feedback effects. In the case of the model specified by Duan *et al.* (2002), the volatility can be one of a finite set of values, depending on the unobserved state. The feedback is between the asset return at the previous time point and the state-dependent value of volatility at the current time point. The observation equation is, under the objective probability measure:

$$\ln\left(\frac{S_{t+1}}{S_t}\right) = r + \lambda \sigma_{t+1} - \frac{1}{2} \sigma_{t+1}^2 + \sigma_{t+1} \varepsilon_{t+1} \quad \text{where } \sigma_{t+1} \text{ is the conditional volatility of the}$$

asset return process from time  $t$  to  $t+1$  and  $\lambda$  is the market price of risk, i.e. the amount of compensation per unit volatility demanded by investors to induce them to take on the risk associated with  $\sigma_{t+1}$ .  $\varepsilon_{t+1}$  is a standard normal random variable.

The feedback between return innovations and the volatility is induced as follows:  $\sigma_{t+1} = \delta_i$  if  $C_{i-1}(\sigma_t) \leq F(\varepsilon_t, \xi_t) < C_i(\sigma_t)$ . The  $C_i(\sigma_t)$  are threshold values and  $F(.,.)$  is a function of the most recent asset return innovation  $\varepsilon_t$  and a volatility innovation  $\xi_t$  that is assumed to be orthogonal to the asset return innovation. Duan *et al.* (2002) specify  $F(.,.)$  as:  $F = q_1(\varepsilon_t - \omega)^+ + q_2(\varepsilon_t - \omega)^- + (1 - q_1 - q_2)|\xi_t|$  with  $q_1 + q_2 \leq 1$  and  $q_1, q_2 \geq 0$ .  $(.)^+$  is as defined previously:  $(.)^+ = \max(., 0)$ , while  $(.)^- = \max(-., 0)$ .  $\omega$  is a bias adjustment that can be used to introduce an asymmetry into the manner in which volatility responds to return innovations. This can be used to account for the leverage effect identified by Black (1976). This effect can also be captured by requiring  $q_1$  and  $q_2$  to have different values, since these act as weights in  $F(.,.)$  which is a weighted sum of the positive and negative components of the asset return innovation, as well as a term depending on the orthogonal volatility innovation. This framework has the effect of causing the transition probabilities for the unobserved state process to become time-varying. This is because they now depend on the most recent value of the asset return innovation. Thus:  $p_{ij} = \Pr[\sigma_{t+1} = \delta_j | \sigma_t = \delta_i, \varepsilon_t]$ , which is time-varying because  $\varepsilon_t$  is time-varying.

The model can also be rewritten in a risk neutral context, using the Locally Risk Neutral Valuation Relationship of Duan (1996). The reader is referred to Duan *et al.* (2002) for this formulation, which is not very different to that presented above. Finally, it is worth noting that if  $q_1 = q_2 = 0$ , then there are no feedback effects between the asset returns and the volatility and the model reduces to the 'standard' model of Chapter 6.

## 7.4 Modelling Information Effects in the Market

Asset prices, and thus asset price returns, change when market participants buy and sell the asset in question for a price other than that prevailing in the market at the time of the given transaction. They would be willing to do so if they believed that the underlying value of the asset was different to that implied by the prevailing market price. This would be due to new information, which impacts the value of the asset, becoming available to market participants. An example of such information would be the case of an equity security where a takeover bid for the issuing company has been launched. The price of the equity will rise, in anticipation of the fact that the bidder will have to pay a premium to the prevailing market price to induce holders of the equity to sell their holdings. This represents a gain in value to the holders of the equity that did not previously exist and, accordingly, the price of the equity will rise to reflect this. Furthermore, in the case where a rival bidder emerges, there is a further potential gain in value as a result of the possibility of a bidding war between rival bidders. This would cause the eventual sale price to rise to an even higher level than that implied by the premium contained in the initial bid price. Another example would be a change of market interest rates in the case of a debt security. Different interest rates imply different discount factors for the payments promised by the security, which means that the value to the holder has changed. The price of the debt security in the market should then change as well.

In some cases, not all market participants are aware of information relevant to the value of the assets that they hold when it first becomes available. This is most frequently often the case with equity securities, where those employed by the issuing company will become aware of information relating to corporate events, such as a takeover bid, before this information is announced to all market participants. In such a situation such employees, referred to as 'insiders', could exploit this privileged access to information to make a profit at the expense of the market participants with whom they transacted. This would be possible due to the external market participant being unaware of the 'inside information' that the company employee was. It is thus possible that the existence of 'inside information' could have a significant effect on market prices, particularly in markets where regulation against the use of such information, to protect those who do not have access to it from suffering an unfair loss at the hands of one who did. Guo (2001) presents a form of the Hidden Markov Option Pricing Model that seeks to take into account the effect that 'inside information' can have on market prices.

Guo (2001) postulates two unobservable states, State 0 where no inside information is influencing market prices and State 1 where inside information is having an effect on prices. The model is a continuous time model and thus a further assumption is necessary about the distribution of the arrival times of the insider information in the market. Guo (2001) assumes that these times are exponentially distributed. This point is needed to derive the probability density functions (pdf) of the occupation times in each state, which are required for computing the option price. By a process similar to that outlined in Chapter 3, the option price is then, conditional on the value of the initial unobserved state:

$$V_i(\tau, K, r) = e^{-r\tau} E^Q[(S_\tau - K)^+ | X_0 = i]$$

$$= e^{-r\tau} \int_{\ln K}^{\infty} \int_0^{\tau} \ln S_t \phi(m(t), v(t)) f_i(t, \tau) dt d \ln S_t$$

where  $\phi(m(t), v(t))$  is the normal density function with mean  $m(t)$  and variance  $v(t)$  and  $f_i(t, \tau)$  is the pdf of the occupation time in state 0 between times  $t$  and  $\tau$ , starting from state  $i$ . In the case of the model presented by Guo (2001), these parameters take on the following values:

$$m(t) = (d_1 - d_0 - \frac{1}{2}(\sigma_0^2 - \sigma_1^2))t + (r - d_1 - \frac{1}{2}\sigma_1^2)\tau$$

$$v(t) = (\sigma_0^2 - \sigma_1^2)t + \sigma_1^2\tau$$

$$f_1(t, \tau) = e^{-\lambda_1\tau} e^{(\lambda_1 - \lambda_0)t} \left[ \left[ \frac{\tau - t}{\lambda_0 \lambda_1 t} \right]^{-\frac{1}{2}} J_{-1} \left[ 2(-\lambda_1 \lambda_0 t \tau + \lambda_0 \lambda_1 t^2)^{\frac{1}{2}} \right] + \lambda_1 \lambda_0 J_0 \left[ 2(-\lambda_0 \lambda_1 t \tau + \lambda_0 \lambda_1 t^2)^{\frac{1}{2}} \right] \right]$$

and similarly for  $f_0(t, t)$ . The variables  $\sigma_i$  and  $d_i$  are those corresponding to state  $i$ , with  $d_i$  being a variable such that if  $d_1 - d_0 > 0$ , then the return available in the market is greater than  $r$ , the risk free rate of return, which implies the existence of an 'inside information effect'. This is because pricing occurs in a risk neutral world and thus the return available should only be  $r$ . Any excess return, under this model, will be due to the effects of inside information. The functions  $J_a(z)$  are Bessel functions of the form:

$$J_a(z) = \left( \frac{1}{2z} \right)^a \sum_{n=0}^{\infty} \frac{(-1)^n \left( \frac{z}{2} \right)^{2n}}{n! \Gamma(a + n + 1)}$$

For proofs of the above and a more detailed examination of this model, the reader is referred to Guo (2001).

By reviewing the wide scope of modifications to the 'standard' HMM Option Pricing Model presented in this chapter, it is possible to gain an understanding of the flexibility and extensibility of this modelling framework. Such extensibility allows for the study of many effects of interest, such as volatility persistence with GARCH-type models, or the effects of insider information as above. The relative ease with which such extensions can be made to the Hidden Markov Option Pricing Model are one of the key advantages of this model.

## Chapter 8. Pricing Bermudan-style Options using the Hidden Markov Option Pricing Model

A Bermudan-style option, or 'Mid-Atlantic' option, is an option that shares some of the exercise features of American-style and European-style options. It lies somewhere between the two, in terms of its exercise characteristics, hence the name. A Bermudan-style option can be exercised at any of a finite number of deterministic, discrete dates up to and including expiry. In this way, by allowing exercise prior to expiry, the option is similar to an American option. However, by pre-specifying the dates on which exercise is permissible, it is like a European option. Bermudan options are most commonly used in interest rate markets, where they are written on bonds, interest rates and swaps. Bermudan option prices can also be used to approximate American option prices, with the approximation improving as the number of possible exercise dates for the Bermudan option is increased. This is because an American option may be exercised at any time prior to the expiry of the option. Thus, a Bermudan option with many possible exercise dates will mimic the behaviour of an American option.

The Hidden Markov Option Pricing Model, in discrete-time, lends itself to the evaluation of such options, which depend on the value of the variable that they are based on at a number of discrete time points. Some, or all, of the discrete time points at which the observed value of the variable underlying the option contract is evaluated by the HMM can be set to coincide with the possible exercise dates of the Bermudan option. This provides a relatively parsimonious and simple framework from which the price of the Bermudan-style option can be evaluated, that recognises the discrete nature of the time index that the option's exercise structure implies.

Other methods that have been used to price Bermudan options include Monte Carlo methods, finite difference methods and backwards recursions using compound option pricing methods. These methods are essentially the same as those used to value American options, which is not surprising given the optional exercise date characteristic that American and Bermudan options share. This feature has substantial value to the holder of the option, by allowing them to choose the 'best' time to exercise the option, i.e. the most profitable time. It will thus have a major impact on the price of the option and will thus influence the pricing method used. The advantages of the HMM framework, as discussed in Chapter 4, provide a justification for using it as an alternative to these methods. It is a more flexible and extensible framework than these other methods, in general, whilst remaining tractable and easy to understand.

The new HMM framework for pricing Bermudan options, then, is as follows. Let  $T_1, T_2, \dots, T_n$  be the discrete, deterministic, possible exercise dates of the option.  $T_n$  is thus the expiry date of the option,  $\tau$ . These dates will often be pre-specified in the option contract. At each possible exercise date, the holder of the option effectively has an implicit compound option, i.e. an option on an option. The choice that the holder faces is whether to keep the option that they hold 'alive', or to exercise it and bank the proceeds of exercise. Effectively, the holder thus has an option on an option with a time to expiry of the time remaining between the given possible exercise date and the expiry date of the Bermudan option, and a strike price equal to the value/price of the option at the given possible exercise date. The decision of whether or not to exercise the option is similar to the decision faced by the holder of an American option, although in that case the decision of whether or not to exercise applies at every instant of time and not merely at one of a discrete set of dates. Thus, a value for the option price is required at every possible exercise date, for comparison with the payoff that could be gained at that time by exercising the option. These option prices will be computed using the discrete-time Hidden Markov Option Pricing Model.

The price of the option at this, and prior, exercise dates will, thus, be evaluated using the methods described in 6.1. However, since the time interval of interest is now from the given exercise date to the expiry date, the form of the occupation times and occupation time probabilities that will be required to price the option will be slightly different to those presented in 6.1. The occupation times now required are those for the time period from the given exercise date up to expiry.

Using the notation of 6.1, the required occupation times are:  $O_{T_j \rightarrow \tau}^i = \sum_{m=T_j}^{\tau} \langle X_m, e_i \rangle$ .

The occupation time probabilities associated with these values can be calculated using (23), (24) and (25) from Chapter 6, assuming that the transition probability matrix of the Markov Chain being used to describe the unobserved state process is time-homogeneous. This assumption, together with the fact that the dependence structure of the model is Markov, implies that the calculation of occupation times from any given time point onwards is the same as the calculation starting from time 0 onwards. The key variable is the length of the time interval over which the occupation times are to be computed. The difference between the calculations from 6.1 and those that will be used here, then, is that the starting date for each set of calculations will be the potential exercise date of the Bermudan option in question. Thus, the first passage probabilities required are the probability of first passage to a given unobservable state after  $j$  time steps, starting from the exercise date in question. Similarly, the occupation times will be measured as starting from each possible exercise date, for each set of calculations relating to each possible exercise date.

Once the occupation time probabilities,  $\Pr(O_{T_j}^1 = \zeta_1, K, O_{T_j}^{N-1} = \zeta_{N-1} | X_{T_j} = e_h)$  have been calculated, the option price at exercise time  $T_j$  can be evaluated using (19) and (20) in Chapter 6. A slight modification is required, to allow for the fact that the evaluation time is  $T_j$  and not time 0:

$$C(T_j, \tau, S) = \sum_{h=1}^N \Pr(X_{T_j} = e_h | \mathfrak{F}_{T_j-1}) C(T_j, \tau, S_0, X_{T_j} = e_h) \quad (34)$$

where the notation is as in 6.1. The probabilities  $\Pr(X_{T_j} = e_h | \mathfrak{F}_{T_j-1})$  will be computed using a forecast generated using the transition probability matrix of the Markov Chain. The values for the filtered probabilities of being in each state, at the final time point in the sample that the model is fitted to, will be updated by multiplying the final row of the matrix storing these probabilities by the transition probability matrix as many times as there are days between  $T_j$  and the last date in the sample. This is because the options expire in the future, beyond the end date of the sample. Multiplying the filtered probabilities by the transition probability matrix forecasts these probabilities by 1 day into the future, taking into account the fact that the options pricing is being performed 'out of sample'. Then the procedure for evaluating the option price at each possible exercise date is very similar to the procedure presented in 6.1, with the calculations presented there repeated for each possible exercise date. Note that the  $\zeta_i$  increase with every step back through the exercise dates, by the amount of time between the possible exercise dates.

Once the option price at a given exercise date has been calculated, a value for the variable underlying the option contract must be simulated, for that date. This is done using the observation equation of the HMM for the underlying variable, (15). Since the observed value will depend on the value of the unobserved state variable, at the given date there will be  $N$  possible observed values. The value used for comparison with the option price calculated for that date, to determine whether the option is exercised or not, will thus differ from state to state.

The value of the option at the given exercise date is then the maximum of  $(S_{\text{comp}} - K)$  and the option price that was evaluated at that exercise date, where  $S_{\text{comp}}$  is the simulated observed variable value. An expectation is then taken over the  $N$  option values that this comparison produces, using the forecast probabilities of being in each state, at the given exercise date. This expected value is then discounted to time 0, using the risk-free rate of return. This is because it is assumed that pricing occurs under risk neutral conditions. The overall option price is then obtained by summing over the  $n$  discounted values, one for each exercise date. The value of the option on the expiry date is, of course, merely the difference between  $S_{\text{comp}}$ , at that date, and  $K$ . Thus, as was the case in 6.1, the calculations are lengthy but straightforward, in that they are linear.

## Chapter 9. A Lattice Pricing Method for the HMM Framework

Lattice pricing methods for options provide a simple framework for evaluating the price of relatively complex options, such as American options and exotic options such as barrier options. This is part of the reason that these methods are presented here, to provide access to the simplicity of this framework within the HMM framework. These methods will also be for comparison purposes, to calculate prices to compare against the values computed for Bermudan-style options using the methods presented in the previous chapter. Lattice methods include finite difference methods as well as binomial and trinomial tree methods. In the case where there are many assets underlying the option in question, these methods become intractable and other numerical procedures, such as Monte Carlo simulation, must be used to evaluate these more complex options. The form that these methods can take when unobserved states are present is somewhat limited due to the additional complexity that the state variables introduce. In this chapter a trinomial tree method with an unobserved state process that can take one of two possible values is considered.

The common binomial tree method, of Cox *et al.* (1979), uses a two dimensional tree to approximate option values. The time to expiry,  $\tau$  assuming that now is time 0, is divided into  $n$  sub-intervals of length  $\Delta t$ . At each time step there are a set of nodes in the binomial tree. At each node in the tree it is assumed that the asset price can either decrease by a given amount, or increase by a given amount. This amount, which can be different for the up and down steps but is assumed to be the same by Cox *et al.* (1979), is known as the 'step size'. There are thus  $n$  nodes in the binomial tree at the  $n$ -th time step. Essentially, the tree approximates the dynamics of the asset price, in the Black-Scholes market a Geometric Brownian Motion, by a binomial distribution. It thus represents a discrete approximation of the continuous dynamics commonly assumed for the asset price, as in the Black-Scholes model.

The tree is calibrated to the dynamics of the asset price by moment matching. The parameters required are the probabilities of an up and down move,  $p$  and  $q$  respectively, and the step size. The first and second moments of the asset price, at a given node and under risk neutral probabilities, are calculated from the specified dynamics. The moments implied by the parameters of the tree are then also computed and equated to the moments of the asset price implied by the given dynamics. Together with the requirement that the probabilities add up to 1, these equations form a system, which can be solved to produce values for the parameters. In order for a unique solution to the system to exist, a further constraint is required. Cox *et al.* (1979) use:  $ud = 1$ , where  $u$  is the increase factor relating to the up move in the asset price and  $d$  is the decrease factor associated with a down move. So to generate an up move, the asset price at the

current node is multiplied by  $u$  and similarly for  $d$  and a down move. This is another way of writing the condition that the step size of an up move is the same as the step size of a down move. The parameters that Cox *et al.* (1979) thus arrive at are:

$$\begin{aligned} u &= e^{\sigma \sqrt{\Delta t}} \\ d &= e^{-\sigma \sqrt{\Delta t}} \\ p &= \frac{e^{r\Delta t} - d}{u - d} \\ q &= 1 - p \end{aligned} \tag{35}$$

with  $r$  being the risk-free rate of return and  $\sigma$  the volatility of the asset price returns.

In the case where the asset price is modelled by an HMM, a separate tree is required for each unobserved state since the evolution of the asset price under the influence of a particular given state will be different to that under the other possible states. For example, if there are different states relating to differing levels of volatility then the step size of the tree relating to a higher level of volatility will be larger than that relating to a lower level of volatility. This is because the asset price volatility is a measure of the average variability of the asset price. Bollen (1998) studies the case where there are two unobserved states, one relating to high volatility and one relating to low volatility. The two resulting binomial trees are superimposed to produce what Bollen (1998) calls a 'quadrinomial' lattice. The inner two branches, at a given node in the tree, represent the tree relating to the low volatility state, since the asset price varies less when volatility is lower, whilst the outer branches represent the tree relating to the high-volatility state.

This form of lattice is somewhat problematic, however, since it does not recombine very efficiently, as discussed by Bollen (1998). Unless the volatility of the high-volatility state is exactly half that of the low-volatility state, then the number of nodes at time step  $t$  will be  $t^2$ . This means that, in general, this lattice method will only be feasible for a low number of time steps. This is undesirable because a low number of time steps will produce a poor approximation of the continuous dynamics of the asset price process. Bollen (1998) proposes a 'pentanomial' lattice scheme, with five branches, to overcome this difficulty. This consists of two superimposed trinomial trees, one for each volatility state, that share the middle branch along which the asset price remains unchanged. Again, the outer branches correspond to the high-volatility state and the inner branches to the low-volatility state. The shared fifth branch is added to improve the recombining properties of the lattice. The following diagram, taken from Bollen (1998), illustrates this lattice system.

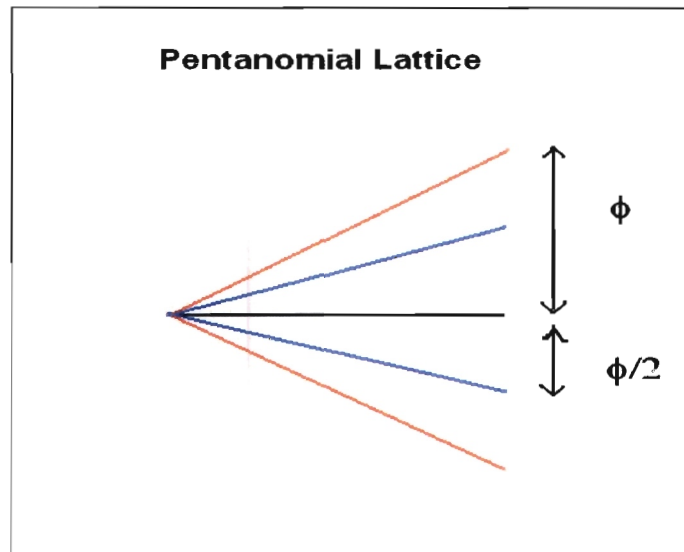


Figure 4: The pentanomial lattice structure of Bollen (1998)

The red branches represent the tree relating to the high-volatility state, whilst the blue branches represent those of the low-volatility state. The black branch is that which both trees share when they are superimposed. By further requiring that the step size  $\phi$  of the high-volatility state tree is twice that of the low-volatility state tree, Bollen (1998) creates an equally spaced 'pentanomial' lattice. This lattice then, as shown by Bollen (1998), recombines more efficiently than the 'quadrinomial' version. At each time step there are now  $(4t-3)$  nodes, as opposed to  $t^2$  in the 'quadrinomial' case. This is illustrated by the diagram below, also taken from Bollen (1998):

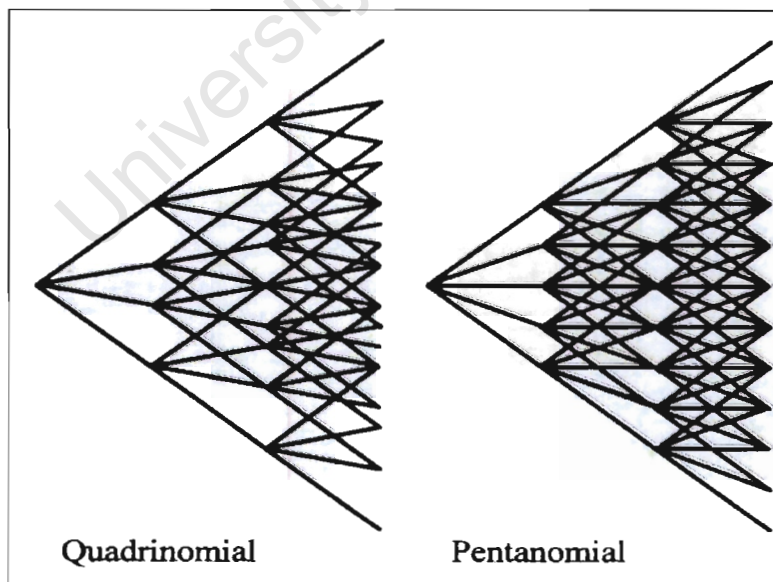


Figure 5: A comparison of the recombining properties of the 'quadrinomial' and 'pentanomial' lattices

The step size  $\phi$  is found by calculating the binomial tree parameters for both the low-volatility and high-volatility states. In the case of the specification used by Cox *et al.* (1979) the step size is  $(e^{\sigma\sqrt{\Delta t}} - 1)$ . The step sizes for the two states are then compared and adjusted to meet the requirement of Bollen (1998) that they are in the ratio 1:2. Bollen (1998) proposes increasing the step size of the low-volatility state tree if it is less than half of the high-volatility state tree, and increasing the step size of the high-volatility state tree if it is less than double that of the low-volatility state tree. The probabilities of an up and down move, for the high-volatility-state, are then kept as those calculated using the binomial tree equations. This implies that the probability of a move along the middle branch is 0 for the high-volatility state. Given that a high level of volatility will produce a very variable asset price; this is not an unreasonable assumption. The probabilities for the low-volatility state tree are then, according to Bollen (1998), computed using the following set of equations. As in the case of the binomial tree equations, these equations arise from a moment matching procedure. The reader is referred to Bollen (1998) for a derivation.

$$p_{l,u} = \frac{e^{r\Delta t} - e^{-\frac{\phi_h}{2}} - p_{l,m}(1 - e^{-\frac{\phi_h}{2}})}{e^{\frac{\phi_h}{2}} - e^{-\frac{\phi_h}{2}}}$$

$$p_{l,m} = 1 - 4\left(\frac{\phi_l}{\phi_h}\right)^2 \quad (36)$$

$$p_{l,d} = 1 - p_{l,u} - p_{l,m}$$

$p_{l,u}$  is the probability of an up move in the low-volatility state,  $p_{l,m}$  is the probability of no change in value (i.e. a sideways move in the asset price) in the low-volatility state and  $p_{l,d}$  is the probability of a down move in the low-volatility state.  $\phi_l$  is the step size for the low-volatility state and  $\phi_h$  is the step size for the high volatility state.

These equations are problematic, however. Since  $\phi_h$  has already been set to half of  $\phi_l$ , it follows that  $p_{l,m}$  will be 0. Since this is the low-volatility state, it seems unreasonable that the probability of no change in the asset price will be 0. It should be reasonably large, given that a low level of volatility should be associated with a lower amount of variation in the asset price. Referring to the numerical example that Bollen (1998) provides, it appears that he assigns a probability of 0.5 to the event of no change in the asset price, for each time step. The equations from the moment matching procedure then assign the remaining transition probability, at each time step, to an up move and a down move in the asset price, using the above equations for  $p_{l,u}$  and  $p_{l,d}$ , respectively. This rather arbitrary assignment of a value for  $p_{l,m}$  is necessary because the moment matching equations provide no specification for this value. It seems reasonable, however, given that this is the low-volatility state that a relatively large probability should be assigned to the event of no change in the asset price. This is the approach that will therefore be used.

Pricing of options under this lattice system is then similar to that of the binomial tree of Cox *et al.* (1979). In this case the payoffs under the option contract are calculated at time  $\tau$ , the expiry date of the option, being the maximum of 0 and the difference between the strike price of the option and the price of the asset, the exact form of which depends on whether the option is a call or a put. The price of the option at time 0 is then computed using a backwards recursion through the tree. The price at each node is calculated as the discounted expected payoff from the nodes at the ends of the branches extending from that node. The expectation is calculated using the probabilities calculated from the binomial tree equations and the discounting occurs at the risk-free rate of return. This is because the binomial tree is a discrete approximation of the Black-Scholes market, where pricing occurs on a risk neutral basis. Thus if  $price(i,j)$  is the price at the  $i$ -th time step and the  $j$ -th space step, then the value of  $price(i,j)$  is:

$$e^{-r\Delta t} \{p(price(i+1, j+1)) + q(price(i+1, j-1))\}. \quad (37)$$

For American-style options, this value must be compared to the value that could be gained by exercising the option at the given node. The option value, at the given node, is then the maximum of this exercise value and that calculated by the expectation.

Allowing the asset price to depend on unobserved states introduces a further complication into the calculations. The probability that the unobserved state, at a given time step, was a particular value varies over time. Furthermore, the value of the unobserved state can change at each node in the lattice. To deal with this complexity, Bollen (1998) proposes calculating two conditional option prices at each node, conditional on the value of the unobserved state at the previous time step. This is because there are two possible values for the unobserved state value. In terms of the superposed architecture that is being used, an option price is calculated for each tree, conditional on the fact that the asset price is evolving according to that tree at the given node. Thus, if  $C(t_i^h)$  is the value of the option at time step  $t$ , conditional on the unobserved state being the high-volatility state, and  $\chi = p_{hh}$ ,  $\psi = p_{ll}$  then  $C(t_i^h) = e^{-r\Delta t} \{ \chi E[C(t+1_i^h)] + (1-\chi) E[C(t+1_i^l)] \}$  with the expectations being taken using the computed lattice probabilities, evaluated using the probability equations for the high and low-volatility states, respectively. A similar result is true for  $C(t_i^l)$ . The option price at time 0, conditional on the unobserved state value at time 0, is then calculated using a backwards recursion procedure, as in the binomial tree of Cox *et al.* (1979).

There are thus two possible option prices at time 0, or at the time at which the price is evaluated. Bollen (1998) assumes two possibilities at this point. Either market participants are aware of the state value at the time of pricing, in which case they know which price to use, or the state value remains unobserved. In this case market participants will have to make a probabilistic inference about the state value at this time. They would do so using Bayes' Rule, as in (31), for in-

sample option prices and would forecast the probability of being in a given state using the transition probability matrix of the Markov Chain otherwise. The option price that they would then calculate would be a weighted average of the two option prices from the lattice. The weights would be the conditional probabilities that the state value at the time of evaluation was that conditional on which the given option price was calculated.

A final point that needs to be noted about the use of this type of lattice relates to changes in the time step size. The accuracy of the approximation of an option price that is calculated using lattice methods increases as the step size decreases. Thus, it may be desirable to alter the time step size of the lattice. However, the price of an option at a given node in the lattice depends on  $p_{hh}$  and  $p_{ll}$ . These, in turn, depend on the size of the time step between the nodes as they are measures of the persistence of the high-volatility and low-volatility states, respectively. Thus, if the size of the time step changes, then these values will need to be updated as well. This is done by using the updating equations in Bollen (1998). These are reproduced below.

$$\begin{aligned} \Xi &= \chi^2 + (1 - \chi)(1 - \psi) \\ \text{and} \\ \Psi &= \psi^2 + (1 - \psi)(1 - \chi) \end{aligned} \tag{38}$$

$$\Rightarrow \begin{cases} \psi^2 + 1 - \psi - (1 - \psi)\sqrt{\psi^2 + \Xi - \Psi} - \Psi = 0 \\ \chi = \sqrt{\psi^2 + \Xi - \Psi} \end{cases}$$

where  $\Xi$  and  $\Psi$  are the parameters relating to the lattice prior to reducing the time step size, with  $\Xi = p_{hh}$  and  $\Psi = p_{ll}$ . The corresponding parameters for the new lattice are  $\chi$  and  $\psi$ , respectively.

The fact that the equation for  $\psi$  is quadratic in  $\psi$  implies the existence of two possible roots to this equation. Bollen (1998) suggests using the larger root as this will be consistent with the inverse relationship that exists between the length of the time interval, implied by the size of the time step, and persistence in the unobserved state value. The form of the equation for  $\psi$  does not lend itself to easy analytical solution and is probably most easily solved using a numerical method such as Newton's method. Having determined the new values for  $\chi$  and  $\psi$ , the lattice for the new time step size can then be calibrated using the equations for the lattice probabilities, as previously. The implementation then follows as described for lattices where the time step size has not been changed.

## Chapter 10. Fitting the HMM to the Variable Underlying the Option Contract

The Hidden Markov Option Pricing Model has been fully described in previous chapters and possible advantages that can be gained from using this model have been reviewed. In this chapter, the model will be tested by fitting it to data. The discrete-time version of the model will be used for this purpose. The data chosen was the Johannesburg Securities Exchange (JSE) All Share Index, from South Africa, and the Financial Times Stock Exchange (FTSE) 100 Index from the United Kingdom. As noted in Chapter 1, ten years of daily data was used, over the period 15 March 1996 - 15 March 2006. All data was sourced from Thomson Datastream. Weekend days are not included in the dataset, i.e. only weekdays are considered. Public holidays are assigned the index value of the previous day, i.e. the index is assumed to have a zero return over public holidays. The index values that were used were the "Market values", according to the Datastream classification. These indices were chosen as they are both key indices in the markets that they are drawn from. They are thus very likely to be used as the underlying variable in options written on indices in those markets. The fact that these indices are from different markets, with the FTSE 100 based in a developed economy and the JSE index in an emerging market, should allow for a comparison to determine if the model behaves differently under these different conditions. Both indices were measured in their local currency to avoid any distorting effects arising from currency fluctuations.

In general, Hidden Markov Models perform better with larger datasets because the larger dataset is more likely to undergo more structural changes than a smaller dataset. Such a larger dataset allows for a more accurate estimation of the transition probabilities of the Markov Chain component of the model, which measure the likelihood of a switch to another hidden state from the hidden state in force at a given time point. The reason for this better estimation is because such a switch results in a structural change in the observed time series. This effect of producing better estimates is the main reason for selecting a fairly large dataset to estimate the parameters of the model from. In the case of these equity market indices, data from a longer period of time is more likely to include several bull and bear market episodes. This information will allow for a more accurate estimation of the probability of the market switching between these two main states.

### 10.1 A Basic Summary Analysis of the Data

Before the model was fitted to the data the log returns of both indices, being the logarithmic differences between successive index values, were calculated. Effectively, the HMM will be fitted to these log returns which are assumed to be normally distributed. Of course, the structure of the HMM allows for the

assumption that these log returns follow a dependent mixture of normal distributions, as was discussed in Chapter 4. The assumption of normality follows from the assumption that successive market transactions that lead to changes in the value of the index are independent of each other. An application of the central limit theorem then leads to the conclusion that log returns of stock indices are normally distributed. A discussion of this argument can be found in Fama (1976). It is perhaps worth noting that the assumption of independence between successive transactions follows from the well-known Efficient Markets Hypothesis. This hypothesis assumes that the prices of securities, and thus the value of any indices calculated from those securities prices, include all available information relevant to the value of those securities. Any new information relevant to the value of those securities that then enters the market and causes the price to change is then assumed to arrive in a random fashion. Further, the new information events are assumed to be independent, which in turn means that any price changes are independent of each other.

Thus, any basic analysis of the log return data that has been calculated should investigate whether it is in fact normally distributed. If not, then an attempt should be made to determine the possible nature of the components of a mixture of normal distributions that might be used to describe the data. An HMM could then be used to fit a similar mixture distribution, where the components of the mixture are the distributions associated with each possible hidden state value.

### 10.1.1 JSE All Share Index Data

Summary Statistics	
Mean	0.000456
Std Dev	0.011834
Skewness	-0.79827
Kurtosis	12.24689
JB	9568.532
p-value	0
Data points	2608

The above statistics clearly show that the log returns of the JSE All Share index are not normally distributed. They are negatively skewed, whereas the normal distribution is symmetric, and have a very high kurtosis whereas the normal distribution has a kurtosis of 3. This evidence for non-normality is confirmed by the high Jarque-Bera statistic (Jarque and Bera (1987), p166). This statistic follows a chi-squared distribution with two degrees of freedom and as the p-value shows, the null hypothesis of the normality of the distribution of the data can easily be rejected. It thus remains to determine what form of distribution this data could possibly follow. A histogram of the data will be used for this purpose.

As can be seen in the histogram below, the data appears to be almost symmetrically distributed around the daily mean return of approximately 0. The

negative skewness of the data thus appears to arise from some extremely low returns in the left tail of the observed distribution. The high kurtosis of the data is immediately apparent in the sharp peak of the histogram. Based on the brief discussion of mixture distributions in Chapter 4, it thus appears that a possible mixture of normal distributions that may fit this data would include distributions with different variances and different means. The distribution with the lower mean return would be associated with the higher variance of return, in keeping with the principle of the leverage effect. The number of component distributions in the mixture would depend on the number of states chosen for the HMM that is fitted to the log returns.

Note that the above discussion is intended as a starting point in the fitting of an HMM to the log return data. It is not meant to be an exact analysis of the data. The aim is to determine what combinations of means and variances might be used as starting values for the various hidden states in the model, in the algorithm that is used to estimate the parameters of the model. Further, the above indications, taken from the histogram below, also point to the combinations of parameters that could or should be expected once the HMM estimation is completed. A further expectation will be that at least one of the hidden states will have a negative mean rate of return associated with it, to account for bear market conditions.

Please note that the full dataset, together with the statistics presented above can be found on the CD-ROM submitted with this work, in the "Source data" folder, for both the JSE and FTSE 100 datasets.

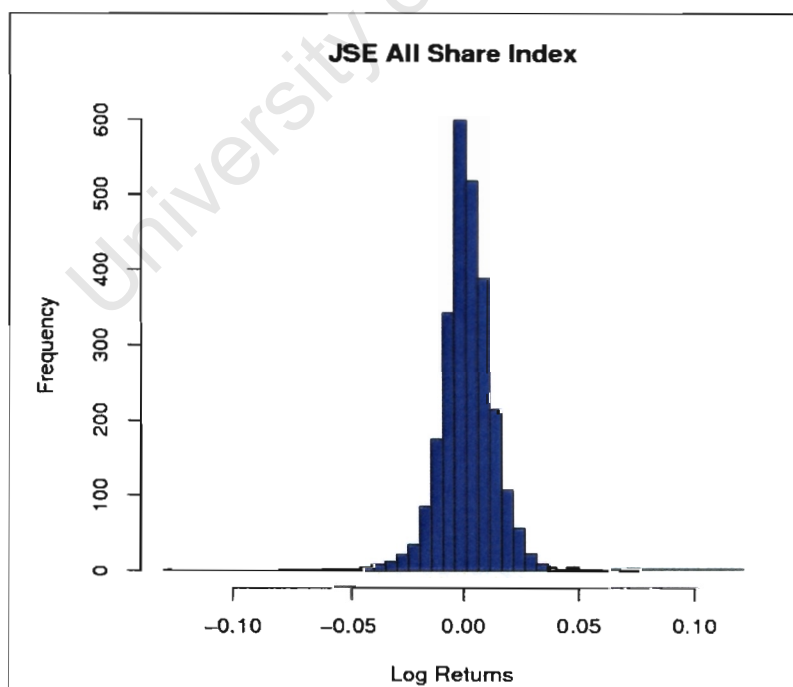


Figure 6: Histogram of the log returns of the JSE All Share Index

## 10.1.2 FTSE 100 Index Data

Summary Statistics	
Mean	0.000189
Std Dev	0.011162
Skewness	-0.18244
Kurtosis	5.873925
JB	911.9942
p-value	0
Data points	2608

Although both the skewness and kurtosis of the log returns of the FTSE 100 are lower than those of the JSE All Share Index, the data is clearly still not normally distributed. In this case, the log returns are also negatively skewed as was the case with the JSE data. The kurtosis of 5.87 is closer to 3 than that of the JSE, at 12.25, but is still higher than that of a normal distribution. The Jarque-Bera statistic is again large and has a very low p-value, virtually indistinguishable from 0. The data is thus not normally distributed, by this test. The fact that the FTSE 100 data is closer to being normally distributed than the JSE data may be explained, in part, by the possibility that the FTSE 100 is calculated from share prices set in a more efficient market than the JSE. This could result from the UK market having more participants, whose trading activity introduces more information into the market and thus moves it closer to being fully efficient than the JSE.

Turning to the possible distributional form of the FTSE 100 data, the histogram below indicates that the data is distributed around the daily mean return of approximately 0. The higher than normal kurtosis can be seen in the peaked form of the histogram. As in the case of the JSE data, the possible elements of a mixture of normal distributions will include distributions with differing means and variances. This will take account of the skewness and kurtosis of the FTSE 100 data. As before, the lower mean rates of return will be associated with higher standard deviations of return, to incorporate the leverage effect. It is also again expected that at least one of the hidden states will be associated with a negative mean rate of return, to account for bear market conditions.

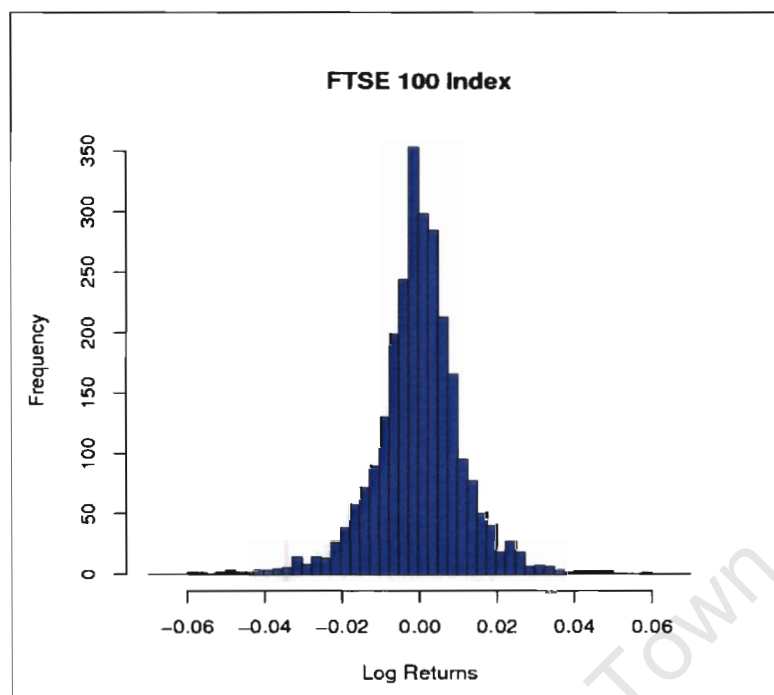


Figure 7: Histogram of the log returns of the FTSE 100 index

## 10.2 Estimating the Parameters of the HMM of the Log Returns Data

The method chosen for fitting an HMM to the log return data is that based upon the Baum-Welch algorithm, from Baum *et al.* (1970). It turns out that this, in turn, is a form of the Expectation-Maximisation (EM) algorithm of Dempster *et al.* (1977). Since the information about the hidden states is unavailable to the observer, it can be regarded as 'missing data'. The EM algorithm is designed for estimation of parameters for models, by maximum likelihood, where there is missing data. The data may be missing due to problems with data capture, or model choice as is the case here. The method is essentially to choose 'start-up' parameters for the model, find the expected value of the missing values according to the model using these values and then find the parameters that maximise the likelihood of the 'pseudo-complete' dataset, which includes all the data and expected values. These new parameters are then substituted into the model and the process is repeated until successive sets of estimates have converged, according to some convergence criterion. In the case of the estimation used here, the criterion chosen was that estimates should not differ by more than  $10^{-8}$ .

Since there are a large number of parameters to be estimated, including the means and variances of the normal distributions of the log returns associated with each hidden state, as well as the transition probabilities for the Markov Chain that describes the hidden state process, it is to be expected that the

likelihood function will have multiple local maxima. The EM algorithm only ensures that a local maximum will be attained when successive sets of parameter estimates converge. This is because successive estimates produced by the EM algorithm result in points on the likelihood surface that are higher than, or equal in height to, the points associated with the previous estimates (Dempster *et al.* (1977), p7). There is no guarantee that the global maximum has been attained. To try and find the global maximum, it is necessary to use several sets of 'start-up' values for the algorithm to try and find several of the local maxima. The value of the likelihood at these maxima can then be compared and the parameter set associated with the highest maximum will then be used for further computations.

### 10.2.1 The Updating Equations used for Estimating the Parameters and their Standard Errors

The first set of equations required for the estimation of the HMM parameters are those to perform the Expectation step of the EM algorithm. These will provide the expected values for the 'missing data' relating to which state the observed values are produced by. Since this data is merely whether or not a particular state was in force at a given time point, it is thus an indicator variable which will be 1 if the state in question was in force and 0 otherwise. The expected value of the data relating to the hidden state at a given time point will thus be the probability that the given state was in force at the given time point, conditional on the entire dataset. The equations used to determine these probabilities, for all time points in the dataset, will be those of the 'forward-backward' algorithm of Baum *et al.* (1970), p168. The specific form of these equations that will be used will be those with scaling, as in Rabiner (1989), to avoid numerical underflow in the computations, which could arise due to the successive multiplication of various probabilities. This scaling will also ensure that the probabilities calculated at each successive time step will add up to 1. A complete discussion of these methods can be found in Cappé *et al.* (2005), which is where the scheme outlined below was sourced.

#### Forward Filtering:

$$\Pr(X_0 = i | Y_{-1}) = \pi(i) = \Pr(X_0 = i)$$

$$\Pr(X_u = j | \mathfrak{S}_u) = \phi_u(j)$$

$$C_u = \sum_{i=1}^N \Pr(X_u = i | Y_{u-1}) f(Y_u | X_{u-1}) \quad (39)$$

$$\phi_u(j) = \frac{\Pr(X_u = j | Y_{u-1}) f(Y_u | X_u = j)}{C_u}$$

$$\Pr(X_{u+1} = j | \mathfrak{S}_u) = \sum_{i=1}^N \phi_u(i) p_{ij}$$

$$u = 1, 2, \dots, T$$

(Cappé *et al.* (2005), p122)

Notice that the above algorithm is initialised using the unconditional initial distribution of  $\{X_t\}$ . This can be exogenously specified, or estimated as part of the model. For a large dataset, as was used here, this should not make too much difference since the effect of the starting values in these computations should become very small as the calculations progress over the dataset. The choice made here was to parameterise the initial distribution of states as part of the model. Note as well that the above filtering algorithm can also be found in Hamilton (1989) and Hamilton (1990), among others. The  $C_u$  are the scaling constants and  $\Pr(X_{u+1} = j | \mathfrak{F}_u)$  is then the input for the next step in the filtering algorithm. As a useful by-product of the filtering algorithm, the value of the log-likelihood at each time point, for the given set of parameter estimates being used,

can be calculated. This is:  $\text{Log-Likelihood} = \sum_{u=1}^T \ln(C_u)$ , which is robust to both numerical underflow and overflow (Cappé *et al.* (2005), p122). This value will be compared, for different sets of starting values for the EM algorithm, to determine which set of estimates is associated with the largest local maximum.

#### Backward Smoothing:

$$\begin{aligned} \beta_{T|T}(j) &= C_T^{-1} \\ \Pr(X_T = j | \mathfrak{F}_T) &= \phi_{T|T}(j) = \phi_T(j) \\ \beta_{u|T}(i) &= C_u^{-1} \sum_{j=1}^N p_{ij} f(Y_{u+1} | X_{u+1} = j) \beta_{u+1|T}(j) \\ \phi_{u|T}(i) &= \frac{\phi_u(i) \beta_{u|T}(i)}{\sum_{j=1}^N \phi_u(j) \beta_{u|T}(j)} \end{aligned} \quad (40)$$

$u = T, (T-1), \dots, 1$  and  $i, j = 1, 2, \dots, N$

(Cappé *et al.* (2005), p123)

This smoothing algorithm can also be found, albeit in vector form, in Hamilton (1994). Due to the inputs required for the smoothing algorithm, it is recommended that the scaling constants and the conditional density values  $f(Y_u | X_u)$ , as well as the filtered probabilities  $\phi_u(j)$ , are stored during the filtering algorithm. This will improve the efficiency of the above smoothing algorithm (Cappé *et al.* (2005), p123).

Some further values that will be required for estimating the transition probabilities, at each step in the EM algorithm, are  $\phi_{u,u+1|T}(i, j) = \Pr(X_{u+1} = j | X_u = i, \mathfrak{F}_T)$ . These can be computed using the following:

$$B_u(j, i) = \frac{\phi_u(i) p_{ij}}{\sum_{m=1}^N \phi_u(m) p_{mj}} \quad (41)$$

$$\phi_{u:u+1}(i, j) = \phi_{u+1|T}(j) B_u(j, i)$$

$$u = T, (T-1), \dots, 1 \text{ and } i, j = 1, 2, \dots, N$$

where the  $B_u(j, i)$  are the backwards transition probabilities:  $\Pr(X_u = i | X_{u+1} = j)$ . It is interesting to note that due to the construction of the  $\phi_{u:u+1}(i, j)$ , an alternative method of calculating the smoothed probabilities, as above, arises:  $\phi_{u|T}(i) = \sum_{j=1}^N \phi_{u:u+1}(i, j)$ , which can be used to check the smoothed probabilities computed in the smoothing algorithm above, if required.

Having performed the 'E-step' of the EM algorithm by computing the smoothed probabilities of being in a given state at a given time point, it remains to perform the maximisation step to produce the updated set of parameters.

#### Maximisation of the 'pseudo-complete' data:

The log-likelihood of the 'pseudo-complete' dataset, including the expected value substitutes for the 'missing data', can be written as follows:

$$Q(\theta; \theta') = Cst - \frac{1}{2} \sum_{u=1}^T \sum_{i=1}^N \phi_{u|T}(i; \theta') \left[ \log(\sigma_i^2) + \left( \frac{Y_u - \mu_i}{\sigma_i^2} \right)^2 \right] + \sum_{v=2}^T \sum_{i=1}^N \sum_{j=1}^N \phi_{v:v+1}(i, j; \theta') \log(p_{ij})$$

where  $\theta'$  is the vector of parameter estimates from the previous iteration of the EM algorithm and  $\theta$  is the vector of parameters to be estimated from the current step of the algorithm.  $\theta$  contains the mean and variance pairs, associated with each state, as well as the transition probabilities of the Markov Chain that describes the hidden state process and the unconditional distribution of the initial state values, in this case (Cappé *et al.* (2005), p368).

Furthermore, since  $\sum_{j=1}^N p_{ij} = 1$ , it is necessary to add this constraint to the log-likelihood function above, to produce the Lagrangian:

$$L(\theta; \lambda; \theta') = Q(\theta; \theta') + \sum_{i=1}^N \lambda_i \left( 1 - \sum_{j=1}^N p_{ij} \right) \quad (i, j = 1, \dots, N)$$

where the  $\lambda_i$  are Lagrange multipliers

(Cappé *et al.* (2005), p368).

To maximise this log-likelihood function, under the given constraint, it is necessary to calculate the elements of the gradient vector, or score, of the Lagrangian above. These are found by differentiation and are as follows:

$$\begin{aligned}\frac{\partial L}{\partial \mu_i} &= \frac{1}{\sigma_i^2} \sum_{u=1}^T \phi_{u|T}(i; \theta') (Y_u - \mu_i) \\ \frac{\partial L}{\partial \sigma_i^2} &= -\frac{1}{2} \sum_{u=1}^T \phi_{u|T}(i; \theta') \left[ \frac{1}{\sigma_i^2} - \frac{(Y_u - \mu_i)^2}{(\sigma_i^2)^2} \right] \\ \frac{\partial L}{\partial p_{ij}} &= \sum_{u=2}^T \frac{\phi_{u,u+1}(i, j; \theta')}{p_{ij}} - \lambda_i \quad i, j = 1, \dots, N \\ \frac{\partial L}{\partial \lambda_i} &= (1 - \sum_{j=1}^N p_{ij})\end{aligned}$$

The constrained log-likelihood achieves a maximum when these derivatives are set to 0, so performing this calculation will result in the updated parameters for the HMM, for this iteration of the EM algorithm. The required updating equations are then:

$$\begin{aligned}\mu_i^* &= \frac{\sum_{u=1}^T \phi_{u|T}(i; \theta') Y_u}{\sum_{u=1}^T \phi_{u|T}(i; \theta')} \\ \sigma_i^{2*} &= \frac{\sum_{u=1}^T \phi_{u|T}(i; \theta') (Y_u - \mu_i^*)^2}{\sum_{u=1}^T \phi_{u|T}(i; \theta')} \\ p_{ij}^* &= \frac{\sum_{u=2}^T \phi_{u,u+1|T}(i, j; \theta')}{\sum_{u=1}^{T-1} \phi_{u|T}(i; \theta')} \\ \pi^*(i) &= \phi_{1|T}(i; \theta')\end{aligned} \tag{42}$$

$$i, j = 1, \dots, N$$

(Cappé *et al.* (2005), pp. 369-370)

The new estimates computed using (36) are then used as inputs for the next step of the EM algorithm. The smoothed probabilities are then calculated using (33), (34) and (35), and the new estimates for the parameters are found using (36). The algorithm then repeats until convergence is achieved.

### Standard Errors of the Estimates:

Computing the standard errors of the estimated values requires computing the Observed Information Matrix:  $I(\theta)$ . This is  $-\mathcal{H}$ , where  $\mathcal{H}$  is the Hessian of the constrained log-likelihood function, i.e. the matrix of second derivatives of the log-likelihood with respect to the parameters of the model. The covariance matrix of the parameters is then  $I^{-1}(\theta)$ . The standard errors are then the square roots of diagonal elements of this matrix. The EM algorithm does not automatically calculate  $I(\theta)$  and so this needs to be done separately. This set of calculations will be performed after the algorithm has converged. The method that will be used here to compute  $I(\theta)$  will be that proposed by Louis (1982), who suggests computing  $I(\theta)$  from the Hessian of the log-likelihood of the 'pseudo-complete' dataset.

This requires the calculation of the second derivatives of the constrained log-likelihood with respect to the various parameters of the model. These second derivatives are:

$$\frac{\partial^2 L}{\partial \mu_i^2} = -\frac{1}{\sigma_i^2} \sum_{u=1}^T \phi_{u|T}(i; \theta')$$

$$\frac{\partial^2 L}{\partial (\sigma_i^2)^2} = -\frac{1}{2} \sum_{u=1}^T \phi_{u|T}(i; \theta') \left[ \frac{-1}{(\sigma_i^2)^2} + \frac{2(Y_u - \mu_i^*)^2}{(\sigma_i^2)^3} \right]$$

$$\frac{\partial^2 L}{\partial p_{ij}^2} = \sum_{u=2}^T -\frac{\phi_{uu+1|T}(i, j; \theta')}{(p_{ij})^2}$$

$$\frac{\partial^2 L}{\partial \lambda_i^2} = 0$$

$$\frac{\partial^2 L}{\partial \mu_i \partial \mu_j} = 0$$

$$\frac{\partial^2 L}{\partial \mu_i \partial \sigma_j^2} = -\frac{1}{(\sigma_j^2)^2} \sum_{u=1}^T \phi_{u|T}(i; \theta') (Y_u - \mu_i^*)$$

$$\frac{\partial^2 L}{\partial p_{ij} \partial \lambda_i} = -1$$

$$\frac{\partial^2 L}{\partial \lambda_i \partial p_{ij}} = -N$$

$$\frac{\partial^2 L}{\partial \sigma_i^2 \partial \lambda_i} = \frac{\partial^2 L}{\partial \lambda_i \partial \sigma_j^2} = \frac{\partial^2 L}{\partial \mu_i \partial \lambda_i} = \frac{\partial^2 L}{\partial \lambda_i \partial \mu_j} = \frac{\partial^2 L}{\partial p_{ij} \partial \mu_j} = \frac{\partial^2 L}{\partial \mu_i \partial p_{ij}} = \frac{\partial^2 L}{\partial \sigma_i^2 \partial p_{ij}} = \frac{\partial^2 L}{\partial p_{ij} \partial \sigma_j^2} = 0$$

These values are then placed into the Hessian, with the subscript  $i$  ranging over the rows of the matrix and the subscript  $j$  ranging over the columns. The Hessian

is then multiplied by -1 and inverted to produce the covariance matrix of the estimated parameters. It may be the case, especially where the dataset is large and there are a large number of parameters to be estimated, that the Observed Information Matrix is singular. This means that it will no longer be possible to find the covariance matrix of the estimators. In such a case, standard errors for the estimates will not be available. Since a large dataset is being used here, this situation could conceivably arise, especially for models with several hidden state variables.

### 10.2.2 Selecting the Number of States and the Start-up Parameters

As was discussed in Chapter 4, there appears to be no agreement in the literature on how to select the number of hidden states to use in an HMM. Thus, it is necessary to include this choice as part of the specific model specification that is being used in a given situation. For the purposes of this work, the aim would be to select the number of states so that the model provides a good description of the data whilst avoiding model overfit, but still retaining an economic justification for the number of states chosen. It is thus proposed that models with differing numbers of states would then be fitted and compared, on the basis of some criterion. In this case, Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) will be used. The model with the lowest AIC and/or BIC value will then be selected for further use.

There is, however, a complicating factor relating to the number of states used in the HMM. To calculate the prices of options written on the variable that is modelled by the HMM, it is necessary to calculate the occupation times and the probabilities for the joint occupation times, as in Chapter 5, using (23), (24) and (25). The joint occupation times, for the first  $(N-1)$  states, are a permutation with replacement. If there are  $j$  periods to the expiry of the option, there are then  $j^{(N-1)}$  possible sets of joint occupation times and thus joint occupation time probabilities. The  $N$ -th state occupation time is, of course, uniquely determined by the first  $(N-1)$  states and so does not need to be included in the analysis. In the case of an option with  $\tau$  as the time to expiry,  $j = (\tau+1)$  since  $j$  includes the occupation time of 0 as well.

The most convenient way to store these occupation times and occupation time probabilities is in  $(\tau+1) \times N$  matrices. If the occupation time of the first state is allowed to vary from 0 to  $\tau$  in each matrix, and the  $N$ -th state occupation time is determined by the first  $(N-1)$  states, then there are  $(N-2)$  states for which occupation times must still be specified, to calculate all the possible joint occupation times and joint occupation time probabilities. There are thus  $(\tau+1)^{(N-2)}$  matrices of joint occupation times. Now the complication is clear. If  $N$  is greater than 3, then the number of matrices required becomes exponential in  $(\tau+1)$ . Since daily data is being used here, and thus  $\tau$  will be measured in days, it follows that the calculations will only remain feasible for small  $\tau$  (5 days or less)

when  $N$  is greater than 3. This is unreasonable since it is unlikely that the time series of the variable underlying the option contract will undergo structural changes over such a short time period. There would thus be little justification in using a more complicated model, such as the Hidden Markov Option Pricing Model, to take account of any possible structural changes in the time series of the underlying variable.

It is therefore suggested that models including only 2 or 3 hidden states be used when fitting an HMM that is to be used for pricing options, as will be done here. This is a somewhat arbitrary constraint, introduced by the structure of the option pricing component of the model. It does, however, act as a guard against model overfit by keeping the number of parameters in the model reasonably low. It is also possible to provide an economic justification for 2 and 3 states, which was suggested as one of the criteria that should be taken into account when selecting the number of hidden states. In the case of equity market indices, as used here, 2 hidden states could represent 'good' market conditions and 'bad' market conditions, i.e. a bull market state and a bear market state, respectively. When the model has 3 hidden states then these could represent a bull market state, a 'random walk' state that represents market conditions that are neither bull market nor bear market, and a bear market state. This will be the interpretation used here. This will, in turn, influence the start up values that will be used and the values that the final estimates will be expected to take.

#### Sample Start-up Values for the JSE All Share Index:

State 1	State 2	State 3
$\mu = 0.0008, \sigma = 0.009$	$\mu = 0.0003, \sigma = 0.002$	$\mu = -0.0005, \sigma = 0.014$
$\mu = 0.001, \sigma = 0.008$	$\mu = 0.0005, \sigma = 0.001$	$\mu = -0.0002, \sigma = 0.011$
$\mu = 0.0008, \sigma = 0.007$	$\mu = 0.004, \sigma = 0.0009$	$\mu = -0.0002, \sigma = 0.012$
$\mu = 0.0006, \sigma = 0.01$	$\mu = 0.0002, \sigma = 0.004$	$\mu = -0.0001, \sigma = 0.0105$

These values were chosen so as to remain reasonably close to the overall mean and standard deviation of the dataset, as presented in 10.1, whilst keeping the interpretation of the states outlined above. Further, states with a lower return associated with them should have a higher standard deviation associated with them, to satisfy the leverage effect. In the model with only 2 hidden states, the start-up values for State 3 were used for State 2, as these are the bear market state start-up values.

#### Sample Start-up Values for the FTSE 100 Index:

State 1	State 2	State 3
$\mu = 0.0007, \sigma = 0.001$	$\mu = 0.0003, \sigma = 0.004$	$\mu = -0.0001, \sigma = 0.012$
$\mu = 0.0004, \sigma = 0.009$	$\mu = 0.0001, \sigma = 0.006$	$\mu = -0.0005, \sigma = 0.004$
$\mu = 0.0003, \sigma = 0.01$	$\mu = 0.0005, \sigma = 0.007$	$\mu = -0.002, \sigma = 0.03$

In both cases, the start-up values for the initial distribution of states assumed that each state was equally likely. Thus under the 2-state model each state had an initial probability of  $\frac{1}{2}$  and under the 3-state model each state had an initial probability of  $\frac{1}{3}$ .

The start-up values for the transition probabilities of the Markov Chain component of the model should be chosen so that the probability of remaining in a given state is greater than 0.5. This is because other studies, such as Ishijima and Kihara (2005), have found the states to be very persistent. These probabilities also need to add up to 1, for each state. This means that only  $(N-1)$  transition probabilities need to be estimated for each state. Thus, for the 2-state model only two transition probabilities need to be directly estimated and for the 3-state model only 6. In this case  $p_{11}$  and  $p_{22}$  were estimated for the 2-state model and  $p_{11}$ ,  $p_{12}$ ,  $p_{21}$ ,  $p_{22}$ ,  $p_{31}$  and  $p_{32}$  were directly estimated for the 3-state model.

### 10.2.3 Results of Fitting the HMM to the Log Return Data

Estimated Values for the 2-State Model of the JSE Data:

	Estimated Values	Standard Errors		
$\mu_1$	0.2879011	0.000179	AIC	-16362.07
$\mu_2$	-0.1838494	0.000735	BIC	-5850.579
$\sigma_1$	0.1506186	2.00E-06	Interest rate: State 1	0.07
$\sigma_2$	0.3641046	1.98E-05	Interest rate: State 2	0.1
			Expected Duration: 1	41.209023
			Expected Duration: 2	14.561814

Transition Probabilities for the Markov Chain

To From	State 1	Std Error	State 2	Std Error
State 1	0.9757335	0.006688	0.024267	0.006688
State 2	0.0686728	0.006965	0.931327	0.006965

Note that the mean rates of return and the standard deviations of return are reported as annualised values. This is because that is the form of the values required for the option pricing formula. The reasonably low standard error values suggest that all of these parameter estimates are probably statistically significant. The values that were estimated for  $\mu_1$ ,  $\mu_2$ ,  $\sigma_1$  and  $\sigma_2$  were as expected, with State 1 representing bull market conditions and State 2 representing bear market conditions. The leverage effect has been incorporated, since the state with the lowest mean rate of return associated with it also has the highest standard deviation of return associated with it. Notice also that the states are very persistent with the probability of remaining in a given state being greater than 0.9.

The AIC was calculated as  $AIC = (-2 \times \log\text{-likelihood}) + (2 \times \eta)$ , where  $\eta$  is the number of independently adjusted parameters in the model (Akaike (1974), 716).

The BIC, on the other hand, is  $BIC = -\log\text{-likelihood} + (\frac{1}{2} \times \log(n) \times \eta)$  where  $n$  is the number of values in the sample (Schwarz (1978), p461). In both the AIC and the BIC calculations, the log-likelihood value used is that of the log-likelihood at the maximum where the parameters were selected. For this 2-state model, the number of independently adjusted parameters is 6.

The interest rates chosen for each state were selected after reviewing the yield available from 91-Day Treasury Bills over the period of the sample. These slightly arbitrary figures represent a rough 'average high' and 'average low' level over the period in question. The higher interest rate was assigned to the lower mean rate of return state, since equities (and thus equity indices) achieve a higher rate of return under conditions of lower interest rates. These interest rates will be used as the risk-free rate of return input required for option pricing. Short-term government debt, in the local currency is the closest to being risk-free, in a given economy. This is why the yield on Treasury Bills was used to determine these interest rates.

The expected duration in each state is measured in days and is calculated as:

Duration  $i = \frac{1}{(1 - p_{ii})}$ . It is interesting to note that a bull market phase is expected

to last almost three times as long as a bear market phase. This is unlikely to be an artefact of this dataset alone, since it includes the bear market induced by the Emerging Markets Crisis of 1997-1998 as well as that resulting from the bursting of the tech-stock bubble in 2000. Both of these bear market phases were severe and considered prolonged. For example, the bear market that began in 2000 only abated completely in 2003. In the long run, it would thus appear that investing in equity markets is likely to be beneficial.

To give an indication of the state dynamics over the dataset, graphs of the smoothed probability of being in a given state are provided below.

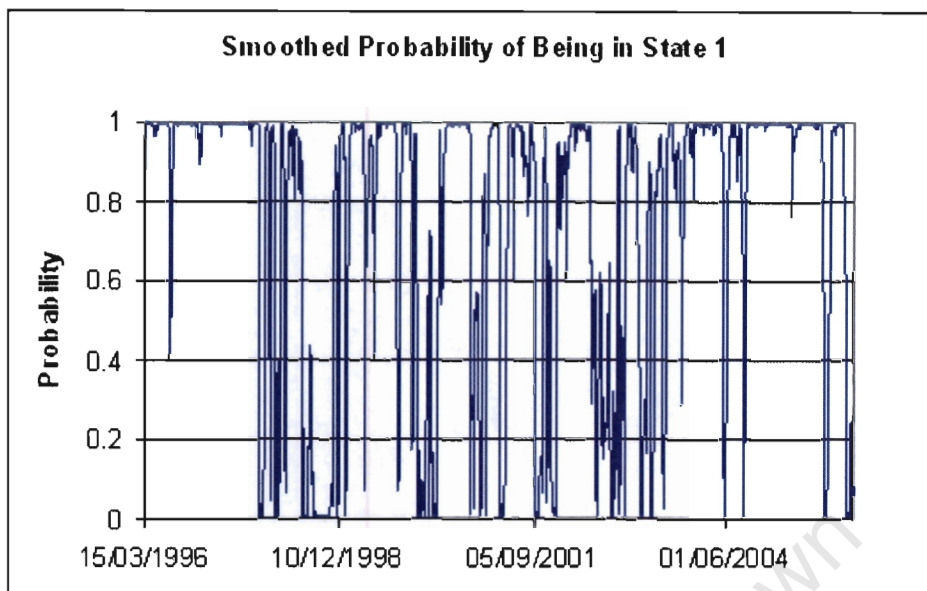


Figure 8: Smoothed Probability of being in State 1 in the 2-state model of the JSE Data

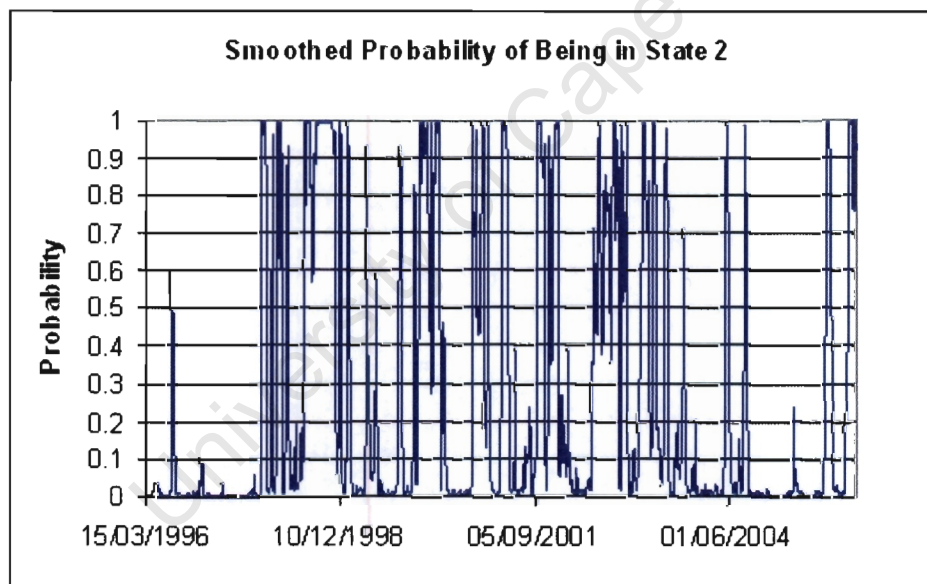


Figure 9: Smoothed Probability of being in State 2 in the 2-state model of the JSE Data

Estimated Values for the 3-State Model of the JSE Data:

Estimated Values		AIC	-16504.85
$\mu_1$	0.29128021	BIC	-8217.225
$\mu_2$	0.27912532	Interest rate: State 1	0.07
$\mu_3$	-1.7043765	Interest rate: State 2	0.1
$\sigma_1$	0.12735316	Interest rate: State 3	0.13
$\sigma_2$	0.22329181	Expected Duration: 1	51.951326
$\sigma_3$	0.55170675	Expected Duration: 2	40.89026
		Expected Duration: 3	10.921308

## Transition Probabilities for the Markov Chain

To / From	State 1	State 2	State 3
State 1	0.980751	0.017087	0.002162
State 2	0.015544	0.975544	0.008911
State 3	2.84E-25	0.091564	0.908436

Note that the Observed Information Matrix was singular and thus standard errors of the estimates are not available here. All values are again presented as annualised rates. The mean rates of return do appear to conform to the model structure expected, with State 1 having the highest mean rate of return and State 3 the lowest. The lower levels of mean rates of return are also associated with higher standard deviations of returns, thus incorporating the leverage effect. It is interesting to note that the mean returns of State 1 and State 2 are relatively close to each other, whilst the standard deviations of return differ by a far larger margin. The mean rate of return is far worse in State 3 than in State 2 for the 2-state model. It thus appears that this version of the model estimates the effect of the bear market to be much more severe than the 2-state model. The states are again all very persistent, with the probability of remaining in a given state estimated at above 0.9 for all states. It is also interesting that the probability of moving from State 3 to State 1 is extremely low. This makes economic sense, since it is unlikely that the market will recover immediately from a deep bear market to a full bull market without moving through a 'recovery' phase first.

The AIC and the BIC of this model are both significantly lower than for the 2-state model, suggesting that this model provides a better description of the JSE data than the 2-state model. Given that the dataset includes two severe bear markets, this seems reasonable, since the bear market state for this model has a much lower mean rate of return than is the case in the 2-state model.

The interest rates were again chosen in a fairly ad hoc fashion, based on a simple overview of the yield on 91-Day Treasury Bills over the given period. The lowest interest rate was again matched to the state with the highest mean rate of

return, whilst the highest interest rate was associated with the state of the lowest mean rate of return. The highest possible interest rate is higher for this model, to reflect the more severe conditions of the bear market state. The expected duration in each state, in days, is again shown. The bear market state, although very severe, again has a much lower expected duration than the other two states with positive mean rates of return.

Below are graphs of the smoothed probabilities of being in each state, at a given time, to give an idea of the state dynamics under this model.

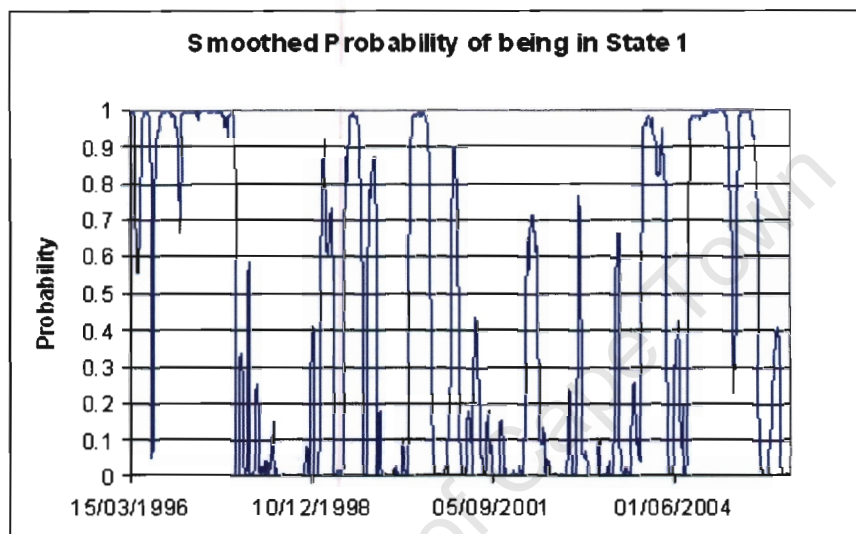


Figure 10: Smoothed Probability of being in State 1 in the 3-state model of the JSE Data

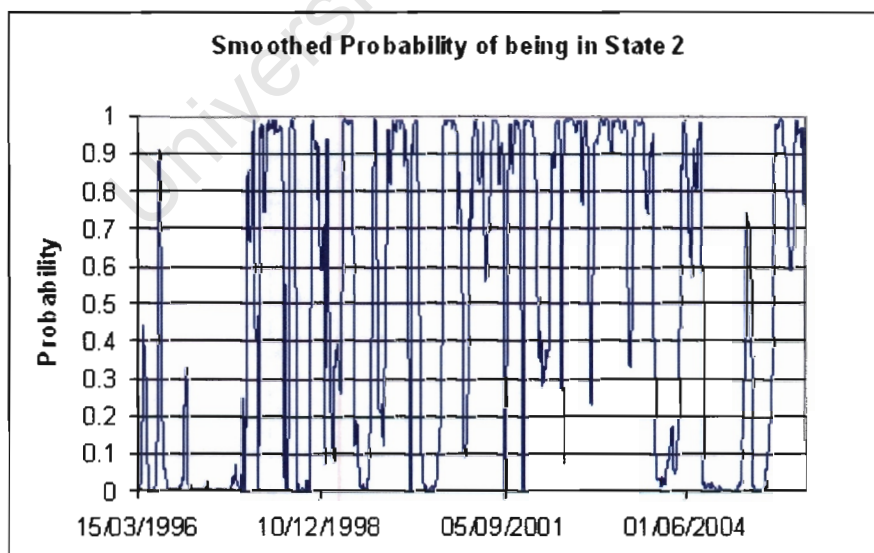


Figure 11: Smoothed Probability of being in State 2 in the 3-state model of the JSE Data

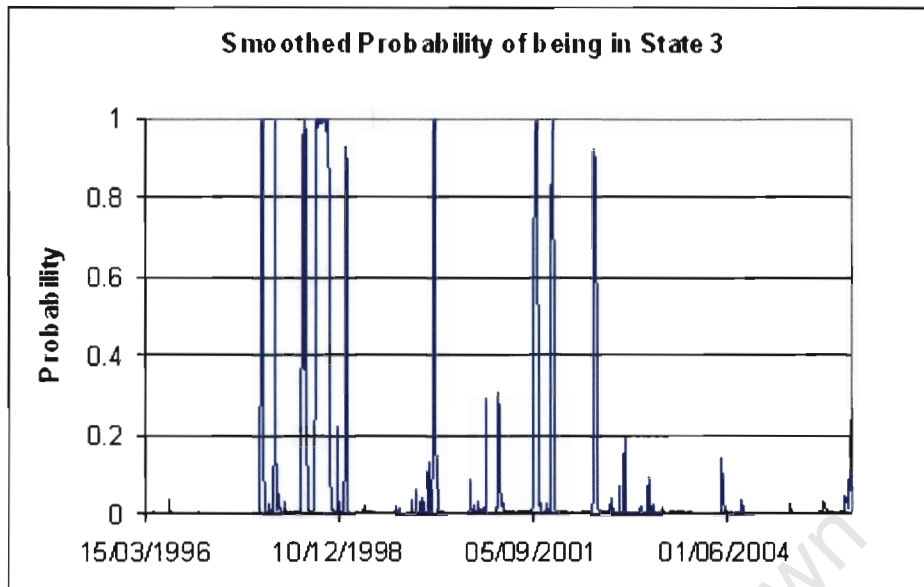


Figure 12: Smoothed Probability of being in State 3 in the 3-state model of the JSE Data

Estimated Values for the 2-State Model of the FTSE 100 Data:

	Estimated Values	Standard Errors		
$\mu_1$	0.204703	0.000171	AIC	-16744.82
$\mu_2$	-0.138812	0.000486	BIC	-8354.813
$\sigma_1$	0.1299058	1.65E-06	Interest Rate: State 1	0.04
$\sigma_2$	0.2983325	1.07E-05	Interest Rate: State 2	0.05
			Expected Duration: 1	94.231489
			Expected Duration: 2	61.65584

## Transition Probabilities for the Markov Chain

To \ From	State 1	Std Error	State 2	Std Error
State 1	0.9893878	0.003843	0.010612	0.003843
State 2	0.0162191	0.003773	0.983781	0.003773

As was the case with the JSE data, the estimates appear to account for both bull and bear market conditions, as well as incorporating the leverage effect by associating the higher standard deviation of return with the lower mean rate of return. The standard errors of the estimates are again relatively low, which indicates that these estimates are probably all statistically significant. The states are also again very persistent, with the probability of remaining in a given state being more than 0.98. This is higher than the probability of remaining in a given state for the 2-state model of the JSE data. This, together with the fact that the expected duration in each state is also longer than in the case of the JSE data, suggests that the FTSE 100 index does not switch between states as often as the JSE does. This is not surprising, since the FTSE is based in a developed

economy and is therefore less risky than the JSE. It will thus exhibit less volatility than the JSE, which would include less structural change in the time series than the JSE.

The AIC and BIC are both lower for this model than for the 2-state JSE model, indicating that the 2-state model fits the FTSE 100 data better than the JSE data. This is because the number of independently adjusted parameters and the sample size are the same for both 2-state models.

The interest rates for the states were again chosen by examining the yield on Treasury Bills, although in this case these were 1-month UK T-Bills. The lower maturity was chosen simply because it was available. In the case of the South African data, the lowest maturity T-Bill available was the 91-Day bill. As in the previous models, the lower rate of interest was associated with the higher mean rate of return state. Notice, however, that the level of interest rates is lower in the UK than in South Africa, as are the mean rates of return on the index. This is probably due to the lower risk associated with the developed UK economy, as compared to the South African economy. This lower risk will lead to lower rates of return being demanded by investors in that economy, as is evidenced by the lower interest rates and mean rates of return on the FTSE 100 index.

The graphs of the smoothed probabilities are again presented, to give an idea of the state dynamics. Notice that the conclusions about the fact that fewer state changes occur for this data than for the JSE data are confirmed by the fact that these series are less variable than those for the JSE 2 state model.

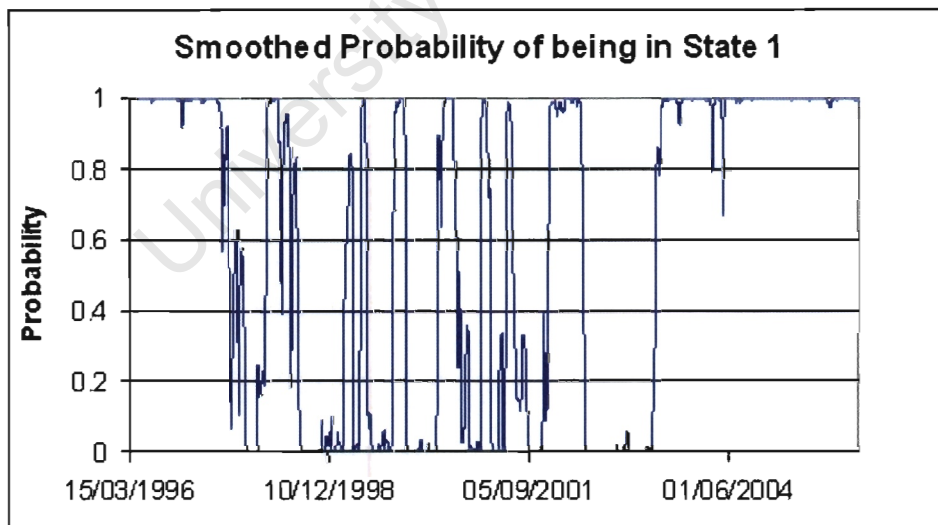


Figure 13: Smoothed Probability of being in State 1 in the 2-state model of the FTSE 100 data

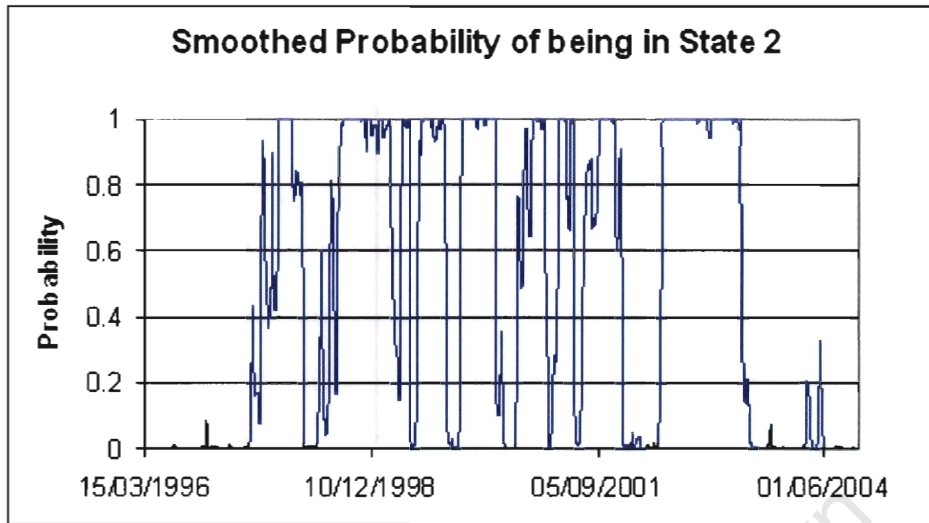


Figure 14: Smoothed Probability of being in State 2 in the 2-state model of the FTSE 100 data

#### Estimated Values for the 3-State Model of the FTSE 100 Data:

	Estimated Values
$\mu_1$	0.2404335
$\mu_2$	0.0399867
$\mu_3$	-0.441779
$\sigma_1$	0.1133451
$\sigma_2$	0.2081526
$\sigma_3$	0.4173382

AIC	-16895.94
BIC	-8412.774
Interest rate: State 1	0.04
Interest rate: State 2	0.05
Interest rate: State 3	0.06
Expected Duration: 1	129.6504
Expected Duration: 2	77.110887
Expected Duration: 3	37.077924

#### Transition Probabilities for the Markov Chain

From \ To	State 1	State 2	State 3
State 1	0.992287	0.007713	0
State 2	0.006653	0.987032	0.006315
State 3	9.44E-57	0.02697	0.97303

As was the case for the 3-state model of the JSE data, the Observed Information Matrix was singular for this model and so standard errors for these estimates are again not available. The mean returns seem to fit the rationale of the 3-state model very well, with a high mean return that would correspond to a bull market state, a low rate of return that could be associated with a 'random walk' type state and a negative mean rate of return that would be associated with a bear market state. This pattern, which was expected due to economic considerations, is much more apparent for the FTSE 100 data than for the JSE data. A possible

reason for this result could be that the HMM appears to describe the FTSE 100 data better than the JSE data, as can be seen by the fact that both the AIC and BIC values are lower for this model than for the 3-state model of the JSE data. The AIC and BIC values are also lower than those for the 2-state model of the FTSE 100 returns, suggesting that the 3-state model provides a better description of the data, as was found in the case of the JSE data.

The leverage effect is also present, with the lower mean rates of return being associated with higher standard deviations of return. It is also interesting to note that the change in the estimated mean rates of return and standard deviations of return between the 2-state and 3-state models was not as extreme as for the JSE data. The highest mean rate of return is slightly higher than in the 2-state model, at 24.04% compared to 20.47%, whilst for the JSE data it was 29.13% as compared to 28.79%. The real difference arises when comparing the mean rates of return for the lowest return state, between the 2-state and the 3-state models. In the case of the FTSE 100 data this was -13.88% compared to -44.18%, whilst for the JSE data this was -18.38% compared to -170.44% (!). This seems to indicate that the JSE experiences more extreme negative events than the London market, where the FTSE 100 is taken from. This makes economic sense, since the London market is based in a developed economy and is therefore less risky than the JSE, which is based in an Emerging Market. The more extended parameterisation of the 3-state model appears to capture this information better than the 2-state model. This further confirms that the 3-state model better describes the dataset for both the 3-state and 2-state model.

As was found in the case of the 2-state models, the FTSE 100 data does not transit between states as much as the JSE data does. Evidence for this conclusion can be found in the fact that the probability of remaining in a given state is higher in the case of the FTSE 100 data than for the JSE data, at over 0.97 for each state. The expected duration in each state is also greater for the FTSE 100 data than for the JSE data, which confirms this conclusion. As was the case for the JSE data, the expected duration in each state is also greater for the 3-state model of the FTSE 100 data than for the 2-state model. This may be due to the fact that the 3-state model describes the data better than the 2-state model and so does not need to switch between states as often, to explain the features of the data.

The probability of moving from State 3 to State 1 is also very low, as was observed for the 3-state model of the JSE data. Again, this is probably due to the fact that it is unlikely that the market will move directly from a bear market phase to a bull market phase, without first entering a recovery phase. Another interesting feature here is that the probability of moving from State 1 to State 3 is 0. Thus it is not possible to move directly from a bull market phase to a bear market phase, according to this model. This means that reversals of fortune in this market will not be as sudden as on the JSE, where a switch from a bull market phase directly to a bear market phase is possible. In the case of the

London market, returns first reduce to lower average levels before becoming negative in a bear market phase. This is further evidence for the London market being less risky than the JSE.

Finally, the graphs of the smoothed probabilities of being in a given state, at a given time point, are presented for this model. They confirm the conclusion that the FTSE 100 data transits between states less often than the JSE data, by being less variable than the graphs for the 3-state model of the JSE data.

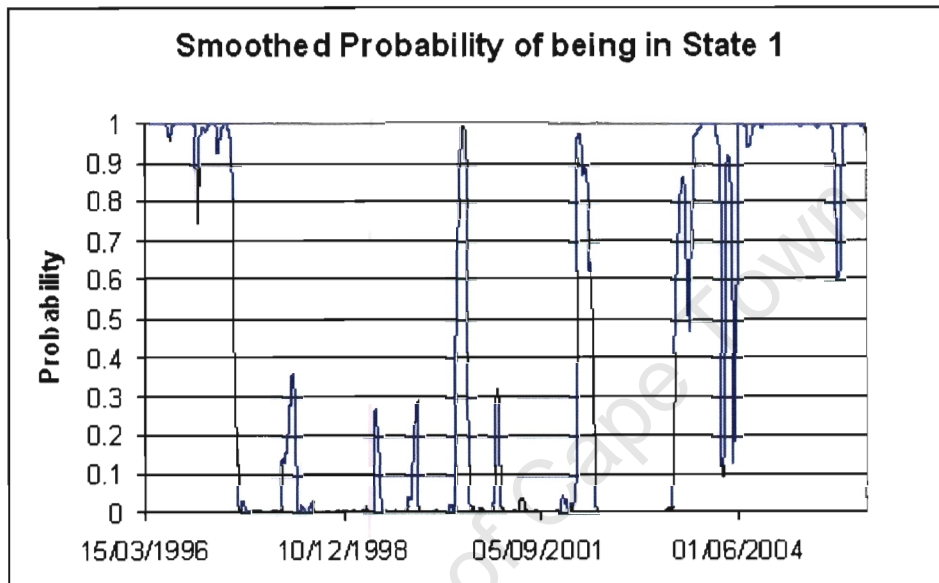


Figure 15: Smoothed Probability of being in State 1 in the 3-state model of the FTSE 100 data

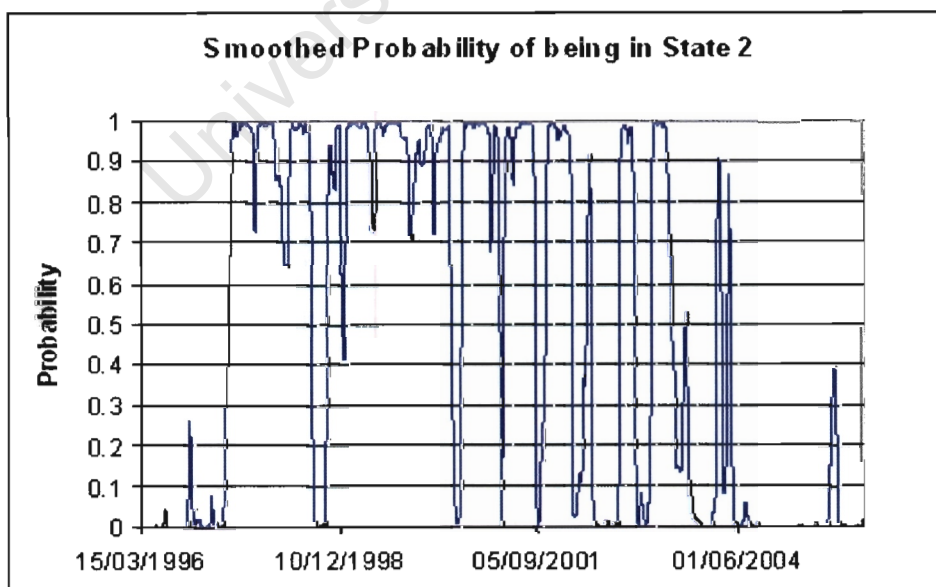


Figure 16: Smoothed Probability of being in State 2 in the 3-state model of the FTSE 100 data

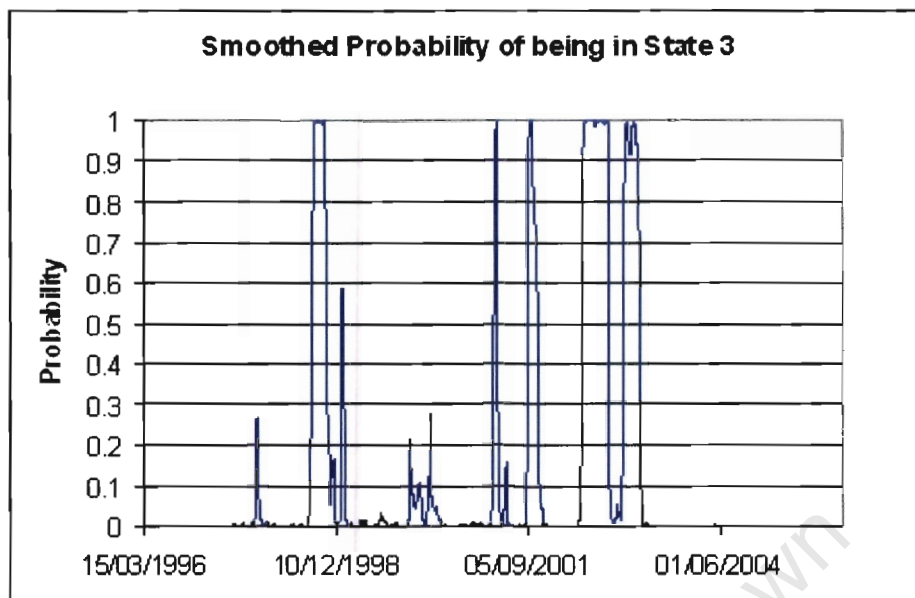


Figure 17: Smoothed Probability of being in State 3 in the 3-state model of the FTSE 100 data

## Chapter 11. Pricing Options on the JSE All Share and FTSE 100 indices using the fitted HMM Parameters

The options on the JSE All Share and FTSE 100 indices that will be priced will have maturities of 30, 60 and 91 days. These were chosen so as to allow for a spread of maturities, whilst still keeping the computations manageable. In any case, most options have reasonably short maturities and so working with a maximum maturity of 3 months is thus not excessively restrictive. Option prices will be computed for both the 2-state and 3-state models of both indices. Several different combinations of initial index values and strike prices for the options to be priced will be used, to determine if the Hidden Markov Option Pricing Model incorporates the volatility smile effect. This will be tested by comparing graphs of the implied volatilities of options of different strike and initial value combinations, for the various maturities. For each maturity, the implied volatility curve should decrease from out-the-money levels of strike and initial index value, to in-the-money combinations.

Prices for put and call options on the indices will be calculated as in Chapter 6, using (19) and (20). The inputs to these formulae will be those presented in the previous chapter. Note that it was assumed that either no dividends were paid by the indices, or alternatively that all dividends paid were immediately reinvested in the index. This would result in all returns from the index being of a capital nature. The net effect is the assumption that no dividend adjustment is needed in the option pricing formula. Having made this assumption, the following results were obtained.

### 11.1 Option Prices Computed using the 2-state model of the JSE Data

#### Call Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
Initial State is State 1	0.148	2.5753	10.286	2.9768	6.7084	11.478	1.2533	5.2587	12.3385
B-S Prices: State 1	0.0147	2.0197	10.094	0.1551	3.0368	10.7676	0.4226	3.9183	11.4924
Initial State is State 2	0.9088	4.5225	11.284	5.5304	9.1229	13.3232	3.6072	8.3711	14.8674
B-S Prices: State 2	0.9354	4.5676	11.317	2.3305	6.6873	13.221	3.6745	8.4519	14.9416

### Put Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
Initial State is State 1	9.5301	1.9573	0.16817	3.7262	1.4578	0.5274	9.3521	3.3576	0.9373
B-S Prices: State 1	9.441	1.446	0.01992	9.011	1.8927	0.1235	8.6925	2.1883	0.2623
Initial State is State 2	10.096	3.7099	0.97179	5.9117	3.5042	2.0046	11.163	5.9265	2.9228
B-S Prices: State 2	10.117	3.749	0.99857	10.7	5.0569	2.0906	11.212	5.9895	2.9793

The Black-Scholes prices that were calculated for comparison with the prices generated by the Hidden Markov Option Pricing Model assume that the volatility and risk-free rate remain constant over the life of the options, at the values associated with each initial state. They thus form a base case to compare with the HMM prices, which allow these quantities to vary over the life of the option.

As can be seen from the above tables, the model prices where the initial state was State 1 were mostly greater than the Black-Scholes prices for both call and put options. The prices for the case when the initial state was State 2 were mostly lower than the Black-Scholes prices. This is a reasonable outcome because State 1 has a lower volatility associated with it, whilst State 2 has a higher volatility. All other conditions being equal, options written on indices with a higher volatility will have a higher price than options written on indices with a lower volatility. Since the HMM method allows the volatility to change over the life of the option, as the state value changes, the options that start in State 1 (with a lower volatility) will experience some periods of higher volatility before expiry, as the state value changes to State 2. This should have the effect of increasing the price to levels above those generated by the Black-Scholes formula, which assumes a constant (in this case lower) level of volatility. A complicating factor here, though, is that the risk-free rate is also not constant. All else being equal, options priced using a higher risk-free rate will have lower prices than options priced using lower risk-free rates. Since the higher risk-free rate is associated with the state with a higher volatility, on economic grounds, this will have a retarding effect on the increase in price due to the higher level of volatility. This may explain why the 60-Day option prices appear not to conform to the general pattern of prices described above. Please note that a more complete set of prices can be found on the CD-ROM accompanying this work, in the "Output" folder.

To determine whether the Hidden Markov Option Pricing Model can incorporate the volatility smile effect, the implied volatilities of options with varying combinations of initial index values and strike prices must be examined. A graph of these implied volatilities should then have a smile-like shape. Implied volatility is the volatility input required for the Black-Scholes formula that will set the Black-Scholes price of the option equal to that observed in the market, or that generated by another option pricing method, as is the case here. It is computed

by inverting the Black-Scholes formula, to solve for volatility. This requires some kind of nonlinear equation solution method. In this case the Solver routine from Microsoft Excel was used to solve for the various implied volatilities. Some of the results obtained are presented below. The remaining results can be found together with the option prices on the CD-ROM.

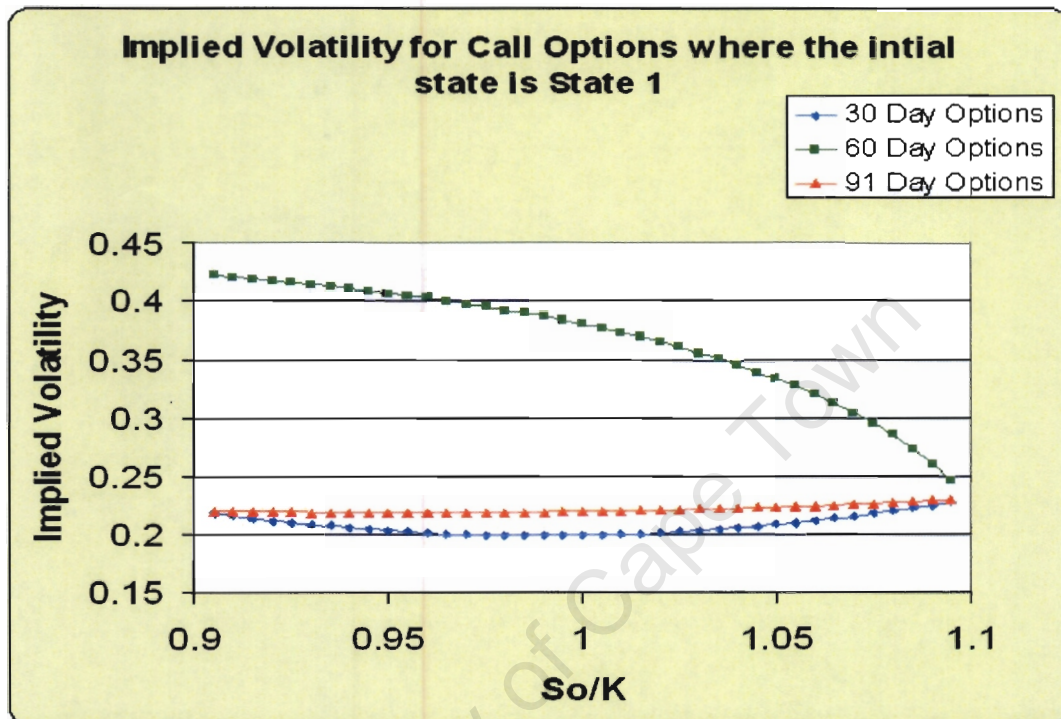


Figure 18: Implied Volatility for Call Options under the JSE 2-State model

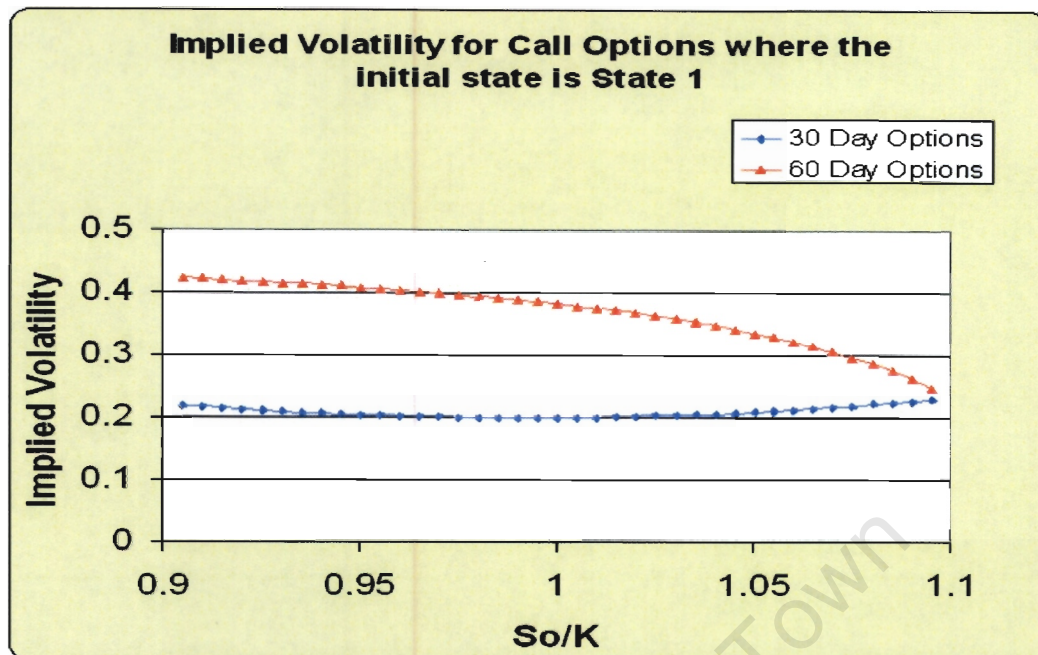


Figure 19: Implied Volatility for 30 and 60-Day Call Options under the 2-State JSE model

Call options of all maturities do seem to exhibit at least a 'skew-like' behaviour in their implied volatility graphs, with out-of-the-money options having a greater implied volatility and thus value. Clearly, the 30-Day options seem to exhibit an implied volatility structure that is closest to that of a smile. The implied volatilities of the 60 and 91-Day options appear to exhibit more of a concave skew. It is interesting to note that the implied volatilities for the 60-Day options are the highest, for all combinations of initial index value and strike prices of options. This is somewhat counterintuitive, since options with a greater maturity are expected to have higher prices and thus higher implied volatilities, all else being equal. Thus, this might suggest that the 2-state model is performing poorly in pricing 60-Day call options on the JSE All Share index. It may also be the case that the implied volatility measure is not suitable for use with option prices generated using the HMM. This is because it is assumed that this implied volatility is a constant value over the life of the option, whilst the HMM method allows the volatility used to price the option to vary over the life of the option. An example of a very poor performance of implied volatility is presented below.

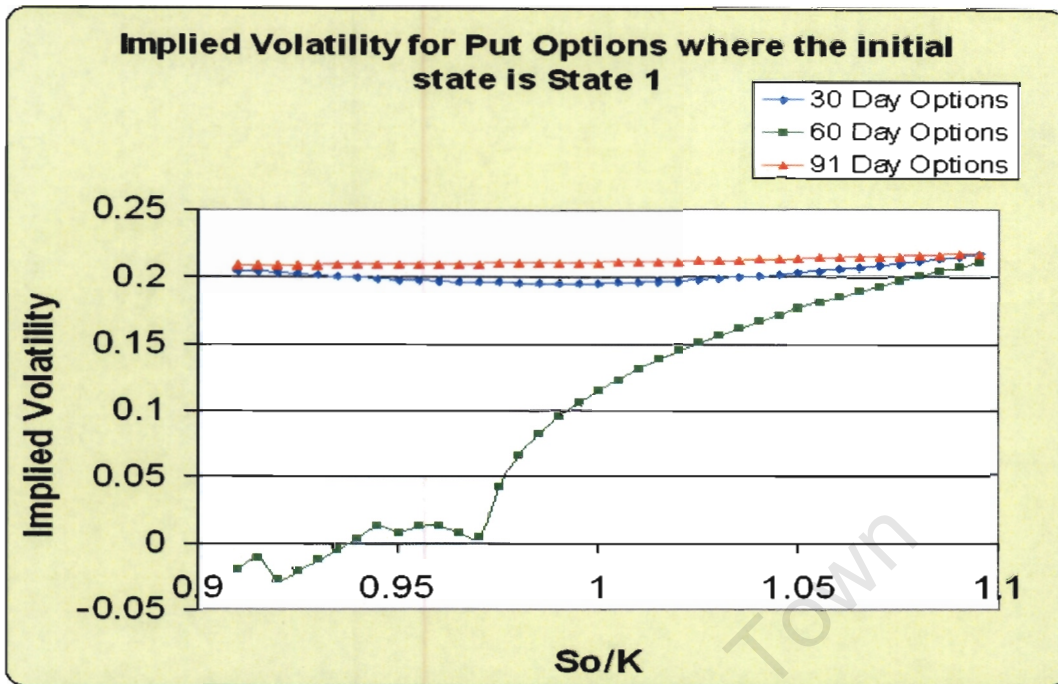


Figure 20: Implied Volatility for Put Options under the JSE 2-state model

As can clearly be seen, the implied volatility curve for 60-Day put options is very odd. For out-of-the-money options the implied volatility is even negative, which represents an impossible situation. Volatility can never be negative. Further investigation showed that these values may be due to the Solver optimisation. It is possible to find a set of implied volatilities that are all positive, for these option prices. The graph is virtually identical to that above, with the only difference being that the 60-Day line is slightly above the '0' level at the out-of-the-money end. It is interesting that it is again the 60-Day options for which the implied volatility statistic fails. The 30 and 60-Day curves appear to behave in a fashion closer to what would be expected, with a smile-like shape, although this is slightly difficult to observe on this scale.

## 11.2 Option Prices Computed using the 3-state model of the JSE Data

### Call Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
Initial State is State 1	0.6361	3.1928	8.5493	2.0716	5.5044	10.6941	3.7091	7.8721	13.4418
B-S Prices: State 1	0.0033	1.7572	10.079	0.2278	3.4736	11.3502	0.2278	3.4736	11.3502
Initial State is State 2	0.5121	2.8394	9.0027	1.681	5.1406	11.0199	3.0014	7.1622	13.2784
B-S Prices: State 2	0.1703	2.9732	10.494	0.7336	4.4551	11.6747	1.438	5.7313	12.8546
Initial State is State 3	2.2698	6.489	12.864	4.5947	9.4326	15.7143	6.6184	11.818	18.1255
B-S Prices: State 3	2.5191	6.8157	13.153	7.1816	12.449	18.7478	7.1816	12.449	18.7478

### Put Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
Initial State is State 1	7.5629	2.5205	0.6579	8.4791	4.0764	1.8224	9.8655	5.5011	2.9696
B-S Prices: State 1	8.9224	1.5265	0.048	8.9224	1.5265	0.048	8.4978	1.7435	0.1201
Initial State is State 2	8.2166	2.2556	0.5451	9.0256	3.8512	1.5284	10.06	5.0609	2.4758
B-S Prices: State 2	9.1032	2.8247	0.5443	9.5186	3.0467	0.609	8.9756	3.2689	0.8923
Initial State is State 3	11.254	5.473	2.3477	12.574	7.4118	4.1935	13.579	8.7785	5.5857
B-S Prices: State 3	12.889	7.7939	4.5571	12.889	7.7939	4.5571	13.992	9.26	6.0587

The relationship between the prices generated using the HMM method and the Black-Scholes formula is no longer as clear cut as in the case of the 2-state model above. Put options with State 1 as the initial state all have higher prices than those generated by the Black-Scholes formula. Call options that had State 2 as their initial state, on the other hand, have prices that are greater than, or less than, the Black-Scholes values depending on whether the options were out-of-the-money or in-the-money, respectively. Put and call options that had State 3 as the initial state all had prices less than the equivalent Black-Scholes values. This lack of a clear pattern may be due to the fact that the 3-state model of the JSE returns produced much more extreme estimates than the 2-state model. Since the AIC and BIC suggested that the 3-state model was a better model of the JSE data, this may be because it includes the more extreme effects present in this market. The HMM option prices are thus different from the Black-Scholes prices because they incorporate the variability and more extreme effects present in the JSE market. There is no clear pattern because these effects combine in different ways to influence the option prices, depending on the value of the initial state.

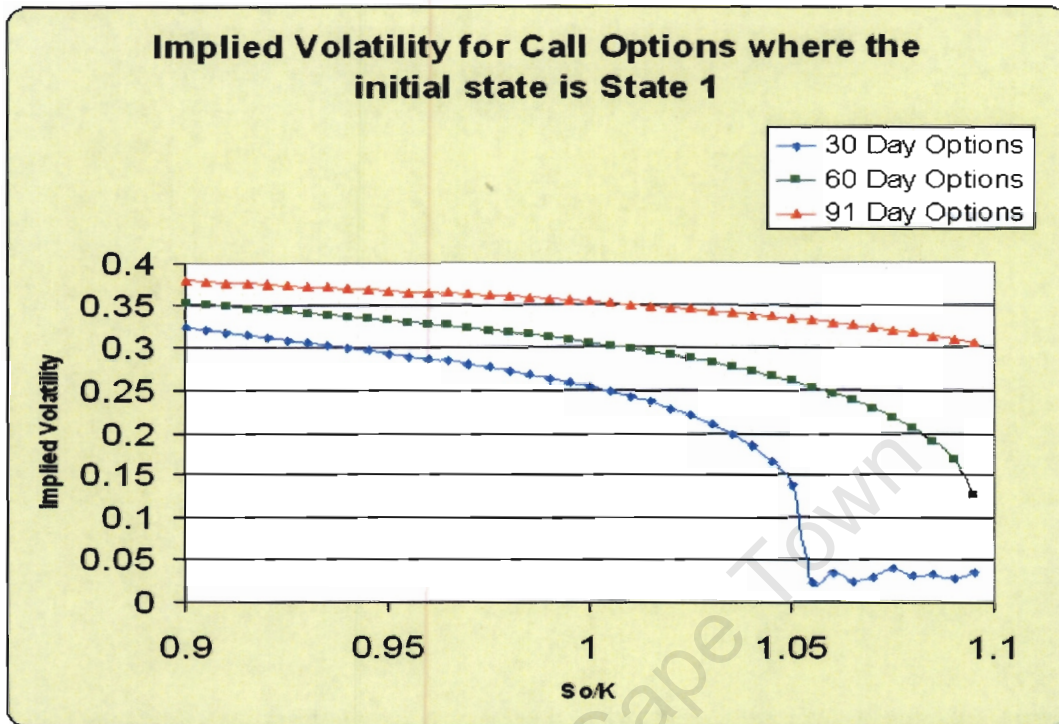


Figure 21: Implied Volatility for Call Options under the 3-state model of the JSE data

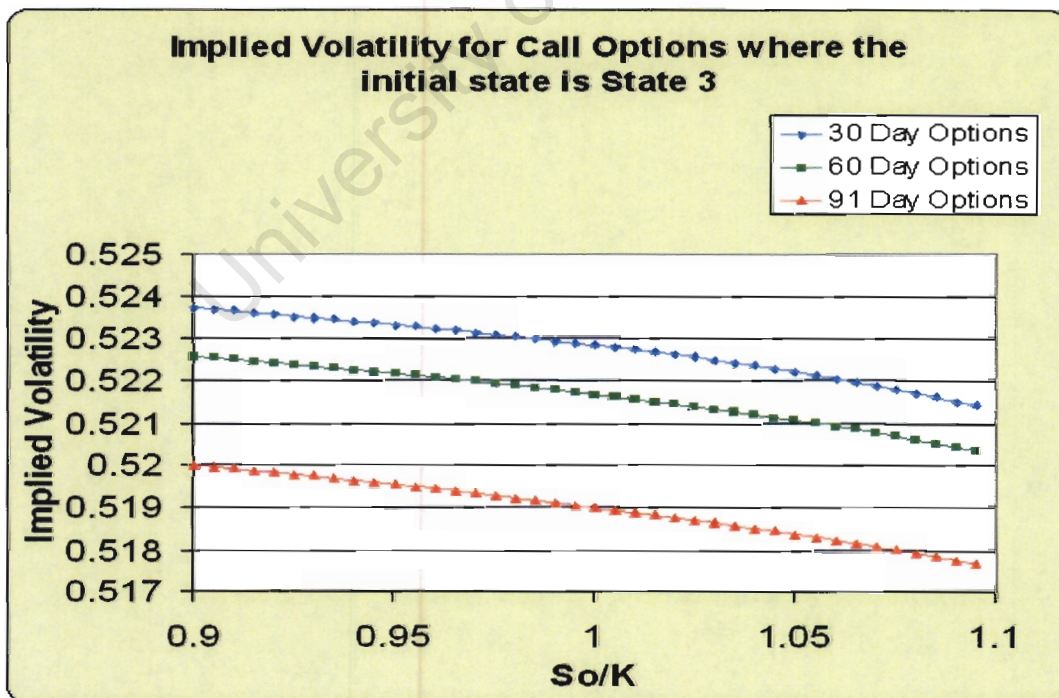


Figure 22: Implied Volatility for Call Options under the 3-State Model of the JSE data (2)

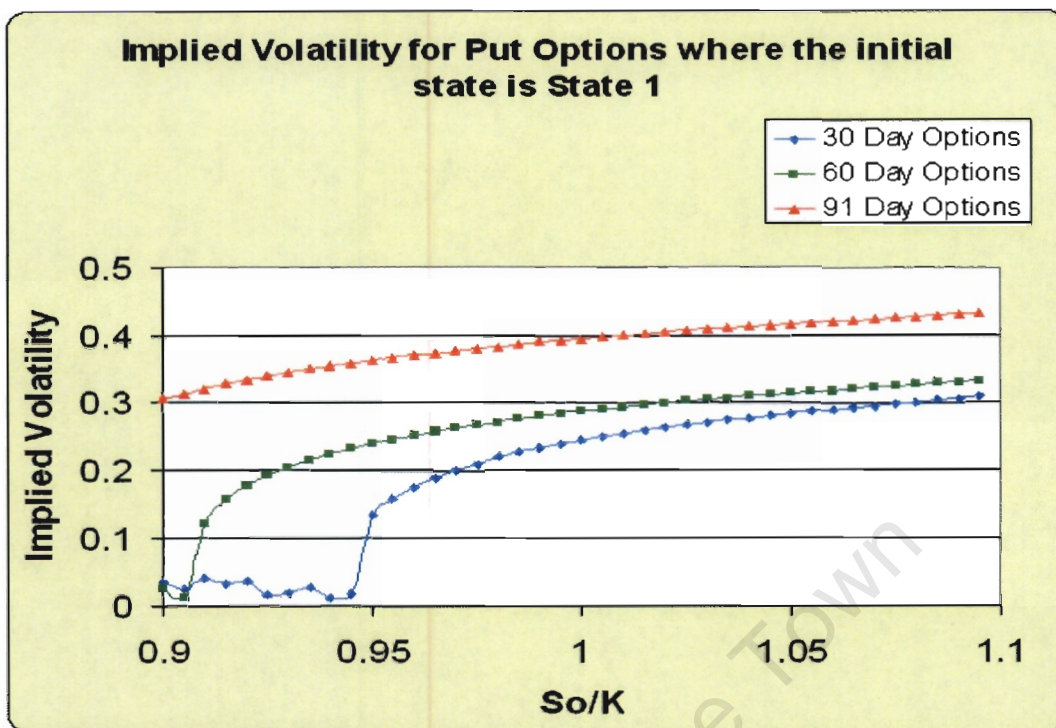


Figure 23: Implied Volatility for Put Options under the 3-state model of the JSE data

As was the case for the 2-state model of the JSE, the implied volatility curves for the various options and maturities of options either appear to exhibit smile-like behaviour, or not, depending on the initial state value. Call options that have State 3 as the initial state, for example, appear to have a slightly concave volatility skew. An anomaly in this graph, however, is that the options with longer maturities have lower implied volatility curves than options with shorter maturities. As was discussed earlier, this is the opposite of what would ordinarily be expected. Call options with State 1 as the initial state also appear to exhibit volatility behaviour consistent with the smile. As in the case of the 2-state model prices, the put options with an initial state of State 1 include implied option volatility curves that slope upward over the whole range of initial index value and strike price combinations. This is in complete contrast to the behaviour expected, which would be a fall in implied volatility followed by a rise. This suggests that the Hidden Markov Option Pricing Model may be under-pricing in-the-money put options.

As noted previously, it is unclear whether the odd results using implied volatility for some options are a result of the Hidden Markov Option Pricing Model performing poorly, or the lack of applicability of the implied volatility method to this situation. A final verdict on the matter may require the comparison of model generated option prices and those found trading in the market.

## 11.3 Option Prices Computed using the 2-state model of the FTSE 100 Data

### Call Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
Initial State is State 1	0.0498	1.9735	9.9699	0.308	3.145	10.622	0.7281	4.1775	11.367
B-S Prices: State 1	0.0033	1.6531	9.8354	0.0033	1.6531	9.8354	0.0033	1.6531	9.83544
Initial State is State 2	0.4537	3.5578	10.434	1.3197	5.1216	11.643	2.1975	6.393	12.7763
B-S Prices: State 2	0.4693	3.6131	10.512	0.4693	3.6131	10.512	0.4693	3.6131	10.5116

### Put Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
Initial State is State 1	9.6502	1.574	0.0703	9.5226	2.3596	0.3362	9.5511	3.0006	0.6901
B-S Prices: State 1	9.67504	1.3249	0.0072	9.4033	1.7827	0.0694	9.1994	2.1086	0.1752
Initial State is State 2	10.1246	3.2287	0.6049	10.663	4.4645	1.4864	11.203	5.3981	2.2814
B-S Prices: State 2	10.0592	3.203	0.6015	10.546	4.4046	1.4655	11.037	5.3016	2.2364

Unlike the case of the 2-state model of the JSE data, the prices generated using the 2-state model of the FTSE 100 data are all greater than the Black-Scholes equivalents, with the exception of 30-Day call options with an initial state of State 1. This is interesting because this means that even though the options priced include periods of occupying the lower volatility state, State 1, they still produce prices that are higher than the Black-Scholes prices. This is even true when the Black-Scholes prices use the higher level of volatility as a constant, over the life of the options.

Notice, however, that the prices generated by the HMM method are much closer to the Black-Scholes prices than was the case for the 2-state model of the JSE data. This may be due to the fact that the 2-state model provided a better description of the FTSE data than the JSE data, with lower AIC and BIC values, as was noted in the previous chapter. This may, in turn, allow the Hidden Markov Option Pricing Model to produce 'better' prices for the various options. Another explanation may be that the London market is less variable than the JSE and also switches between states less frequently than the JSE. This would mean that the options are priced using volatility and risk-free rate sequences that are closer to being constant than those associated with the JSE. This would have the effect of producing prices that are closer to the Black-Scholes prices, which assume that the volatility and risk-free rate remain constant over the life of the option.

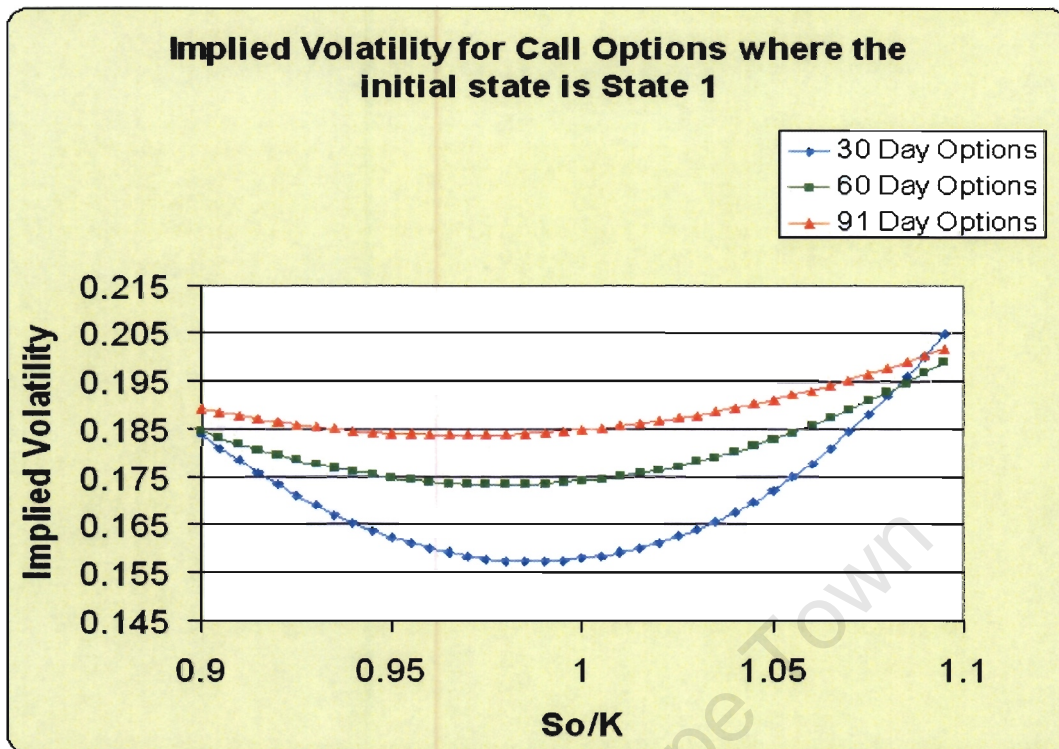


Figure 24: Implied Volatility for Call Options under the 2-state model of the FTSE 100 data

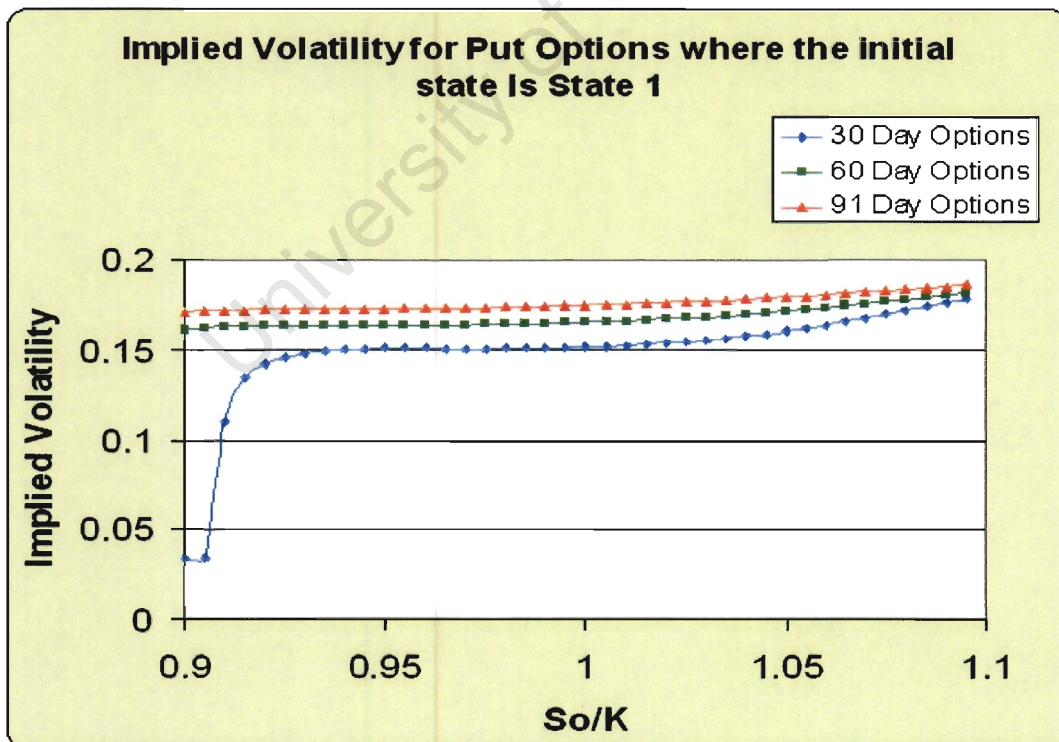


Figure 25: Implied Volatility for Put Options under the 2-state model of the FTSE 100 data

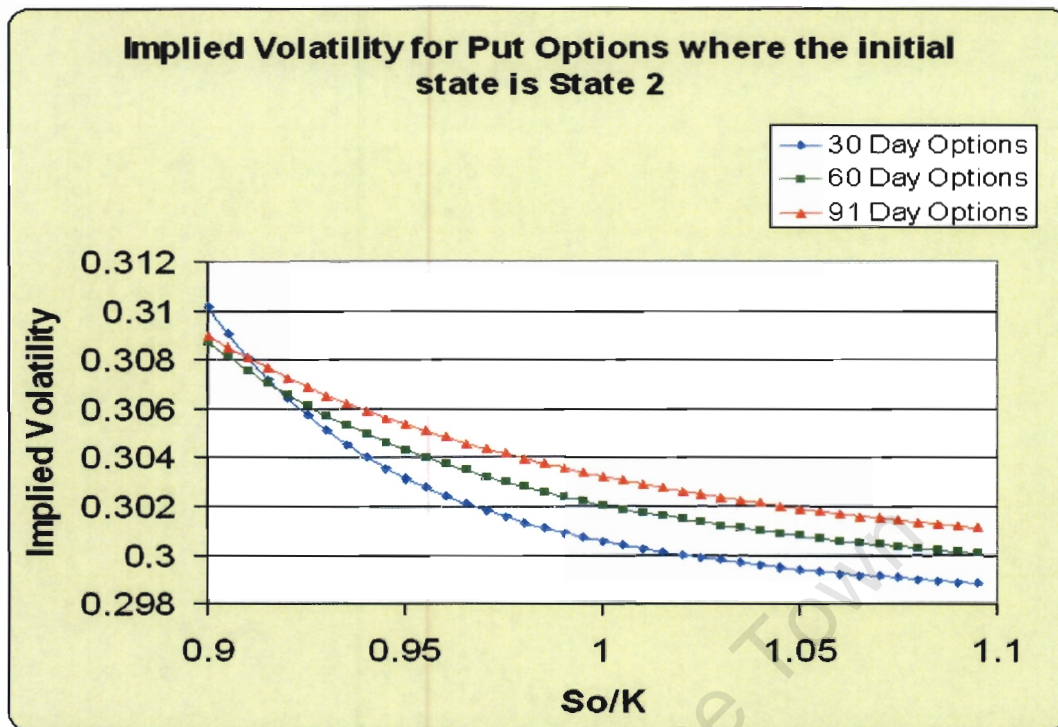


Figure 26: Implied Volatility for Put Options under the 2-state model of the FTSE 100 data (2)

The implied volatility curves for this model are much closer to what would be expected under the volatility smile effect than those of the 2-state model of the JSE data. The implied volatilities for call options with State 1 as the initial state produce a textbook volatility smile, for all maturities. The level of the implied volatility curves increases with increasing maturity of the options as well, as would be expected. In fact, this set of implied volatility curves was the closest to the expected volatility smile for all options priced. The put prices with State 2 as the initial state produce more of a convex skew than a smile, but this is still not extremely far from the expected behaviour. Even in the case of an apparently poor set of curves, as for put options with an initial state of State 1, are mostly close to the volatility smile. The only distorting effect here is due to a few very small implied volatilities for deep out-of-the-money puts.

This better performance of the model, in terms of implied volatilities, may be due to the fact that the prices produced by the model for options written on the FTSE 100 index are closer to Black-Scholes prices than for the JSE. This, in turn, makes the assumption of a constant volatility over the life of the option more reasonable. The volatility smile can then be more clearly determined than in the case of the JSE options, which nevertheless did exhibit differing levels of implied volatilities for options of differing initial index value and strike price combinations.

## 11.4 Option Prices Computed using the 3-state model of the FTSE 100 Data

### Call Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
Initial State is State 1	0.2018	2.4374	8.5486	0.8486	3.6429	8.8285	1.5869	4.7007	9.62715
B-S Prices: State 1	0.0007	1.4649	9.8301	0.0245	2.1735	10.185	0.1012	2.7772	11.6504
Initial State is State 2	0.1625	1.9474	8.5486	0.6851	3.3442	9.579	1.4052	4.6937	10.7088
B-S Prices: State 2	0.1041	2.5862	10.058	0.4938	3.7772	10.84	0.9945	4.7678	11.6504
Initial State is State 3	1.225	4.9429	11.441	2.7721	7.1154	13.387	4.1835	8.8799	15.0917
B-S Prices: State 3	1.2678	5.0087	11.492	2.8483	7.2113	13.474	4.2856	9.0002	15.2054

### Put Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
Initial State is State 1	8.2059	2.086389	0.2604	7.8495	2.9851	0.895	8.0893	3.7149	1.5275
B-S Prices: State 1	9.6725	1.136684	0.0019	9.3691	1.5181	0.0295	9.1089	1.7849	0.0885
Initial State is State 2	8.699	2.086389	0.2028	8.6735	2.7003	0.7345	9.1434	3.6789	1.3787
B-S Prices: State 2	9.694	2.176045	0.1484	9.6753	2.9587	0.5213	9.7557	3.529	0.9116
Initial State is State 3	10.738	4.455507	1.4536	11.8	6.1432	2.9149	12.713	7.4092	4.121
B-S Prices: State 3	10.776	4.516719	1.5004	11.867	6.2299	2.9921	12.801	7.5154	4.2206

As was found for the option prices computed using the 3-state model of the JSE Data, the relationship between HMM prices and Black-Scholes prices is no longer as clear cut as is the case for prices computed using the 2-state model. For example, call options with State 1 as an initial state all have prices higher than the Black-Scholes equivalent, whilst put options with State 1 as the initial state have prices that are greater, or less, than the Black-Scholes values, depending on whether the options are out-of-the-money, or in-the-money, respectively. This is due to the complicating effect of the extra state on the dynamics of the model. This is particularly true for the model of the FTSE 100 data, which would only transit from State 1 to State 3 via State 2, the 'random walk' type phase of the market. This conclusion was reached by examining the estimated transition probabilities of the Markov Chain.

The 3-state model also indicated that the FTSE 100 index was slightly more likely to transit between states under the 3-state model, with slightly lower probabilities of remaining in a given state than under the 2-state model. The prices produced are thus likely to be more different from the equivalent Black-Scholes values than those computed using the 2-state model of the FTSE 100 data. This can be seen

by comparing the tables of prices above. The differences are still not as severe as those for the JSE models, since the overall variability of the FTSE 100 data, including the number of state transitions, is less than that of the JSE. This means that price calculations have 'more constant' inputs than in the JSE case, and are thus more similar to the Black-Scholes values, which assumes that all inputs are constant over the life of the option.

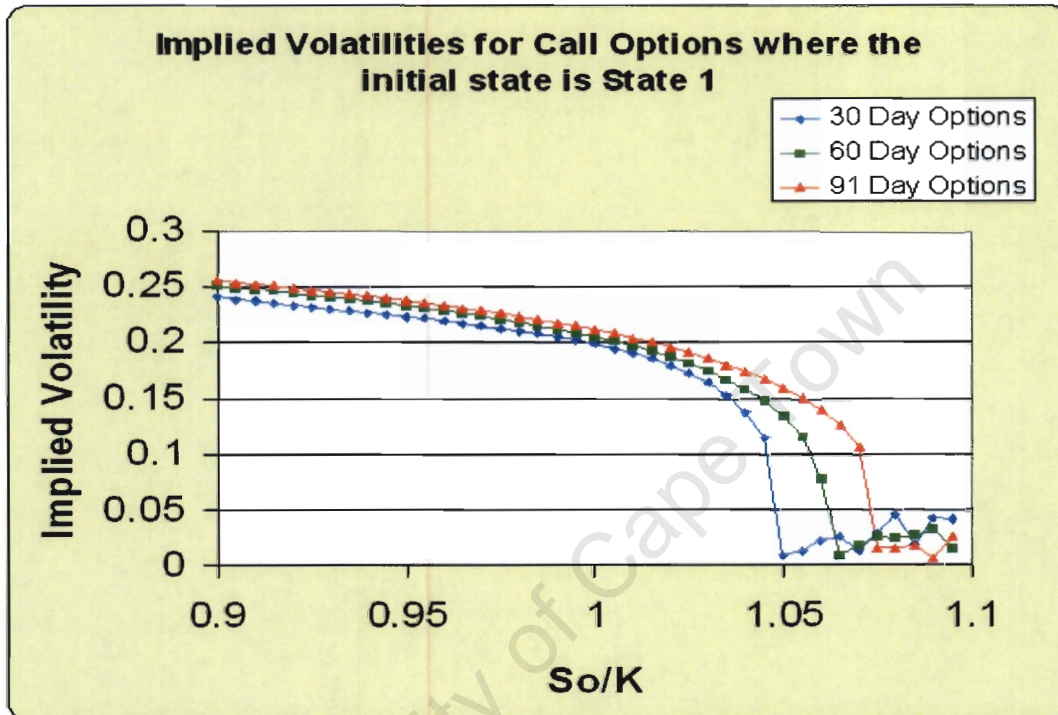


Figure 27: Implied Volatilities for Call Options under the 3-state model of the FTSE 100 data

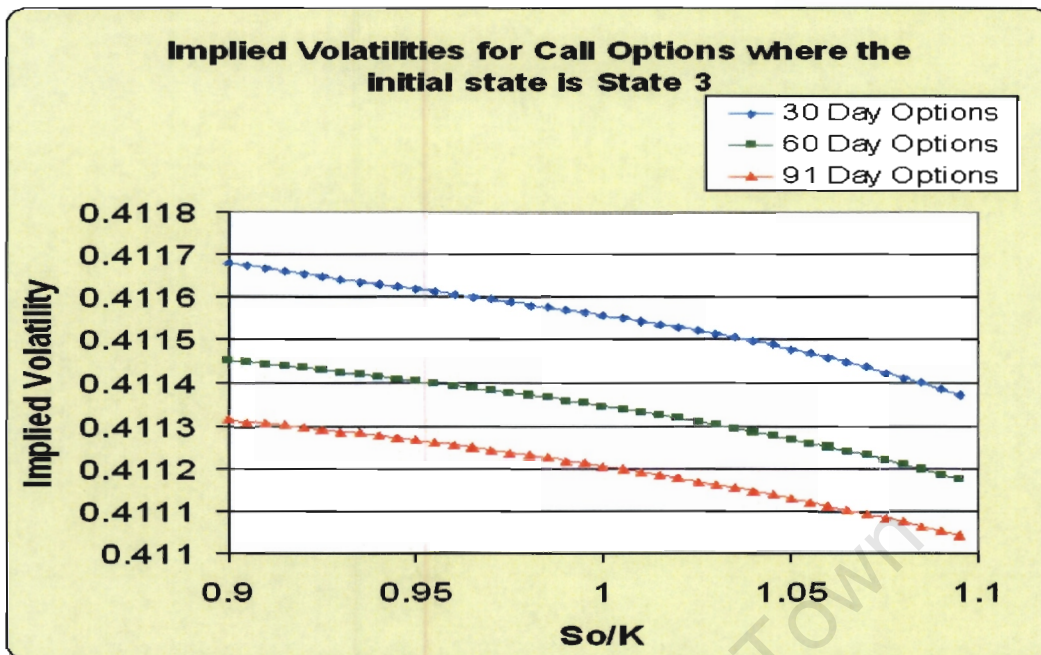


Figure 28: Implied Volatilities for Call Options under the 3-state model of the FTSE 100 data (2)

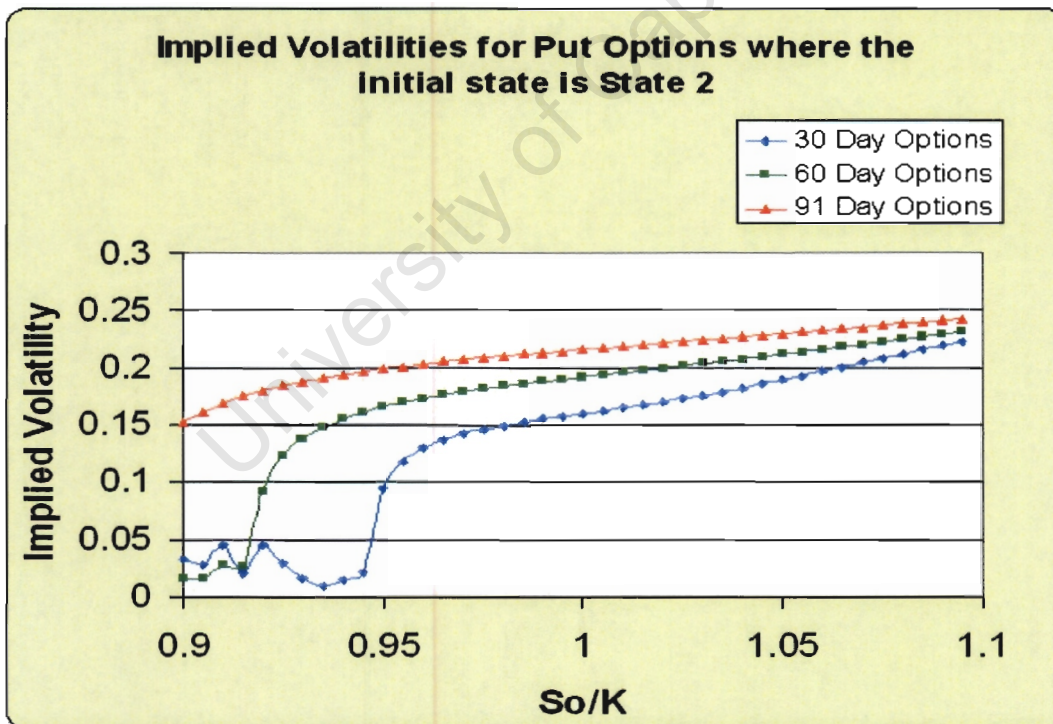


Figure 29: Implied Volatilities for Put Options under the 3-state model of the FTSE 100 data

Although the prices generated by the 3-state model of the FTSE 100 data are only slightly more different than the Black-Scholes prices than for the 2-state

case, the implied volatility curves perform far worse. Both the put options with State 2 as the initial state, as well as the call options with State 1 as the initial state, produce results that are other than what would be expected. Both in the case of the put options and call options, the curves exhibit anomalous behaviour for deep in-the-money options. These are on opposite sides of the graph because call options are in-the-money when the index value is high compared to the strike value of the option, whilst the reverse is true for put options. Only in the case of options with State 3 as the initial state does the shape of the implied volatility curves appear to conform to the volatility smile effect. As was found in other cases, however, the level of the curves for the differing maturities is contrary to what would be expected. The lower maturity options have the higher level of implied volatility, whilst the higher maturity options have a lower level of implied volatility. As was discussed previously, this is the opposite of what would be expected.

The departure from the behaviour expected from the implied volatility is greater for the prices generated by the 3-state model of the FTSE 100 data than those for prices generated by the 2-state model. This could be due to the fact that the prices generated by the 3-state model differed from the Black-Scholes prices by more than those from the 2-state model. Alternatively, the 3-state model could have been able to include effects that the 2-state model did not, which result in a better description of the dynamics of the FTSE 100 data. This greater influence of the dynamics could result in implied volatility becoming less applicable, as was discussed previously.

## 11.5 Brief General Conclusions

The model fitting procedure in the previous chapter indicated, both for the FTSE 100 data and the JSE data, that the 3-state model described the data more accurately than the 2-state model. Despite this, in both cases the implied volatility curves for prices generated by the 3-state model were more different from what would be expected, in the form of the volatility smile effect. The option prices were also more different from the Black-Scholes equivalent for the 3-state model prices than for the 2-state model prices, for both datasets. This result was particularly extreme in the case of the JSE data. This could be due to the fact that the 3-state model better describes the dynamics of the two indices, and thus renders the implied volatility measure less applicable with its assumption of constant volatility over the life of the options.

This greater incorporation of dynamic effects may also explain the difference between the option prices generated by the model and the Black-Scholes prices. This is because the Black-Scholes formula assumes that the inputs to the formula, volatilities and risk-free rates, remain constant over the life of the option. If the inputs do in fact vary considerably often, as was found especially to be the case with the JSE data, then the model prices that take account of this variability will differ from the Black-Scholes prices, as was found here. Since the FTSE 100

data exhibited less overall variability, in both volatility and state transitions, than the JSE data, the model prices generated for options written on the FTSE 100 index were less different from the Black-Scholes equivalent prices than options written on the JSE. A final check on whether the model option prices are, in fact, more accurate than the Black-Scholes prices, as well as being different from them, will require a comparison with traded option prices.

Chapter 12.	Calculating Prices for Bermudan-style Options using the Hidden Markov Option Pricing Model
-------------	--

The example that will be used to illustrate the Hidden Markov Option Pricing Model for Bermudan-style options, empirically, will be that of a payer swaption. A swaption is an option that grants the holder the right to enter the pay-fixed, or the receive-fixed, side of a swap agreement at or before the expiry date of the swaption, at a contracted swap rate. A swap is an agreement to exchange a set of cashflows calculated based on a fixed rate of interest, called the swap rate, for a set calculated using a floating rate of interest. This floating rate is usually a function of some benchmark short term interest rate, such as the London Inter-Bank Offer Rate (LIBOR), which is the rate at which London banks will lend US dollars to each other. Thus, the variable underlying the swaption contract is the swap rate for the swap agreement that the swaption is based on. This swap could have any maturity, starting from the expiry date of the swaption, or the exercise date if the swaption is exercised prior to expiry. Typically, swap maturities run from 1 year to 30 years, although any maturity can be contracted between willing parties.

Payer swaptions grant the holder the right to pay a given fixed rate of interest in the swap underlying the contract. These swaptions are, effectively, call options on the swap rate for the swap in question (Hull (2003), p521) and are the variety of swaption that will be considered here. Receiver swaptions grant the right to receive a fixed rate of interest and are put options on the swap rate. The calculation of prices for these swaptions is thus very similar to that for those presented below. The size of the payments that will be made under the swap depends on the 'notional principal' of the swap. This is the amount that the interest rates, specified by the swap agreement, will be multiplied by to calculate the payments in the same way that the principal in a loan agreement is used to calculate the interest payments for the loan. The difference between a swap and a loan, in this regard, is that the 'principal' of the swap is used only for calculation purposes. Since this value is essentially a scaling factor, it will be assumed to be 1 for the purposes of the analysis that follows below. Prices for swaptions with notional principal values other than 1 can be obtained by multiplying the prices presented below by the notional principal desired.

The choice of swaptions for the illustration of the Bermudan case of the Hidden Option Pricing Model was made partly for variety's sake and partly for the reason that the swaption market is one option market where Bermudan-style options are actively traded. This is thus an application of the Bermudan case of the model that is likely to be applicable in a practical setting.

## 12.1 Basic Summary Analysis of the Data

Summary Statistics	
Mean	-0.00019
Std Dev	0.008164
Skewness	0.536768
Kurtosis	5.495302
JB	801.8527
p-value	0
Data points	2608

The data that was used to fit the model to was the swap rate for 5-year swaps in the UK market. As noted in Chapter 1, daily data from the period 15/03/1996 – 15/03/2006 was used and was sourced from Thomson Datastream. As was the case for the index data, weekend days are not included in the dataset. Only week days are considered. Public holidays are assigned the value of the swap rate on the previous day. It is therefore assumed that the swap rate remains constant over public holidays. The swap rates used were “Market values”, according to the Datastream classification. The full dataset, together with these statistics can also be found in the “Source data” folder on the CD-ROM that accompanies this work. As was the case for the index data used in the previous chapters, it is assumed that the swap rate is lognormally distributed. The argument in favour of this assumption is very similar to that presented for the index data in Chapter 10 and will thus not be repeated here. This assumption implies that a Hidden Markov Model where, conditional on the unobserved state value, the variable modelled is considered to be normally distributed will be fitted to the log returns of the swap rate. To determine the approximate nature of the component normal distributions of the HMM mixture distribution, the summary statistics of the log returns of the swap rate are examined below.

As was the case for the index data, the swap rate is clearly not normally distributed. It is positively skewed and has a high kurtosis. Furthermore, the Jarque-Bera statistic (Jarque and Bera (1987), p166) is very large and thus the p-value for the Jarque-Bera test is very low. This is further confirmation that the data is not normally distributed. To determine the possible nature of the components of a mixture of normal distributions that might account for the observed data distribution, the above statistics will be considered in conjunction with the histogram of the data below.

As can be seen from the histogram, the data is distributed largely around the mean, which is close to zero. Thus the component distributions used in the HMM should have mean values that do not differ by large amounts. They should differ slightly, to allow for the positive skewness observed in the data. The high

kurtosis, at 5.495, can be seen in the peaked nature of the histogram with the peak occurring near the mean. This can be accounted for in the HMM by allowing the variances of the hidden states to be different, but with means that are very similar. If the state with the higher mean value associated with it were to have the higher variance as well, then the overall mixture distribution of the data would resemble the observed data below relatively closely.

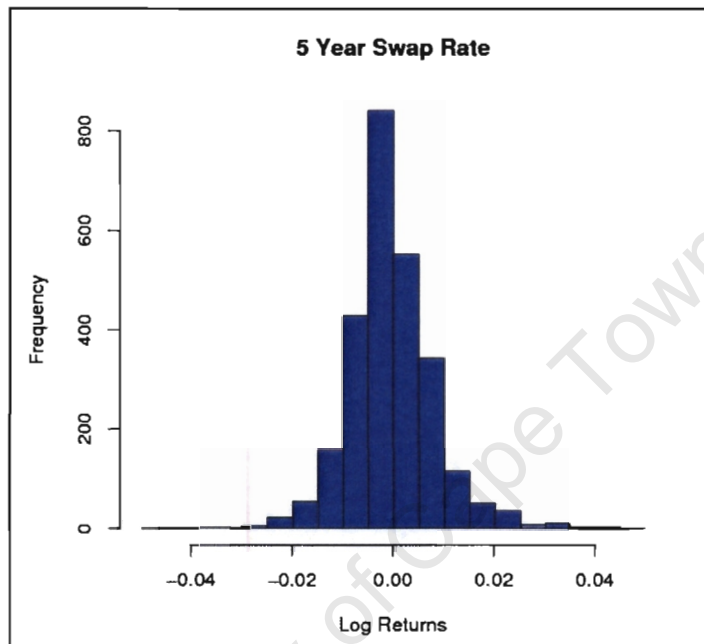


Figure 30: Histogram of the log returns of the 5-year UK swap rate

## 12.2 Estimated values for the 2-state Model of the Data

	Estimated Values	Standard Errors
$\mu_1$	-0.228189	0.000135
$\mu_2$	0.2933498	0.00042
$\sigma_1$	0.1100345	1.10E-06
$\sigma_2$	0.2268307	7.05E-06

AIC	-17952.4
BIC	-8958.61
Interest Rate: State 1	0.04
Interest Rate: State 2	0.05
Expected Duration: 1	16.29918
Expected Duration: 2	7.208037

Transition Probabilities for the Markov Chain

To From	State 1	Std Error	State 2	Std Error
State 1	0.9386472	0.009002	0.061353	0.009002
State 2	0.138734	0.009326	0.861266	0.009326

The parameters were estimated using the same procedure as described in Chapter 10, as were the AIC, BIC and expected duration in each state. As previously, the R code required to perform these estimations can be found on the CD-ROM that accompanies this work, in the "R Code" folder. The interest rates were selected in the same way as in Chapter 9 and are the same as those used for the 2-state model of the FTSE 100 data. This is because both the FTSE 100 data and the swap data are from the UK market. As in Chapter 10, these estimates are reported as annualised figures, since these are the values that are used in the option pricing formula.

As was expected, the distribution with the higher mean value also has the higher variance, which will result in the overall mixture distribution being close to that observed in the histogram above. As was the case for the index data, the standard errors of the estimates (as per Louis (1982)) were reasonably low. This implies that the parameter estimates are all likely to be statistically significant. The expected duration and transition probabilities both indicate that State 1 is the more persistent of the two states. This is the state with a negative mean return associated with it, suggesting that on average 5-year swap rates decreased for longer periods of time, once they began to decrease, than they increased once they began increasing. Despite this, the expected duration in each state is fairly low, being approximately one week for State 2 and two weeks for State 1. Comparing these values to those obtained for the index data in Chapter 10, it appears that the 5-year swap rate data series tended to undergo structural changes more often than the index data. This is confirmed by the greater variability in the values obtained for the smoothed probability of being in each state, presented below. The probabilities are also less often 0 or 1 than was the case for the index data, which also points to a greater amount of switching

between states since it is less often certain what the value of the unobservable state is.

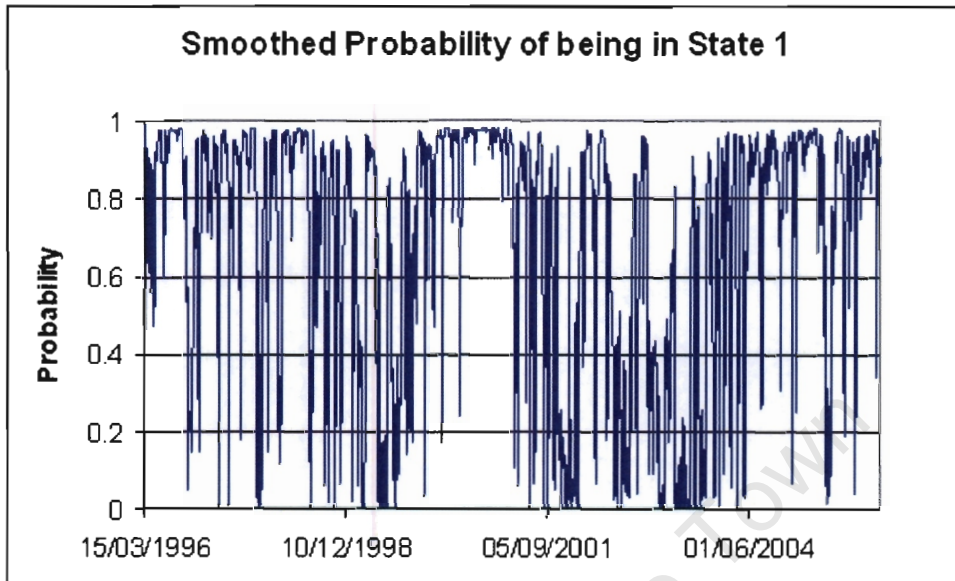


Figure 31: Smoothed Probability of being in State 1 in the 2-state model of the swap data

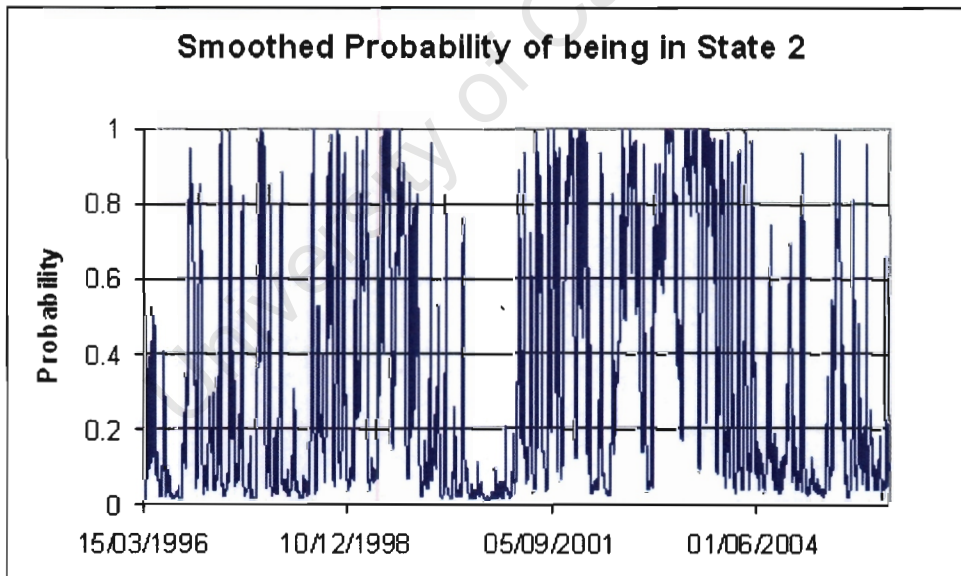


Figure 32: Smoothed Probability of being in State 2 in the 2-state model of the swap data

## 12.3 Estimated values for the 3-state Model of the Data

	Estimated Values
$\mu_1$	0.3857731
$\mu_2$	-0.26298
$\mu_3$	0.162197
$\sigma_1$	0.2119024
$\sigma_2$	0.1005515
$\sigma_3$	0.2260941

AIC	-18013.73
BIC	-8971.669
Interest Rate: State 1	0.05
Interest Rate: State 2	0.04
Interest Rate: State 3	0.06
Expected Duration:1	1.2667601
Expected Duration:2	4.6833196
Expected Duration:3	65.6774

Transition Probabilities for the Markov Chain

To From	State 1	State 2	State 3
State 1	0.210585	0.789415	8.33E-13
State 2	0.208775	0.786476	0.004749
State 3	0.015225	6.61E-07	0.984774

As was found to be the case for the index data in Chapter 10, the 3-state model appears to fit the data more accurately than the 2-state model. This is suggested by the fact that both the AIC and BIC values are lower for this version of the model than was the case for the 2-state model. This could be due to the fact that the inclusion of an extra state variable allows for the inclusion of effects in the model that the 2-state model was unable to incorporate, due to its sparser parameterisation. The Observed Information Matrix was singular, as for the 3-state models of the index data, resulting in no standard errors for these estimates being available.

It is interesting to note that the state that has the highest probability of persisting, from one time point to the next, in this model, by quite a large margin, is State 3, which has a positive rate of return associated with it. This is evidenced by the probability of remaining in the state being over 0.9 and the expected duration being more than two months. This seems to contradict the result obtained for the 2-state version of the model. The overall mean rate of return for the data was negative, which further casts doubt on this result. However, this is not the highest rate of return that was estimated. It was the second highest, with the lowest mean rate of return being a fairly large negative value. Furthermore, the state with the highest mean rate of return is also the least persistent, with an expected duration of less than two days. It thus appears that the 'medium' case, at least in terms of mean rate of return, was the most likely with more extreme cases of increase and decrease in the swap rate persisting for less amounts of time. This

makes economic sense, since these swap rates are taken from a developed economy, where economic variables are likely to be more stable over time.

The smoothed probabilities of being in a given state, presented below, confirm that the model is less likely to remain in States 1 and 2 for prolonged periods, since the probabilities are more variable for these states. They are also less often 1 or 0, which means that it is less often certain that State 1 or State 2 is the unobserved state in force at the given time point.

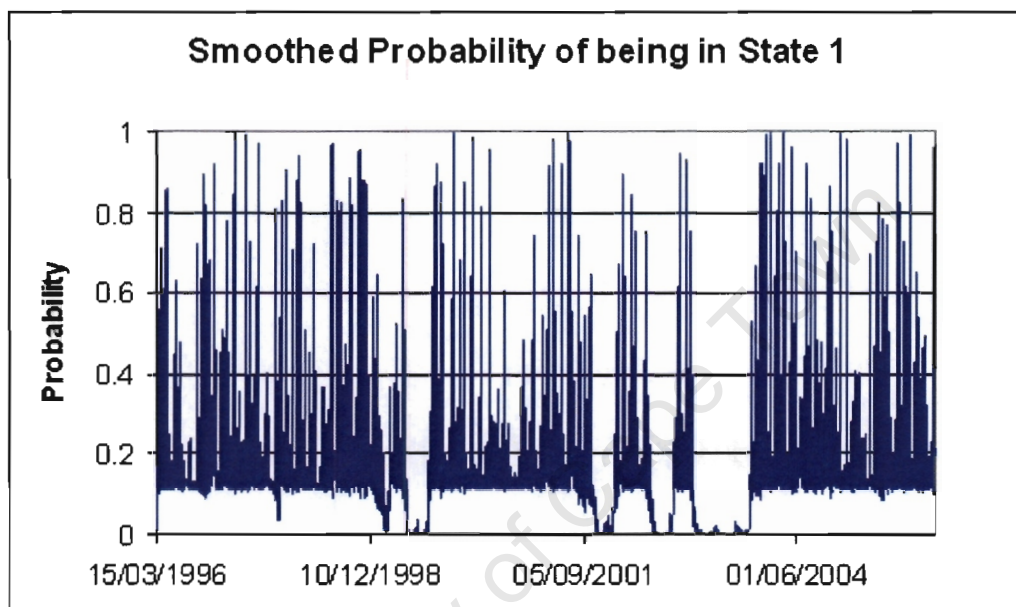


Figure 33: Smoothed Probability of being in State 1 under the 3-state model of the swap data

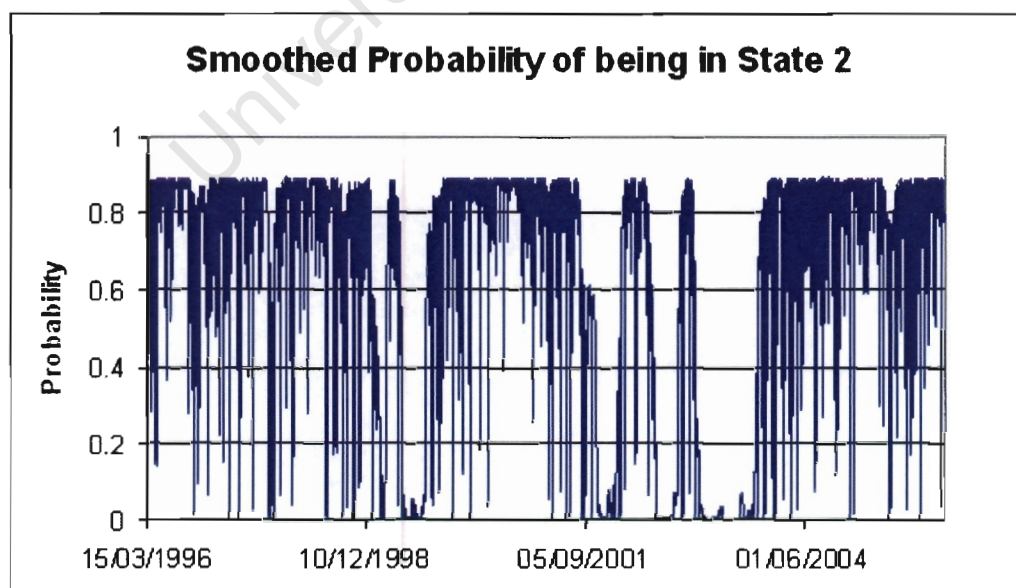


Figure 34: Smoothed Probability of being in State 2 under the 3-state model of the swap data

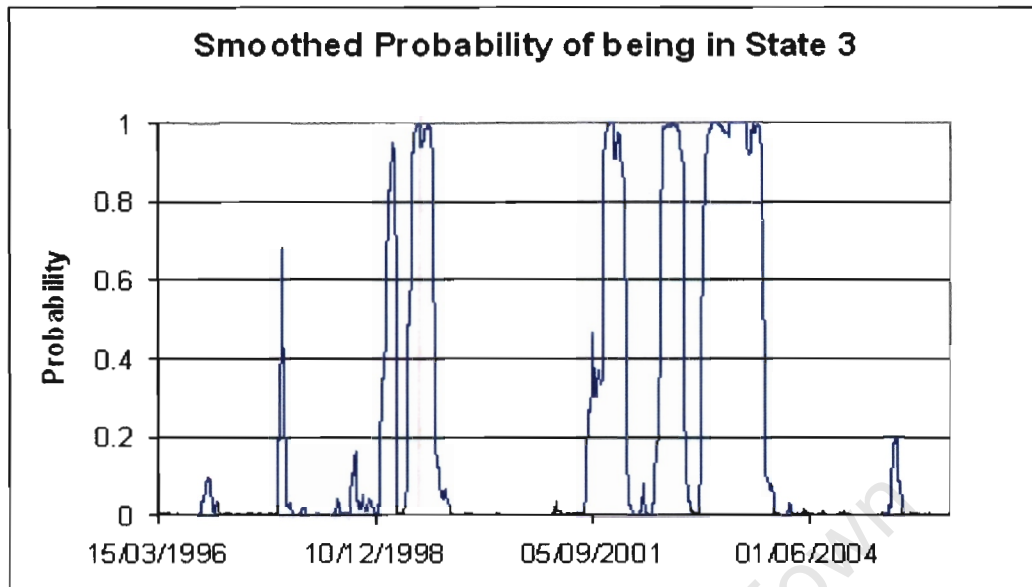


Figure 35: Smoothed Probability of being in State 3 under the 3-state model of the swap data

## 12.4 Pricing Bermudan-style Payer Swaptions

The swaptions on 5-year swaps that will be priced here will have maturities of 365, 182 and 91 days for the 2-state model and 91 days and 60 days for the 3-state model. The reason for the shorter maturities in the case of the 3-state model is purely to ensure that the calculations did not become too cumbersome, recalling that  $(\tau+1)$  sets of occupation times will need to be stored for options of maturity  $\tau$  days. Differing numbers of possible exercise dates will be considered for swaptions of each maturity, to determine whether this has any effect on the price of the swaptions. For the 365-day swaptions, prices will be computed for 4, 6 and 12 possible exercise dates. For the 182-day swaptions, prices for 3, 6 and 9 possible exercise dates will be computed. For 91-day swaptions prices for 3 and 6 possible exercise dates and for the 60-day swaptions, prices for 2 and 3 possible exercise dates will be computed.

Prices for these swaptions will also be computed for several different strike rates, to determine the difference (if any) in the behaviour of prices for options that are in- or out-of-the-money. A more detailed analysis of this effect using implied volatilities, as was performed in the previous chapter, is not possible here since there is no standard formula available for pricing Bermudan-style options from which implied volatilities could be computed.

The lattice method of Bollen (1998) will be used to compute prices for swaptions based on the 2-state model of the swap rate, for comparison against those generated using the HMM method, described in Chapter 8. This lattice method can only be used for a Hidden Markov framework with two possible values for the

unobserved state variable. Visual Basic for Applications (VBA) code for pricing these swaptions using this lattice method can be found on the CD-ROM accompanying this work, in the "R Code" folder. There will thus be no available prices for comparison with those generated using the 3-state model. Prices will, however, be computed for the swaptions specified using a trinomial tree approach, with no hidden states. This will allow for a comparison against the prices computed using both the lattice method and the HMM approach outlined in Chapter 8, which incorporate the effect of the hidden states. For the case of the 3-state models, prices will also be computed using finite difference methods to allow for a more complete comparison, since the lattice method of Bollen (1998) does not apply there. An indication of the possible value of the information contained in the hidden states can then be obtained.

### 12.4.1 Swaption prices computed using the 2-state model

(The initial value of the swap rate was taken to be 4.7625%, the final value in the time series used for estimating the HMM parameters above)

	Strike	4.85	4.8	4.7625	4.7	4.65
Maturity	Possible Exercise Dates					
365 days	12	7.258676	7.474409	7.645474	7.949904	8.212228
	6	5.87744	5.974418	6.050433	6.183916	6.297246
	4	5.68183	5.754557	5.811314	5.910153	5.993057
182 days	3	0.639144	0.691495	0.733218	0.808289	0.874107
	6	0.892734	0.983904	1.058045	1.1939	1.314769
	9	2.217462	2.44959	2.631443	2.948484	3.213865
91 days	3	0.579765	0.629652	0.670177	0.744717	0.811593
	6	0.801767	0.89572	0.969856	1.099705	1.208598

Lattice Prices:

	Strike	4.85	4.8	4.7625	4.7	4.65
Maturity	Possible Exercise Dates					
365 days	12	37.78608	37.78608	37.78608	37.78608	37.78608
	6	37.78608	37.78608	37.78608	37.78608	37.78608
	4	37.44771	37.44771	37.44771	37.44771	37.44771
182 days	3	14.22918	14.22918	14.22919	14.22919	14.22919
	6	14.22918	14.22918	14.22919	14.22919	14.22919
	9	14.22918	14.22918	14.22919	14.22919	14.22919
91 days	3	8.675563	8.675703	8.675808	8.675983	8.676123
	6	8.737465	8.737611	8.737721	8.737903	8.738049

Prices computed with no hidden states:

	Strike	4.85	4.8	4.7625	4.7	4.65
Maturity	Possible Exercise Dates					
365 days	12	0.368582	0.395489	0.41567	0.452231	0.482299
	6	0.368582	0.395489	0.41567	0.452231	0.482299
	4	0.368582	0.395489	0.41567	0.452231	0.482299
182 days	3	0.223972	0.24981	0.269189	0.305826	0.336332
	6	0.223972	0.24981	0.269189	0.305826	0.336332
	9	0.223972	0.24981	0.269189	0.305826	0.336332
91 days	3	0.134511	0.158998	0.177363	0.214261	0.245484
	6	0.134511	0.158998	0.177363	0.214261	0.245484

As might be expected, the prices computed without taking into account the effects of any hidden state variables were the lowest, although it is perhaps somewhat surprising that the difference in values is so large. This is especially pronounced for the swaptions with a maturity of 365 days. The difference in price for the shorter maturity swaptions is less extreme, which is perhaps to be expected since a longer maturity would allow time for more changes of the unobserved state, implying a structural change in the time series, to occur. The inclusion of this effect in the price computed using the HMM approach will thus contribute more to the value of the swaption for longer maturities, and thus

account for the greater difference between prices computed with and without hidden state variables.

Although this reasoning can explain the difference between the prices computed using the HMM approach and those from the trinomial tree, this does not explain the reason for the very high prices computed using the lattice method of Bollen (1998). Further investigation showed that the reason for these results was that the size of an up move for the high-volatility state was very large, with the swap rate doubling over a single day. For a lattice with many time steps, as was used when swaption prices for 365-day swaptions were computed, this results in extremely large values for the swap rate at expiry of the option. This, in turn, results in very large values of the swaption, at the final time step. These large values are then carried back through the lattice to the pricing date, during the backwards recursion through the tree. Since the high-volatility state was the less persistent of the two unobserved states, it seems unreasonable that this state should have such a pronounced effect on the price of the swaption.

This swamping effect of the high volatility state on the swaption price suggests that the lattice method should only be used when the volatility in the high-volatility state is relatively low, or the maturity of the option being priced is relatively short. These conditions would reduce the impact that the high-volatility state has on the option values at expiry, by generating large values of the underlying variable. If these conditions do not hold, then prices generated using this method are likely to be unreliable due to the distorting effect of the high volatility state on the option values.

It is also interesting to note that the prices computed using the trinomial tree approach, without using hidden states, do not appear to find evidence in favour of early exercise of the swaptions. This can be seen from the fact that the prices, for each strike rate, are the same for a given maturity no matter how many possible exercise dates there are. Thus, the price does not depend on the number of possible exercise dates and so the possibility of early exercise has no value. This could only be the case if early exercise of the swaptions was never optimal. Otherwise the existence of more possible exercise dates would have value to the holder of the swaption and this value should then be expressed in higher swaption prices for swaptions with more possible exercise dates. In contrast to this result, both pricing methods that took account of unobserved state behaviour produced prices that vary with the number of possible exercise dates. In the case of the lattice method, this is only true for swaptions with shorter maturities. This is probably due to the swamping effect of the high option values generated at expiry in the larger lattices, used for pricing swaptions with longer maturities.

The result that only swaptions priced taking into account the effect of unobserved states show evidence of optimal early exercise is probably due to the effects of the high-volatility state. A higher volatility level would result in higher simulated future market swap rates, which would result in a greater probability that early

exercise of the Bermudan swaption would be optimal. This is because early exercise of the option is optimal when the value that could be gained by immediate exercise, being a function of the difference between the market swap rate and the strike swap rate, is larger than the option price at the exercise date. The higher the market swap rate, the more likely, then, that it will be optimal to exercise the swaption at the given early exercise date.

Below are graphs of swaption prices, for differing strike rates and for the different methods of calculation, for given swaption maturities and numbers of possible exercise dates. A more comprehensive set of such graphs, together with the prices presented above can be found on the CD-ROM accompanying this work, in the "Swap Output" file.

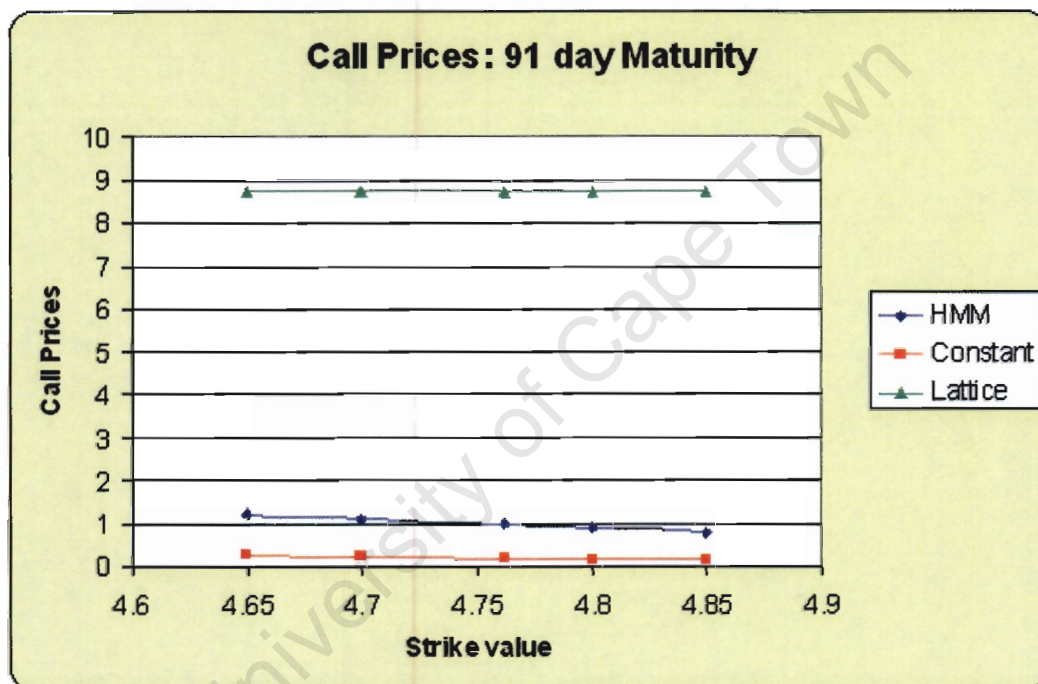


Figure 36: Swaption prices for 91-day swaptions under the 2-state model of the swap rate with 6 possible exercise dates

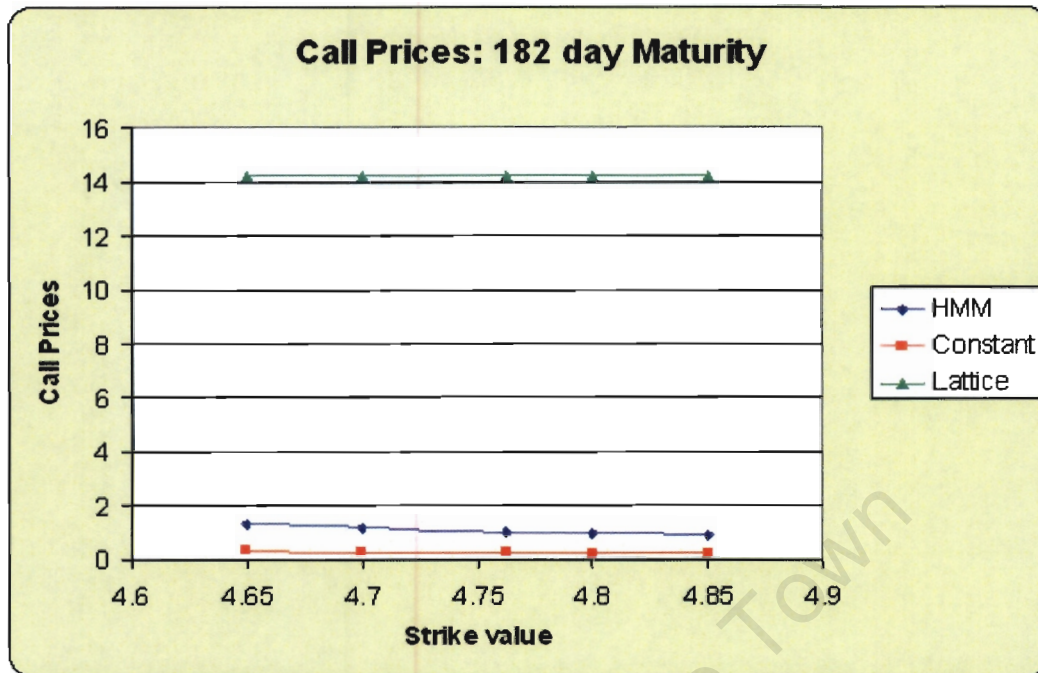


Figure 37: Swaption prices for 182-day swaptions under the 2-state model of the swap data with 6 possible exercise dates

As was to be expected, it can be seen that the options with a lower strike rate had a higher price than those with a higher strike rate. In-the-money options have more value to the holder than out-of-the-money options and this is reflected in a higher price. The values that the model is producing are thus consistent with what would be expected. The very high values produced by the lattice method for the swaption prices does make the effect of strike rate on price slightly difficult to make out, given the scale that this results in for the y-axis. It is just noticeable, however.

#### 12.4.2 Swaption prices computed using the 3-state model

	Strike	4.85	4.8	4.7625	4.7	4.65
Maturity	Possible Exercise Dates					
91 Days	3	0.596478	0.667198	0.723875	0.825314	0.912734
	6	3.173027	3.386278	3.538421	3.765473	3.910978
60 Days	2	0.39996	0.447304	0.485242	0.553133	0.611635
	3	0.380355	0.433854	0.477507	0.557152	0.627151

Prices computed with no hidden states using the trinomial tree:

	Strike	4.85	4.8	4.7625	4.7	4.65
Maturity	Possible Exercise Dates					
91 Days	3	0.134511	0.158998	0.177363	0.214261	0.245484
	6	0.134511	0.158998	0.177363	0.214261	0.245484
60 Days	2	0.097517	0.121073	0.138739	0.176007	0.207918
	3	0.097517	0.121073	0.138739	0.176007	0.207918

Prices computed with no hidden states using a finite difference method:

	Strike	4.85	4.8	4.7625	4.7	4.65
Maturity	Possible Exercise Dates					
91 Days	3	0.13488	0.155255	0.170537	0.212381	0.245857
	6	0.13488	0.155255	0.170537	0.212381	0.245857
60 Days	2	0.097932	0.116002	0.129554	0.173339	0.208366
	3	0.097932	0.116002	0.129554	0.173339	0.208366

As was found to be the case for prices generated using the 2-state model, the prices that include the effect of hidden states are higher than those that do not. This is probably due to the fact that the extra information included in the hidden states contains value, which is then reflected in the higher option prices that take account of this information. The difference in values is not as severe as in some of the cases for the 2-state model prices, but this is due to the fact that the maturities of the swaptions considered here are shorter than those priced using the 2-state model. This would result in less time being available for the hidden states to influence the price of the swaption, by changing value.

Even for swaptions with a comparable maturity, the prices generated by the 3-state model differ by less from the swaption prices calculated without taking into account any hidden state effects than those of the 2-state model. This could be due to the influence of State 3, which was very persistent and had a very long expected duration. If there was very little change in the value of the underlying state variable, with the state process remaining in State 3 for long periods of time, then the price of the swaption should more closely resemble the price of an option with no hidden state effects. This is because the long sojourns in State 3 would effectively produce behaviour that was similar to that of a time series with no hidden state changes. Any difference in price between this pricing method

and one that does not take account of any hidden state effects would then be due to differences in the parameter values used for these two pricing methods, if any. Notice, however, the large difference for 91-day swaptions with 6 possible exercise dates. This would seem to indicate that an early exercise of the option is providing substantial value to the holder, a fact that the pricing method that does not employ hidden states fails to identify.

As previously, the swaption prices computed without taking into account any hidden state effects, using a trinomial tree and finite difference methods, again provide no evidence that early exercise of the swaptions is optimal prior to expiry. This is because prices computed using this method do not vary as the number of possible exercise dates changes, for each swaption. The prices computed using the HMM method, on the other hand, do vary with the number of possible exercise dates, thus suggesting that early exercise of the swaptions may be optimal and that this value to the holder of the swaption is then included in the price. As before, this discrepancy between the prices generated by the HMM method and those that do not take account of hidden state effects can probably be explained by the influence of the higher volatility states. The influence that these states have on the simulated swap rates used to price the option with the HMM method, raising the probability of early exercise of the swaption, produces this result. This would suggest that not taking into account this effect results in underpricing of the swaptions, which will be subject to periods of higher volatility in the market. This then provides an argument in favour of using the Hidden Markov framework to price Bermudan swaptions.

As before, graphs of swaption prices for varying strike rates are presented below to determine the effects of differing strike rates on the swaption prices. As is to be expected, prices for swaptions with lower strike rates are higher. The effect is clearer in these graphs than in those for the 2-state model, due to the absence of the distorting effect of the lattice prices. It is interesting to note that the difference in price between in-the-money options and out-of-the-money options is more marked for the prices computed using the HMM method. This may be due to the HMM including information that has value that the other method cannot. Confirmation of this supposition would require a comparison between the prices generated by the HMM method and those available in the market, if any were available. This would allow for determining whether the HMM prices were more accurate.

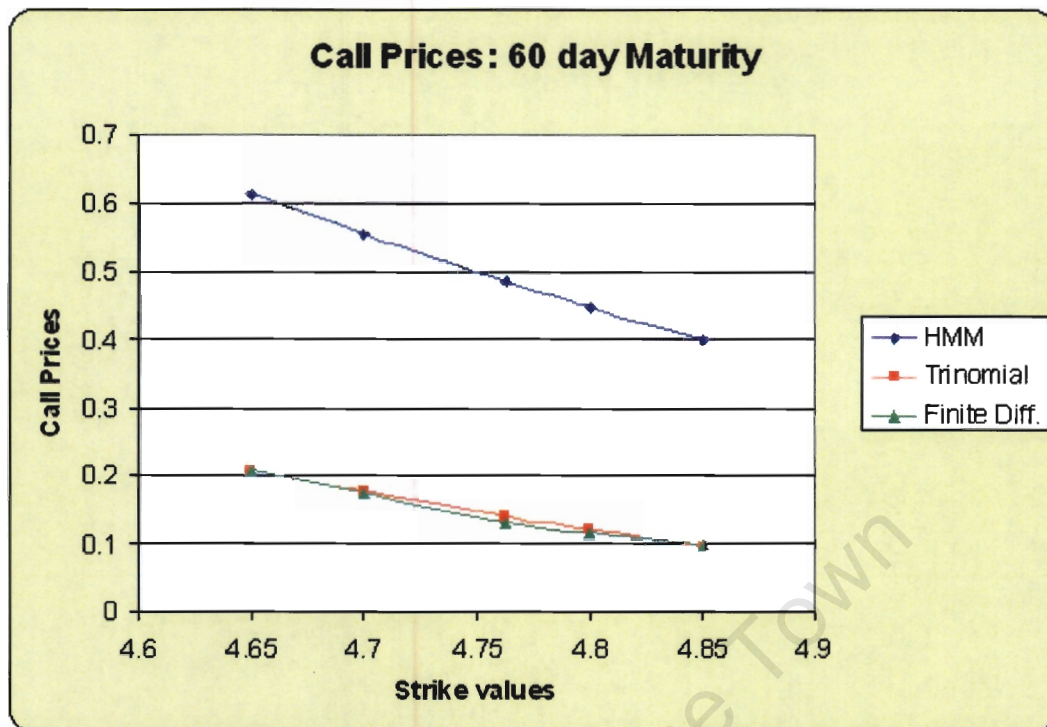


Figure 38: Swaption prices for 60-day swaptions under the 3-state model of the swap data, with 2 possible exercise dates

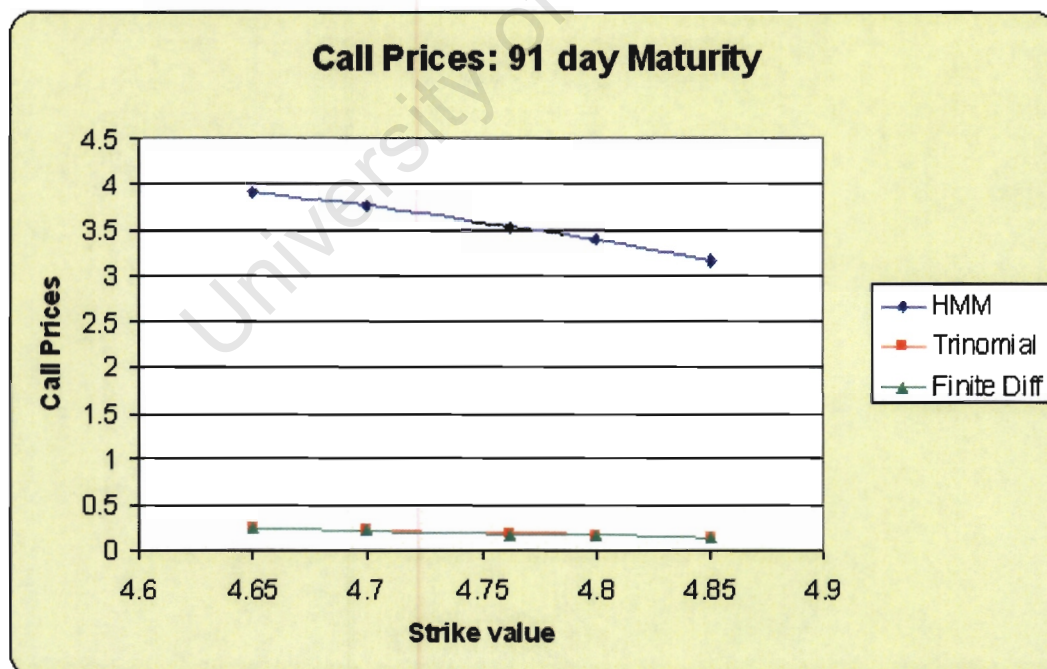


Figure 39: Swaption prices for 91-day swaptions under the 3-state model of the swap data with 6 possible exercise dates

## Chapter 13. Pricing Options on the JSE and FTSE 100 indices using an HMM with Time-varying Transition Probabilities

The results of the option pricing in Chapter 11 exhibited some odd behaviour, particularly in the case of the implied volatilities that were calculated. This may have been due to the fact that the models of the indices used there did not adequately capture the dynamics of the indices that the options were based on. To investigate whether this was in fact the case, the options will be re-priced using a model with time-varying transition probabilities, as outlined in Chapter 7. By allowing the transition probabilities of the Markov Chain component of the Hidden Markov Option Pricing Model to vary over time, it is possible that further dynamic effects in the index series may be included in the model. This, in turn, could lead to better pricing of the options and perhaps a correction of the abnormal results for the implied volatilities of the options.

To determine whether the above hypothesis is correct, the same options as were specified in Chapter 11 will be priced using time-varying transition probabilities. As before, prices for options using several combinations of strike prices and initial index values for each option maturity will be computed, to determine whether the model incorporates the volatility smile effect. To this end, implied volatilities for all the options priced will be computed and compared. As in Chapter 11, the graph of implied volatilities against the strike price of the options should resemble a smile shape. Note that, as before, it is assumed that the index either pays no dividends, or all dividends are immediately reinvested in the index as they are paid. In this way, all returns to the investor are due to changes in the index value and the option pricing formula does not need to take account of dividend payments.

### 13.1 Estimating the Model Parameters and Occupation Time Probabilities

The form of the model chosen for the transition probabilities is similar to that proposed by Tzavalis and Chourdakis (1999). However, since there is no model for the mean rate of return, there will be no term containing an error component in the model used here. The equations used are thus:  $p_{ij,t} = g(\xi_t) = N(\xi_t)$  where  $N(\cdot)$  is the cumulative standard normal distribution function and  $\xi_t = a_{ij} + k_{ij}(p_{ij,t-1})$ . This is similar to a GARCH model for volatility, in that the current value of the transition probability,  $p_{ij,t}$ , is a function of a linear combination of past values of the  $p_{ij}$ . To initialise the model, it is assumed that there is an

equal probability of remaining in a given unobservable state, or moving to any of the other possible states. Thus, if there are  $N$  possible states, then the transition probabilities are set to  $p_{ij,0} = \frac{1}{N}$ , for all  $i$  and  $j$ .

The fact that the transition probabilities are no longer constant over time complicates an evaluation of the maximum likelihood estimates for the model parameters using the EM algorithm. This is largely due to the fact that the maximisation step of this method involves a very complicated expression for the parameters relating to the transition probabilities, i.e. the  $a_{ij}$  and  $k_{ij}$  values. It is thus proposed that the log-likelihood be directly evaluated and then numerically maximised. This can be done using the calculation of the 'filtered' probabilities that the unobserved state, at a given time point, was a particular value. As was mentioned in Chapter 10, these calculations, presented in (33), can be used to evaluate the log-likelihood directly. This function was then maximised using the nonlinear minimisation routine in R to determine the maximum likelihood estimates of the parameters and the related standard errors. The maximum was found by minimising the negative log-likelihood.

As was discussed in Chapter 7, the occupation times and occupation time probabilities can no longer be directly computed using (21), (22), (23), (24) and (25). This is due to the nonlinear nature of the function  $g(\cdot)$  and the serial dependence of the  $p_{ij,t}$  induced by the choice of function for  $\xi_t$ . The occupation time probabilities required for option pricing will thus be computed using the Monte Carlo simulation method proposed by Tzavalis and Chourdakis (1999). The occupation times were simulated 20 000 times and the simulated occupation time probabilities were then computed. Thus, in the notation of Chapter 7,  $M = 20\ 000$ . Using the above methodology, the following results were then obtained.

## 13.2 Results for the JSE Data

### 13.2.1 Parameter Estimates for the 2-state Model of the JSE Data

	Estimated Values	Standard Errors
$\mu_1$	0.288011	0.014208
$\mu_2$	-0.184855	0.028127
$\sigma_1$	0.150612	0.015006
$\sigma_2$	0.364129	0.030316
$a_{11}$	1.380874	2.203597
$k_{11}$	0.605539	2.231286
$a_{22}$	1.235256	-
$k_{22}$	0.266374	-

AIC	-16356.86
BIC	-8154.966
Interest Rate: State 1	0.07
Interest Rate: State 2	0.1

As was the case for the model with constant transition probabilities, one of the unobservable states was associated with a positive mean rate of return and one with a negative mean rate of return. The model thus includes both bull and bear market effects. Further, the state with a lower mean rate of return associated with it also had the higher standard deviation of return associated with it, thus incorporating the leverage effect. Parameters for the transition probability equations were only estimated for the probability of remaining in each state,  $p_{ii}$ , since the remaining transition probabilities can be computed using the fact that the probabilities must add up to one, for each state. The estimates for the parameters relating to State 1 have very large standard errors, whilst the standard errors for the parameters relating to State 2 could not be computed. This suggests that the estimates for the parameters relating to the transition probabilities are not statistically significant and that the 2-state model with constant transition probabilities is thus a better model of the data.

However, both the AIC and BIC values are lower for the model with time-varying transition probabilities, as above. This result is particularly pronounced for the BIC value, at -8154.966 compared to -5850.799. It is thus suggested that the choice of model used should depend on the behaviour of the option prices calculated using the model. It may be the case that the dataset used was not large enough to produce statistically significant parameter estimates. It could then still be the case that the model with time-varying transition probabilities still incorporates enough additional information so as to produce 'better' option prices than the model with constant transition probabilities.

### 13.2.2 Option prices computed using the 2-state model of the JSE Data

#### Call Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
HMM Prices	0.3802	3.4597	10.6154	1.2235	5.1377	11.9694	2.1375	6.5362	13.2651
B-S Prices	0.5517	3.8254	10.7258	1.559	5.5708	12.1736	2.5775	7.0142	13.5197

#### Put Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
HMM Prices	9.6891	2.7691	0.4253	9.84297	3.7571	1.0888	10.0496	4.4482	1.6772
B-S Prices	9.978	3.2517	0.6521	9.9780	3.2518	0.6521	9.9780	3.2517	0.6521

The Black-Scholes prices were calculated using a volatility value that was a weighted average of the two volatilities associated with the unobserved states. The weights used were the smoothed probabilities that the unobserved state, at

the final time point in the sample, was that associated with the given volatility value. It is interesting to note that the call option prices were all lower than the corresponding Black-Scholes prices, whilst this is only the case for put options of shorter maturities. Longer dated put options have higher prices than the Black-Scholes equivalent. This could, as in the case of the model with constant transition probabilities, be due to the effect of the higher volatility state on the option prices. The longer dated options would allow more opportunities for the unobserved state value to change. This, in turn, would result in the index value spending more time under the influence of the higher volatility state, which would lead to higher option prices. This is because higher volatility in the value of the variable underlying the option contract implies a greater probability that the value of that variable, at the exercise date of the option, would be such that the option would be in-the-money. This represents value to the holder of the option and thus results in a higher price. Since the Black-Scholes prices are based on a constant volatility value, and thus do not include the effects of the higher volatility state, they are lower than the HMM prices that do include this effect.

As was the case for the model with constant transition probabilities, the implied volatilities for options with various combinations of strike price and initial index values were computed. This was done using Microsoft Excel Solver, as before, to solve for the volatility value in the Black-Scholes equation. The graphs of the resulting values should resemble a smile-like shape, due to the higher value placed on out-of-the-money options by market participants. The results for the 2-state model of the JSE data are presented below.

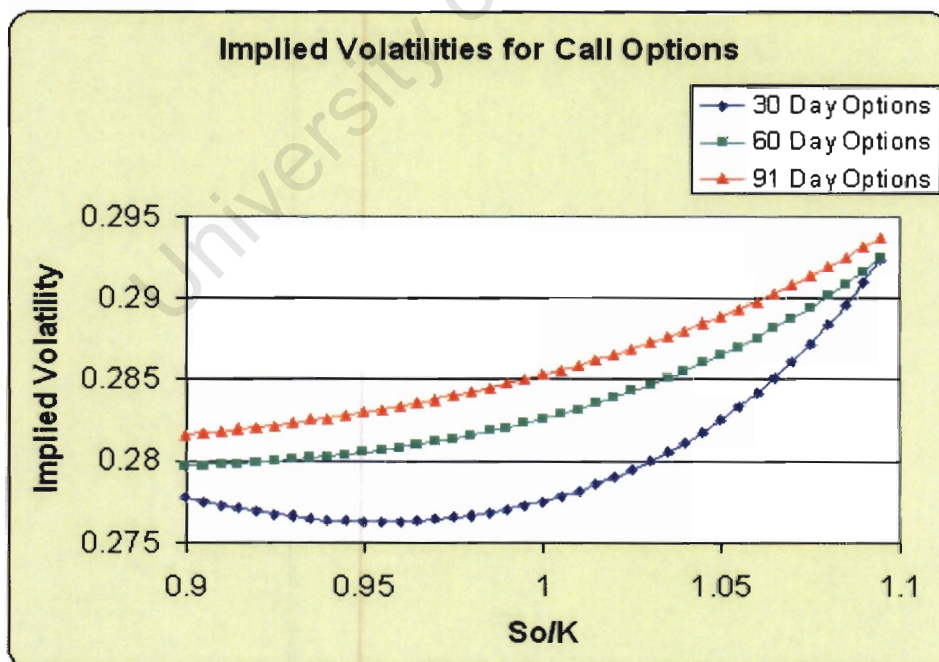


Figure 40: Implied Volatility for Call Options under the JSE 2-state model

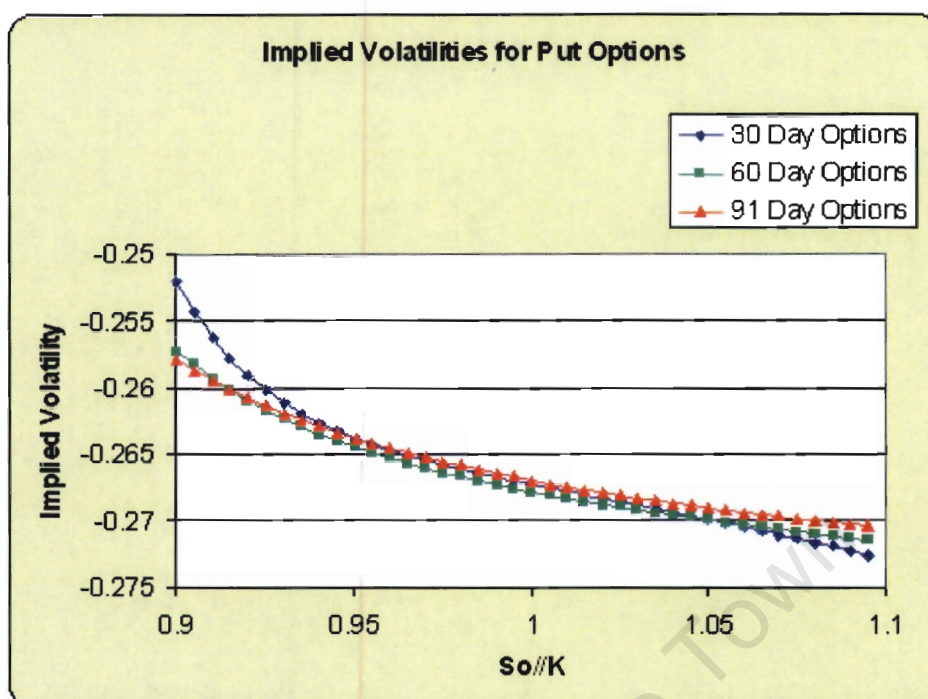


Figure 41: Implied Volatility for Put Options under the JSE 2-state model

The graph of the implied volatilities for call options follows a textbook smile pattern, with the highest implied volatilities being calculated for options with the lowest and highest strike prices, relative to the initial index value. The volatilities for the put prices exhibit a skew-like behaviour, which could also be expected from market prices. However, the implied volatilities are all negative, which is nonsensical since volatility is always a positive value. Upon further investigation, it was found that the values calculated for the implied volatilities depended quite strongly on the values used to start the optimisation used to compute them. It is possible to find a positive set of implied probabilities for these option prices, which are presented in Figure 42 below. However, the behaviour that these implied volatility values exhibit is still very unorthodox, with little similarity to the expected smile effect. This suggests that the HMM method used to price these options, with time-varying transition probabilities, prices call options better than it does put options. A similar conclusion could be drawn from the implied volatility graphs for the prices calculated using constant transition probabilities, on the basis of the implied volatility performance.

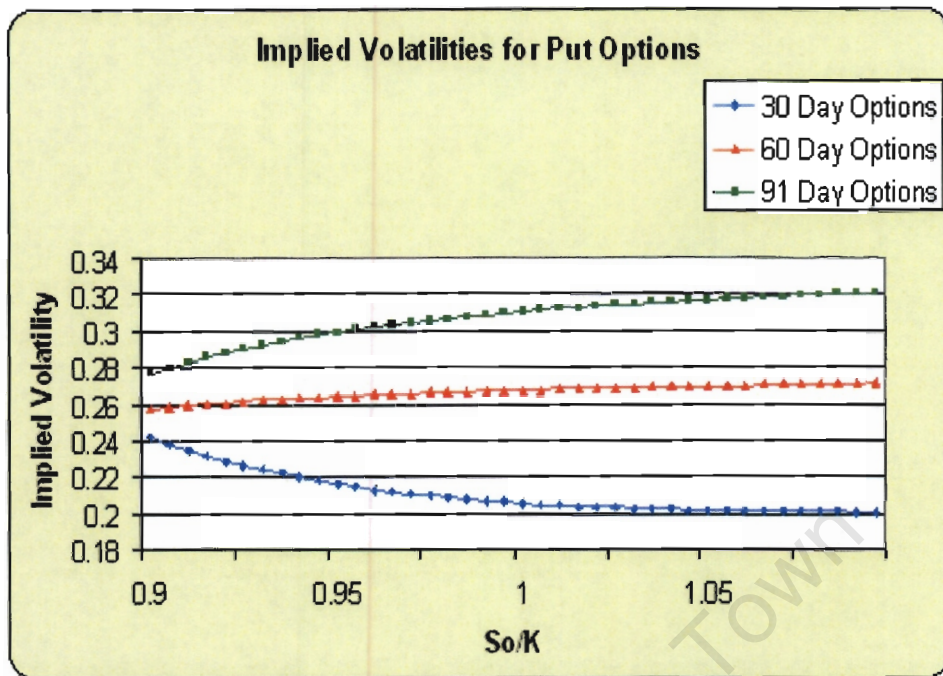


Figure 42: Implied Volatility for Put Options under the JSE 2-state model, positive solution

### 13.2.3 Parameter Estimates for the 3-state model of the JSE Data

	Estimated Values
$\mu_1$	0.7407566
$\mu_2$	0.73
$\mu_3$	6.9522006
$\sigma_1$	0.2146132
$\sigma_2$	2.2925968
$\sigma_3$	1.03E-04
$a_{11}$	0.008
$k_{11}$	0.01
$a_{12}$	0.002
$k_{12}$	0.07
$a_{21}$	0.002
$k_{21}$	0.05
$a_{22}$	0.004
$k_{22}$	0.03
$a_{31}$	0.003
$k_{31}$	0.05
$a_{32}$	100.00288
$k_{32}$	0.0327239

AIC	-15637.9
BIC	-7766.154
Interest Rate: State 1	0.1
Interest Rate: State 2	0.13
Interest Rate: State 3	0.07

There are no standard errors available for these estimates, as the Observed Information Matrix was singular. The mean rates of return that were estimated are all very large, particularly that related to State 3. The estimated volatilities are also fairly large, with the largest value being associated with State 2, which also had the lowest mean rate of return. Thus, there is some incorporation of a leverage-type effect. It is unusual, however, that none of the unobserved states has a negative mean rate of return associated with it. This would suggest that there were no bear market conditions experienced over the entire sample. Since the sample covered 10 years of data, with two known bear market phases, this result appears nonsensical. The fact that the parameters associated with the probability of a move from State 3 to State 2 suggest that such a move is highly likely is also odd. This is because State 2 is the highest volatility state, whilst State 3 is the lowest volatility state. The model thus appears to be suggesting that the market will jump directly from the lowest volatility state to the highest volatility state whenever it enters the low volatility state.

The smoothed probabilities of being in a given unobserved state, at a given time point also exhibit strange behaviour. They suggest that the probability of being in

State 1 is 1 for virtually the entire sample, with the only exceptions being found right at the end of the sample. This is a very strange situation, suggesting no change in the value of the unobserved state value over almost the entire sample of data. Given that a relatively large sample was used, relating to a relatively long time period, this is unlikely to be the case. Altogether, this 3-state version of the model appears to exhibit very strange behaviour, which might suggest that the 2-state model of the data would be a better choice. This is confirmed by the fact that both the AIC and BIC for this version of the model are higher than those for the 2-state version, which suggests that that is the better model. Furthermore, the 3-state model of the JSE data using constant transition probabilities also has lower AIC and BIC values than this version of the model, suggesting that if a 3-state model is required, then the version with constant transition probabilities should be used.

### 13.2.4 Option prices computed using the 3-state model of the JSE data

#### Call Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
HMM Prices	2.3208	4.4639	12.6302	2.5946	5.117	13.4414	2.6995	5.5989	14.0163
B-S Prices	0.0019	1.6785	10.0765	0.0479	2.5611	10.6773	0.1802	3.341	11.3189

#### Put Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
HMM Prices	11.6217	3.8277	2.5538	11.3746	3.8991	2.7254	10.8941	3.7936	2.711
B-S Prices	9.4282	1.4171	0.0028	8.9038	1.4171	0.0333	8.4501	1.611	0.0889

As was the case for the 2-state model of the data, the Black-Scholes prices were calculated using a weighted average of the estimated volatilities. The weights were again the smoothed probability of being in the unobserved state associated with the given volatility value. In the case of both calls and puts, the HMM prices are much larger than the corresponding Black-Scholes prices. This is probably due to the effect of the highest volatility state on the HMM prices, since the smoothed probabilities give most of the weight to the volatility associated with State 1, which has the second highest volatility level. Under the Black-Scholes assumptions, this volatility level remains constant over the life of the options being priced. The influence of the volatility levels associated with the other unobserved states is thus not taken into account by the Black-Scholes prices, including the highest volatility.

The estimated parameters suggested that this model includes some unusual behaviour and is in fact inferior to the 2-state model of the data. This may also account for the large departure from the Black-Scholes prices. The graphs of the implied volatilities, calculated for the option prices presented above, are shown below. These also exhibit some strange results.

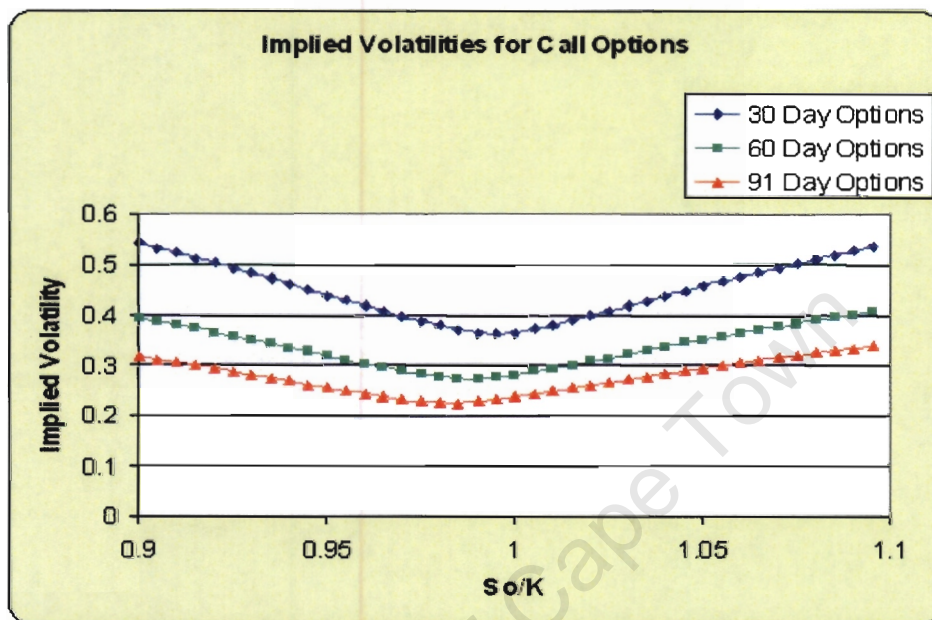


Figure 43: Implied Volatility for Call Options under the 3-State model of the JSE data

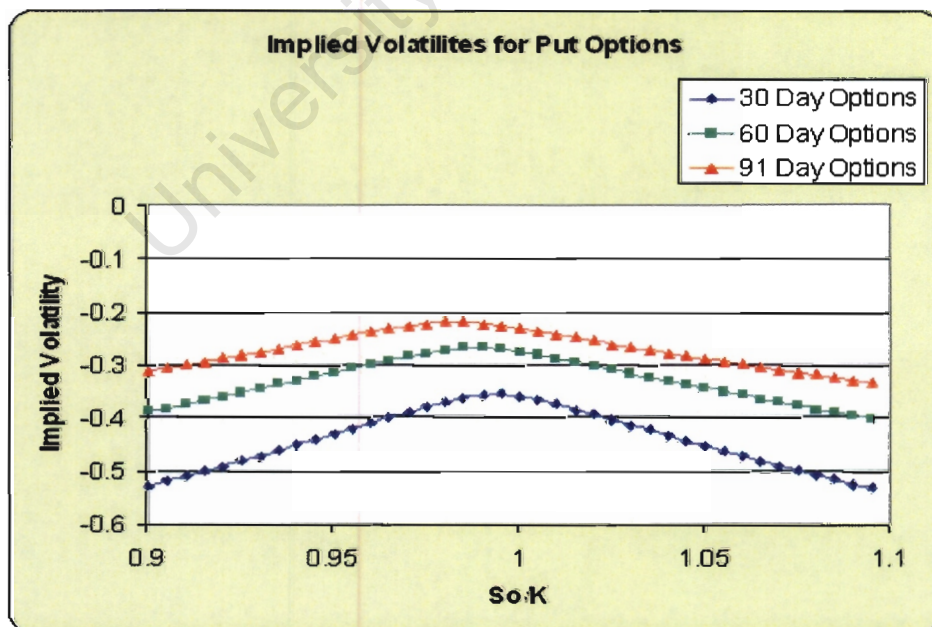


Figure 44: Implied Volatility for Put Options under the 3-State model of the JSE data

The implied volatility curves for the call prices do exhibit a smile-like behaviour. However, the options with the shortest maturity have the highest levels of implied volatility, whilst the options with the longest maturity have the lowest levels of implied volatility. This is completely contrary to what would be expected. This is because, all other things being equal, options with longer maturities should have higher prices than options with shorter maturities. This would translate into higher implied volatilities for options with longer maturities, which is the complete opposite to the pattern observed above. As was the case for the put options under the 2-state model, the implied volatilities for the put options here are again negative. This suggests that the model still prices call options better than put options, the additional state notwithstanding. The pattern of implied volatilities for the put options is more smile-like than was the case for the 2-state model; although a 'negative smile' is observed since the implied volatilities are negative. Again, this may be due to the computation process, as was noted for the 2-state case.

On balance, it is suggested that the 2-state model of the data is 'better', both in that it produces less odd results such as very large mean returns and volatilities of returns, and due to the AIC and BIC values. The option prices produced by the 2-state model also exhibit a less extreme departure from the Black-Scholes values, which may suggest that they are more reasonable. They include the extra information relating to the unobserved states that the Black-Scholes values do not, without being radically different from the Black-Scholes values, which would be unexpected.

### 13.3 Results for the FTSE 100 Data

#### 13.3.1 Parameter Estimates for the 2-state model of the FTSE 100 Data

	Estimated Values	Standard Errors
$\mu_1$	0.2045631	0.013562
$\mu_2$	-0.138872	0.0223213
$\sigma_1$	0.1299005	0.01552
$\sigma_2$	0.2983439	0.0231647
$a_{11}$	1.5858163	2.3839882
$k_{11}$	0.7257232	2.3975779
$a_{22}$	1.5644333	-
$k_{22}$	0.5828674	-

AIC	-16739.47
BIC	-8346.268
Interest Rate: State 1	0.04
Interest Rate: State 2	0.05

As was found for the 2-state model of the JSE data, the above estimates include both bull and bear market effects. These are accounted for by the positive and negative mean rates of return that are associated with the respective unobserved state values. The state with the negative mean rate of return associated with it, State 2, also has the higher volatility of returns associated with it in keeping with the leverage effect. As was also observed in the case of the JSE data, the standard errors of the estimates of the parameters relating to the transition probabilities are very large, or do not exist. Again, this may suggest that the inclusion of time-varying transition probabilities in the model is not optimal. This is confirmed by the fact that the AIC and BIC values for the 2-state model of the FTSE 100 data with constant transition probabilities are lower than those computed for the model above. Overall, then, the 2-state model of the FTSE 100 data using constant transition probabilities is a better model of the data than the model with time-varying transition probabilities.

### 13.3.2 Option Prices calculated using the 2-state model of the FTSE 100 data

#### Call Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
HMM Prices	0.16142	2.7608	10.094	0.6494	4.0423	10.9484	1.2349	5.0947	11.8176
B-S Prices	0.02715	2.0785	9.8754	0.2073	3.0367	10.3953	0.4973	3.8345	10.974

#### Put Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
HMM Prices	9.7935	2.3934	0.2268	9.9146	3.3074	0.7135	10.1217	3.9815	1.2044
B-S Prices	9.6989	1.7503	0.0472	9.5519	2.3814	0.2399	9.50501	2.8422	0.4817

Both call and put option prices, for all combinations of strike prices and maturities are higher than the Black-Scholes equivalent values. The differences appear to be larger for out-of-the-money options than for in-the-money options, which may be due to the influence of the volatility smile. This effect is largely produced by market option prices being higher than the Black-Scholes equivalent value for out-of-the-money options. Confirmation of this speculation can be obtained by viewing the graphs of implied volatility, calculated for the options above, which are presented below. The fact that the option prices are higher than the Black-Scholes equivalent values, as before, is likely to be due to the effect of the high volatility state on the HMM prices. Higher levels of volatility, which come into effect as the unobserved state changes to the high volatility state, make the options more valuable by increasing the probability that they will expire in-the-

money. This is reflected in the higher option prices. Since the Black-Scholes prices are calculated based on a constant level of volatility, they do not include this effect and so are lower.

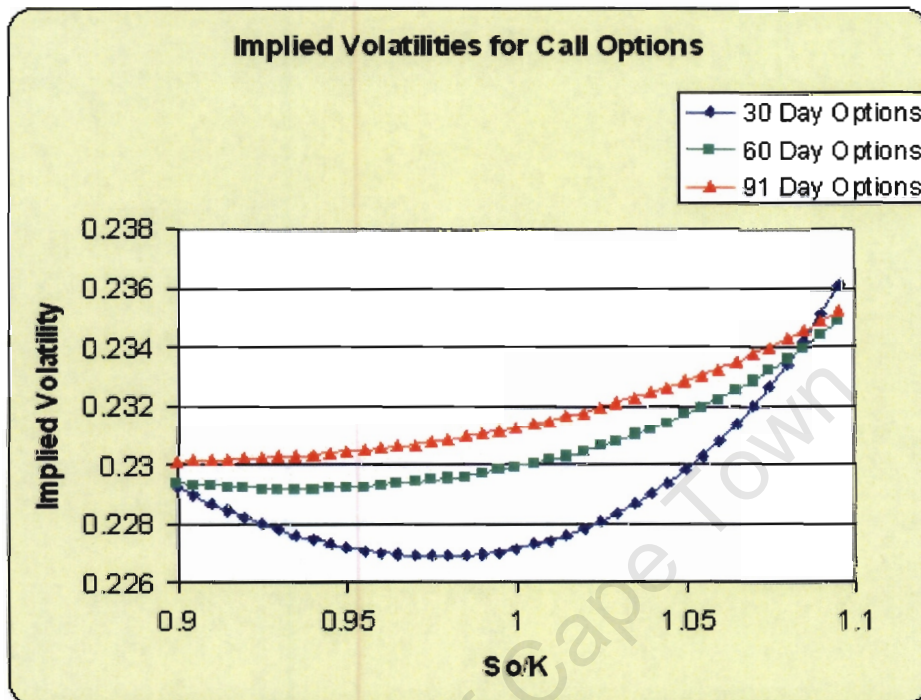


Figure 45: Implied Volatilities for Call Options under the 2-state model of the FTSE 100 Data

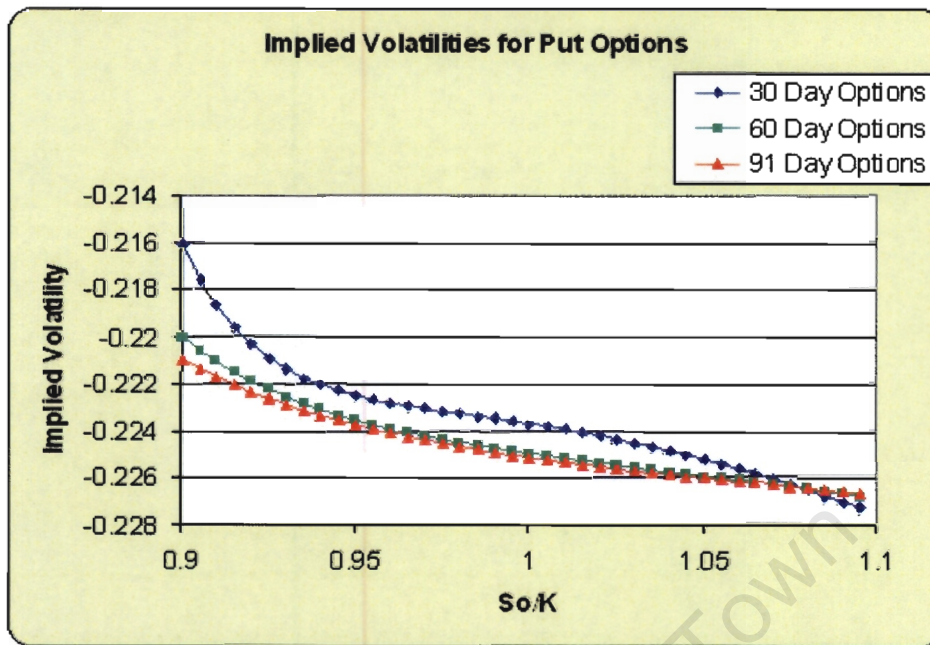


Figure 46: Implied Volatilities for Put Options under the 2-state model of the FTSE 100 Data

As was the case for the JSE data, the implied volatilities for the call options produce a classic smile pattern. This confirms the suspicion that the higher HMM option prices for options that were out-of-the-money were due to the volatility smile effect.

### 13.3.3 Parameter Estimates for the 3-state model of the FTSE 100 Data

	Estimated Values
$\mu_1$	0.0687613
$\mu_2$	0.73
$\mu_3$	0.2158817
$\sigma_1$	0.2131987
$\sigma_2$	2.2925968
$\sigma_3$	2.816503
$a_{11}$	0.008
$k_{11}$	0.01
$a_{12}$	0.002
$k_{12}$	0.07
$a_{21}$	0.002
$k_{21}$	0.05
$a_{22}$	0.004
$k_{22}$	0.03
$a_{31}$	0.003
$k_{31}$	0.05
$a_{32}$	100.00268
$k_{32}$	0.0327197

AIC	-16008.806
BIC	-7951.6062
Interest rate: State 1	0.06
Interest rate: State 2	0.04
Interest rate: State 3	0.05

As in the case of the JSE data, the Observed Information Matrix was singular for the 3-state model of the FTSE 100 data and thus standard errors are not available for these estimates. The same odd behaviour, with very high mean returns and volatility of returns, as was observed for the 3-state model of the JSE data are present in the above results as well. This, together with the fact that the AIC and BIC values are higher than for the 3-state model of the data with constant transition probabilities suggests that this is not a good model of the data. The AIC and BIC values are also higher than those of the 2-state model of the data, with time-varying transition probabilities, suggesting that this is the worst possible model for this data, out of those considered. It is thus suggested that the 2- or 3-state model with constant transition probabilities should be used to model this dataset. This is because the 2-state model with time-varying transition probabilities has a higher AIC and BIC value than the corresponding model with constant transition probabilities. This suggests that although the 2-state model with time-varying transition probabilities is a better model of the data than the 3-state model, it is still inferior to the corresponding model with constant transition probabilities.

### 13.3.4 Option Prices calculated using the 3-state model of the FTSE 100 Data

#### Call Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
HMM Prices	12.784	16.9116	21.2422	19.402	24.0756	28.8035	24.9286	30.0639	35.1752
B-S Prices	5.4275	10.1635	16.1066	9.2071	14.4099	20.3703	12.2724	17.7641	23.8012

#### Put Option Prices

So/K:	30 Day Options			60 Day Options			91 Day Options		
	0.9	1	1.1	0.9	1	1.1	0.9	1	1.1
HMM Prices	19.5188	16.643	14.32	26.0897	23.5267	21.3801	31.5666	29.2061	27.1966
B-S Prices	15.0993	9.83532	6.2784	18.5517	13.7546	10.2149	21.2801	16.7718	13.3089

Both the call and put prices are significantly higher than the Black-Scholes equivalent values. This is most likely due to the influence of the very high volatility states on the HMM prices. This is because the volatility used for the Black-Scholes computations was a weighted average of the volatilities estimated for the various unobserved states, with the weights being the smoothed probability of the index being under the influence of the given state, at the final time point in the sample. However, these probabilities gave virtually all of the weight to State 1, which had the lowest level of volatility associated with it. This was then assumed to be the constant level of volatility in the Black-Scholes computations. The high volatility levels associated with State 2 and State 3 thus had virtually no impact on the Black-Scholes prices. They would have had an influence on the HMM prices, however, resulting in these prices being much higher than the corresponding Black-Scholes values.

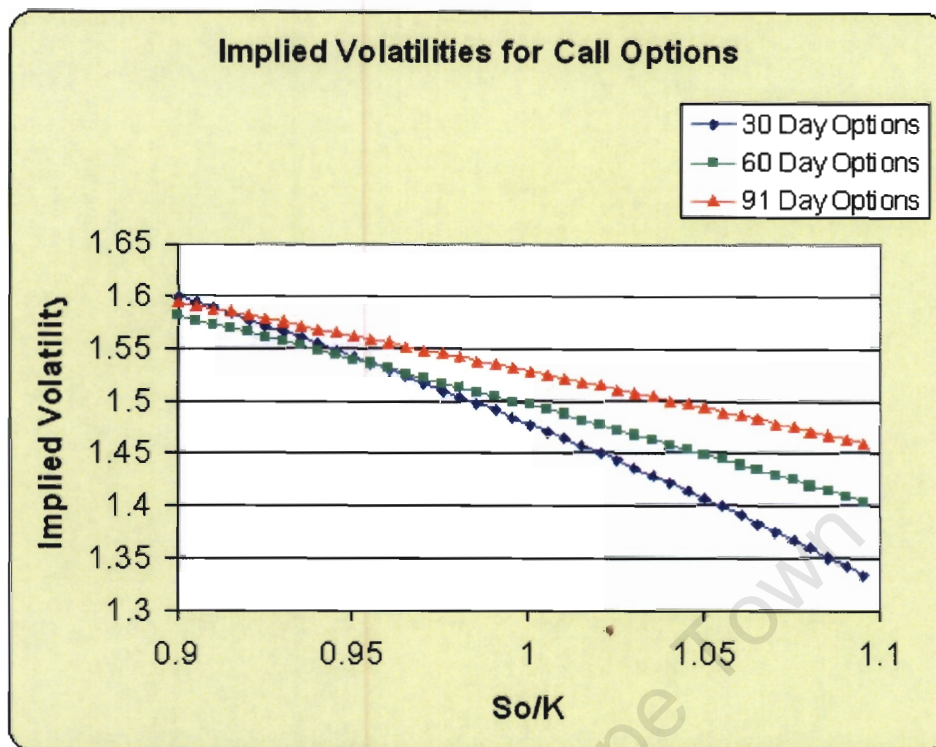


Figure 47: Implied Volatilities for Call Options under the 3-state model of the FTSE 100 Data

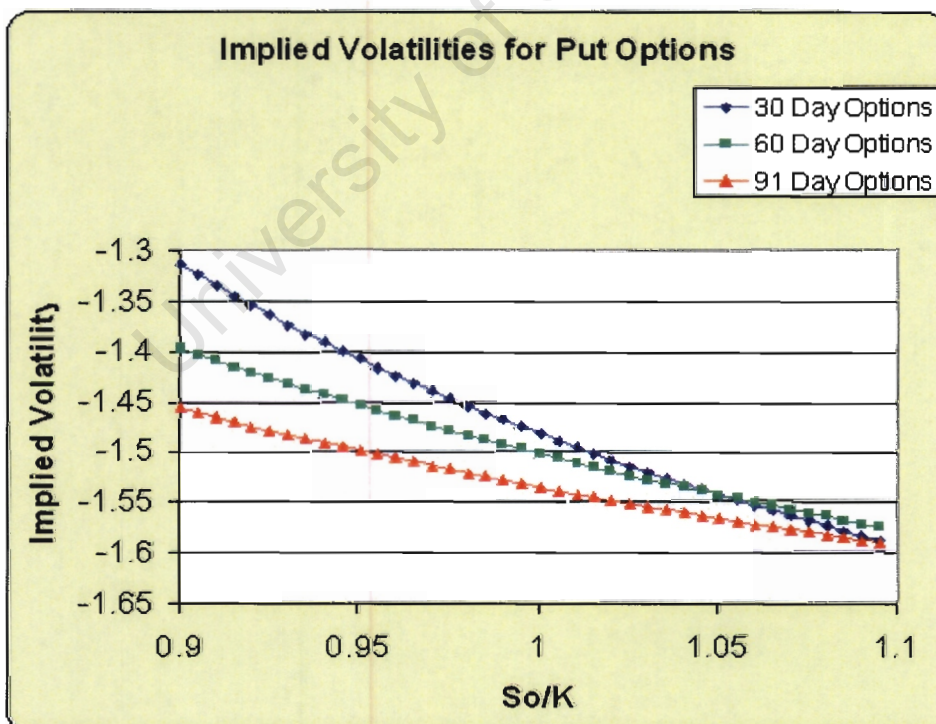


Figure 48: Implied Volatilities for Put Options under the 3-state model of the FTSE 100 Data

The graph of implied volatilities for the call options does resemble a skew, which could possibly be expected from market prices. The volatility values are all very high, however, due to the large difference in prices between the HMM and Black-Scholes prices for these options. The put options again have negative implied volatilities which seems to suggest that the model does not price put options very well. The 'better' behaviour of the implied volatilities for the 2-state model of the data, in terms of producing implied volatilities more consistent with the volatility smile effect, adds weight to the conclusion that that is a better model to use than the 3-state model for this data.

## 13.4 Some Conclusions

Overall, when reviewing the results above, it appears that the inclusion of time-varying transition probabilities into the model of the index returns is unnecessary. The large standard errors produced for the parameters relating to the transition probabilities, when these were available, as well as the larger AIC and BIC values, relative to the models with constant transition probabilities, all support this view. The implied volatility curves for the 2-state model of both the JSE and FTSE 100 data might appear to be more consistent with the volatility smile effect, but this is only the case for call options. The implied volatilities for the put options exhibited some very odd behaviour, which was not always the case for the models with constant transition probabilities.

Furthermore, this is a much more complicated version of the model than that with constant transition probabilities, with more parameters to be estimated. It also requires the occupation time probabilities to be calculated using a Monte Carlo simulation due to the nonlinear equations used to compute the transition probabilities. The model would thus have to perform considerably better than the version with constant transition probabilities to be selected instead of that model. This would be in the interests of retaining computational simplicity and a parsimonious model framework. This was not found to be the case here, and so the model with constant transition probabilities is recommended as the 'best' model to use for these datasets.

## Chapter 14. Conclusions

### 14.1 Advantages in using the Model

The Hidden Markov Option Pricing Model has several features that recommend it as a good model for pricing options. Key among these is the fact that it allows the parameters that are input into the option pricing formula to vary over time. This resolves a key problem with the Black-Scholes formula for option pricing, which assumes that these parameters remain constant over time. This constancy is not observed in practice, making this a serious deficiency of the Black-Scholes model. There are other methods available for pricing options that accommodate time-varying parameters, such as stochastic volatility models and GARCH option pricing models. Both of these methods only allow the volatility parameter to vary, whilst in practice the risk-free rate of return varies stochastically as well. This effect is incorporated by the Hidden Markov Option Pricing Model.

Furthermore, the stochastic volatility models do not admit a closed-form expression for option prices and must be evaluated using Monte Carlo simulation. This is a disadvantage since this method of computation can be very inefficient. The Hidden Markov approach does result in a closed-form expression for option prices, which is both efficient and easy to understand. The GARCH option pricing models can, under certain conditions, admit closed-form expressions for option prices. They suffer, however, from the problem that all GARCH models face in that they tend to overestimate the persistence of volatility values. This effect, documented by Lamoureux and Lastrapes (1990), is due to the assumption made implicitly by GARCH models that there is only one level of unconditional volatility. The Hidden Markov approach avoids this problem by assuming several levels of unconditional volatility. In the simplest form of the model, this is the model for volatility that is used, i.e. several discrete values of volatility which the asset being modelled switches between, according to the value of the unobservable state variable.

The unconditional distribution of asset returns implied by the Hidden Markov Option Pricing Model also is an advantage of the model. It is a dependent mixture of normal distributions, which can replicate the skewness and leptokurtosis that are often observed in empirical asset return distributions. This means that the model could provide a more accurate description of the asset return series under consideration than a model that assumes that returns are normally distributed, such as the Geometric Brownian Motion model. Details of how a mixture distribution can account for high levels of skew and kurtosis can be found in McLachlan and Peel (2000) as well as in Chapter 4. This feature of the model will be particularly important in a market with negative skewness and leptokurtosis, since a model that does not take these effects into account runs the risk of underestimating the probabilities of extreme negative returns. This is because the high kurtosis and skewness result in a greater probability weighting

to events in the negative tail of the returns distribution. Any model that does not take account of this could produce inaccurate option prices, especially for put options which benefit the holder if asset prices fall below the strike price of the option.

Lastly, a case can be made that any changes in asset prices or events that influence asset prices are discrete in nature. The Hidden Markov Option Pricing Model assumes that the asset return process undergoes discrete structural changes, and that asset prices change in a discrete fashion. It can thus be suggested that the Hidden Markov approach provides a more theoretically appealing model of asset returns than the traditional continuous-change models.

## 14.2 Problems with the Model

The main problem with the Hidden Markov Option Pricing model is the problem of deciding how many possible values there should be for the unobservable state process. This is highly problematic, since the choice of the number of possible states has a large effect on the nature and accuracy of the model. Too many states could result in model overfit, whilst too few could result in a very inaccurate model. There is no agreement in the literature on how to resolve this issue. Some authors suggest using some form of Bayesian approach, often making use of a Dirichlet process to model the parameters that must be chosen, including the number of possible hidden state values. Another possible approach is to try and test various hypotheses about a model that has already been fitted, to determine whether the number of states chosen is correct. The problem here, however, is that there are unspecified parameters under the null hypothesis, due to the fact that the states are unobservable (Krolzig (1997) p141). This results in difficulties in performing the hypothesis tests. A range of possible solutions to this problem has been offered, yet there does not appear to be any agreement in the literature as to which of these to use.

As a last resort, which has been used in this work, several models with varying numbers of possible unobservable states can be fitted, and then compared on the basis of some criterion, such as the AIC or BIC. The option pricing stage of the model does provide a restriction on the maximum number of possible state values, however. The pricing method is only feasible for three or fewer hidden states. This is because of the storage requirement in calculating the occupation times of the model in each of the states, which becomes exponential in the time to maturity of the options being priced when there are more than three possible state values. Thus, models with two and three possible hidden state values were fitted and compared. This somewhat arbitrary restriction does have the advantage of reducing the possibility of model overfit that could arise from using too many possible hidden state values.

A further problem with the model is that, in its most basic form, it can be argued as being a too-simplistic description of the parameter processes. A more complicated model of the parameter values, such as an ARIMA-type model may

be more suitable. This can be accommodated as an extension to the model, which would then use a switching ARIMA-type model to model the parameter processes. The problem with such an approach is that it adds a significant complication to the estimation procedure for the model. This is because the parameter values would then be path dependent, losing the Markov property assumed by the 'base case' of the model. This makes estimation much more difficult. This is because the later parameter values depend on earlier parameter values, which depend on the hidden states, which are unobservable and thus unknown. A solution to this problem has been suggested by Gray (1996), although it will result in a less efficient estimation procedure than that of the 'base case' model. The key point here is that such extensions to the model, in the interests of improving accuracy, may prove to be difficult to implement.

### 14.3 Performance of the Model in Pricing European-style Index Options

The prices generated by the Hidden Markov Option Pricing Model were, for the most part, significantly different to those produced by the Black-Scholes formula. This was largely due to the influence of the hidden state/s associated with higher levels of volatility. Since higher volatility, all other things being equal, leads to higher option prices, it follows that the model prices were generally higher than the Black-Scholes equivalent values. By allowing the volatility parameter to vary over time, the model accounts for the fact that the market does experience periods of higher than average volatility from time to time. These periods of higher volatility, in turn, then have an impact on the value of options written on assets in the market that are 'alive' (i.e. have not expired, or been exercised) during these periods. Thus, the model was able to incorporate information into the prices that it generated, relating to the periods of high volatility, that the Black-Scholes model did not. In this case, by assuming a constant level of volatility over the life of the option, the Black-Scholes model underprices options. This could be of concern to a seller of options, since they would be undercharging for the options that they sold, if they were using the Black-Scholes formula to price these options.

In general, it was also found that option prices for options written on the JSE All Share Index tended to differ more from their Black-Scholes equivalents than those of options written on the FTSE 100 Index. This is probably due to the fact that the FTSE 100 Index is measured in a market based in a developed economy, whilst the JSE Index is based in an emerging market. Emerging markets are typically associated with higher levels of risk than developed economies and thus subject to more extreme changes in behaviour. Evidence of this is found in the higher "High" levels of volatility that were estimated for the JSE than for the FTSE 100. These higher volatility levels will then cause option prices evaluated using them to differ more from a Black-Scholes price calculated using a constant, lower, level of volatility. The greater price differences can also be partially explained by the fact that the South African market is likely to undergo more structural changes, over a given time period, than the UK market.

This is due to the greater overall variability of the South African market, which is a result of the perceived higher level of risk associated with it. This feature of the market will result in more variability in the hidden state value of the Markov Chain component of the Hidden Markov Option Pricing Model. This, in turn, will allow for more opportunities for the High-volatility state to be in force, which will cause it to have a greater impact on option prices than in the case of a less variable hidden state series. This is another way of saying that, by being more variable, the JSE market is less like the constant parameter world assumed by the Black-Scholes model. This results in prices generated taking account of this greater variability being more different from the Black-Scholes equivalents than in a market that is subject to less variability, and is thus more like a market that satisfies the Black-Scholes assumptions, such as the UK equity market.

Prices calculated using the 3-state version of the model were also found to exhibit a greater difference from the Black-Scholes equivalents than prices calculated using the 2-state version of the model. This could be due to the fact that the 3-state version of the model was able to incorporate more information and/or effects than the 2-state version of the model, by virtue of having more parameters. In this case, the extra information or effects would lead to prices that were more different from the Black-Scholes prices, with their assumptions of constancy, than those generated by a model that could not incorporate the extra effects. An alternative explanation is that the model is overspecified and is producing spurious results due to a problem related to model overfit. This question highlights the importance of correctly specifying the number of possible hidden state values in the model. As has already been pointed out above, however, there is no clear way to do this. This reinforces the point already made that this is a key weakness of the Hidden Markov Option Pricing Model. Ultimately, a key test of which version of the model is the 'best' would be to compare the prices that it generates to those observed in the market. The most accurate model would then be selected for further use.

In terms of incorporating the 'volatility smile' effect, the model seemed largely to perform poorly in this regard. There were one or two instances where a near-perfect smile was produced by option prices generated by the model, for a given option maturity and number of hidden states. For the most part, however, it appears that the model was not able to incorporate this effect, using the given datasets. This was particularly true for the case of put options, where some very unusual patterns of implied volatility were observed. In some cases the implied volatilities for some put options were found to be negative, which is a nonsensical result. As has been mentioned previously, further investigation did seem to indicate that this problem might be due to problems with the optimisation used to compute the implied volatilities. The volatility values calculated depended quite strongly on the initial values used to start the optimisation, both for Microsoft's Solver routine and the nonlinear minimisation routine in R. It was possible to generate positive values for the implied volatilities, but the behaviour of these values is still very unusual. In particular, it deviates widely from the expected smile effect to a fairly large degree, as Figure 42 in Chapter 13 shows. Thus, on the basis of implied volatility, the put option prices generated by the Hidden

Markov Option Pricing Model performed poorly, even after the suspect negative implied volatilities had been explained.

At the very least, call options did tend to exhibit a volatility 'skew' effect, which can sometimes be observed in the market. It was also found that the more reasonable implied volatility patterns were produced by options assumed to be written on the FTSE 100 Index and those whose prices were calculated using the 2-state version of the model. This could be due to the fact that both of these cases represent situations where less extreme events and changes in market behaviour might be experienced. This, in turn, could result in option prices that are more likely to conform to 'normal' behaviour, which would result in more conventional implied volatility patterns. Whether the observed volatility behaviour of the model option prices was a good or bad indication would depend on how well the model approximated market option prices. If the volatility behaviour is an indication of a departure from market values and behaviour, then this is a bad sign and vice versa.

#### 14.4 Performance of the Model when used to Evaluate Bermudan-style Option Prices

As was found in the case of the European-style index options, the Bermudan option prices generated by the Hidden Markov Option Pricing Model were all different to those calculated using the assumption of constant parameters. Since the Black-Scholes formula does not apply to Bermudan options, these 'constant parameter' prices were computed using trinomial tree and finite difference methods. Both of these methods produced prices that were very close to each other in value. The model option prices were all higher than the 'constant parameter' prices, which is probably due to the influence of the high-volatility states, as was the case for the European-style index options. This result is also partially due to the fact that the model option prices showed evidence of optimal early (i.e. prior to expiry of the option) exercise, which has value to the holder of the option, whilst the 'constant parameter' prices did not. The extra value that accrues to the holder of the options as a result of being able to exercise them prior to expiry, if and when it was optimal to do so, then results in higher prices for these options. As was the case for the European-style options, the 'constant parameter' methods appear to be under-pricing options. This would be of particular concern to option sellers, who may not be charging enough for the value that they are providing to option buyers, if they use a 'constant parameter' method to calculate their prices.

The lattice pricing method, of Bollen (1998) that was used to provide prices for comparison with those produced by the model was not found to be a useful method. The problem was largely due to the impact of the high-volatility state on the simulated lattice swap rates. Over the time periods used for the maturities of the options that were priced, the high volatility tended to result in very high simulated swap rates. Since the options that were being priced were call options, this led to very high option prices, compared to those of the model, and the

'constant parameter' methods. This 'swamping effect' of the high volatility state on the option prices tends to suggest that the lattice method is only useful for options of short maturity, or those priced with a relatively low 'high' level of volatility.

## 14.5 Performance of the Model with Time-varying Transition Probabilities

Except for the 2-state model of the JSE Index, all of the models that were fitted under the assumption of time-varying transition probabilities provided a poorer fit to the data than those that assumed that the transition probabilities were constant, according to the AIC and BIC. Further evidence that this form of the model performed poorly on the given datasets can be found in the form of the very large standard errors of the parameters relating to the transition probabilities. This is an indication that these parameters are not statistically significant and that this model is thus not justified, for these datasets. The 3-state version of the model provided a poorer fit to the data than the 2-state version of the model, with the parameter estimates exhibiting some very strange behaviour. For example, there was no hidden state that represented bear market conditions despite the fact that the data used contained several known bear markets. This could be due to the 3-state model being over-specified and thus producing spurious results.

Although the implied volatility behaviour for call option prices computed using the 2-state version of the model was much closer to the conventional volatility smile, the behaviour for put options was very unusual. This fact, combined with the poor model fit to the data suggests that this form of the model should not be used for this data. This is reinforced when considering that this is a less parsimonious version of the Hidden Markov Option Pricing Model, which also results in a less efficient method of computing the occupation time probabilities. This is because these values have to be computed by Monte Carlo simulation, due to the dependence between successive values of the transition probabilities. This was not a good model, based on the evidence produced by the datasets used here.

## 14.6 Further Areas of Work on this Model

When reviewing all of the above, it is clear that evidence for the usefulness of the Hidden Markov Option Pricing Model is mixed. The model appears to offer several advantages, such as accommodating time-varying parameters and the skewness and leptokurtosis often found in the empirical distribution of asset returns. It also has some disadvantages, largely relating to problems associated with specifying the number of possible hidden states. When implemented, the model does appear to capture extra information and effects that the Black-Scholes model cannot. It does not, however, appear to be able to incorporate the commonly experienced volatility smile effect. This set of mixed results suggests

that further work will be necessary to determine the worth of the model. Some suggestions for such further efforts follow.

One possibility is to test the model against some other model that accommodates time-varying parameters, such as the stochastic volatility models or the GARCH option pricing model. Performance could be judged according to a similar scheme to that used in this work, comparing prices and implied volatility behaviour and the difference between model prices and the Black-Scholes equivalents. Another possibility is to test the model on different datasets. It may be the case that the poor performance of the model, in some areas, was due to a quirk of the markets that were considered here, or due to the time period that data was sampled in. For example, Ishijima and Kihara (2005) find much better implied volatility performance using data from the Topix index, which is based on the Japanese market. A final possible method, and perhaps the best for deciding the issue, would be to compare prices generated using the model to those found in the market. If the model prices were able to approximate the market prices closely, for any maturity or strike price of options, then it would be proven to be a good model. This test would have to take place over a period of time, or with datasets covering several disjoint time periods, to ensure that the model performs well under different market conditions and that any good performance was not due to chance. Until such a test has been performed, the question of the usefulness of the Hidden Markov Option Pricing Model for pricing options will not be fully answered.

## Glossary

**American-style option:** Option that grants the holder the right to enter into the transaction specified in the option contract at any time up to and including the expiry date of the option

**ARIMA:** AutoRegressive Integrated Moving Average. A form of time series model where each value in the series is expressed as a linear combination of certain previous values in the series, and certain previous innovation terms

**ARMA:** AutoRegressive Moving Average. A form of ARIMA model.

**Bear market:** Set of market conditions characterised by falling prices and, by the leverage effect, high levels of volatility of prices and returns

**Bermudan-style option:** Option that grants the holder the right to enter into the transaction specified in the option contract at any of a finite number of discrete, deterministic, exercise dates prior to the expiry date of the option.

**Bull market:** The opposite of a bear market. Set of market conditions characterised by rising prices and low levels of volatility

**Call Option:** Option that grants the holder the right to buy the asset specified in the option contract, for the strike price

**European-style option:** Option that grants the holder the right to enter into the transaction specified in the option contract on the expiry date of the option only.

**Exercise date:** Date on which the transaction specified in an option contract may be entered into by the holder of the option, should they wish to do so. If the holder has chosen to enter into the transaction, then this may refer to the date on which they did so

**Expiry:** Date after which an option contract is no longer valid

**GARCH:** Generalised Autoregressive Conditional Heteroscedasticity. A model for volatility time series that is a form of ARMA model

- Implied volatility:** Value required to set the option price generated by the Black-Scholes formula equal to that observed in the market, or produced by some other option pricing model. This has to be evaluated by solving for the volatility parameter in the Black-Scholes formula, using some nonlinear equation solution method
- Markov Chain:** Discrete stochastic process that consists of a set of discrete, qualitative, values that change randomly over time. The model is fully specified by the matrix of transition probabilities, being the probabilities of moving from a given state value to another, and the distribution of the values at the initial time point. The dependence structure of the model is Markov, allowing for a very tractable framework
- Martingale:** A stochastic process that satisfies the property that the conditional expectation of any future value of the process, conditional on the information available up to the present, is equal to the current value of the process
- Maturity:** The amount of time remaining until expiry for a given option contract. This term may also be used to refer to the expiry date itself. This can usually be determined from the context within which it is used
- Put Option:** Option that grants the holder the right to sell the asset specified in the option contract, for the strike price
- Risk:** The potential for loss that an investor faces by investing in a given asset. Usually this is measured as a combination of the possible loss that the investor faces if the future value of the asset is unfavourable to them, and the probability of this event/s arising. For investment analysis the standard deviation of returns on the asset is used as a risk measure, since a larger standard deviation implies a greater possibility that the value of the asset in the future will be lower than it is today. This represents a risk to the investor, who would then face a loss.
- Risk averse:** An investor is said to be risk averse if they prefer not to take on risk. Unless otherwise stated, all investors are typically assumed to be risk averse. In order to earn a return on an investment, financial theory requires that risk be taken on by the investor. A risk averse investor requires a more than corresponding level of return on a

risky asset, for a given level of risk, to compensate them for bearing that risk, to cause them to invest in that risky asset.

**Risk neutral:** An investor is said to be risk neutral if they have no preferences with regard to risk. They do not require extra compensation for taking on more risk, and so require a rate of return equal to the risk-free rate of return on any asset that they invest in.

**Risk seeking:** An investor is said to be risk seeking if they prefer to take on more risk rather than less. Taking on the risk is, in itself, a form of 'compensation' to them. They require a less than corresponding level of return, for a given level of risk on an asset, to cause them to invest in that asset.

**Strike Price:** Price that the transaction specified in the option contract will be transacted at, should the option be exercised.

**Volatility:** The annualised standard deviation of returns on a particular financial variable. The annualisation factor is equal to the square root of the amount of time units that the standard deviation was calculated based on. So, for example, if the standard deviation was calculated using daily data, the factor would be the square root of 250: the number of trading days in a year. Or, if monthly data was used, then the factor would be the square root of 12, etc.

## Bibliography

- [1] Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control* **19**, No. 6, 716-722
- [2] Barone-Adesi, G. and Whaley, R.E. (1987). Efficient Analytic Approximation of American Option Values. *The Journal of Finance* **42**, No.2, 301-320
- [3] Baum, L. and Petrie, T. (1966). Statistical Inference for Probabilistic Functions of Finite State Markov Chains. *The Annals of Mathematical Statistics* **37**, No. 6, 1554-1563
- [4] Baum, L., Petrie, T., Soules, G. and Weiss, N. (1970). A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. *The Annals of Mathematical Statistics* **41**, No. 1, 164-171
- [5] Bengio, Y. (1997). Markov Models for Sequential Data. Available at: <http://www.iro.umontreal.ca/~bengioy/ift6265/hmms/hmms2.html>  
Last visited: 05/06/2006
- [6] Bhar, R. and Hamori, S. (2004). Hidden Markov Models: Applications to Financial Economics. *Advanced Studies in Theoretical and Applied Econometrics* **40**, Kluwer Academic Publishers, Boston
- [7] Bickel, P.J., Ritov, Y., and Ryden, T. (1998). Asymptotic Normality of the Maximum Likelihood Estimator For General Hidden Markov Models. *The Annals of Statistics* **26**, No. 4, 1614-1635
- [8] Björk, T. (2004). Arbitrage Theory in Continuous Time (2<sup>nd</sup> Ed.). Oxford University Press, Oxford.
- [9] Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the 1976 meetings of the business and economic statistics section, American Statistical Association*, 177-181
- [10] Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* **81**, 637-659
- [11] Bollen, N. (1998). Valuing Options in Regime-Switching Models. *Journal of Derivatives* **6**, No. 1, 38-49
- [12] Bollen, N., Gray, S.F. and Whaley, R.E. (2000). Regime switching in foreign exchange rates: Evidence from currency option prices. *Journal of Econometrics* **94**, 239-276

- [13] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* **31**, 307-327
- [14] Bollerslev, T., Chou, R.Y and Kroner, K.F. (1992) ARCH modeling in finance. *Journal of Econometrics* **52**, 5-59
- [15] Britten-Jones, M. and Neuberger, A. (2000). Option Prices, Implied Price Processes and Stochastic Volatility. *Journal of Finance* **55**, No. 2, 839-866
- [16] Bühlmann, H., Embrechts, P., Delbaen, F. and Shiryaev, A.N. No-arbitrage, change of measure and conditional Esscher transforms.  
Available at: <http://www.math.ethz.ch/~delbaen/ftp/preprints/BDES~CWI.pdf>.  
Last visited: 05/06/2006
- [17] Buraschi, A. and Jackwerth, J. (2001). The price of a smile: Hedging and Spanning in options markets. *The Review of Financial Studies* **14**, No. 2, 495-527
- [18] Campa, J.M., Reider, R. and Chang, P.H. (1997). Implied Exchange Rate Distributions: Evidence from OTC Options Markets.  
Available at:  
<http://finance.math.biu.ac.il/readings/bonds/3%02implied%20exch%20rates.pdf>  
Last visited: 05/06/2006
- [19] Cappé, O., Moulines, E. and Ryden, T. (2005). Inference in Hidden Markov Models. Springer, New York
- [20] Casarin, R. (2003). Bayesian Inference for Markov Switching Stochastic Volatility Models.  
Available at: <http://www.ceremade.dauphine.fr/preprints/CMD/2004-14.pdf>.  
Last visited: 05/06/2006
- [21] Chopin, N. and Pelgrin, F. Bayesian inference and state determination for hidden Markov models: an application to the information content of the yield curve about inflation  
Available at: <http://www.crest.fr/doctravail/document/2001-44.pdf>.  
Last visited: 05/06/2006
- [22] Cosslett, S.R. and Lee, L-F. (1985). Serial Correlation in Latent Discrete Variable Models. *Journal of Econometrics* **27**, 79-97
- [23] Cox, J., Ross, S. and Rubinstein, M. (1979). Option Pricing: A Simplified Approach. *Journal of Financial Economics* **7**, 229-264
- [24] Delyon, B., Lavielle, M. and Moulines, E. (1999). Convergence of a Stochastic Approximation Version of the EM Algorithm. *The Annals of Statistics* **27**, No. 1, 94-128

- [25] Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum Likelihood Estimation from Incomplete Data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, **39**, 1-38
- [26] Derman, E. and Kani, I. (1994a). The volatility smile and its implied tree. Quantitative Strategies Research Notes. Goldman-Sachs and Company.
- [27] Derman, E. and Kani, I. (1994b). Riding on the smile. *Risk* **7**, 32-39
- [28] Driffill, J., Kenc, T. and Sola, M. (2002). Merton-Style Option Pricing Under Regime Switching.  
Available at: <http://www.ismacentre.rdg.ac.uk/pdf/seminar/kenc1.pdf>. Last visited: 05/06/2006
- [29] Duan, J-C. (1995). The GARCH Option Pricing Model. *Mathematical Finance* **5**, 13-32
- [30] Duan, J-C. (1996). A Unified Theory of Option Pricing Under Stochastic Volatility - from GARCH to Diffusion. Hong Kong University of Science and Technology: unpublished.
- [31] Duan, J-C., Popova, I. and Ritchken, P., (2002). Option Pricing under Regime Switching. *Quantitative Finance* **2**, 1-17
- [32] Elliott, R.J., Aggoun, L. and Moore, J.B. (1995). Hidden Markov Models: Estimation and Control, *Applications of Mathematics* **29**, Springer-Verlag, New York
- [33] Elliott, R.J. and Buffington, J. (2002). American Options with Regime Switching. *International Journal of Theoretical and Applied Finance* **5**, No. 5, 497-514
- [34] Elliott, R.J., Chan, L. and Siu, T.K. (2005). Option pricing and Esscher transform under regime switching. *Annals of Finance* **1**, 423-432
- [35] Fama, E.F. (1976). Foundations of Finance. Basil Blackwell, Oxford, 17-38
- [36] Gerber, H.U and Shiu, E.S.W. (1994). Option Pricing by Esscher Transforms. *Transactions of the Society of Actuaries* **46**, 99-191
- [37] Gray, S.F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics* **42**, 27-62
- [38] Guo, X. (2001). Information and option pricing. *Quantitative Finance* **1**, 38-44

- [39] Hamilton, J.D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica* **57**, No. 2, 357-384
- [40] Hamilton, J.D. (1990). Analysis of Time Series Subject to Changes in Regime. *Journal of Econometrics* **45**, 39-70
- [41] Hamilton, J.D. (1994). Modeling Time Series with Changes in Regime. *Time Series Analysis*, Princeton University Press, Princeton, New Jersey, 677-703
- [42] Hamilton, J.D. (1996). Specification testing in Markov-switching time-series models. *Journal of Econometrics* **70**, 127-157
- [43] Hamilton, J.D. and Susmel, R. (1998). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics* **64**, 307-333
- [44] Harrison, M. and Kreps, D. (1979). Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory* **20**, 381-408
- [45] Harrison, M. and Pliska, S. (1981). Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochastic Processes and Their Applications* **11**, 215-260
- [46] Heston, S.L. (1993). A Closed-Form Solution For Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies* **6**, No. 2, 327-343
- [47] Heston, S.L. and Nandi, S. (2000). A Closed-Form GARCH Option Pricing Model. *Review of Financial Studies* **13**, No. 3, 585-625
- [48] Hull, J.C. (2003). Options, Futures and Other Derivatives (5<sup>th</sup> Ed.). Prentice Hall, New Jersey.
- [49] Hull, J.C. and White, A. (1987). The pricing of options on assets with stochastic volatility. *The Journal of Finance* **2**, 281-300
- [50] Ishijima, H. and Kihara, T. (2005). Option Pricing with Hidden Markov Models. Waseda University Institute of Finance Working Paper Series, Available at: [http://www.waseda.jp/wnfs/pdf/labo5\\_2005/wnif05-004.pdf](http://www.waseda.jp/wnfs/pdf/labo5_2005/wnif05-004.pdf)  
Last visited: 05/06/2006
- [51] Jarque, C.M. and Bera, A.K. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review* **55**, No. 2, 163-172
- [52] Juang, B.H. and Rabiner, L.R. (1991). Hidden Markov Models for speech recognition. *Technometrics* **33**, No.3, 251-272

- [53] Karatzas, I. and Shreve, S.E. (2000). *Brownian Motion and Stochastic Calculus* (2<sup>nd</sup> Ed.). Springer, Berlin.
- [54] Konikov, M. and Maden, D.B. (2002). Option Pricing using Variance Gamma Markov Chains. *Review of Derivatives Research* **5**, No.1, 81-115
- [55] Krogh, A., Brown, M., Mian, I.S., Sjolander, K. and Haussler, D. (1994). Hidden Markov Models in computational biology: an application to protein modeling. *Journal of Molecular Biology* **235**, No. 5, 1501-1531
- [56] Krolzig, H.M. (1997). *Markov-switching Vector Autoregressions: Modelling, Statistical Inference, and Applications to Business Cycle Analysis*. Springer, Berlin.
- [57] Lamoureux, C.G. and Lastrapes, W.D. (1990). Persistence in Variance, Structural Change and the GARCH Model. *Journal of Business and Economic Statistics* **8**, No. 2, 225-234
- [58] Laverly, W.H., Miket, M.J. and Kelly, I.W. (2002). Simulation of hidden Markov models with EXCEL. *The Statistician* **51**, No. 1, 31-40
- [59] Louis, T.A. (1982). Finding the Observed Information Matrix when Using the EM Algorithm. *Journal of the Royal Statistical Society, Series B*, **44**, No. 2, 226-233
- [60] MacDonald, I.L. and Zucchini W. (1997). *Hidden Markov and Other Models for Discrete-valued Time Series*. Monographs on Statistics and Applied Probability **70**, Chapman and Hall, London
- [61] McLachlan, G. and Peel, D. (2000). *Finite Mixture Models*. Wiley Series in Probability and Statistics. John Wiley and Sons, New York
- [61] Meyer, M. (2001). *Continuous Stochastic Calculus with Applications to Finance*. Chapman and Hall, Florida.
- [62] Musiela, M. and Rutkowski, M. (1998). *Methods in Financial Modeling*. Springer, Berlin.
- [63] Nefian, A.V. and Hayes, M.H. (1998). Hidden Markov Models for face recognition. *Acoustics, Speech and Signal Processing* **5**, 2721-2724

- [64] Otranto, E. and Gallo, G.M. (2001). A Nonparametric Bayesian Approach to Detect the Number of Regimes in Markov Switching Models. Working Paper 2001/04 Dipartimento di Statistica "G. Parenti", Universita di Firenze,  
*Available at:*  
[http://www.ds.unifi.it/ricerca/pubblicazioni/working\\_papers/2001/wp2001\\_04.pdf](http://www.ds.unifi.it/ricerca/pubblicazioni/working_papers/2001/wp2001_04.pdf)  
Last visited: 05/06/2006
- [65] Rabiner, L.R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE* **77**, No. 2, 257-284
- [66] Ritchken, P. and Trevor, R. (1999). Pricing Options under Generalized GARCH and Stochastic Volatility Models. *The Journal of Finance* **54**, No.1, 377-402
- [67] Robert, C.P., Ryden, T. and Titterton, D.M. (2000). Bayesian Inference in Hidden Markov Models through the Reversible Jump Markov Chain Monte Carlo Method. *Journal of the Royal Statistical Society, Series B*, **62**, No.1, 57-75
- [68] Rossi, A. and Gallo, G. (2002). Volatility Estimation via Hidden Markov Models.  
*Available at:*  
[http://www.ds.unifi.it/ricerca/pubblicazioni/working\\_papers/2002/wp2002\\_14.pdf](http://www.ds.unifi.it/ricerca/pubblicazioni/working_papers/2002/wp2002_14.pdf)  
Last visited: 06/06/2006
- [69] Ryden, T. (1994). Consistent and Asymptotically Normal Parameter Estimates For Hidden Markov Models. *The Annals of Statistics* **22**, No. 4, 1884-1895
- [70] Sclove, S.L. (1983). Time-Series Segmentation: A Model and a Method. *Information Sciences* **29**, 7-25
- [71] Tzavalis, E. and Chourdakis, K.M. (1999). Option Pricing Under Discrete Shifts in Stock Returns. *Available at:*  
<http://www.qmul.ac.uk/~ugte184/pdf/ct.pdf> . Last visited: 05/06/2006
- [72] Wiggins, J.B. (1987). Option Values Under Stochastic Volatility. *Journal of Financial Economics* **19**, 351-372